aws summit

INDIA | MAY 25, 2023

ANA007

How Gupshup built their multi-tenant messaging analytics platform on Amazon Redshift

Soujanya Konka (she/her)
Sr. Solutions Architect-ISV
AWS India

Pankaj Bisen (he/him)
Director - Engineering
Gupshup



Agenda

- Amazon Redshift powering Analytics for Gupshup
- What's new on Amazon Redshift
- Gupshup's migration journey
- Architecture supporting multi-tenet conversational AI platform



Conversational platforms



"Through Gupshup, companies have made communication an integral part of successful customer interactions."



Analyze all your data

WITH AWS SERVICES INTEGRATION BI and analytics apps Connect apps to analyze and visualize your data SQL Federated query **Operational** Amazon Materialized views databases Redshift ML Query live data, maintain materialized views Data sharing **Amazon Redshift** ML & analytics Data lake export Spectrum query S3 services Analyze open standards-**Data marketplaces** based data formats for third-party data Amazon S3 data lake Keep up to exabytes of data in S3



Listening to our customers for the last 10 years

INVENTING AND REINVENTING CLOUD DATA WAREHOUSING WITH AMAZON REDSHIFT

2012

Terabytes to petabytes to exabytes

2022

Customer Needs



a cost-effective way to run fast analytic queries on large volumes of data

- Data lake analytics and query of open data formats
- Scale storage independently of compute
- Elastic scaling
- Query operational data
- Query without data movement, data sharing and collaboration

- Data science needs beyond highly skilled data scientists
- Simplified query interface
- Minimal user involvement in performance acceleration, concurrent scaling
- Third-party data analytics

- More data users, needing to abstract complexity
- Real-time analytics in the warehouse
- Granular governance and security controls
- No code data pipelines, ingestion
- Apache Spark

Amazon Redshift



Amazon Redshift launched as the first MPP cloud data warehouse at a price performance unattainable in onpremises MPP DWs



Price-performance at any scale

Concurrency Scaling
Materialized Views
Workload Management
Short Query Acceleration
Vectorized Scans



Analyze all your data

Data Sharing
ADX Integration
Federated Query
Data Lake Query / Spectrum
Streaming Ingestion
Redshift ML
SUPER Data Type
Geospatial Analysis



Easy, secure, and reliable

Serverless Query Editor V2 RBAC Column-level Security Row-level Security



Amazon Redshift innovates to meet your needs



Easy analytics for everyone



Serverless



Query editor v2



Automated DW management



views

Automatic materialized



Data API



Amazon Redshift Advisor



AWS CloudFormation templates



Grafana Plugin



Analyze all your data



Data sharing



AWS Data Exchange integration



Amazon Redshift ML



Federated query



Geospatial enhancements



SUPER data type with JSON



Best price performance at any scale



RA3 nodes & managed storage



AQUA



Concurrency scaling for writes



SQL enhancements & migration support



Security, governance & compliance



Workload management enhancements



What's new on Amazon Redshift



- Apache Spark Integration
- Streaming Ingestion Support
- Dynamic Data Masking
- Auto-Copy From Amazon S3
- Multi-AZ deployment (preview)

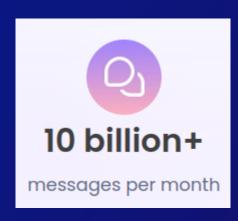
Background & Context

• 10 billion+ messages per month

More than 10 life cycle events per message

Constantly evolving eco-system

Need for near-real-time incremental analytics



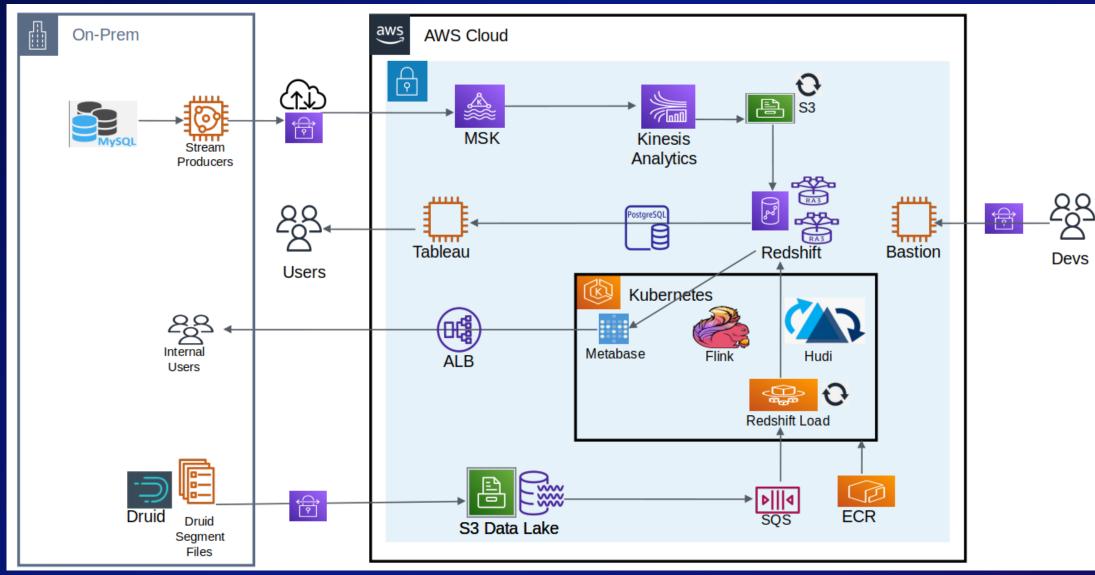


Live Migration – Stranger pattern

- Large legacy data warehouse on Apache Druid
- Live migration to cloud & Amazon Redshift
- Parallel run and gradual legacy sunset
- Live prototyping and roll-outs of new analytics
- Incremental analytics



Architecture - Conversational Analytics



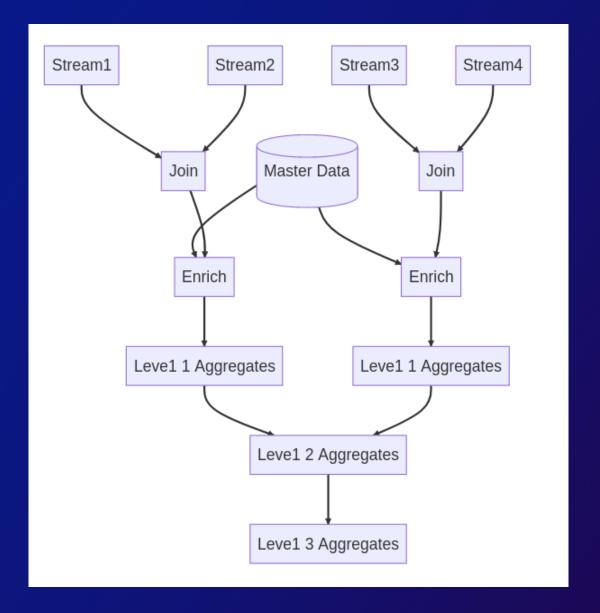
Incremental and Operational Analytics

Operational

- On-Demand Scaling
- Snapshot and Restore
- Blue-Green Deployments
- Workload Management
- On-Demand Re-processing of historical data
- Concurrency Scaling

Incremental

- Streaming data
- Enrichment using master data
- Hierarchy of Aggregates
- Live prototyping
- Incremental Processing using Materialized Views
- ~30 minutes of freshness





Your time is now

Build in-demand cloud skills your way



Thank you!

Soujanya Konka Sr. Solutions Architect-ISV AWS India Pankaj Bisen Director – Engineering Gupshup



Please complete the session survey

