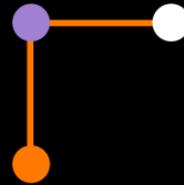


# Flexible Engine Webinaire

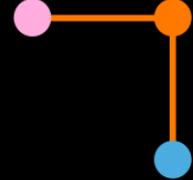
Data lake : révéléz la valeur  
de vos données grâce au cloud



**Merci pour votre présence  
aujourd'hui !**



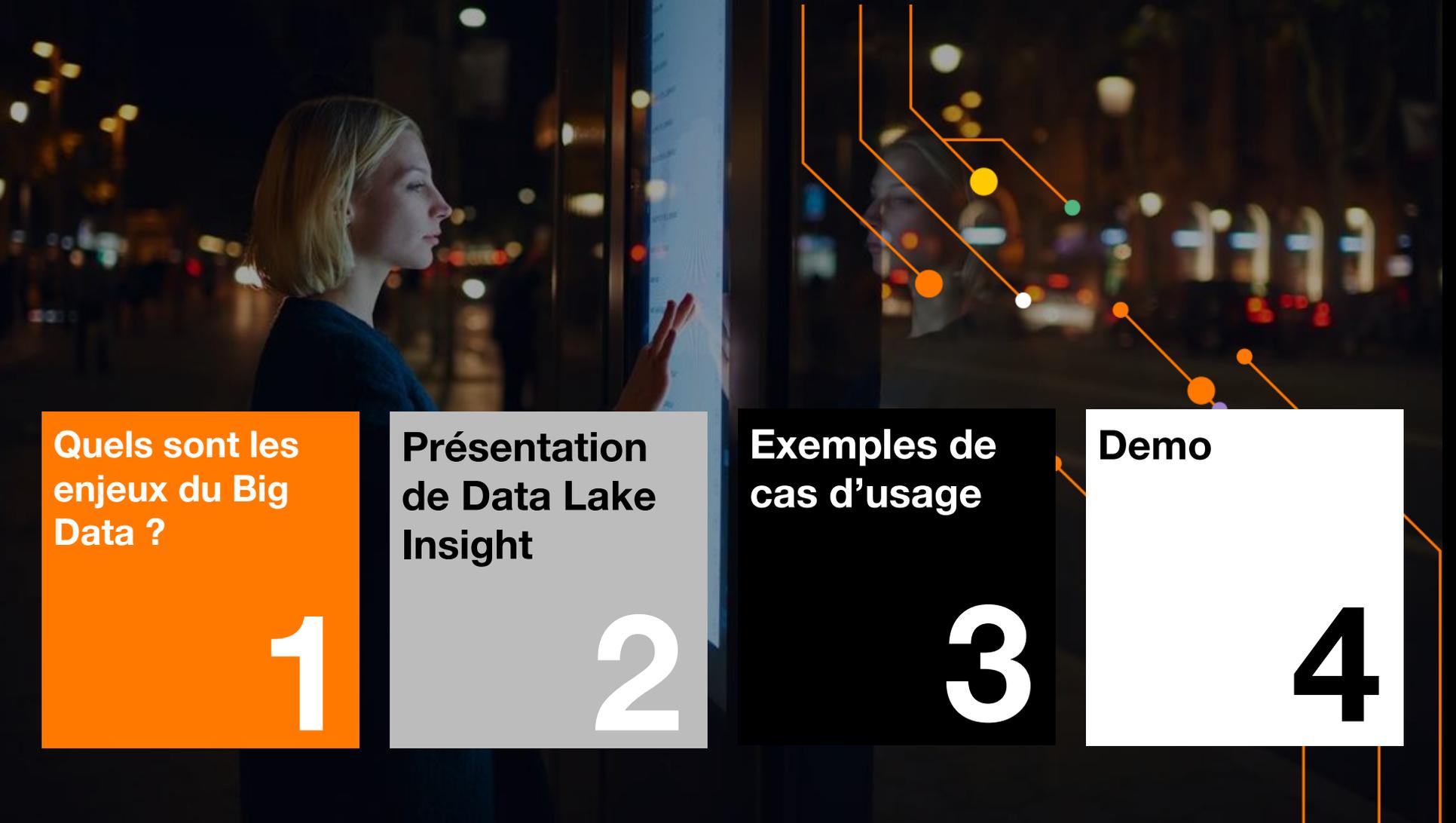
**Business  
Services**



**Wissam Ben Boubaker**  
**Cloud Solution Architect**  
[wissam.benboubaker@orange.com](mailto:wissam.benboubaker@orange.com)



