



Complex Event Processing

KLIPFFEL Tararaina
Polytech Grenoble - RICM 5

Sommaire

I/ Complex Event Processing

- a. Qu'est ce qu'un CEP
- b. Architecture EDA

II/ Domaine d'application

III/ Exemple de produit

- a. Apache Storm
- b. Spark Streaming

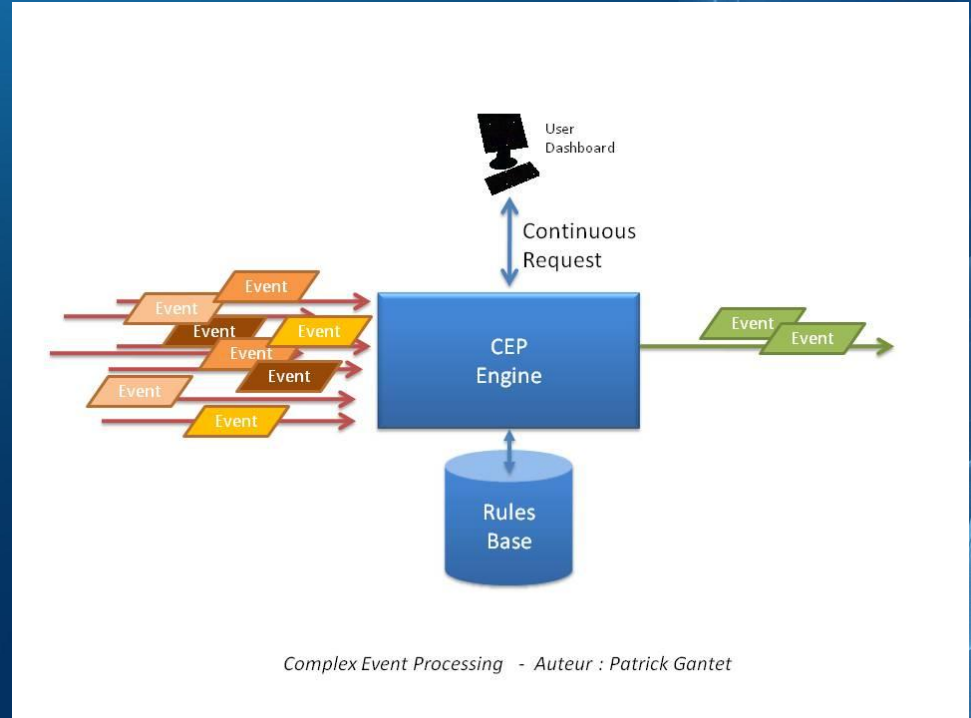
IV/ Positionnement : MapReduce Vs Spark

Qu'est ce qu'un Complex Event Stream Processing

- concept d'analyse d'événements
- calcul et/ou traitement des événements

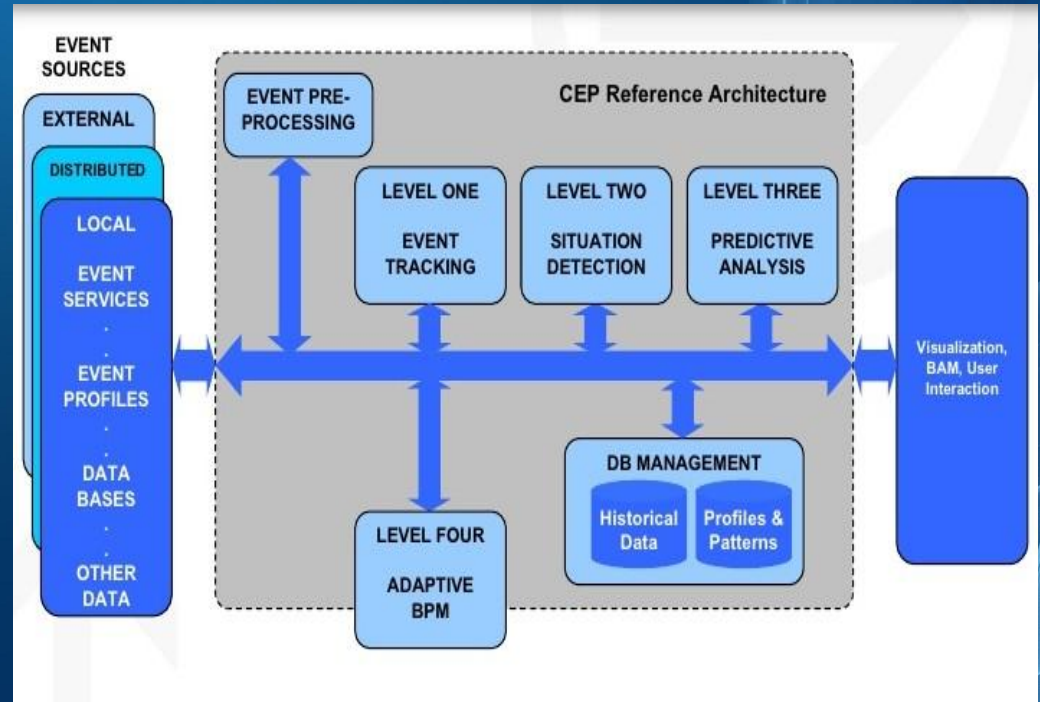
But :

