

Migrating from BigQuery to AWS Lakehouse

How to modernize your Data & AI platform

Karim Berrah

Sr. Analytics GTM, AWS
kaberrah@amazon.com

Spiros Potamitis

Sr. Analytics Specialist, AWS
spirosp@amazon.com

Edijs Drezovs

CEO, GoStack
edijs.drezovs@gostack.eu



Next event Wednesday 16th 2026

Deliver Data Seamlessly from Amazon MSK Topics and KDS to Apache Iceberg on S3 Tables

<https://aws-experience.com/emea/de-central-growth/e/d5a5e/whats-new-with-amazon-streaming-services>

Migrating from BigQuery to AWS: Yggdrasil Gaming's Modern Lakehouse Journey

GOSTack partnered with Yggdrasil Gaming to migrate their data analytics platform from Google BigQuery to a modern, open lakehouse architecture on AWS. The new platform,...

[Read case study](#)

60%

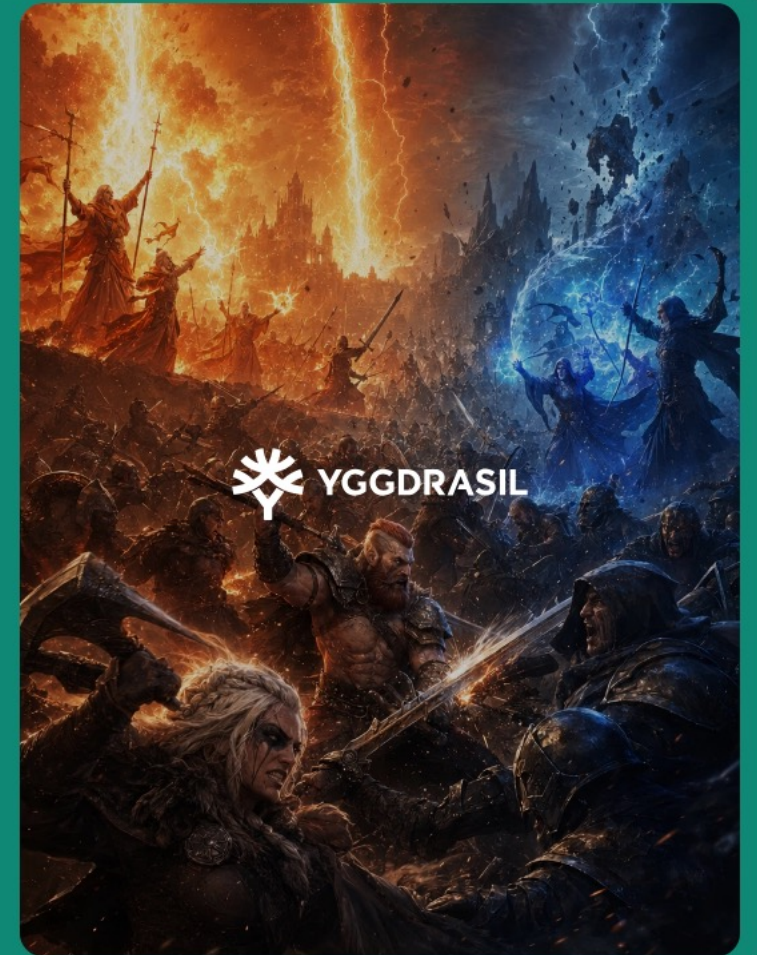
Data Processing Costs Reduced

4x

Processing Time Improved

100%

Sovereign



What do we need in a modern Data & AI platform?



Open data foundation for multi-engine compatibility

A single place to find all their structured and unstructured data, models, dashboards, and AI assets that's interoperable with compute engines and 3P. A single version of truth without data duplication.



Consistent governance and semantic layer for AI readiness

A governed data foundation becomes the context enabling AI agents to reason over trusted, traceable, and access-controlled data at scale and comply with regulations.



Flexibility for developers and agents

Equip developers and agents with a full suite of tools based on open standards, purpose-built to handle the full range of modern data and AI workloads.



AWS Open Lakehouse Architecture

S3 – A Solid Foundation for Data Lakes and lakehouse architectures



S3 is the solid and durable foundation on which data lakes and lakehouse architectures are built. A lakehouse architecture also requires:

Catalog & Governance

Discoverable, trusted, access-controlled data

Streaming Services

Real-time data ingestion and event-driven pipelines at any scale

Query & Transform

SQL, Spark, and serverless engines at any scale

AI & ML Integration

From raw bytes to intelligent predictions

AWS Lakehouse is an intentional architecture

Scale from Gigabytes to Exabytes, from SQL to Agentic AI



S3 and S3 Tables

- Durable storage with ACID semantics via Iceberg
- 10x higher TPS & 3x faster query performance
- Up to 10 GB/s streaming ingestion via MSK Express
- Variant data type (Iceberg V3) to store semi-structured JSON/variant data natively in Iceberg tables



