

Transformez vos données IBM i en leviers stratégiques

La Réplication de données



L'imagination et l'expertise au cœur
de vos infrastructures et solutions

Francois Martin
+33 6 73 86 99 85

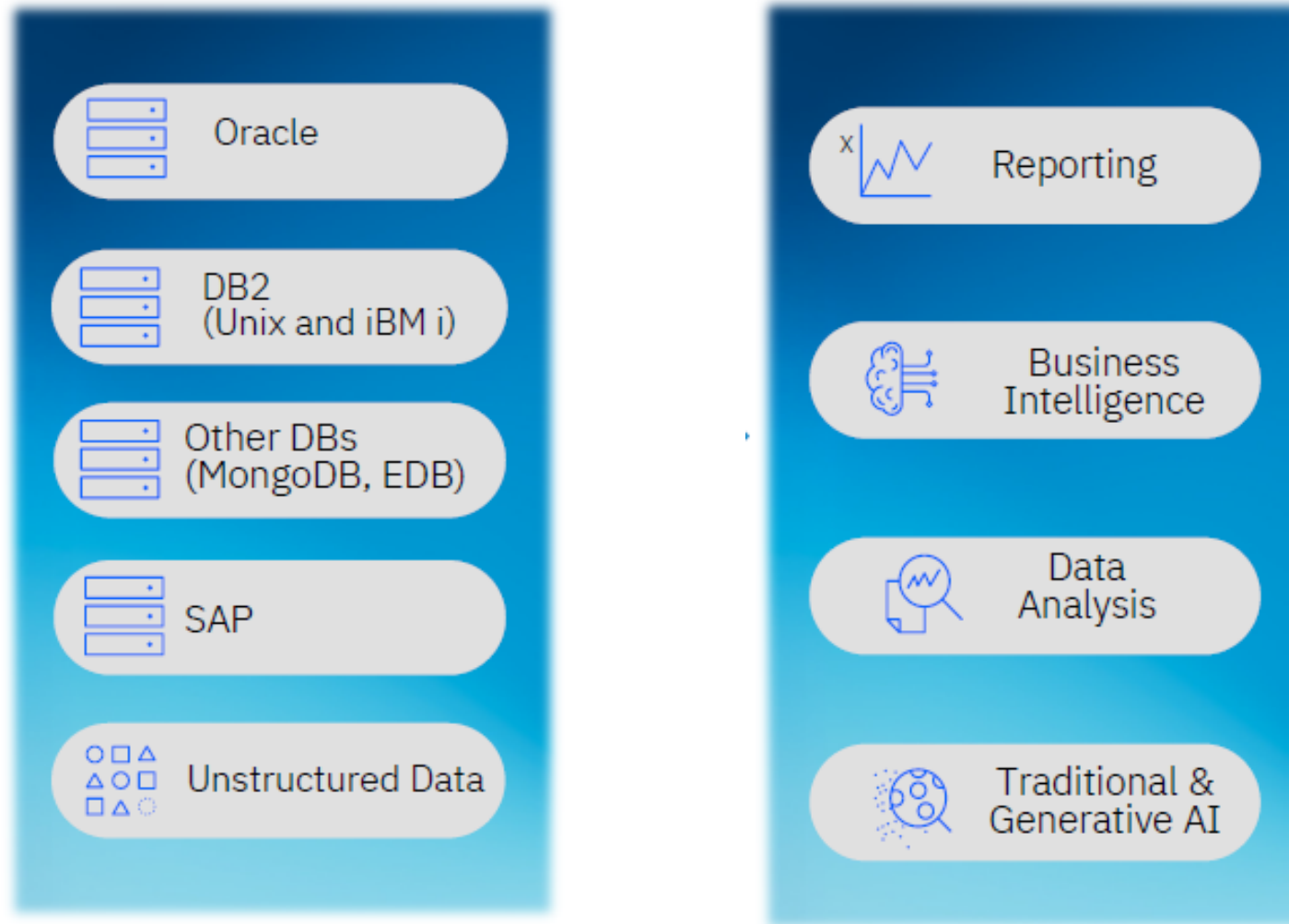
fmartin@acmi.fr

Octobre 2025

- Données, Bases de données, Réplication et Valorisation des données
- Types et besoins de Réplication
 - Réplication PRA, logicielle ou logicielle et stockage
 - PRA pour IBM i avec Mimix ou QuickEDD
- Cas d'usage: Réplication applicative (outils CDC temps réel ou ETL mode batch)
 - Data Lake (Cas le plus courant Analytique/Big Data, ex: Snowflake),.
 - Kafka (déclenchement évènement)
 - Copie pour site Web (vers SQL/NoSQL)
 - Migration IBM i vers autre plateforme
 - Solutions de réplication IBM i proposées par ACMI
 - Precisely CDC Connect / DI Suite
 - IBM IDR
 - GlueSync
 - Syniti /GammaSoft (fin de commercialisation/support)
 - Un exemple client pour une Mise en production rapide de modifications
- ACMI votre partenaire réplication

La valorisation de la donnée

Sources de données → Consommateurs de données



Exemples pour l'analyse exploratoire:

- Visualiser la corrélation entre le prix et la quantité vendue.
- Identifier les pics de vente liés à la météo
- Détecter des valeurs aberrantes, grouper des clients par comportement d'achat

Exemples pour la Business Intelligence:

- Création d'un tableau de bord des ventes
- Analyse du pipeline commercial et du comportement client
- Suivi des campagnes publicitaires
- Évaluation RH des performances annuelles et des ressources
- Des analyses géospatiales, temporelles, ...

Challenges:²

- Gérer le volume et la diversité croissants des données
- Simplifier la connectivité des données
- Garantir la qualité des données grâce à une gouvernance intégrée
- Optimiser les performances et les coûts

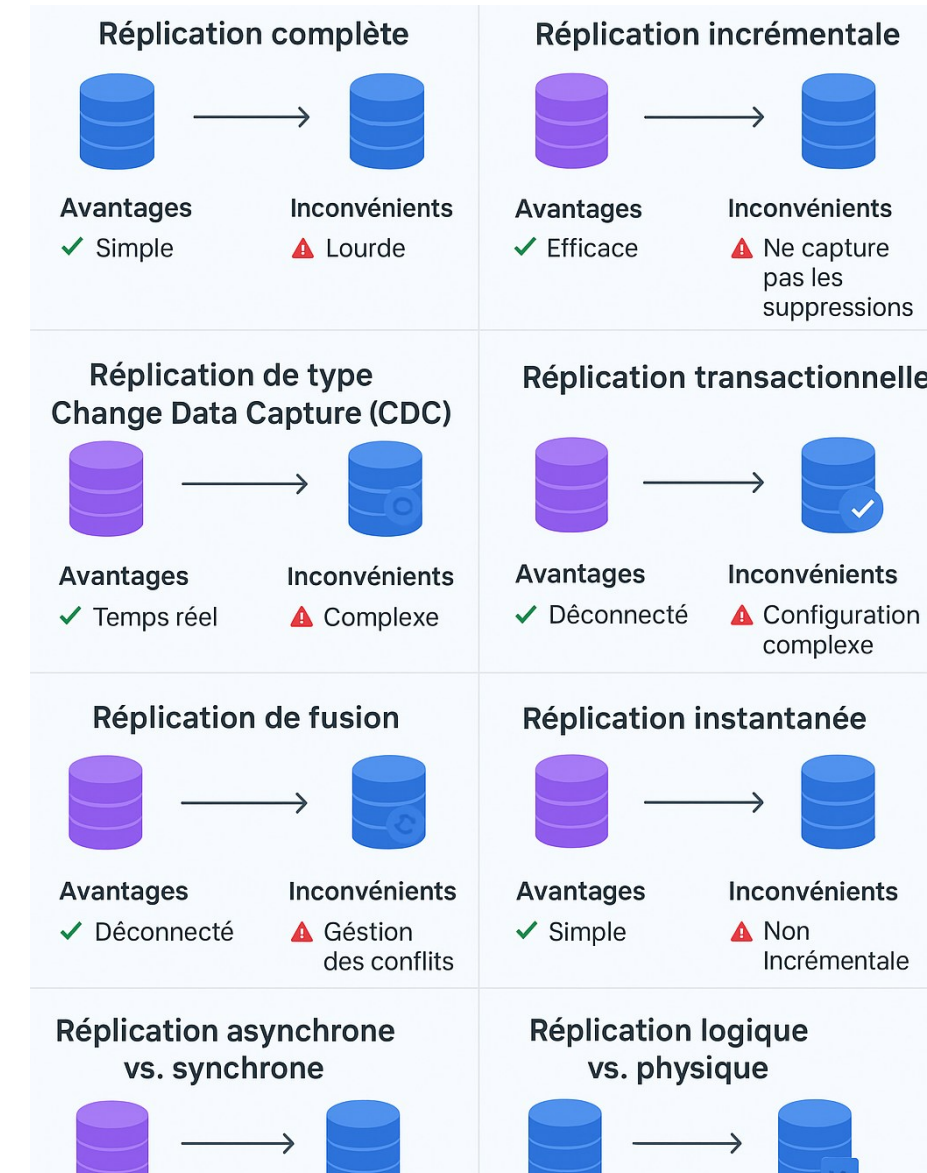
La réplication s'applique aux bases de données, et aussi à d'autres objets de données dans le cadre de l'IBM i

