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

ANT212

How JPMC and LexisNexis modernize analytics with Amazon Redshift

Anand Singh (he/him)

Executive Director
JPMorgan

Deepak Senthilkumar (he/him)

Director, Software Engineering
LexisNexis

Manan Goel (he/him)

Principal Product Manager
Amazon Redshift

aws

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

Agenda

Business value of data

Why Amazon Redshift

Analytics at scale at JPMorgan Chase & Co. with Amazon Redshift

Modernizing analytics at LexisNexis with Amazon Redshift

Q&A

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

Your data is growing exponentially

More data is created **per hour now** than in an entire year just two decades ago*

*Source: The Seagate Rethink Data Survey, IDC, January 2020

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

This data growth is a challenge and an opportunity



Challenge

Sluggish performance
Security breaches
Cost overruns



Opportunity

Customer experiences
New products & services
Outsmart competition

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

Amazon Redshift

BEST PRICE-PERFORMANCE CLOUD DATA WAREHOUSE DELIVERING POWERFUL, SECURE INSIGHTS ON YOUR DATA AT SCALE



Analyze all your data

Break through data silos and gain real-time and predictive insights on all your data in a few clicks, with no data movement or data transformation



Price-performance at any scale

Delivering best price performance, with performance innovation out of the box at no additional cost to you



Easy, secure, and reliable

Get data to insights in seconds without worrying about infrastructure management, with secure and reliable data warehouse service

aws

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

Tens of thousands of customers analyze exabytes of data every day with Amazon Redshift

Use case	Business intelligence & dashboards | Self-service analytics | Real-time analytics | Data science & machine learning | Data sharing & collaboration

Media & entertainment	EA, zynga, WB, DREAM11, SONY, ROVIO, redhill GAMES, playrix, gumi
Financial and professional services	coinbase, JPMorganChase, FINANCIAL TIMES, intuit, Klarna., LexisNexis, robinhood, Goldman Sachs, ORION
Healthcare	23andMe, AMGEN, Johnson & Johnson, nextgen healthcare, AstraZeneca, MERCK, moderna
Consumer services	iflix, Kimberly-Clark, QANTAS, SCHOLASTIC, SOUNDCLOUD, Sysco, yelp, PELOTON
Web and technology	amazon, CIVIS, coursera, EQUINOX, FOURSQUARE, HEARST, Pinterest, Prezi, ADP

aws

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

Analytics at scale using Amazon Redshift

Anand Singh
Executive Director
JPMorgan Chase & Co.

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

“Scale breaks everything.”



Dr. Werner Vogels
VP and CTO, Amazon.com

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

Operating at scale

JPMorgan Chase & Co.

$33T
in assets under custody

500PB+
Total data stored

40B
Daily security events

60M
active digital customers

>50%
or 66 million US households have a relationship with Chase

>90%
of Fortune 500 companies do business with JPMC

55K+
technologists

6.4K+
applications

$10T
in payments processed daily

Corporate & Investment Bank: Markets business

#1
in markets revenue

100+
markets globally

12.2%
market share

$31.7B
2021 markets and securities services revenue

XIB platform scale

300+
Market feeds

4K+
Daily reports

5M
Trade events per day

$31.7Bn
34k clients

8Bn
Records in management

80TB+
Storage

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

Problem domain and solution architecture



Analytics at scale

AWS lake house architecture
Unified analytics for differentiated insights



Performance

Amazon Redshift RA3 scalable compute & storage with ML based performance optimizations



