
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

ANT324

Modernize your data warehouse

Neeraja Rentachintala
Sr. Manager
Product Management
Amazon Redshift

Shruti Worlikar
Sr. Manager
Analytics Specialist SA
AWS

Shyam Mohapatra
Director
Janssen NA Commercial Data & Insights
Johnson & Johnson

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

- Why data warehouse modernization?
- Use cases of modern data warehouse
- Amazon Redshift for modern data warehousing
- Demo
- Data warehouse modernization journey of J&J

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Traditional warehousing is not built for the volume, variety, and velocity of data coming into organizations



"Analytics requirements are coming in from all corners. We don't have enough specialized skills."



"How do I derive value from growing, diverse data, living in data silos?"



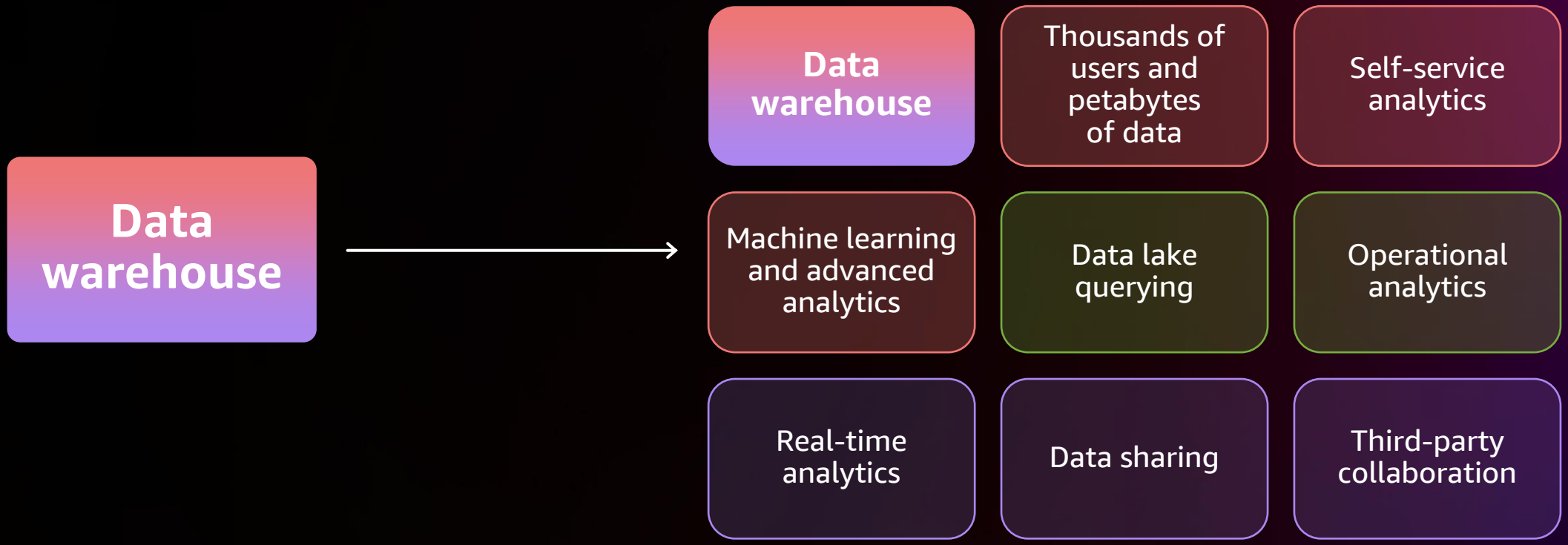
"Performance has to be consistently fast. Costs are hitting the roof."

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

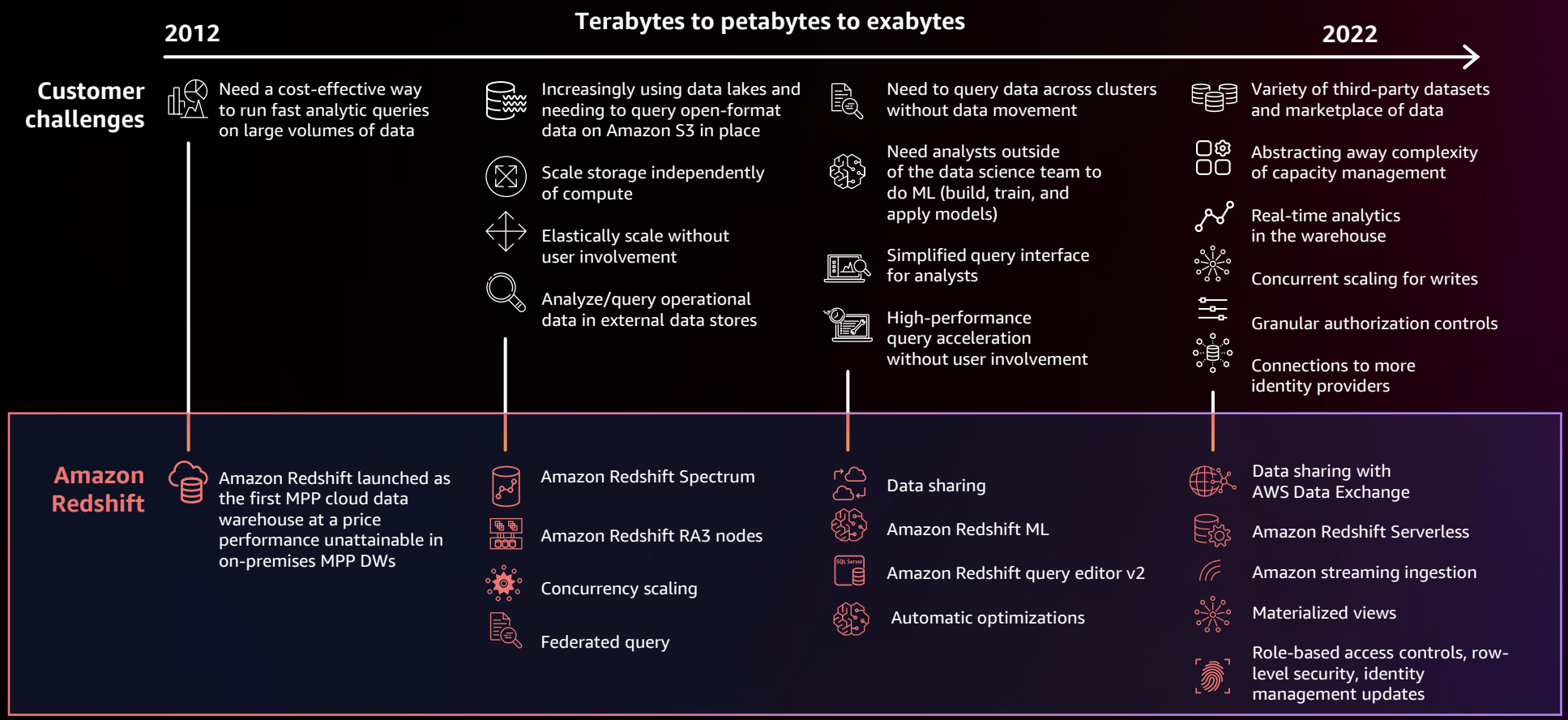
Modern data warehouse is not a silo

ENABLES A VARIETY OF NEW ANALYTICS USE CASES WITHOUT WORRYING ABOUT INFRASTRUCTURE



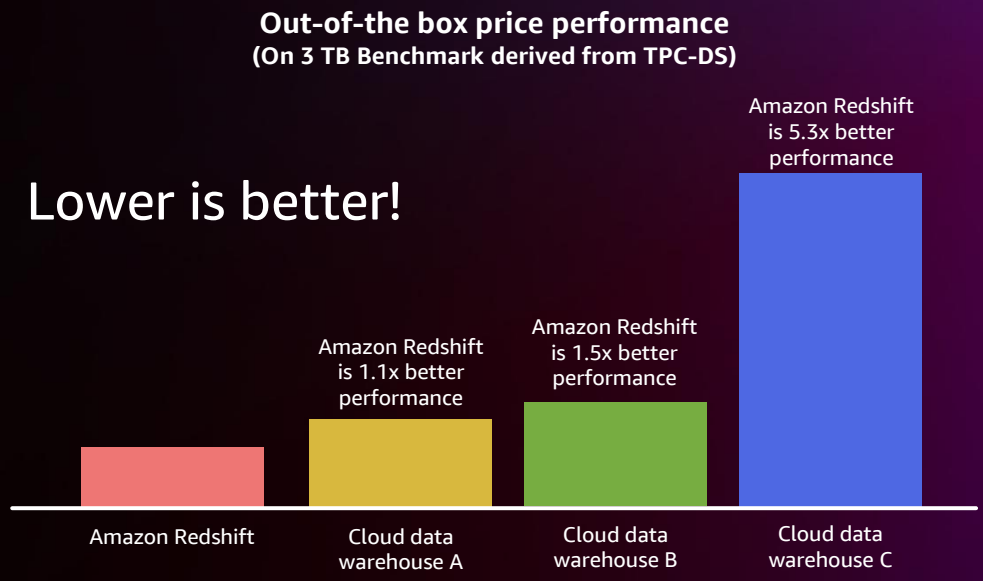
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

