SOLUTION BRIEF

Converged Analytics

Closing the insights-action gap on a data foundation you own

Executive summary

Modern enterprises increasingly operate across four separate data systems—a transactional database (Postgres®), a real-time analytics layer, a data warehouse, and the lakehouse. While each performs well individually, the costs of fragmentation are compounding: extract, transform, and load (ETL) pipelines consume a substantial share of engineering time, insight latency leaves decisions running on yesterday's data, AI-driven workflows reason from stale context, and most tools built to modernize the stack introduce new constraints, including proprietary formats, cloud-bound infrastructure, and complicated governance.

EDB Postgres AI (EDB PG AI) powers a converged analytics approach, unifying the full data lifecycle on an open foundation that's deployable anywhere. The EDB Postgres Analytics Accelerator (PGAA) extension eliminates ETL pipelines—Postgres transactions replicate continuously to Apache Iceberg, with ClickHouse® delivering sub-second results on live event data, WarehousePG handling petabyte-scale historical analytics, and GPU-accelerated Spark powering the most intensive workloads for organizations with existing Spark infrastructure. Security rules defined at the transactional source are automatically inherited across every engine and every AI agent querying the data.

The business case is compelling: Benchmarks indicate up to 30x faster insights on operational data, up to 99x GPU-accelerated throughput, and up to 58% lower total cost of ownership compared to Snowflake, Databricks, and Redshift. Real-world deployments at MNTN, Kyobo Book Centre, and Euronext FX demonstrate the platform's ability to deliver petabyte-scale performance, cost reductions, and seamless migration.

Rather than replacing one fragmented stack with another, EDB PG AI converges the full data lifecycle on open formats the organization owns—deployable anywhere, and ready for AI agent workflows from day one.

The problem every analytics leader recognizes

You've done the hard work. Postgres handles your operational data. A real-time analytics layer handles event streams and time-sensitive queries. A warehouse stores your historical data for batch analytics. A lakehouse holds your ML training data and archival records. Each system is good at what it does—and each one adds constraints the next has to work around.

Running them together creates two compounding problems.

The first is structural: Most analytical estates have been assembled, layer by layer, on proprietary infrastructure. Your warehouse stores data in a format tied to that vendor's platform—moving it requires migration. Your lakehouse catalog is bound to a platform not chosen with portability in mind. Every compliance audit asks the same questions: Where does the data actually live, and who controls access to it? Governance complexity grows with every ETL pipeline in the stack—each one is another point of failure, another audit question, another surface to secure. The cost of replatforming grows with every year the estate runs on someone else's terms.

The second is operational: ETL pipelines connect the layers—and those pipelines now consume a substantial share of your team's time and budget. According to a Fivetran study, 44% of enterprise data engineering time is spent maintaining data movement pipelines rather than building analytical value. By the time data completes the journey from transaction to insight, it reflects yesterday's reality. And as your organization moves toward AI-driven workflows, that lag creates an entirely new problem: Agents can't act on stale context.

The core challenge: Four stacks, one data estate—no coherent path to action

For VPs and directors of data and analytics, the costs of fragmentation are familiar:

- **Vendor lock-in accumulates with every layer added.** Proprietary data formats require migration to move, infrastructure tied to vendor-controlled regions may not satisfy data residency requirements, and the cost of replatforming grows each year the estate expands on these terms.
- **ETL maintenance crowds out analytical work.** Pipelines break. Schema changes cascade. Engineers who should be building analytical capabilities spend their time keeping data moving instead.
- **Insight latency undermines decisions.** When the most current data available is hours or days old, the analytical layer loses credibility with business stakeholders. Real-time decisions get made on gut feel because the data can't keep up.
- **Separate compute and storage models inflate costs.** Scaling analytical capacity on warehouse or lakehouse infrastructure means scaling both compute and storage together—even when only one is the actual bottleneck. Organizations pay for capacity they aren't using.
- **ETL pipelines are a compliance liability.** Every pipeline that moves data out of the secured Postgres source is a point where governance state can be lost, inconsistencies can go undetected, and compliance exposure accumulates.
- **AI workloads have no coherent data context layer.** Agents and AI applications need a governed, semantically discoverable interface to query analytical data. Fragmented systems can't provide one.

