ANT201

What's new with Amazon Redshift

Eugene Kawamoto
Director, Amazon Redshift Product Management
AWS

Aurelie Bergugnat
Chief Data Officer
Senior Vice President, Data and Performance Management
Schneider Electric

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

Agenda

Challenges and business opportunities with data

Why Amazon Redshift?

Key innovations and new features

Schneider Electric

Summary

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

“In 2020, 64.2ZB 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, Worldwide Global DataSphere Forecast, 2021–2025, Doc # US46410421, March 2021



“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?”

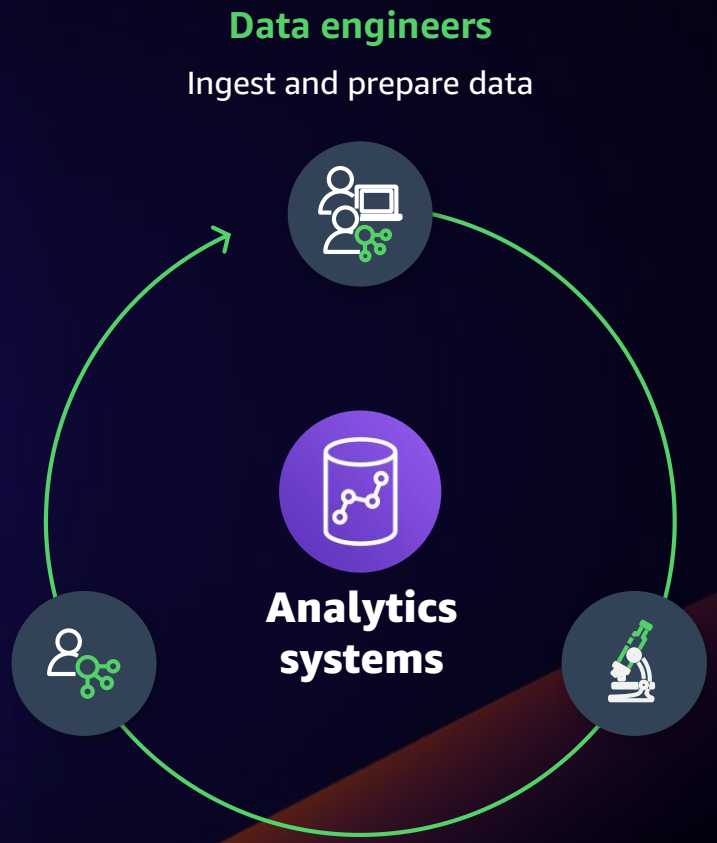


“Performance has to be consistently fast; costs are hitting the roof”

aws

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

Data can work more effectively for diverse users through easy-to-deploy, self-serve, and auto-scaling analytics systems



Improve operational efficiency; make more informed decisions; accelerate innovation

By making 10% more data accessible, a typical Fortune 1000 company will see a ***$65 million increase in net income***[1]

[1] Dykes, "The Four Key Pillars to Fostering a Data-Driven Culture"

aws

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

Future proof your organization with a modern data strategy

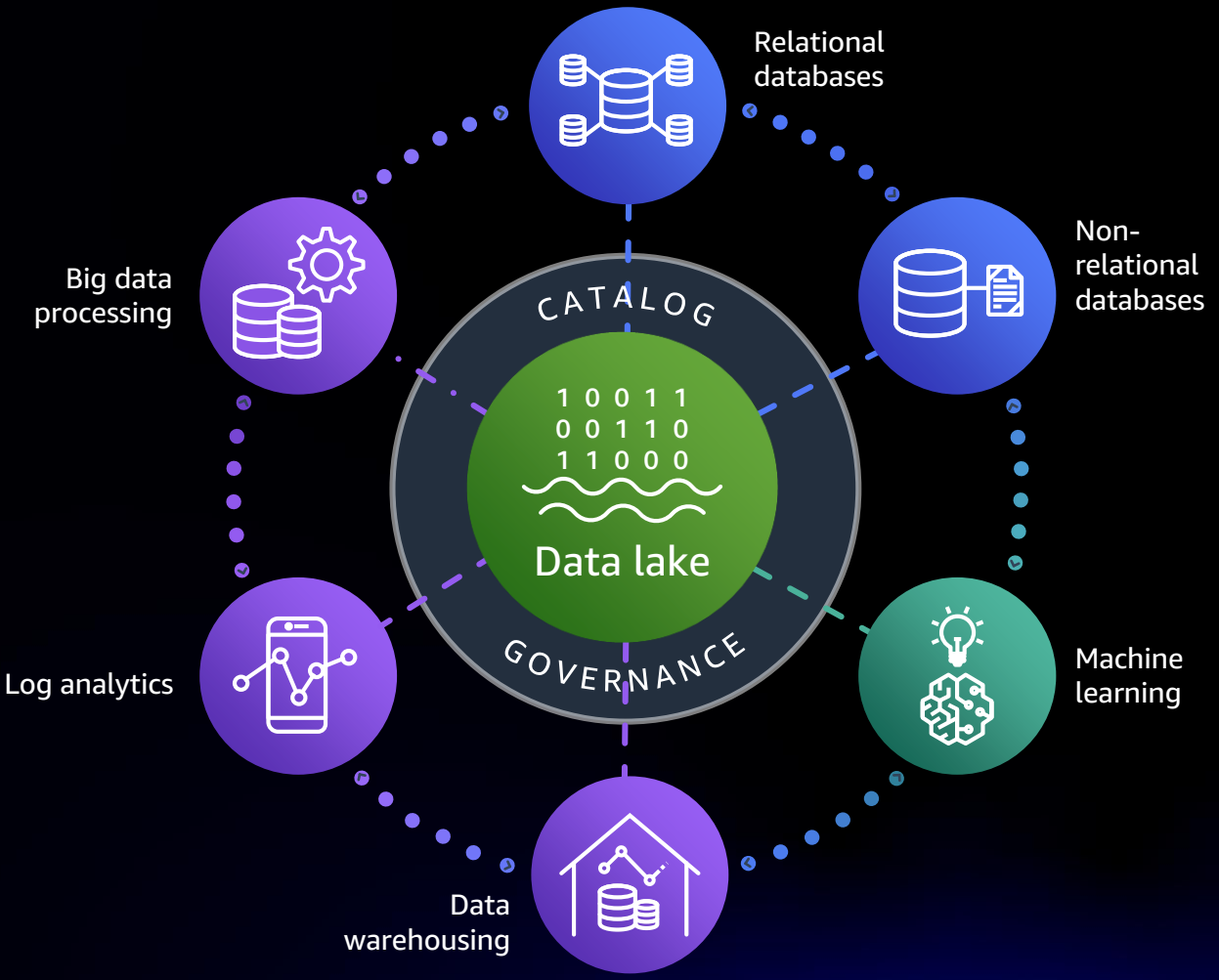
Data at any scale

The best price performance

Seamless data access

Unified governance

AI and ML to solve business challenges



aws

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

Why Amazon Redshift for your data needs?

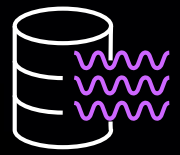
FAST, EASY, AND SECURE ANALYTICS AT SCALE

Easy analytics for everyone



Focus on getting from data to insights in seconds without worrying about infrastructure

Analyze all your data



Get insights running real-time and predictive analytics on complex, scaled data across your operational databases, data lake, and data warehouse

Best price performance at any scale



Gain up to 3x better price performance than other cloud data warehouses, and dynamically scale to improve query speed for complex and critical workloads

aws

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

Tens of thousands of customers process exabytes of data with Amazon Redshift daily

Schneider Electric

Supports tens of thousands of users through Redshift concurrency scaling and RA3 nodes to support a green future

Nasdaq

Accommodated the jump from 30 billion records to 70 billion records a day because of the flexibility and scalability of Amazon S3 and Amazon Redshift

Magellan Rx Management

Redshift ML for predictive analytics to forecast future drug costs and identify trends

zynga

ETL performance improved 2x and scaled to process > 5.3 TB of daily game data

jobcase

Performs model inference on 2–3 billion job search recommendations in 15–20 minutes, down from 2–3 hours, with no additional cost

AstraZeneca AMGEN ancestry coursera DOW JONES EQUINOX GE yelp EA FINANCIAL TIMES docomo

intuit London Stock Exchange Liberty Mutual FOX Pfizer QANTAS SCHOLASTIC Sysco tinder jobcase WB

aws

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

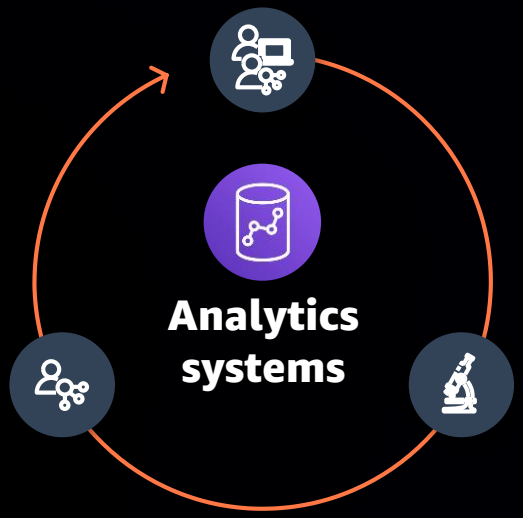


Easy analytics for everyone

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

Easy analytics for everyone

FOCUS ON GETTING FROM DATA TO INSIGHTS IN SECONDS



Data engineers

Ingest and prepare data

Business leaders

Make business decisions and deliver on outcomes

Developers, analysts, data scientists

Build applications and analyze data

Automatic provisioning and scaling

Automatically provisions and scales the underlying compute resources to deliver high performance for demanding and unpredictable workloads

Visualize your data

Single, visual interface for querying data to improve productivity through one-click visual analytics, collaboration, version control, and scheduling

Bypass administrative tasks

Take advantage of automated provisioning, backup, patching, tuning, and monitoring in Amazon Redshift clusters

aws

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

Today, Amazon Redshift allows you to customize capacity according to your needs

1. Flexible instance types for granular control of price performance
2. Manual and automatic workload management with query priorities
3. Vertical and horizontal scaling for performance and concurrency
4. Pause/resume clusters to save on costs during idle periods

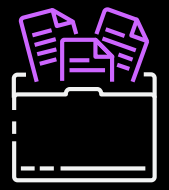
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Customers want to empower more users with Amazon Redshift...



