Amazon Redshift & Spectrum

Big Data Immersion Day

November, 2018

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

AWS Databases and Analytics

Business Intelligence & Machine Learning

QuickSight | SageMaker | Comprehend/Polly…

Relational Databases

Aurora

RDS

Non-Relational Databases

DynamoDB

ElastiCache
(Redis, Memcached)

Neptune
(Graph)

Analytics

DW | Big Data Processing | Interactive

Redshift | EMR | Athena

Real-time

Elasticsearch Service | Kinesis Analytics

Data Lake

S3/Glacier | Glue
(ETL & Data Catalog)

Data Movement

Database Migration Service | Snowball | Snowmobile | Kinesis Data Firehose | Kinesis Data Streams

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

aws

Amazon Redshift

더 빠르고
더 간단하고
더 싸게

관계형 데이터 웨어하우스

대용량 병렬 처리 - 페타 바이트급

관리형 서비스

$1,000/TB/year

starts at $0.25/hour

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

aws

Amazon Redshift 주요 고객

NTT Docomo | Telecom

FINRA | Financial Svcs

Philips | Healthcare

Yelp | Technology

NASDAQ | Financial Svcs

The Weather Company | Media

Nokia | Telecom

Pinterest | Technology

Foursquare | Technology

Coursera | Education

Coinbase | Bitcoin

Amazon | E-Commerce

Etix | Entertainment

Spuul | Entertainment

Vivaki | Ad Tech


Z2 | Gaming

Neustar | Ad Tech

SoundCloud | Technology

BeachMint | E-Commerce

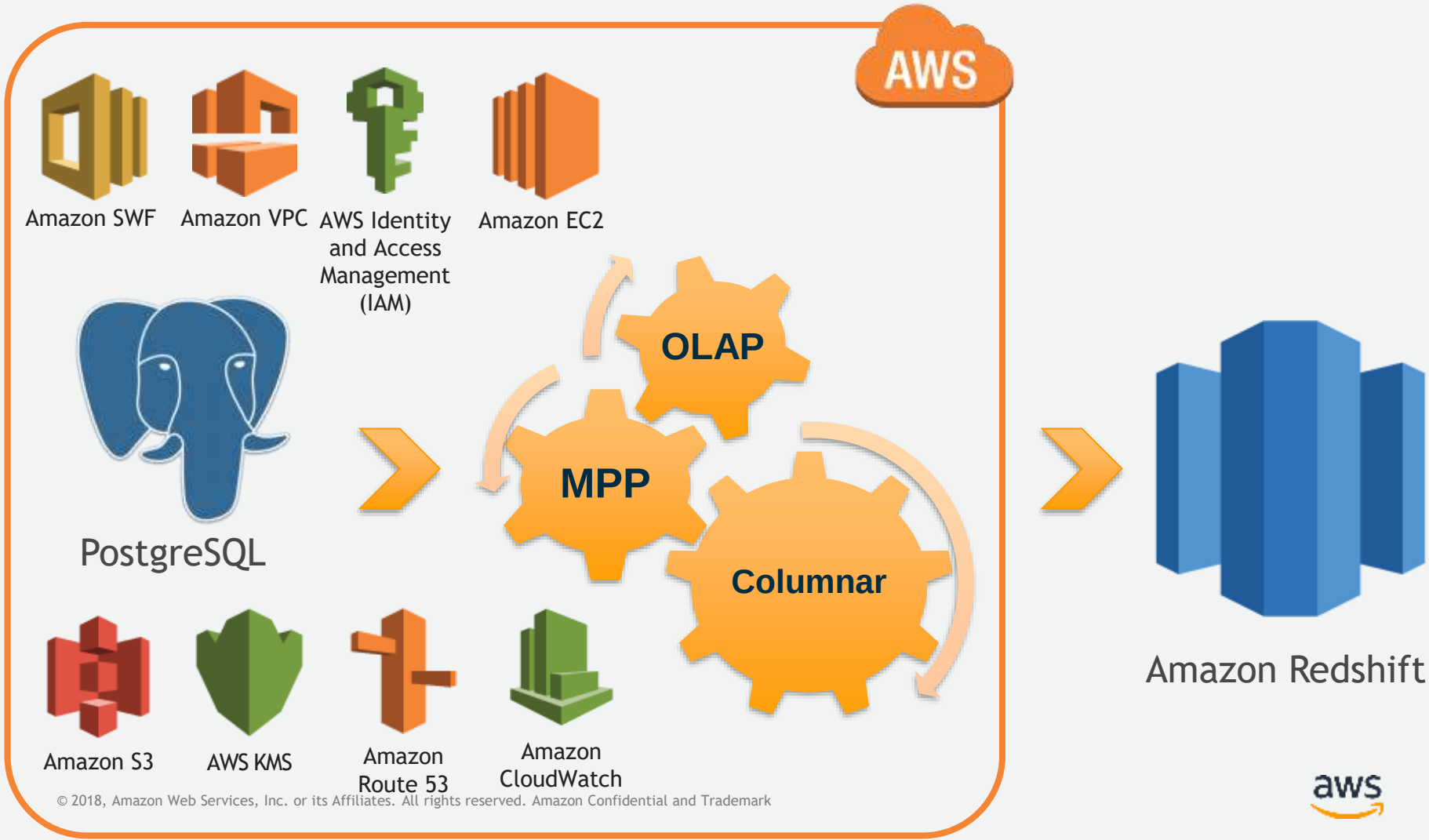
Civis | Technology

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

aws

Redshift

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark



© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

Amazon Redshift 아키텍처

리더(Leader) 노드

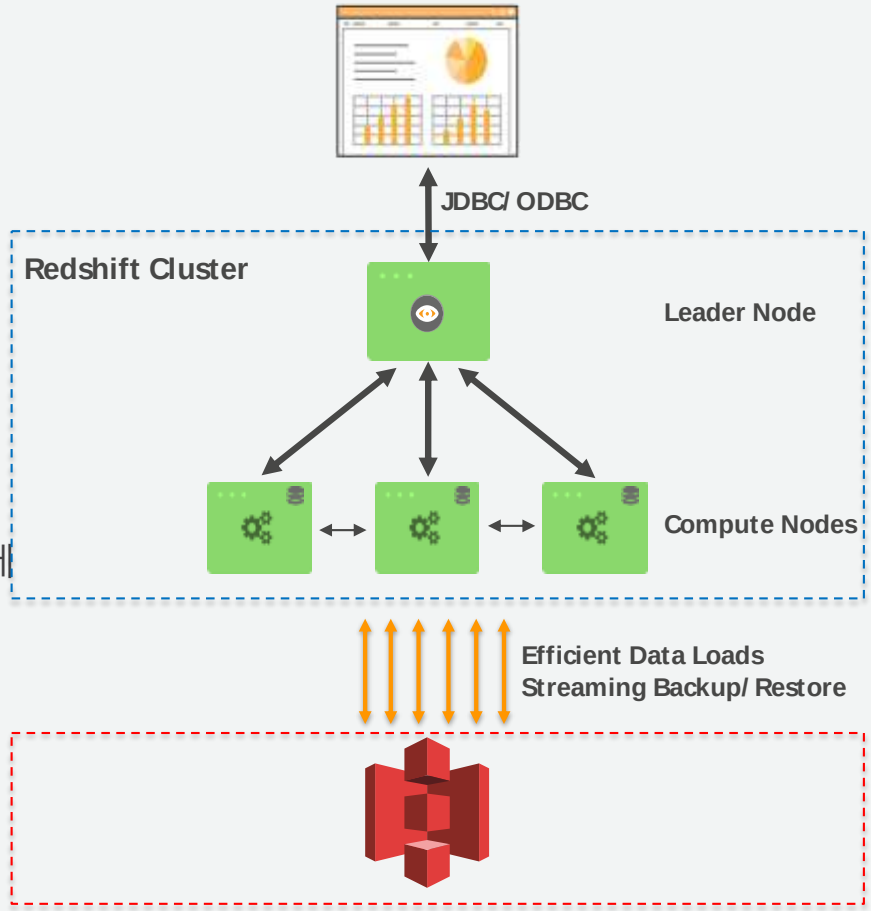
- 클라이언트, 컴퓨팅 노드와의 모든 통신 주관
- 메타 데이터 관리
- SQL문 컴파일 및 쿼리 플랜 작성
- 쿼리 실행 주관

컴퓨팅(Compute) 노드

- 로컬 열 기반 스토리지
- 모든 데이터 적재/쿼리/백업,복원/리사이징 등에 대한 병렬 분산 처리 주체
- 노드별 2, 16, 32개의 조각(Slice)