Reduced TTM

Simplified architecture with ability to leverage rich AWS services ecosystem



Data governance

AWS centralized data lake governance/data mesh



Maintenance

AWS Managed Service allows us to focus on business goals



Cost

Excellent price performance with 36% cost savings realized in year one

aws

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

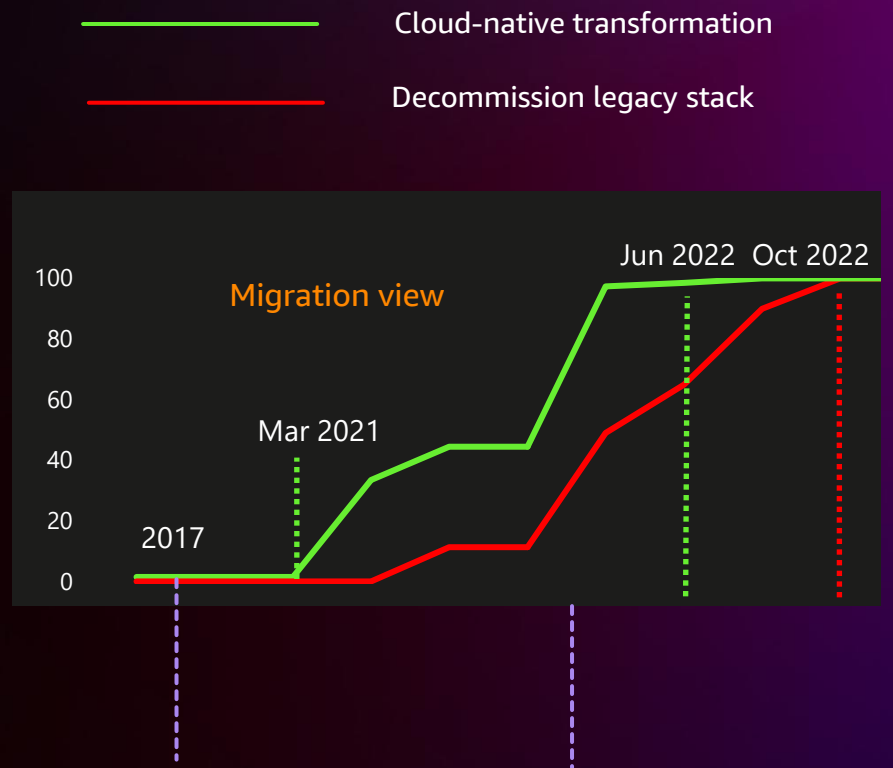
Journey to AWS

1. Self-managed Kubernetes
2. Legacy ETL tooling
3. Legacy on-premises data warehouse
4. Relational database

TO >

Amazon Redshift

Amazon EKS



2017: Challenges faced during lift-shift approach

Compliance restriction led to hybrid deployment
Data aggregation challenges for reports

2021/2022: Rearchitect

Consolidate to ETL and reporting on Amazon Redshift
Collaboration with AWS on evolving Amazon Redshift as a product

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

Achieved outcomes



~4k daily reports, jobs finish 70% faster

80+ TB data processed

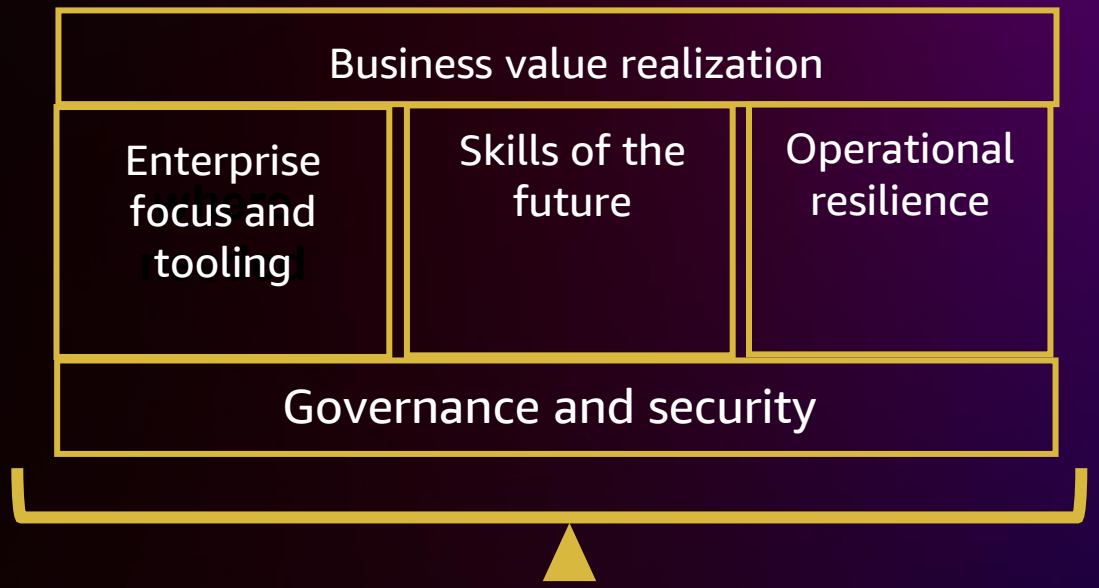
Platform stability up by 34%

36% reduction in infra cost (~1m/yr.)