**N'hésitez pas à poser vos  
questions dans le chat**



**Quels sont les enjeux du Big Data ?**

**1**

**Présentation de Data Lake Insight**

**2**

**Exemples de cas d'usage**

**3**

**Demo**

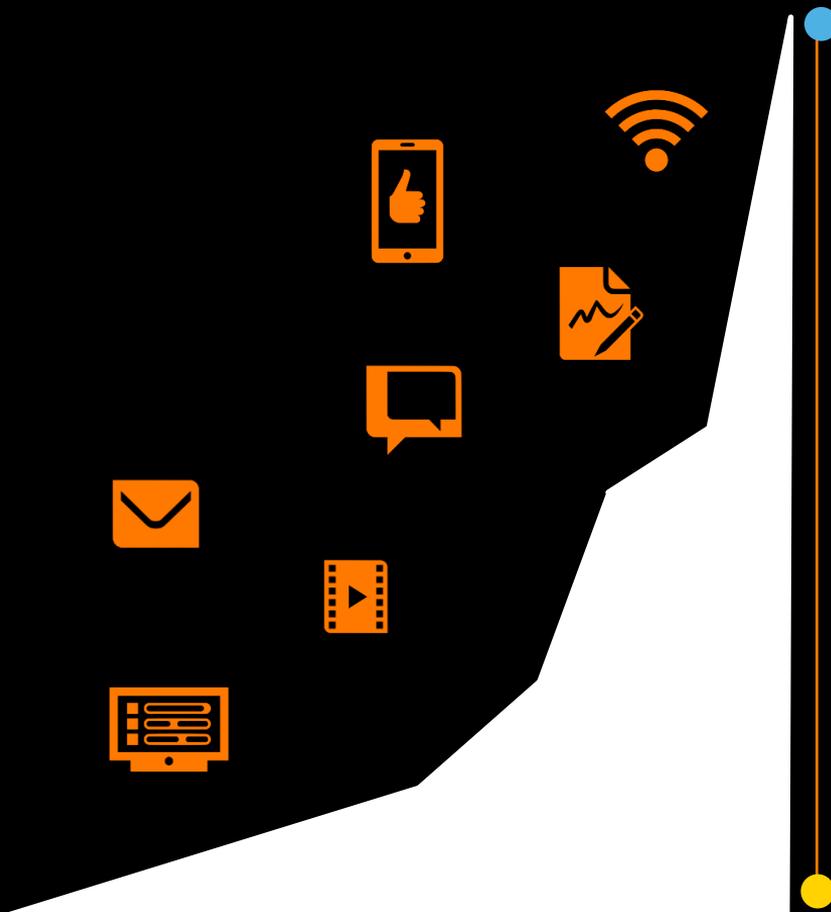
**4**

# Big Data

Quels sont les enjeux ?



