
ENKIDU

MIOVSKI GROUP

undefined

Relation avec le document compagnon. Le présent document décrit le concentrateur de données. La plateforme de données de référence (partie 1 MDM et partie 2 entrepôt de données de référence, MDW) fait l'objet d'une architecture distincte. Les deux sont complémentaires : le concentrateur est la couche d'intégration et de médiation et la passerelle vers la plateforme ; la plateforme est l'autorité en matière d'identité et la couche de consolidation analytique.

1. Contexte et énoncé du problème

1.1 Le problème d'intégration

L'organisation exploite plusieurs systèmes ministériels qui détiennent chacun une partie du portrait opérationnel (RH, ERP, CRM, planification des quarts). La paie est différente : elle est centralisée et exploitée à l'externe par SPAC. Une paie correcte dépend de données exactes, ponctuelles et correctement séquencées circulant vers le système de paie, et du retour des résultats et des exceptions pour rapprochement.

En interfaces directes point à point, ce type de parc se dégrade vite : le nombre d'interfaces croît de façon combinatoire et il n'existe aucun endroit unique pour imposer la qualité ou prouver que ce qui a été envoyé correspond à ce qui a été reçu. Le concentrateur remplace ce rapprochement par paires par une seule couche de médiation gouvernée : chaque système s'intègre une fois, avec le concentrateur.

1.2 Ce qu'est — et n'est pas — le concentrateur

Le concentrateur est la couche d'intégration, de médiation et de passerelle. Il ingère depuis les systèmes ministériels, harmonise et valide selon un modèle canonique, séquence et rapproche l'échange avec la paie, et distribue des données fiables. Il n'est ni le maître de l'identité ni l'entrepôt analytique : ces responsabilités appartiennent à la plateforme de données de référence.

1.3 Portée

Dans la portée : les couches et sous-systèmes du concentrateur ; l'intégration avec RH, ERP, CRM et la planification comme sources et avec la paie SPAC comme consommateur centralisé ; le rôle de passerelle vers MDM et le MDW ; la discipline de mise en étape, le rapprochement, la sécurité, la traçabilité et la gouvernance.

Hors portée : la conception interne des systèmes sources et de la paie SPAC ; la conception interne de MDM et du MDW (document compagnon) ; la conception physique, le choix des produits, le dimensionnement ; la planification de projet et la migration.

1.4 Relation avec la plateforme de données de référence

La plateforme résout l'identité et publie l'enregistrement de référence (golden record), la référence croisée (XREF) qui associe l'identifiant de référence à chaque clé source, et les données de référence gouvernées. Le concentrateur les consomme comme intrants conformes : il consulte l'identité auprès de la plateforme plutôt que d'en décider lui-même — principe de l'autorité unique pour l'identité.

2. Principes et décisions d'architecture

L'architecture repose sur un petit nombre de décisions structurantes ; toute conception qui en contredit une doit être documentée et approuvée.

2.1 Décisions structurantes

D1 • Médiation en étoile (hub-and-spoke)

Tout échange intersystèmes passe par le concentrateur — N interfaces gouvernées, non $N \times (N-1)$ fragiles. Un seul endroit pour gouverner, sécuriser, observer et rapprocher.

D2 • Modèle de données canonique

Chaque source se mappe une fois vers un modèle neutre ; la sémantique est imposée centralement, non négociée par paires.

D3 • Déférer l'identité au MDM (courtier, jamais maître)

Le concentrateur résout les identifiants de référence en les consultant dans la XREF du MDM ; il n'exécute pas son propre appariement. Une seule autorité pour l'identité.

D4 • Rapprochement et gestion des exceptions de premier plan

Chaque transfert est prouvable et rien n'est perdu silencieusement : totaux de contrôle, accusés de réception, file d'attente morte et rejeu.

D5 • Atterrissage immuable et mise en étape rejouable

Les données passent par Atterrissage -> Travail -> Publiée. L'Atterrissage est immuable, rendant le pipeline rejouable quand les règles changent.

D6 • Sous-systèmes explicites et bornés

Chaque capacité est nommée, bornée et dotée d'une seule responsabilité, rendant la propriété et les interfaces sans ambiguïté.