Les besoins peuvent être différents selon qu'il s'agit de:

- Backup, copies pour CyberSécurité ou clones fréquents pour servir de PRA avec perte
- Réplication pour du PRA
- Réplication « applicative »
 - copier les données vers d'autres applications, consolider plusieurs sources, alimenter des applications analytiques.
- Réplication pour un besoin spécifique de production ou de test

Les technologies:

- Copie snapshot mais n'est pas de la réplication, c'est plutôt orienté backup
 - « photo » complète des données à un instant T et appliquée sur l'environnement cible. (copie de blocs physique ou logique)
- Réplication complète et continue
 - synchronisation, transfert des données, modifiées (ajoutées ou mises à jour), ou transfert basé sur un horodatage selon la technologie
- Réplication CDC « Change Data Capture »
 - Capture temps réel les opérations INSERT, UPDATE et DELETE directement à partir du journal de transactions (Base de données)
 - CDC n'est pas une technique d'ETL mais peut parfois être utilisé pour quelques transformations (ex: champs de date) en évitant de perturber la réplication temps réel, ou alors en mode batch (vacation)



La copie snapshot avec stockage : ce n'est pas de la réplication

La copie de blocs de données est souvent physique au niveau du stockage,
parfois logique au niveau de la base de données

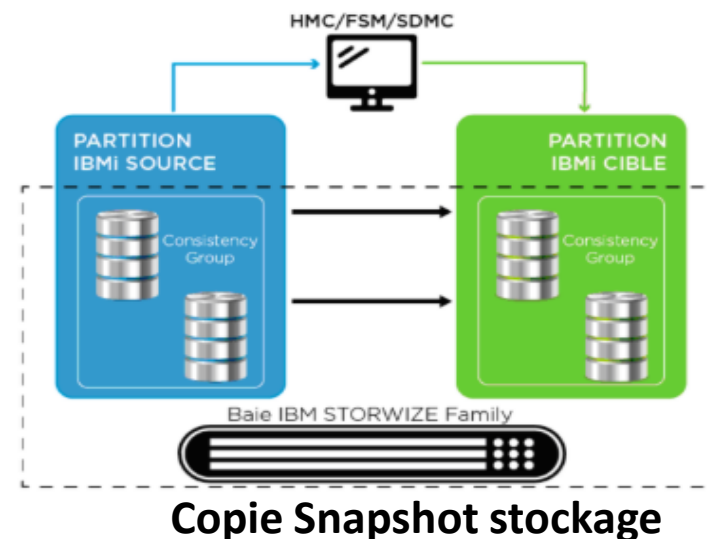
- La copie au niveau stockage et une méthode de backup, de migration ou de clonage
 - ce n'est pas de la réplication de données
- Basée sur une référence temporelle, « Photo » à un instant t, pas de donnée « temps réel »
- Peut être « lourde » en fonction du volume de données snapshot et selon la technologie de stockage et les options
- FLASH for i (M81) permet de faire un snapshot destiné aux sauvegardes ou de environnements de test
- Migration d'IBM i d'un stockage vers un autre stockage (avec arrêt): 2 méthodes supportées et 1 méthode non supportée

Migrating IBM i Data from Current Disk Storage to Different External Disk Storage

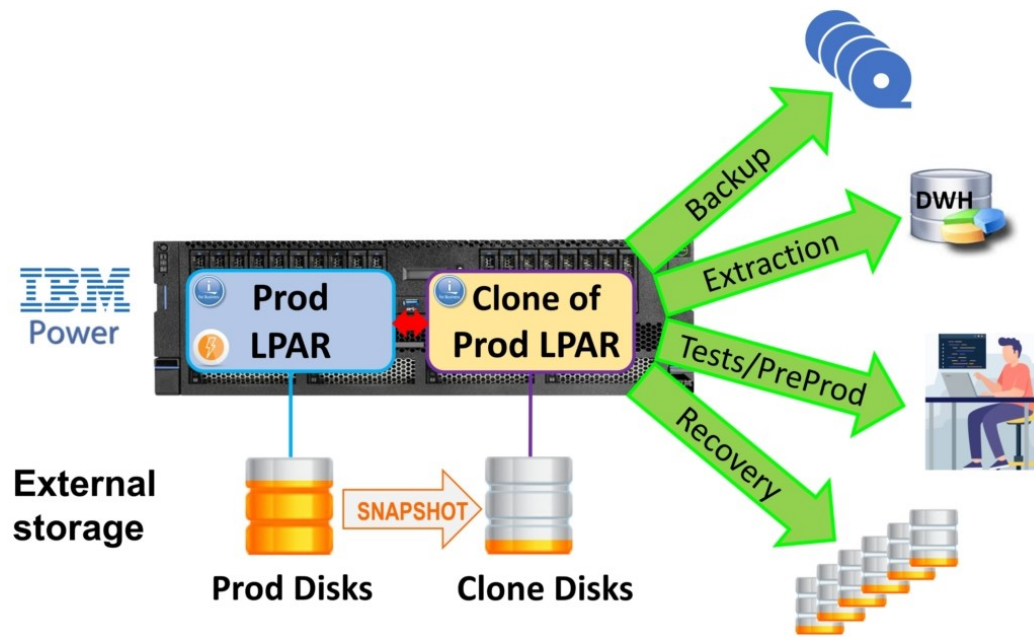
<https://www.ibm.com/support/pages/migrating-ibm-i-data-current-disk-storage-different-external-disk-storage>

- Migration vers nouveau server stockage: scratch/rechargement est la seule méthode supportée (option A)
- Migration vers nouveaux disques (option B)
- Copie/Clone via le stockage externe, pas supporté sur IBM i

Flash for i (M81)



Le produit FLASH for i a pour but principal de réduire le temps de sauvegarde ou créer des environnements de tests en quelques minutes en dupliquant une partition IBM i



Besoin:

- Une copie de toutes les données IBM i pour un redémarrage (plan de reprise) en cas de désastre de l'environnement primaire
 - y compris les tables et objets non journalisés / fichiers IFS
- Une réplication proche du « temps réel » afin d'obtenir un RPO proche de zéro. (RPO=0, pas de perte de transaction)
- Des mécanismes automatiques de correction pour maintenir la cohérence, audits de synchronisation
- Un processus de bascule (« switch over » et « switch back ») personnalisables, automatisables (en mode batch ou interactif)
- Une faible consommation CPU (quelques %) et une optimisation du trafic réseau (compression, minimisation des envois)