# La donnée est exponentielle



Il y a **+** de données  
que les gens ne le pensent

Les données sont plus **diverses**

Les données  
augmentent

**>10x**

tous les 5 ans

**12,8%**  
par an d'ici  
2025, selon  
IDC

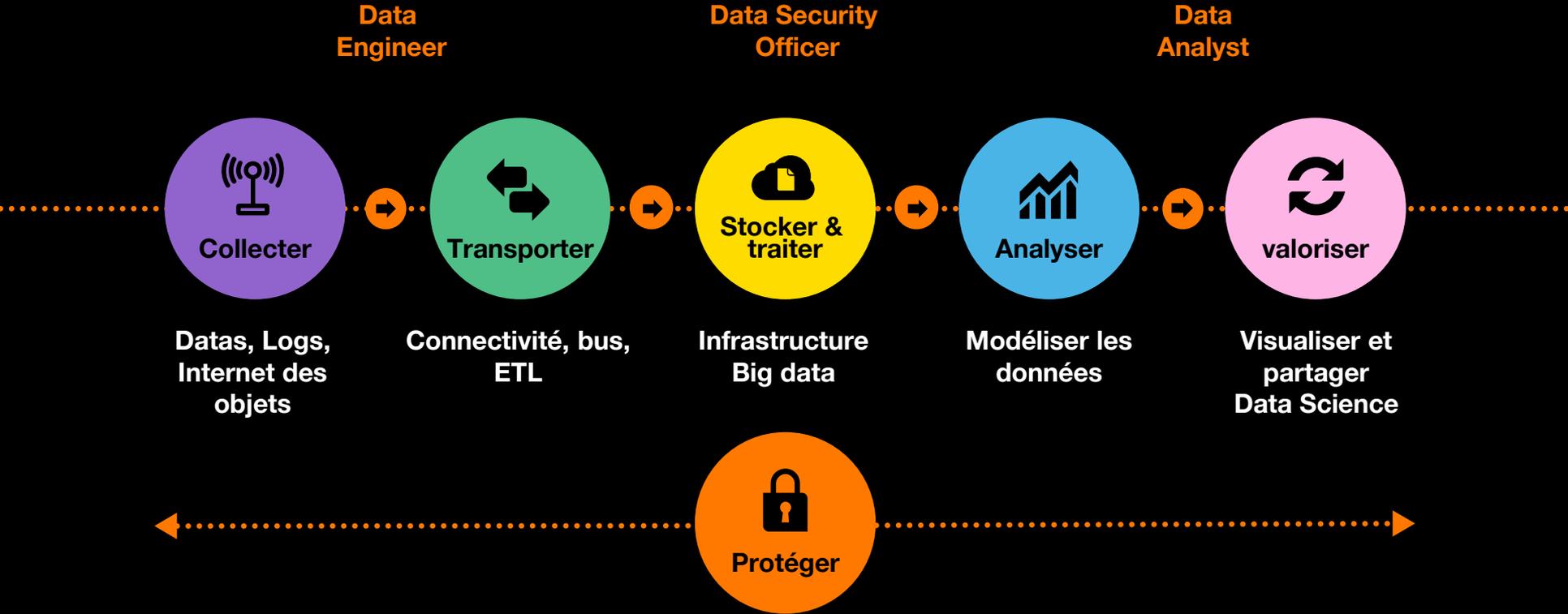
Les plateformes  
de données doivent

durer **15** ans

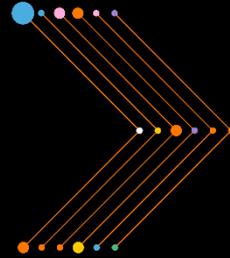
et être scalables

**1 000x**

# Le Big data au cœur du voyage de la donnée



# L'entreprise de demain sera data driven



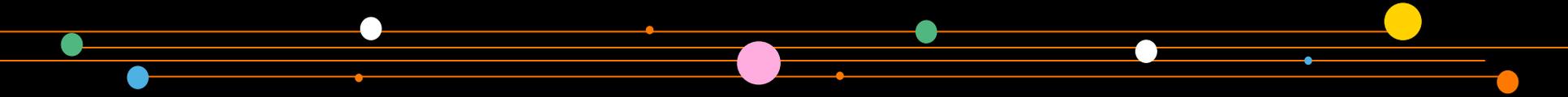
- Efficacité opérationnelle
- Connaissance Client
- Modernisation de l'architecture
- Gestion des risques

