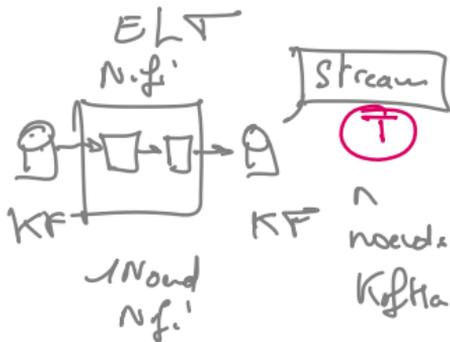
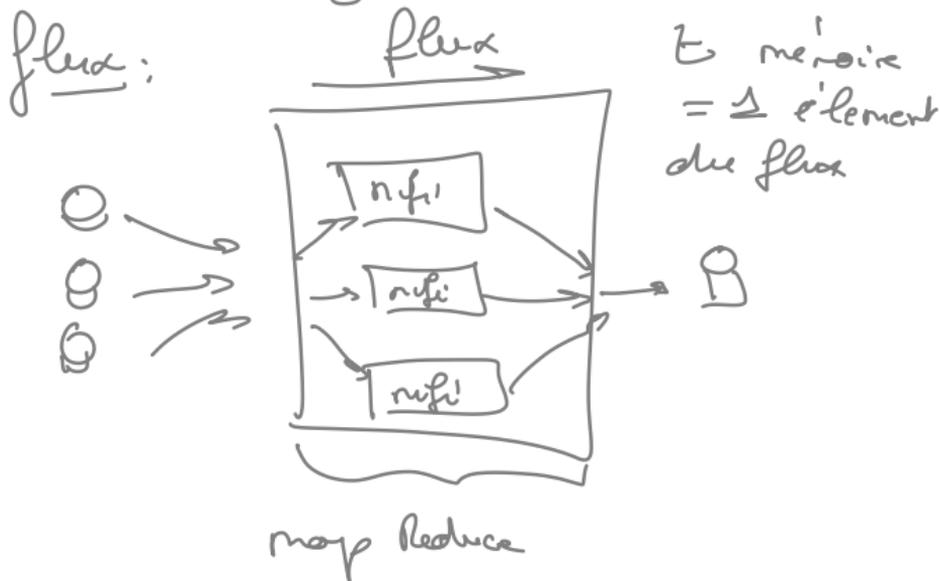


Bonjour tout le
monde


Téléchargement des fichiers du cours : <https://we.tl/t-V71unJ8Mhs>



Kafka



Implémentation d'un serveur Kafka



Ce support est la propriété d'IB SA



- Installation et architecture
 - Installation et options avancées
 - Fichiers journaux de Kafka
 - Réplication et fiabilité
 - Chemins d'écriture et de lecture de Kafka
 - Partitions, consumer groups et scalability
- Développement et mise en œuvre
 - Les outils en ligne de commande
 - Les APIs
 - Cas d'utilisation de Kafka Streams



INTRODUCTION À KAFKA

4





Introduction à Kafka

- Historique et complexité des systèmes
 - Les serveurs de type message broker (courtiers de messages), ou « message queues », ou « messagerie applicative » ne sont pas nouveaux
 - Ils permettent l'échange de messages asynchrones entre applications.
 - Les serveurs historiques :
 - IBM MQSeries
 - IBM WebSphere MQ
 - Microsoft MSMQ
 - Oracle Service Bus
 - Les serveurs openSource traditionnels
 - RabbitMQ
 - Apache ActiveMQ
 - Open JMS
- Kafka est initialement issu de l'épine dorsale de LinkedIn

5

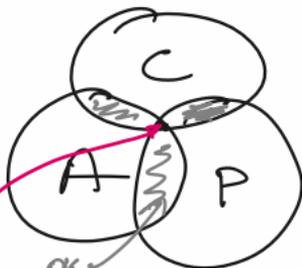


CAP

Consistency

Availability

Partitionning



Pas
Possible

OK
C, A, P en même temps → impossible

Kafka = Haute Dispo au
détriment de la cohérence
(AP)

