AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

ANT 302-R2

New use cases for Amazon Redshift

Harshida Patel
Senior Analytics Specialist SA
AWS

Madhu Muppagowni (he/him)
Senior Staff Software Engineer
GE Digital

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

Modern data architecture

Amazon Redshift features

GE Digital manufacturing re-platform use case

Q&A

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

"In 2020, 64.2 ZB of data was created or replicated. … The amount of digital data created over the next five years will be greater than twice the amount of data created since the advent of digital storage.“

Source: IDC Modern Data Platform 2021



“Analytics requirements are coming in from all corners. We don’t have enough specialized skills for data management, or machine learning.”



“How do I derive value from such large swaths of diverse data, living in these data silos in my organization?”



“How can we keep performance consistently fast? Costs are hitting the roof as performance degrades.”

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

A modern data strategy

Data at any scale

The best price-performance

Seamless data access

AI and ML to solve business challenges



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Modern data architecture

Scalable data lakes

Purpose-built data services

Seamless data movement

Unified governance

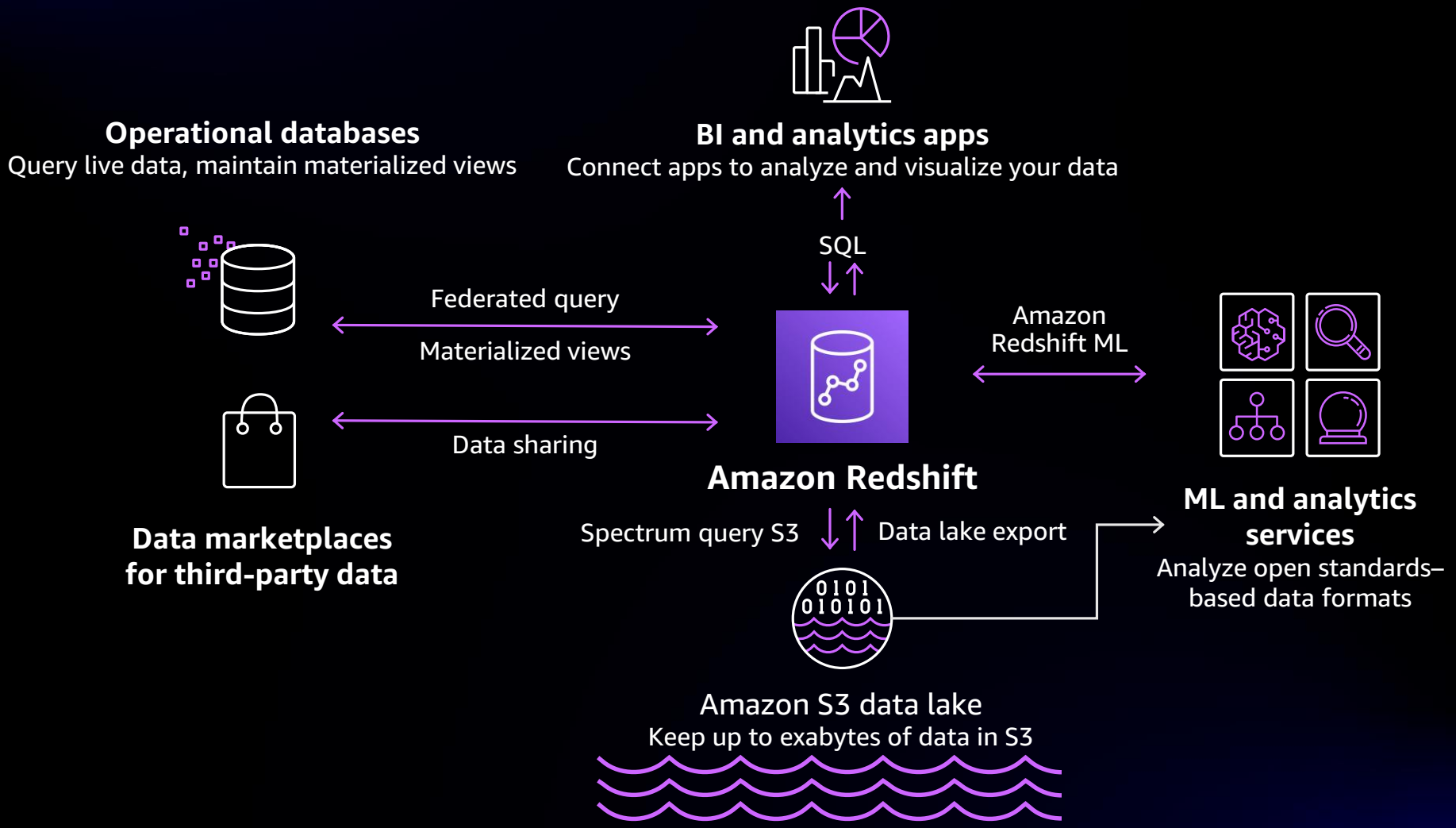
Performant and cost-effective



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Connected insights on all your data

WITH AWS SERVICES INTEGRATION

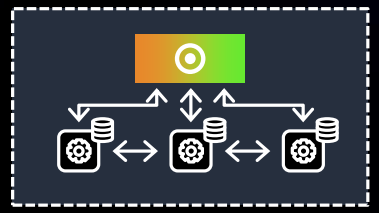


© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

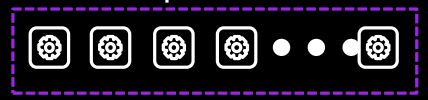
Evolving architecture (2017–2021)

INCREMENTAL FEATURES RELEASED IN THE LAST FEW YEARS

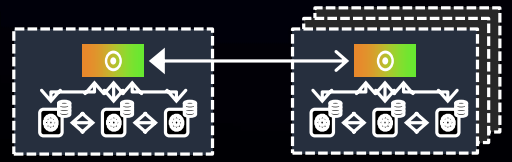
Redshift Spectrum for data lake analytics