# Les défis à relever pour une stratégie data optimale

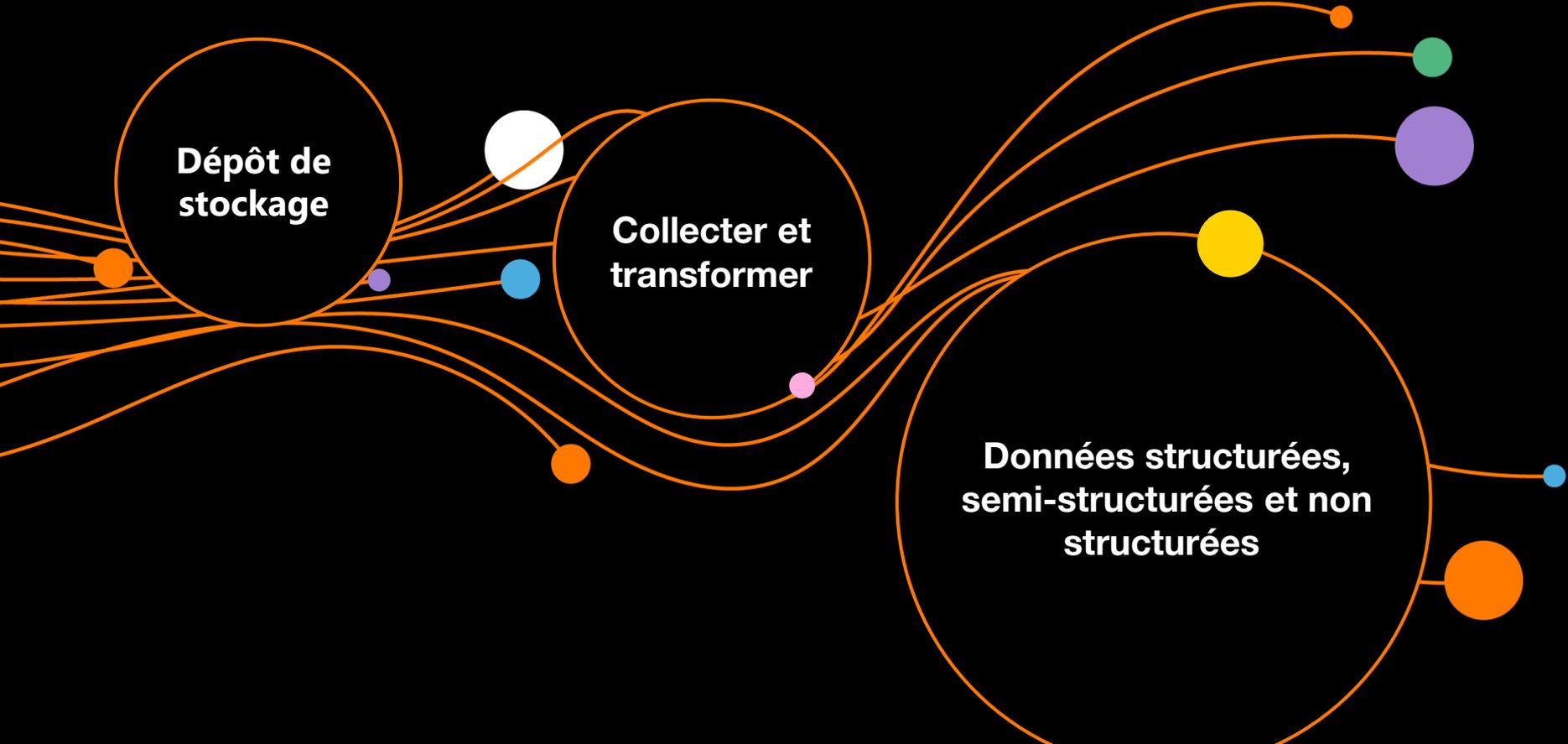
- || Présence de silos
- ||| Plusieurs SI ou/et bases de données
- Manque de visibilité
- ↻ Peu d'automatisation



# Approche traditionnelle du traitement de la donnée



# Qu'est-ce qu'un Data Lake ?



# Data Lake Insight

## Présentation



# Pourquoi mettre en place un Data Lake sur Flexible Engine ?

## Données variées :

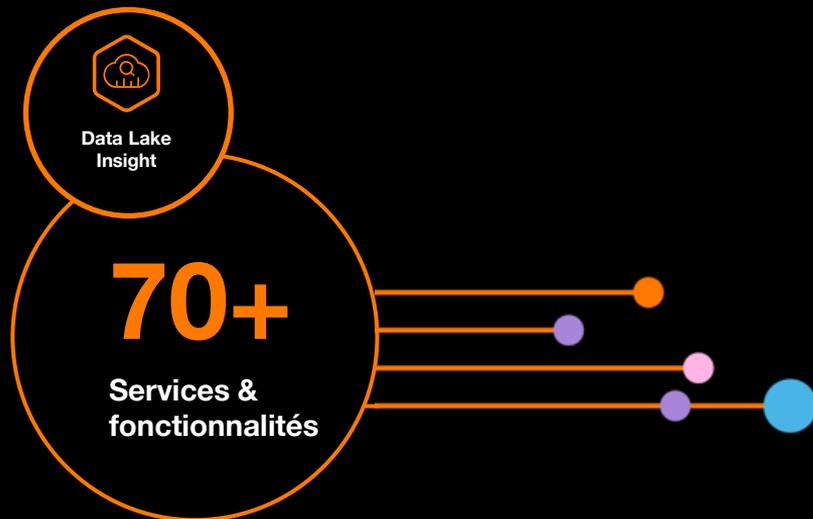
- structurées (base de données relationnelles)
- issues de bases NoSQL
- semi-structurées (csv, xml, json)
- non structurées (email, documents, pdf)
- de type blob, binary large object, (images, vidéos)

