
ENKIDU

MIOVSKI GROUP

undefined

Relación con el documento complementario. Este documento describe el concentrador de datos. La Plataforma de Datos Maestros (parte 1 MDM y parte 2 almacén de datos maestros, MDW) se describe en una arquitectura aparte. Ambos son complementarios: el concentrador es la capa de integración y mediación y la pasarela a la plataforma; la plataforma es la autoridad de identidad y la capa de consolidación analítica.

1. Contexto y planteamiento del problema

1.1 El problema de integración

La organización opera varios sistemas departamentales que poseen cada uno parte del panorama operativo (RR. HH., ERP, CRM, planificación de turnos). La nómina es distinta: está centralizada y la opera externamente PSPC. Una paga correcta depende de datos exactos, oportunos y correctamente secuenciados que fluyen al sistema de pago, y del retorno de resultados y excepciones para su conciliación.

Con interfaces directas punto a punto, este tipo de parque se degrada rápido: el número de interfaces crece de forma combinatoria y no hay un lugar único para imponer calidad ni para probar que lo enviado coincide con lo recibido. El concentrador sustituye esa conciliación por pares con una única capa de mediación gobernada: cada sistema se integra una vez, con el concentrador.

1.2 Lo que es — y no es — el concentrador

El concentrador es la capa de integración, mediación y pasarela. Ingiere de los sistemas departamentales, armoniza y valida contra un modelo canónico, secuencia y concilia el intercambio con la nómina, y distribuye datos fiables. No es el maestro de la identidad ni el almacén analítico: esas responsabilidades pertenecen a la Plataforma de Datos Maestros.

1.3 Alcance

En alcance: las capas y subsistemas del concentrador; la integración con RR. HH., ERP, CRM y planificación como fuentes y con la nómina PSPC como consumidor centralizado; el rol de pasarela a MDM y al MDW; la disciplina de escalonamiento, la conciliación, la seguridad, el linaje y la gobernanza.

Fuera de alcance: el diseño interno de los sistemas fuente y de la nómina PSPC; el diseño interno de MDM y del MDW (documento complementario); el diseño físico, la selección de productos y el dimensionamiento; la planificación del proyecto y la migración.

1.4 Relación con la Plataforma de Datos Maestros

La plataforma resuelve la identidad y publica el registro maestro (golden record), la referencia cruzada (XREF) que asocia el ID maestro con cada clave de origen, y los datos de referencia gobernados. El concentrador los consume como insumos conformados: consulta la identidad en la plataforma en lugar de decidirla él mismo — principio de autoridad única para la identidad.

2. Principios y decisiones de arquitectura

La arquitectura descansa en un pequeño número de decisiones estructurales; todo diseño que contradiga una de ellas debe documentarse y aprobarse.

2.1 Decisiones estructurales

D1 · Mediación en estrella (hub-and-spoke)

Todo intercambio entre sistemas pasa por el concentrador — N interfaces gobernadas, no $N \times (N-1)$ frágiles. Un solo lugar para gobernar, asegurar, observar y conciliar.

D2 · Modelo de datos canónico

Cada fuente se mapea una vez a un modelo neutral; la semántica se impone centralmente, no se negocia por pares.

D3 · Deferir la identidad al MDM (intermediar, nunca dominar)

El concentrador resuelve los ID maestros consultándolos en la XREF del MDM; no ejecuta su propio emparejamiento. Una única autoridad para la identidad.

D4 · Conciliación y gestión de excepciones de primer nivel

Cada transferencia es demostrable y nada se pierde en silencio: totales de control, acuses, cola muerta y reproceso.

D5 · Aterrizaje inmutable y escalonamiento reproducible

Los datos pasan por Aterrizaje -> Trabajo -> Publicada. El Aterrizaje es inmutable, lo que hace el pipeline reproducible cuando cambian las reglas.

D6 · Subsistemas explícitos y acotados

Cada capacidad tiene nombre, límites y una única responsabilidad, de modo que la propiedad y las interfaces son inequívocas.

D7 · Procedencia y cadena de custodia

Para cada valor, el concentrador registra su origen y la secuencia ordenada de decisiones — suficiente para reconstruir y defender cualquier registro.