180% increase in cloud certification and 14 AWS workshops delivered

Enabled data for advanced analytics

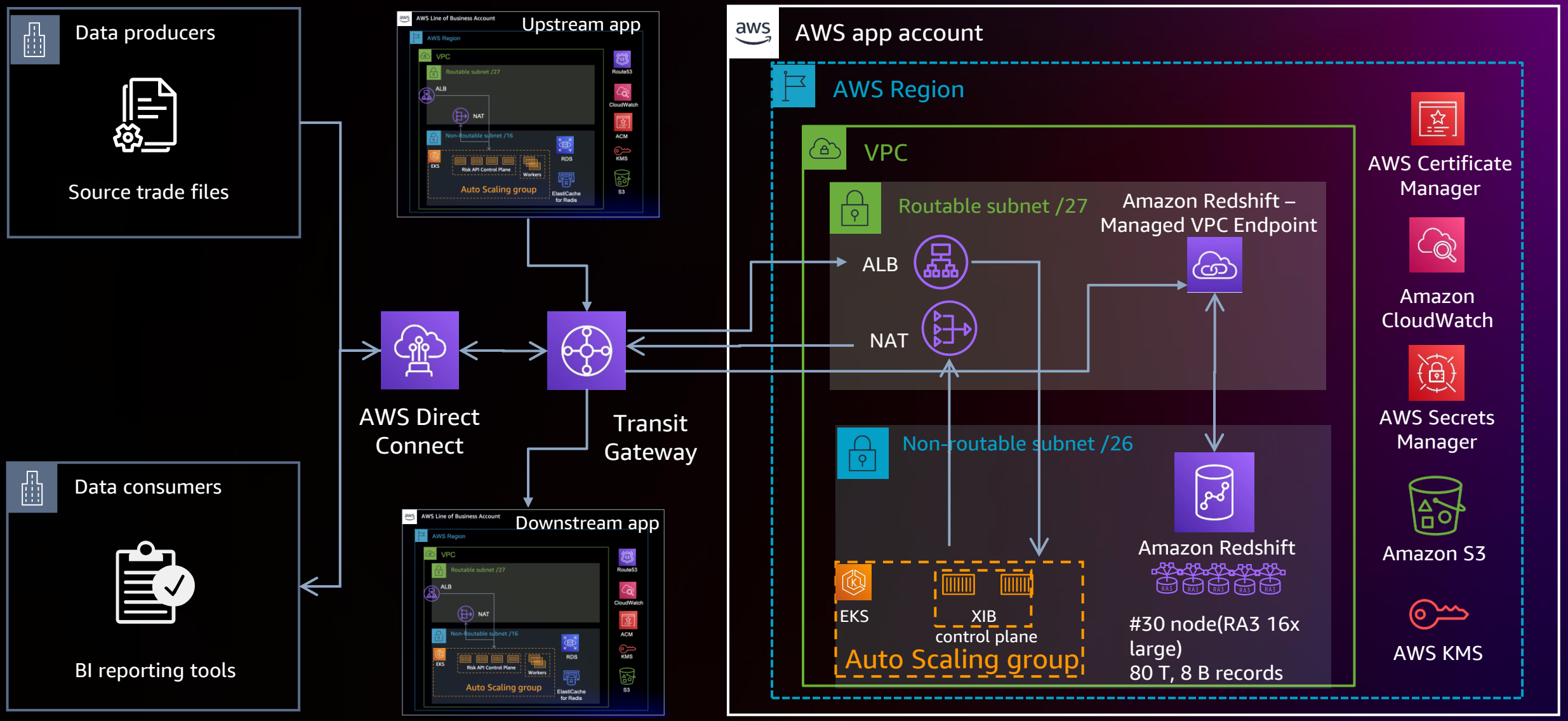
Strategy for large-scale complex migration:



aws

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

XIB AWS architecture



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

Cloud learnings and next steps

Cost improvements

Year RI, LightSwitch, financial discipline

Programmatic integration

Surrogate keys
Global temp table to parallelize
Copy command with static constant
UDF