## 3 avantages principaux :

- optimisation des coûts
- un emplacement unique pour les données
- facilitation d'analyse des données



# Un service Big Data disponible depuis le cloud Flexible Engine



intégrés dans **8** Solutions

-  Migration vers le cloud
-  Sauvegarde & archivage
-  Virtual Desktop (VDI)
-  Diffusion de contenu média
-  Web & mobile DevOps
-  Big data & Entreprise Intelligence
-  Calcul haute performance
-  Internet des objets

Construit sur  
**Openstack**

**2 100** domaines clients

**500k** vCPU

**99,998%** de disponibilité  
compute en 2021

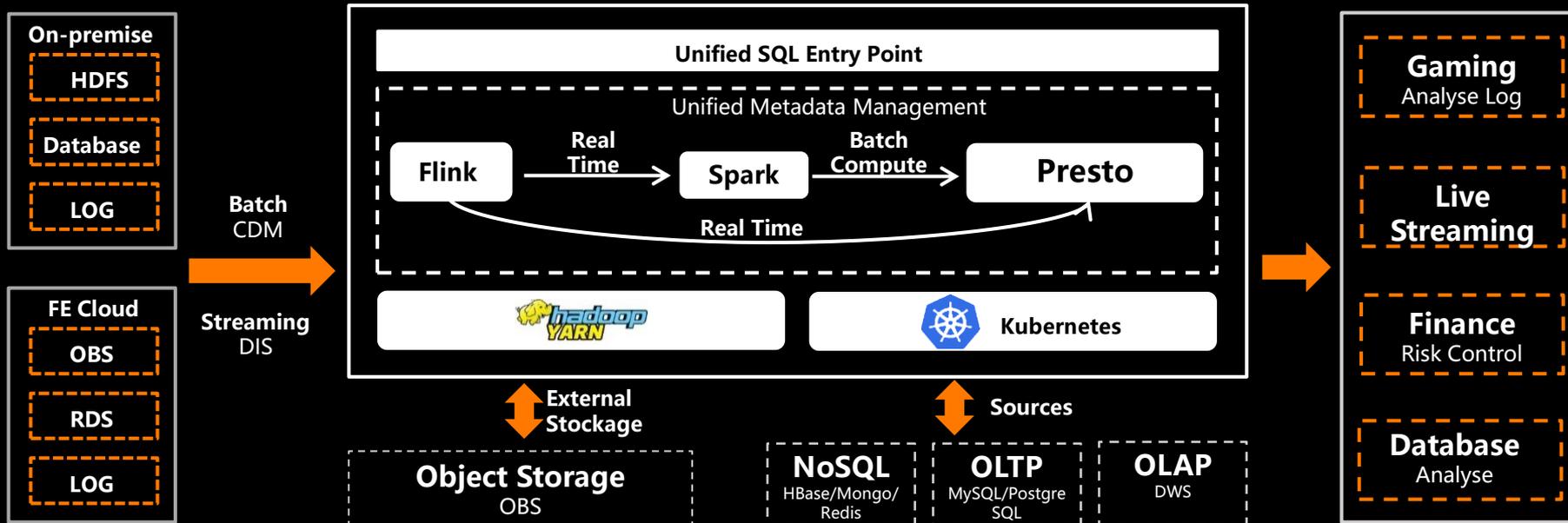
# Un écosystème de services d'analyse de données sur Flexible Engine



# Data Lake Insight

## Un service Serverless : Entièrement compatible avec l'écosystème Open Source

Data Lake Insight (DLI) = Apache Flink + Apache Spark + Presto



### All in SQL

Prend en charge des normes ANSI SQL 2003 et CEP SQL. SQL peut-être utilisé pour les analyses big data.

### Cross-Source Query

DLI peut requêter des données stockées sur OBS sans migration (utilisation des Metadata). DLI peut être utilisé pour requêter des données issues de sources variées

### Enterprise-class multi-tenancy

Prend en charge la gestion fine des autorisations d'accès aux ressources informatiques et des données par tenant.