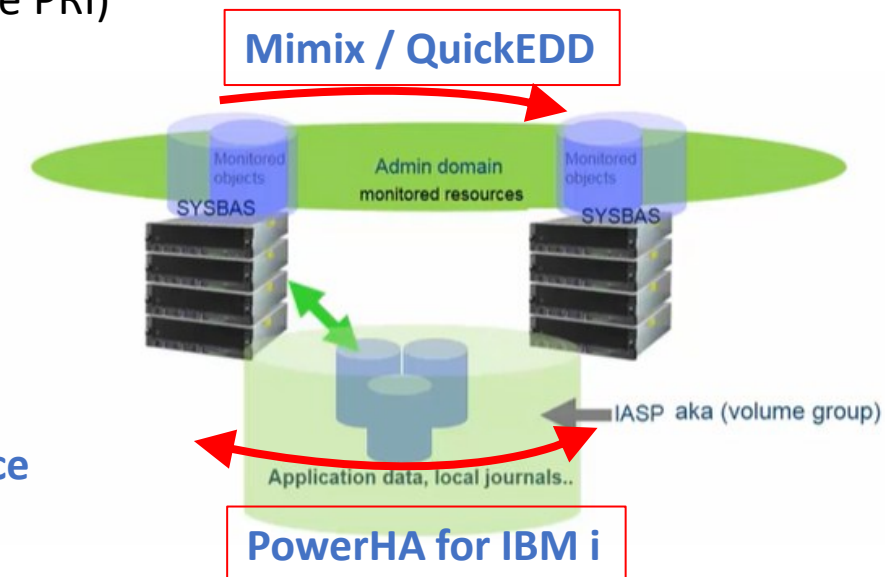
Eventuellement:

- Le support de topologies multiples (1 à n serveurs IBM i) et la compatibilité entre différentes versions matérielles et IBM i
- La possibilité de cryptage, d'anonymisation des données sensibles avant réplication (exigences de confidentialité RGPD)
- Des copies immuables en cas de cyber-cryptage. (Plusieurs copies horodatées, on parle de PRI)

Solutions: toutes les tables, objets, IFS, doivent être copiées à l'identique

- Logicielles: Mimix ou QuickEDD. ACMI gère ces 2 solutions (Precisely)
- Logicielles et matérielles: PowerHA for IBM i
- Basés sur un cluster de stockage partagé avec réplication
 - il n'y a qu'une seule copie logique des données et non pas 2 copies)

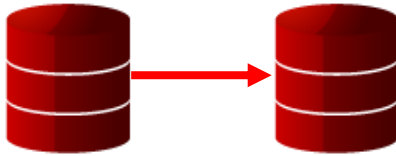
Ces solutions répondent à des besoins de haute disponibilité, reprise d'activité ou maintenance concurrente. Ce n'est pas de la réplication « applicative » pour alimenter une cible avec des données spécifiques



Les options de réplication permettent des implémentations flexibles afin de réduire les duplications, simplifier la gouvernance des données, garantir la qualité, et optimiser les coûts

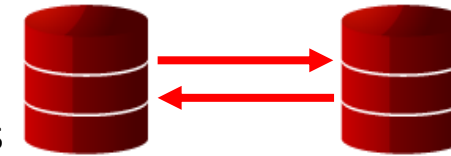
Sens Unique

PRA, Analytique, ..



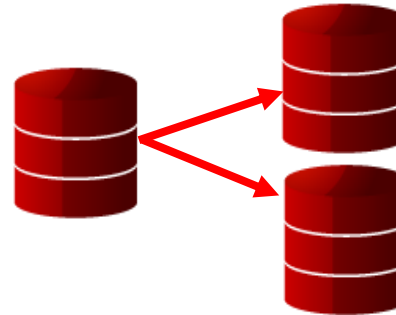
Bidirectionnel

Utilisateurs actifs sur plusieurs sites



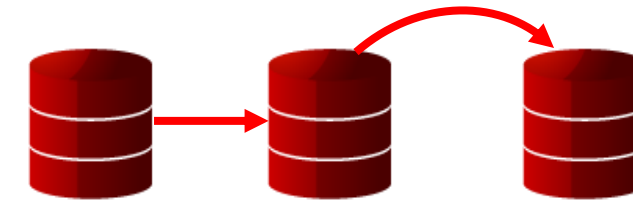
Distribution

Alimenter plusieurs applications (ERP, web)



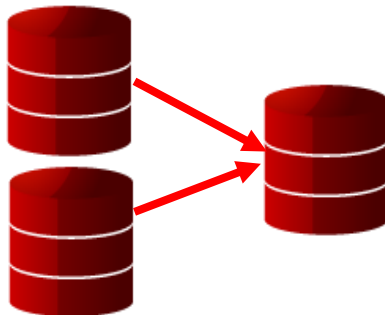
Cascade

tests, backups, répartition géographique



Consolidation

Données de plusieurs agences vers siège central



Choix d'une topologie ou d'une combinaison de plusieurs pour satisfaire les besoins de partage de données

La réplication logicielle « applicative », le principe du CDC

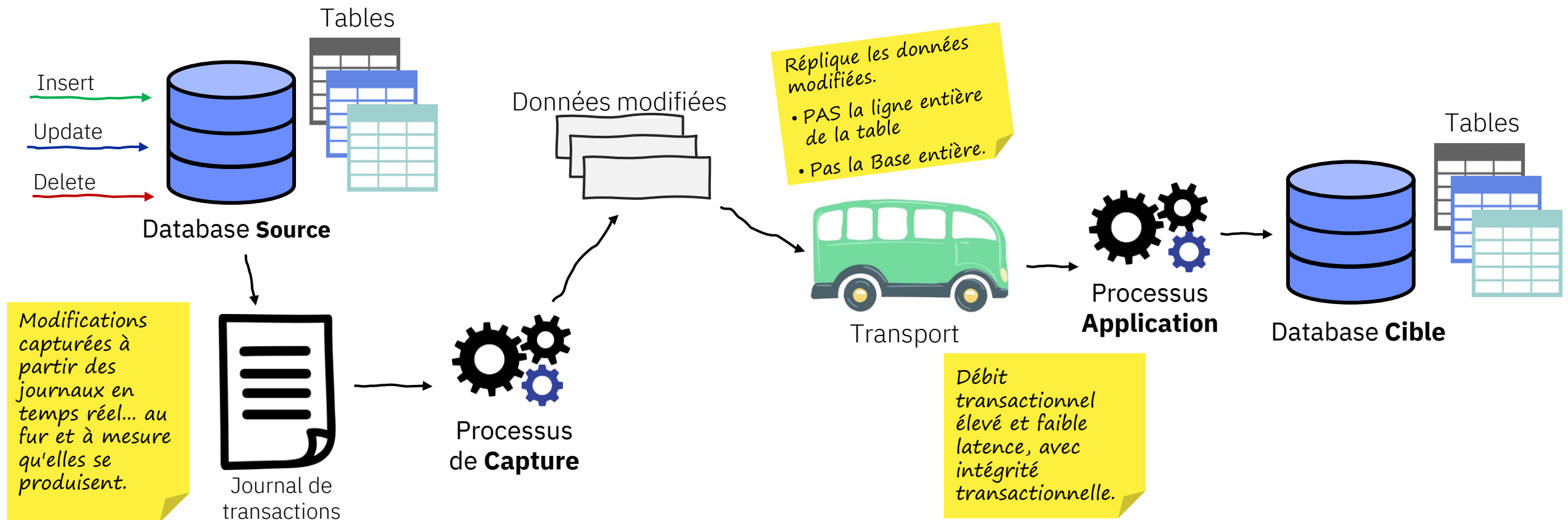
Le besoin est fonctionnel par rapport au contenu et à l'usage des données, différent du besoin PRA

La réplication capture les opérations directement au niveau de la base de données à partir du journal de transactions