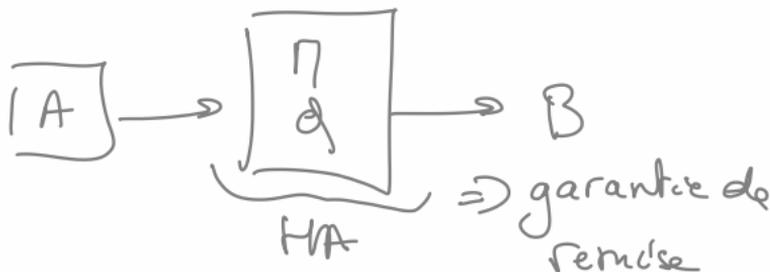


- Le traitement en temps réel : un important business
 - La messagerie applicative permet la garantie de remise des messages, de manière asynchrone, voir semi synchrone
 - Elle permet de résoudre les problèmes de connectivité entre les systèmes dont les appels synchrones ne sont pas possibles (RPC)
 - Différents scénarios évolués existent : transactions, message de retour, corrélations, niveaux de priorités, ...
 - Ce sont des MDM (Message Oriented Middlewares)
 - Les MQ sont très utilisés dans le domaine bancaire, ...
 - Ces serveurs sont passifs (pas de traitement), ne sont pas prévus pour fonctionner facilement en cluster, impliquent une charge de travail administrative.
 - Ils sont l'équivalent des serveurs relationnels par rapport au monde BigData