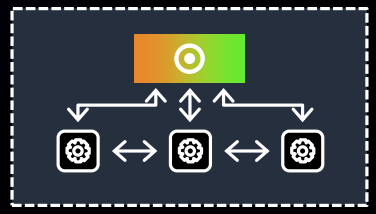
Spectrum



Concurrency Scaling for bursty workloads

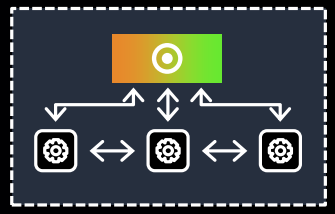


RA3 with independent compute and storage scaling

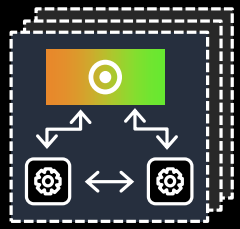


Data sharing across clusters

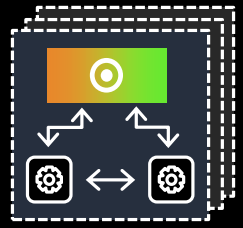
Producer



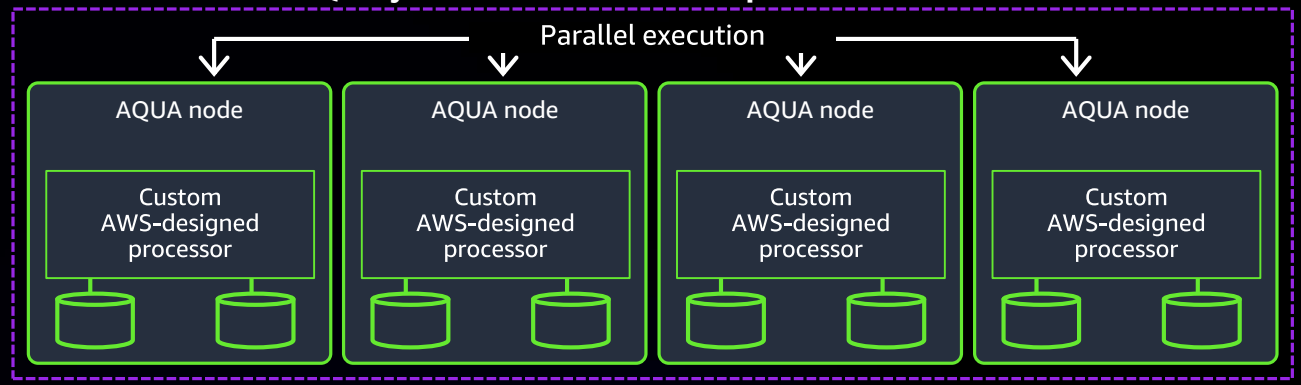
Consumers



Redshift Serverless



Query acceleration with computational cache







aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Enhanced concurrency scaling

READ AND WRITE QUERIES SUPPORT

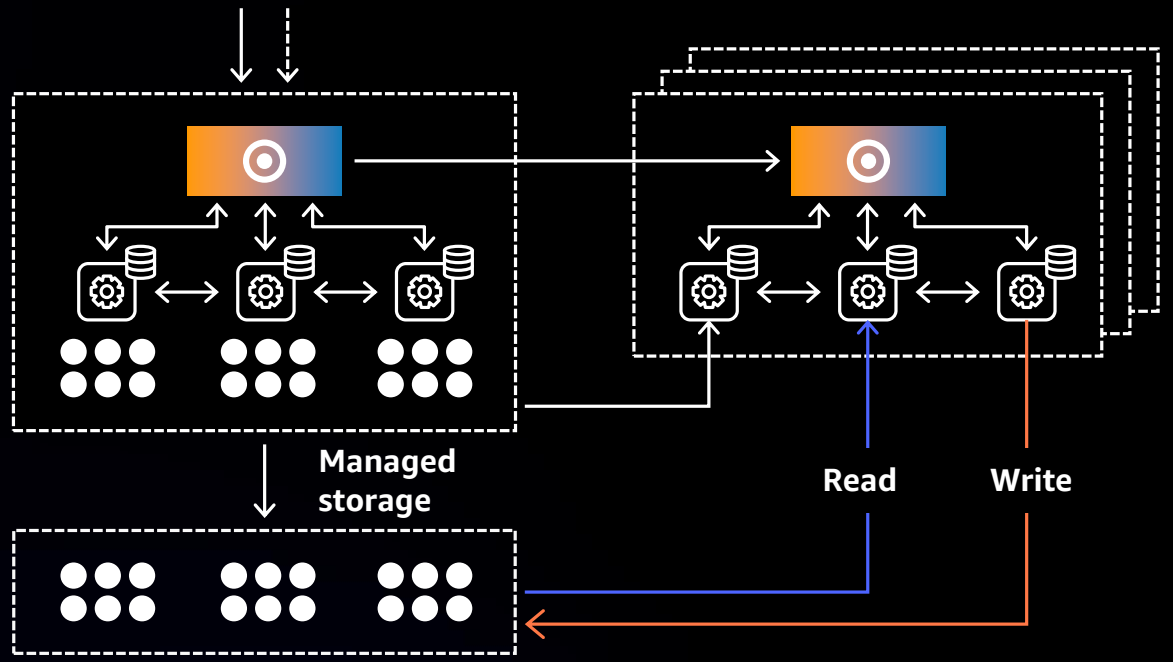
Use case
Auto scale write queries, in addition to read queries, for virtually unlimited concurrent users and queries

---

Example write operations include
INSERT, DELETE, UPDATE, COPY, CTAS,
INSERT INTO SELECT FROM

Write queries scale out to multiple Redshift clusters automatically when concurrency scaling is enabled—no new user action necessary

Per second billing and cost controls apply



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon Redshift Federated Query

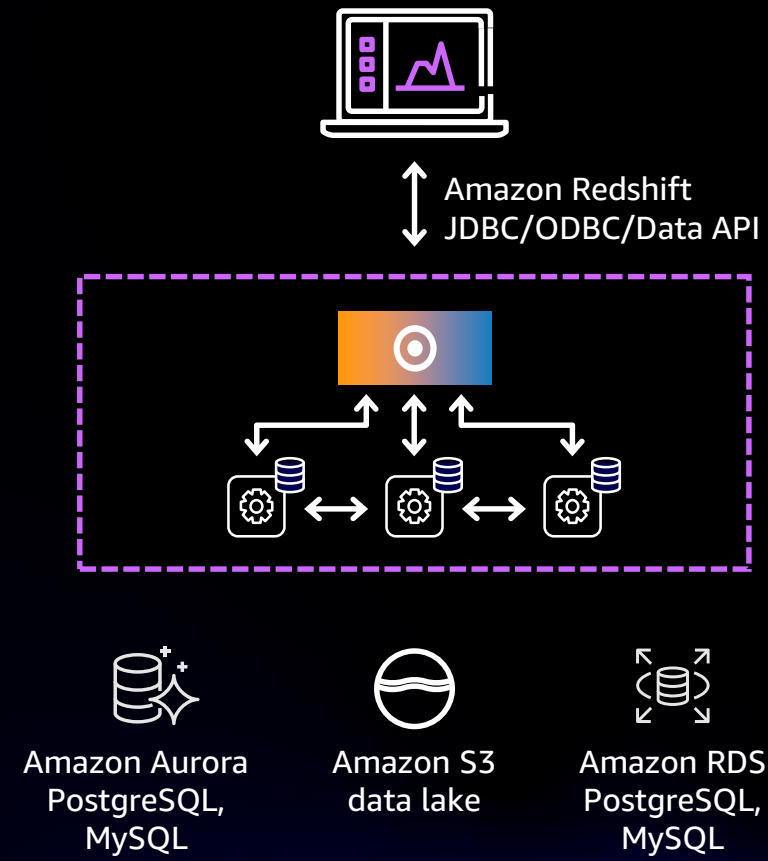
UNIFIED ANALYTICS ACROSS DATABASES, DATA WAREHOUSE, AND DATA LAKE

Use case: Integrate operational data with data warehouse and data lake for real-time analytics