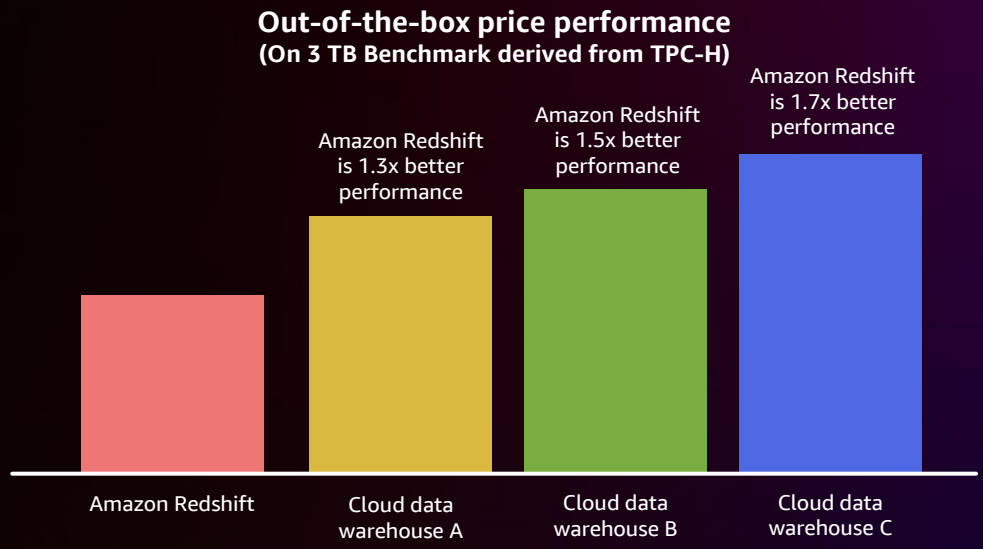
Amazon Redshift is reinventing data warehousing with continuous innovation by listening to customers



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Up to 5X better price performance





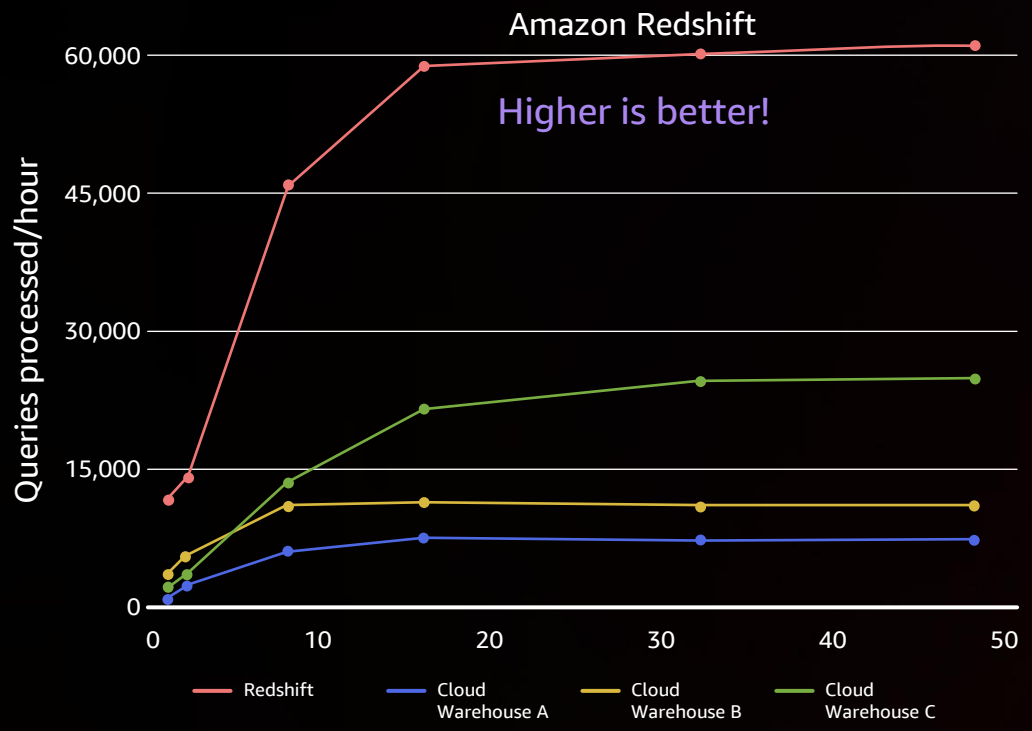
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Price performance at scale with Amazon Redshift

7x better price performance!

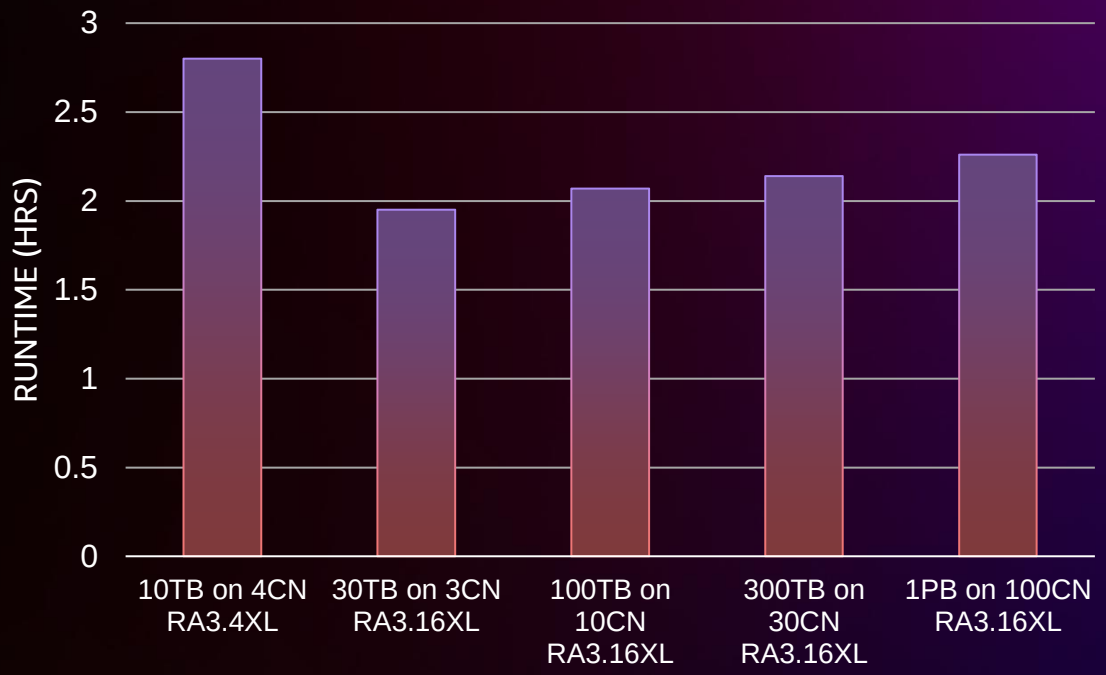
Query throughput for short queries

(On 10 GB benchmark derived from TPC-DS)



Linear scalability

(On benchmark derived from TPC-DS)

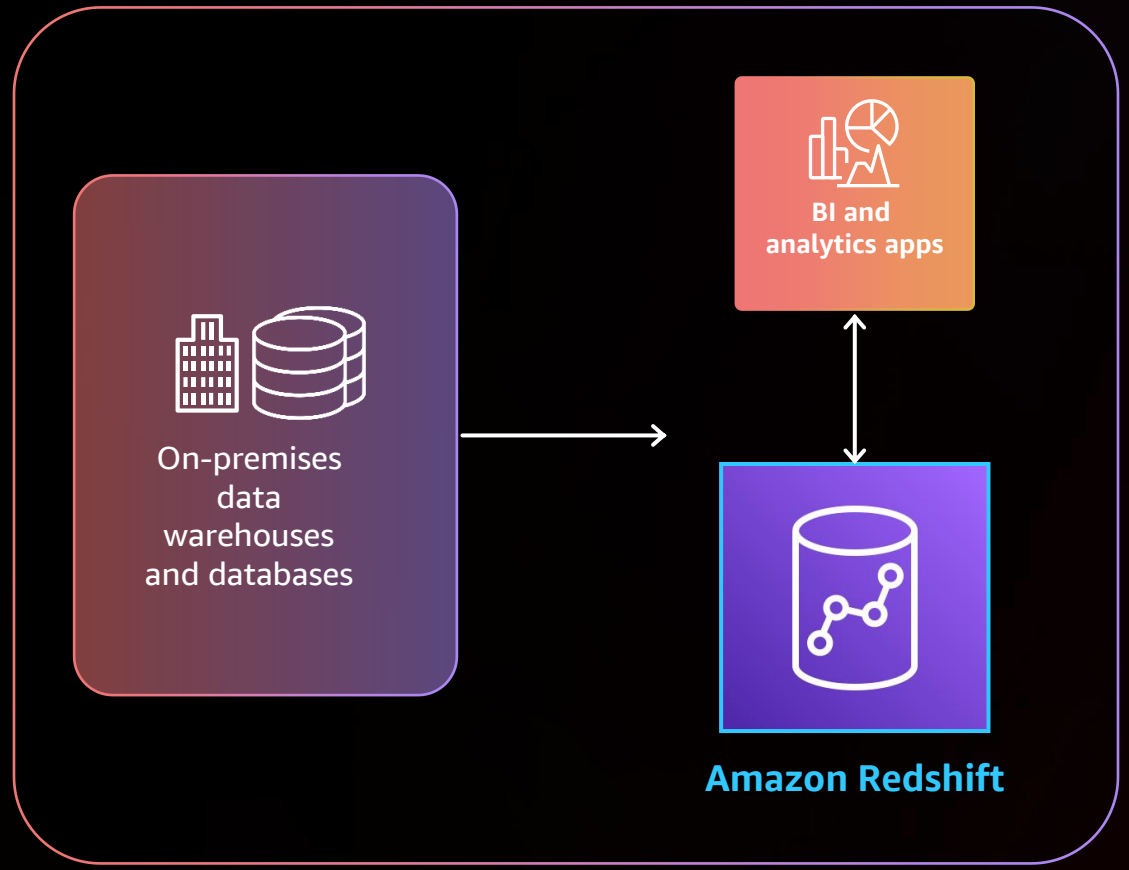


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Modern data warehousing with Amazon Redshift

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data warehouse migration with Amazon Redshift



Rich ANSI SQL capabilities and tool connectivity

Simpler application development with data API

Native semi-structured and spatial data support

Easy extensibility with SQL, Python, and Lambda based UDFs

New! Merge, OLAP functions (Rollup, Cube, Grouping sets)

Industry-leading security and access management features, out-of-the box at no additional cost

AWS security standards and compliance certifications

Granular authorization at row and column level

New!