Glue + Lake Formation

- Unified catalog and fine-grained governance across all engines
- Glue business context and semantic search API
- Glue 6.0 offers 30% lower price and full Apache Iceberg v3 support



Athena + EMR + Redshift

- Right-sized compute for every query pattern and team
- EMR delivers 4.5x faster performance than OSS Spark
- Redshift RG instances deliver up to 4.2x better price-performance than other data warehouses and eliminate Redshift Spectrum scan charges



Quick + SageMaker + Kiro + Bedrock

- Conversational BI, ML and agentic AI grounded in your governed lake data
- OpenAI models + Codex + Claude Desktop now available on Amazon Bedrock
- Generative AI Inference Recommendation to right-size model hosting directly in SageMaker AI

Context Intelligence for Data & AI Agents

AWS Context

Coming Soon

- Maps data relationships into a knowledge graph for AI agents
- Governed access to business rules and domain knowledge at runtime
- Identity-aware: inherits IAM + Lake Formation permissions
- Metadata published in Apache Iceberg format to S3
- Self-improving: learns from agent usage patterns

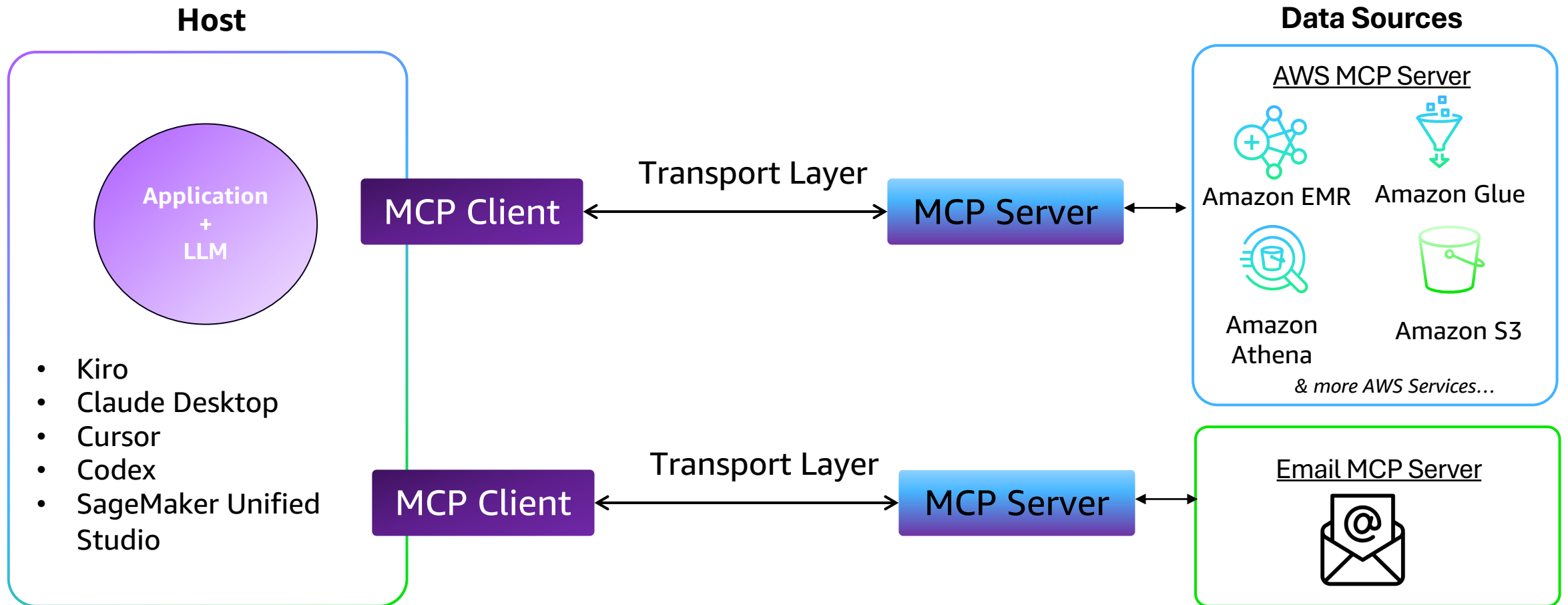
Glue Data Catalog: Business Context

Preview

- Enrich tables with glossary terms, business descriptions, custom metadata
- New Semantic Search API: find data by business meaning
- Skill assets: attach runbooks, query patterns, and usage rules
- Agents ground reasoning in trusted definitions vs. inferred context
- Compatible with any MCP agent via aws-data-analytics plugin

Why it matters: These services give AI agents governed access to organizational knowledge, eliminating per-agent curation while maintaining security and auditability.

Conversational Analytics Architectures



Thank you!



<https://pulse.amazon/survey/9JZR9KPA>