Analytics on operational data without data movement and ETL delays

Query and join data from one or more Amazon RDS and Aurora databases

NEW Amazon RDS and Aurora MySQL support



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data sharing

A SECURE AND EASY WAY TO SHARE DATA ACROSS AMAZON REDSHIFT CLUSTERS

Instant, granular, high-performance data access without data copies / movement

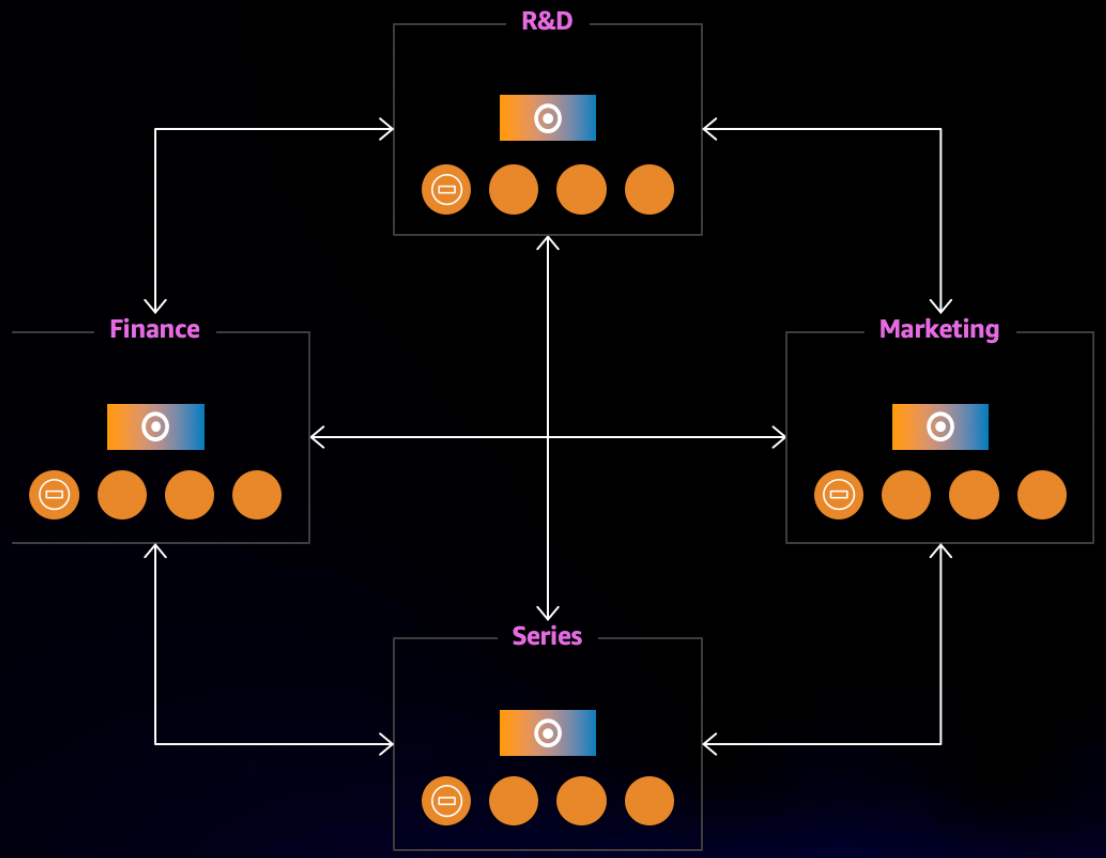
Live and consistently updating views of data across all consumers

Secure and governed collaboration within and across organizations

Workloads accessing shared data are isolated from each other

NEW Cross-account data sharing

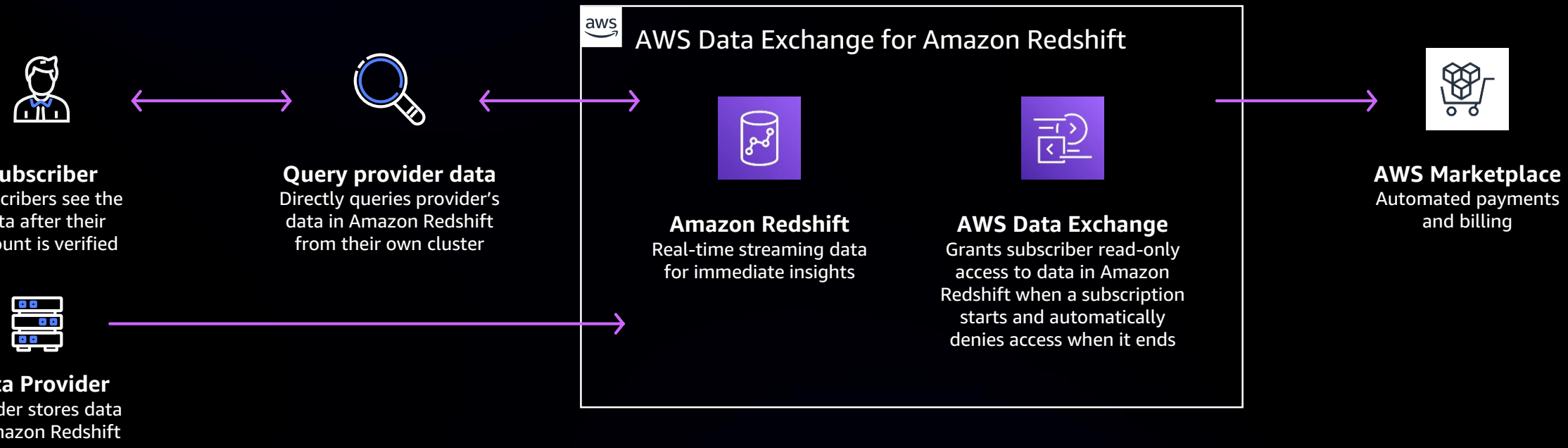
NEW Cross-region data sharing



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon Redshift and Amazon Data Exchange integration



Data providers can package their Redshift data as data products, with pricing terms and conditions

