AWS re:Invent

ANT237-R

Zero to hero: Quick starts with Amazon Redshift

Matt Scaer

Principal DW Specialist Solutions Architect

Amazon Web Services

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Refresher on Amazon Redshift

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws

Amazon Redshift architecture

Massively parallel, shared-nothing columnar architecture

Leader node

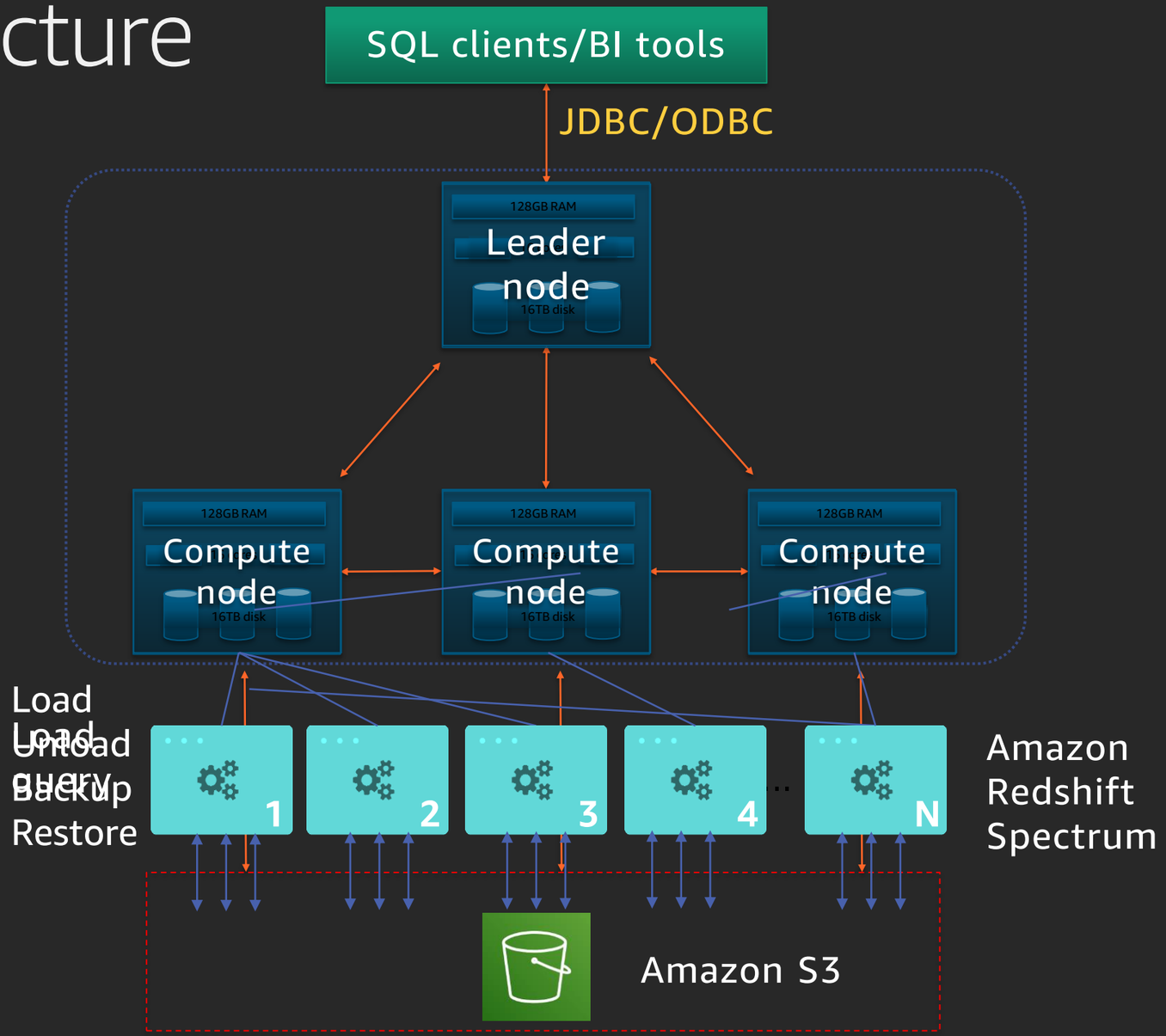
- SQL endpoint
- Stores metadata
- Coordinates parallel SQL processing

Compute nodes

- Local, columnar storage
- Executes queries in parallel
- Load, unload, backup, restore

Amazon Redshift Spectrum nodes

- Execute queries directly against Amazon Simple Storage Service (Amazon S3)