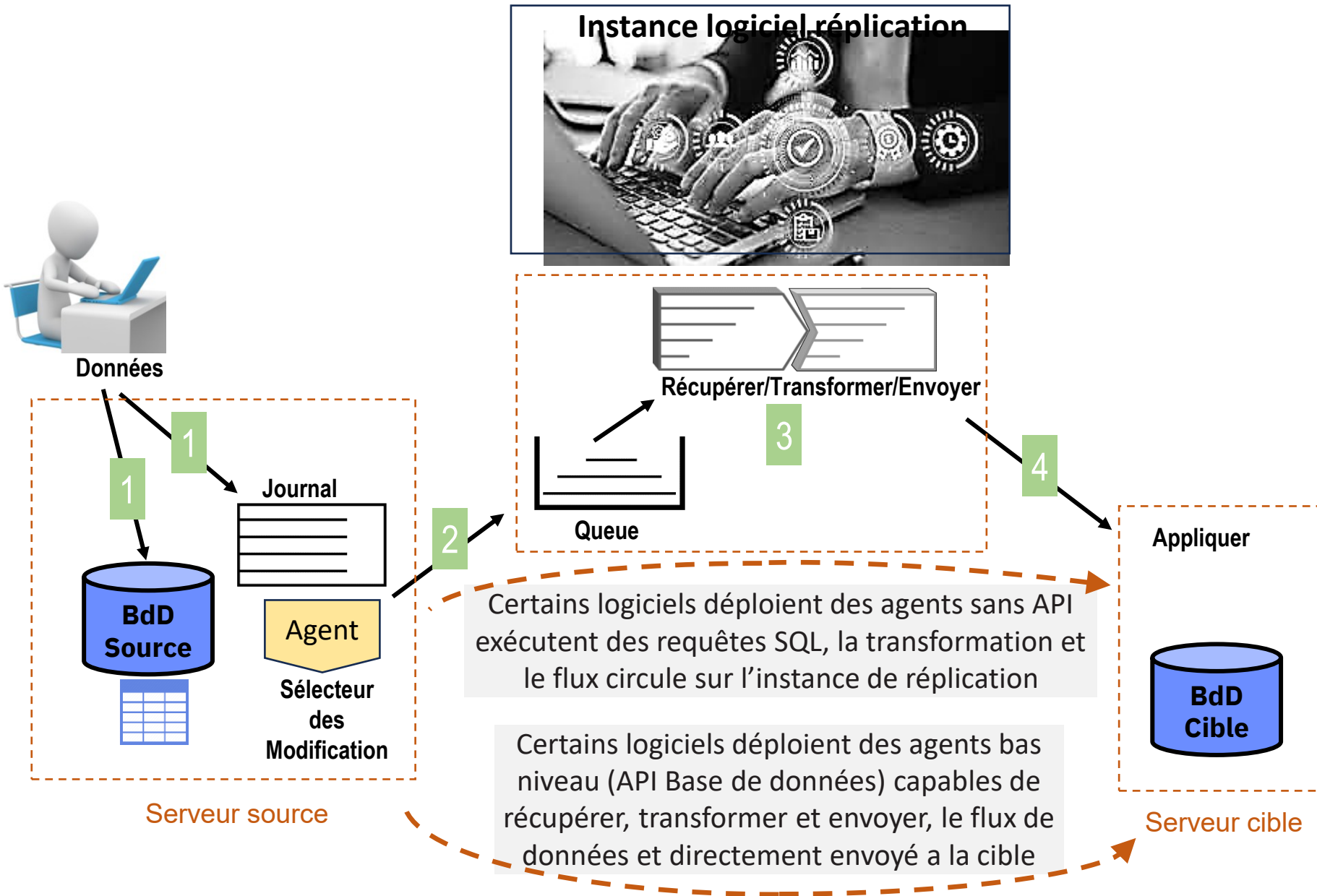
Pourquoi au niveau de la base de données: CDC – Change Data Capture

- Elle permet de travailler en « temps réel » plutôt qu'en mode incrémental ou batch basé sur un horodatage
- Elle permet d'alimenter des bases de données avec n'importe quel type de moteur

Vue conceptuelle du fonctionnement de la réplication des données



Principe de fonctionnement de CDC sur la base d'un journal: le logiciel déploie des agents sur la source et sur la cible



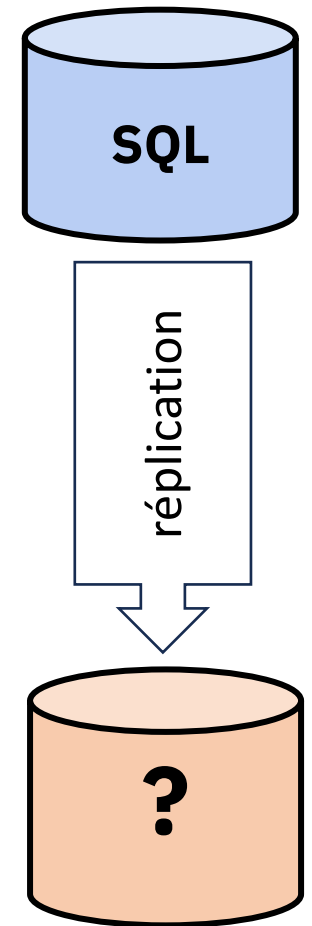
- L'utilisation de journaux de transactions ou de déclencheurs élimine la nécessité d'actions invasives sur la base de données
- Des extraits sélectifs des journaux et un espace d'attente défini garantissent l'intégrité des données.
- Dans de nombreux cas, la transformation peut être effectuée en dehors de l'environnement source et peut réduire l'impact sur la production
- Le processus d'application renvoie l'accusé de réception à la file d'attente pour compléter le commit en deux phases.

La réplication CDC: plusieurs **types de moteurs de base de données** **pour l'Analytique, le reporting, l'IA, ... autres applications**

Les données sont « consommées » de manières différentes, il existe des types de moteurs en fonction de l'usage

➔ Les cibles de la réplication sont multiples et exploitent les données en fonction de l'application cible

