



جامعة عبد المالك السعدي
تanger | مکانیکی عالیہ | دارالفنون
Université Abdelmalek Essaadi



RAPPORT BIG DATA & DATA ENGINEERING

ANALYSE DE LOGS EN TEMPS RÉEL

Gestion des Événements

Réalisée par :

Sara El-Otmani

Encadrée par :

M. Hassan BADIR

1. Introduction

Dans le cadre du module Big Data / Data Engineering, nous avons développé un projet de traitement et analyse de logs système en temps réel.

Objectifs du projet :

1. Mettre en pratique les concepts Spark (RDD, DataFrame, Spark SQL, Spark Streaming).
2. Comprendre le fonctionnement d'un pipeline Big Data complet : ingestion → stockage → traitement → visualisation.
3. Utiliser des technologies distribuées : Kafka, HDFS et Spark.

Cas d'usage choisi :

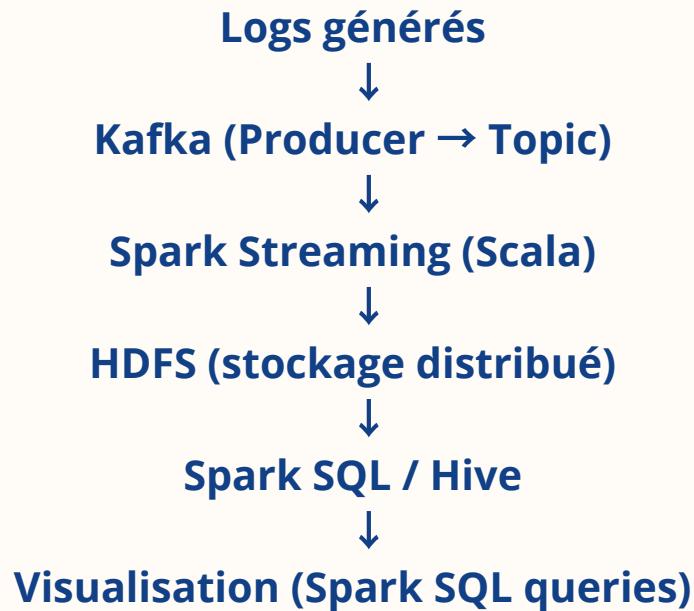
- Analyse des logs système pour détecter en temps réel les erreurs, warnings et événements critiques.

2. Architecture du pipeline

Le pipeline se compose des étapes suivantes :

Étape	Technologie	Rôle
Génération de logs	Kafka Producer	Simuler l'arrivée continue de logs
Ingestion	Kafka Topic	Stockage temporaire des messages
Traitement en temps réel	Spark Streaming (Scala)	Transformation, filtrage, comptage
Stockage distribué	HDFS	Stockage des logs filtrés et résultats
Analyse / Visualisation	Spark SQL / DataFrame	Requêtes et agrégations

Diagramme du pipeline :



3. Technologies utilisées

3.1 Apache Kafka

- Rôle : Gestion de messages en temps réel (logs)
- Concepts : Producer, Consumer, Topic

3.2 Apache Spark

- Modules utilisés : RDD, DataFrame, Spark SQL, Spark Streaming
- Rôle : Lecture, transformation et traitement des flux de données

3.3 HDFS

- Rôle : Stockage distribué des logs
- Avantages : Scalabilité et tolérance aux pannes