인스턴스 유형

- DS(고밀도 스토리지): EC2 d2 기반, HDD
- DC(고밀도 컴퓨팅): EC2 i3 기반, SSD



© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

인스턴스별 노드 크기

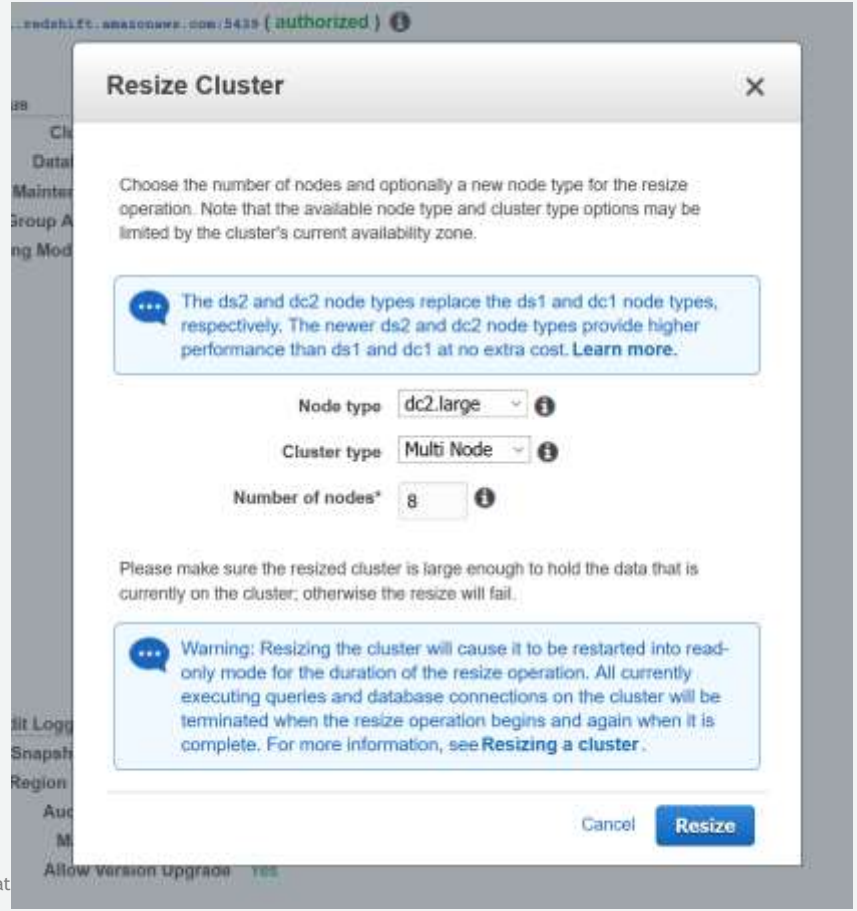
노드 크기	vCPU	RAM(GiB)	노드당 조각 수	노드당 스토리지	노드 범위	총 용량
ds2.xlarge	4	31	2	2TB HDD	1~32	64 TB
ds2.8xlarge	36	244	16	16TB HDD	2~128	2 PB

노드 크기	vCPU	RAM(GiB)	노드당 조각 수	노드당 스토리지	노드 범위	총 용량
dc1.large	2	15	2	160 GB SSD	1~32	5.12 TB
dc1.8xlarge	32	244	32	2.56 TB SSD	2~128	326 TB
dc2.large	2	15.25	2	160 GB NVMe-SSD	1~32	5.12 TB
dc2.8xlarge	32	244	16	2.56 TB NVMe-SSD	2~128	326 TB