D8 · Observabilidad con aplicación de cierres

Puntualidad, integridad y rendimiento se miden contra los SLA, con alertas ligadas a las ventanas de cierre de la nómina.

2.2 Decisión del estilo de integración

El concentrador adopta un estilo de mediación orientada a eventos con modelo canónico, tras API gestionadas. El punto a punto se rechaza (proliferación de interfaces); el ETL por lotes es parcial (bueno para analítica, lento para la nómina); el bus de servicios empresarial acopla demasiado. Se elige el concentrador orientado a eventos: desacopla los sistemas, admite casi tiempo real y lotes, y hace las transferencias conciliables y reproducibles.

2.3 Principios de apoyo

- Programa, no proyecto: la gobernanza y la administración son de primer nivel y continuas.
- Los datos brutos son inmutables: la zona de Aterrizaje nunca se edita in situ.
- Cada decisión es atribuible: ningún valor se publica sin origen y justificación registrados.
- Gobernar una vez, reutilizar en todas partes: el concentrador es la pasarela sancionada a los

datos maestros y la analítica.

- Acoplamiento débil: los sistemas se integran mediante contratos versionados y eventos publicados.
- Diseñado para la recuperabilidad: el trabajo en curso puede reproducirse de forma idempotente.

3. Arquitectura de referencia de alto nivel

La arquitectura se lee de arriba abajo como un flujo: los radios fuente alimentan una capa de integración; los datos aterrizan, se trabajan y se publican; los datos fiables se distribuyen a los consumidores, de los que la nómina es el más crítico y MDM/MDW los socios de pasarela. Dos planos transversales abarcan cada capa: gobernanza y seguridad; observabilidad, linaje y auditoría.

Fuente de datos fiables, no solo una única fuente de verdad. El concentrador publica los datos que todos aceptan usar, ensamblados y validados contra el modelo canónico y las identidades maestras intermediadas desde el MDM.

4. Catálogo de subsistemas

Cada subsistema tiene una responsabilidad clara. Juntos cubren todo el ciclo de vida de los datos — capturados, conformados, secuenciados, resueltos contra el MDM, conciliados, persistidos y distribuidos — bajo gobernanza, seguridad y observabilidad continuas.

HUB-01 · Integración y conectores

Capturar cada fuente (CDC, API, lote/archivo, eventos) hacia el Aterrizaje y syndicar datos fiables de vuelta.

HUB-02 · Modelo canónico y mapeo

Definir el modelo neutral de entidades/atributos y los mapeos de fuente a canónico.

HUB-03 · Motor de calidad de datos

Perfilar, estandarizar, validar y enriquecer según las seis dimensiones de calidad.

HUB-04 · Orquestación y flujo de trabajo

Secuenciar operaciones, gestionar dependencias y aplicar reintento, compensación (saga) y planificación.

HUB-05 · Intermediación de datos maestros

Resolver claves de negocio a ID maestros vía la XREF del MDM. Intermedia identidad; nunca la domina.

HUB-06 · Conciliación y control

Comparar totales de control y recuentos con los acuses para probar cada transferencia.

HUB-07 · Gestión de excepciones y reproceso

Capturar fallos en cola muerta, clasificar y enrutar, reprocesar o reproducir de forma idempotente.

HUB-08 · Escalonamiento y persistencia

Mantener las zonas Aterrizaje (inmutable), Trabajo y Publicada, más el almacén operativo y el

estado de reproceso.

HUB-09 · Salida y sindicación

Distribuir datos fiables por eventos y API, y resincronizar valores corregidos a las fuentes.

HUB-10 · Gestión de API y pasarela

Imponer contratos, versionado, limitación, seguridad y descubrimiento — la puerta de entrada gestionada.

HUB-11 · Seguridad, linaje y auditabilidad

Control de acceso, procedencia a nivel de atributo y registro de cadena de custodia.

HUB-12 · Observabilidad y monitoreo

Medir puntualidad, integridad, rendimiento y salud según los SLA, con alertas ligadas a los cierres de nómina.

HUB-13 · Metadatos y catálogo de referencia

Mantener definiciones, propiedad, sensibilidad y reglas de calidad; cachear las listas de códigos gobernadas por el MDM.

