2018/11/08 - PostgreSQL Migration

Wednesday, September 5, 2018   3:17 PM

Oracle - PostgreSQL migration

- NCSS Oracle Architecture:
    - 3 DC DBs containing full data (PI,HT,IR,WF). 2 at UCB (1 for serving data to public) and 1 in MP. DC DBs are used for permanent archiving (other catalogs import/export, metadata import/export, continuous & event waveforms import/export, continuous waveforms QC).
    - 2 RT DBs containing recent data for a subset of tables (PI,IR). 1 at UCB and 1 in MP. Event and processing information are replicated to DC DBs. One of the RT DB is flagged as master.
    - Our AQMS system is not typical of other installations. Pete introduced some NC specific components. Long distance between replicated nodes.
    - If connectivity goes down between Berkeley and Menlo Park, we need the ability to have the two RT DBs being master on each side simultaneously. When the link goes back up, at the database level, replication should reconcile all master DBs automatically.
- UW PostgreSQL Architecture:
    - 2 PP DBs & 2 RT DBs.
    - The master PP DB subscribes to the 2 RT DBs. One of the RT DB is flagged as master.
    - The 2nd PP DB does streaming replication of the master PP DB. It does not subscribe to the RT DBs and is read-only.
    - When doing a role switch, the new master PP DB subscribes to the RT DBs and the new slave PP DB starts streaming replication from the new master DB.
- Oracle Advanced Replication:
    - All our databases are currently running Oracle 12.1. Premier support for this version ended on July 2018 but Oracle waived Extended support until July 2019, possibly until July 2021.
    - Starting in Oracle Database 12c release 2 (12.2), the Advanced Replication feature is desupported in its entirety. Oracle recommends to migrate to Oracle GoldenGate which requires a separate license. Often, desupported features are still there and can be used at your own risk with no support. Premier support for Oracle 12.2 is available until March 2022 (Extended support until March 2025).
- Oracle Features:
    - Snapshot replication from RT DBs to DC DBs. Master can have multiple snapshot sites.
    - Multi-master replication. New nodes can be added without suspending replication.
    - Can PostgreSQL replication (subscribe, streaming) queue up transactions if link is down?
    - Partitioning.
    - Stored procedures (not supported in PostgreSQL).
    - "setuid" for stored procedures (caller vs owner).
    - Synonyms (used extensively for EGS data set, not supported in PostgreSQL).
    - Views, sequences, database links.
    - Performance, scalability, robustness.
- DB Language Interface:
    - Schema differences (data types).
    - OTL (C++).
    - DB stored procedures (PL/SQL, java body).
    - Perl modules (DBI/DBD).
    - Pro*C (internal tools e.g. dbselect, Populate_WF, …).
    - Java (JDBC) & SQLJ.
    - PHP.
- How To Migrate:

- Implement full set of parallel systems. Doing it as a piece meal is not an option.
- Implement hybrid system:
    - Migrate only the RT DB in MP with no DC DB in MP. Implement application based replication between Oracle and PostgreSQL (when an event is declared done).
    - Migrate the RT DB in MP to Oracle standard edition. Replicate it to DC DB at UCB.
- Time frame:
    - Linux migration took several years to implement.
    - On the data center side, there are tons of applications accessing the DB that will need to be ported.