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

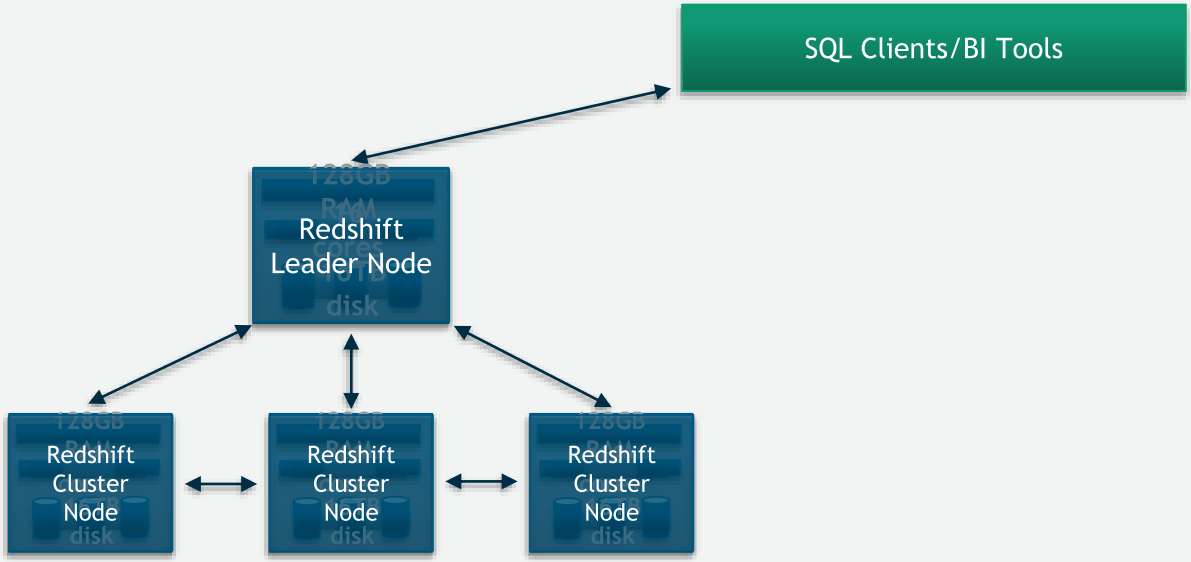
aws

노드 타입 및 클러스터 사이즈 변경



© 2018, Amazon Web Services, Inc. or its Affiliat

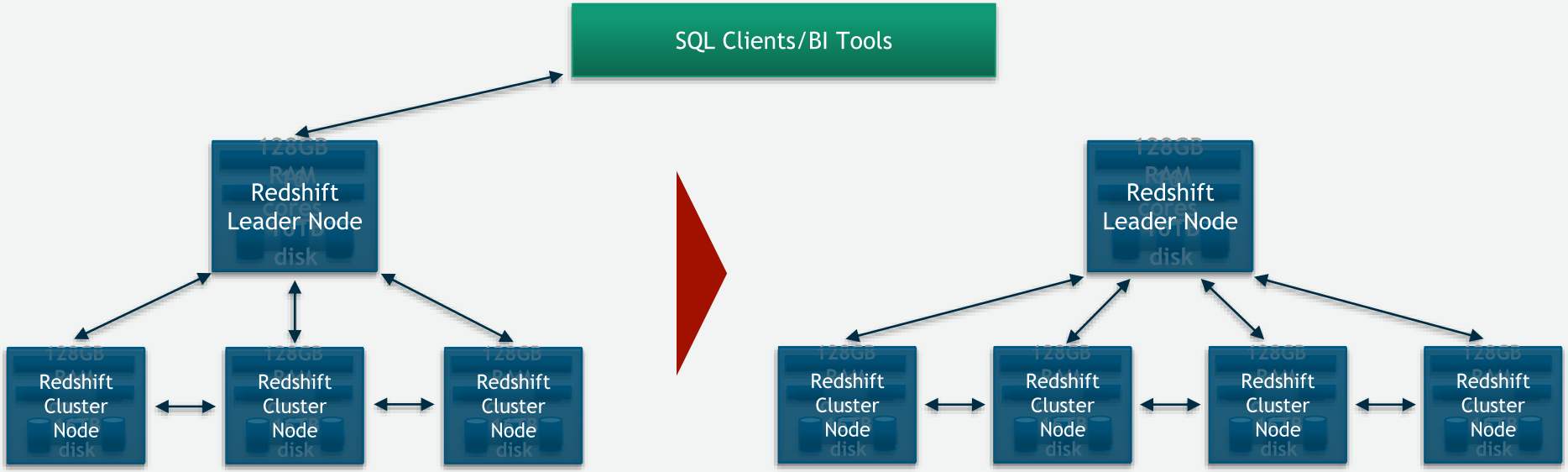
노드 타입 및 클러스터 사이즈 변경