Sensitive data protection with dynamic data masking

Amazon Virtual Private Cloud (VPC) support

Identity management and federation

Auditing and compliance

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

We simplify and help through all phases of your migration journey

AWS tools



AWS Schema Conversion Tool (AWS SCT)



AWS Database Migration Service (AWS DMS)



Amazon S3 Transfer Acceleration



AWS DataSync



AWS Snow Family



Amazon Kinesis Data Firehose

And many more

Comprehensive AWS Partner ecosystem

2nd Watch, accenture, AGILISIUM BIG ON CLOUD. BIG ON DATA., bizanalytica

brillio, Cloudwick, Cloudreach, Deloitte.

Infosys, LTI, MACTORES, MERKLE

mission CLOUD SERVICES, ONICA by rackspace technology., quantiphi, slalom

softserve, TEKsystems Own change, virtusa, ZS

This is not a complete list. To view all AWS Partners for this category, visit AWS Partner Solutions Finder.
This list of partners is current as of October 28, 2022.

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NTT docomo

RetailMeNot FACTSET ROVIO
intuit amazon Jack in the box
Magellan HEALTH VOO EQUINOX
nielsen LexisNexis

Migrating to Amazon Redshift allowed us to scale to over 10 PB of uncompressed data with a 10x performance improvement over our prior on-premises system. Today, it is the center of our analytics environment. Since we started using Amazon Redshift, both our data and number of users have increased dramatically. We are impressed with the flexibility and ease of use, even as we scale users and data. We are excited about the new Amazon Redshift RA3 instances with managed storage, enabling us to scale compute and storage separately. We appreciate AWS's continual innovation on behalf of its customers.

Takaaki Sato
General Manager of Service
Innovation Department, NTT Docomo

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Self-service analytics



NOW [GA]

Amazon Redshift Serverless

Get insights from data in seconds

Experience consistently high performance

Get started with no modifications

Pay for what you use

YOU

focus on insights

Automatic compute provisioning

Automatic scaling

Automated patching

Automatic workload management

Advanced monitoring

Backup and recovery

Performance at scale

Security and industry compliance

aws

takes care of the rest

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



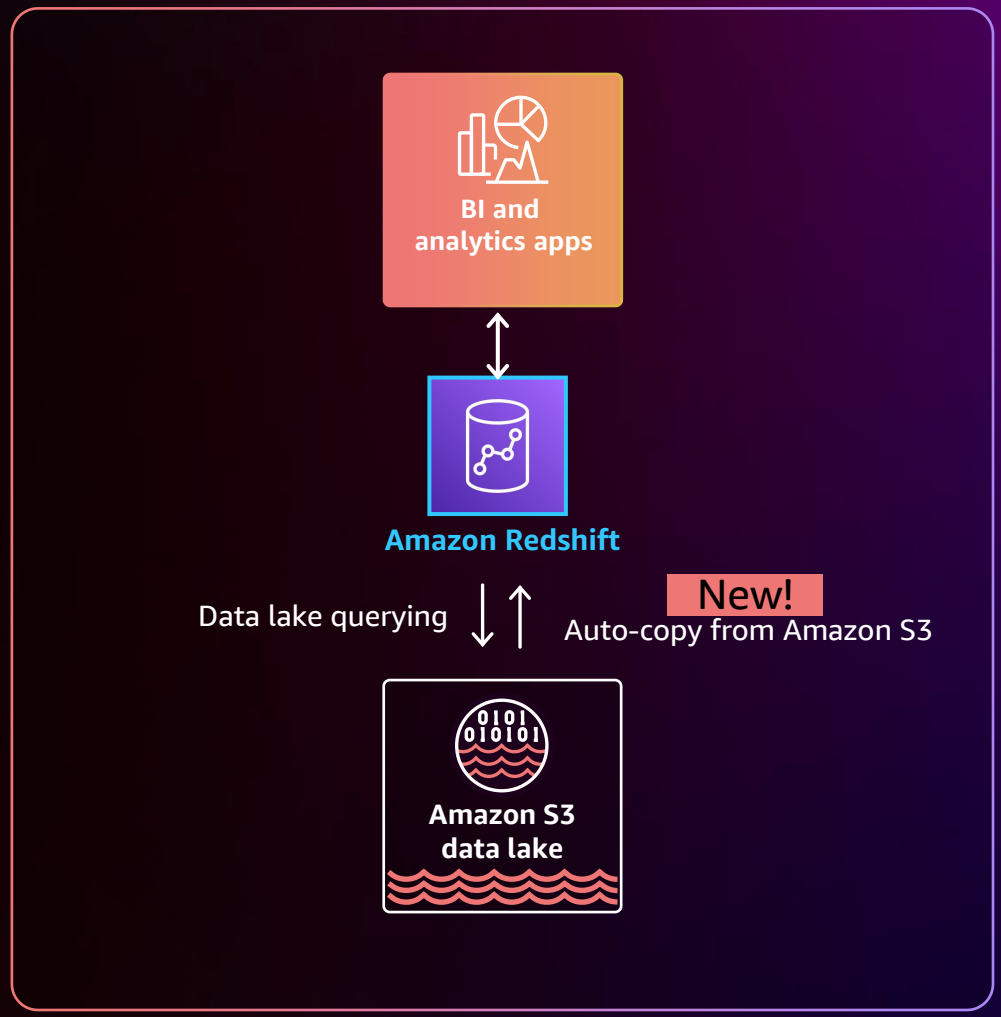
Amazon Redshift Serverless helps us complete our data management without having to manage clusters and optimizes our cost by provisioning just the right amount of capacity to meet demand. Amazon Redshift Serverless is reducing the operational burden, lowering costs, and enabling scale for the Roche go-to-market domain. This simplification is a game changer, helping us rapidly onboard and support a variety of analytics-heavy use cases without friction.

Dr. Yannick Misteli
Lead, Cloud Platform and ML Engine, Roche

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data lake querying with Amazon Redshift

- Query structured and semi-structured data in Amazon S3 data lake
- Massive parallelism and high-performance querying on exabytes of datasets
- Join data across data warehouse and data lake
- Comprehensive open file format support
- Automatically ingest data as needed from S3 without manual data pipelines
- Deep integration with AWS Glue Data Catalog and AWS Lake Formation for centralized granular security and governance



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



We were able to easily support the jump from 30 billion records to 70 billion records a day because of the flexibility and scalability of Amazon S3 and Amazon Redshift.

Robert Hunt
Vice President of Software Engineering, Nasdaq

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

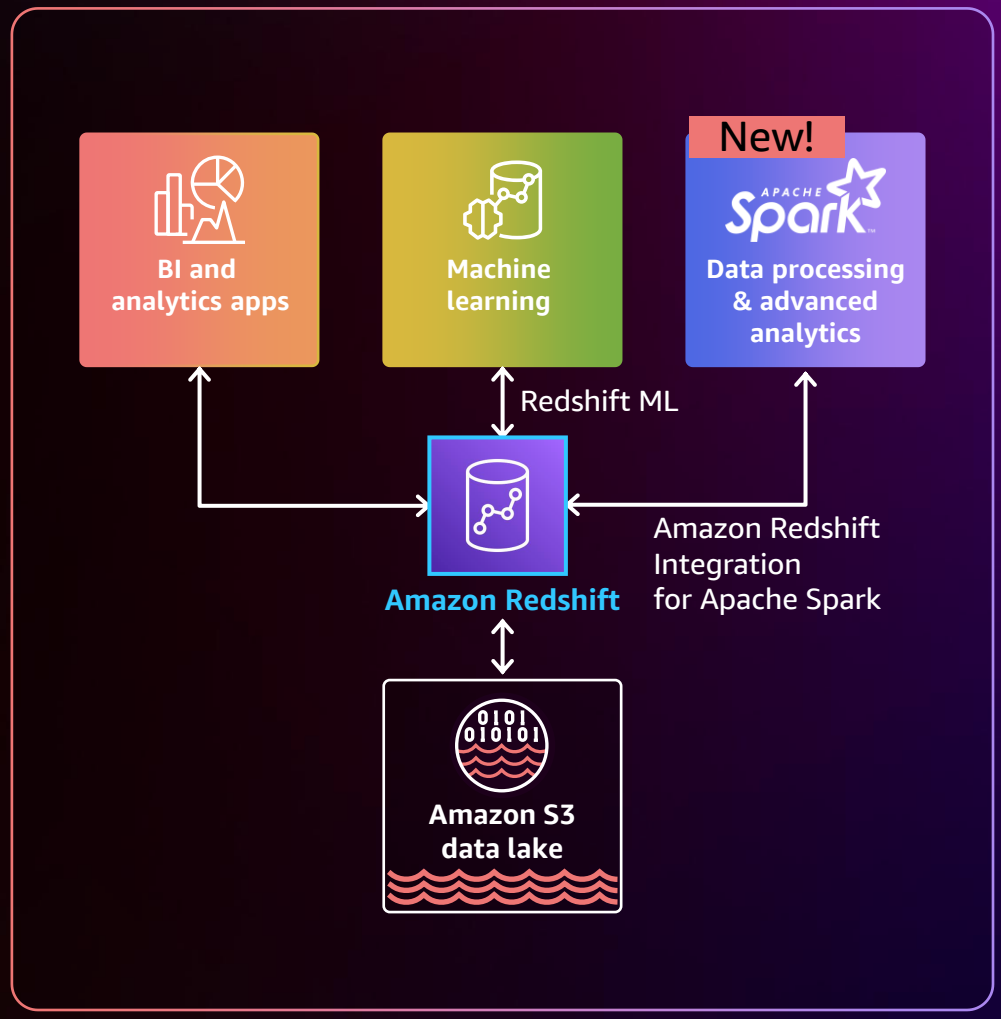
Machine learning and advanced analytics with Amazon Redshift