3.4 Scala

- Rôle : Langage principal pour coder Spark
- Avantage : Intégration directe avec Spark

```
C:\Windows\System32\cmd.exe - .\bin\windows\zookeeper-server-start.bat .\config\zookeeper.properties
r.server.quorum.QuorumPeerConfig)
[2025-12-28 23:14:14,393] INFO clientPortAddress is 0.0.0.0:2181 (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2025-12-28 23:14:14,393] INFO secureClientPort is not set (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2025-12-28 23:14:14,393] INFO observerMasterPort is not set (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2025-12-28 23:14:14,394] INFO metricsProvider.className is org.apache.zookeeper.metrics.impl.DefaultMetricsProvider (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2025-12-28 23:14:14,394] INFO Starting server (org.apache.zookeeper.server.ZooKeeperServerMain)
[2025-12-28 23:14:14,451] INFO ServerMetrics initialized with provider org.apache.zookeeper.metrics.impl.DefaultMetricsProvider@1936f0f5 (org.apache.zookeeper.server.ServerMetrics)
[2025-12-28 23:14:14,458] INFO ACL digest algorithm is: SHA1 (org.apache.zookeeper.server.auth.DigestAuthenticationProvider)
[2025-12-28 23:14:14,459] INFO zookeeper.DigestAuthenticationProvider.enabled = true (org.apache.zookeeper.server.auth.DigestAuthenticationProvider)
[2025-12-28 23:14:14,473] INFO zookeeper.snapshot.trust.empty : false (org.apache.zookeeper.server.persistence.FileTxnSnapLog)
[2025-12-28 23:14:14,507] INFO (org.apache.zookeeper.server.ZooKeeperServer)
[2025-12-28 23:14:14,508] INFO _____ - (org.apache.zookeeper.server.ZooKeeperServer)
[2025-12-28 23:14:14,510] INFO |__ / ||| (org.apache.zookeeper.server.ZooKeeperServer)
[2025-12-28 23:14:14,516] INFO / / __ _ ||| - _ _ _ _ - _ _ _ _ - _ _ (org.apache.zookeeper.server.ZooKeeperServer)
[2025-12-28 23:14:14,519] INFO / / / _ \ / _ \ | | // / _ \ / _ \ / _ \ / _ \ | ' | (org.apache.zookeeper.server.ZooKeeperServer)
[2025-12-28 23:14:14,521] INFO / / _ | ( ) | | ( ) | | < | _ / | _ / | | | | (org.apache.zookeeper.server.ZooKeeperServer)
[2025-12-28 23:14:14,522] INFO /____| \____/ \____/ |_|\_ \ \____| \____| | . / \____| | (org.apache.zookeeper.server.ZooKeeperServer)
[2025-12-28 23:14:14,523] INFO | | (org.apache.zookeeper.server.ZooKeeperServer)
```

- Lancement du ZOOKEEPER

```
Invite de commandes - .\bin\windows\kafka-server-start.bat .\config\server.properties
Microsoft Windows [version 10.0.19045.6456]
(c) Microsoft Corporation. Tous droits réservés.

C:\Users\HP>cd C:\kafka

C:\kafka>.\bin\windows\kafka-server-start.bat .\config\server.properties
[2025-12-28 23:14:48,083] INFO Registered kafka:type=kafka.Log4jController MBean (kafka.utils.Log4jControllerRegistration$)
[2025-12-28 23:14:48,822] INFO Setting -D jdk.tls.rejectClientInitiatedRenegotiation=true to disable client-initiated TLS renegotiation (org.apache.zookeeper.common.X509Util)
[2025-12-28 23:14:49,095] INFO starting (kafka.server.KafkaServer)
[2025-12-28 23:14:49,096] INFO Connecting to zookeeper on localhost:2181 (kafka.server.KafkaServer)
[2025-12-28 23:14:49,166] INFO [ZooKeeperClient Kafka server] Initializing a new session to localhost:2181. (kafka.zookeeper.ZooKeeperClient)
[2025-12-28 23:14:58,274] INFO Client environment:zookeeper.version=3.8.4-9316c2a7a97e1666d8f4593f34dd6fc36ecc436c, built on 2024-02-12 22:16 UTC (org.apache.zookeeper.ZooKeeper)
[2025-12-28 23:14:58,275] INFO Client environment:host.name=DESKTOP-4FU74TK (org.apache.zookeeper.ZooKeeper)
[2025-12-28 23:14:58,276] INFO Client environment:java.version=1.8.0_202 (org.apache.zookeeper.ZooKeeper)
[2025-12-28 23:14:58,277] INFO Client environment:java.vendor=Oracle Corporation (org.apache.zookeeper.ZooKeeper)
[2025-12-28 23:14:58,277] INFO Client environment:java.home=C:\java\jdk1.8.0_202\jre (org.apache.zookeeper.ZooKeeper)
[2025-12-28 23:14:58,278] INFO Client environment:java.class.path=C:\kafka\libs\activation-1.1.1.jar;C:\kafka\libs\aaopalliance-repackaged-2.6.1.jar;C:\kafka\libs\argparse4j-0.7.0.jar;C:\kafka\libs\audience-annotations-0.12.0.jar;C:\kafka\libs\caffeine-2.9.3.jar;C:\kafka\libs\commons-beanutils-1.9.4.jar;C:\kafka\libs\commons-cli-1.4.jar;C:\kafka\libs\commons-collections-3.2.2.jar;C:\kafka\libs\commons-digester-2.1.jar;C:\kafka\libs\commons-io-2.14.0.jar;C:\kafka\libs\commons-lang3-3.12.0.jar;C:\kafka\libs\commons-logging-1.2.jar;C:\kafka\libs\commons-validator-1.7.jar;C:\kafka\libs\connect-api-3.9.1.jar;C:\kafka\libs\connect-basic-auth-extension-3.9.1.jar;C:\kafka\libs\connect-file-3.9.1.jar;C:\kafka\libs\connect-json-3.9.1.jar;C:\kafka\libs\connect-mirror-3.9.1.jar;C:\kafka\libs\connect-mirror-client-3.9.1.jar;C:\kafka\libs\connect-runtime-3.9.1.jar;C:\kafka\libs\connect-transforms-3.9.1.jar;C:\kafka\libs\error_prone_annotations-2.10.0.jar;C:\kafka\libs\hk2-api-2.6.1.jar;C:\kafka\libs\hk2-locator-2.6.1.jar;C:\kafka\libs\hk2-utils-2.6.1.jar;C:\kafka\libs\jackson-annotations-2.16.2.jar;C:\kafka\libs\jackson-core-2.16.2.jar;C:\kafka\libs\jackson-databind-2.16.2.jar;C:\kafka\libs\jacks
```