1. 기존 클러스터를 Read Only 모드로 변경

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

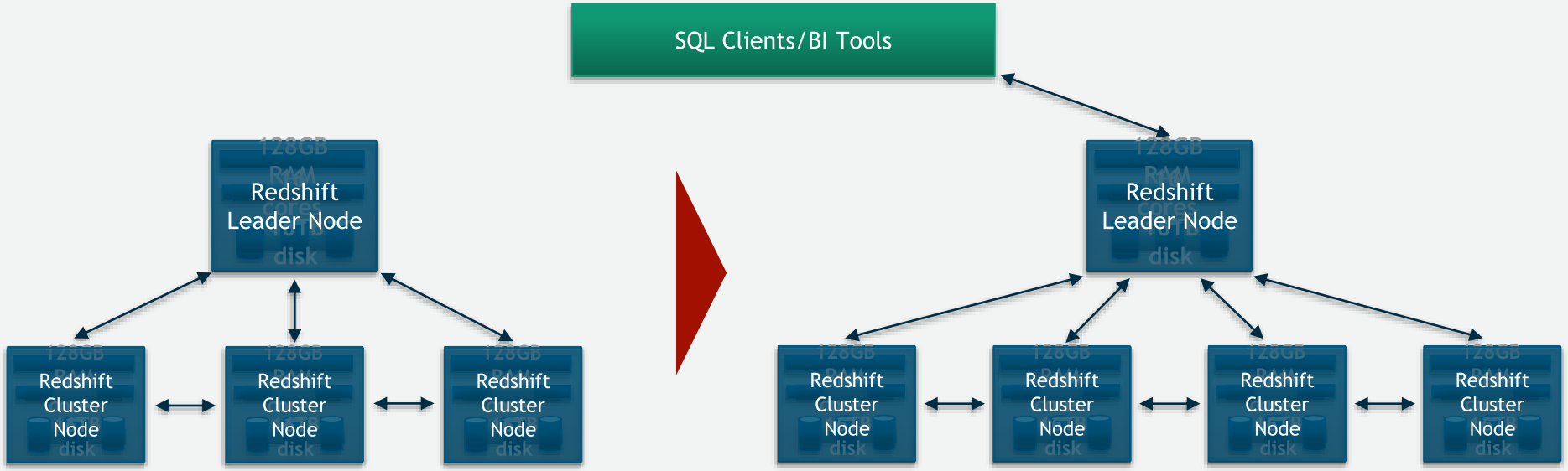
노드 타입 및 클러스터 사이즈 변경



1. 기존 클러스터를 Read Only 모드로 변경
2. 변경된 노드 수량 및 노드 타입에 맞춰 신규 클러스터 생성
3. 클러스터의 데이터를 신규 클러스터로 병렬 전송