# Data Lake Insight

## Cas d'usage



# Cas d'usage #1 : organisation en entreprise

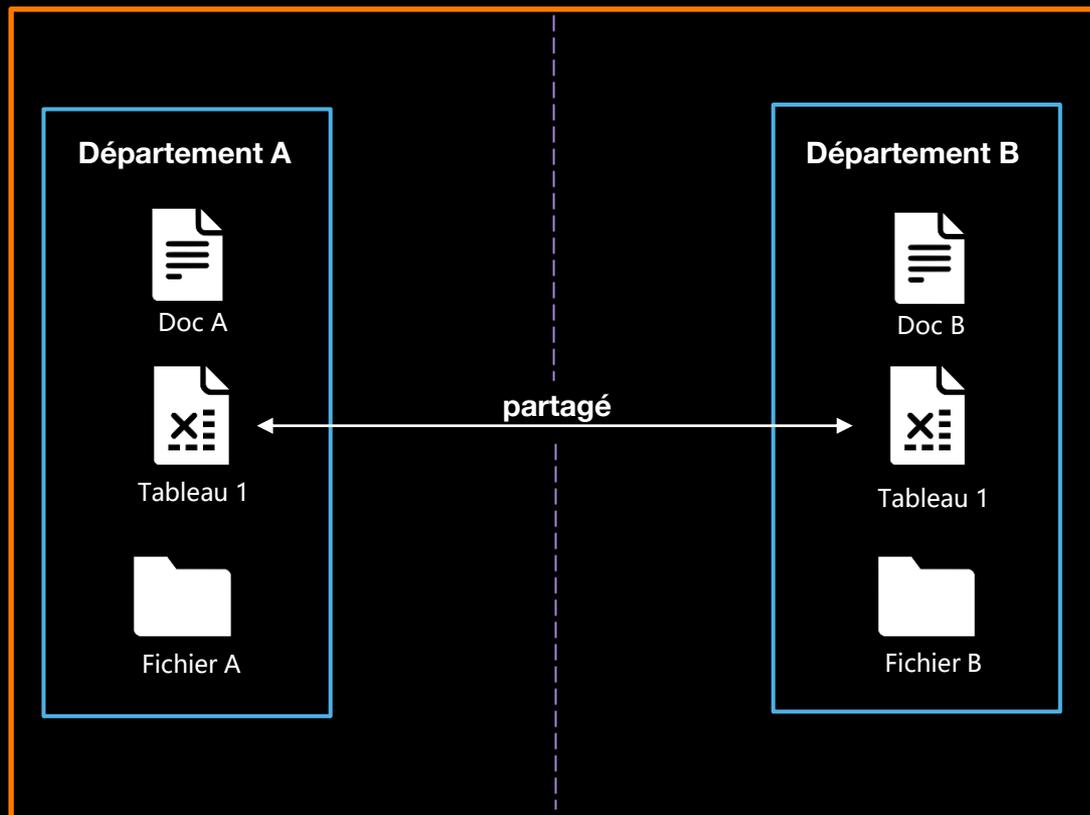


Data Lake  
Insight

Vous dirigez une entreprise, différents départements de la société peuvent avoir besoin de gérer les ressources de manière indépendante. Vous avez besoin d'une gestion fine des permissions pour assurer la sécurité des données et améliorer l'efficacité de la gestion

## Avantages :

- Gestion fine des autorisations
- Un seul compte IAM pour gérer les multiples employés de votre entreprise



