

Richtlijnen aansluiten van databronnen op het VfPf Datawarehouse

Datum: 9 maart 2026

Inleiding

Dit document geeft richtlijnen voor het tot stand brengen van een datakoppeling tussen een databronstelsel en het datawarehouse van VfPf.

Uitgangspunten

- De databron is technisch koppelbaar via Matillion of direct via Snowflake;
- Het ontkoppelpunt dat de voorkeur heeft, is een pull-mechanisme. Daar is een (ODS) omgeving ingericht door leverancier van de data die hun data zelf in de ODS zet. Periodiek doet het datawarehouse een pull om de data in de eigen omgeving in te laden en te verwerken;
- De frequentie is flexibel en afhankelijk van de informatiebehoefte. De kleinste frequentie is dagelijks. Wekelijks, maandelijks, jaarlijks en ad-hoc zijn ook mogelijk;
- Verwerking van de data moet binnen het maintenance window zijn afgerond. Dit geldt voor alle aangesloten databronnen. Dagelijks om 7:30 uur moet de data verwerkt zijn door het data warehouse;
- API koppelingen doen we bij voorkeur niet. API's zijn veelal gebouwd voor het snel ophalen van informatie over één case en niet om alle data van alle cases in één keer op te halen. De snelheid van een API is voor een datawarehouse over het algemeen niet optimaal;
- Kennis van het bronsysteem en de werking daarvan is belegd bij de leverancier van het bronsysteem;
- Afhankelijk van de mogelijkheden bij de leverancier wordt de methode van de datakoppeling bepaald.

Proces aansluiten bronsysteem

Bij een bronontsluiting zijn er gesprekken (onderhandelingen) over de wijze van koppelen, inhoud en SLA. Bij een nieuwe bronontsluiting is het noodzakelijk om de volgende partijen/rollen hierbij te hebben:

- VfPf Enterprise Architect
Verantwoordelijk voor waarborgen van de VfPf architectuurstandaarden en de voorkeursaansluiting;
- VfPf Data Platform Architect
Verantwoordelijk voor de ontsluitingsmethoden waarin de bronnen kunnen aanhaken om data te kunnen interfaceren;
- VfPf Data Analyst
Representeert de VfPf Business en de gerelateerde user story's;
- Bronsysteem afvaardiging
Heeft de technische en functionele kennis van de data in het bronsysteem. Is verantwoordelijk voor de realisatie van de technische en inhoudelijke handshake tussen het bronsysteem en het ontkoppelpunt van het VfPf data platform;
- Data Specialist/Data Officer
Bepaalt of de datakoppeling voldoet aan het databeleid van VfPf.

Tezamen met de hierboven benoemde partijen worden de mogelijkheden besproken en de beste aansluitmethode gekozen.

Soorten datakoppelingen

Ontkoppelpunt: Push versus Pull

1. Push
De data wordt door het bronsysteem zelf op een specifieke locatie neergezet;
2. Pull
VfPf haalt de gegevens op uit het bronsysteem.

Aanvullende afspraken moeten gemaakt worden tussen partijen over het moment van neerzetten en oppakken van de dataset en de verwijdering van eerdere data.

Data inhoud

1. Full
De complete dataset inclusief alle historie wordt verwerkt. De data is een snapshot van de gehele toestand op moment van ontkoppelen;
2. Delta
Alleen de data welke verschilt ten opzichte van de vorige dataset wordt verwerkt. De bepaling van de delta omvat mogelijk extra effort aan de zijde van het bronsysteem omdat

deze afhankelijk is van de vorige aanlevering. Bij een delta is het advies om ook periodiek een full load te kunnen doen om eventueel niet verwerkte delta's te corrigeren;

3. Transactie

De dataset betreft alle inserts, deletes en updates transacties ('changed data capture') vanaf het voorgaand ontkoppelmoment. Afnemende partij berekent zelf de nieuwe toestand op basis van deze inserts, deletes en updates.

De keuze wordt gemaakt op basis van de kenmerken rondom de data zoals verwacht datavolume, en achterliggend datamodel.

Manier van ontsluiten

1. API

Via een API call wordt de dataset ontsloten;

2. ODS

Er is een aparte omgeving aanwezig (Operational Data Store) waar de data gestructureerd is ontsloten. Dit is een relationele database.

3. Blob

Dit is een andere vorm van een ODS maar dan in minder gestructureerde vorm. Dit zijn bestanden in een folder op een container in een Azure storage account.