D7 • Provenance et chaîne de possession

Pour chaque valeur, le concentrateur enregistre son origine et la séquence ordonnée des décisions — de quoi reconstituer et défendre tout enregistrement.

D8 • Observabilité avec application des clôtures

Ponctualité, exhaustivité et débit sont mesurés par rapport aux SLA, avec alertes liées aux fenêtres de clôture de la paie.

2.2 Décision de style d'intégration

Le concentrateur adopte un style de médiation événementielle avec modèle canonique, en façade d'API gérées. Le point à point est rejeté (prolifération d'interfaces) ; l'ETL par lots est partiel (bon pour l'analytique, trop lent pour la paie) ; le bus de services d'entreprise est trop couplé. Le concentrateur événementiel est retenu : il découple les systèmes, prend en charge le quasi temps réel et les lots, et rend les transferts rapprochables et rejouables.

2.3 Principes de soutien

- Programme, pas projet : la gouvernance et l'intendance sont de premier plan et continues.
- Les données brutes sont immuables : la zone Atterrissage n'est jamais modifiée sur place.
- Chaque décision est attribuable : aucune valeur publiée sans source et justification enregistrées.
- Gouverner une fois, réutiliser partout : le concentrateur est la passerelle sanctionnée vers les données de référence et l'analytique.
- Couplage lâche : les systèmes s'intègrent via des contrats versionnés et des événements publiés.
- Conçu pour la récupérabilité : le travail en cours peut être rejoué de façon idempotente.

3. Architecture de référence de haut niveau

L'architecture se lit de haut en bas comme un flux : les rayons sources alimentent une couche d'intégration ; les données atterrissent, sont traitées et publiées ; les données fiables sont distribuées aux consommateurs, dont la paie est le plus critique et MDM/MDW les partenaires de passerelle. Deux plans transversaux couvrent chaque couche : gouvernance et sécurité ; observabilité, traçabilité et audit.

Source de données fiables, pas seulement une source unique de vérité. Le concentrateur publie les données que tous acceptent d'utiliser, assemblées et validées selon le modèle canonique et les identités de référence courtisées auprès du MDM.

4. Catalogue des sous-systèmes

Chaque sous-système a une responsabilité claire. Ensemble, ils couvrent tout le cycle de vie des données — capturées, conformées, séquencées, résolues contre le MDM, rapprochées, persistées et distribuées — sous gouvernance, sécurité et observabilité continues.

HUB-01 • Intégration et connecteurs

Capter chaque source (CDC, API, lot/fichier, événements) vers l'Atterrissage, et syndiquer les données fiables en retour.

HUB-02 • Modèle canonique et mappage

Définir le modèle neutre d'entités/attributs et les mappages source-vers-canonique.

HUB-03 • Moteur de qualité des données

Profiler, standardiser, valider et enrichir selon les six dimensions de qualité.

HUB-04 • Orchestration et flux de travail

Séquencer les opérations, gérer les dépendances, appliquer reprise, compensation (saga) et planification.

HUB-05 • Courtage des données de référence

Résoudre les clés métier en identifiants de référence via la XREF du MDM. Courtier d'identité ; jamais maître.

HUB-06 • Rapprochement et contrôle

Comparer totaux de contrôle et comptes aux accusés de réception pour prouver chaque transfert.

HUB-07 · Gestion des exceptions et du rejeu

Capturer les échecs en file morte, classer et router, retraiter ou rejouer de façon idempotente.

HUB-08 · Mise en étape et persistance

Détenir les zones Atterrissage (immuable), Travail et Publiée, plus le magasin opérationnel et l'état de rejeu.

HUB-09 · Sortie et syndication

Distribuer les données fiables par événements et API, et resynchroniser les valeurs corrigées vers les sources.

HUB-10 · Gestion des API et passerelle

Imposer contrats, versionnage, limitation, sécurité et découverte — la porte d'entrée gérée.

HUB-11 · Sécurité, traçabilité et auditabilité

Contrôle d'accès, provenance au niveau attribut et journal de chaîne de possession.

HUB-12 · Observabilité et surveillance

Mesurer ponctualité, exhaustivité, débit et santé selon les SLA, avec alertes liées aux clôtures de paie.