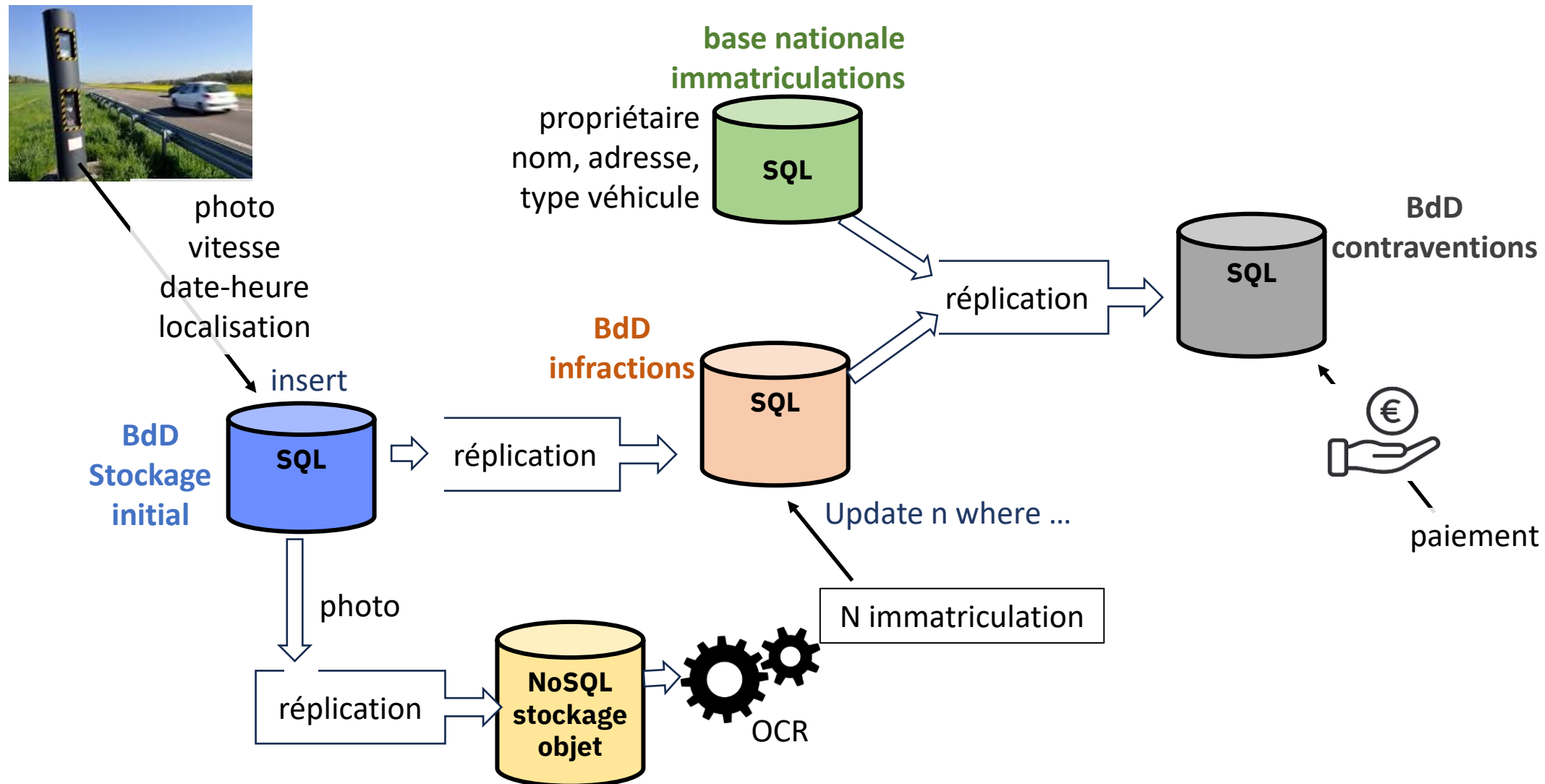
- Bases de données **relationnelles**
 - Données organisées en tables avec lignes et colonnes, liées par des clés
 - Language SQL
- Bases de données orientées **documents**
 - NoSQL (langage JSON, ...)
 - Données non structurées, environnements web, big data, micro-services, ..
- Bases de données orientées **colonnes**
 - Langage forme de SQL (BigQuery SQL, Cassandra Query Language,)
 - Big Data, Data warehousing, lecture rapide de données agrégées
- Bases de données orientées **graphes**
 - Données modélisées sous forme de nœuds (entités) et arêtes (relations) Ex: Neo4j
- Bases de données orientées **séries temporelles**
 - données chronologiques (IoT, monitoring,..) ex: Prometheus
- Bases de données **géospatiales**
 - coordonnées, cartes, et données spatiales (ex PostGIS)
- Bases de données **distribuées / cloud-native**
 - Données répliquées sur plusieurs nœuds (scalabilité)
- Bases de données « **mixtes** », Orientée colonnes + Cloud-native / Distribuée + SQL standard
 - analytique massif, agrégations, data warehouse, dans le Cloud, scalable (stockage séparé du calcul), SQL
 - Ex: Snowflake



Traitement de donnée, types de moteurs et réplication CDC

Les chaines de traitement des données utilisent plusieurs bases de données avec différents types de moteurs pas nécessairement SQL

Chaîne de traitement de l'information des radars autoroutiers

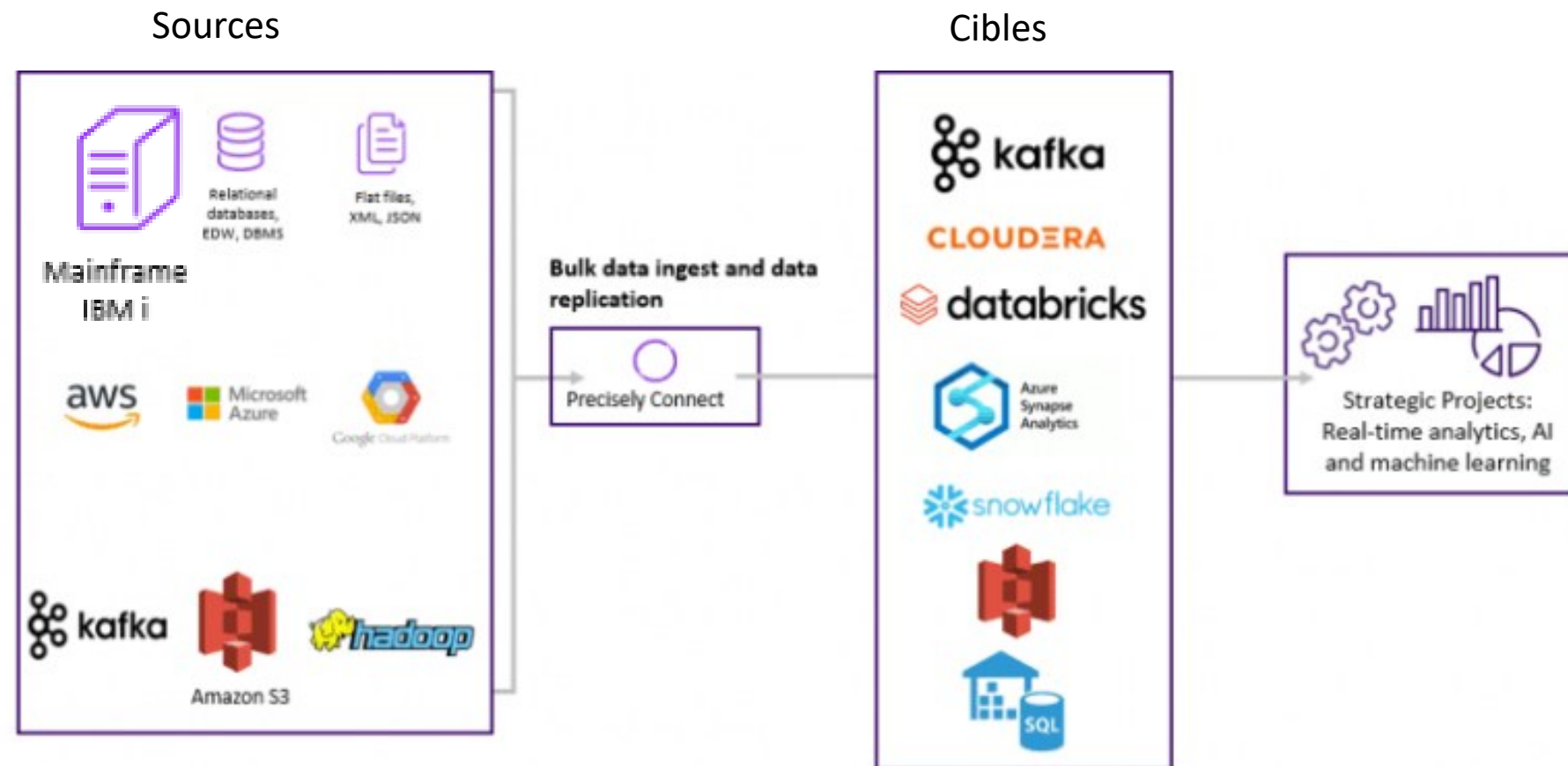


De nombreux logiciels de réplication sont disponibles sur le marché, ACMi maîtrise plusieurs logiciels capables de s'adapter à votre besoin, afin de mettre en place une solution durable et flexible

- ✓ Sinity: arrêt des versions, fin de support juin 2026 ???
- ✓ Connect CDC (ex Mimix share)
- ✓ DI Suite (Precisely)
- ✓ IBM Data Replication
- ✓ Gluesync (Molo17)

Exemples de projets ACMi

- IBM i vers IBM i
- Migration IBM i vers autre plateforme
- Data Lake
 - (Analytique/Big Data, ex: Snowflake),
 - SQL vers NoSQL
- Kafka (déclenchement événement)
- Copie distribuée: site Web...