Users with little or no knowledge in data warehouse management

Access data without thinking about selecting instances, sizing, scaling, and tuning



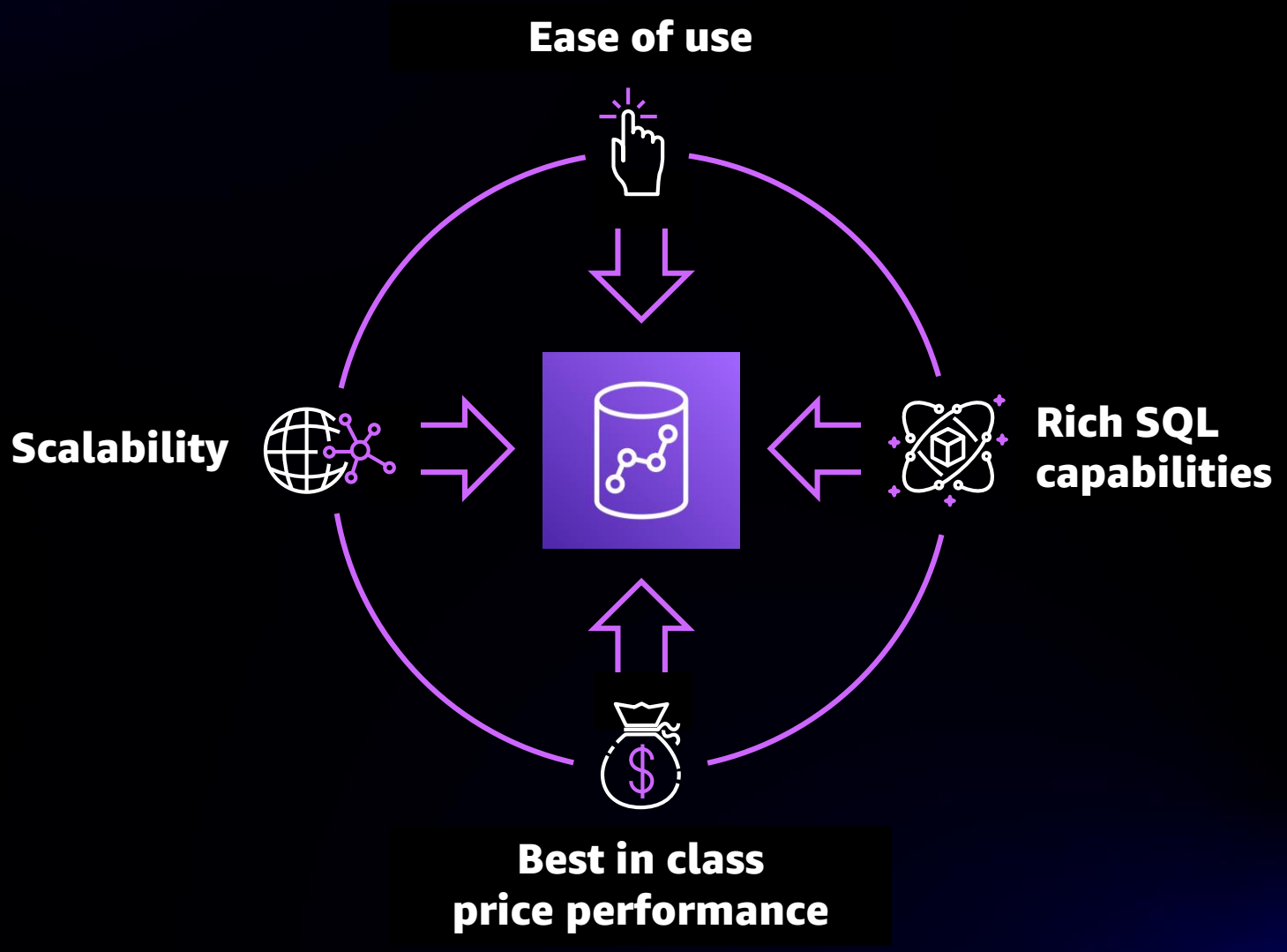
Unpredictable and varying workloads

Diverse use cases with heterogeneous compute requirements

aws

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

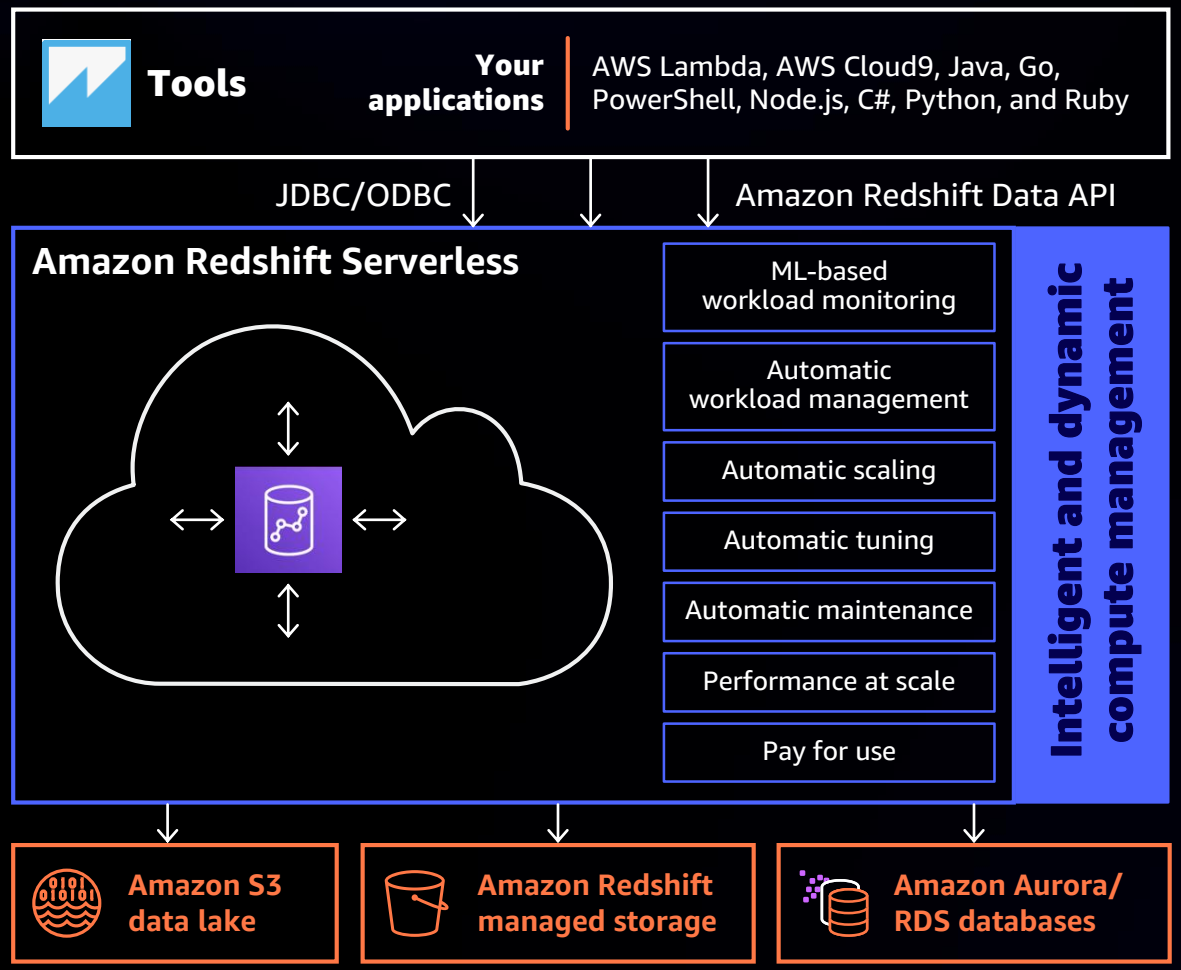
Need for analytics without compromises



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

Introducing Amazon Redshift Serverless

NEW PREVIEW



Simply point applications to the Amazon Redshift Serverless endpoint and start running

All Amazon Redshift SQL functionality applies

- Security and user management
- Complex joins
- Semi-structured data
- Data sharing
- Machine learning functions
- Data lake queries
- Federated query
- Durability and transactional guarantees
- JDBC/ODBC and Data API
- And more

aws

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

Benefits of Amazon Redshift Serverless

Simplified user experience



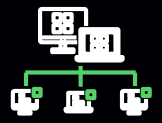
Run and scale analytics without having to manage data warehouse clusters

Intelligent and dynamic compute



Automatically provisions and scales data warehouse capacity to deliver consistently fast performance

All Amazon Redshift functionality and performance



Leverage Amazon Redshift's rich SQL capabilities, seamless data lake integration, and industry-leading price performance at scale

Pay for use



Pay for the compute capacity only for the workload duration on a per-second basis; no charges for idleness

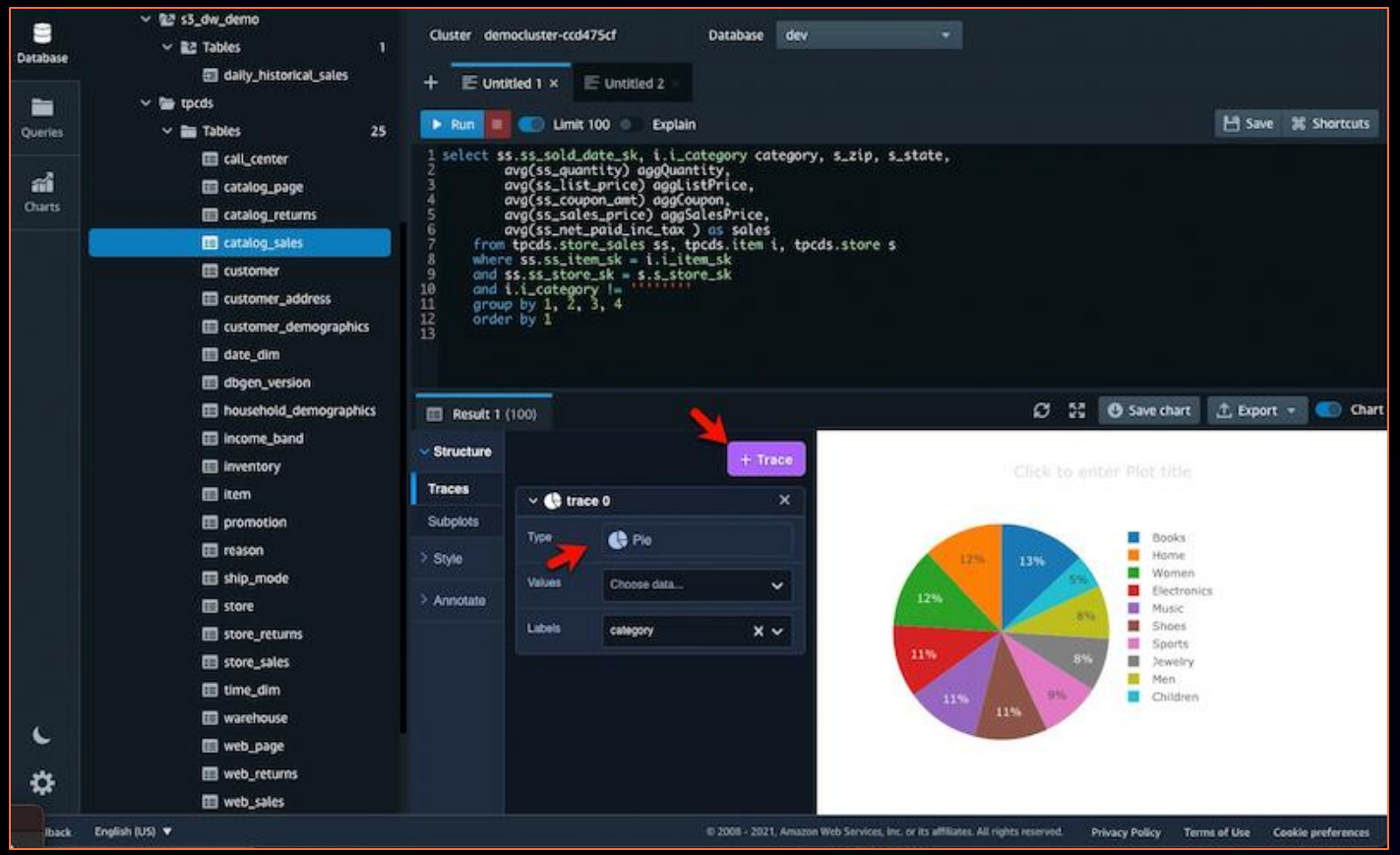
aws

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

Amazon Redshift Query Editor V2

NEW

FREE WEB-BASED ANALYST WORKBENCH FOR DATA EXPLORATION AND ANALYSIS USING SQL



Web-based tool for data analysts, data scientists, and database developers to explore, analyze, and collaborate on data