HUB-13 · Métadonnées et catalogue de référence

Détenir définitions, propriété, sensibilité et règles de qualité ; mettre en cache les listes de codes gouvernées par le MDM.

HUB-14 · Console de gouvernance et d'intendance

La surface opérationnelle pour gérer exceptions, contrats, définitions et politiques.

5. Pipeline de traitement et discipline de mise en étape

Les données passent par trois zones nommées — Atterrissage, Travail et Publiée — reflétant la discipline Bronze -> Argent -> Or. La promotion vers l'avant n'a lieu qu'au franchissement du seuil de la zone suivante ; rien n'est supprimé de l'Atterrissage, donc tout le pipeline est rejouable.

5.1 Atterrissage — brut, immuable

L'Atterrissage reçoit les données telles qu'envoyées, sans les modifier ; chaque enregistrement est estampillé (source, horodatage, lot, somme de contrôle). C'est le registre rejouable de tout ce qui est reçu : Travail et Publiée peuvent être reconstruits sans réextraire des sources.

5.2 Travail — nettoyer, résoudre, rapprocher

Le Travail harmonise vers le modèle canonique, valide et enrichit, résout chaque clé en identifiant de référence via la XREF (garant tout inconnu — section 9), orchestre et rapproche. Tout ce qui ne peut être résolu avec confiance part en gestion des exceptions.

5.3 Publiée — fiable, prête à consommer

La Publiée ne détient que des enregistrements fiables et finis ; la promotion versionne l'enregistrement et estampille sa provenance. La sortie publie événements et API et resynchronise

les valeurs corrigées vers les sources. Les fenêtres de clôture de la paie contraignent le moment où la Publiée doit être prête.

PROPRIÉTÉS DES ZONES

Contenu Atterrissage : brut ; Travail : en cours ; Publiée : enregistrements fiables finis.

Mutabilité Atterrissage : immuable ; Travail : re-dérivable ; Publiée : versionnée, contrôlée.

Accès Atterrissage : pipeline et audit ; Travail : moteur et intendants ; Publiée : consommateurs (lecture), resynchronisation.

6. Intégration et conception des interfaces

Le style est choisi par interface selon la ponctualité, le volume et les capacités ; dans tous les cas, l'interface se mappe au modèle canonique via un contrat versionné.

Lot / fichier planifié

Transferts périodiques à fort volume, dont de nombreuses interfaces de paie ; doivent se rapprocher dans les fenêtres de clôture.

Événement / publication-abonnement

Propagation quasi temps réel de changements discrets ; exige des consommateurs idempotents.

API requête / réponse

Récupération à la demande, en façade de la gestion des API (HUB-10).

Capture des changements (CDC)

Capter les changements incrémentaux sans extraction complète ; gérer suppressions et ordre.

Contrats canoniques. Chaque interface est un contrat versionné contre le modèle canonique (HUB-02), ce qui permet à une source ou un consommateur d'évoluer sans reprendre toutes les autres interfaces.

7. Rapprochement et gestion des exceptions

Comme un point de terminaison pilote la paie, prouver l'exhaustivité et gérer proprement les échecs est la capacité signature du concentrateur. Distinct du nettoyage : le rapprochement prouve que l'ensemble envoyé correspond à ce que le récepteur accuse.

7.1 Contrôles de rapprochement

- Totaux et comptes de contrôle comparés entre le concentrateur et le récepteur.
- Gestion des accusés : chaque transfert sortant est apparié à son reçu et à tout message de retour de la paie.
- Rapports de rapprochement : les écarts sont exposés comme rapports et exceptions, non absorbés en silence.

7.2 Routage des exceptions

Échec de validation auto-corrigeable -> auto-correction + journal

Corrigé par règle avec justification enregistrée ; se poursuit.

Échec non auto-corrigeable -> file morte, intendant

Retenu pour résolution humaine via la console d'intendance.

Identité non résolue -> attente, transfert au MDM

Garé ; entité soumise au MDM ; retraité à la résolution (section 9).

Écart de comptes/totaux -> exception de rapprochement

Transfert retenu et investigué avant libération.

