

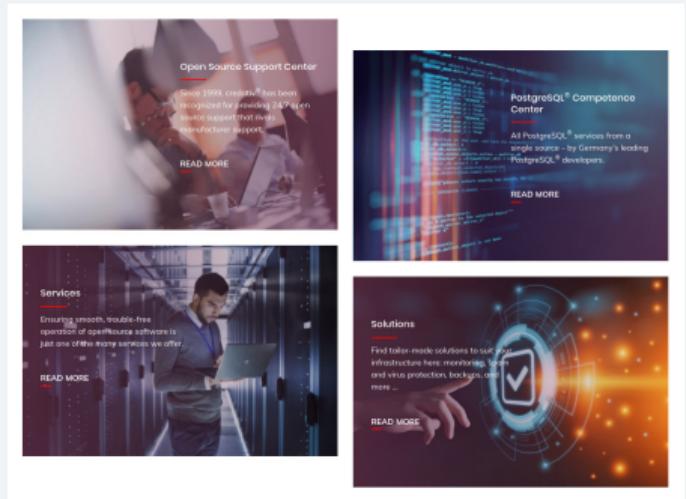
Migrate to PostgreSQL with credativ-pg-migrator

A unified offline migration framework
for Oracle, Informix, Sybase, DB2,
MS SQL, MySQL/MariaDB, SQL Anywhere

Josef Machytka <josef.machytka@credativ.de>

2025-10-31 - credativ Tech Talk

- Founded 1999 in Jülich, Germany
- Close ties to Open-Source Community
- More than 40 Open-Source experts
- Consulting, development, training, support (3rd-level / 24x7)
- Open-Source infrastructure with Linux, Kubernetes, Proxmox
- Open-Source databases with PostgreSQL
- DevSecOps with Ansible, Puppet, Terraform and others
- Since 2025 independent owner-managed company again



- Professional Service Consultant - PostgreSQL specialist at credativ GmbH
- 33+ years of experience with different databases
- PostgreSQL (13y), BigQuery (7y), Oracle (15y), MySQL (12y), Elasticsearch (5y), MS SQL (5y)
- 10+ years of experience with Data Ingestion pipelines, Data Analysis, Data Lake and Data Warehouse
- 3+ years of practical experience with different LLMs / AI / ML including architecture and principles
- From Czechia, living now 12 years in Berlin

- **LinkedIn:** linkedin.com/in/josef-machytka
- **Medium:** medium.com/@josef.machytka
- **YouTube:** youtube.com/@JosefMachytka
- **GitHub:** github.com/josmac69/conferences_slides
- **ResearchGate:** researchgate.net/profile/Josef-Machytka
- **Academia.edu:** academia.edu/JosefMachytka
- **Sessionize:** sessionize.com/josefmachytka

All My Slides:



Recorded talks:



Main Reasons for Migration

- **Cost pressure from proprietary databases**
 - Per-core licensing, options/add-ons, support contracts
 - Growth in CPU cores or cloud instances directly increases license cost
 - Vendors pushing expensive cloud offerings
- **Desire to avoid vendor lock-in**
 - Flexibility to choose hosting environments (on-prem, cloud, hybrid)
 - Freedom to customize and extend the database
 - Avoiding dependence on a single vendor's roadmap and pricing



AI Images by DeepDreamGenerator

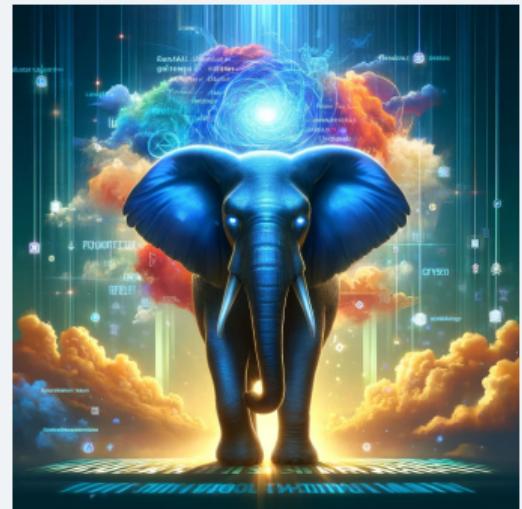
Main Reasons for Migration

- **Legacy database versions & missing features**
 - End of support, security fixes only, no new functionality
 - Some engines are effectively “frozen”, no modern features
 - Vendors focus on different flagship products
 - Aging infrastructure, Databases tied to old hardware / OS versions
- **Knowledge erosion**
 - Vendor and community resources slowly disappearing
 - Scarcity of experts for legacy systems
 - Original developers and DBAs are gone
 - Business logic encoded in stored procedures and triggers



Why yet another migration tool?