Amazon Redshift machine learning

- Create, train, and deploy models using SQL
- Automatic pre-processing, creation, training, deployment, and inferencing of models
- Amazon SageMaker models for in-database or remote inference
- Supervised and unsupervised trainings

New! Amazon Redshift integration for Apache Spark

- Build Apache Spark applications with easy access to data in your data warehouse
- Get started quickly with a fully packaged and certified integration
- Advanced pushdown optimizations to speed up query processing
- Integrated with Amazon EMR and AWS Glue



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



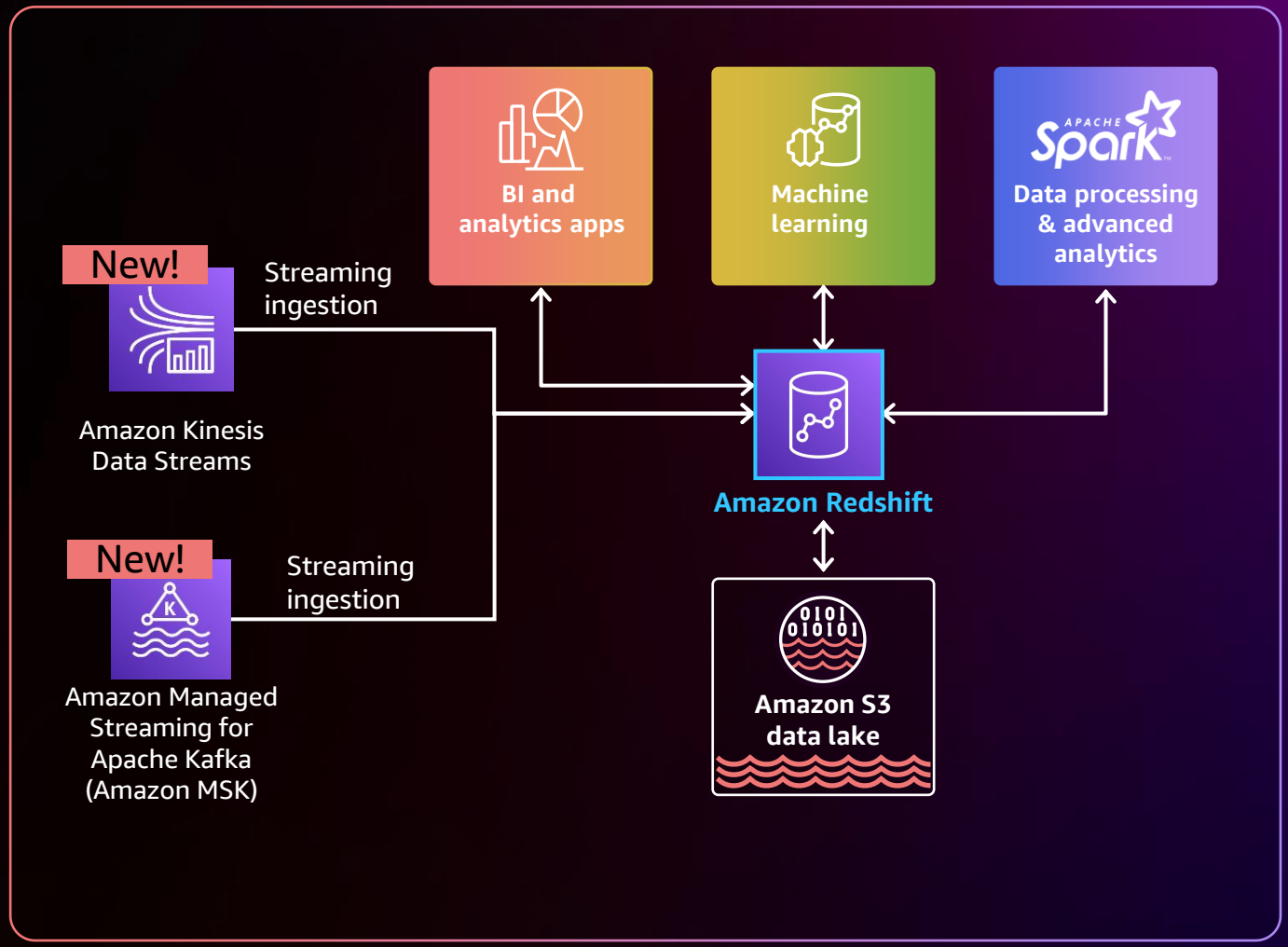
Jobcase has several models in production using Amazon Redshift machine learning. Each model performs billions of predictions in minutes directly on our Amazon Redshift data warehouse with no data pipelines required. With Amazon Redshift ML, we have evolved to model architectures that generate a 5–10% improvement in revenue and member engagement rates across several different types with no increase in inference costs.

Mike Griffin
EVP Optimization and Analytics, Jobcase

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GA – Real-time analytics with Amazon Redshift on streaming data

- Ingest streaming data into your data warehouse for real-time analytics in seconds
- Handle large volumes of data with low latency and high throughput
- Use familiar SQL – no need for complex pipelines and infrastructure
- Generally available for Amazon Kinesis Data Streams and Amazon MSK



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

JOYME

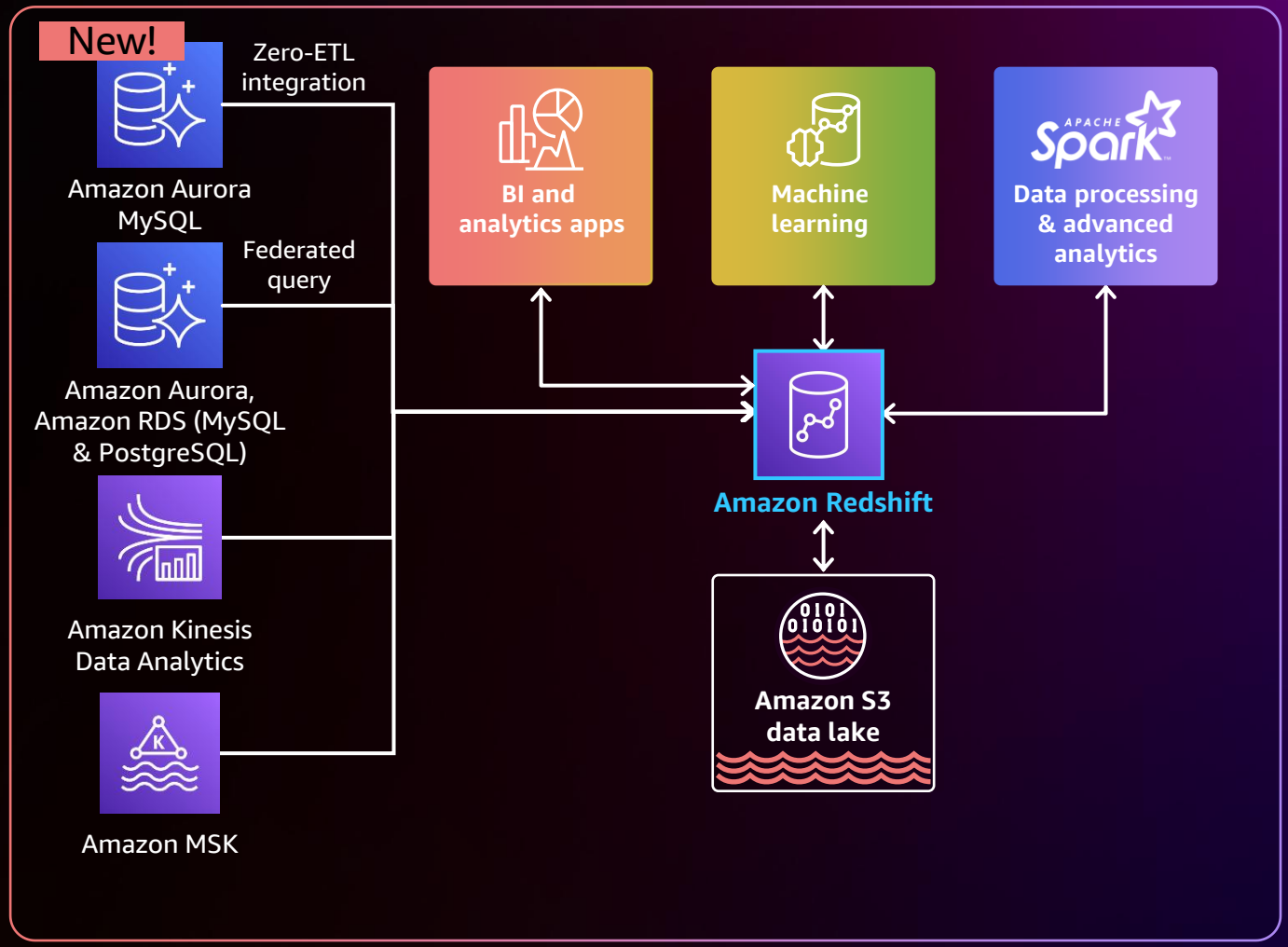
JOYME uses Amazon Redshift's streaming ingestion and other Amazon services for risk control over users' financial activity such as recharge, refund, and rewards. With Amazon Redshift, we are able to view risk control reports and data in near real time – instead of on an hourly basis. Amazon Redshift significantly improved our business efficiency.

