AWS re:Invent

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

STP210

Designing the right data strategy for your startup

Julia Soscia (she/her)
Startup Solutions Architect Manager
Amazon Web Services

Jason Hughes (he/him)
Product Director
Dremio

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

Agenda

1. Business goals
2. Data strategy challenges
3. Build vs. buy – which is right for you?
4. How Dremio provides interactive BI directly on Amazon S3

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

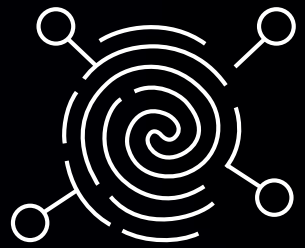
Business goals

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

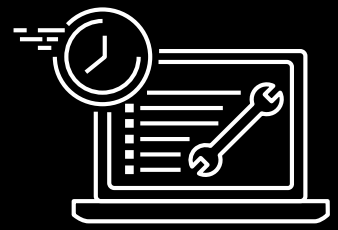
Business goals



Improve operational efficiency



Make more informed decisions



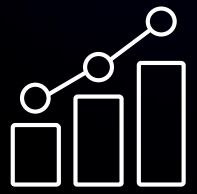
Accelerate innovation

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

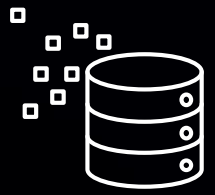
Data strategy challenges

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

Data strategy challenges



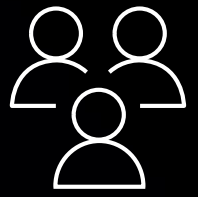
Growing exponentially



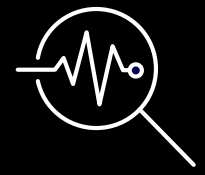
From new sources