HUB-14 · Consola de gobernanza y administración

La superficie operativa para gestionar excepciones, contratos, definiciones y políticas.

5. Pipeline de procesamiento y disciplina de escalonamiento

Los datos pasan por tres zonas — Aterrizaje, Trabajo y Publicada — reflejando la disciplina Bronce -> Plata -> Oro. La promoción solo ocurre al superar el umbral de la siguiente zona; nada se borra del Aterrizaje, por lo que todo el pipeline es reproducible.

5.1 Aterrizaje — bruto, inmutable

El Aterrizaje recibe los datos tal como los envió la fuente, sin modificarlos; cada registro se sella (origen, marca de tiempo, lote, suma de comprobación). Es el registro reproducible de todo lo recibido: Trabajo y Publicada pueden reconstruirse sin volver a extraer de las fuentes.

5.2 Trabajo — limpiar, resolver, conciliar

El Trabajo armoniza al modelo canónico, valida y enriquece, resuelve cada clave a un ID maestro vía la XREF (aparcando lo desconocido — sección 9), orquesta y concilia. Todo lo que no pueda resolverse con confianza va a la gestión de excepciones.

5.3 Publicada — fiable, lista para consumir

La Publicada solo contiene registros fiables y terminados; la promoción versiona el registro y sella su procedencia. La salida publica eventos y API y resincroniza valores corregidos a las fuentes. Las ventanas de cierre de la nómina limitan cuándo debe estar lista la Publicada.

PROPIEDADES DE LAS ZONAS

Contenido Aterrizaje: bruto; Trabajo: en curso; Publicada: registros fiables terminados.

Mutabilidad Aterrizaje: inmutable; Trabajo: rederivable; Publicada: versionada, controlada.

Acceso Aterrizaje: pipeline y auditoría; Trabajo: motor y administradores; Publicada: consumidores (lectura), resincronización.

6. Integración y diseño de interfaces

El estilo se elige por interfaz según puntualidad, volumen y capacidades; en todos los casos la interfaz se mapea al modelo canónico mediante un contrato versionado.

Lote / archivo programado

Transferencias periódicas de alto volumen, incluidas muchas interfaces de nómina; deben conciliar dentro de las ventanas de cierre.

Evento / publicación-suscripción

Propagación casi en tiempo real de cambios discretos; exige consumidores idempotentes.

API petición / respuesta

Recuperación bajo demanda, tras la gestión de API (HUB-10).

Captura de cambios (CDC)

Capturar cambios incrementales sin extracción completa; gestionar borrados y orden.

Contratos canónicos. Cada interfaz es un contrato versionado contra el modelo canónico (HUB-02), lo que permite que una fuente o un consumidor cambie sin renegociar todas las demás interfaces.

7. Conciliación y gestión de excepciones

Como un punto final impulsa la paga, probar la integridad y gestionar los fallos limpiamente es la capacidad distintiva del concentrador. Distinto de la limpieza: la conciliación prueba que el conjunto enviado coincide con lo que el receptor acusa.

7.1 Controles de conciliación

- Totales y recuentos de control comparados entre el concentrador y el receptor.
- Gestión de acuses: cada transferencia saliente se empareja con su recibo y con cualquier mensaje de retorno de la nómina.
- Informes de conciliación: las discrepancias se exponen como informes y excepciones, no se absorben en silencio.

7.2 Enrutamiento de excepciones

Fallo de validación autocorregible -> autocorrección + registro

Corregido por regla con justificación registrada; continúa.

Fallo no autocorregible -> cola muerta, administrador

Retenido para resolución humana vía la consola de administración.

Identidad no resuelta -> retener, traspaso al MDM

Aparcado; entidad enviada al MDM; reprocesado al resolverse (sección 9).

Discrepancia de recuentos/totales -> excepción de conciliación

Transferencia retenida e investigada antes de liberar.

No acusado o rechazado -> reintento, luego cola muerta

Reintento idempotente; si persiste, se eleva como excepción.

Error transitorio (tiempo de espera) -> reintento idempotente + espera

Reintentado automáticamente; sin efecto duplicado.