PengBo Yang
CTO, JOYME

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Analytics with Amazon Redshift on transactional data

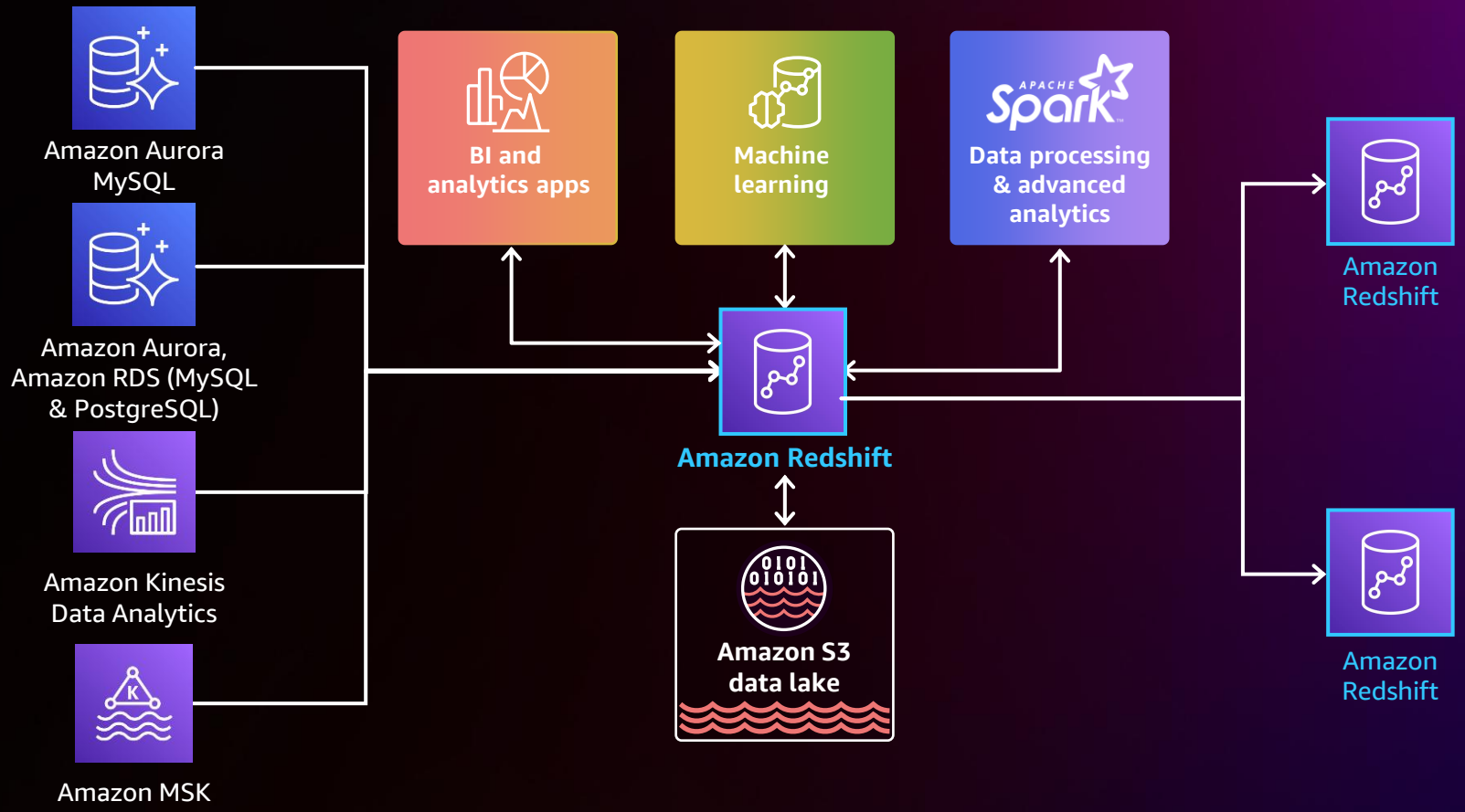
- Rich analytics and machine learning on transactional data without the complexity of data pipelines
- Leverage federated querying directly on transactional data for exploratory analytics
- **New!** Use zero-ETL integration to seamlessly replicate near real-time transactional data into the data warehouse



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data sharing with Amazon Redshift

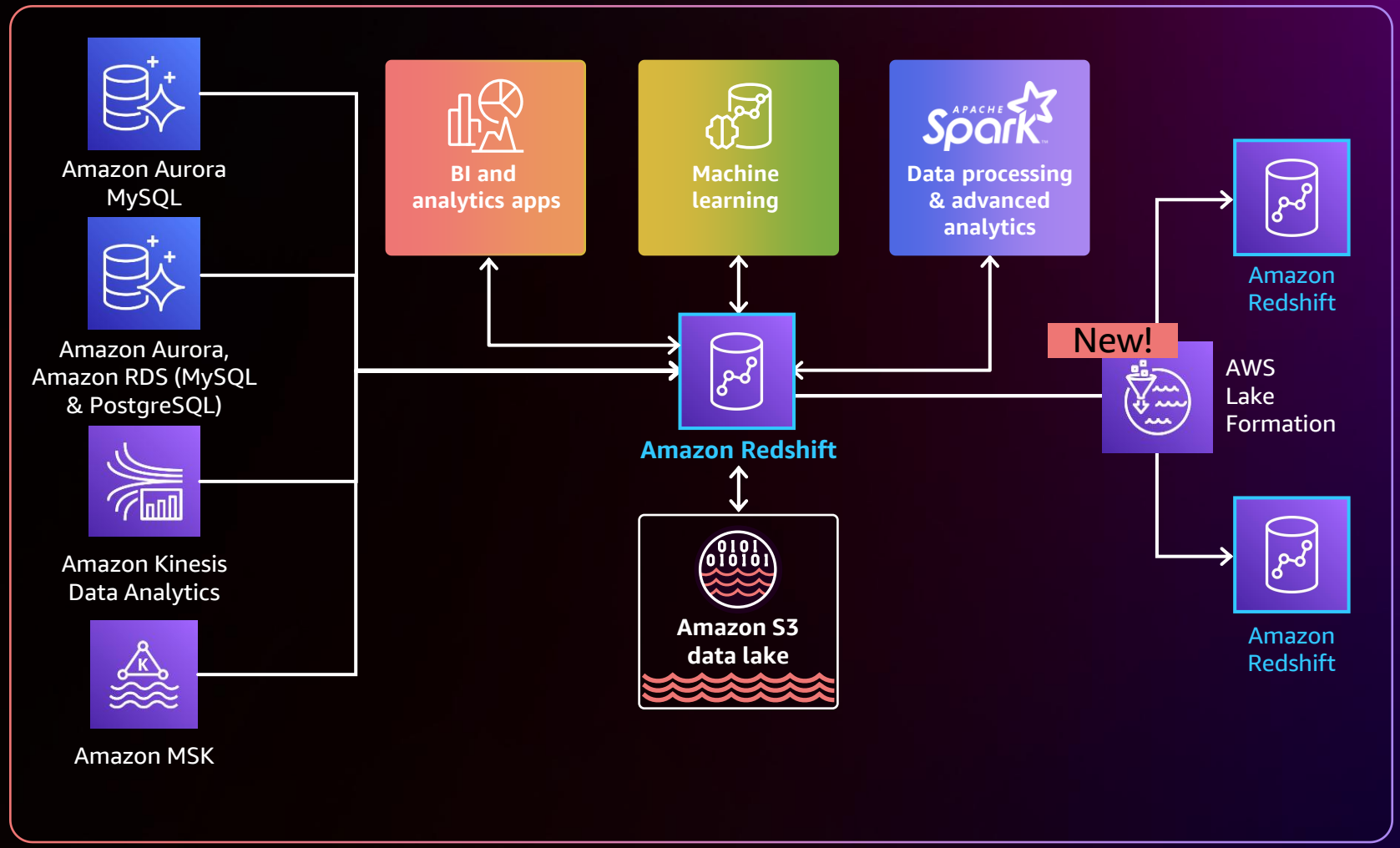
- Instant, secure, and live data sharing across Redshift data warehouses
- Within and across AWS accounts and across AWS Regions
- Live and transactionally consistent
- Flexible multi-cluster and data mesh architectures



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data sharing integration with AWS Lake Formation

- Centrally manage access to data shared across consumers
- Security and governance with granular row-level and column-level controls