- Open source vs. commercial solutions with vendor support
- Existing tools address different migration scenarios
- Traditional tools: planned downtime required
- CDC-based tools: near-zero downtime migrations
- In production, minimum 75% of migrations are offline-based
- Limited flexibility for custom schema transformations
- Complex business logic often requires manual adaptation
- Open source project development cycles can be slow
- Tool maintenance and long-term support concerns



Meet credativ-pg-migrator

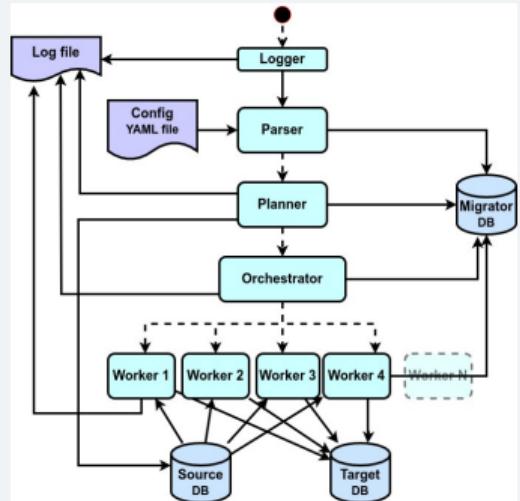


- Inspired by pgloader, designed for enterprise migrations
- Offline migration tool optimized for large datasets
- Python-based with JDBC/ODBC and native DB connectors
- Modern languages have limited support for older DBs
- Modular architecture with pluggable DB adapters
- Built on stable, well-documented libraries
- Uses pyodbc, JayDeBeApi, cx_Oracle, psycopg2, etc.
- Single CLI tool for multiple source databases



credativ-pg-migrator Architecture

- Classes: Logger, Parser, Planner, Orchestrator, Workers
- YAML configuration file, text log file
- Runs multiple parallel workers, one reader and writer per table
- Performance depends mostly on source database I/O capacity
- Migration protocol tables track all object states
- Detailed progress monitoring and error reporting



credativ-pg-migrator Features



- Target database is always PostgreSQL
- Supports 7 different source databases
- Oracle, Informix, Sybase ASE, SQL Anywhere, IBM DB2 LUW
- MS SQL Server, MySQL/MariaDB, (+ PostgreSQL)
- Complete schema migration: tables, views, constraints
- Configurable migration scope: schema-only, data-only, both
- Granular object control: include/exclude by pattern
- Function and trigger migration (database-dependent)



credativ-pg-migrator - Customizations

- Built-in data type and SQL function mapping engine
- Flexible customization options:
 - Custom data type transformations
 - Custom default value mappings
 - Custom SQL functions substitutions
 - Cross-database object reference resolution
 - Selective data migration with WHERE conditions:
 - - Time-based filtering (created_at >= date)
 - - Primary key range filtering (ID BETWEEN x AND y)



- Automated conversion of Informix functions/procedures/triggers
- Configurable conversion scope with include/exclude patterns
- Success rate of code conversion 80 to 90%
- Common conversion issues: missing objects, renamed columns
- Manual review required for complex SQL statements
- Conversion of code can be added for other DBs too



Future plans for credativ-pg-migrator



- Active development roadmap includes:
 - Automated migration testing framework
 - Pre-migration source database analysis
 - Table partitioning analysis and recommendations
 - Configurable target table partitioning
 - Materialized view migration support
 - Extended code conversion for stored procedures
 - Additional source database connectors on demand



PostgreSQL to PostgreSQL migration

- Intended for special PostgreSQL-to-PostgreSQL scenarios
- Enables custom data transformations and selective migrations
- Migrates database subsets with precision control
- Consolidates multiple databases into schema-based structure
- Splits schemas into independent database instances
- Restructures data models across schemas and databases
- Streamlines complex migrations vs. manual approaches



Source Databases – Oracle

- Oracle V2 released in 1979 was the first commercial SQL RDBMS
- Oracle pioneered many features now common in other databases
- Currently at version 23ai with AI capabilities
- Advanced features: vector search, JSON, XML, advanced PL/SQL language
- Main issues: licensing payments - expensive per core licensing
- Many cases of old Oracle versions running on old hardware
- Old apps contain many deprecated legacy features like old data types
- Closest to PostgreSQL in terms of features and SQL syntax
- PostgreSQL took the best from Oracle, simplified a lot of things



Source Databases – Informix



- Created by Informix (Information on Unix) Software in 1980s
- Acquired by IBM in 2001, development now by HCL
- Version 15 released in November 2024
- Improvements mainly in partitioning, size of data chunks
- Adds new functionality through extensions
- Very similar to Oracle, very good database for OLTP
- credativ-pg-migrator supports full migration from Informix
- Including conversion of Functions, Procedures, Triggers