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

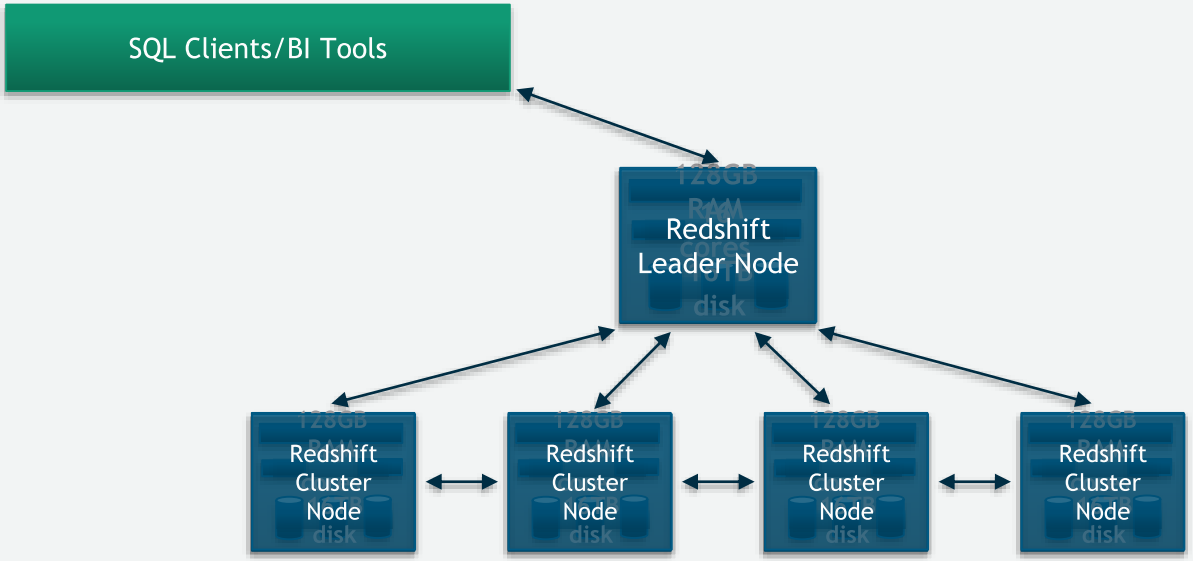
노드 타입 및 클러스터 사이즈 변경



1. 기존 클러스터를 Read Only 모드로 변경
2. 변경된 노드 수량 및 노드 타입에 맞춰 신규 클러스터 생성
3. 클러스터의 데이터를 신규 클러스터로 병렬 전송
4. 기존 클러스터의 SQL Endpoint 변경 (DNS 활용)

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

노드 타입 및 클러스터 사이즈 변경



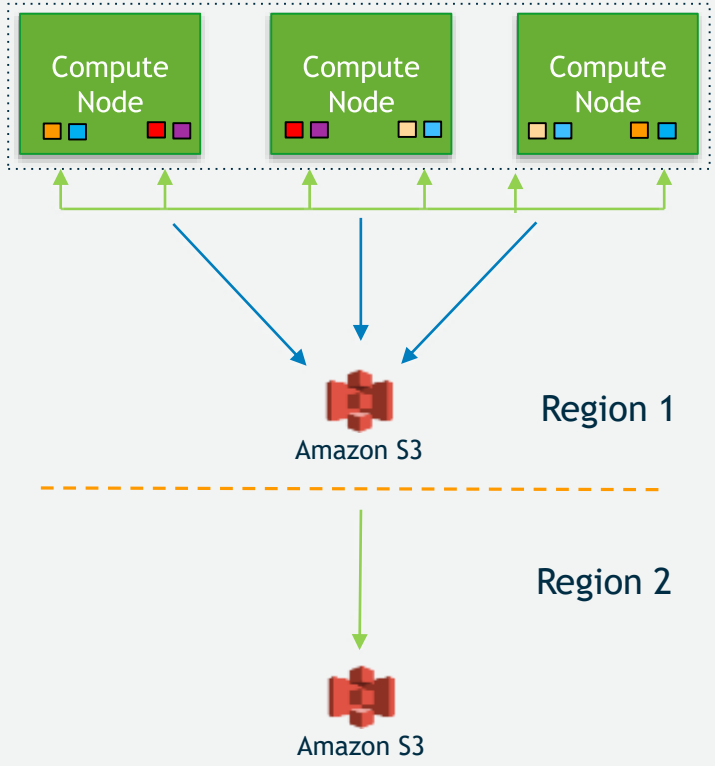
1. 기존 클러스터를 Read Only 모드로 변경
2. 변경된 노드 수량 및 노드 타입에 맞춰 신규 클러스터 생성
3. 클러스터의 데이터를 신규 클러스터로 병렬 전송
4. 기존 클러스터의 SQL Endpoint 변경 (DNS 활용)
5. 기존 클러스터 제거

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

완전 관리형 서비스

지속적/증분 백업

- 노드 간 카피 본 지원
- 지속적 증분 백업을 안전한 S3(Amazon Simple Storage Service)에 저장
- 지속적 증분 백업을 다른 리전으로 복제 가능
- Streaming 복구 지원으로 빠른 사용 가능