- faire ressortir les informations contenues dans les événements
- analyser son impact et décider du plan d'action en temps réel



Architecture EDA (Event Driven Architecture)

- modèle interaction applicative met en oeuvre des services répondant aux demandes
- pas langage particulier
- pas de plateforme particulière
- communication asynchrone
- bus : PublishedSubscribe



Domaine d'application

- réseaux de capteurs (IoT)
- journaux système (détection de fraudes, ...)
- finance (quotations boursières)
- réseaux sociaux (tweeter)
- Business Activity Monitoring (BAM)
- Sécurité des systèmes d'information par corrélation d'événements. (alertes remontées par les composants logiciels et serveurs)
- Optimisation de tournées de flottes de véhicule (Optimisation en temps réel par corrélation de positions GPS et état du trafic)
- cellules de crise (aviation)

Apache Storm



- Storm : 2011
- CEP concept
- Mode de traitement : par stream (One at time)
- Language
 - Java
 - Scala
 - Ruby
 - Python
 -
- Tolérance aux pannes (travail perdu, l'état d'opérateur)

Spark Streaming



- Spark : 2010
- CEP concepts
- mode de traitement : micro-batch
- Langage
 - Java
 - Scala
 - Python
- Tolérance aux pannes (travail perdu, l'état d'opérateur)
- Exécution plateforme : Hadoop, Mesos, standalone, ou sur le cloud

Positionnement : MapReduce VS Spark

MapReduce

- MapReduce : 2004
- patron d'architecture de développement informatique
- inventé par Google
- calculs parallèles et distribués
- but : recherche web
- résolution non optimiser

Spark

- + Rapide (au niveau des applications)
 - facteur 100 entre MapReduce (Hadoop) et Spark (Cassandra)
 - opérations en mémoire, copie des données d'un système de stockage physique vers de la mémoire RAM
- facilité d'utilisation API
- flexibilité
 - requêtes interactives
 - algorithmes avancés comme le Machine Learning.

Merci, pour votre attention.

Avez-vous des questions ?



Démonstration 1 : Storm Apache

“topologie wordcount”

Démonstration 2 : SaaS IFTTT



Bibliographie

- <http://fabulous-lab.com/blog/2015/01/17/spark-streaming-affichez-vos-donnees-en-presque-temps-reel/>
- <http://blog.octo.com/complex-event-processing-cep-de-quoi-sagit-il/>
- <http://www.infoq.com/fr/articles/apache-spark-introduction>
- <http://searchsoa.techtarget.com/definition/complex-event-processing>
- https://upload.wikimedia.org/wikipedia/commons/3/39/Complex_Event_Processing.jpg
- <http://www.lemagit.fr/conseil/Spark-contre-MapReduce-quelle-solution-pour-les-entreprises>
- <http://www.infoq.com/fr/articles/apache-spark-introduction>