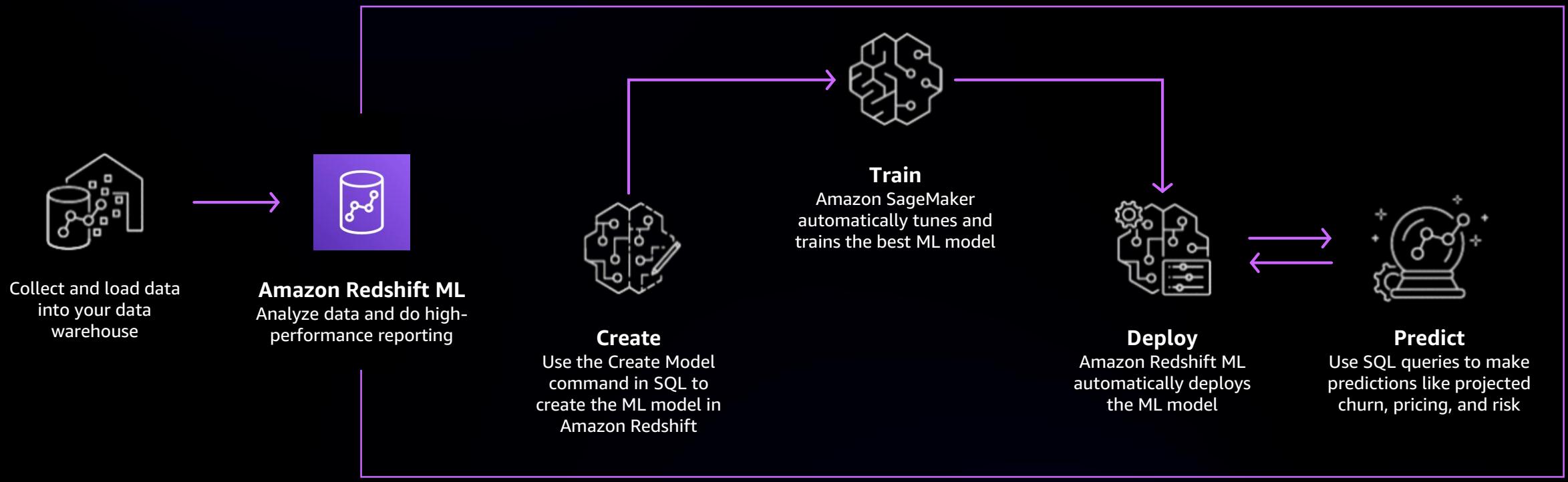
Data subscribers can search and subscribe to data products available in AWS Data Exchange

Get automated "live access" to the data in your Redshift cluster or data lake (via Glue); no ETL or data ingestion necessary

Data providers can update their data directly at the source without having to ETL or generate a new file

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon Redshift ML



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GE Digital Proficy Manufacturing Data Cloud

Madhu Muppagowni

Senior Staff Software Engineer
GE Digital

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

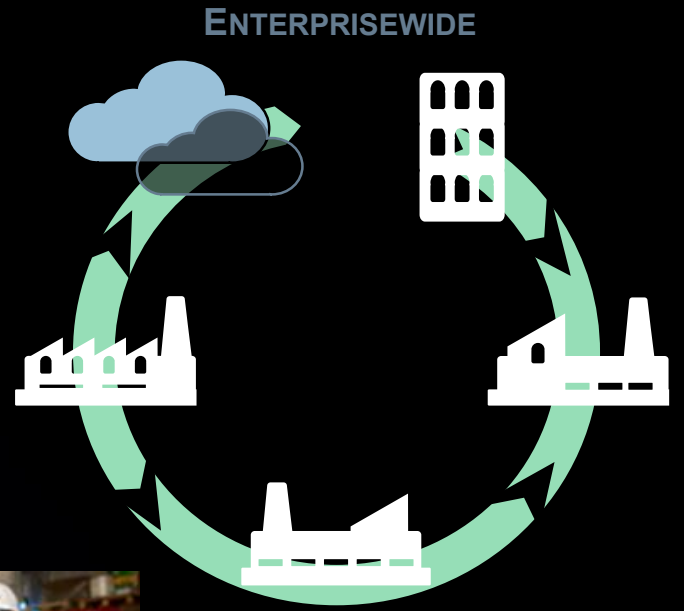
GE Digital manufacturing software portfolio

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GE Digital manufacturing and digital plant

Preproduction ▸ Production ▸ Postproduction

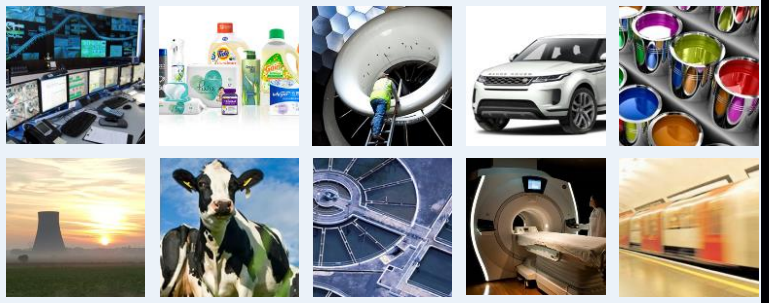
Transforming how industries solve their toughest challenges by putting industrial data to work



Our manufacturing and digital plant software

keeps the world running

We provide control and monitoring solutions and manufacturing execution systems for thousands of major brands around the globe, continuously evolving each of our customer's digital transformation journeys…





Proficy HMI / SCADA
Drive smarter operator decisions with model-based high-performance HMI for faster response and development



Proficy Historian
Powerful industrial time-series and A&E data collection for on-premise and cloud storage and analysis



Proficy Manufacturing Execution Systems (MES)
Enable lean manufacturing through IIoT insights and intelligence



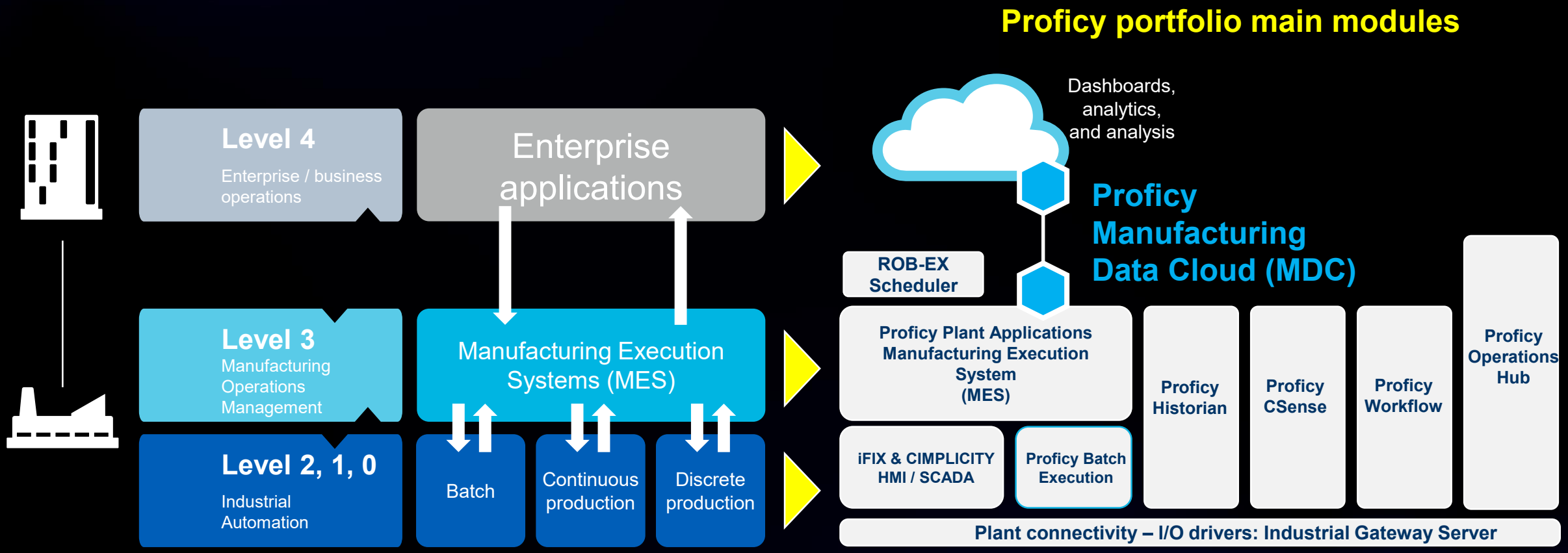
Proficy Analytics
Industrial analytics software that improves asset and process performance with a Process Digital Twin

aws

Proficy Manufacturing Data Cloud

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Proficy MDC and GE Digital manufacturing software portfolio