Non accusé ou rejeté -> reprise, puis file morte

Reprise idempotente ; si persistant, soulevé en exception.

Erreur transitoire (délai) -> reprise idempotente + délai

Repris automatiquement ; sans effet en double.

7.3 Rejeu et idempotence

Comme l'Atterrissage est immuable, tout élément peut être retraité par rejeu à travers Travail et Publiée. Le traitement est idempotent : réingérer le même enregistrement donne le même résultat, sans perte ni doublon.

8. Sécurité, traçabilité et auditabilité

Cette capacité transversale répond, pour toute valeur : qui peut la voir ou la modifier (sécurité), d'où elle vient (provenance), et quelle séquence de décisions l'a produite et routée (chaîne de possession).

8.1 Sécurité et contrôle d'accès

- Accès par rôle et attribut sur les zones et champs ; moindre privilège par zone.
- Chiffrement et masquage des attributs sensibles (identifiants nationaux, données de paie) en transit, au repos et dans les vues d'intendance.

8.2 Traçabilité et provenance

Pour chaque valeur, le concentrateur enregistre la source contributrice, sa transformation et l'identifiant de référence résolu. La provenance est au niveau attribut, non seulement enregistrement ; la traçabilité va de la copie Atterrissage aux transformations Travail jusqu'à la valeur Publiée et son accusé.

8.3 Audit et chaîne de possession

Chaque transformation et décision — automatisée ou humaine — est écrite dans un journal d'audit en ajout seul, suffisant pour reconstituer l'état d'un enregistrement et le défendre. Les régimes de vie privée (LPRPDE, Loi sur la protection des renseignements personnels) exigent de savoir où vivent les données d'une personne ; la provenance au niveau attribut rend cela requêtable.

9. Passerelle vers les données de référence et l'analytique

Le concentrateur est la passerelle unique et gouvernée vers la plateforme (MDM et MDW).

9.1 Le rôle de passerelle

Plutôt que de laisser chaque système atteindre MDM ou MDW directement, le concentrateur médie l'accès : il consulte l'identité dans la XREF et attache l'enregistrement de référence (HUB-05) ; pour l'analytique, il publie des données fiables vers le MDW. Cela préserve traçabilité, sécurité et contrats.

9.2 Entités non résolues — transfert au MDM

Face à une entité sans enregistrement de référence, le concentrateur ne doit pas inventer d'identité : il gare l'enregistrement, soumet l'entité au MDM, et attend. Une fois le golden record publié et la XREF mise à jour, les enregistrements retenus sont réévalués automatiquement et poursuivent vers la Publiée.

9.3 Distribution analytique vers le MDW

La responsabilité du concentrateur s'achève à la livraison de données fiables et conformées au MDW par sortie gouvernée ; la consolidation interne Bronze/Argent/Or du MDW est hors portée (document compagnon).

10. Gouvernance et intendance

Le concentrateur centralise le mouvement des données et donc la responsabilité. Des rôles et processus clairs rendent les données publiées dignes de confiance.

Propriétaire des données

Autorité métier d'un domaine ; approuve définitions, contrats et précédence de source de vérité.

Intendant des données

Gardien quotidien de la qualité ; traite la file des exceptions et approuve les changements que le concentrateur n'applique pas seul.

Dépositaire des données

Côté TI ; exploite et sécurise le concentrateur, ses magasins et la plomberie ; gère incidents et rejeu.

Conseil de gouvernance

Organe transversal ; fixe les normes, arbitre les différends et approuve seuils, contrats et politiques.

10.2 Processus opérationnels

- Gestion du changement : contrats versionnés et changements contrôlés aux mappages, validations et règles.
- Gestion de la qualité : règles, seuils et rapports, exceptions triées aux intendants.
- Tri des exceptions : un flux défini de la détection à la résolution et au rejeu.

- Surveillance opérationnelle : suivi continu des SLA et alertes contre les fenêtres de paie.

11. Exigences non fonctionnelles

Qualités que la mise en ?uvre doit atteindre ; les cibles sont à confirmer avec les parties prenantes et le calendrier de paie.