Future improvements

Serverless
Cross-region relocation
Amazon Redshift ML/Omni AI
Data sharing
Data API



Performance and scalability

SQA for Sales UI
QMR for enterprise reports/Tab Extracts
WLM Auto – Tiering
Concurrency scaling – #30 credit for UI
Autonomics
MV auto-refresh & query rewrite

Operational efficiency

Managed service/autonomics
AWS SCT & multi-part upload for migration
Cluster relocation (resiliency)
Amazon Redshift advisor

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

LexisNexis modernizes analytics using Amazon Redshift

Deepak Senthilkumar (he/him)
Director, Software Engineering
LexisNexis

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

LexisNexis & RELX

130+

Countries and 10K+ employees

65B+

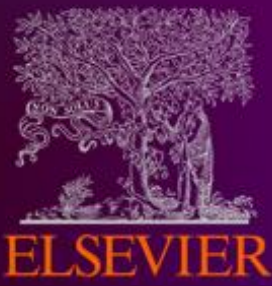
Documents

3

Petabytes of legal news data

150

Times the size of Wikipedia and doubling every three years



SCIENTIFIC, TECHNICAL & MEDICAL



LEGAL



RISK



EXHIBITIONS

aws

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

LexisNexis®

LexisNexis data strategy

The main goal of the data strategy is to establish a trusted global data resource that powers data-driven decision-making for the entire organization, it is also important that we all strive to use a consistent set of technologies.

- **To promote the global sharing of data resources across divisions**
- **To promote high-quality data-driven decision-making across all business units**
- **To develop a more cost-effective data infrastructure**
- **To establish a strong data culture**