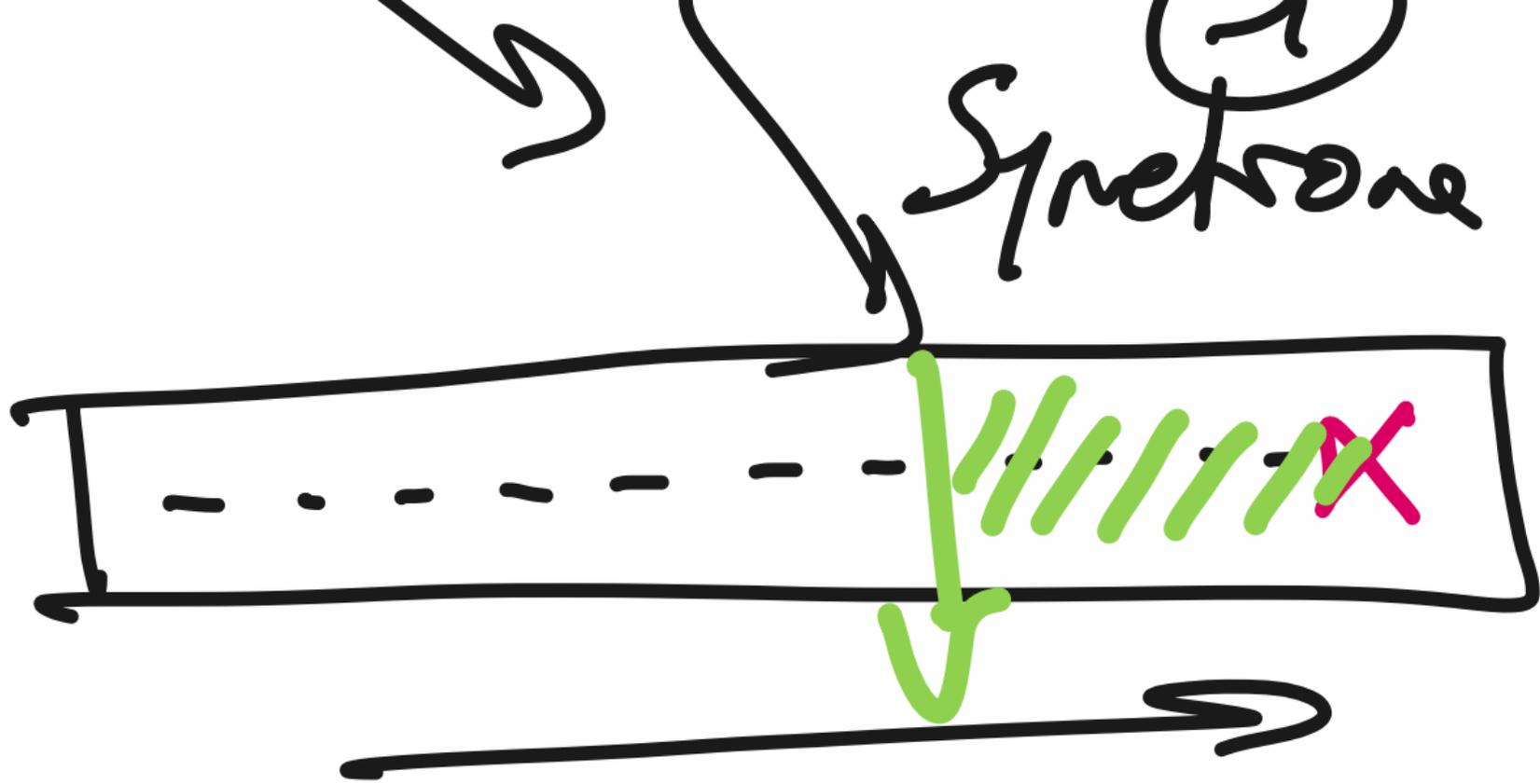
6



Synchrone \rightarrow Pbm de disponibilité



gpus



Synetrons

Layers



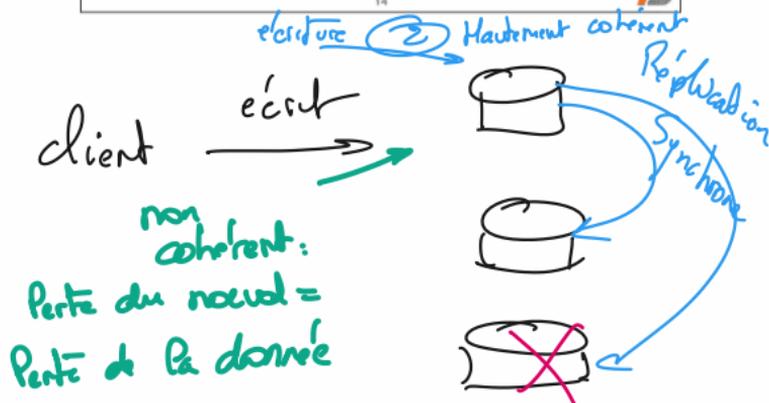
Aperçu de Kafka et de son efficacité

- Kafka peut être considéré comme un serveur de base de données NoSql orienté messages
 - Compatible avec la règle des 3V (Vitesse, Volume, Variabilité de schéma)
 - Sa mise l'échelle horizontale (Scale out) est nativement conçue
 - Les niveaux de cohérence en réplcation sont pris en charge
 - Il peut traiter et stocker des volumes énormes de données

(Data Lake)

14

ib



Les fonctionnalités de Kafka

- Système hautement disponible distribué
- Broker
 - Kafka est un courtier de messages (Broker)
- Publier / s'Abonner
 - Les éditeurs (publishers) peuvent envoyer des messages à des files de messages (topics), qui peuvent être éclatés en partitions
 - Les abonnés reçoivent des notifications (push), ou récupèrent les messages (pull) depuis les topics
- Stockage
 - Kafka est comme un - commit log - et stocke donc les messages selon des règles de rétention, à travers les nœuds du cluster
- Traitement
 - L'API Streams permet de traiter les messages au fil de leur arrivée, pour des scénarios complexes.



Stockage Big Data = Réplique + partition



Big Compute

→ HPC high perf Computing

→ Big Compute Map Reduce → Traitement de base