Increasingly diverse



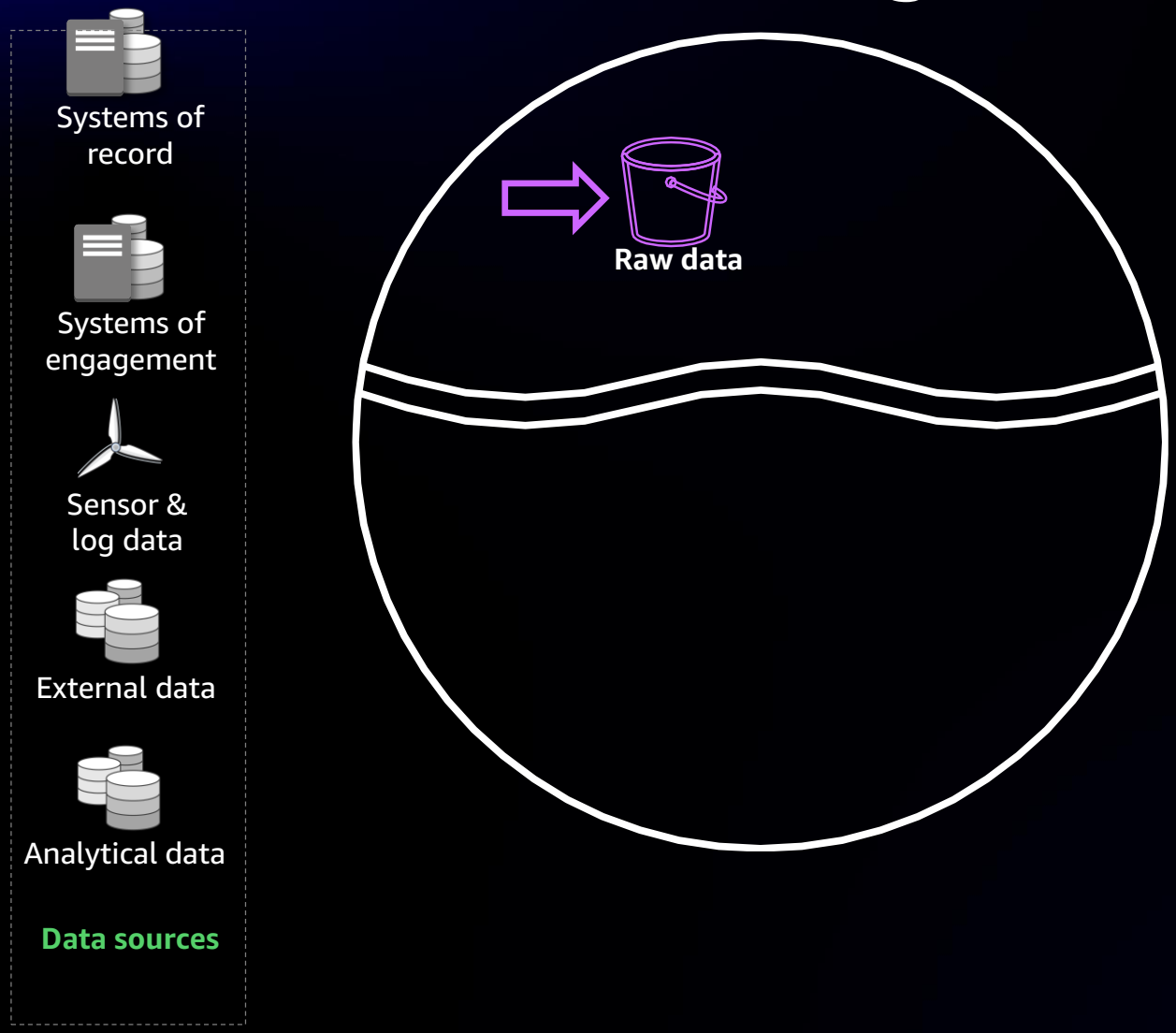
Used by many people



Analytics and machine learning

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

Technical challenges



Raw data – "as-is"

- All source system data for all time
- Structured, semi-structured and non-structured
- No guarantee of data quality
- Schema mismatch between sources

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

Technical challenges





Raw data – "as-is"

- All source system data for all time
- Structured, semi-structured and non-structured
- No guarantee of data quality
- Schema mismatch between sources

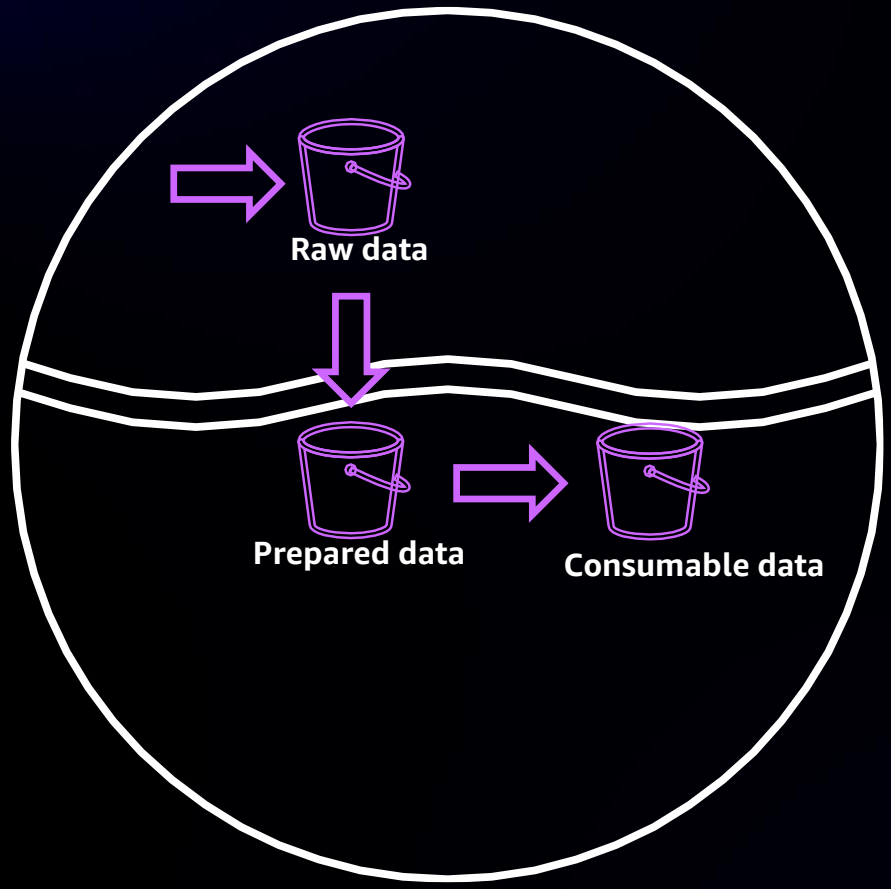
Prepared Data – "integrated"

- Trustworthy
- Easy to use, frictionless
- Business atomic

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

Technical challenges





Raw data – "as-is"

- All source system data for all time
- Structured, semi-structured and non-structured
- No guarantee of data quality
- Schema mismatch between sources

Consumable data – "as-needed"

- Laser focused, application specific
- Right format, granularity
- Fits consumer's technology

Prepared Data – "integrated"