Share and collaborate on queries and analysis, and track changes with built-in version control

Browse, create schema and tables, load data, author SQL queries and stored procedures, and visualize query results with charts

aws

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

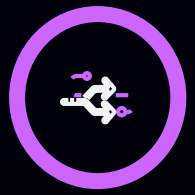
Automated performance tuning

UPDATED

ML-BASED OPTIMIZATIONS TO GET STARTED EASILY AND GET THE FASTEST PERFORMANCE QUICKLY



Automatic vacuum delete



ATO: Automatic distribution keys



ATO: Automatic sort keys



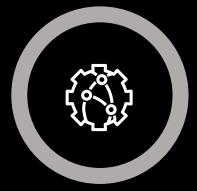
Auto workload manager



Automatic table sort



ATO: Automatic column compression encoding



Auto Analyze



Auto refresh & rewrite Materialized Views

Automates physical data design and optimization

Optimizes for peak performance as data and workloads scale

Leverages machine learning to adapt to shifting workloads

aws

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

Automated materialized views

NEW PREVIEW

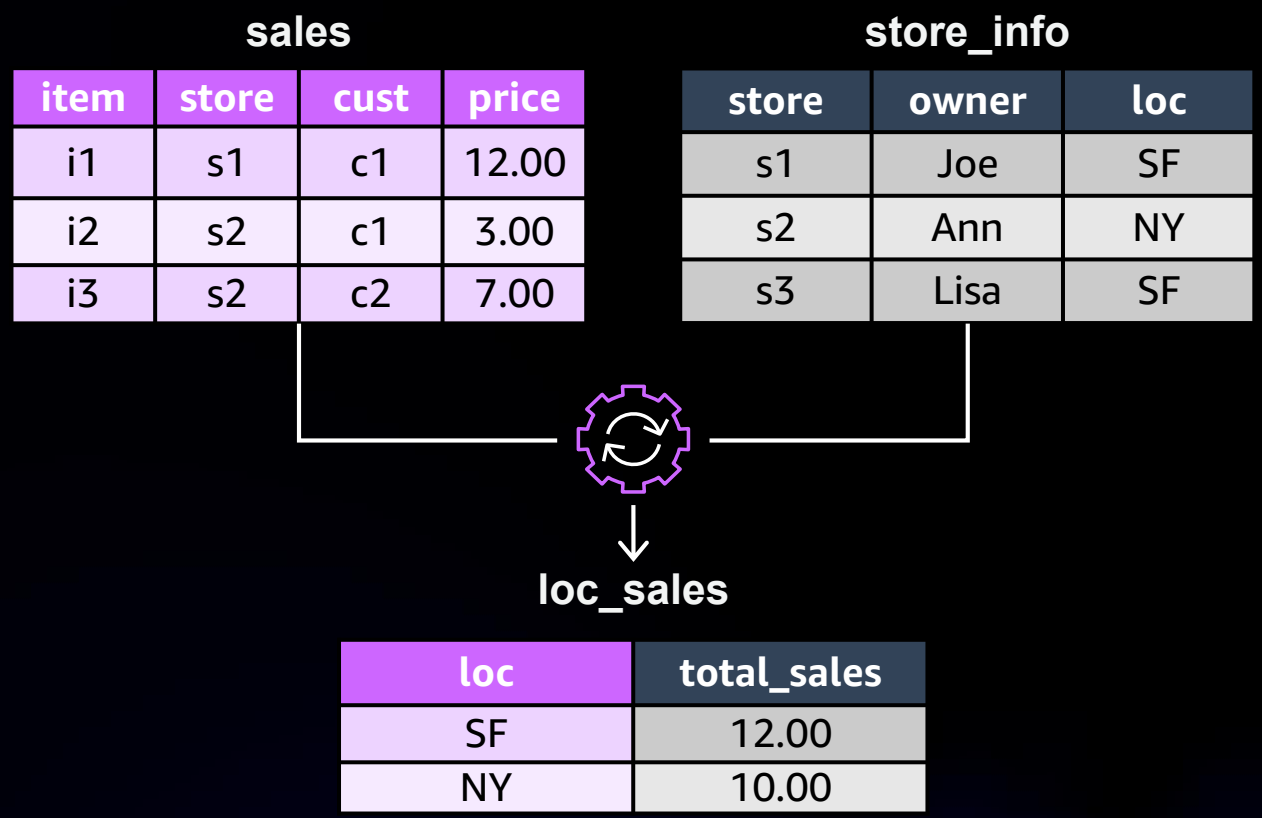
Automated creation and deletion of materialized views, with incremental refresh

Up to orders of magnitude query performance improvement

Continual monitoring to identify queries that will benefit from having a MV and automatically create

Automated query rewrite to leverage MV

Automatically deletes MVs that are no longer useful



aws

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

Amazon Redshift Data API

SIMPLIFIES DATA ACCESS FROM WEB SERVICES–BASED APPLICATIONS

Simplifies data access from web services–based application without configuring drivers and connection pools

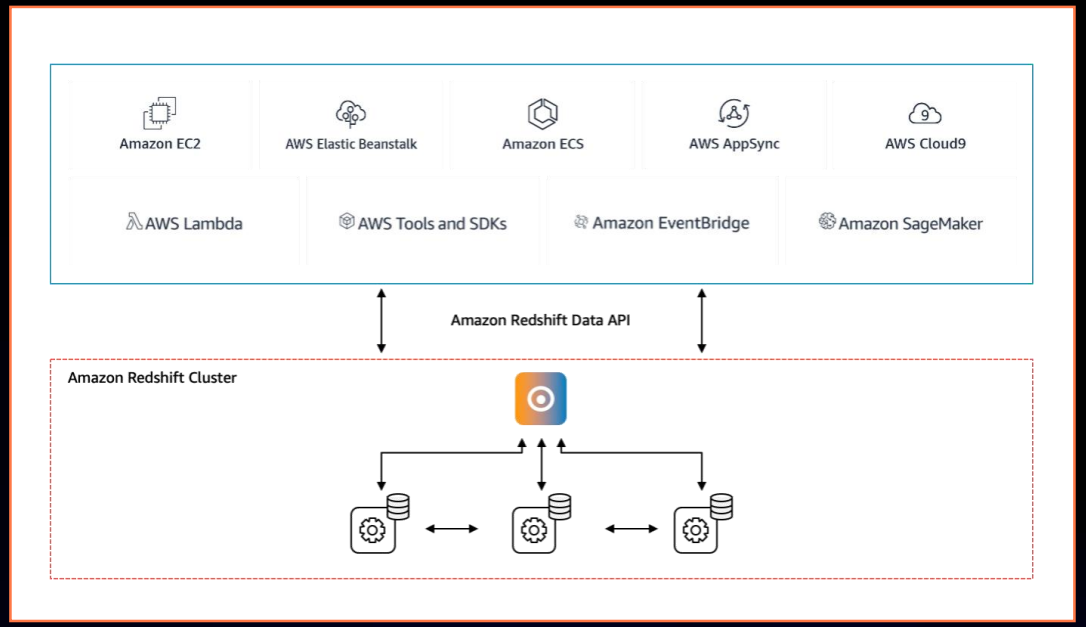
Build your ETL pipeline with AWS Lambda and AWS Step Functions

Build event-driven applications with Amazon EventBridge and Lambda

Simpler access to data from data science tools such as Amazon SageMaker and Jupyter notebooks

Schedule SQL scripts to simplify data load/ unload and refreshing of materialized views

```
aws redshift-data execute-statement
--database [DATABASE]
--query [QUERY]
--secret-arn [CREDENTIALS_ARN]
```



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

Amazon Redshift Grafana plugin

NEW

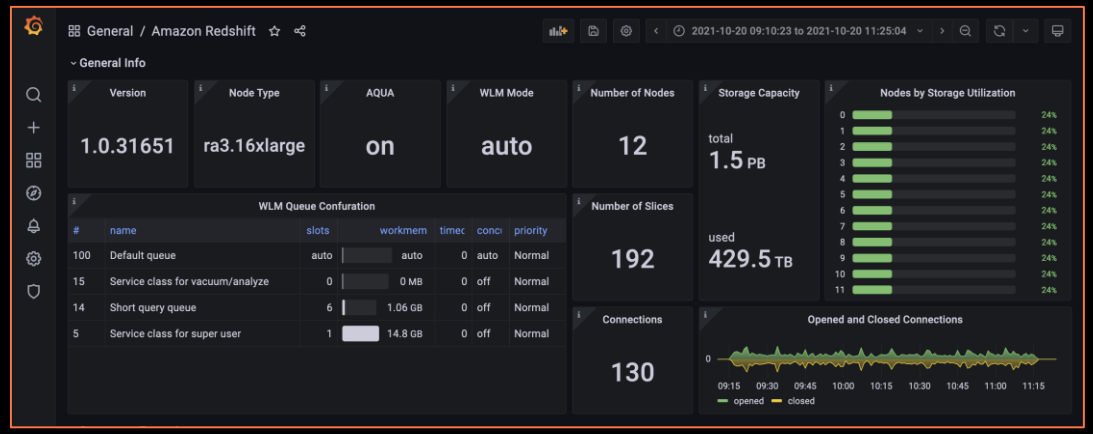
QUERY YOUR OPERATIONAL AND BUSINESS DATA DIRECTLY FROM GRAFANA

Available now in Amazon Managed Grafana and OSS Grafana

Visualize operational metrics and business insights from Amazon Redshift on your Grafana dashboards

Query System Views directly for in-depth system metrics beyond Amazon CloudWatch metrics