Amazon Redshift Advisor: Your DBA's best friend

- Amazon Redshift expert system available in AWS Management Console
- Identifies undesirable end user behaviors for resolution by providing high-impact recommendations to improve performance and reduce cost
- > 96% of clusters have tailored feedback
- Actionable workload management, performance improvement COPY, storage, and system maintenance feedback



5 points of guidance for Amazon Redshift (SET DW)

Sort key (to improve filter performance): Choose up to ~3 columns (compound sort key)

Encoding of columns: Compress all columns **except** for the first sort key column

Table maintenance: Auto VACUUM DELETE and ANALYZE; VACUUM SORT weekly

Distribution key (to improve join performance):

- Follows common join pattern for the table and evenly distributes the data
- DISTSTYLE AUTO is a great go-to for all tables

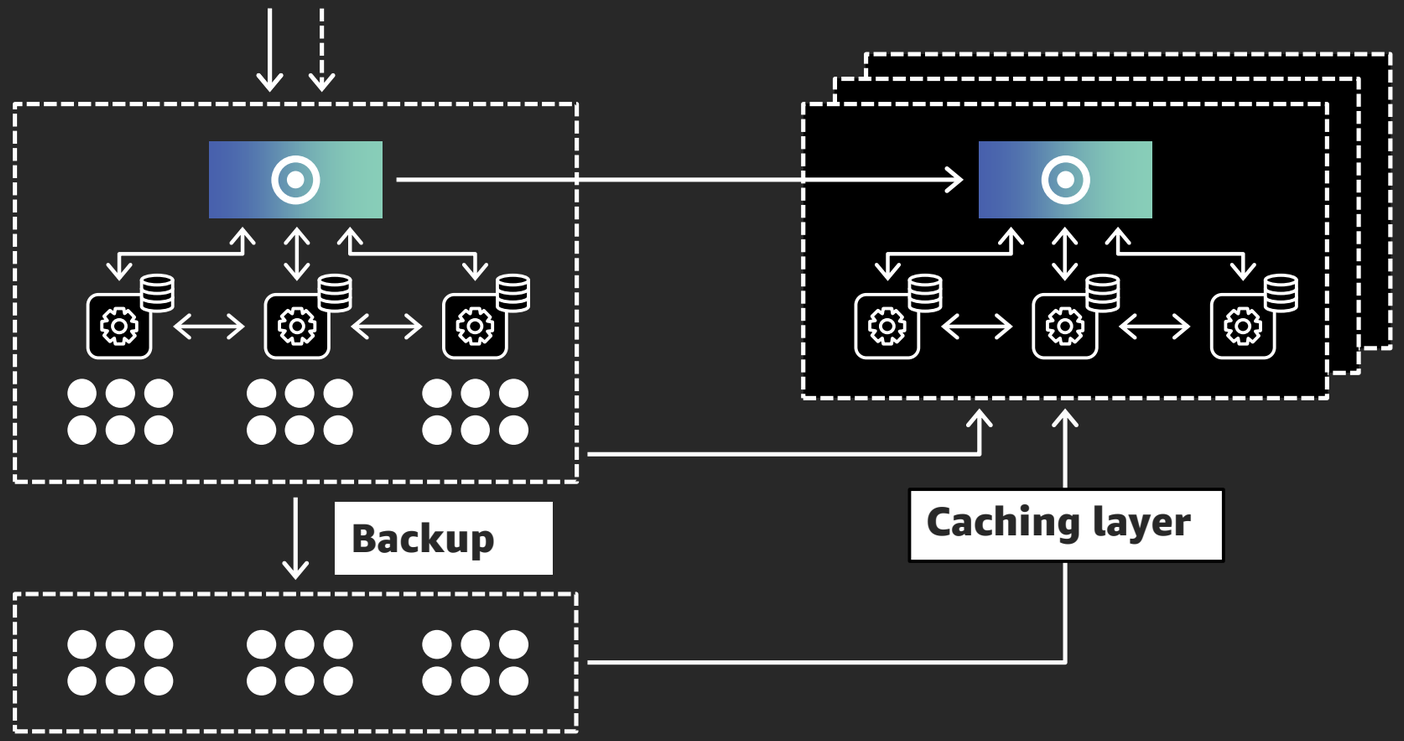
Workload management (WLM) and **query monitoring rules** (QMRs):

- Start with AUTO WLM or with defining up to ~3 queues
- Use QMRs

Concurrency Scaling

GA
March 21, 2019

Amazon Redshift automatically adds transient clusters, in seconds, to serve sudden spikes in concurrent requests with consistently fast performance