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GE Digital hybrid cloud strategy for digital plant

Proficy Manufacturing Data Cloud

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Manufacturing challenges

What we heard …

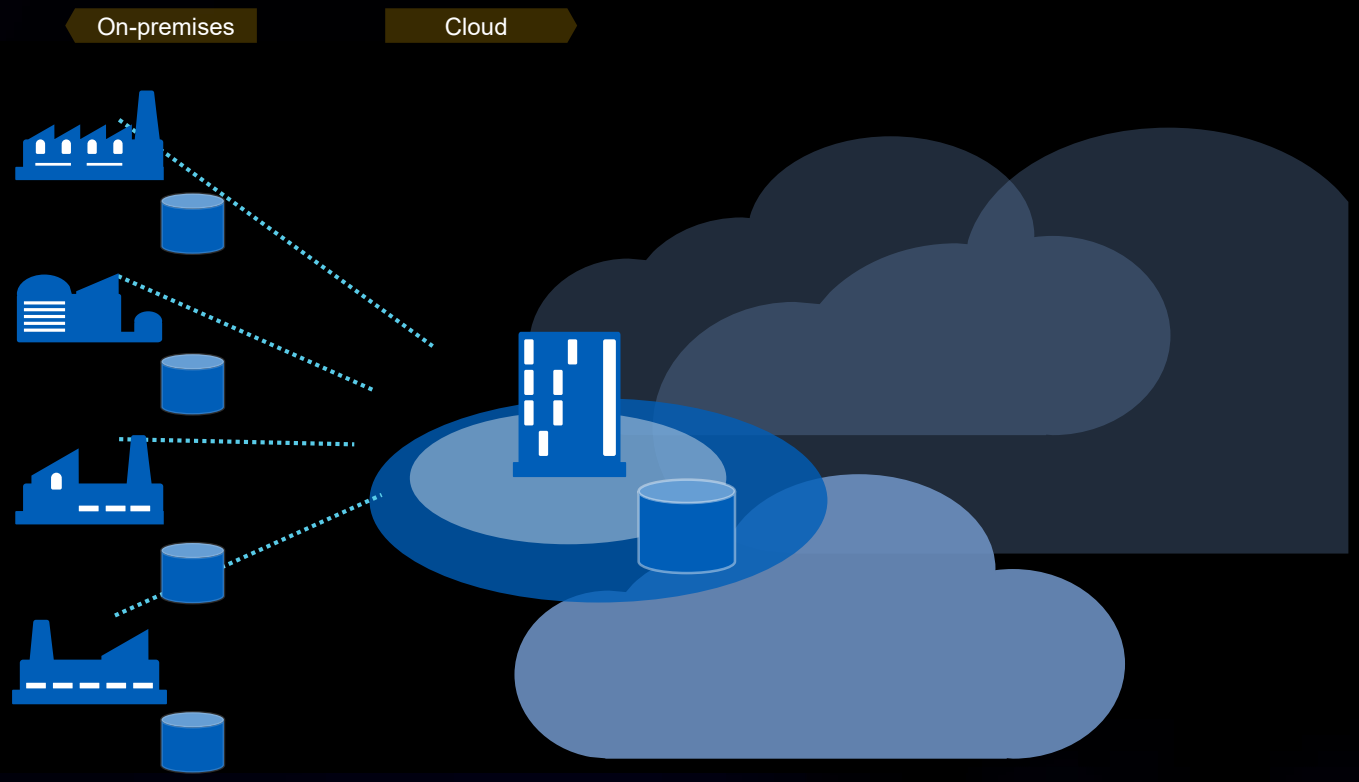
“I need …”

- **Global visibility** *beyond the four walls* – for example, enterprise dashboards
- To perform **advanced analysis and analytics** on my data; compare performance of assets, lines, plants...
- To reduce my IT cost / on-premises footprint
- To support costs, configuration management, and performance issues on multiple back-end data services connecting apps / reporting applications to PA DB
- To ensure **interoperability** with other systems; expose data in a standard way
- Data governance and security

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GE Digital's hybrid cloud strategy

ENTERPRISEWIDE VISIBILITY AND ANALYTICS



Keep key / critical capabilities on-premises

Use the cloud for capabilities like

- Advanced analytics in the cloud
- Enterprise OEE in the cloud

Understand what data to keep and where to keep it, based on criteria like

- Regulatory control
- Critical to operations
- Not required – long-term

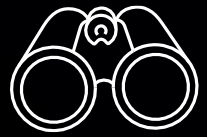
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Add Proficy MDC on top of Proficy Plant Applications

