





























































Copy Job w Fabric

Data Community Kraków, 22.10.2025

Tomasz Libera

libera@kursysql.pl

Data Movement strategy

| Capability |  Mirroring |  Copy job |  Copy Activity (Pipeline) |  Eventstream |
|--|---|--|--|---|
| Sources | Databases + 3rd party integration into Open Mirroring | All supported data sources and formats | All supported data sources and formats | 25+ sources and all formats |
| Destinations | Tabular format in Fabric OneLake (read-only) | All supported destinations and formats | All supported destinations and formats | 4+ destinations |
| Flexibility | Turn-key setup with fixed behavior | Simple to use + Advanced knobs | Advanced and fully customized knobs | Simple and customizable options |
| Custom scheduling |  |  |  |  |
| Table and Column management |  |  |  |  |
| Copy behavior: Append, Upsert, Override |  |  |  |  |
| Advanced observability + auditing |  |  |  |  |
| Copy modes | | | | |
| CDC-based continuous replication |  |  |  |  |
| Batch or bulk copy |  |  |  |  |
| Native support for Incremental copy (watermark-based) |  |  |  |  |
| Copy using user defined query |  |  |  |  |
| Use cases | | | | |
| Continuous Replication for analytics and reporting |  |  |  |  |
| Metadata driven ELT/ETL for data warehousing |  |  |  |  |
| Data consolidation |  |  |  |  |
| Data migration / Data backup / Data sharing |  |  |  |  |
| Free of cost |  |  |  |  |
| Predictable performance |  |  |  |  |

Copy modes

- **Full copy** - za każdym razem wszystkie dane są kopiowane ze źródła do tabeli docelowej
- **Incremental copy** - za pierwszym razem kopiuje wszystko, a następnie tylko nowe bądź zmienione dane
 - Watermark
 - CDC

Demo 1 – Full copy

- Konfiguracja Copy Job
- Monitorowanie
- Column mapping
- Integracja z Pipeline
- Variable Library
- Harmonogram

Summary



Demo 2 – Incremental (watermark)

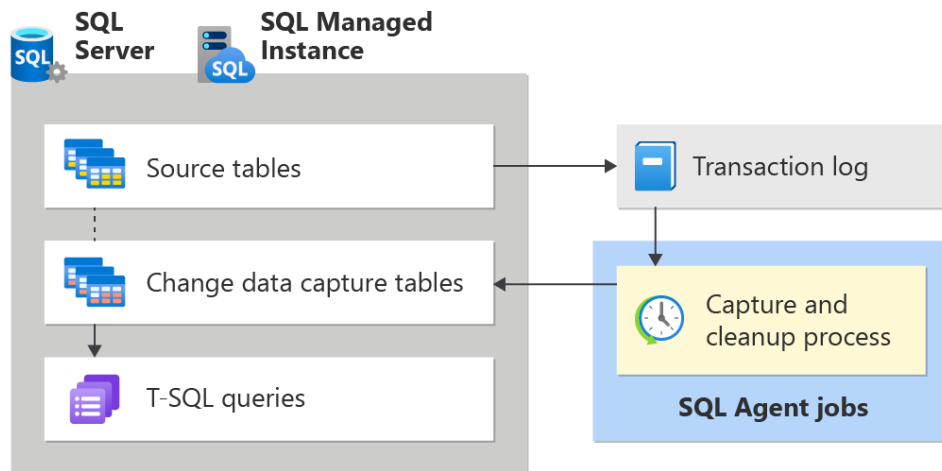
- Key columns
- Incremental column
- Column mapping (guid->string)
- Wstawianie wierszy, modyfikacja
- Kasowanie wierszy z Lakehouse
- Reset zasilania

Summary



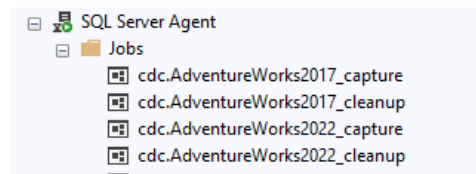
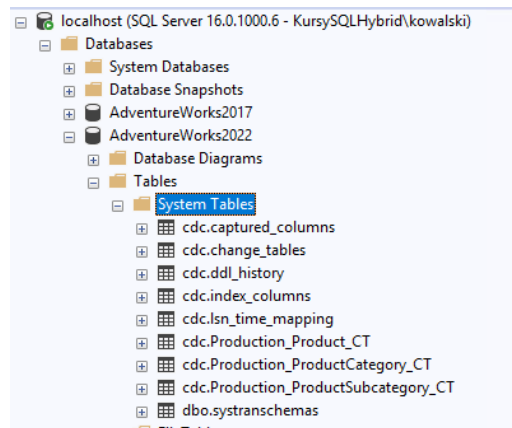
Change Data Capture (CDC)

- SQL Server 2008+, Enterprise Edition
- Zmiany danych są odczytywane z logu transakcyjnego (asynchronicznie)



Change Data Capture Włączenie/ czyszczenie starych danych

- Włączenie śledzenia:
 - Włączenie CDC dla bazy sys.sp_cdc_enable_db
 - Włączenie CDC dla tabeli sys.sp_cdc_enable_table
 - Tylko TSQL (brak wsparcia GUI SSMS)
- Czyszczenie starych (archiwalnych) danych
 - Zadanie SQLServer Agent cdc.AdventureWorks2022_cleanup
 - uruchamia procedurę sys.sp_MScdc_cleanup_job