A default Amazon Redshift operational dashboard is available out of the box





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



Analyze all your data

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

A modern data strategy

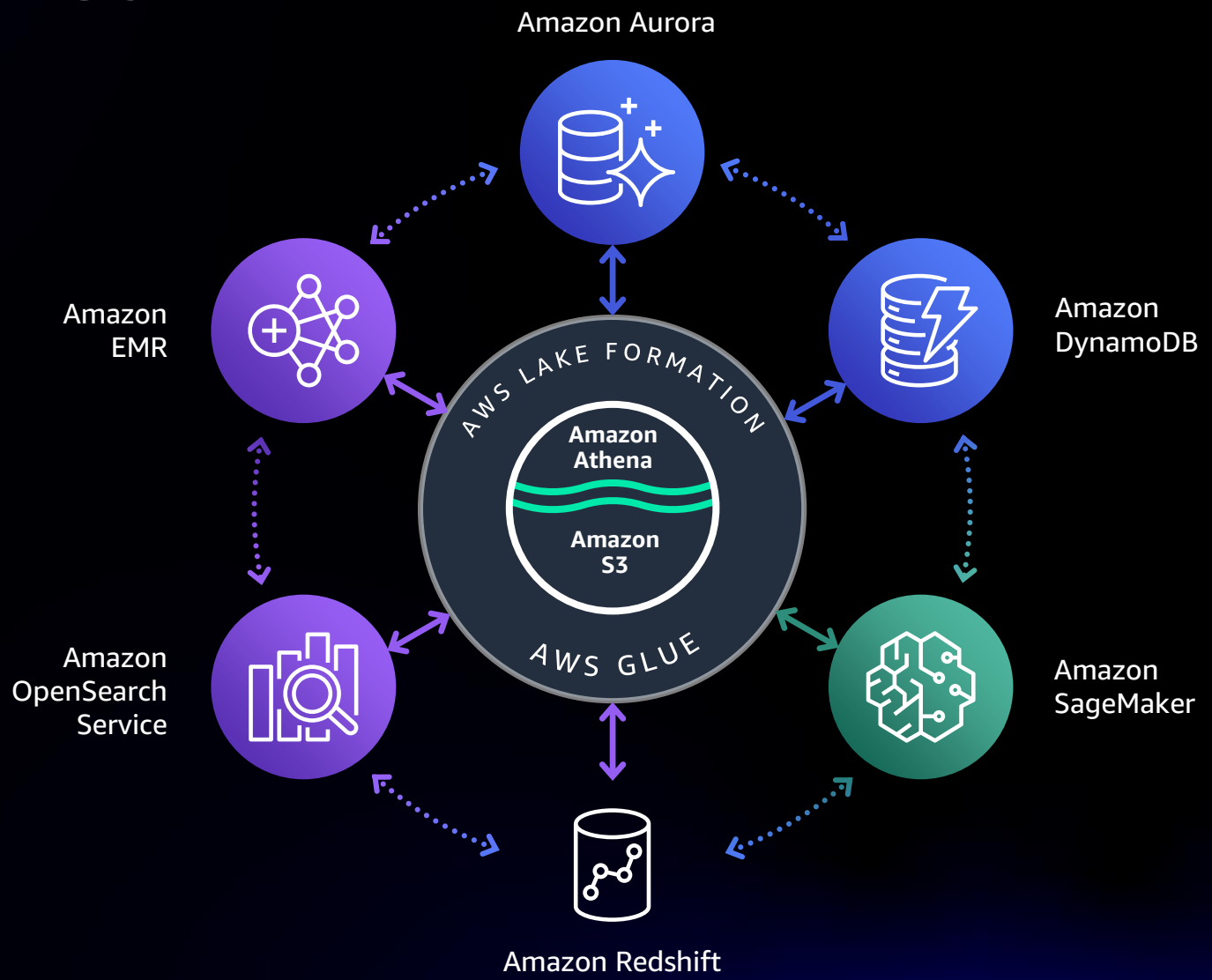
Data at any scale

The best price performance

Seamless data access

Unified governance

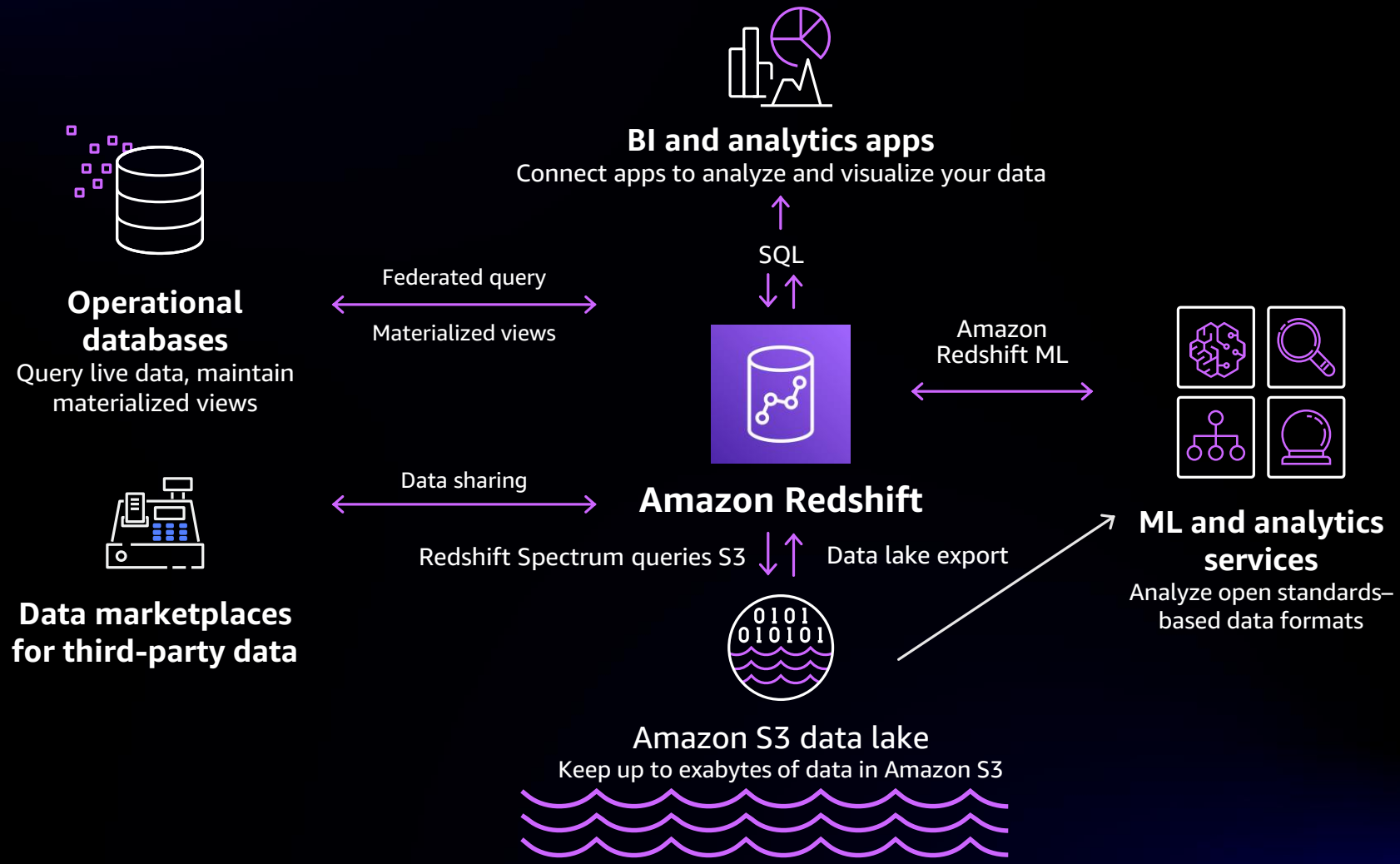
AI and ML to solve business challenges



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

Analyze all your data

WITH AWS SERVICES INTEGRATION



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

Data sharing

UPDATED

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

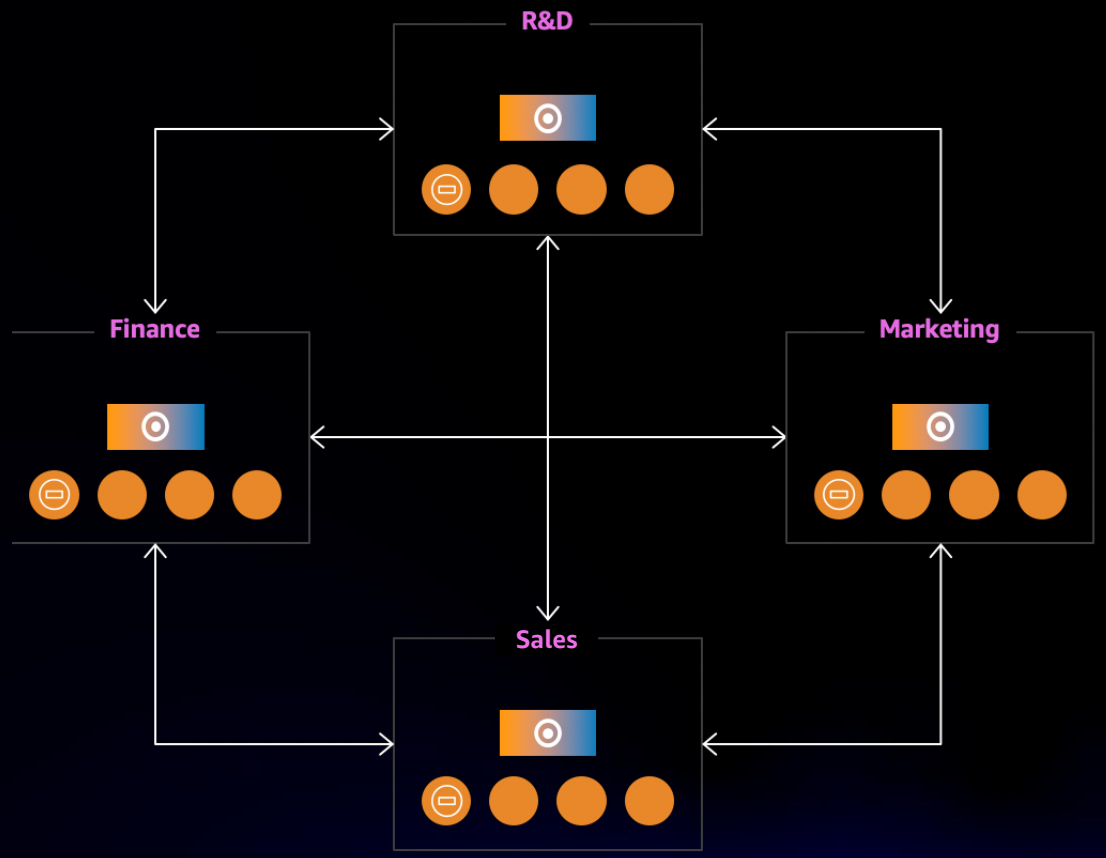
Secure and governed collaboration within and across organizations

Workloads accessing shared data are isolated from each other

NEW Cross-account data sharing

NEW PREVIEW Cross-Region data sharing

NEW Performance Enhancements: Result Caching & Concurrency scaling for consumer clusters



aws

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

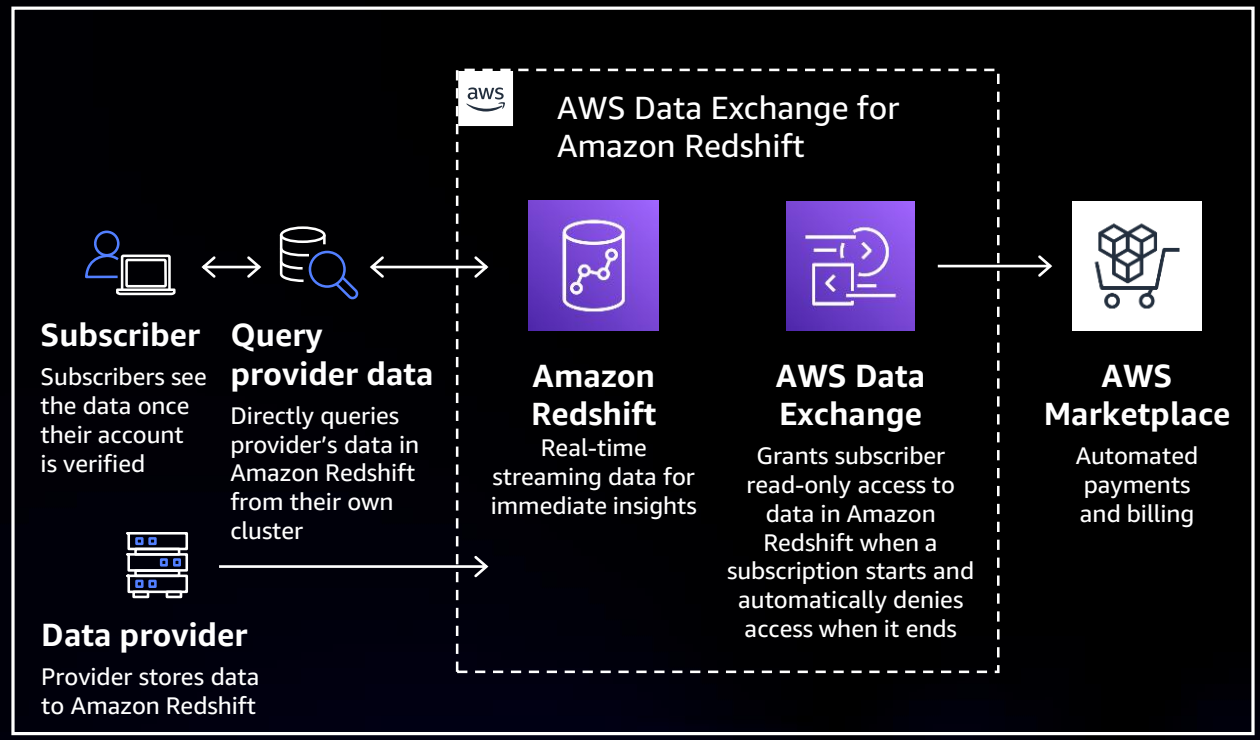
Amazon Redshift and AWS Data Exchange integration