Business outcomes



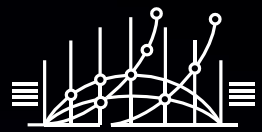
A single source of truth for data



Increased visibility on enterprise performance



Replicate best-in-class practices across sites, fleet, enterprise



Cloud-based analytics



Cost reduction and performance improvement



Increased security

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

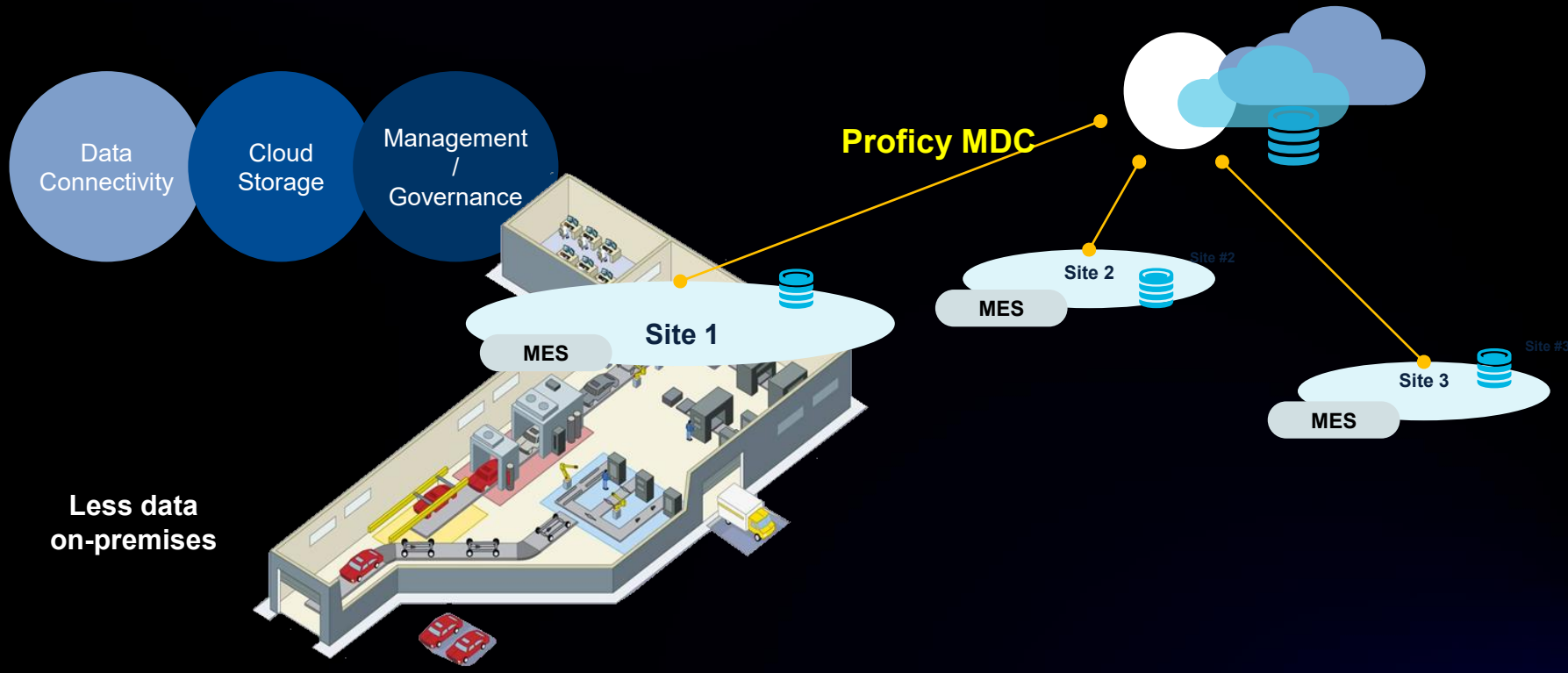
What is Proficy MDC?

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Proficy Manufacturing Data Cloud

UNLOCK THE POWER OF YOUR DATA

Proficy MDC reduces the on-premises database footprint by shifting enterprise data to the cloud and leveraging cloud platform capabilities: secure transfer, storage



Dashboards

Enterprise visualization
- Plant drill-down dashboards
- HQ drill-down dashboards
- Big screen dashboards

Big data analytics

Enterprise analytics
- Analytics services
- Analytics catalog

OPM and APM digital thread

Enterprise events
- Downtimes event and reasons
- Shift / crew events
- Schedule events
- Production events
- Quality events and results

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

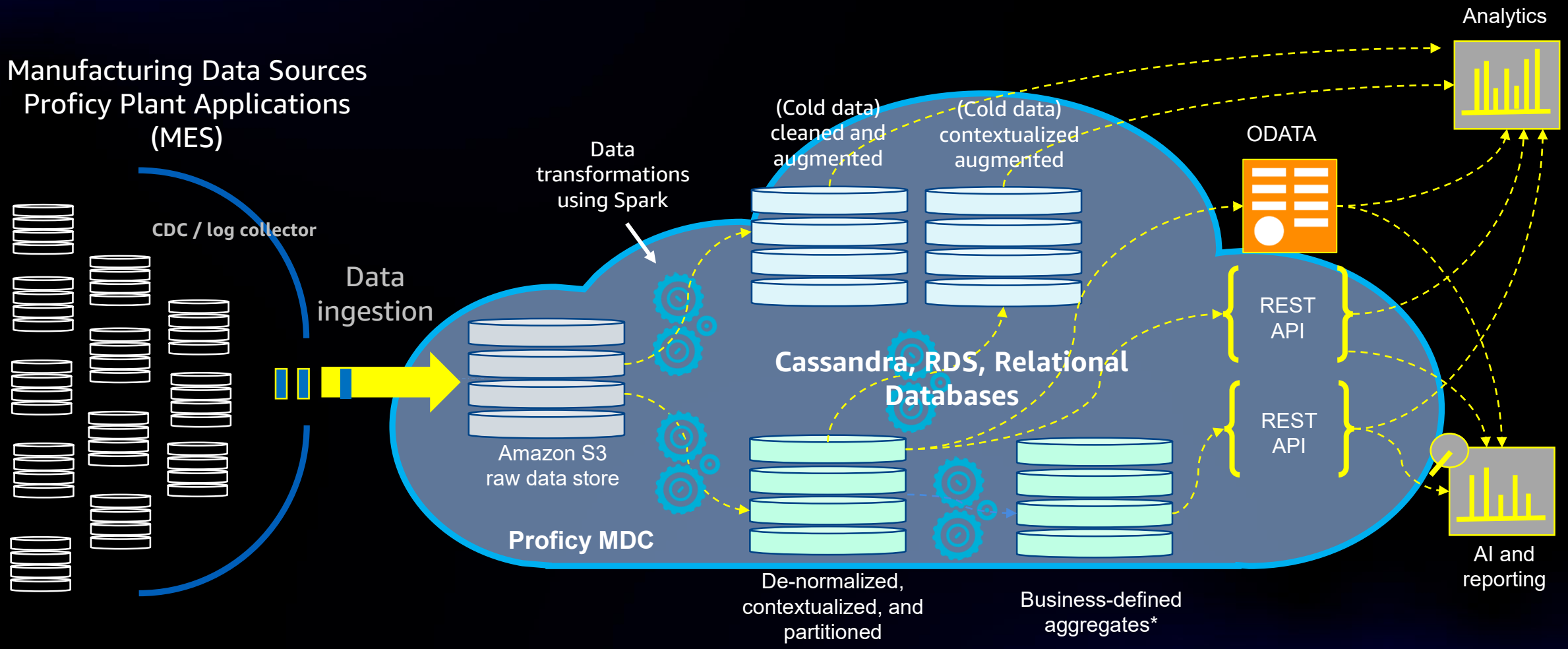
Proficy Manufacturing Data Cloud

Proficy MDC is a plug-and-play cloud BI data store for Proficy Plant Applications

- Securely consolidates and transforms manufacturing data from disparate Plant Applications systems to a cloud-based data repository
- Provides easy access to a standardized set of enterprisewide manufacturing data models for analysis and analytics, dashboarding, manufacturing big data analytics, and connection to other cloud-based solutions
- Provides manufacturing intelligence and insights; for example, improve plant throughput, reduce waste, increase uptime

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Proficy MDC prior data flow



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Proficy MDC metrics

Raw / cold storage

Size: Currently holds a petabyte of data and growing
Uses: Source for all transformations, customer audits