Change Data Capture `sys.sp_cdc_enable_table`

- Włączenie zmian na poziomie tabeli, wymaga SQL Server Agent
 - `@source_schema/name` - informacje o śledzonej tabeli
 - `@capture_instance` - nazwa tabeli do śledzenia zmian
 - domyślna nazwa `cdc.<schema name>_<table name>_CT`
- `@role_name` - nazwa roli bazodanowej która będzie miała dostęp do historii zmian, NULL – nie korzysta z żadnej dodatkowej roli
- `@supports_net_changes` - domyślnie 0, jeśli tabela na PK
 - 1 - tylko ostatnia zmiana wiersza jest zapisywana
 - 0 – wszystkie zmiany.
- `@captured_column_list` - lista kolumn w których

Change Data Capture Tabela zmian

- \$start_Isn - identyfikator transakcji w logu transakcyjnym
- \$end_Isn - Not supported. Future compatibility is not guaranteed
- \$seqval – identyfikacja instrukcji w transakcji
- \$operation
 - 1 – skasowanie wiersza
 - 2 – wstawienie wiersza
 - 3 – przed zmianą
 - 4 – po zmianie
- \$update_mask – które kolumny zostały zmienione

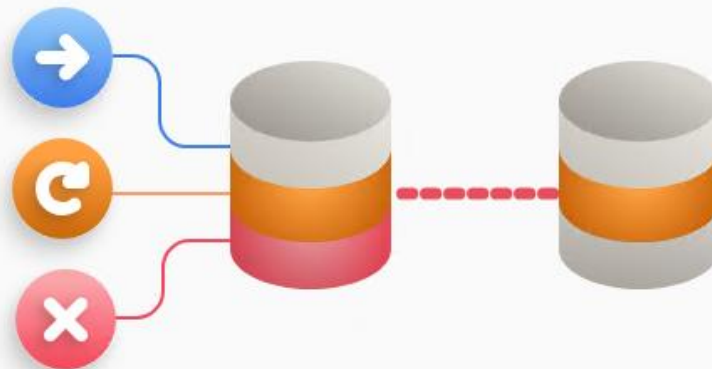
| __Sstart_Isn | __Send_Isn | __Sseqval | __Soperation | __Supdate_mask | ProductID | Name | ProductNumber | MakeFlag | FinishedGoodsFlag | C |
|------------------------|------------|------------------------|--------------|----------------|-----------|----------------------|---------------|----------|-------------------|---|
| 0x0000004B00000F900006 | NULL | 0x0000004B00000F900002 | 3 | 0x00000002 | 2 | Bearing Ball | BA-8327 | 0 | 0 | N |
| 0x0000004B00000F900006 | NULL | 0x0000004B00000F900002 | 4 | 0x00000002 | 2 | Bearing Ball EDITED | BA-8327 | 0 | 0 | N |
| 0x0000004B00000FB00008 | NULL | 0x0000004B00000FB00004 | 2 | 0x01FFFFFF | 1000 | Road Frame Carbon 58 | RF-C58 | 1 | 1 | N |
| 0x0000004B00000FC80011 | NULL | 0x0000004B00000FC80002 | 2 | 0x01FFFFFF | 1001 | Road Frame Carbon 48 | RF-C59 | 1 | 1 | N |
| 0x0000004B00000FC80011 | NULL | 0x0000004B00000FC80006 | 2 | 0x01FFFFFF | 1002 | Mountain Tire 29x2.1 | MT-29X21 | 1 | 1 | N |
| 0x0000004B00000FC80011 | NULL | 0x0000004B00000FC8000B | 2 | 0x01FFFFFF | 1003 | Cycling Helmet Pro | CH-PRO | 1 | 1 | N |
| 0x0000004B00001020000C | NULL | 0x0000004B000010200009 | 1 | 0x01FFFFFF | 1003 | Cycling Helmet Pro | CH-PRO | 1 | 1 | N |

Connectors

| Connector | Source | Destination | Read - Full load | Read - Incremental load (watermark based) | Read - CDC (Preview) | Write - Append | Write - Override | Write - Merge |
|------------------------------|--------|-------------|------------------|---|----------------------|----------------|------------------|---------------|
| Azure SQL DB | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✓ |
| Oracle | ✓ | ✓ | ✓ | ✓ | ✗ | ✓ | ✗ | ✗ |
| On-premises SQL Server | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✓ |
| Fabric Warehouse | ✓ | ✓ | ✓ | ✓ | ✗ | ✓ | ✗ | ✗ |
| Fabric Lakehouse table | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ |
| Fabric Lakehouse file | ✓ | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | ✗ |
| Amazon S3 | ✓ | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | ✗ |
| Azure Data Lake Storage Gen2 | ✓ | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | ✗ |
| Azure Blob Storage | ✓ | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | ✗ |
| Azure SQL Managed | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✓ |

Copy job CDC – connectors

- Źródła
 - Azure SQL DB
 - On-premises SQL Server
 - Azure SQL Managed Instance
 - Fabric Lakehouse table
- Miejsca docelowe
 - Azure SQL DB
 - On-premises SQL Server
 - Azure SQL Managed Instance
 - SQL Database in Fabric (Preview)
 - Snowflake



All tables have CDC enabled.

Copy job uses CDC to automatically detect inserted, updated, and deleted rows. (Preview)

Demo 3 – Incremental (CDC)

- Wprowadzenie do CDC (Change Data Capture)
- Wsparcie dla miejsc docelowych

Summary