Data providers can package their Amazon 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 Amazon Redshift cluster or data lake (via AWS Glue); no ETL or data ingestion necessary

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



aws

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

Amazon Redshift ML

UPDATED

EASILY CREATE AND TRAIN ML MODELS USING SQL QUERIES WITH AMAZON SAGEMAKER

Use case: product recommendations, fraud prevention, reduce customer churn

Train and create ML models using SQL

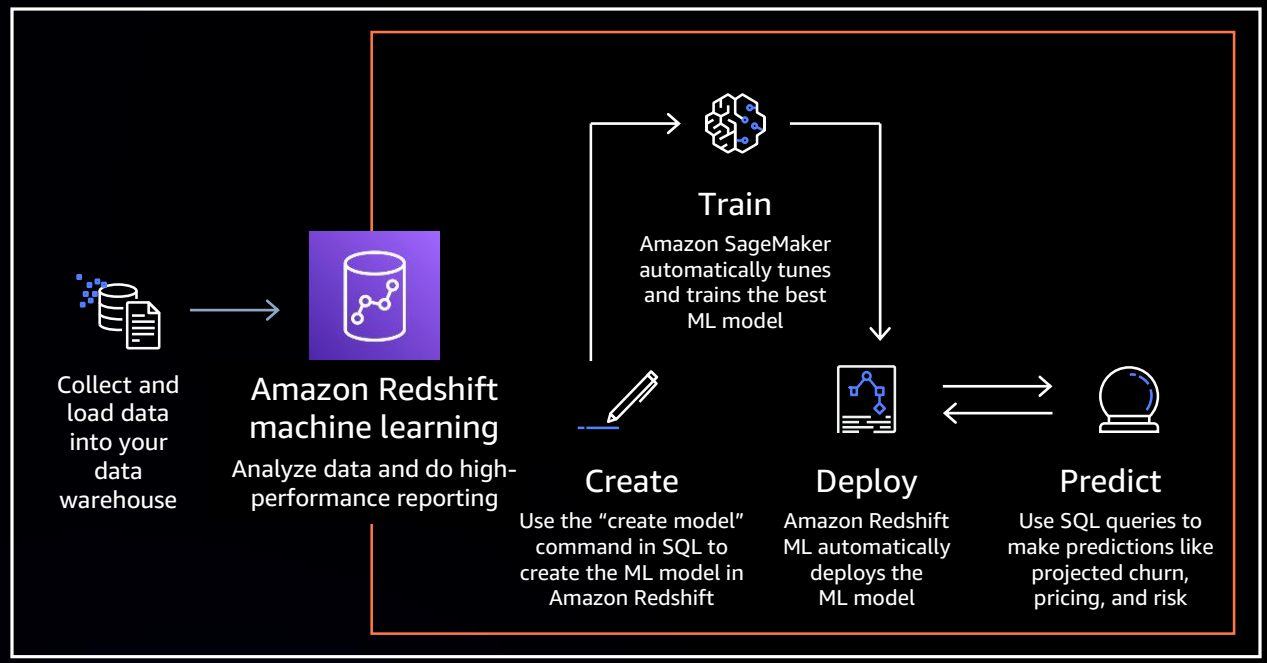
Automatic pre-processing, creation, training, deployment of your model

Deploy inference models locally in Amazon Redshift

Supports ability to bring your SageMaker models for either in-database or remote inference

Supervised and unsupervised trainings

NEW Unsupervised training: K-Means clustering




```
CREATE MODEL customer_churn
FROM (SELECT c.age, c.zip, c.monthly_spend, c.monthly_cases,
c.active FROM customer_info_table c)
TARGET c.active
FUNCTION predict_customer_churn
…;
```

aws jobcase MagellanRx MANAGEMENT

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

Data Store landscape: Supporting the Decision Supply Chain

AWS re:Invent – November 30, 2021



Aurelie Bergugnat
Chief Data Officer
Senior Vice President, Data & Performance Management

© 2021 Schneider Electric, All Rights Reserved

Life Is On | Schneider Electric

COVID-19: accelerator of Digital

Climate change: a code red for humanity

© 2021 Schneider Electric, All Rights Reserved

Life Is On | Schneider Electric



© 2021 Schneider Electric, All Rights Reserved

We need to go at least 3x faster



Efficiency



Electrification



Decarbonization

© 2021 Schneider Electric, All Rights Reserved

Life Is On Schneider Electric

Our mission is to be your digital partner for Sustainability and Efficiency



We deliver decarbonization solutions

c.50% Digital and service
Group revenue

€30 B Energy managed (ESS)

~800 M Energy savings
metric tons of CO_2 saved and avoided on our customers' end by 2025

Life Is On | Schneider Electric

© 2021 Schneider Electric, All Rights Reserved

Empower all to make the most of our energy and resources, bridging progress and sustainability for all