For every 24 hours that your main cluster is in use, you accrue a one-hour credit for Concurrency Scaling—this means that Concurrency Scaling is free for > 97% of customers

How it works:

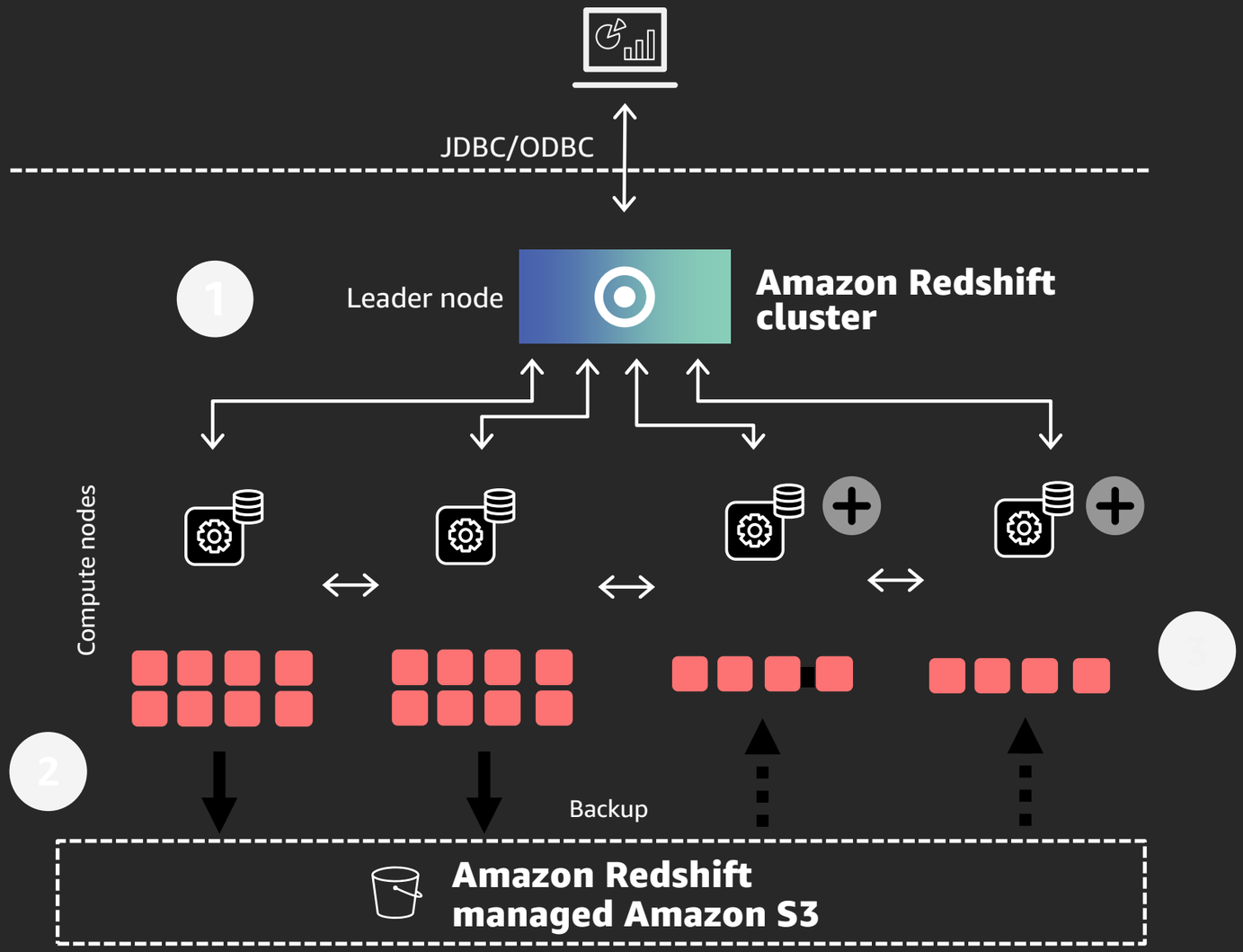
1. All queries go to the leader node; user only sees less waiting for queries
2. When queries in the designated WLM queue begin queuing, Amazon Redshift automatically routes them to the new clusters, also enabling Concurrency Scaling automatically
3. Amazon Redshift automatically spins up a new cluster, processes waiting queries, and automatically shuts down the Concurrency Scaling cluster