- Lancement de Kafka avec succès

4. Implémentation

4.1 Kafka - Création du topic et Producer

```
C:\Windows\system32>cd C:\Kafka  
C:\kafka>.\bin\windows\kafka-topics.bat --create --topic logs-topic --bootstrap-server localhost:9092 --partitions 1 --replication-factor 1  
Created topic logs-topic.  
C:\kafka>
```

```
C:\kafka>.\bin\windows\kafka-console-producer.bat --topic logs-topic --bootstrap-server localhost:9092  
>ERROR Server down  
>INFO User login  
>WARN Disk almost full  
>Terminer le programme de commandes (O/N) ? o
```

Console Kafka Producer avec logs envoyés

```
C:\kafka>.\bin\windows\kafka-console-consumer.bat --topic logs-topic --bootstrap-server localhost:9092 --from-beginning  
ERROR Server down  
INFO User login  
WARN Disk almost full
```

4.2 Spark Streaming (Scala) - Lecture et traitement des logs

5. Architecture Technique

5.1 Langage et Outils

- Langage : Java
- IDE : Android Studio
- SDK minimum : Android 5.0 (API 21)

5.2 Organisation du Projet

Le projet est organisé en plusieurs packages afin d'assurer une bonne lisibilité du code :

- activities : contient les différentes activités
- models : contient la classe Note
- adapters : contient l'adapter personnalisé

6. Adapter Personnalisé

Un **Adapter personnalisé** (NoteAdapter) a été développé en héritant de la classe BaseAdapter.

Rôle de l'Adapter :

- Faire le lien entre la liste de notes (ArrayList<Note>) et la ListView
- Inflater un layout personnalisé (item_note.xml) pour chaque ligne
- Afficher dynamiquement les données de chaque note

Gestion des couleurs :

- Priorité **Haute** → Rouge
- Priorité **Moyenne** → Orange
- Priorité **Basse** → Vert

Cela permet à l'utilisateur d'identifier rapidement l'importance de chaque note.

7. Navigation entre les Activités

La navigation entre les écrans est réalisée à l'aide **d'Intents explicites** :

- Passage de l'écran liste vers l'écran ajout
- Retour à la liste après l'enregistrement grâce à finish()
- Passage des données d'une note vers l'écran de détails via putExtra()

Cette approche respecte le cycle de vie des activités Android.

8. Résultats Obtenus

L'application fonctionne correctement et permet :

- L'ajout de nouvelles notes
- L'affichage dynamique de la liste
- La consultation des détails d'une note
- Une interface simple et intuitive
- Une navigation fluide entre les écrans

9. Conclusion

Ce projet nous a permis de mettre en pratique les concepts fondamentaux du développement Android, notamment :

- La création d'activités
- La gestion des Intents
- L'utilisation des ListView
- Le développement d'un Adapter personnalisé
- L'organisation d'un projet Android professionnel

L'application **MyNotes** répond entièrement aux exigences du cahier des charges et constitue une base solide pour des améliorations futures comme la persistance des données ou l'ajout de fonctionnalités multimédia.