

Styringsmekanismer ved anvendelse af filudtræk

Introduktion

Dette afsnit indeholder anbefalinger til indholdsmæssige styringsmekanismer i forbindelse med vedligeholdelse af anvenderens kopiregister. Der er i nuværende version af guide udelukkende beskrevet styringsmekanismer omkring filudtræk.

Men hvis anvender ønsker at vedligeholde et kopiregister via REST tjenester og hændelser, ligger styringen i anvenderens valg af inputparametre til REST tjenesterne.

- [Navngivning og arkivering af filudtræk](#)
- [Metadatafilen forklaret](#)
- [Deltaudtræk og dataintegritet](#)
- [Kontrol af data efter indlæsning](#)

Sideinformation

Oprindelig kilde	Datafordeleren
Forfatter	Datafordeleren
Oprettet	Jan 08, 2021
Version	1.0
Ændret	Jan 04, 2024
Sidehistorik	• Se sidehistorik

Navngivning og arkivering af filudtræk

[Til toppen](#)

Alle filudtræk skal hentes (pull) fra Datafordelerens filserver, med enten ftp eller sftp. Datafordeleren sender en adviseringsmail, når et givent udtræk er klar til hentning. Alternativt kan man anvende et pull via et tidsstyret program. [Læs mere om filudtræk](#)

Filerne navngives efter følgende regel: <abonnementnavn>_<ååååmmddttmmss>.<format>, eksempelvis: KK_20201030192200.json.

Udover selve datafilen dannes der også en metadatafil, navngivet efter samme regel, med "_Metadata" til sidst i navnet. Eksempelvis: KK_20201030192200_Metadata.json

Anvenderen kan ikke slette filer på Datafordelerens filserver, og sletningen sker automatisk efter en fastsat periode, typisk 30 dage. Perioden er specificeret af registermyndighed der er ansvarlig for udtrækket. Bemærk at filerne kan ikke genproduceres, når de først er slettet.

Disse mekanismer betyder, at anvenderen selv skal holde styr på hvilke filer, der er anvendt og hvilke filer der er nye. Ligeledes skal anvenderen selv gemme en kopi af filerne i eget miljø, hvis der er behov for at gemme filer ældre end 30 dage.

Metadatafilen forklaret

[Til toppen](#)

Metadatafilerne indeholder en række nyttige oplysninger, der kan anvendes til automatisering og kontrol af datafilerne. Filerne er forholdsvis små, så anvenderen kan hurtigt downloade *Metadata.json (*Metadata.XML) filerne og ud fra indholdet i disse filer, downloade og behandle datafilerne i korrekt rækkefølge.

Nedenfor er vist et udsnit af oplysningerne fra metadatafilen, som det anbefales at anvende i forbindelse med automatisering af filudtræk. Oplysningerne gennemgås kort:

Metadatafil - udsnit af oplysninger
"leveranceNavn": "KK_20201030192200.json",
"miljoe": "PROD01",
"MD5Checksum": "d41d8cd98f00b204e9800998ecf8427e",
"DatafordelerUdtraekstidspunkt": [
{
"deltavindueStart": "1900-01-01T00:00:00.000+00:00",
"deltavindueSlut": "2020-10-30T19:22:00.715+01:00"}],
"AbonnementsOplysninger": [
{
"webBrugernavn": "AlfredJensen3Y7Z9T",
"tjenesteBrugernavn": "IOTLGYRZGP",
"abonnementnavn": "KK",
"tjenestenavn": "MUEjerlejlighed",
"tjenesteverision": "1.0.1",
"oprettelsesdato": "2020-10-29T19:49:25.991+01:00",
"senesteAbonnementRedigeringsdato": "2020-10-29T19:49:25.991+01:00"}],

leveranceNavn indeholder navnet på datafilen og kan således anvendes til at downloade den korrekte fil.

miljoe indeholder navnet på det miljø, data er udtrukket fra. Det kan anvendes til at kontrollere at data er fra test henholdsvis produktion, således at man som anvender indlæser data i det korrekte miljø.

MD5Checksum indeholder MD5 checksummen på datafilen. Anvenderen kan bruge MD5 checksummen til at kontrollere at filen er blevet downloadet korrekt, ved selv at generere en MD5 checksum på datafilen efter download og kontrollere at den stemmer overens med denne værdi.

deltavindueStart indeholder timestamper for den ældste ændring i datafilen. Dette felt er mest relevant, hvis der anvendes deltaudtræk. Ved totaludtræk står der altid 1900-01-01 som ovenfor.

deltavindueSlut indeholder timestamper for den nyeste mulige ændring i datafilen. I tilfælde af at der i samme proces downloades flere filer fra samme tjeneste, er det vigtigt at behandle filerne i korrekt dato-orden, med den ældste fil først. Medmindre man kun er interesseret i aktuelle data, i så fald, kan man nøjes med at behandle den nyeste fil.

Attributterne under **AbonnementsOplysninger**, kan anvendes til at kontrollere om det er den forventede tjeneste, tjenesteverision og lignende. Dette kan være relevant i situationer hvor dataformatet i tjenesterne ændres i en ny version.

Deltaudtræk og dataintegritet

[Til toppen](#)

Opdatering af forretningsobjekter på Datafordeleren fra et register der anvender bitemporalitet omfatter i størsteparten af tilfældene behandling (opdatering/indsættelse) af flere rækker i Datafordelerens databaser. De opdaterede data bliver først gjort tilgængelig for anvendere, når alle opdateringer i en opdateringsfil eller opdateringer i et SOAP-kald til Datafordeleren er indlæst. Herved sikres dataintegritet på Datafordeleren i forhold til de dataleverancer der modtages fra registre.

Det kan oplyses at opdateringer indlæses på Datafordeleren i eksakt samme rækkefølge som de modtages fra Registre, således at den datamæssige og forretningsmæssige integritet i data opretholdes. Dette fremgår også af aftalen mellem Datafordeleren og registrene.

Da opdateringer af et forretningsobjekt kan spænde over flere tabeller, bør anvendere altid betragte det fulde deltaudtræk som en atomisk transaktion, for at dataintegritet for forretningsobjekter sikres. Det er ikke nok at fokusere på en tabel ad gangen, da man derved kan få inkonsistente forretningsobjekter i kortere eller længere periode.

Kontrol af data efter indlæsning

[Til toppen](#)

Efter succesfuld download og indlæsning af data, kan antallet af de enkelte forretningsobjekter i databasen for kopi-registret, kontrolleres ved at anvende COUNT funktionen på REST tjenesterne. Der er count på samtlige REST tjenester.

Da kopiregistre over tid kommer ud af synk med master registret, anbefales det også at anvendere periodisk, eller ved mistanke om fejl i kopidata, kontrollerer data i kopiregistret mod Datafordeleren.

Dette kan gøres på flere måder, men det anbefales at lave totaludtræk med jævne mellemrum og sammenligne disse udtræk med kopiregistret.

Denne metode kan eventuelt suppleres med oftere kontroller baseret på kald til registrenes REST services, hvor det returnerede data også sammenlignes med kopiregistret.

Sidstnævnte metode kan anvendes til hurtigt at synkroniserer specifikke objekter, hvis man har mistanke om fejl i disse.