- Trustworthy
- Easy to use, frictionless
- Business atomic

aws

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

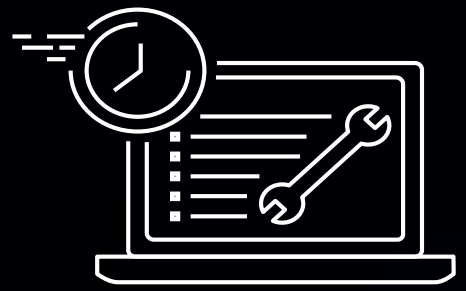
Build vs. buy use cases

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

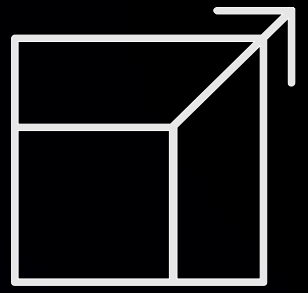
Build vs. buy considerations



Team size and skill



Timeline and go-to-market



Flexibility



Budget

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

Amazon S3-centric ETL pipeline – Data lake

Unmatched durability, availability, and scalability

Many object-level controls



Easy to use with cost optimization: intelligent tiering

Broad portfolio of analytics tools

Many ways to get data in

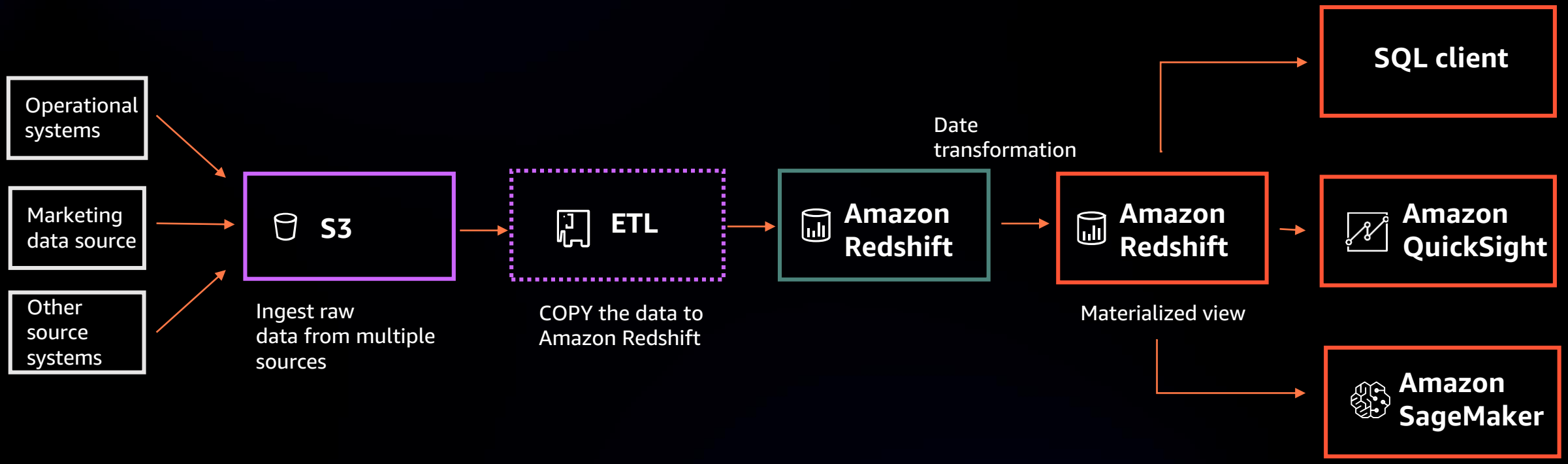
Unmatched security, compliance, and audit capabilities

