Postgresql

Matthias COLIN

Histoire SQL

- 1972-1974: Algèbre Relationnelle (Recherche IBM)
- 1975 : SEQUEL => SQL (Structured Query Language), IBM
- 1979 : Oracle Database, 1er SGBDR (RDBMS) commercial
    - Système de Gestion de Base de Données Relationnelle

Autres :

- IBM DB2, Informix (presque disparu)
- Microsoft SQL Server
- MySQL (racheté par Oracle) / MariaDB (OpenSource)

Standard SQL

- 1986 1ere version du standard
- …
- 2023 dernière version du standard

Les éditeurs suivent +/- ce standard

https://en.wikipedia.org/wiki/SQL#Standardization_history

History

Michael Stonebraker, University of California, Berkeley
Developed by PostgreSQL Global Development Group

- 1985 : post-ingres project
- 1989 : postgres v1
- 1994 : postgres95 with SQL instead of POSTQUEL, client psql
- 1996 : project renamed PostgreSQL + website postgresql.org
- 2016 : PostgreSQL 9.6
- 2017 : PostgreSQL 10
- 2018 : PostgreSQL 11
- 2019 : PostgreSQL 12
- 2020 : PostgreSQL 13
- 2021 : PostgreSQL 14
- 2022 : PostgreSQL 15
- 2023 : PostgreSQL 16
- 2024 : PostgreSQL 17

Maintained versions

Assets

- RDBMS (SGBDR)
- Compliance standard SQL (last version 2023): 99.99%
- Open-source (C/C++)
- Multi-platform
- Object, XML, JSon, Array
- Programming Languages
- Extensions : PostGIS, …
- Clients : PgAdmin4, psql (CLI)

Architecture PostgreSQL

- Logic:
  - Database(s) : postgres by default
    - schema(s) : public (default, transparent, non secure)
      - Variable search_path = “$user”, public (by default)
      - Objects: table, view, sequence, index, …
        - stored code : function, trigger (2 parts): language plpgsql, python, C, perl, ...
        - types
        - materialized view, foreign table
        - FDW (Foreign Data Wrapper)
      - NB: no synonym (schema+view)
  - User/role : postgres (user admin)
- Physique : processus, système de fichiers

Types, functions, operators

https://www.postgresql.org/docs/current/datatype.html

- Numeric types :
  - smallint/smallserial, integer/serial, bigint/bigserial : intervalles
  - numeric/decimal : integer or decimal with number of digits fixed
  - real/double precision : float numbers IEEE754
- temporal : timestamp, date, time, interval, with/without timezone
- boolean
- char(n), varchar(n), text
- xml, json/jsonb
- bytea : binary data
- geometric types (and more with PostGis)
- array version of each type : integer[]

Max size of data : 1GB (clob/blob : provided by an extension)

Primary Key

- smallserial, serial, bigserial (resp smallint, int, bigint)
  - sequence created and ids taken from the sequence by default
- Identity constraint (standard SQL)
  - sequence created

Functions and operators

Complete list here:
https://www.postgresql.org/docs/current/functions.html

- User : CURRENT_USER
- Conditional : case, coalesce, nullif
- Temporal : CURRENT_DATE, CURRENT_TIME, CURRENT_TIMESTAMP
  - DATE_STYLE parameter
  - Formats de types temporels et numériques : https://www.postgresql.org/docs/current/functions-formatting.html

Server PostgreSQL

- 1st process :
  - postgres.exe -D directory_data (windows)
  - postgres -D directory_data (unix/linux)
- Multi-process server with shared memory
- Life cycle controlled by a service or tool pg_ctl : start/stop/restart/reload
- Create new server: initdb
  - encoding
  - locale
  - wal size
- Documentation:
  - https://www.postgresql.org/docs/current/app-pg-ctl.html
  - https://www.postgresql.org/docs/current/app-initdb.html

Data directory

Files :

- postgresql.conf : main configuration
- pg_hba.conf : access filters (user, host, database, authentication)

Directories :

- base : data of each database (oid as its name)
- global : global catalog (databases, users, tablespaces)
- log
- pg_tblspc : links to external tablespaces
- pg_wal : journaling files

Backup/Restore

https://www.postgresql.org/docs/current/backup.html

- Dump: backup, extraction
  - pg_dump (1 database), pg_dumpall (server, multi-databases, users)
  - psql or pg_restore to restore dump
- External File System Tool (zip, rar, tar, …, cp, ...) : système arrêté
- PITR : Point In Time Recovery (uses files WAL)
  - Setup archiving WAL files
  - Backup database : pg_backup (copy data directory + tablespaces + tag WAL)
  - Recover database : 9 points process to follow