We play on both sides of the equation



Leading by example in our ecosystem

Being part of the solution for our customers

© 2021 Schneider Electric, All Rights Reserved

Life Is On | Schneider Electric

Our distinctive DNA is externally recognized in H1 2021

MEANINGFUL

#1 the world's most sustainable company

2021 GLOBAL100 The Voice for Clean Capitalism Corporate Knights

NATIONAL ABILITY AWARDS

2021 WORLD'S MOST ETHICAL COMPANIES® WWW.ETHISPHERE.COM

Renewable Energy Markets Asia AWARD WINNER 2021

CRS

CLEAN200

Winner Microsoft Partner of the Year Sustainability Changemaker

GREEN STAR SOSTENIBILITÀ la Repubblica 2021

GSSC 2021 GLOBAL SUSTAINABLE SUPPLY CHAIN SUMMIT

INCLUSIVE

TOP CEOs 2021 EMPLOYEES' CHOICE

LinkedIn TOP 2021 COMPANIES France

glassdoor 2021 BEST PLACES TO WORK

Great Place To Work. Certified 2021 USA

Bloomberg Gender-Equality Index 2021

WOMEN IN GOVERNANCE PARITY CERTIFIED

FT DIVERSITY LEADERS 2021 statista

EMPOWERED

The Gartner Supply Chain Top 25 for 2021 Gartner

CONTROL ENGINEERING 2021 ENGINEERS' CHOICE AWARDS

THE CHANNEL CO. CRN PARTNER PROGRAM GUIDE 2021

VISA DE SÉCURITÉ

CES INNOVATION AWARDS 2021 HONOREE

Club des Trente

The Drum Search Awards

DESIGN AWARD 2021

reddot winner 2021

NAHB National Association of Home Builders

BRITISH SAFETY COUNCIL International Safety Awards

© 2021 Schneider Electric, All Rights Reserved

Life Is On Schneider Electric

Internal

Digital partner for Sustainability and Efficiency

Key figures for 2020

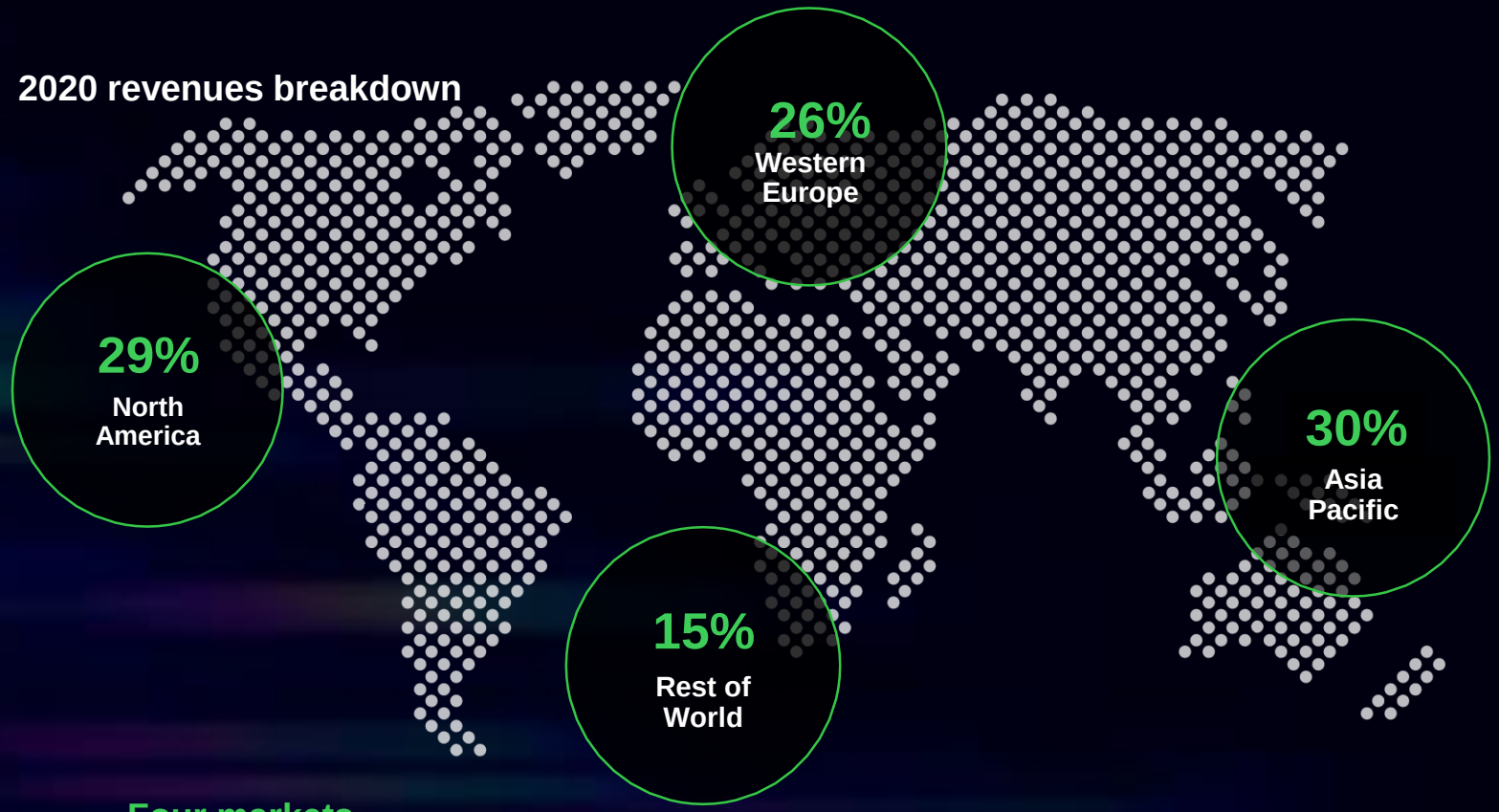
5% of revenues devoted to R&D

€25.2 billion
2020 revenues

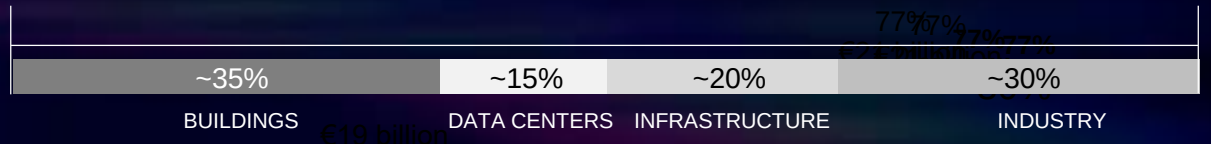
41%
of revenues in new economies

128,000+
employees in over 100 countries

2020 revenues breakdown

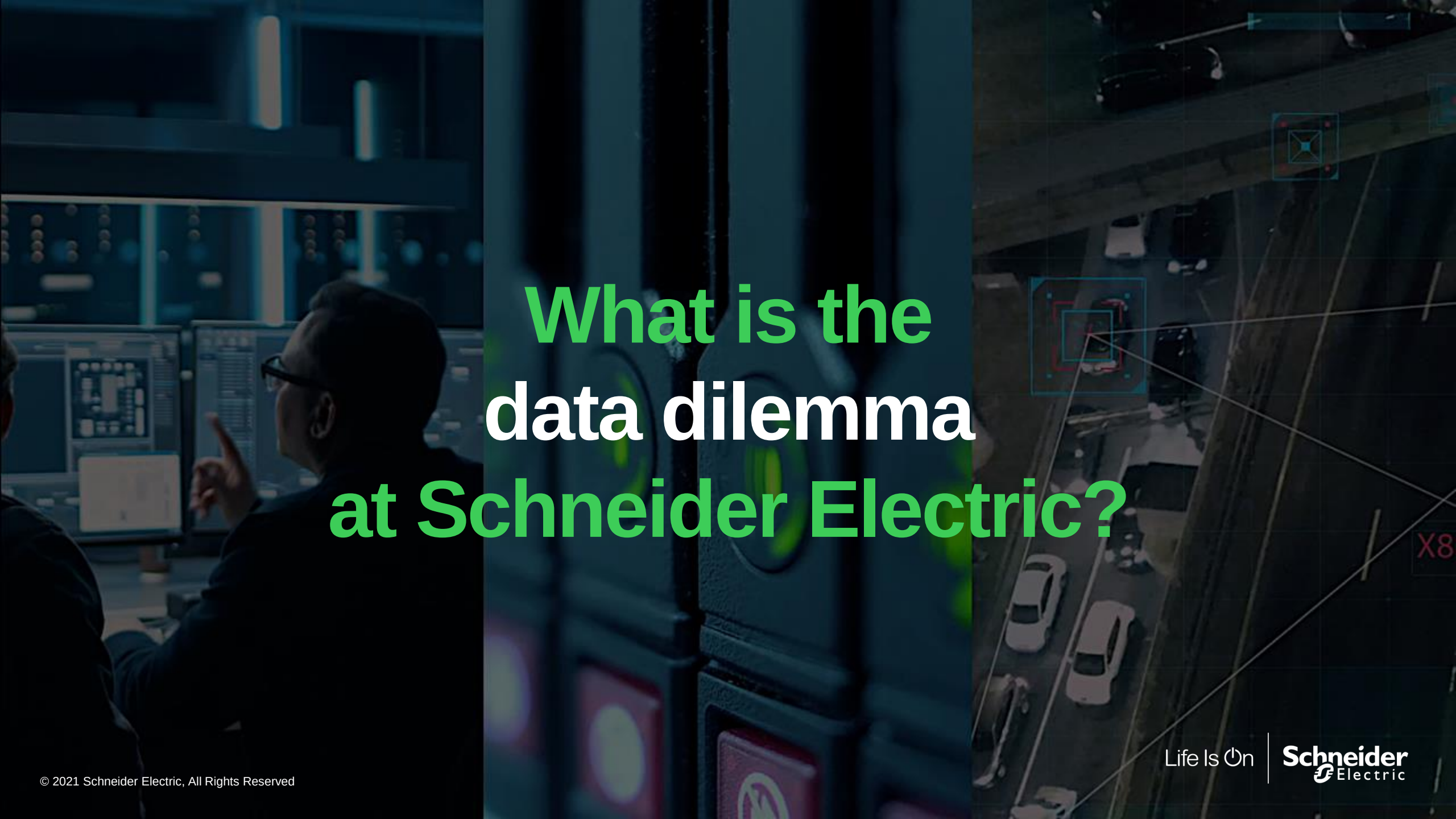


Four markets



Life Is On | Schneider Electric

© 2021 Schneider Electric, All Rights Reserved



What is the data dilemma at Schneider Electric?

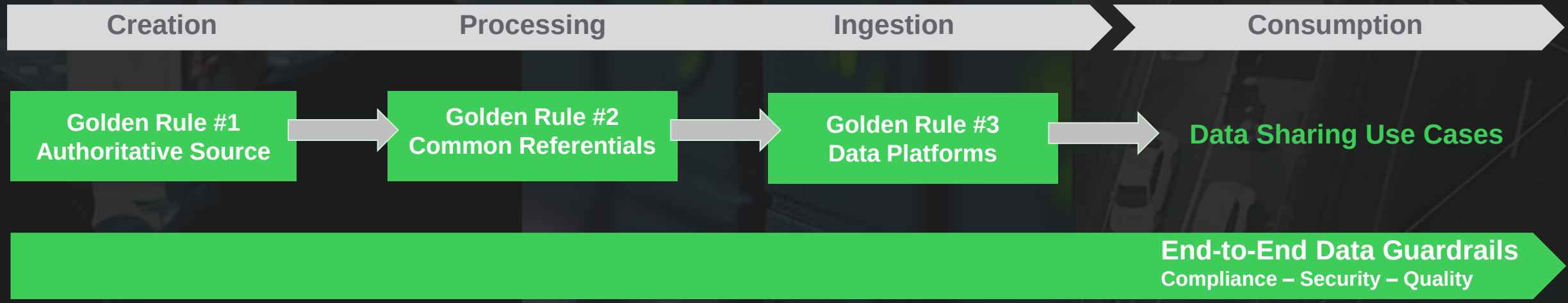
© 2021 Schneider Electric, All Rights Reserved

Life Is On | Schneider Electric

Our data strategy establishes a Decision Supply Chain

Vision	Data-backed execution of strategy compliant w/ regulations
Strategy	Trusted Data @ Scale

Data Supply Chain



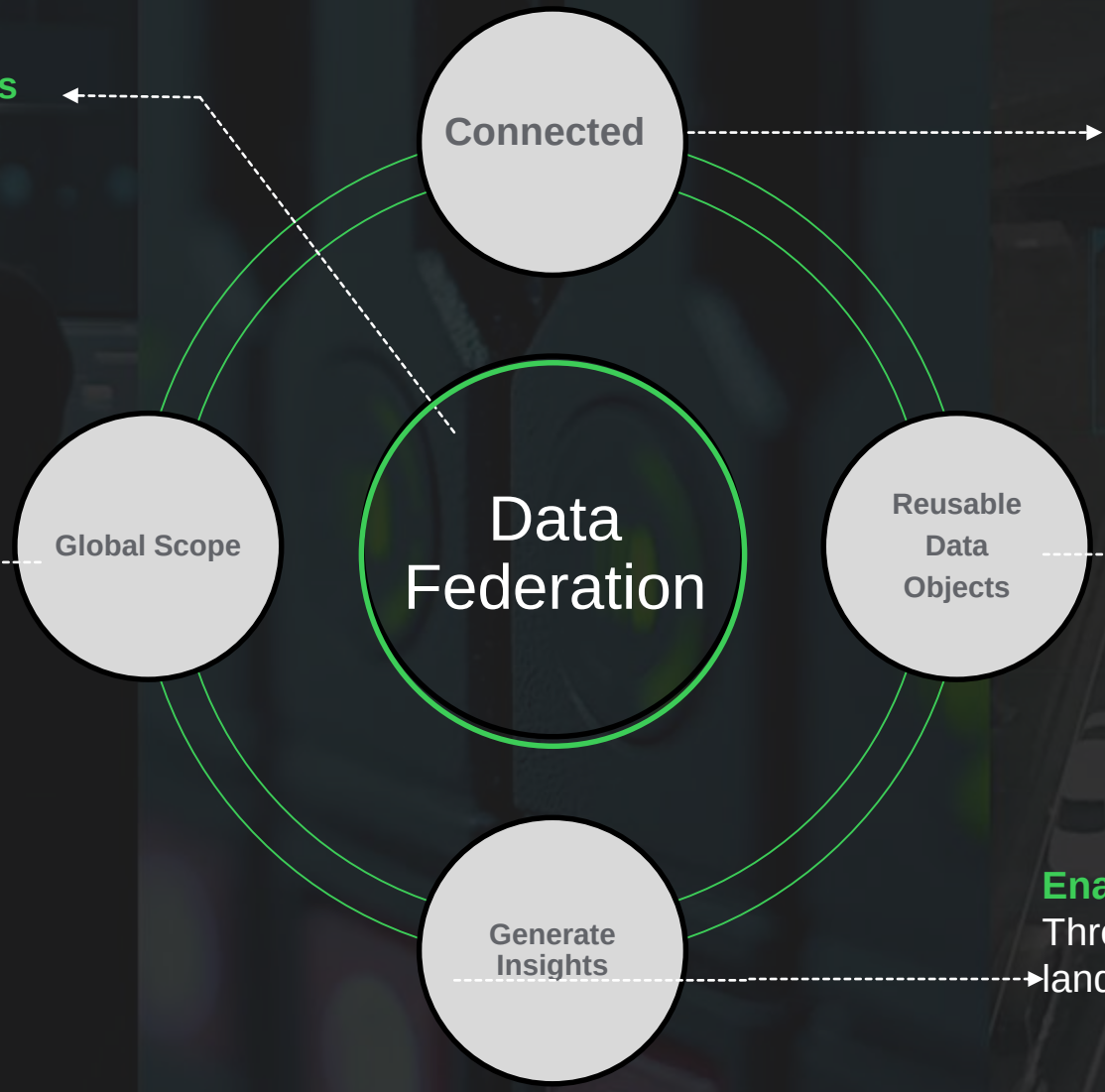
© 2021 Schneider Electric, All Rights Reserved

Life Is On Schneider Electric

Schneider Electric´s Data Store Landscape

Multi-vendor, industry best-in-class technology suite, including latest AWS big data technologies

Multi-domain data & data exchange
Sales, Finance, HR, Web Analytics …



Data extracted from 100+ operational sources (ERPs, CRM, ++transactional tool …)

Accelerates delivery of quality data by unifying global data & preparing it for reuse