7.3 Reproceso e idempotencia

Como el Aterrizaje es inmutable, cualquier elemento puede reprocesarse reproduciéndolo por Trabajo y Publicada. El procesamiento es idempotente: reingerir el mismo registro produce el mismo resultado, sin pérdida ni duplicación.

8. Seguridad, linaje y auditabilidad

Esta capacidad transversal responde, para cualquier valor: quién puede verlo o cambiarlo (seguridad), de dónde vino (procedencia), y qué secuencia de decisiones lo produjo y enrutó (cadena de custodia).

8.1 Seguridad y control de acceso

- Acceso por rol y atributo sobre zonas y campos; mínimo privilegio por zona.
- Cifrado y enmascaramiento de atributos sensibles (identificadores nacionales, datos de paga) en tránsito, en reposo y en las vistas de administración.

8.2 Linaje y procedencia

Para cada valor, el concentrador registra la fuente que lo aportó, cómo se transformó y a qué ID maestro se resolvió. La procedencia es a nivel de atributo, no solo de registro; el linaje va de la copia de Aterrizaje a las transformaciones de Trabajo hasta el valor Publicado y su acuse.

8.3 Auditoría y cadena de custodia

Cada transformación y decisión — automatizada o humana — se escribe en un registro de auditoría de solo anexo, suficiente para reconstruir el estado de un registro y defenderlo. Los regímenes de privacidad (PIPEDA, Ley de Privacidad) exigen saber dónde viven los datos de una persona; la procedencia a nivel de atributo lo convierte en una consulta.

9. Pasarela a datos maestros y analítica

El concentrador es la pasarela única y gobernada a la plataforma (MDM y MDW).

9.1 El rol de pasarela

En lugar de dejar que cada sistema acceda a MDM o MDW directamente, el concentrador media el acceso: consulta la identidad en la XREF y adjunta el registro maestro (HUB-05); para la analítica, publica datos fiables al MDW. Esto mantiene intactos linaje, seguridad y contratos.

9.2 Entidades no resueltas — traspaso al MDM

Ante una entidad sin registro maestro, el concentrador no debe inventar identidad: aparca el registro, envía la entidad al MDM y espera. Una vez que el MDM publica el golden record y actualiza la XREF, los registros retenidos se reevalúan automáticamente y continúan a Publicada.

9.3 Distribución analítica al MDW

La responsabilidad del concentrador termina al entregar datos fiables y conformados al MDW por salida gobernada; la consolidación interna Bronce/Plata/Oro del MDW queda fuera de alcance (documento complementario).

10. Gobernanza y administración

El concentrador centraliza el movimiento de datos y, por tanto, la rendición de cuentas. Roles y procesos claros hacen que los datos publicados sean fiables.

Propietario de datos

Autoridad de negocio de un dominio; aprueba definiciones, contratos y precedencia de fuente de verdad.

Administrador de datos (steward)

Guardián diario de la calidad; trabaja la cola de excepciones y aprueba cambios que el concentrador no aplica por sí solo.

Custodio de datos

Lado de TI; opera y asegura el concentrador, sus almacenes y la plomería; gestiona incidentes y reproceso.

Consejo de gobernanza

Órgano transversal; fija normas, arbitra disputas y aprueba umbrales, contratos y políticas.

10.2 Procesos operativos

- Gestión del cambio: contratos versionados y cambios controlados a mapeos, validaciones y reglas.
- Gestión de la calidad: reglas, umbrales e informes, con excepciones derivadas a administradores.
- Triage de excepciones: un flujo definido de la detección a la resolución y el reproceso.
- Monitoreo operativo: seguimiento continuo de SLA y alertas contra las ventanas de nómina.

11. Requisitos no funcionales

Cualidades que la implementación debe cumplir; los objetivos se confirmarán con las partes interesadas y el calendario de nómina.

- Disponibilidad: alta disponibilidad en la ruta crítica de la paga, sin punto único de fallo.
- Recuperabilidad: RTO/RPO definidos; reproceso sin pérdida ni duplicación.
- Puntualidad: transferencias conciliadas dentro de las ventanas de nómina, con alertas de riesgo.