Tools : pg_dump/pg_dumpall

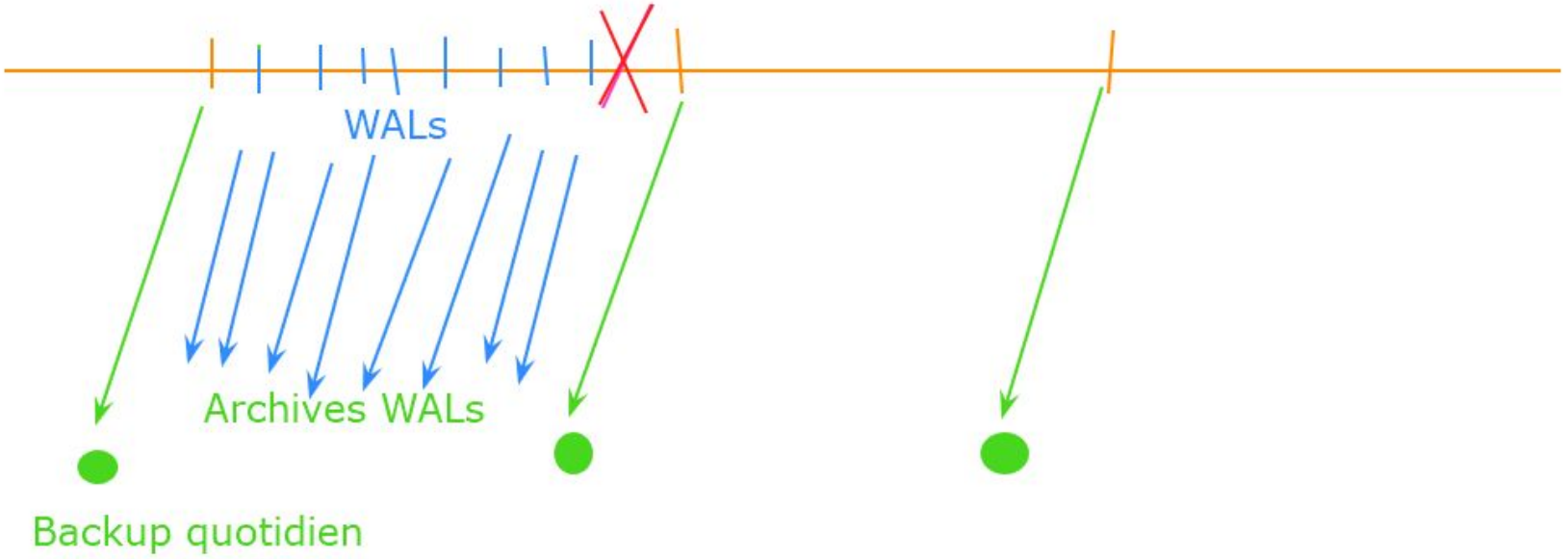
- Format : plain (SQL)
  - Backup with pg_dump => restore with psql
- Format : custom, directory, tar
  - Backup with pg_dump => restore with pg_restore
- Main switches :
  - -F format : plain, tar, directory, custom
  - -E encoding
  - -c --clean + --if-exists : drop objects before recreating
  - Data only : -a --data-only + --disable-triggers (--inserts/--column-inserts for insert vs copy)
  - DDL only : -s --schema-only
  - Selection : -n -N (schema) -t -T (table), include or Exclude

Continuous Archiving Logs

https://www.postgresql.org/docs/current/wal-configuration.html

- WAL : Write Ahead Log, journalization files
  - Switch on archiving in postgresql.conf
  - Choose command or script to archive : copy, cp, scp, ftp, …
  - Choose command or script to restore archives
  - pg_switch_wal function
- Current WAL is tagged backup while backuping database
- File recovery.conf => recovery.done (recovery.signal PG 12+)
  - control recover process
  - parameterized in postgresql since PG 12

Backup + Continuous archive WALs



Vacuum

- https://www.postgresql.org/docs/current/routine-vacuuming.html
- Each table, index, sequence has at least:
  - 1 file of data (relfilenode)
  - 1 file FSM : free space map (fragmentation)
  - 1 file VM : visibility map (current tuples used in transactions)
- Big table, index can use:
  - Several files
  - Toast technique
- Vacuum : simple or combined with analyze and full mode
- Autovacuum: parameters in postgresql.conf

Indexation

https://www.postgresql.org/docs/current/indexes.html
https://www.postgresql.org/docs/current/sql-createindex.html
https://www.postgresql.org/docs/current/sql-reindex.html

- B-tree: =, <, between, like, ~ (begin pattern fix: 'Star%', '^Star')
- hash: =
- GiST: GiS index with GiS operators
- SP-GiST: more complex GiS index
- GIN: inverted index (for array, full text search)
- BRIN: Block Range Indexes

Query Planner

https://www.postgresql.org/docs/current/runtime-config-query.html
https://www.postgresql.org/docs/current/planner-stats.html
https://www.postgresql.org/docs/current/runtime-config-query.html

- Explain, Explain Analyze
- Statistics:
  - reltuples, relpages: in pg_class
  - others in pg_statistics (pg_stats for human read)
- Just In Time compilation

Optimize Plan

- Optimize plan:
  - Adjust costs
  - Increase default_statistics_target
  - ALTER TABLE SET STATISTICS
  - Create statistics (dependencies, ndistinct)
- Supervision Queries:
  - extension/view: pg_stat_statements
    https://www.postgresql.org/docs/current/pgstatstatements.html

Supervision

- Database:
  - service or pg_ctl tool or signal: start|stop|restart|reload
  - pg_reload_conf function
  - pg_rotate_logfile function
- Transactions : pg_lock
- Sessions : pg_stat_activity
  - Kill session : pg_terminate_backend
  - Kill request : pg_cancel_backend
  - Tutorial: https://wiki.postgresql.org/wiki/Lock_Monitoring
- Disk : https://www.postgresql.org/docs/current/functions-admin.html
  - functions computing size of database, schema, tablespace, table, index
  - functions to obtain filepath or oid from object name (table, index, sequence, …)

Configuration

- Postgresql.conf
  - listener/port
  - demory
  - wal, disk writing
  - directories, logs
  - statistics, costs
  - autovacuum
  - restore/recover
- Pg_hba.conf : protocol, database, user, computer, authent

https://www.postgresql.org/docs/current/config-setting.html
https://wiki.postgresql.org/wiki/Performance_Optimization