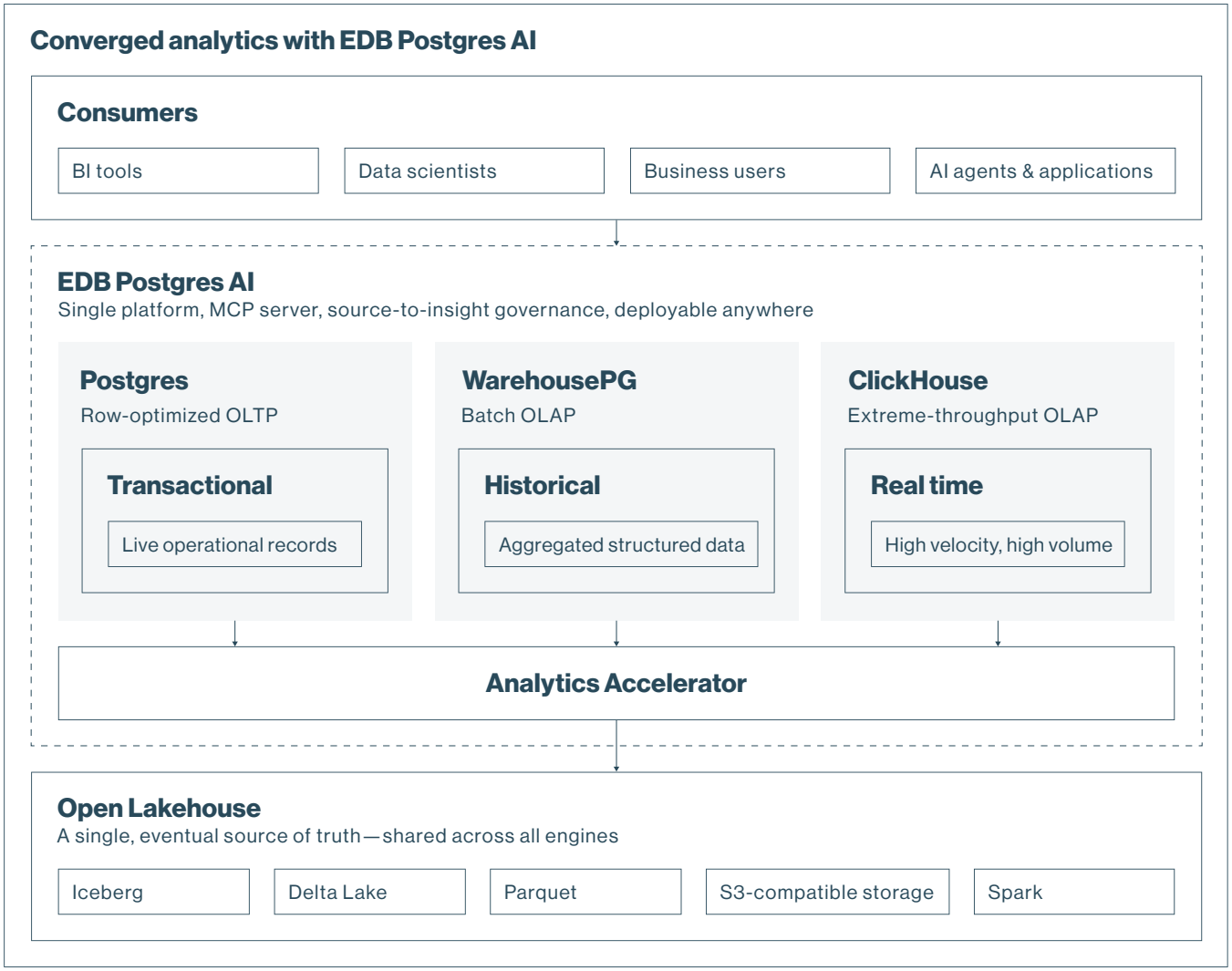
The approach: The entire data estate converged on open, sovereign foundations

EDB PG AI addresses these challenges with a converged analytics approach built on vendor-neutral open standards. Every component reads from a shared Apache Iceberg catalog, deployable on-premises, in sovereign cloud regions, or hybrid.

PGAA keeps that catalog current without ETL—live Postgres transactions replicate continuously, with no batch windows or pipeline maintenance. ClickHouse delivers sub-second results on live event data; WarehousePG handles petabyte-scale batch analytics, reporting, and machine learning. All engines share the same catalog, so every query reflects the current state.

AI agents access the full estate through the MCP server—a standardized, metadata-aware SQL interface that requires no custom integration work.

The result is an open analytical foundation—governed at the source, deployable anywhere, and owned by the organization that built it.



