Debezium 2.2 and 2.3

New features and breaking changes

Vojtěch Juránek

Red Hat

Agenda

- New features
- Breaking changes

New features - core

- Topic naming strategies
- Surrogate keys for incremental snapshot
- Notifications (sink, log, JMX)
- Signalling channels (source, Kafka, JMX)
- Partition routing

Topic naming strategies

- Replaces `sanitize.field.names`
- Topic naming strategy
  - `field.name.adjustment.mode`
- Schema naming strategy
  - `schema.name.adjustment.mode`
- Implemented strategies
  - `None` - no adjustments
  - `avro` - replace chars forbidden by Avro by _
  - `avro_unicode` - replace chars forbidden by Avro and _ by Unicode chars

See also https://debezium.io/blog/2023/02/16/debezium-2-2-alpha2-released/

Surrogate keys for incremental snapshot

- Table must have a column which is unique.
- This column can be used as a surrogate key.
- Using multiple columns as surrogate keys is not possible.
- Can be used also for tables which do have primary key, e.g. when the primary key consists of multiple columns.
- Example:

```
{
  "data-collections": [ "public.mytab" ],
  "surrogate-key": "customer_ref"
}
```

See also https://debezium.io/blog/2023/03/08/debezium-2-2-alpha3-released/

Notifications

- Possible notifications channels, set via `notification.enabled.channels`:
  - sink channel (can be used also outside Kafka)
  - log
  - JMX
  - custom notification channel
- Currently notification available for initial and incremental snapshots.

See also:

- https://debezium.io/blog/2023/06/27/Debezium-signaling-and-notifications/
- https://debezium.io/blog/2023/07/10/custom-http-signaling-notification/
- https://debezium.io/blog/2023/10/05/Debezium-JMX-signaling-and-notifications/
- https://debezium.io/blog/2023/05/15/debezium-2-3-alpha1-released/
- https://debezium.io/blog/2023/06/21/debezium-2-3-final-released/

Signalling channels

- Possible signal channels, set via `signal.enabled.channels`:
  - database table
  - Kafka topic
  - JMX
  - file
  - custom channel

- For incremental snapshot, signalling table is still needed (except read-only MySQL).

See also:

- https://debezium.io/blog/2023/06/27/Debezium-signaling-and-notifications/
- https://debezium.io/blog/2023/07/10/custom-http-signaling-notification/
- https://debezium.io/blog/2023/10/05/Debezium-JMX-signaling-and-notifications/
- https://debezium.io/blog/2023/05/15/debezium-2-3-alpha1-released/
- https://debezium.io/blog/2023/06/21/debezium-2-3-final-released/

Partition routing

- SMT which allows routing of the records to a partition based on specified (one or more) payload fields.
- Config options:
  - `partition.payload.field` - payload fields to be used to computing destination partition
  - `partition.topic.num` - number of partitions of destination topic
  - `partition.hash.function` - hash function to use used for computing destination partition

See also https://issues.redhat.com/browse/DBZ-5902

New features - Mongo

- Server-side change stream filtering
  - Network bandwidth and CPU savings
- MongoDB sharding support
  - Connections created over Mongo router (mongos) instead of connecting directly to a shared members.
  - `Mongodb.connection.mode`
    - `sharded` - single connections through router
    - `replica_set` - connections to individual members of replica set
- MongoDB incremental snapshot for shards/replica sets.

See also

- https://debezium.io/blog/2023/03/08/debezium-2-2-alpha3-released/
- https://debezium.io/blog/2023/04/03/debezium-2-2-beta1-released/

New features - MySQL

- Parallel execution of initial snapshot
  - Parallel on a table level - single thread always snapshots whole table.
  - `snapshot.max.threads` - number of threads used for initial snapshot.

See also https://debezium.io/blog/2023/03/08/debezium-2-2-alpha3-released/

New features - Oracle

- Reading from read-only Oracle logical standby
  - `internal.log.mining.read.only=true`
- Server side filtering for Oracle
  - `log.mining.query.filter.mode` - possible values are
    - `none` - no filtering, simple query
    - `in` - filtering based on `in` SQL clause, include/exclude lists shouldn't use any regular expression
    - `regex` - filtering based on Oracle `REGEXP_LIKE`, regular expression can be used in include/exclude lists

See also:

- https://debezium.io/blog/2023/01/19/debezium-2-2-alpha1-released/
- https://debezium.io/blog/2023/06/29/debezium-oracle-series-part-3/

New features - Postgres

- Automated replica configuration for Postgres
  - `replica.identity.autoset.values` - comma-separated list of `<fully-qualified-table-name>:<replica-identity>`
- Exactly-once support for Postgres
  - only for the streaming part
  - `exactly.once.support` set to `required`

See also

- https://debezium.io/blog/2023/05/29/debezium-2-3-beta1-released/
- https://debezium.io/blog/2023/06/22/towards-exactly-once-delivery/
- https://debezium.io/blog/2023/06/21/debezium-2-3-final-released/

JDBC sink connector

- New sink connector

```
{
    "name": "mysql-to-postgres-pipeline",
    "config": {
    "connector.class": "io.debezium.connector.jdbc.JdbcSinkConnector",
    "tasks.max": "1",
    "connection.url": "jdbc:postgresql://postgres:5432/postgres",
    "connection.username": "postgres",
    "connection.password": "postgres",
    "insert.mode": "upsert",
    "delete.enabled": "true",
    "primary.key.mode": "record_key",
    "schema.evolution": "basic",
    "topics": "dbserver1.inventory.customers"
    }
}
```

See also:

- https://debezium.io/documentation/reference/2.3/connectors/jdbc.html
- https://debezium.io/blog/2023/04/03/debezium-2-2-beta1-released/

Minor breaking changes

- Temporal column using `ZonedTimestamp` doesn't truncate zero micro/nanoseconds:
  `2023-01-19T12:30:00.123Z -> 2023-01-19T12:30:00.123000Z`
- `ssn` in Oracle connector source info block has changed from `int32` to `int64`.
- MySQL `database.ssl.mode` and Postgres `database.ssl.mode` defaults to `preferred` resp. `prefer` - Debezium tries to connect over SSL and eventually fails back to non-secure connection.

See also

- https://debezium.io/blog/2023/04/20/debezium-2-2-final-released/
- https://debezium.io/blog/2023/06/21/debezium-2-3-final-released/

Thank you!

Red Hat