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Fannie Mae

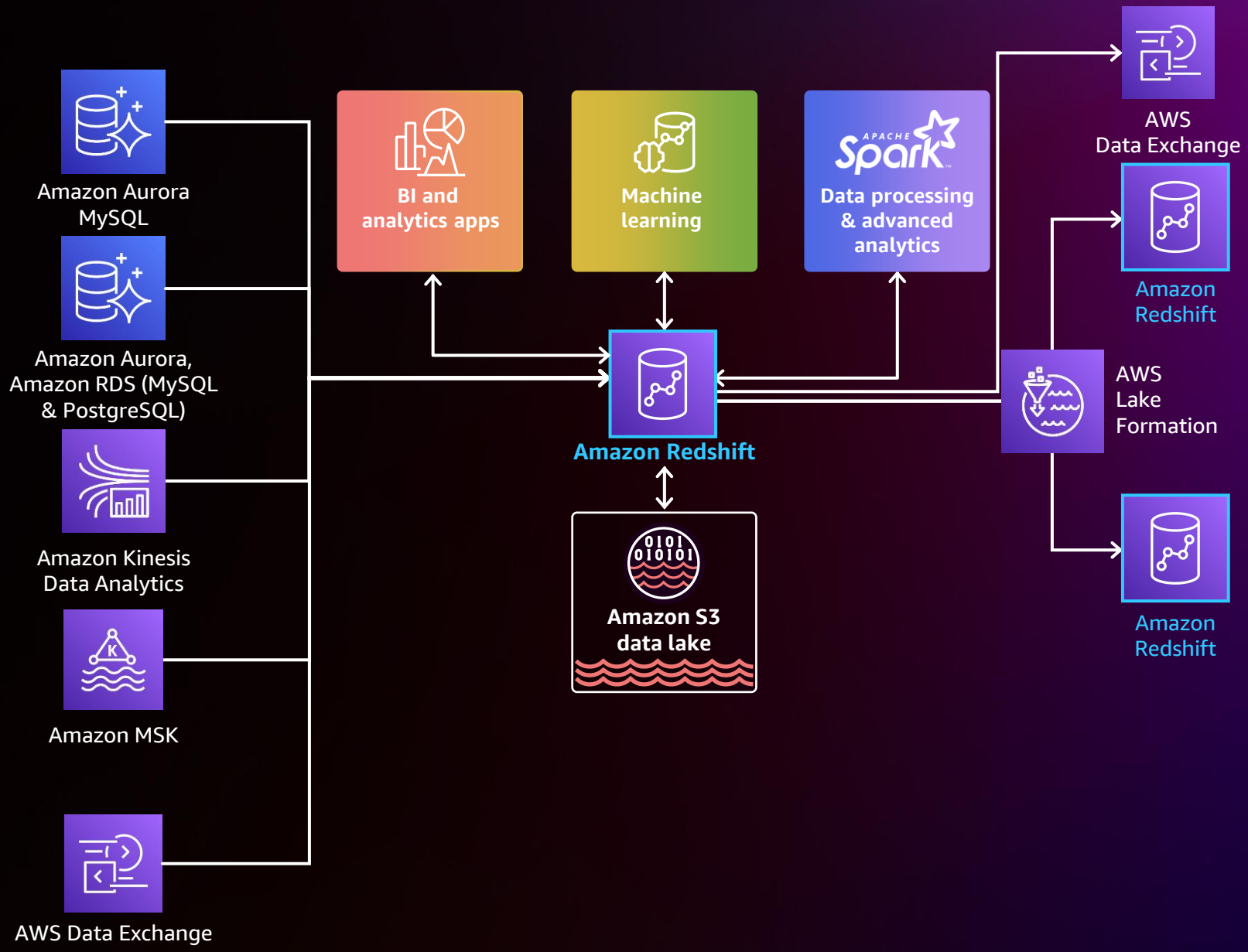
At Fannie Mae, we adopted a decentralized approach to data warehouse management with tens of Amazon Redshift clusters across many applications. While each team manages their own dataset, we often have use cases where an application needs to query the datasets from other applications and join with the data available locally. With data sharing, we can enable seamless sharing of data across application teams and give them common views of data without having to do ETL. We are also able to avoid the data copies between pre-prod, research, and production environments for each application. Data sharing made us more agile and gave us the flexibility to scale analytics in highly distributed environments like Fannie Mae.

Amy Tseng
Enterprise Databases Manager, Fannie Mae

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon Redshift and AWS Data Exchange integration

- Easily discover, subscribe, and consume third-party datasets
- License and monetize data by offering data warehouse data a as a service
- Live data access – no ETL and data copies necessary
- AWS Data Exchange handles subscriptions and payments



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



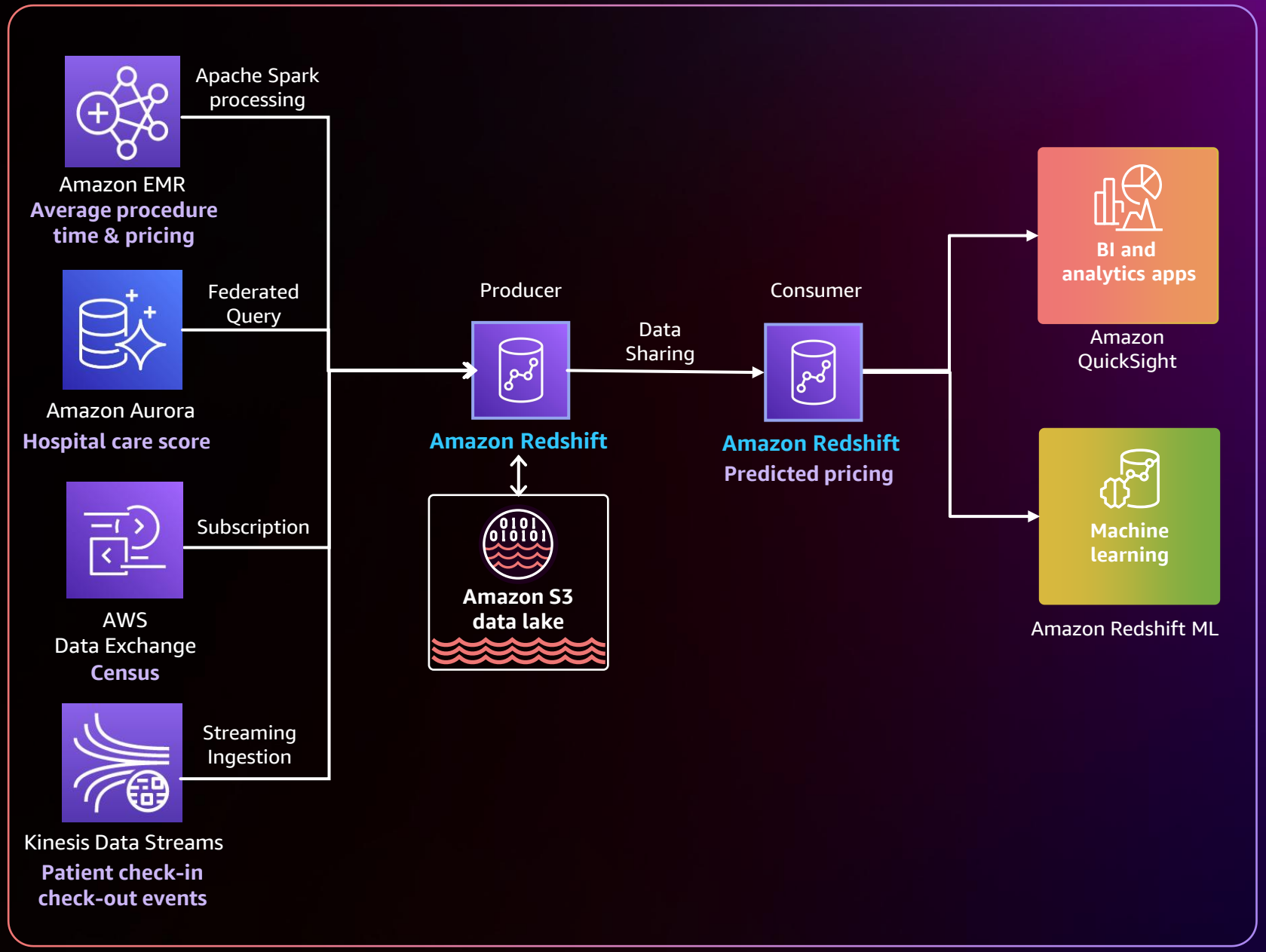
Facteus consistently strives to bring the highest-quality data with the lowest latency to market, and we are proud to work with AWS Data Exchange and the Amazon Redshift team to make the data available even faster, for our Amazon Redshift clients. Delivering our data via the Amazon Redshift share results in operational efficiency and cost savings for both Facteus and our partners and removing potential data ingest errors.

Ben Zweig
CEO, Facteus

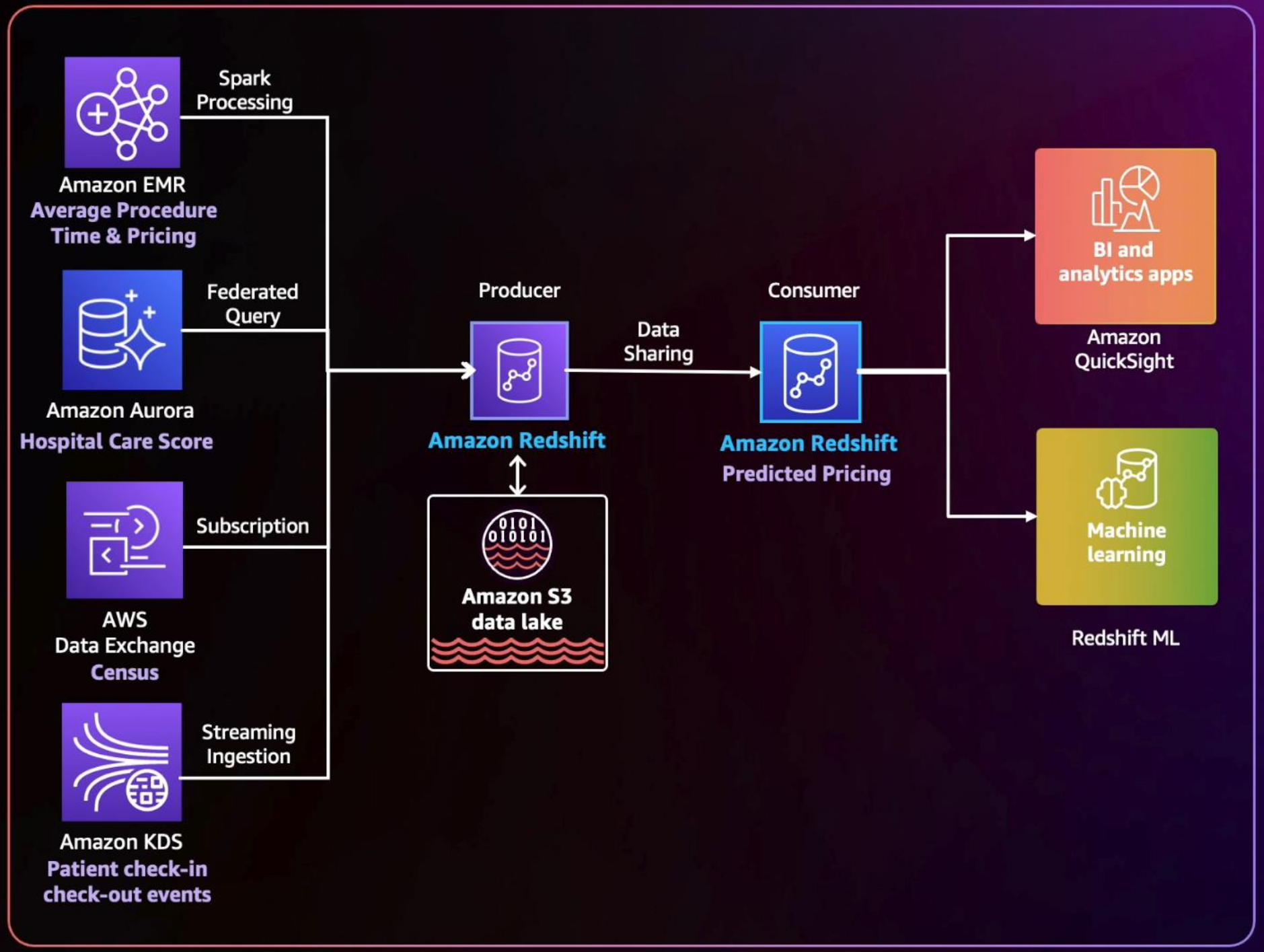
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Demo – Putting this all together