Enables data insights
Through a multi-vendor visualization landscape

Life Is On Schneider Electric

© 2021 Schneider Electric, All Rights Reserved

Our journey to Big Data Architecture



2014

Schneider **D**ata **H**ub for Transactional Data

2017

Big Data Architecture POC

2018

Intel Data Store Launch

Hundreds of users

2019

Data Warehouse AI Pipeline

Thousands of users

2021

Data Federation Landscape

Tens of thousands of users

Life Is On Schneider Electric

© 2021 Schneider Electric, All Rights Reserved

Data Store Landscape: Supporting our Decision Supply Chain

Analogy with Product Lifecycle

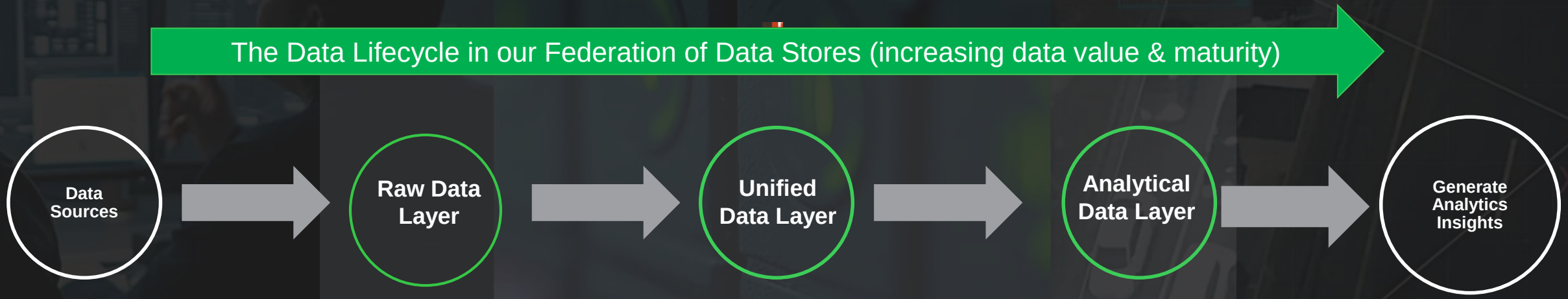
Data FACTORIES
(Production of raw materials)

Data in STOCK
(Assets put in Pallets)

Data PREPARED
(Final Products, labeling & wrap up)

Data in STORES
(Locally distributed to ease their consumption)

Data on SHELVES
(Consumer ready)



© 2021 Schneider Electric, All Rights Reserved

Life Is On | Schneider Electric

Our Technical Advancements in collaboration with AWS

2021

- RA3
- adna
- Concurrency scaling
- Auto-distribution
- Auto-sort key
- Column-encoding

Amazon Redshift

Next

- Data Sharing
- Amazon Redshift Serverless

© 2021 Schneider Electric, All Rights Reserved

Life Is On Schneider Electric

Life Is On | Schneider Electric

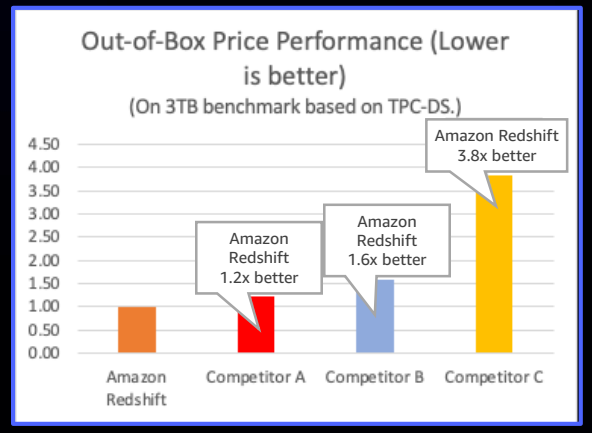


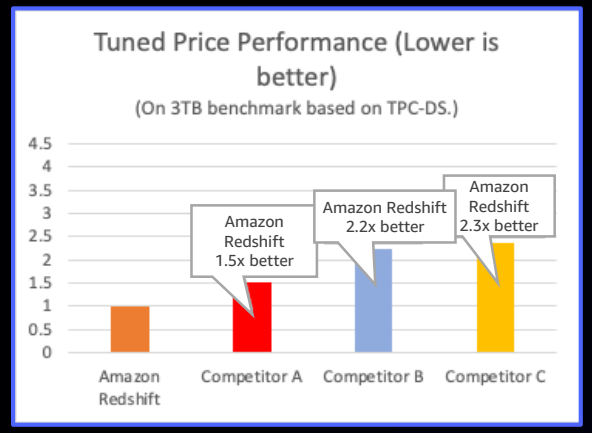
Performance at any scale

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

Performance at any scale

AMAZON REDSHIFT HAS UP TO 3X BETTER PRICE PERFORMANCE THAN OTHER CLOUD DATA WAREHOUSES





Amazon Redshift has up to 3x better price performance than other cloud data warehouses

Scale from GBs to PBs of data with predictable cost and performance

Boost queries up to 10x with AQUA for Amazon Redshift

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

RA3 nodes with managed storage

SCALE COMPUTE AND STORAGE INDEPENDENTLY

Managed storage

Large high-speed cache

High-bandwidth networking

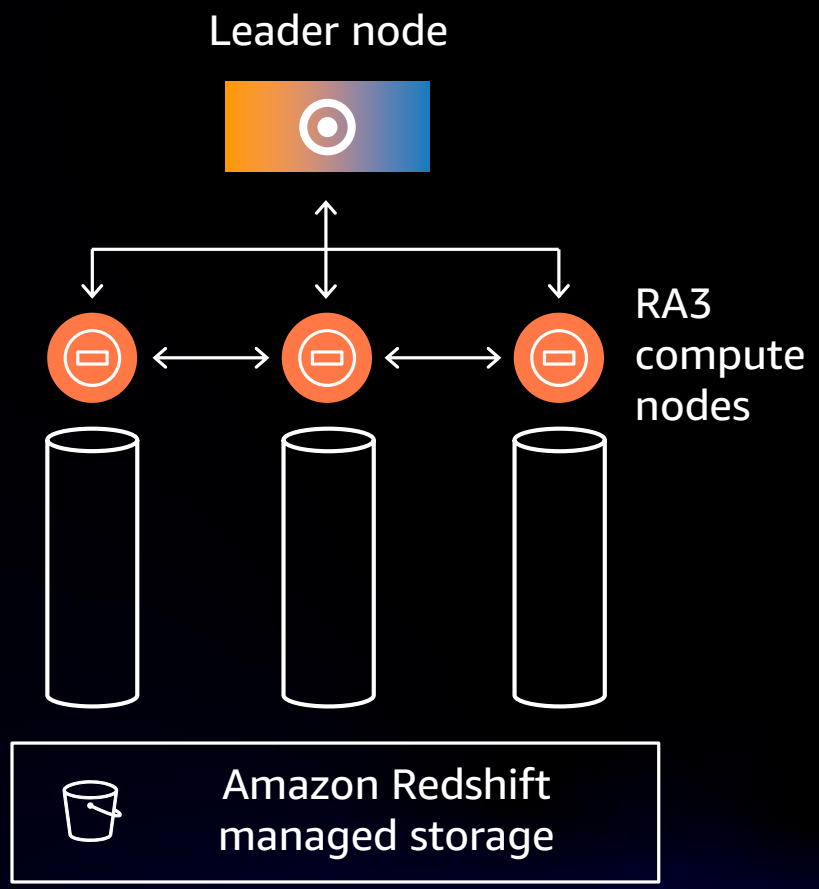
Scale and pay independently for compute and storage

Automatic, no changes to any workflows, no need to manage storage

Available in three denominations for small, medium, and large workloads

NEW RA3.xlplus single node cluster

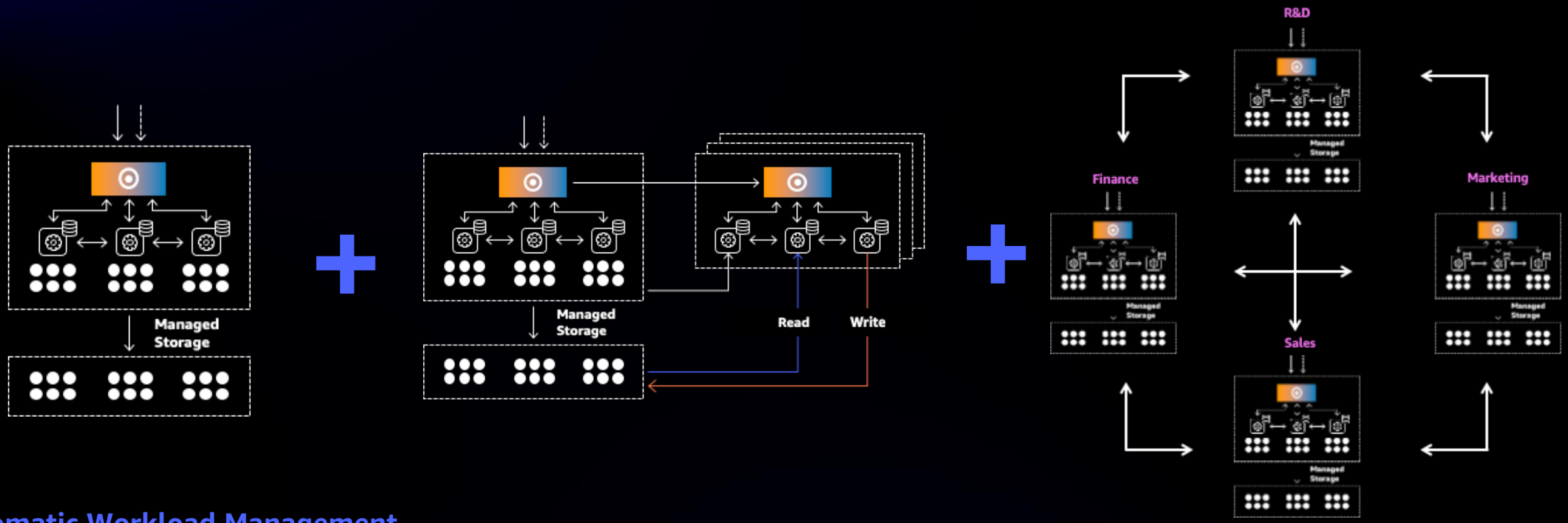
NEW RA3 Reserved Instance (RI) upgrade from DS2s



aws

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

Flexible compute to support your workload



Automatic Workload Management (WLM)