Connect CDC : un partage automatisé des données : les points clés

Supprime les barrières entre les bases de données

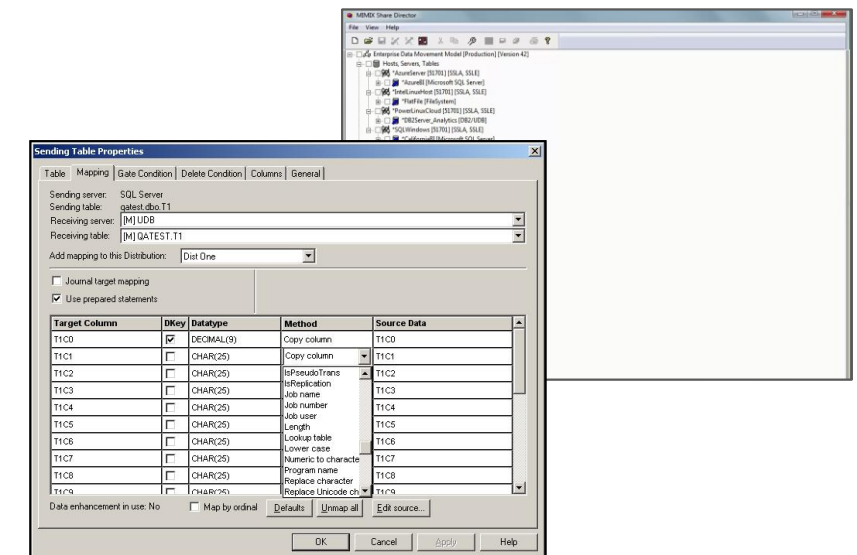
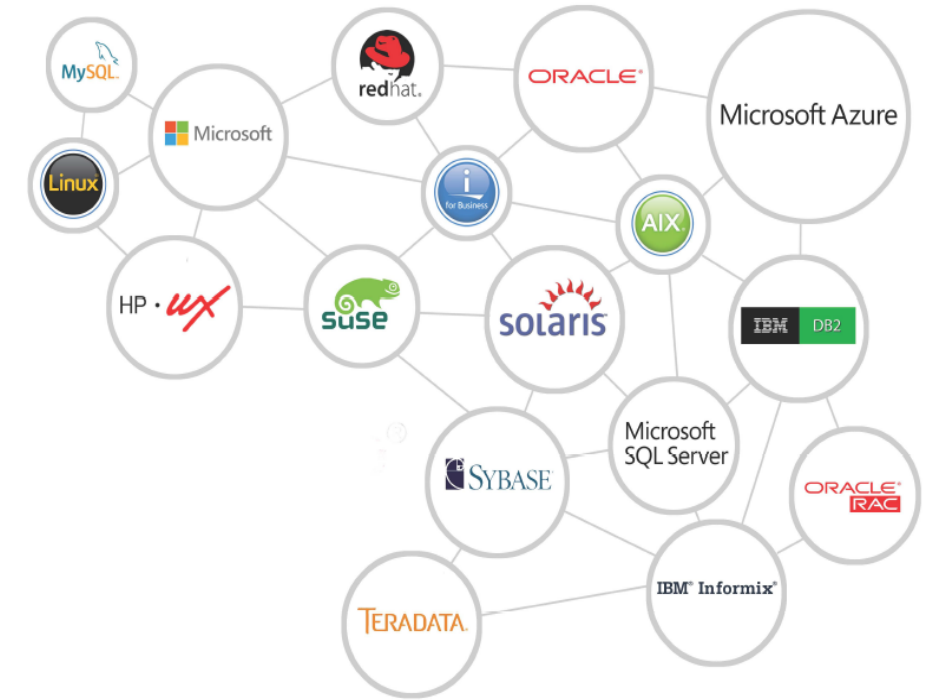
- Systèmes de gestion de bases de données identiques ou différents
- Systèmes d'exploitation identiques ou différents
- Plateformes physiques, virtuelles ou en cloud
- Quelle que soit la distance

Facilite le partage des données

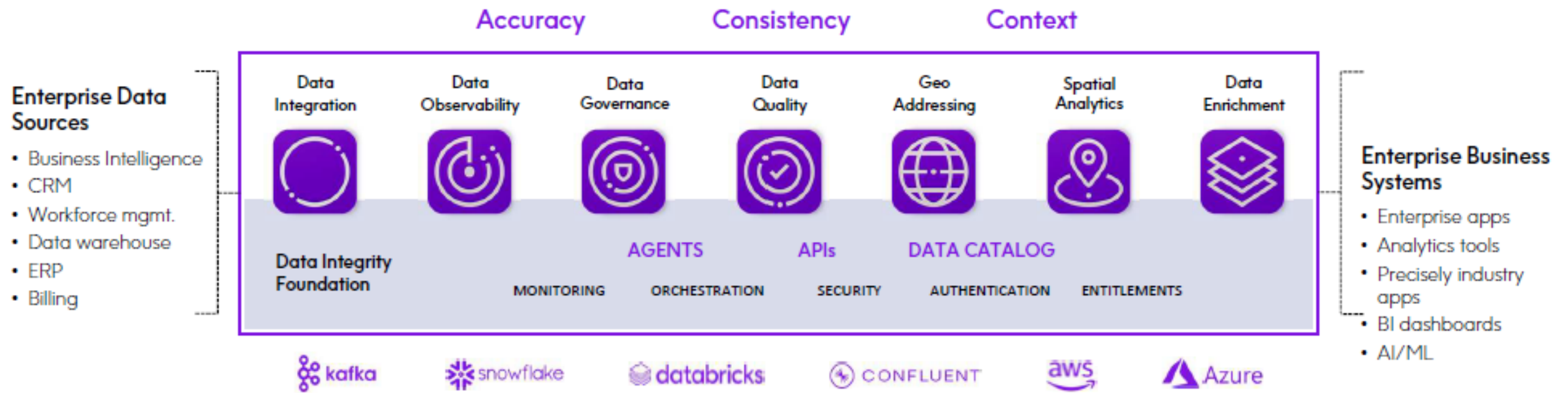
- Réplique les modifications apportées à la base de données en temps réel
- Transforme et améliore les données pendant la réplication
- Prend en charge les principales bases de données et les principaux systèmes d'exploitation
- Offre une variété d'architectures de réplication
- Interface graphique simple - aucune programmation n'est nécessaire !

Retour rapide sur investissement

- Prise de décision plus solide
- Productivité accrue de l'entreprise
- Possibilité de choisir une infrastructure plus rentable
- Libère l'informatique pour qu'elle se concentre sur d'autres initiatives de l'entreprise