Hot and warm storage

Size: 14 TB for the warm duration
(not Including Amazon S3 storage)
Uses: ETL, customer reporting,
analytical applications, BI tools

Ingestion

Up to 1 GB of compressed data per site per day

Processing

> 300 million records processed every day

Consumption

> 100,000 API requests per day

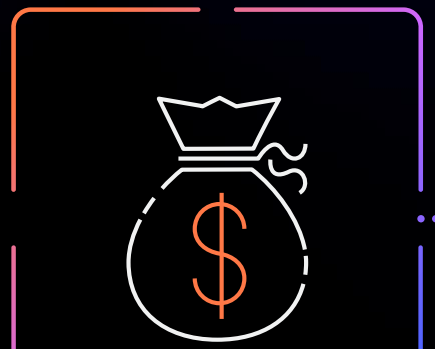
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Proficy MDC challenges and path forward

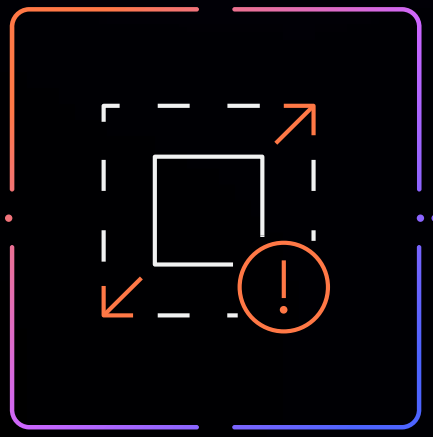
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Where we were

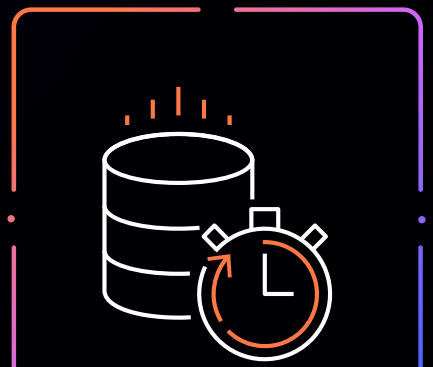
PROBLEMS THAT NEEDED TO BE RESOLVED



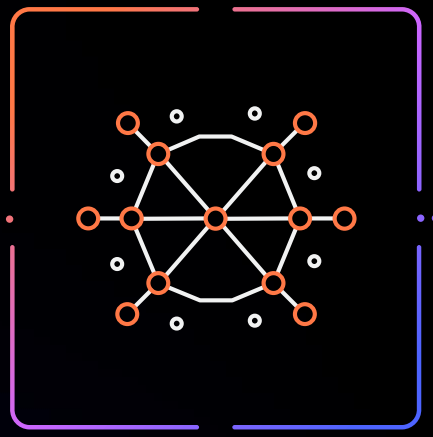
Increased infrastructure and operational costs



Scalability and management overhead (was self-hosted)



Increased data latency and impacted API performance



Complexity of development process

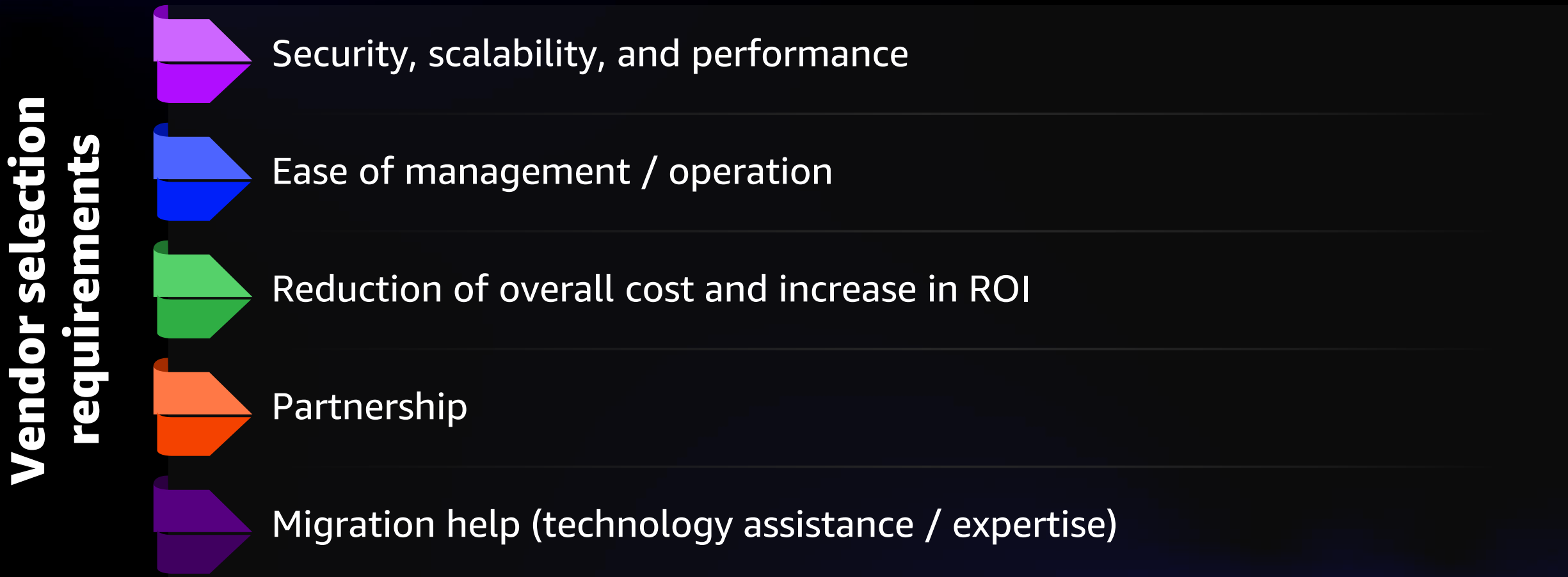


Limited out-of-the-box functionality

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Our road map to a solution

KEY STEPS AND DECISIONS



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Re-platform with Amazon Redshift



Proficy Manufacturing Data Cloud

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Why we chose Amazon Redshift and AWS

KEY STEPS AND DECISIONS



Amazon Redshift

aws

Amazon Redshift is scalable and fully managed, and provides separation of storage and compute

AWS offers many additional value propositions that can be incorporated into our product in the future

Reduction in operational cost compared to other vendors

AWS services are broadly used within GE Digital and have gone through internal security review

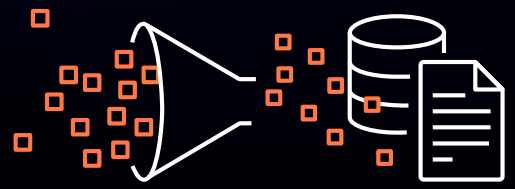
Experienced pool of AWS ProServe team members to help with migration and best practices

