



Greenplum® to WarehousePG: A technical playbook

What a binary-compatible swap to EDB Postgres® AI for WarehousePG actually replaces, what stays exactly where it is, and how to find out whether your Greenplum 6.x or 7.x cluster qualifies.

What this playbook covers

This is for the team that runs the Greenplum cluster. It answers one question in detail: what does a binary swap actually do to my estate? Then two follow-ups: does my version qualify, and what breaks?

Most migration playbooks are about moving data. This one largely is not, which is what makes the engineering conversation different. The business and governance case is a separate companion document.

What a binary swap actually is

EDB Postgres AI for WarehousePG is an Apache 2-licensed fork of the formerly open-source Greenplum Database. EDB's documentation states binary compatibility with the Greenplum versions most widely deployed, 6.x and 7.x.

Because it is a fork rather than a re-implementation, the on-disk data, the catalog and the cluster layout are already in the shape it expects. The swap replaces the executables.

The procedure, in outline

Stop the Greenplum cluster, remove the old packages, install the WarehousePG packages on every host, and restart the database. It is the same stop, replace and restart sequence the WarehousePG documentation gives for its own minor-version upgrades, which replace the binaries without a data migration.

For backward compatibility, current WarehousePG 7 releases also keep symlinks at the old `/usr/local/greenplum-db` paths. Everything else in this document is about establishing whether the swap applies to your cluster and proving that it worked.



Version scope: establish this first

Your Greenplum version decides whether you are looking at a one-step swap, a two-step project, or an open question.

YOUR CLUSTER	PATH	NOTE
Greenplum 7.x	Direct swap	The single-step case. Binary compatibility is documented for this line.
Greenplum 6.x	Two steps	Swap to WarehousePG 6, then upgrade the cluster to WarehousePG 7.x. EDB professional services provide version-upgrade assistance for this path. Plan it as a project, not a maintenance window.
Apache Cloudberry or another fork	No published claim	EDB publishes no binary-compatibility claim for Cloudberry or any other Greenplum fork. That does not mean it is impossible; it means nobody has committed to it in writing, so treat it as an open question for an assessment rather than an assumption.
Greenplum 4.x / 5.x	No published claim	Outside the documented 6.x/7.x scope.

Source: EDB Postgres AI for WarehousePG documentation (binary compatibility with Greenplum 6.x and 7.x; upgrading from WarehousePG 6 to WarehousePG 7).

What step two of the 6.x path involves

Step one is the swap itself, from Greenplum 6.x to WarehousePG 6, which is packaged and documented in its own right. Step two is a major-version upgrade, and it is not a binary swap. The documented path from WarehousePG 6 to 7 initializes a new WarehousePG 7 cluster and moves the data with gpbackup and gprestore, and the documentation warns that a substantial number of changes between 6 and 7 can affect existing applications.

So on a 6.x estate, the no-data-movement property applies to step one only. Budget step two for data movement, storage headroom and application testing, and review the documented 6-to-7 changes, including resource group behavior, before you schedule it.



What stays, and what does not

One row is marked not included. Scope it explicitly: it is the row most likely to change a plan.

ELEMENT	STATUS	NOTE
On-disk data	Unchanged	No data movement. The reason cutover is measured in hours rather than weekends.
Schema, including distribution keys	Unchanged	No redesign, no redistribution, no partitioning rework.
Client applications, ports, drivers	Unchanged	Postgres wire protocol throughout. Applications do not need to know.
BI, ETL and DBA tooling	Unchanged	Including regularly scheduled administration jobs.
MPP topology	Same architecture	Coordinator plus segments, interconnect, segment mirroring, standby coordinator.
MADlib, PostGIS, PL/R, PL/Java, PL/Container	Supported	In-database machine learning, geospatial objects and additional procedural languages, all listed in the WarehousePG documentation.
PXF	Supported	External tables to HDFS and S3-compatible object storage; ORC, Avro, Parquet and other formats. Has its own documentation set.
Resource management	Supported	Both resource queues and resource groups. One scheme is active at a time, set with the <code>gp_resource_manager</code> parameter.
gpfdist	Supported	External-table loading through the <code>gpfdist://</code> and <code>gpfdists://</code> protocols is documented.
Proprietary Tanzu add-on modules	Not included	Commercial modules from the proprietary Tanzu Greenplum suite are not part of an Apache-licensed fork. What replaces each one is a module-by-module answer, so get it from EDB in writing. This is the most likely technical objection; scope it explicitly.

Source: EDB Postgres AI for WarehousePG documentation: key features, platform requirements and supplied modules, workload management (resource queues and resource groups), external tables (gpfdist), PXF.



Staging, then promotion

The swap is short. The preparation is not. Euronext FX describes months of planning before a cutover that took hours. That ratio is the normal shape of this project, and a plan that inverts it is a plan to discover something unpleasant in production.

1

Establish the facts

Greenplum version, fork lineage, extension inventory, commercial add-on dependencies, resource-management configuration, external-table loading method. Six answers, and together they determine whether section 02's direct path applies.

2

Swap in staging

Perform the package swap on a staging cluster that matches production in extension set and configuration. Time it. That measured duration is your cutover window input.

3

Validate in staging

Section 05. Run the real workload, not a smoke test.

4

Promote to production

With the validation evidence, a tested rollback path, and a named go/no-go authority. The governance companion document covers that structure.

On the phrase “zero migration”

It means no data movement and no schema change. It does not mean no project.

Used precisely it is accurate and useful; used loosely it sets an expectation that the first planning meeting will contradict. On a Greenplum 6.x estate it describes step one of the two-step path, not step two.