- Rendimiento y escalabilidad: manejar picos; escalar ingestión y procesamiento de forma independiente.
- Seguridad y privacidad: mínimo privilegio, cifrado, clasificación, auditoría.
- Auditabilidad: linaje completo y pista de auditoría a prueba de manipulaciones.
- Observabilidad: SLA medibles con alertas proactivas y tableros.
- Mantenibilidad: añadir o cambiar un radio sin rediseñar los demás.
- Calidad de datos: seis dimensiones medidas de forma continua.

12. Preocupaciones transversales

- Escalabilidad y reproceso: gracias al Aterrizaje inmutable, Trabajo y Publicada se reconstruyen por reproceso.
- Idempotencia: reingresar el mismo registro produce el mismo resultado.
- Modos de latencia: evento/CDC para casi tiempo real; lotes para grandes volúmenes de nómina.
- Versionado: contratos, reglas, modelo canónico y registros versionados.
- Monitoreo de la calidad: las seis dimensiones medidas de forma continua.
- Despliegue por fases: empezar por los flujos de mayor valor (RR. HH. a paga).

13. Riesgos y decisiones de diseño clave

13.1 Decisiones de diseño clave

Ubicación del MDM

Cambio respecto a v1.0: intermediar el MDM externo. Con una plataforma dedicada, incrustar un segundo maestro crearía decisiones de identidad conflictivas.

Cadencia de interfaces

Mixta: eventos para cambios discretos, lotes para grandes transferencias de nómina ligadas a cierres.

Estado del concentrador

Con estado: Aterrizaje inmutable y etapas Trabajo/Publicada requeridos para conciliación, reproceso y consultas.

Acoplamiento

Mediado por el concentrador vía contratos canónicos y eventos, para evitar la proliferación.

13.2 Registro de riesgos

Datos incorrectos o incompletos llegan a la nómina

Impacto: paga errónea, pérdida de confianza. Mitigación: calidad en la frontera; conciliación por diseño.

Ventana de cierre de nómina perdida

Impacto: paga retrasada. Mitigación: monitoreo de SLA; alta disponibilidad en la ruta crítica.

Decisiones de identidad conflictivas concentrador/MDM

Impacto: golden records divergentes. Mitigación: autoridad única — el concentrador intermedia y defiende al MDM.

Acceso directo no gobernado a fuentes/MDM/MDW

Impacto: versiones divergentes. Mitigación: imponer el concentrador como pasarela única.

Fallo irrecuperable a mitad de transferencia

Impacto: pérdida o duplicación. Mitigación: Aterrizaje inmutable; reproceso idempotente y reproducible.

14. Anexos

Anexo A — Glosario

Modelo canónico Representación neutral de las entidades compartidas a la que se mapea cada fuente.

Golden record El registro compuesto autoritativo de una entidad, producido por el MDM.

XREF (referencia cruzada) La asociación de un ID maestro con cada clave de origen, mantenida por el MDM.

En estrella (hub-and-spoke) Topología en la que todos los sistemas se integran mediante un concentrador central.

Aterrizaje / Trabajo / Publicada Las tres zonas de escalonamiento — bruto-inmutable, en curso, fiable — equivalentes a Bronce/Plata/Oro.

Conciliación Probar que el conjunto enviado coincide con lo que el receptor acusa.

Procedencia De dónde vino un valor — la fuente que lo aportó.

Cadena de custodia La secuencia ordenada y registrada de decisiones que produjo y enrutó un valor.

Idempotencia La propiedad de que reprocesar el mismo elemento no produce efecto duplicado.

Resincronización Publicar valores corregidos de vuelta a los sistemas fuente (coexistencia).

Anexo C — Referencias

- Plataforma de Datos Maestros — Arquitectura de referencia (parte 1 MDM, parte 2 MDW): el documento complementario.
- Método de arquitectura empresarial (por ejemplo TOGAF).
- Cuerpo de conocimiento en gestión de datos (por ejemplo DAMA-DMBOK).
- Modelo de calidad de datos (por ejemplo ISO/IEC 25012).
- Directivas gubernamentales de arquitectura, seguridad y privacidad (por ejemplo PIPEDA y la Ley de Privacidad).

FIN DEL DOCUMENTO