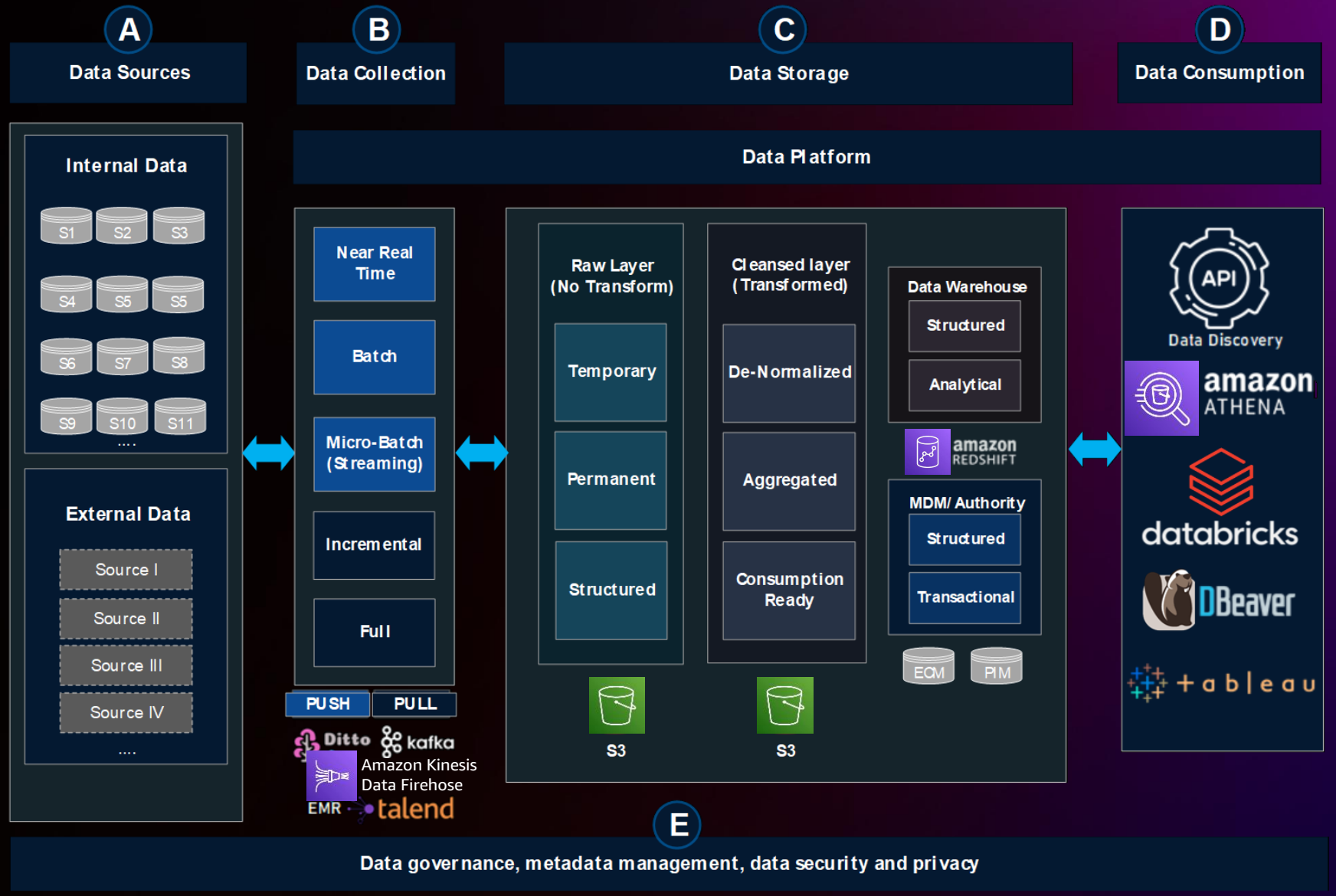
aws

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

LexisNexis®

Data platform architecture

- 50 + data sources
- 2,000+ ETL jobs
- 600+ million records daily
- 3+ TB data processed daily
- 200+ data analysts daily
- 700+ tableau users daily
- 500K transactions per day
- Multiple ML models



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

Journey from EDW to GDW

The on-premises Enterprise Data Warehouse (EDW) is where we store a large volume of data accumulated from a wide range of sources (50+). It acts as the backend for a variety of reports and dashboards and serves ~3,700 end-users across the organization. It was reaching end-of-life.

- **Additional hardware costs + existing software license costs**
- **Existing high maintenance cost with DBAs/product support**
- **Complexities**
  - Architecture: Onboarding new sources/building new features
  - Not scalable due to the fixed on-premises hardware
- **Migrate to a best-of-breed, low-cost, high-performing cloud data warehouse solution**
- **Evaluated multiple vendors including Amazon Redshift**
- **Selected Amazon Redshift in Q4 2020 based on multiple POCs and detailed analysis**

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

Amazon Redshift migration: Benefits and opportunities

Cost savings

- Less license
- Less operation
- Less maintenance
- ~0.2 x Cost of the on-premises cost

Simplify architecture

- Moving to a simplified architecture: Extract + Load + Transform source data is replicated as-is into S3 and then to Amazon Redshift
- Transformation happens in Amazon Redshift
- Quicker onboarding of new sources using a standard framework

Consolidate systems

- Creation of One Global Data Warehouse which comprises of
- Enterprise data warehouse
- UK/CEMEA/PAC warehouse
- Marketing warehouse, quality warehouse

Customer experience

- 3x performance with easier integration within cloud ecosystem
- Faster data refreshes (more frequent loads). 8x improvement
- <40% errors, better quality, all required data is in one place to query
- With all data in one place, helps us get to 360° view of the customer