Key capabilities

- **Sovereign control:** EDB PG AI runs on open source engines and writes to Apache Iceberg, a format no commercial vendor controls. Data stays in formats the organization owns, on infrastructure it chooses, with full control over the migration path.
- **Zero-ETL lakehouse sync:** EDB PG AI replicates live Postgres transactions directly to open Iceberg format—continuously, without pipelines. Analytical data is always current. Engineering cycles previously spent on ETL maintenance are redirected to building value.
- **Real-time analytics:** ClickHouse reads natively from the shared Iceberg catalog, returning sub-second results on live event streams—IoT telemetry, fraud scoring, AdTech attribution, and time-series data at millisecond latency.
- **Batch analytics:** WarehousePG handles petabyte-scale batch BI, model training, and data science. In-database ML via MADlib runs models where the data lives, and a built-in Foreign Data Wrapper lets WarehousePG query ClickHouse directly.
- **Analytical query offload:** Run analytical workloads against live Postgres data without affecting transactional performance—single-node DataFusion for interactive queries, or offload to Spark for distributed execution, with optional GPU acceleration for up to 99x faster queries. Compute and storage scale independently, so you pay for what you actually use.
- **Governed foundation:** Governance is enforced at the Postgres data layer using battle-tested access and authorization controls. Data flows directly from Postgres to open Iceberg format without ETL pipelines, eliminating the compliance gaps that accumulate when data moves through multiple proprietary systems.
- **Agent-native SQL interface:** Every analytical asset is semantically discoverable via MCP. AI agents and analytical workflows submit SQL queries through a consistent, metadata-aware interface regardless of the underlying execution engine. No custom API work, no fragmented context.

Business outcomes

EDB PG AI delivers measurable impact across the dimensions that matter most to analytics organizations:

Outcome	Metric	Source
Faster insights on your live operational data	Up to 30x	PGAA benchmark, EDB 2026 (workload dependent)
GPU-accelerated analytical throughput	Up to 99x	NVIDIA Spark RAPIDS via PGAA, EDB 2026 (workload dependent)
Lower TCO vs. Snowflake, Databricks, and Redshift	Up to 58%	McKnight Consulting Group, 2026
Storage footprint reduction on open formats	Up to 5x	Open table formats vs. standard tables and indexes

*Competitive comparisons are based on publicly available information and are subject to change as vendor offerings evolve and new information is made available. All product names, trademarks, and registered trademarks are the property of their respective owners.

Beyond the metrics, the structural shift matters: Analytics teams move from pipeline maintenance to analytical development. Insight latency collapses from hours or days to the current moment, and the underlying architecture is open and sovereign—data lives in formats that the organization controls, on infrastructure of its choosing.

Why EDB PG AI

Most organizations are running their analytical estate on someone else's terms—data in proprietary formats that raise the cost of replatforming over time, infrastructure they don't control, and ETL pipelines that create compliance exposure every time data moves between systems. Addressing any one of these pains with a point solution relocates the problem rather than resolving it.

Four structural properties explain why EDB PG AI's converged analytics approach holds up:

- **Own your data estate:** EDB PG AI is centered on open source at every layer—engines and formats the organization owns, on infrastructure it controls. The platform deploys wherever the organization needs it. No vendor holds the keys.
- **Control your cost model:** Per-core pricing has delivered up to 58% lower TCO vs. Snowflake, Databricks, and Redshift. Lakehouse interoperability adds a 5x smaller storage footprint and up to 18x more cost-efficient storage.
- **Connect every layer without pipelines:** Live Postgres data stays current across every engine without ETL or batch windows. Adding an engine is a configuration, not a project.
- **Be agent-ready from day one:** The MCP server gives AI agents governed access to the full estate—transactions, real-time events, and historical depth—without custom integration.

The result is a data foundation the organization genuinely owns—closing the insights-action gap on infrastructure it controls.

EDB Postgres AI: The sovereign data and AI platform for the agentic enterprise

EDB PG AI brings together a unified data layer, governance, sovereign control and orchestration, and an agent runtime environment, giving enterprises a trusted foundation for AI on infrastructure they own and control. The platform unifies transactional, analytical, and AI workloads in a single Postgres-based architecture—eliminating ETL, data movement, and operational fragmentation. And you choose where and how to deploy: on-premises, cloud, managed, or certified appliance.

The outcome: production-ready sovereign AI in days or weeks, not months.

EDB POSTGRES AI

EDB Postgres® AI (EDB PG AI) is the sovereign data and AI platform for the agentic enterprise. Built on Postgres, the world's leading open source database, EDB PG AI unifies transactional, analytical, and AI workloads in a single governed architecture, on-premises and across clouds. To learn more, visit www.enterprisedb.com.