Cold storage and archive capabilities

aws

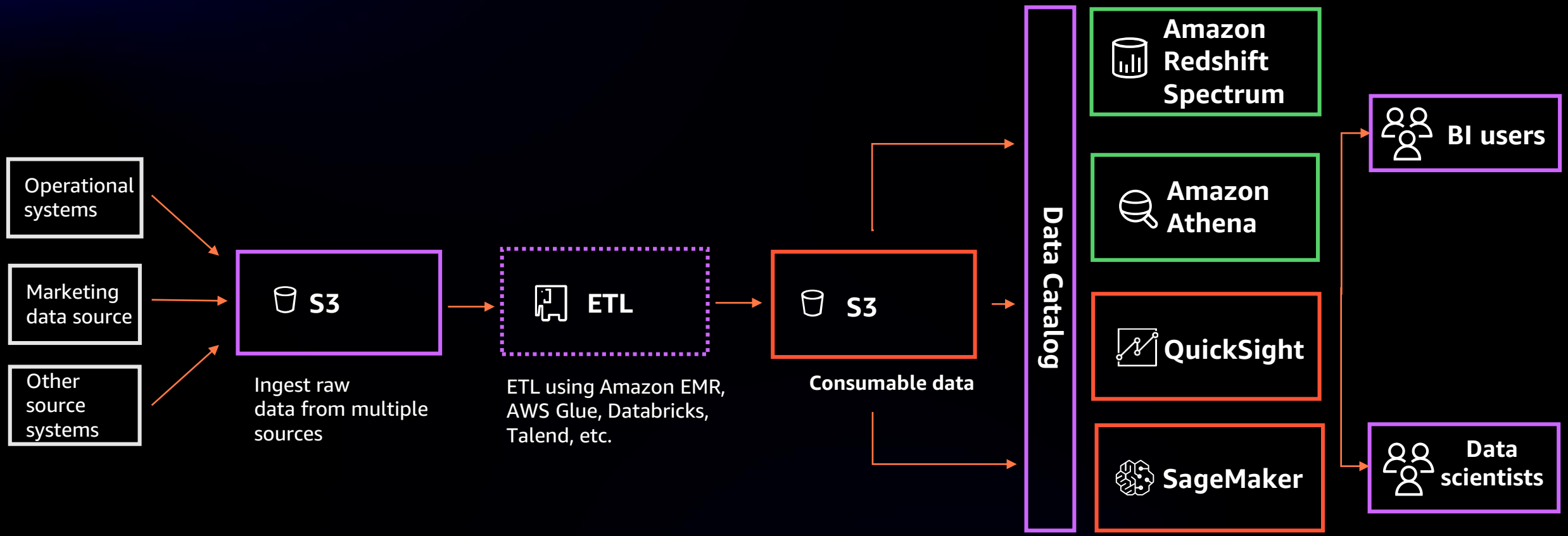
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Build – An Amazon Redshift–centric ELT pipeline



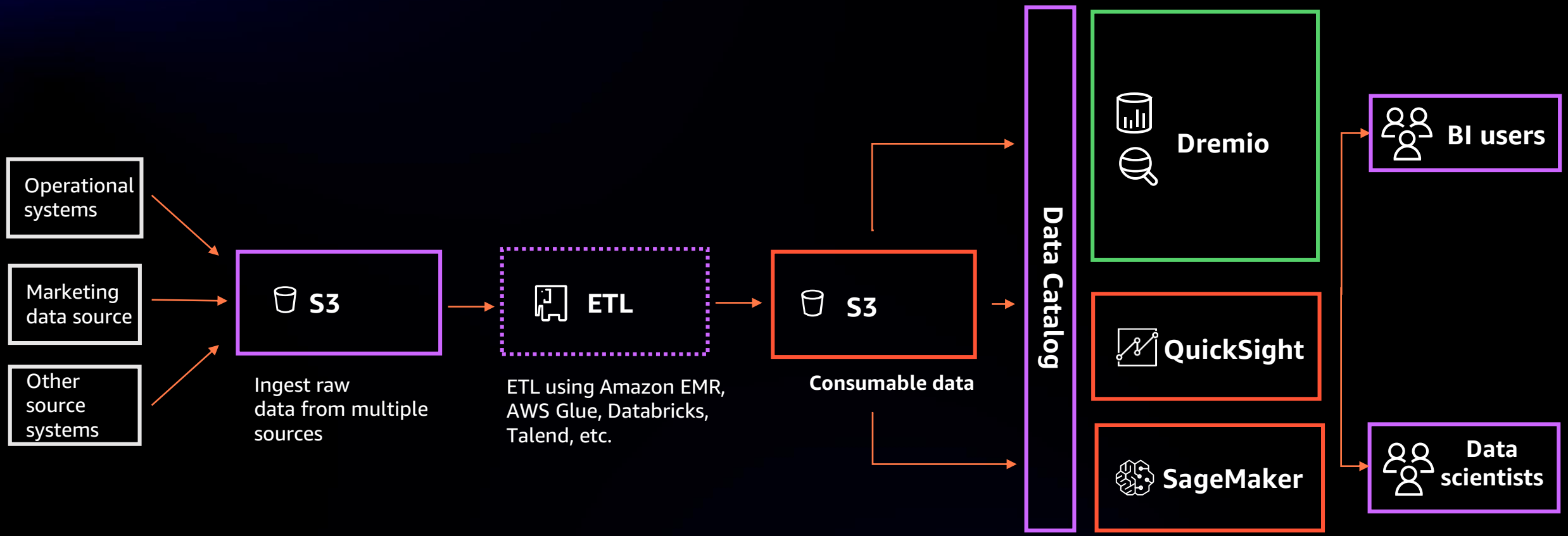
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Build – An S3-centric ETL pipeline