AWS is a trusted GE Digital partner

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Migration decisions

KEY STEPS AND DECISIONS

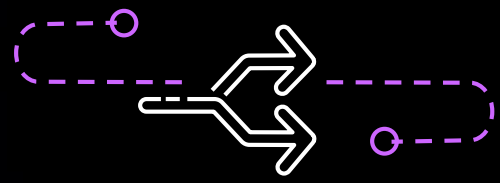


Phased migration of data models:

Apply learnings to later phases

Realize performance and cost benefits as soon as possible

Make the migration transparent to the customer



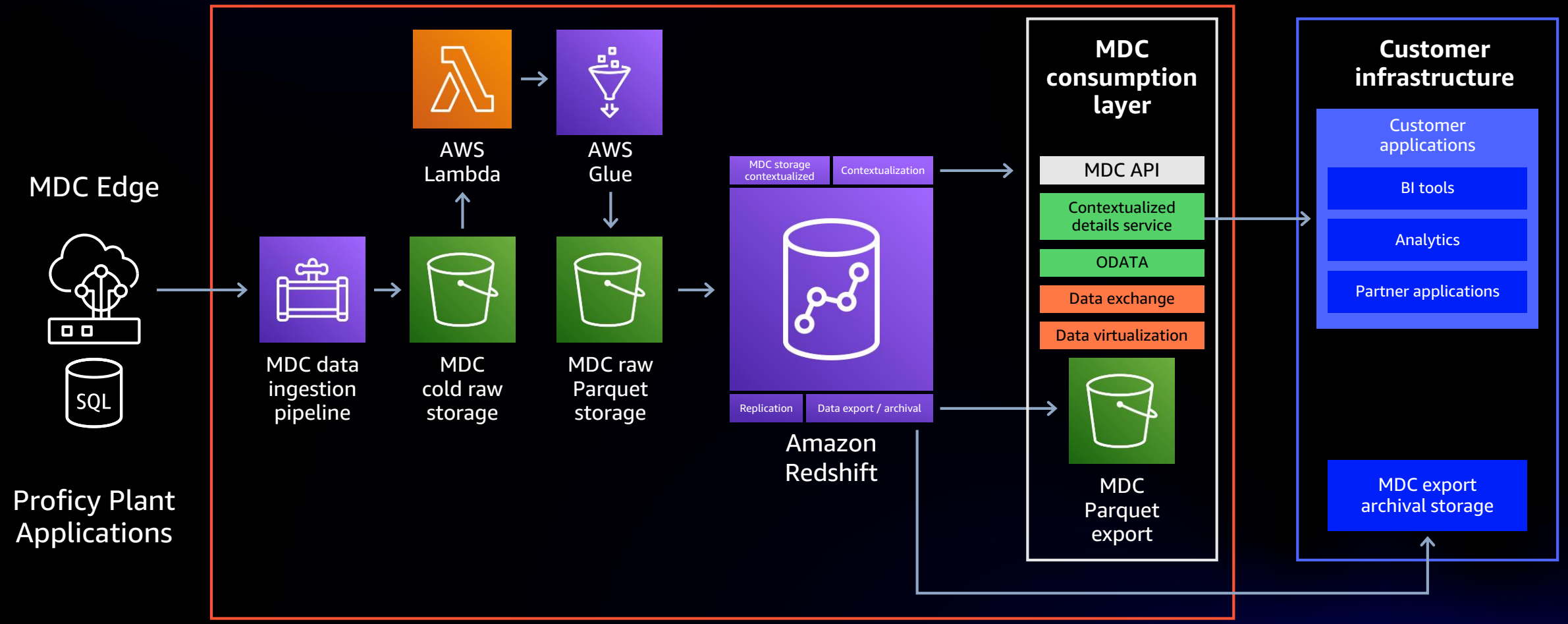
Integration of AWS ProServe team:

Knowledge transfer while migrating and building expertise in-house

Team learns the technology options and their associated tradeoffs

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Current Proficy MDC architecture with Amazon Redshift



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key Redshift and AWS functionality

Federated query – helped us with phased migration

Auto WLM with short-query caching

Spectrum query of S3 objects

Concurrency scaling

Data sharing with open-storage format

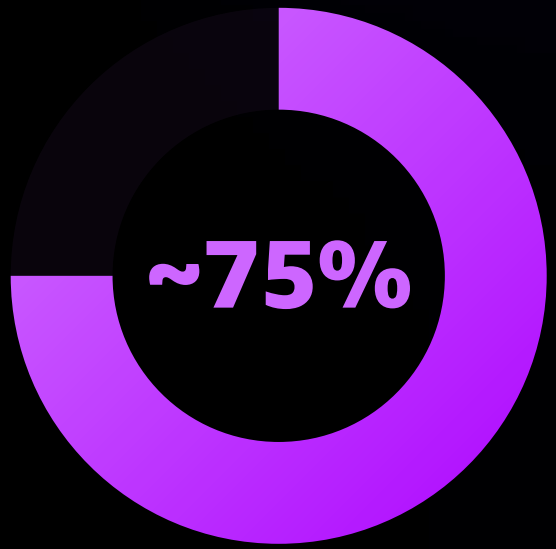
Latest RA3 instance types

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Where we are today with Redshift

WHAT ARE WE SEEING

ELT performance



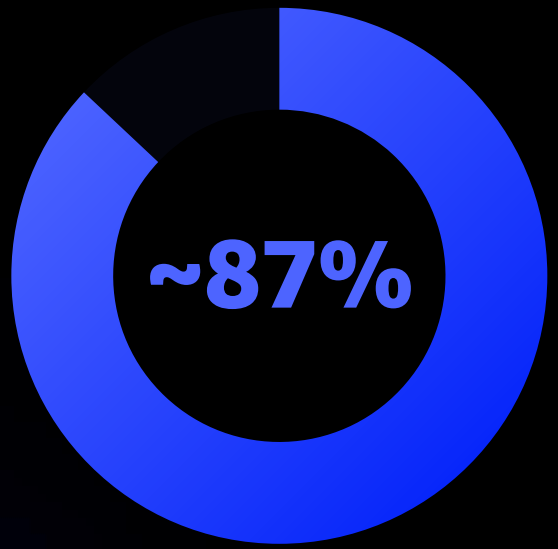
improvement

Lower infrastructure, maintenance, and management costs

Improved customer data accessibility

Improved ease of development

Lower data latency from source to data models



improvement

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

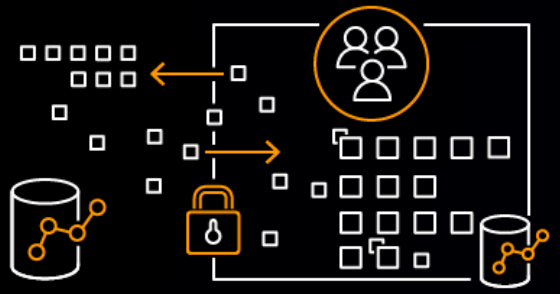
Amazon Redshift use cases



Business intelligence



Operational analytics on events



Data as a service



Predictive analytics

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Harshida Patel
hspatel@amazon.com

Madhu Muppagowni
Madhu.Muppagowni@ge.com

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.