ACCESS, ANALYZE, AND SHARE REAL-TIME DATA



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.


Spark Processing
Amazon EMR
Average Procedure Time & Pricing
Federated Query
Amazon Aurora
Hospital Care Score
Subscription
AWS Data Exchange
Census
Streaming Ingestion
Amazon KDS
Patient check-in check-out events
Producer
Amazon Redshift
Amazon S3 data lake
Data Sharing
Consumer
Amazon Redshift
Predicted Pricing
BI and analytics apps
Amazon QuickSight
Machine learning
Redshift ML

Data warehouse modernization at North America Janssen (Pharmaceutical Company of J&J)

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

- Challenges – voice of customer
- Our journey
- Business benefits
- Data architecture
- Learning & next steps

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

North America Janssen (Pharmaceutical Company of Johnson & Johnson) commercial business technology – data & insights product group

We shape the **future of healthcare** through a diverse team by unlocking the power of data through an insight-driven organization for the purpose of improving efficiency, generating value by empowering data users to deliver a better customer experience.

Group

Our vision

Data users

- Sales & marketing
- Data scientists
- Contracting
- Medical teams
- Market research users
- Evidence generation teams

Key opportunities

1. Growth in business
2. Data growth
3. Data sharing needs
4. New channels
5. New data users & applications
6. Self-service
7. Faster time to delivery
8. Reducing cost

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Johnson & Johnson

Challenges

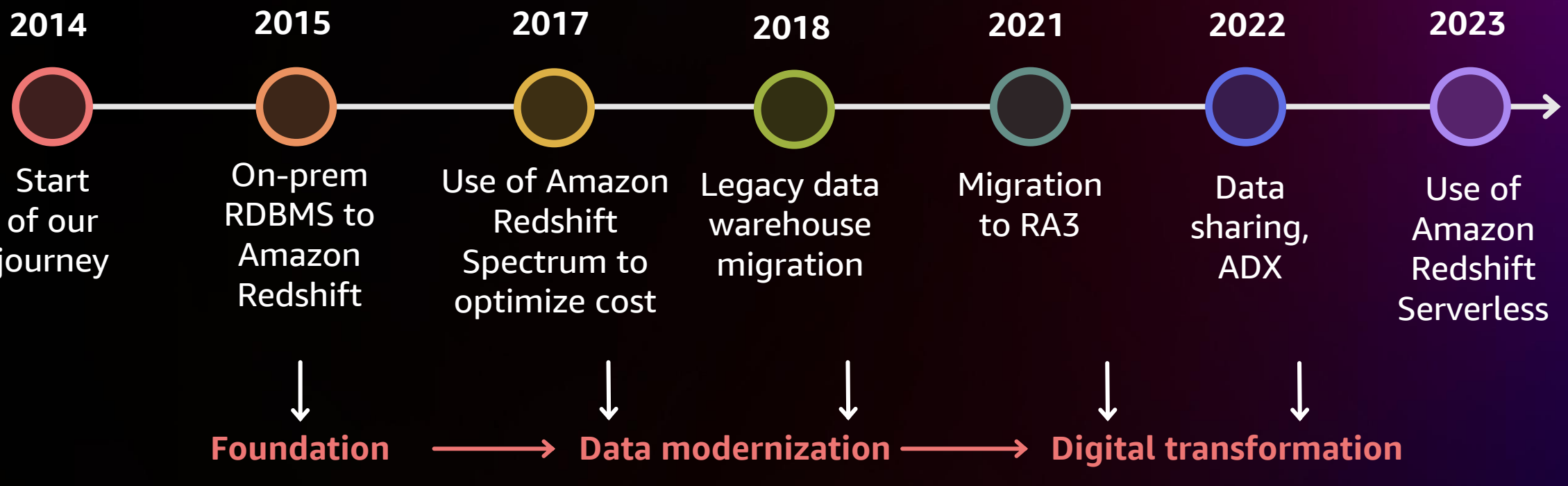
VOICE OF THE CUSTOMER



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Our journey

NA JANSSEN COMMERCIAL



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key business benefits



Compute & storage separation

Support data & business growth

Lower carbon footprints



Data sharing

Better user experience

Reduce data duplication

Data interoperability

Drive innovation



Cluster management

Better capabilities – query, usage, concurrency, elastic resize (increase scalability)

Reduced administration & operational overheads



Data growth

Ability to scale to petabytes (20x growth)

Enable new use cases – omnichannel, AI/ML

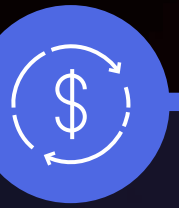


Application & user base

New applications & users

Enabled self-service

Democratizing data & access



Cost & effort

30% reduction on operation cost

50% faster time to delivery



Self-service

Data scientists & citizen data users

AI/ML capabilities, real-time data, data APIs



Scaling

Optimized production workloads

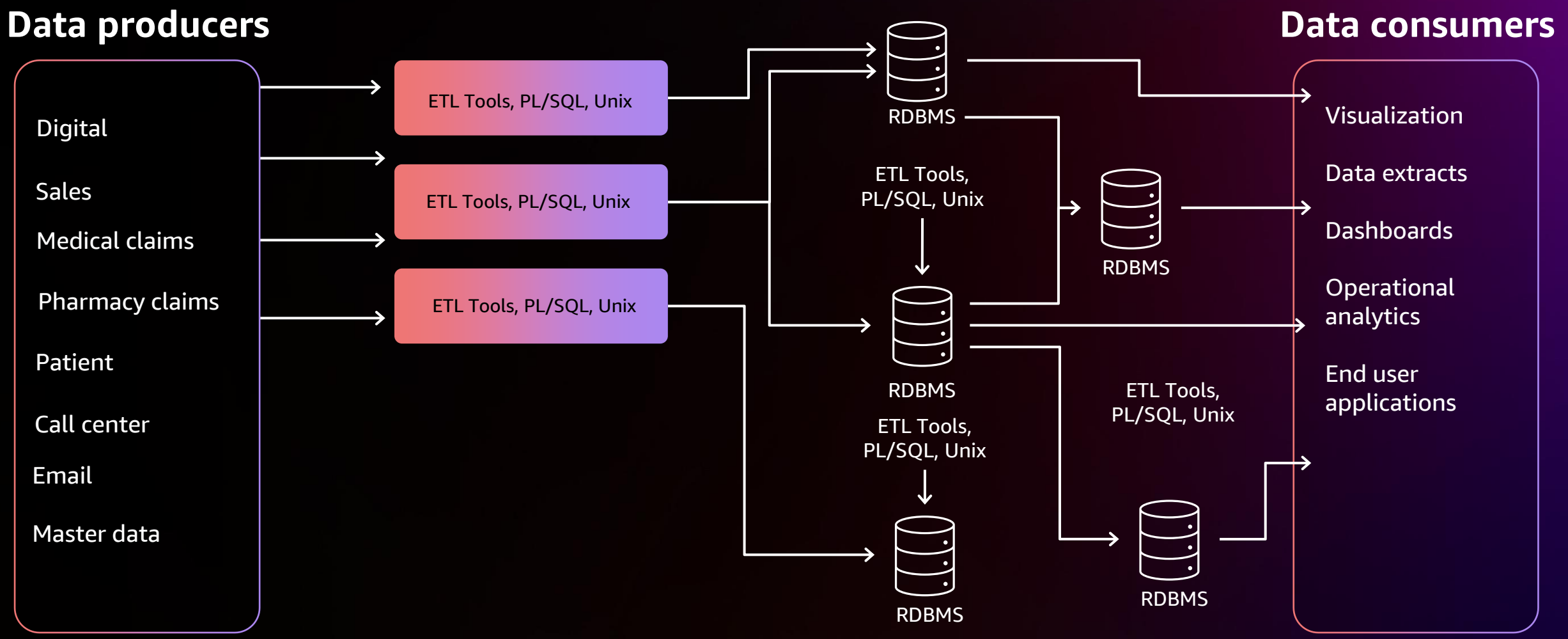
Support new, dynamic, and changing workloads