- Created 1987 by Sybase, acquired by SAP in 2010
- Pioneered client-server architecture, stored procedures
- Very good performance "out of the box"
- Uses special "Tabular Data Stream" (TDS) protocol
- No big new features lately, last version 16.1 in 2022
- Small changes added via Service Packs
- Future unclear, SAP is focusing on SAP HANA



Source Databases – SQL Anywhere



- Created as Watcom SQL in 1992,
- 1995 acquired by Sybase -> Sybase SQL Anywhere
- Sybase invested in synchronization and replication
- DB used mainly in embedded systems/ mobile apps
- Great cross-platform support - Windows, Linux, macOS
- 2010 Sybase acquired by SAP -> SAP SQL Anywhere
- Last stable release 17 from 2015



- First released in 1983, based on System R project
- System R was the source of SQL and relational DBs
- Runs on Linux, Unix, Windows (LUW), and mainframes (z/OS)
- LUW and z/OS are different products with specific features
- Different versions & code base - LUW 11.5, z/OS 13
- SELECT 1 FROM SYSIBM.SYSDUMMY1; - common idiom
- credativ-pg-migrator supports DB2 LUW (tested on 11.5)
- Migrates tables, data, indexes, constraints, views
- Uses native DB2 system tables - SYSCAT.TABLES, COLUMNS,...



- Created as a fork of Sybase ASE for OS/2 in 1989
- Later adapted to Windows NT platform and rewritten
- Many advanced features for memory optimization
- Columnar store for analytics, Hekaton for in-memory
- Vector data types and vector search
- Rich tools for development, reporting, BI
- PolyBase for data virtualization
- Simplifies data access to external data sources
- Main issues: expensive licensing payments



- Created in 1995 by MySQL AB
- Acquired by Sun in 2008, by Oracle in 2010
- Oracle acquired Innobase in 2005, added InnoDB engine
- Forked as MariaDB in 2009 by the original developers
- Over years, paths of both databases diverged
- Specific features depend on storage engine
- InnoDB uses redo log (transactions - changes in pages)
- Atomicity through undo logs (rollback segment - state before)



What is Different

- Special unique row identifiers ROWID (Oracle, Informix)
- Frequent usage of synonyms (for deployment, ETL tasks)
- Big differences in partitioning
- Over time different character sets allowed
- DBs sometimes convert char sets on fly

- Legacy code often enforces strict locking patterns
 - Usually use REDO / UNDO logs, updates in place
 - CURSOR WITH HOLD, FOR UPDATE cursors
 - Explicit COMMITs in loops



What is Different

- Differences in SQL SELECT syntax
 - JOINs without ON clause (Oracle, Informix)
 - Comma delimited lists of tables
 - Join conditions in WHERE clause
 - Special syntax for LEFT OUTER JOIN (Informix)
 - Pseudo tables for selects without FROM
 - - Oracle: SELECT 1 FROM DUAL
 - - IBM DB2: SELECT 1 FROM SYSIBM.SYSDUMMY1
- Hierarchical queries with CONNECT BY PRIOR
- Hints in SQL statements - Oracle, MS SQL Server



What is Different

- No custom schemas, only owners or special schemas
- Simple cross-database queries and references
- Theoretically much longer identifiers up to 128 chars
- But old versions limited 18 or 30 chars
- PostgreSQL shortens identifiers to 63B - no error, just notify
- Higher limits for CLOBs, BLOBs, Large Objects
- Usually much smaller limits for VARCHARs
- Foreign Keys between different data types



What is Different

- All of them have procedural language similar to PL/pgSQL
- Oracle PL/SQL, Informix SPL, Sybase Transact-SQL
- Some have 2 different SQL and PL dialects
- Allowed more freedom in formatting, explicit blocks etc.
- Can return multiple different result sets from a function
- Allow to return multiple values from a function
- Usually allow global variables, but with different scope
- Differences in triggers due to REDO/UNDO vs MVCC



Summary

- Migrates data models from 7 different databases:
- Oracle, Informix, Sybase ASE, SQL Anywhere
- IBM DB2 LUW, MS SQL, MySQL/MariaDB
- Replacements of Data types, Default values, Remote references
- Limitations for Data migrations (based on dates, etc.)
- From Informix migrates also triggers, stored procs/funcs

- If you are interested in using the tool, please contact us
- Every database is somehow special - we can add on demand feature you need
- We help you migrate, you will help improve this tool for the community



GitHub repository - github.com/credativ/credativ-pg-migrator

Released under the GNU General Public License, version 3 (or any later version)

Available also on PyPi - pypi.org/project/credativ-pg-migrator/

I created this tool and many thanks to my colleague Michael Banck
for all the hard work with open sourcing & publishing it!



Thank you for your attention!



Questions?

All my slides



Recorded talks