aws

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

LexisNexis®

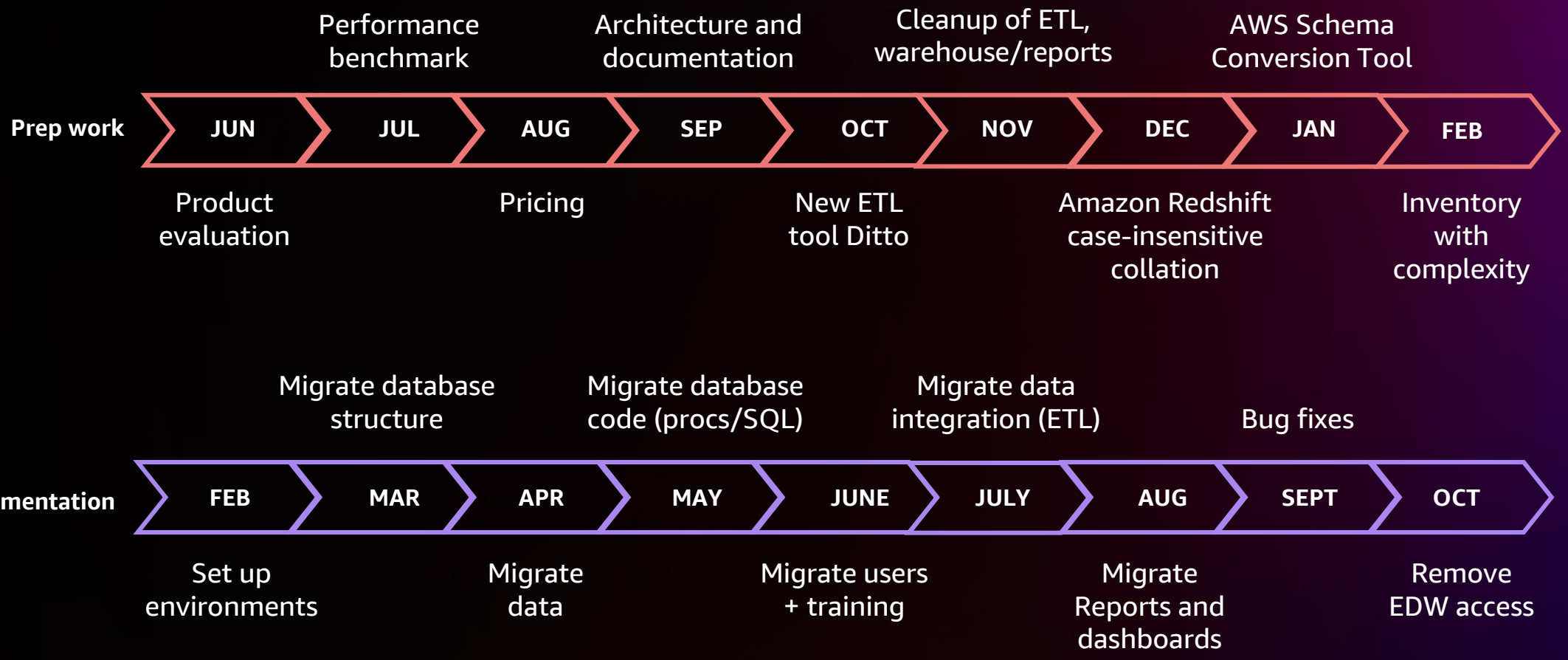
“If I had six hours to chop down a tree, I’d spend the first four hours sharpening the axe.”

Abraham Lincoln

Former US President

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

Amazon Redshift migration:
Timeline (2020–2021)



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

Amazon Redshift migration: Approach

- **Communication/touch-point with end users**
  - User-group kick-off meetings | Countdown timers | Email/MS Teams
  - Office Hours | Migration support hours | ServiceNow tickets
  - Contests with rewards – Bug Bounty/Give up EDW IDs
- **Tools/process**
  - Kanban using Trello board
  - DBeaver-customized open-source version with 20+ features/bug fixes
- **Training**
  - Customized courses (4 hours/90 min) based on persona + materials
- **Engineering – created sub-teams + leveraged AWS ProServe**
- **Status tracking – daily for team + monthly for leadership**





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

Amazon Redshift migration: Learnings + future

- **Longer runway**
- **More emphasis on testing/QA (quality assurance)**
- **Better understanding of Amazon Redshift nuances around**
  - Time zone/timestamp, rounding (banker's rounding especially), serializable isolation
- **All's well that ends well**
  - With the successful completion of the project, we have trained 275+ users, migrated 400+ users, 1400+ BI Reports, 1500+ ETL jobs, 3000+ procedures, 4500+ views, 14500+ tables, and 230+ billion rows of data!
  - It's continued to grow now with over 650+ users
- **Future: Row-level security + serverless + data share**

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

Get started today

New to Amazon Redshift?

- ✓ IDC study on Business value of Amazon Redshift: https://bit.ly/3i36S1i
- ✓ Explore demos, customer stories, and the latest features on aws.amazon.com/redshift
- ✓ Get trained: https://bit.ly/3i36YWI
- ✓ 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: https://go.aws/3AGuBuO

Get help: https://bit.ly/3GEkPx5

aws.amazon.com/redshift

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

Thank you!

Anand Singh
LinkedIn

Deepak Senthilkumar
LinkedIn

Manan Goel
LinkedIn



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

aws

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