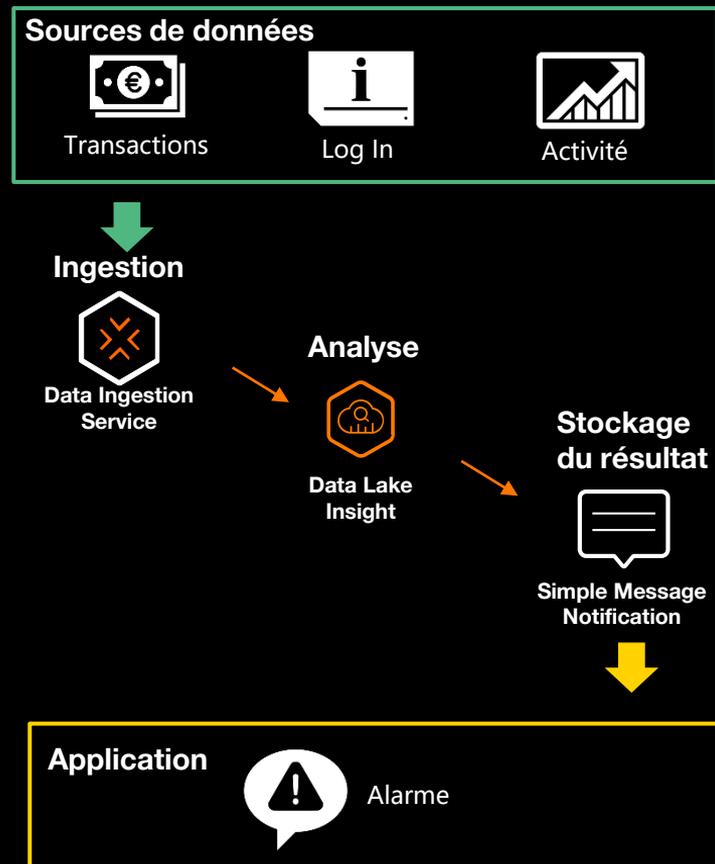
## Cas d'usage #2 : services financiers

### Contrôle du risque en temps réel

Presque tous les aspects des services financiers impliquent une certaine forme de contrôle des risques. Un système complet est nécessaire.

#### Avantages :

- **Analyse des données en temps réel** dans DLI à l'aide d'un modèle de flux de données Apache Flink. Une seule unité centrale peut traiter de 1 000 à 20 000 messages par seconde.
- Possibilité **d'enregistrer des flux** de données en temps réel vers plusieurs services de cloud computing comme SMN.



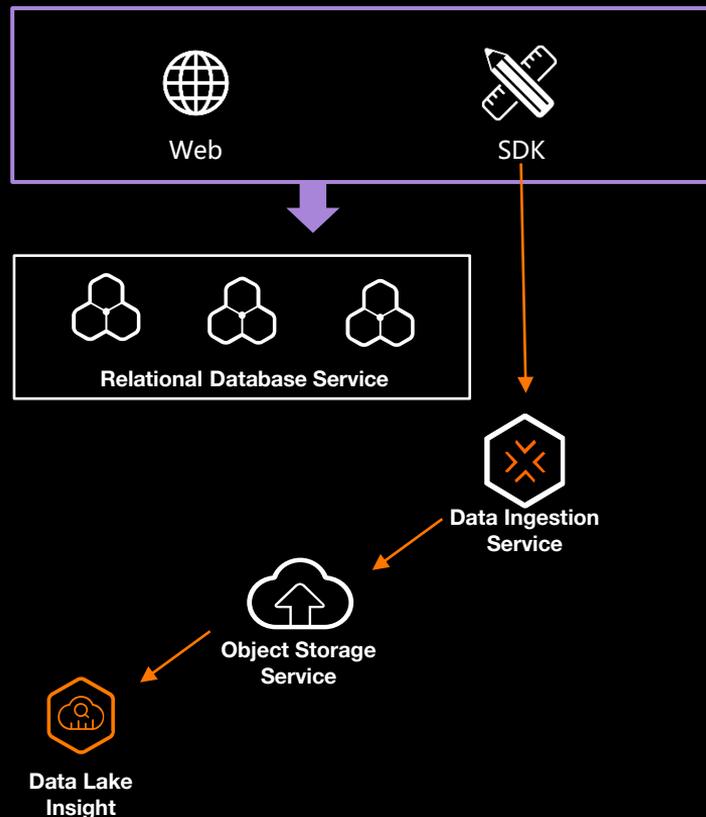
# Cas d'usage #3 : e-commerce

## Marketing de précision

Les informations provenant de plusieurs canaux doivent être combinées pour une analyse combinée afin d'améliorer le taux de conversion.

### Avantages

- **Analyse en croisant** des sources de données
- Possibilité de créer un tableau à l'aide de simples **instructions SQL**





# Démonstration

# Data Lake Insight



À emporter

# Quelques bonnes pratiques

1. **Explorer** les données
2. Simplifier et normaliser l'**accès** aux données
3. Développer un **catalogue**
4. Mettre en place des **droits** d'accès aux données
5. **Evangéliser** en interne



# Merci



orange™  
Business  
Services