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

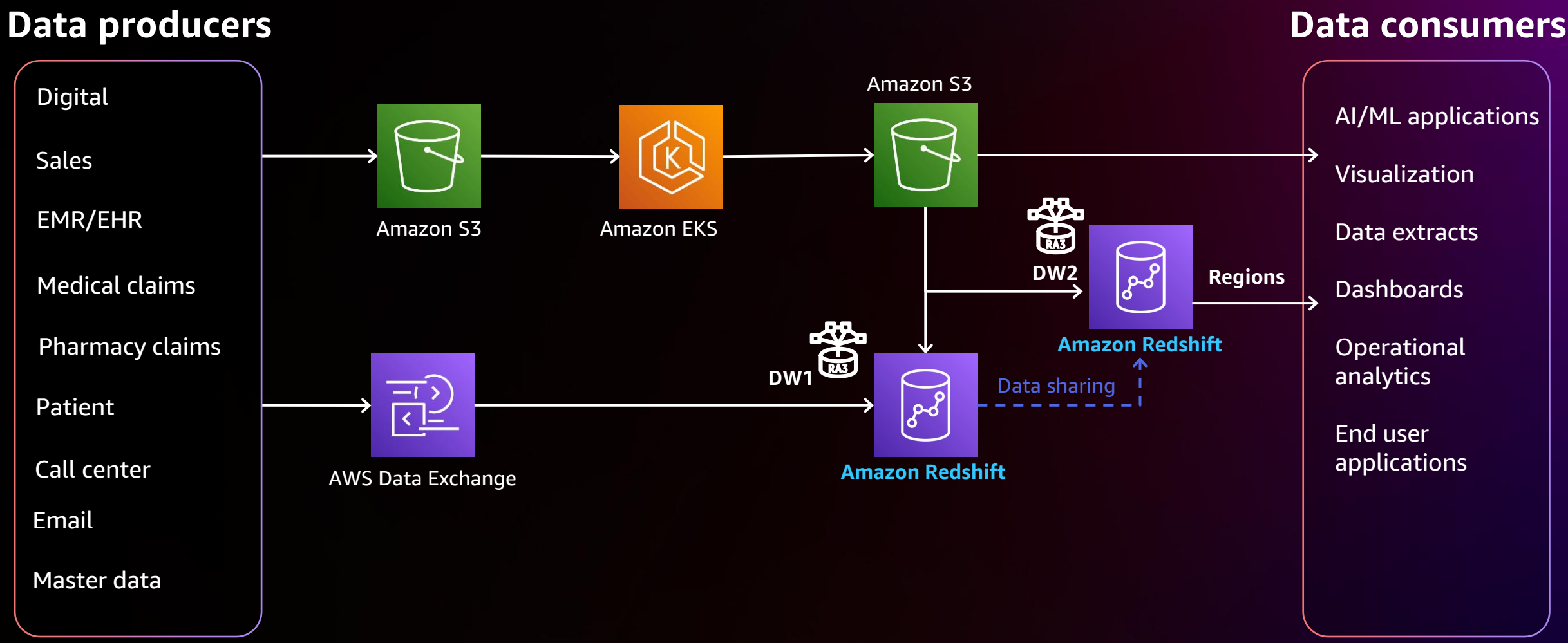
Johnson & Johnson

Data architecture (prior state)



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data architecture

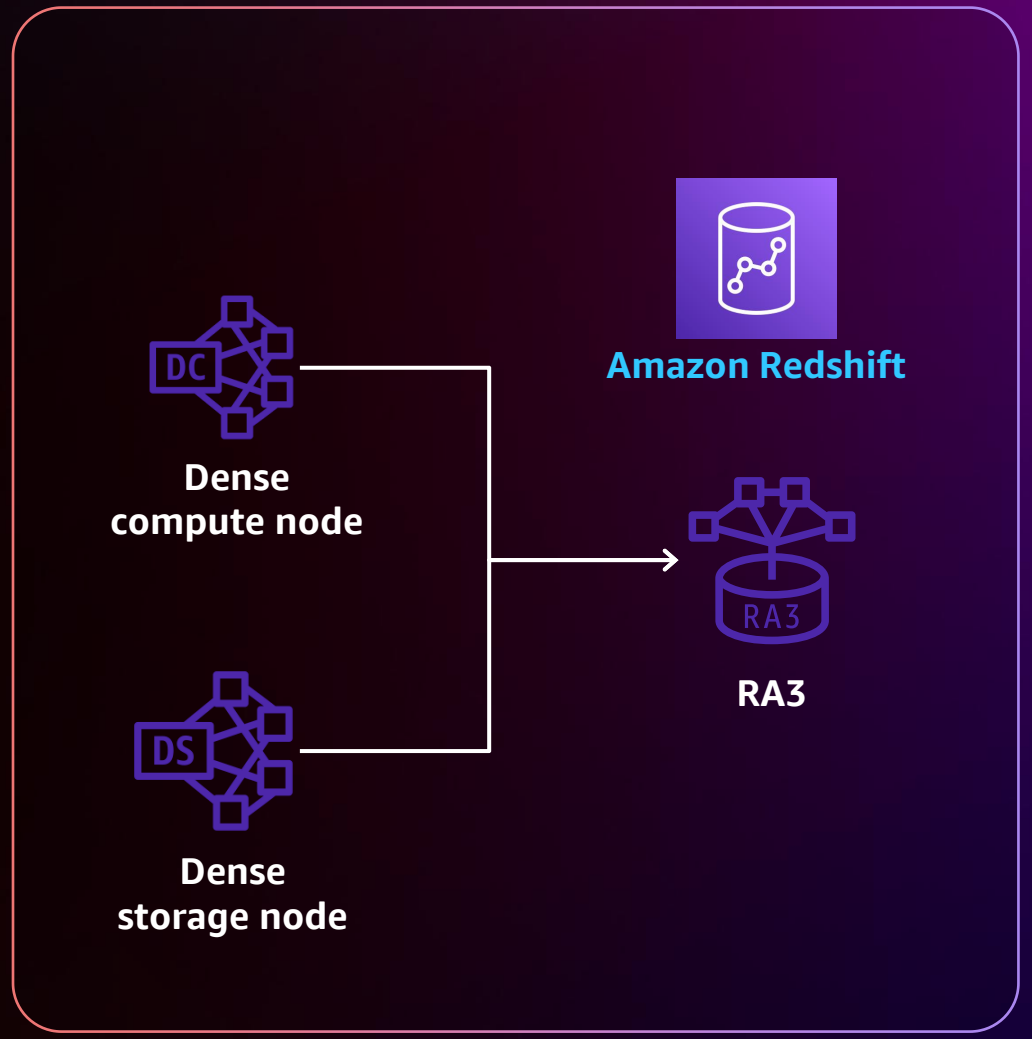


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Modernization effort to Amazon Redshift RA3

JANSSEN COMMERCIAL NA

- Clusters – DS2 & DC2 migrated to RA3
- Data 100+ TB
- Migration time for larger clusters **<2 hrs**
- **Zero** downtime
- $0 net cost increase
- RA3 backward **compatibility** enabled seamless user/app transition
- Minimum migration cost with **least impact** to business
- Migrations were smooth and no issues encountered
- AWS provided **standby support** during migration



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key recent learnings



Right sizing from **DS2/DC2 to RA3**



Start with **smaller clusters & non-prod**



Use of new features like **Advanced WLM, data sharing, etc.**



Encryption for bigger clusters was taking a very long time – required for cross-cluster data sharing; AWS released a feature update to speed up encryption



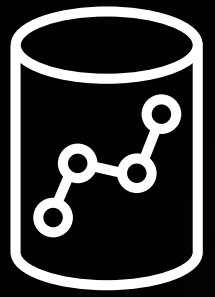
Initially elastic re-sizing was getting stuck with a bug; **AWS team quickly provided standby support to mitigate it**

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

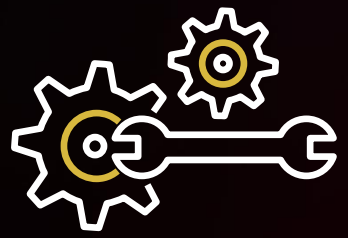
Johnson & Johnson

What's next

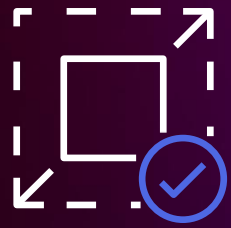


Amazon Redshift Serverless

To run & scale analytics use cases



Expansion of use cases for **ADX** to enable **self-service** & faster data access, insights



Scaling data sharing across organization

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Johnson & Johnson

Get started today

New to Amazon Redshift?

- ✓ IDC study on Business value of Amazon Redshift
- ✓ Explore demos, customer stories, and the latest features on aws.amazon.com/redshift
- ✓ Get trained
- ✓ Ask your account team for a 10-minute demo

Already using Amazon Redshift?

- ✓ Ask your account team for a free Amazon Redshift optimization session
- ✓ Learn more about what's new https://aws.amazon.com/redshift/whats-new/

Thinking of migration?

AWS teams and programs can help you with

Building your business case, creating migration plan

Deciding on budget, scope, etc.

Conduct a proof of concept or architect your data strategy

Get Help

aws.amazon.com/redshift

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Neeraja Rentachintala
neerajre@amazon.com

Shruti Worlikar
shrutiw@amazon.com

Shyam Mohapatra
SMohapa1@its.jnj.com



Please complete the session survey in the **mobile app**

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.