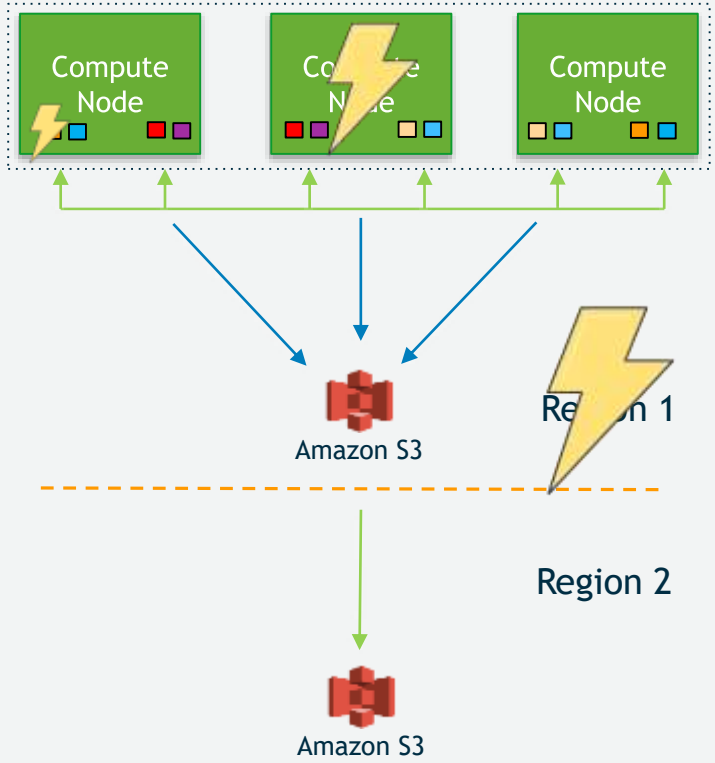


© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

완전 관리형 서비스

내결함을 위한 관리 지원

- 디스크 결함
- 노드 결함
- 네트워크 결함
- Availability Zone/Region 수준의 이벤트 발생 대비



© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

컬럼 기반 저장 방식

"Select ID, Amount, State from Order"

ID	Date	Amount	Handling	Items	State	Payment
1	10/1/2013	$53.50	5	1	VA	Visa
2	10/1/2013	$100.25	8	3	MD	MC
3	10/2/2013	$25.50	3	1	VA	Debit
4	10/2/2013	$150.75	6	10	NC	Visa
5	10/2/2013	$99.90	3	4	PA	ACH
6	10/3/2013	$75.85	7	7	MD	Visa

Row Store

Columnar Store

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

aws

테이블 분산 스타일

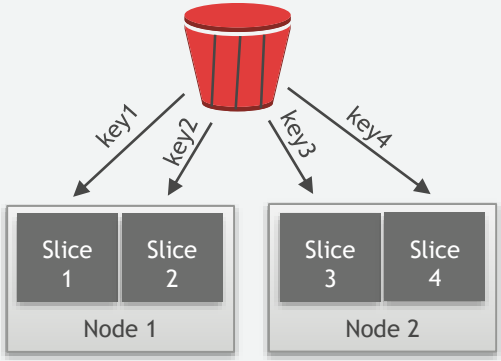
```
create table alldiststyle (col1 int)
diststyle all;

create table evendiststyle (col1 int)
diststyle even;

create table keydiststyle (col1 int)
diststyle key distkey (col1);
```

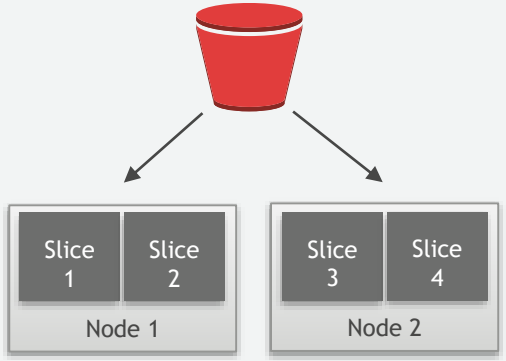
분산 키 (key)

동일 키는 동일한 위치에



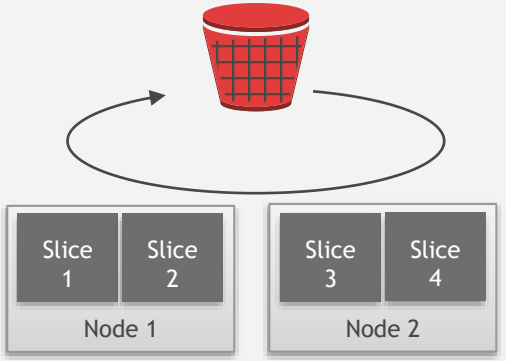
전체 (All)

모든 데이터를 각 노드에 전부



균등 (Even)

라운드 로빈



© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

적절한 분산 키의 선택

목적

- 각 노드에 균등하게 데이터 분산
- 데이터의 이동 최소화 : Co-located Joins & Aggregates

분산 키에 적절한 컬럼

- 가장 큰 테이블에서 조인에 활용되는 컬럼
- Group By 조건에서 사용되는 컬럼
- 높은 Cardinality를 보유한 컬럼

분산 키에 적절치 않은 컬럼

- Equality filter에서 활용되는 칼럼
- 데이터의 몰림을 유발하는 컬럼

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

분산 키 설정 참고 사항

1. 가장 큰 Dimension Table의 Primary키와 Fact Table의 Foreign키를 Dist Key로 선택

2. 나머지 조인이 되는 조건의 Dimension Table은 Distribution ALL을 검토 한다.