Data Integrity Suite

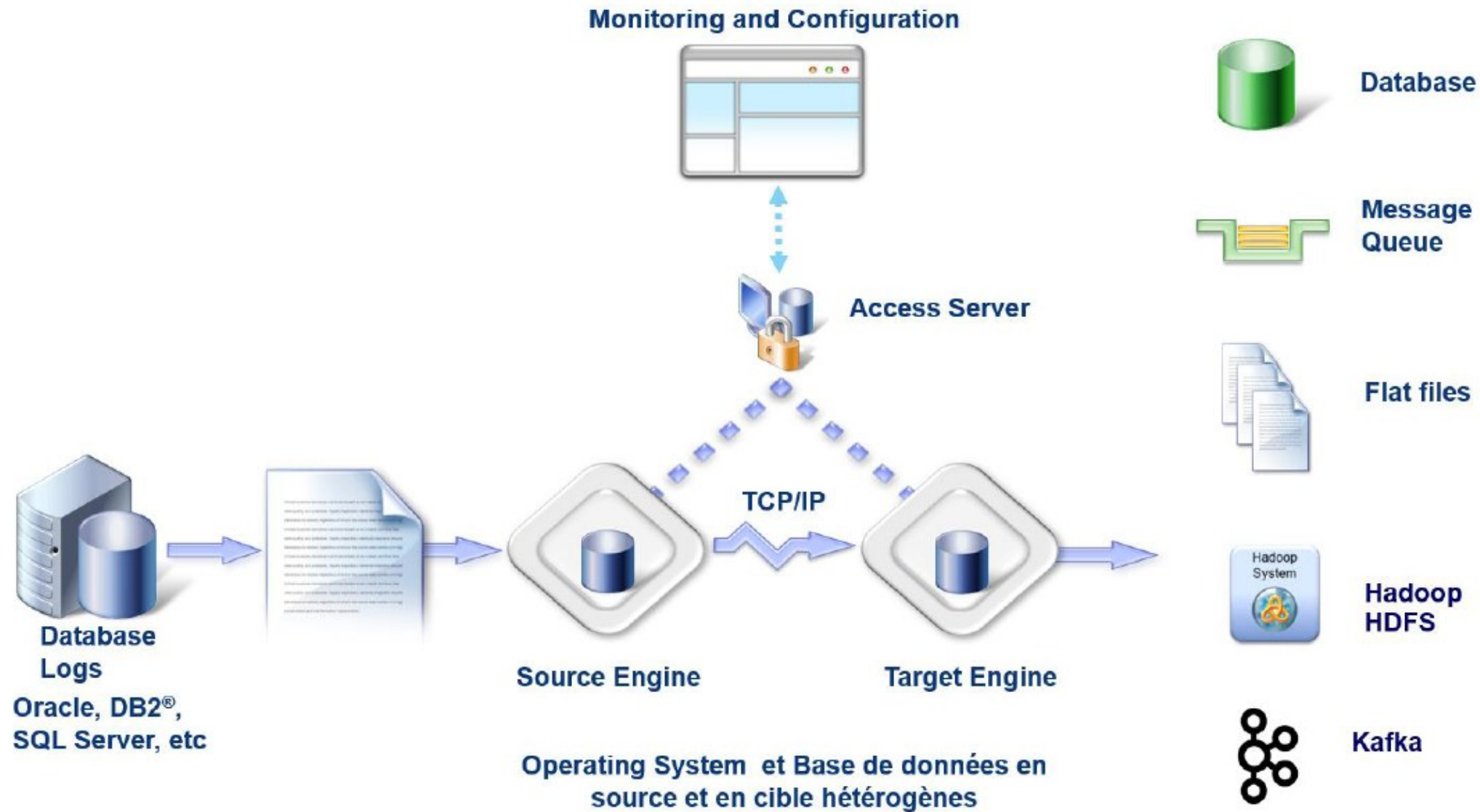


Solution en mode SaaS: « Design and Monitor in the Cloud »

- Connection au cloud.precisely.com (utilisateur/password ou SSO) pour accéder au module d'intégration des données
- Créer et modifier les configurations du pipeline de réplication de données dans l'interface utilisateur du navigateur
- Déployer les configurations dans l'environnement utilisateur pour l'installation des agents
- Démarrez votre pipeline de réplication de données
- Monitorer le statut et les statistiques pour la surveillance et le reporting