Source: Euronext FX customer story, EDB (months of planning, cutover in hours).



Proving the swap worked

A binary swap removes the failure modes that dominate ordinary migrations: there is no data movement to reconcile and no schema conversion to review. What remains is narrower, which means it can be tested properly rather than sampled.

Before promotion

- **Extension inventory parity.** Every extension present and at the expected version. This is where a commercial add-on dependency surfaces if step 1 missed it.
- **Catalog and object counts.** Databases, schemas, tables, views, functions, external table definitions: source versus swapped, expecting exact parity rather than approximate.
- **Query plan comparison.** Capture plans for the queries that matter before the swap and compare after. Same engine lineage means plans should be stable; anything that moves is worth understanding before promotion.
- **Workload run.** The real reporting and ETL workload, not a synthetic benchmark, with runtimes compared against the pre-swap baseline.
- **Connection and tooling check.** BI tools, ETL jobs, scheduled administration tasks, monitoring. These should be unaffected; confirming that is a short exercise and a useful piece of evidence.

At promotion

- **Timed rollback rehearsal.** Reinstalling the previous packages is the rollback. Rehearse it, time it, and record the duration.
- **Backup verified before the window.** Verified, not merely taken.
- **Named go/no-go authority.** One person, with criteria agreed in advance.



What the numbers measure

Two figures from one commissioned study. Read each against the comparison it actually makes.

Up to 58% lower TCO against Snowflake, the study's headline cost comparison	Up to 52% more consistent under high concurrency, same comparator set	10TB TPC-DS benchmark scale under a high-concurrency workload
FIGURE	WHAT IT COMPARES	SOURCE
Up to 58% lower TCO	Annual platform cost: EDB Postgres AI for WarehousePG against Snowflake, the study's headline cost comparison. Not a Databricks or Amazon Redshift figure.	McKnight Consulting Group, 2026. 10TB TPC-DS, high concurrency, vs Snowflake, Databricks and Amazon Redshift. Commissioned by EDB.
Up to 52% more consistent under high concurrency	Same study, same comparator set: how far query performance slows as load rises from 1 to 5 concurrent users.	McKnight Consulting Group, 2026. Commissioned by EDB.

Source: McKnight Consulting Group, 2026, commissioned by EDB. 10TB TPC-DS, high-concurrency workload, measured against Snowflake, Databricks and Amazon Redshift.

Read these for what they are

These are cloud-data-warehouse comparisons, not Greenplum comparisons. They answer “where else could this workload go, and what would it cost?” They do not claim WarehousePG is faster than Greenplum, and EDB has published no benchmark that makes that claim.

If someone quotes you a Greenplum-versus-WarehousePG performance figure, ask for the study.



Where to start

1

Establish version and lineage

Greenplum 7.x, 6.x, or a fork. This single answer determines whether you are looking at a one-step swap, a two-step project, or an open question.

2

Inventory extensions and add-ons

Particularly proprietary Tanzu add-on modules, the one row section 03 marks not included. Record your resource-management scheme and external-table loading method too: both carry over, and both belong in the staging checklist.

3

Get the add-on answer in writing

For each proprietary module you depend on, ask what replaces it. If a module is on your critical path, a documented answer is worth more than a verbal one.

4

Swap in staging and time it

The measured duration is the input every subsequent plan depends on.

Request an assessment

EDB offers a packaged binary swap service for this specific motion. Go to enterprisedb.com and request a migration assessment, or contact your EDB account team. If you received this document from an EDB representative, reply to them directly.

enterprisedb.com

Sources

- EDB Postgres AI for WarehousePG documentation: binary compatibility scope (Greenplum 6.x and 7.x installation and upgrade including the WarehousePG 6 to 7 path, administration, workload management, external tables and gpfdist, supplied modules, PXF).
www.enterprisedb.com/docs/warehousepg/latest
- McKnight Consulting Group, 2026, commissioned by EDB: 10TB TPC-DS high-concurrency benchmark against Snowflake, Databricks and Amazon Redshift.
www.enterprisedb.com/resources/mcknight-predictable-analytics-at-scale
- Euronext FX customer story, EDB: cutover duration and planning ratio.
www.enterprisedb.com/resources/customer-story/euronextfx



Confirm scope against an assessment of your own cluster.

Compatibility scope in this document reflects EDB product documentation at time of writing.

This playbook is engineering planning guidance and does not constitute a contractual commitment or a performance guarantee. Compatibility scope reflects EDB product documentation at time of writing; confirm it against current documentation and against an assessment of your own cluster before planning a cutover or making a purchasing decision. Greenplum® is a registered trademark of Broadcom Inc. Tanzu is a trademark of Broadcom Inc. EDB and EDB Postgres AI are not affiliated with, endorsed by, or sponsored by Broadcom Inc. Any references to Greenplum are for comparative, educational, and interoperability purposes only. Apache, Apache Cloudberry and the Apache feather logo are trademarks of The Apache Software Foundation. Snowflake, Databricks and Amazon Redshift are trademarks of their respective owners. Postgres and PostgreSQL are registered trademarks of the PostgreSQL Community Association of Canada and used with their permission. EDB Postgres® AI is a trademark of EnterpriseDB Corporation. © EnterpriseDB Corporation 2026. All rights reserved.

EDB Postgres® AI is the first open, enterprise-grade sovereign data and AI platform, with a secure, compliant, and fully scalable environment, on premises and across clouds. Supported by a global partner network, EDB Postgres AI unifies transactional, analytical, and AI workloads, enabling organizations to operationalize their data and LLMs where, when, and how they need them. For more information, visit enterprisedb.com.