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

Buy – Dremio

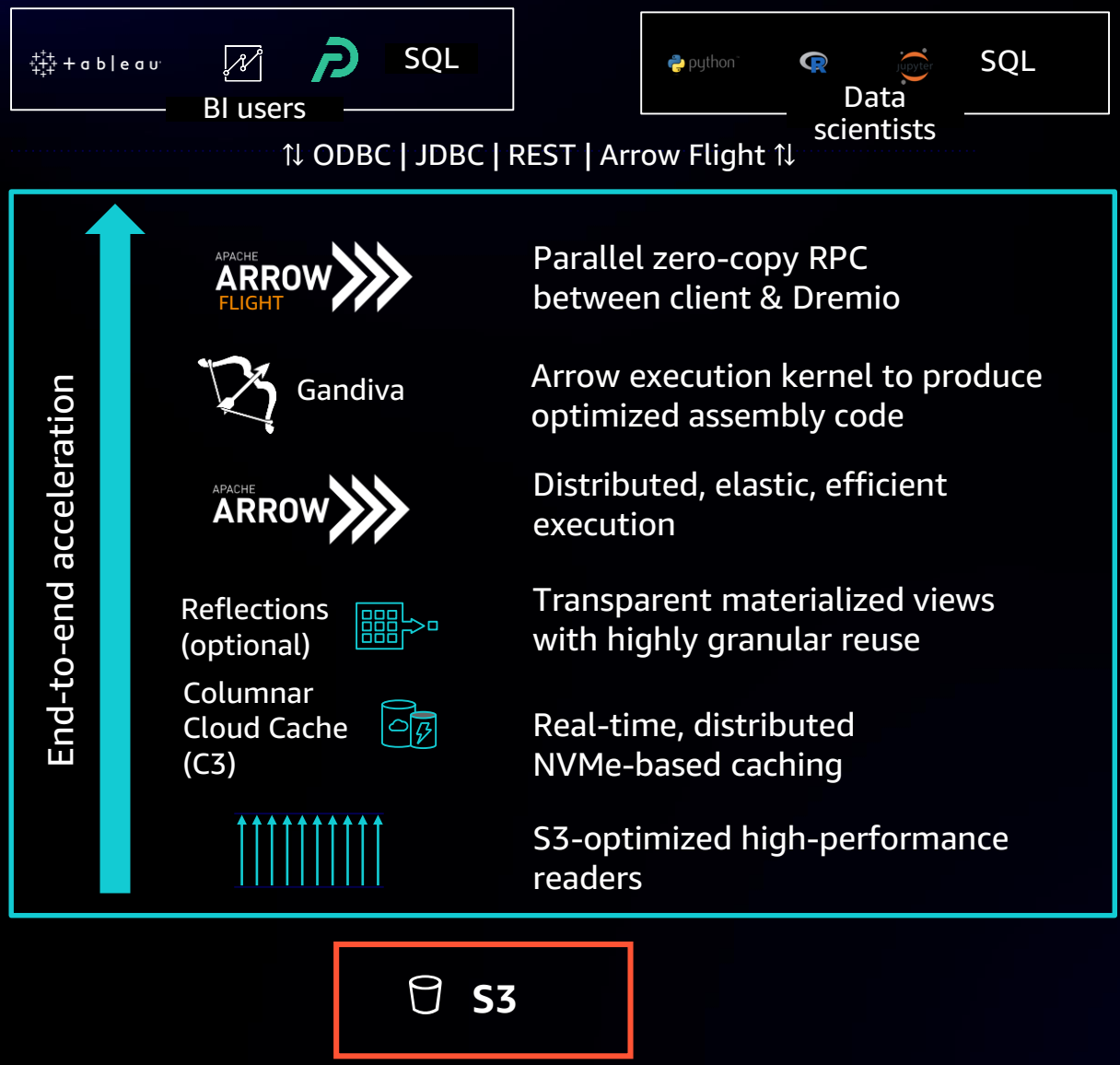


© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How Dremio provides interactive BI directly on Amazon S3

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

Interactive BI on S3 via end-to-end acceleration



- Lightning-fast, interactive queries directly on Amazon S3
- 4x faster one-time (or ad-hoc) queries, and up to 1000x faster for BI vs. a comparable query engine
- Highly cost-efficient
- Elastic scale of the cloud

The R logo is © 2016 The R Foundation / CC-BY-SA 4.0

aws

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

Thank you!

Julia Soscia
soscj@amazon.com

Jason Hughes
jason@dremio.com

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