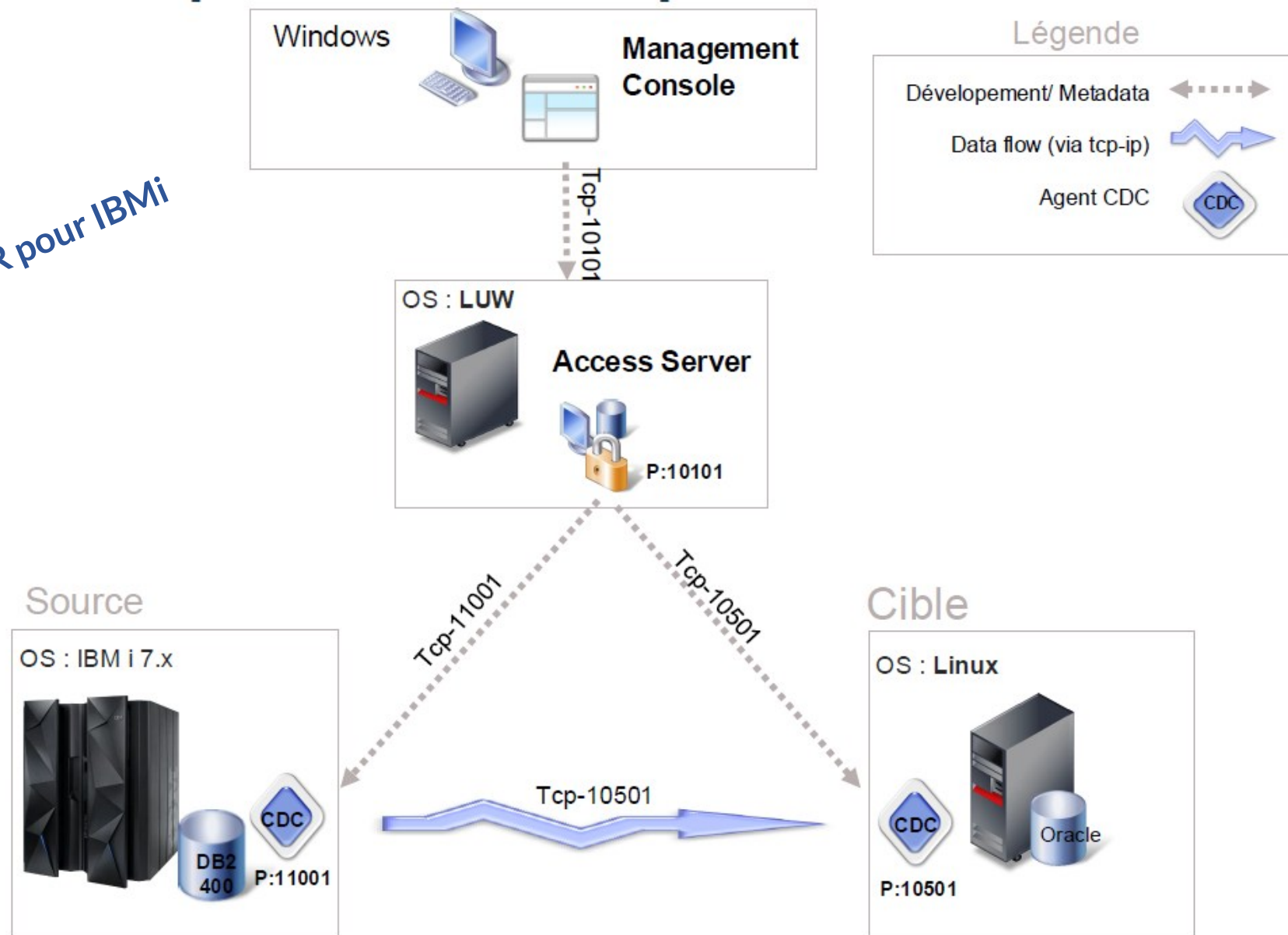
IBM Data Replication

Architecture

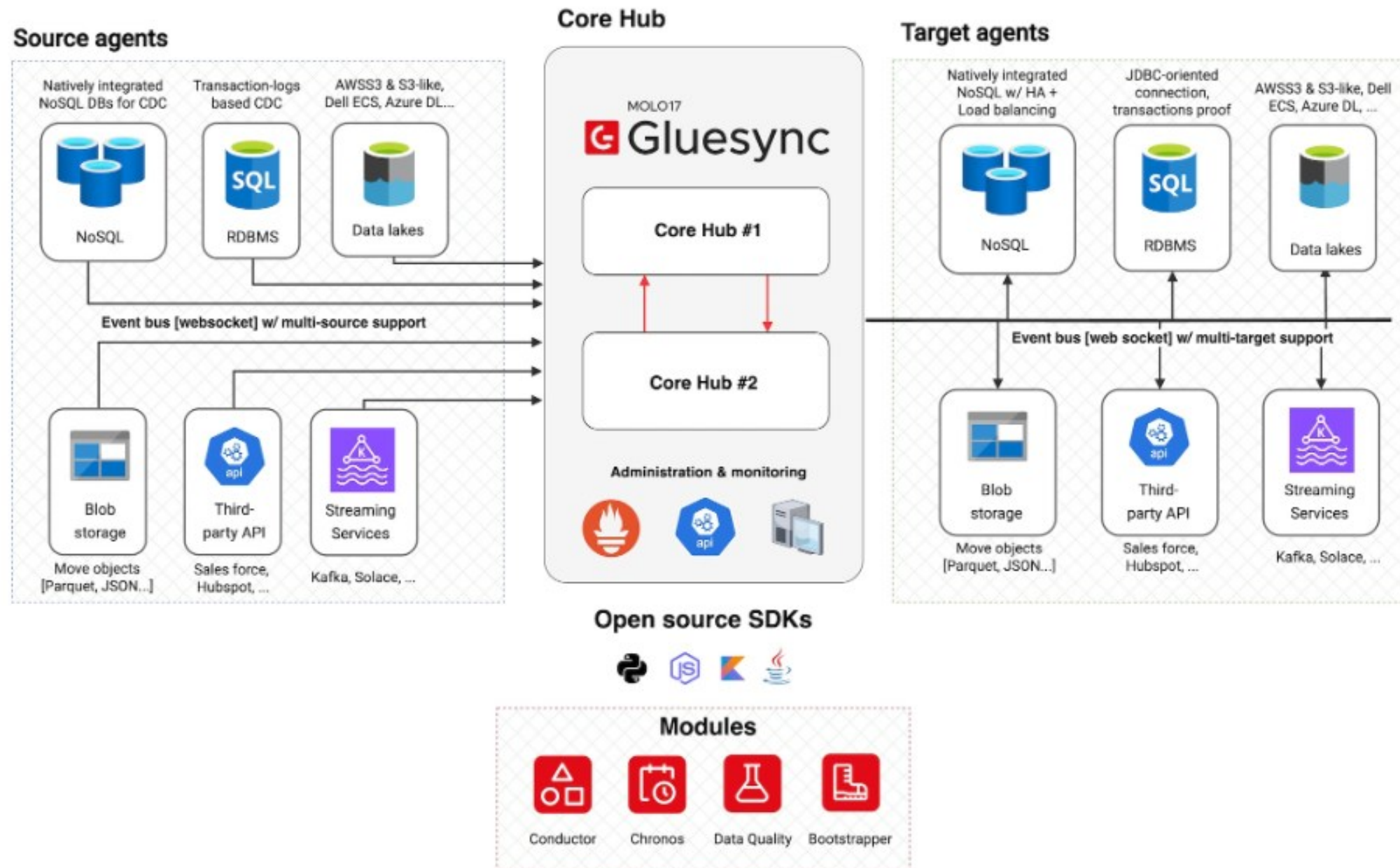


IBM Data Replication

Exemple d'architecture IDR pour IBMi



Gluesync (Molo17), Architecture

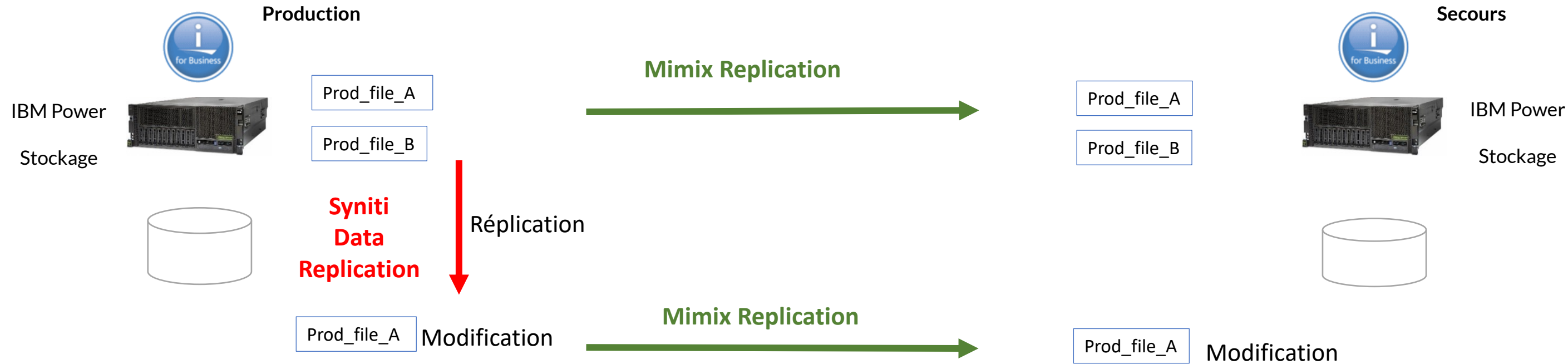


5 composants principaux

- Agents sources : Gère l'extraction des données et la capture des modifications
- Core Hub : Orchestre le flux de données et les opérations système
- Agents cibles : Gère le chargement et la transformation des données
- SDK : Permet l'intégration des développeurs avec Core Hub
- Modules : Étend les fonctionnalités de la plateforme grâce à des composants basés sur le SDK

Un exemple client, un peu différent

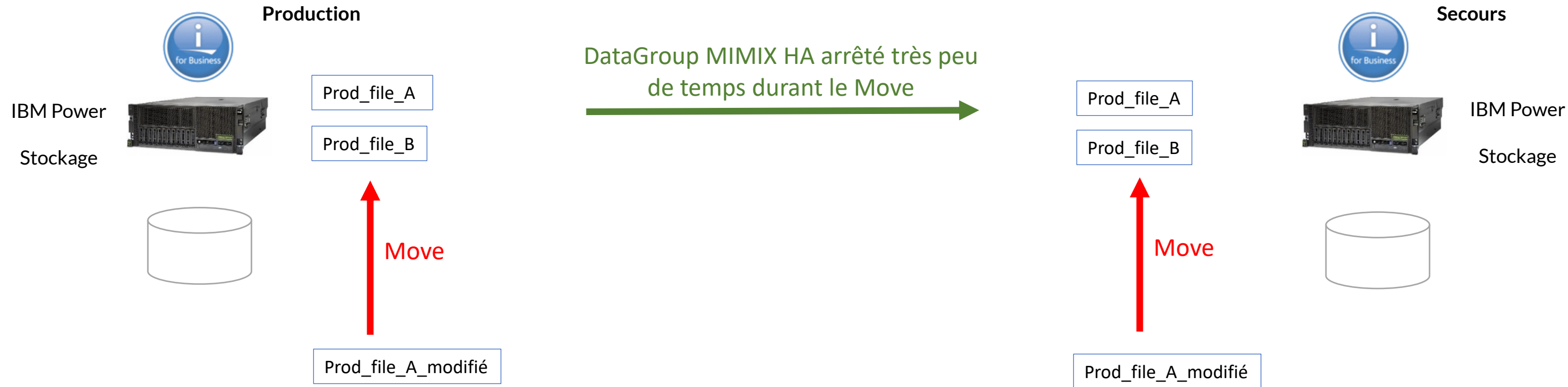
Solution Data Réplication coté production



- Remarques : Data Replication peut être installé coté prod ou coté secours
- Technique Data Replication :
 - Sur la Source : via la journalisation DB2/400 (en *BOTH)
 - Sur la Cible : application des modifications en ODBC.
 - Moteur et interface graphique pour la paramétrage : sur une VM Windows.
- Charge réseau à surveiller

Réplication
puis
Modification

Réplication: Mise en Production des modifications



Avantage : Temps d'arrêt limité à l'arrêt applicatif, move , puis redémarrage

Quelle données, quelle réplication, dans quel but,

Vos Besoins

- Valorisation pour la performance de l'entreprise
- L'intégration fonctionnelle dans votre SI, et la croissance des volumes de données
- Déploiement, Gestion et Outillage
- Gouvernance, accès facile, contrôlé et sécurisé
- Evolution et Modernisation et le contrôle des coûts

La compétence et l'expérience d'ACMI pour

- Choix de la Stratégie réplication
- Le Socle Architecture et Infrastructure
- La solution durable et évolutive



Vos données critiques sont sur IBM i et voulez:

- Les rendre hautement disponibles
- Les valoriser au travers d'autres applications
- Mettre en place une IA
- Innover et Moderniser vos solutions

ACMI possède l'expertise de l'IBM i et de la réplication de données

Nous pouvons qualifier vos choix, définir les moyens d'atteindre vos objectifs, répondre à la gestion de vos données

Contactez-nous

- Julien Bigo jbigo@acmi.fr
- Francois Martin fmartin@acmi.fr