Maximize query throughput in a single cluster

Concurrency Scaling

Auto scale your cluster for high concurrency and spikes

Data Sharing/Workload Isolation

Multi-cluster deployments for workload isolation and data as a service

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

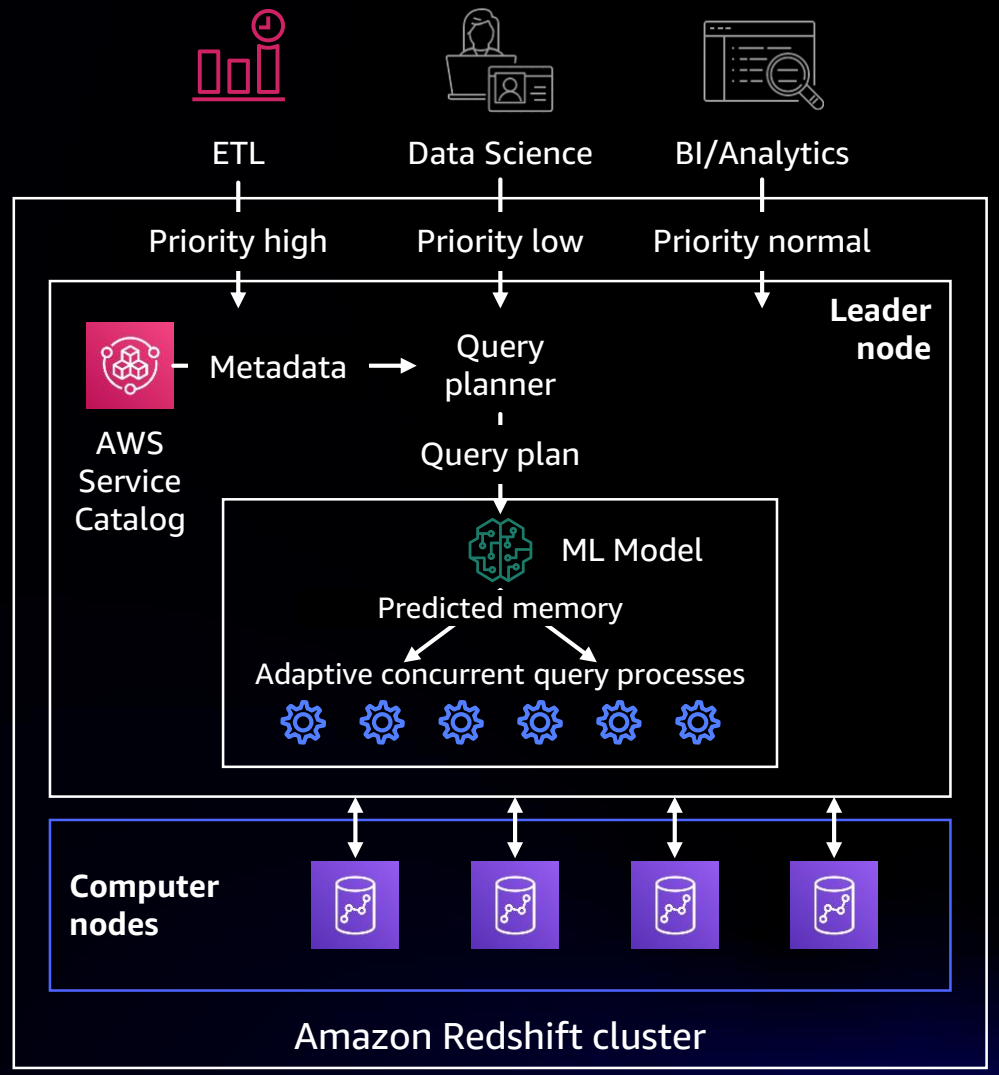
Enhanced auto WLM: Adaptive concurrency

UPDATED

Based on the query traffic and resource utilization, Amazon Redshift automatically determines the number of concurrent queries to **optimize query throughput**

Amazon Redshift's automatic WLM allows you to **create user or query-group queues** based on business function, priority, and SLA

Concurrency scaling based on **business priority and cost** to ensure critical workloads meet their SLA and do not have to wait for main cluster to free up



aws

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

Enhanced concurrency scaling

NEW PREVIEW

WRITE QUERIES SUPPORT

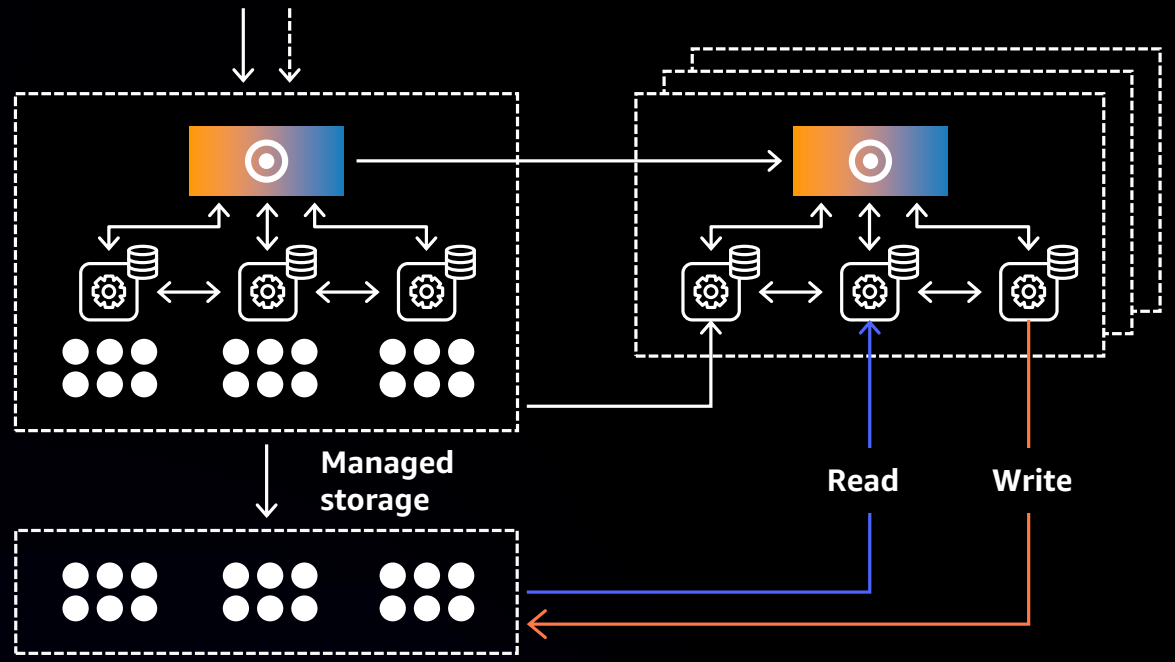
Use case
Auto-scale write queries in addition to read queries for virtually unlimited concurrent users and queries

---

Example write operations include
INSERT, DELETE, UPDATE, COPY

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

Per-second billing and cost controls apply



aws

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

AQUA (Advanced Query Accelerator)

NEW HARDWARE-ACCELERATED CACHE THAT DELIVERS UP TO 10X BETTER QUERY PERFORMANCE THAN OTHER CLOUD DATA WAREHOUSES

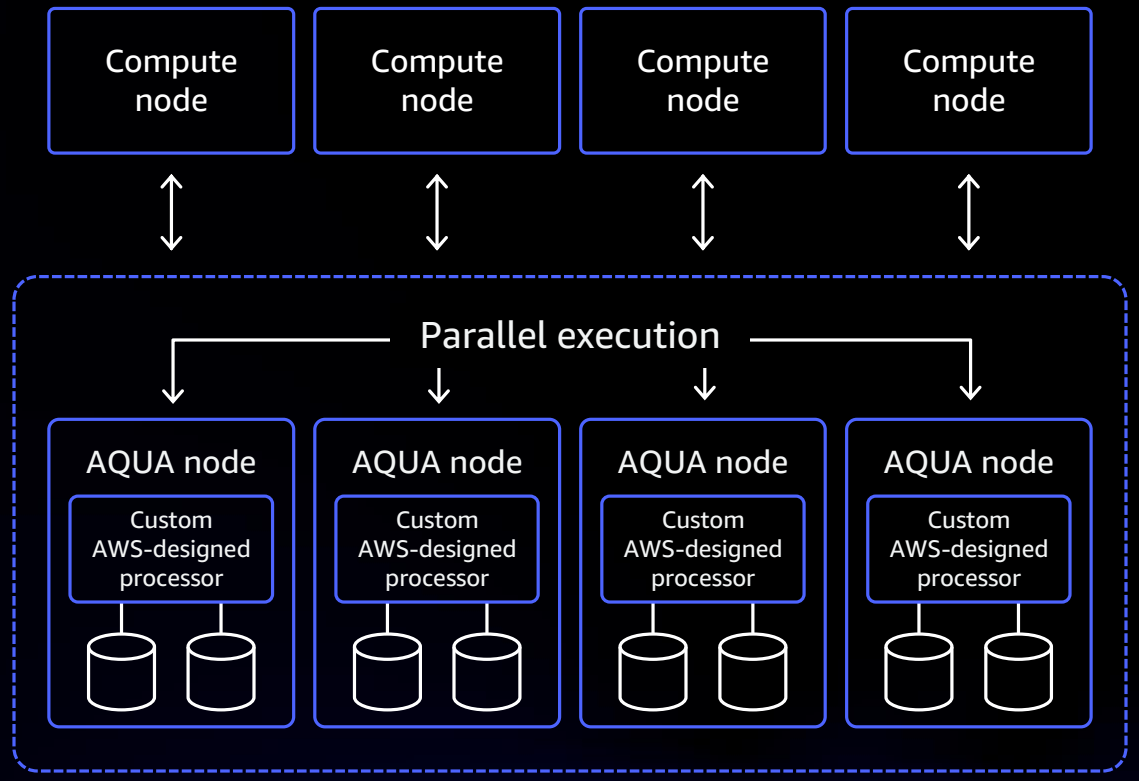
Minimize data movement over the network by pushing down operations to AQUA nodes

---

AQUA nodes with custom AWS-designed analytics processors to make operations (compression, encryption, scans, filtering, and aggregation) faster than traditional CPUs

---

Available with RA3s, no code changes required, no additional cost



FOX accenture SISENSE

aws

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

SQL enhancements & migration support

ACCELERATE MIGRATIONS FROM LEGACY DATA WAREHOUSES

SQL Syntax

- Recursive CTE
- Stored Procedure Enhancements
- HyperLogLog
- PIVOT/UNPIVOT (NEW)
- Case Sensitivity
- Numeric and Textual Casting
- Unload to JSON (NEW)

Types & limits

- 100,000 Tables
  10,000 SPs
- SUPER (semi-structured data)
- Time & TimeTZ
- GEOGRAPHY (NEW)
- VARBYTE (NEW)

Security

- IAM integration
- Azure AD, AD, Okta, Ping Federate
- Column Level Permissions
- Tokenization via Lambda UDFs

Tooling & connectivity

- Open Source Python & JDBC Drivers
- **RSQL** -command line tool with improved control flow and BTEQ support
- SQL Alchemy, Apache Airflow, and QueryBook support

aws

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

Amazon Redshift innovates to meet your needs

Easy analytics for everyone

- NEW! Serverless
- NEW! Query editor v2
- Updated! Automated DW management
- NEW! NEW! Automatic materialized views
- Data API
- Amazon Redshift Advisor
- AWS CloudFormation templates
- NEW! Grafana Plugin

Analyze all your data

- Updated! Data sharing
- NEW! AWS Data Exchange integration
- Updated! Amazon Redshift ML
- Updated! Federated query
- Updated! Geospatial enhancements
- SUPER data type with JSON

Best price performance at any scale

- RA3 nodes & managed storage
- AQUA
- NEW! Concurrency scaling for writes
- Updated! SQL enhancements & migration support
- Security, governance & compliance
- Workload management enhancements

aws

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

Next steps

View related sessions (live or on demand)

Keynote by Adam Selipsky

Keynote by Swami Sivasubramanian

Leadership session by Rahul Pathak

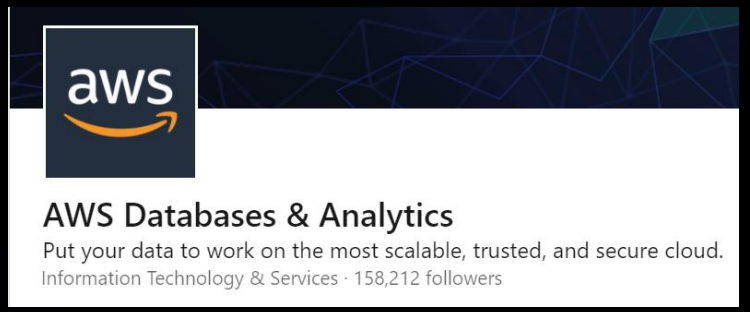
ANT303: Democratizing data for self-service analytics and ML

ANT215: Introduction to AWS Data Exchange for Amazon Redshift

Redshift Serverless

Follow news on LinkedIn

linkedin.com/showcase/aws-databases/



aws.amazon.com/redshift

Please complete the session survey

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

Thank you!

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