Amazon Redshift elastic resize

Scale your Amazon Redshift clusters up and down in minutes to get **the optimal performance**

How it works:

- Amazon Redshift updates the snapshot in Amazon S3 with the most recent data
- New nodes are added (for scaling up) or removed (for scaling down) during this period
- The cluster is fully available for read and write queries; queries that were being held are queued for execution automatically



Amazon Redshift Spectrum

Extend the data warehouse to exabytes of data in an Amazon S3 data lake

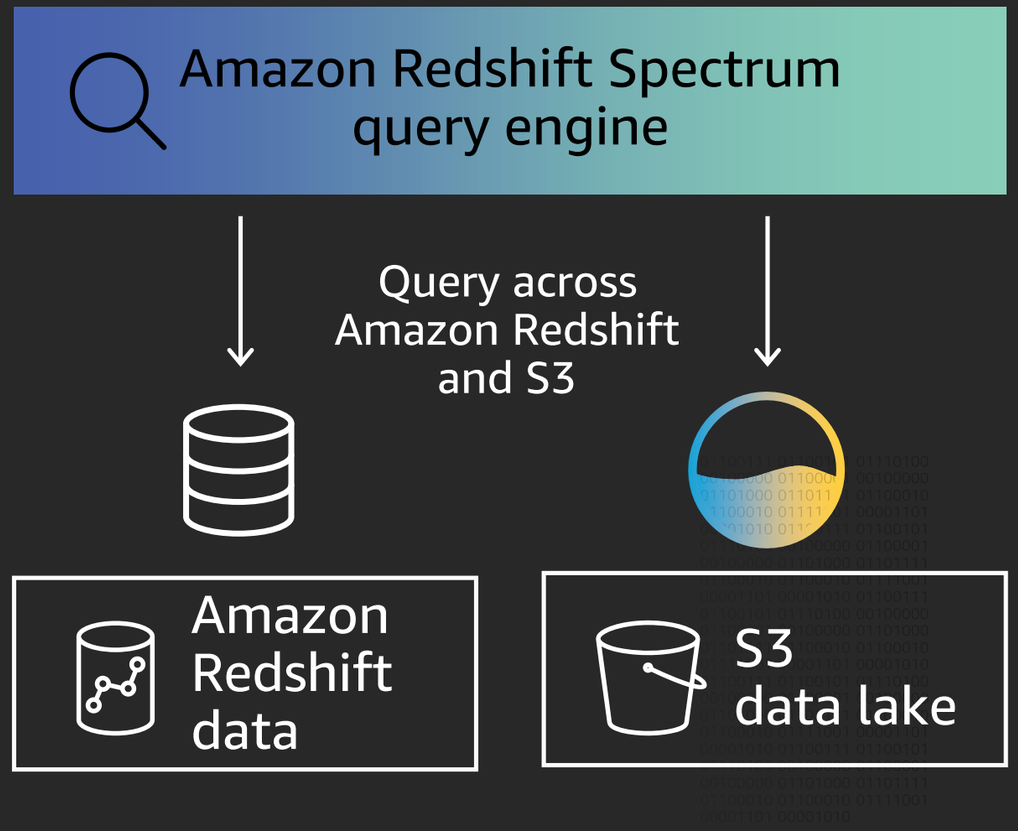
No data loading required

Scale compute and storage separately

Directly query data stored in Amazon S3

Parquet, ORC, Avro, JSON, and CSV data formats

Spectrum Request Accelerator

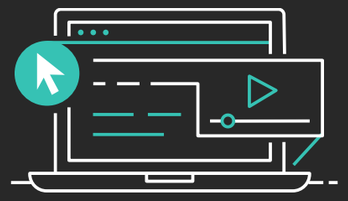


Workshop guide https://tinyurl.com/ant237-2019

1. Log in with the provided workshop credentials at https://dashboard.eventengine.run/login
2. Follow the workshop guide

Learn big data with AWS Training and Certification

Resources created by the experts at AWS to help you build and validate data analytics skills



New free digital course, Data Analytics Fundamentals, introduces Amazon S3, Amazon Kinesis, Amazon EMR, AWS Glue, and Amazon Redshift



Classroom offerings, including Big Data on AWS, feature AWS expert instructors and hands-on labs



Validate expertise with the **AWS Certified Big Data - Specialty** exam or the new **AWS Certified Data Analytics - Specialty** beta exam

Visit aws.amazon.com/training/paths-specialty/

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws training and certification

Thank you!

Matt Scaer

mscaer@amazon.com

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Please complete the session survey in the mobile app.

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.