- Disponibilité : haute disponibilité sur le chemin critique de la paie, sans point unique de défaillance.
- Récupérabilité : RTO/RPO définis ; rejeu sans perte ni doublon.
- Ponctualité : transferts rapprochés dans les fenêtres de paie, avec alertes de risque.
- Performance et évolutivité : gérer les pics ; mise à l'échelle indépendante de l'ingestion et du traitement.
- Sécurité et vie privée : moindre privilège, chiffrement, classification, audit.
- Auditable : traçabilité complète et piste d'audit inviolable.
- Observabilité : SLA mesurables avec alertes proactives et tableaux de bord.
- Maintenabilité : ajouter ou changer un rayon sans re-concevoir les autres.
- Qualité des données : six dimensions mesurées en continu.

12. Préoccupations transversales

- Évolutivité et rejeu : grâce à l'Atterrissage immuable, Travail et Publiée se reconstruisent par rejeu.
- Idempotence : réingérer le même enregistrement donne le même résultat.
- Modes de latence : événement/CDC pour le quasi temps réel ; lots pour les gros volumes de paie.
- Versionnage : contrats, règles, modèle canonique et enregistrements versionnés.
- Surveillance de la qualité : les six dimensions mesurées en continu.
- Déploiement progressif : commencer par les flux à forte valeur (RH-vers-paie).

13. Risques et décisions de conception clés

13.1 Décisions de conception clés

Placement du MDM

Changement depuis v1.0 : courtiser le MDM externe. Avec une plateforme dédiée, intégrer un second maître créerait des décisions d'identité conflictuelles.

Cadence des interfaces

Mixte : événements pour les changements discrets, lots pour les gros transferts de paie liés aux clôtures.

État du concentrateur

Avec état : Atterrissage immuable et étapes Travail/Publiée requis pour rapprochement, rejeu et requêtes.

Couplage

Médié par le concentrateur via contrats canoniques et événements, pour éviter la prolifération.

13.2 Registre des risques

Données incorrectes ou incomplètes vers la paie

Impact : paie erronée, perte de confiance. Atténuation : qualité à la frontière ; rapprochement par conception.

Fenêtre de clôture de paie manquée

Impact : paie retardée. Atténuation : surveillance des SLA ; haute disponibilité sur le chemin critique.

Décisions d'identité conflictuelles concentrateur/MDM

Impact : golden records divergents. Atténuation : autorité unique — le concentrateur courtise et défère au MDM.

Accès direct non gouverné aux sources/MDM/MDW

Impact : versions divergentes. Atténuation : imposer le concentrateur comme passerelle unique.

Échec irrécupérable en cours de transfert

Impact : perte ou doublon. Atténuation : Atterrissage immuable ; retraitement idempotent et rejouable.

14. Annexes

Annexe A — Glossaire

Modèle canonique Représentation neutre des entités partagées vers laquelle chaque source se mappe.

Golden record L'enregistrement composite faisant autorité pour une entité, produit par le MDM.

XREF (référence croisée) L'association d'un identifiant de référence à chaque clé source, détenue par le MDM.

En étoile (hub-and-spoke) Topologie où tous les systèmes s'intègrent via un concentrateur central.

Atterrissage / Travail / Publiée Les trois zones de mise en étape — brut-immuable, en cours, fiable — équivalentes à Bronze/Argent/Or.

Rapprochement Prouver que l'ensemble envoyé correspond à ce que le récepteur accuse.

Provenance D'où vient une valeur — la source qui l'a contribué.

Chaîne de possession La séquence ordonnée et enregistrée des décisions ayant produit et routé une valeur.

Idempotence La propriété selon laquelle retraiter le même élément ne produit aucun effet en double.

Resynchronisation Publier les valeurs corrigées vers les systèmes sources (coexistence).

Annexe C — Références

- Plateforme de données de référence — Architecture de référence (partie 1 MDM, partie 2 MDW) : le document compagnon.
- Méthode d'architecture d'entreprise (par exemple TOGAF).
- Corpus de connaissances en gestion des données (par exemple DAMA-DMBOK).
- Modèle de qualité des données (par exemple ISO/IEC 25012).
- Directives gouvernementales d'architecture, de sécurité et de vie privée (par exemple LPRPDE et Loi sur la protection des renseignements personnels).

FIN DU DOCUMENT