   “300만개 이하의 경우 데이터 분산 타입을 전체 (ALL)로 선택해도 무방 "

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

분산 키에 의한 데이터 분산

분산 키는 레코드가 어느 슬라이스에 저장되게 될 지를 결정함

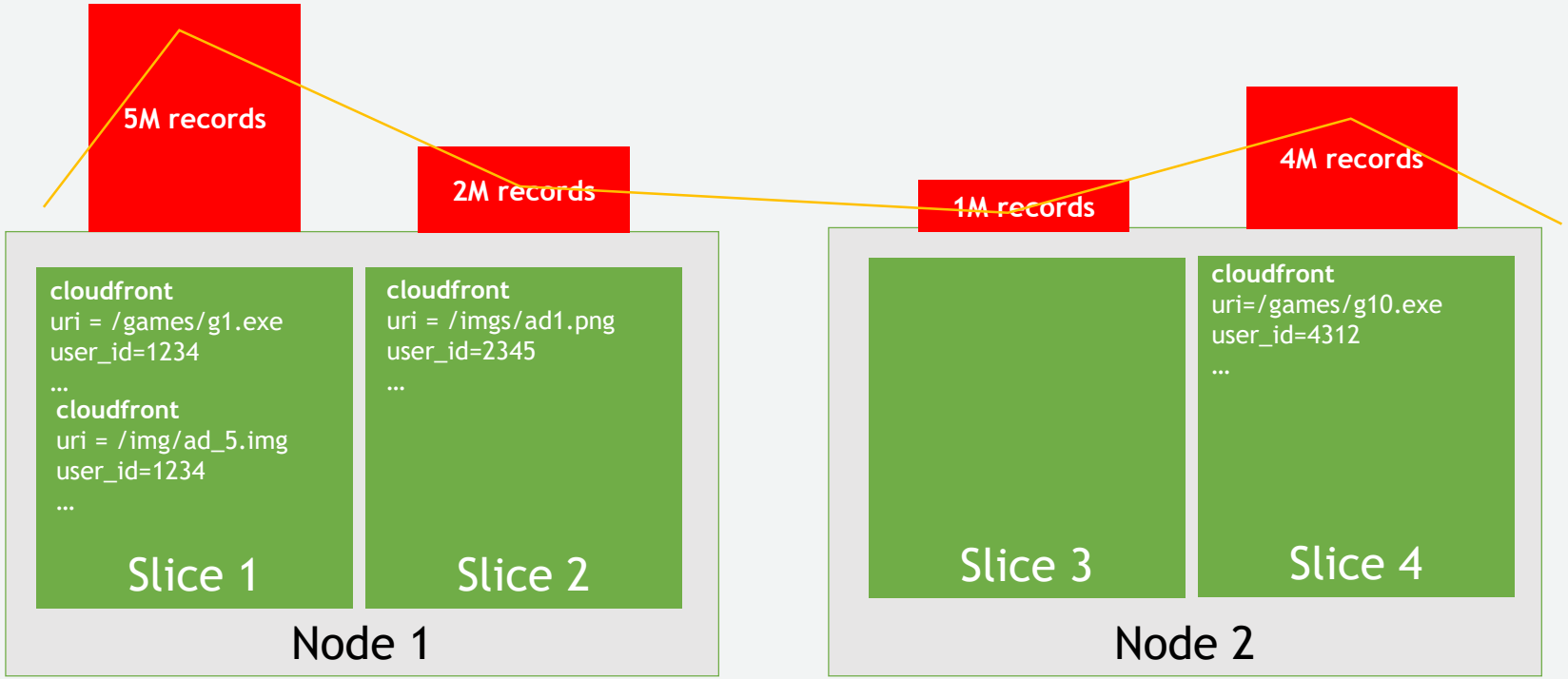
© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

분산 키에 의한 데이터 분산

분산 키를 통한 조인 연산에 대한 데이터 지역성 확보

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

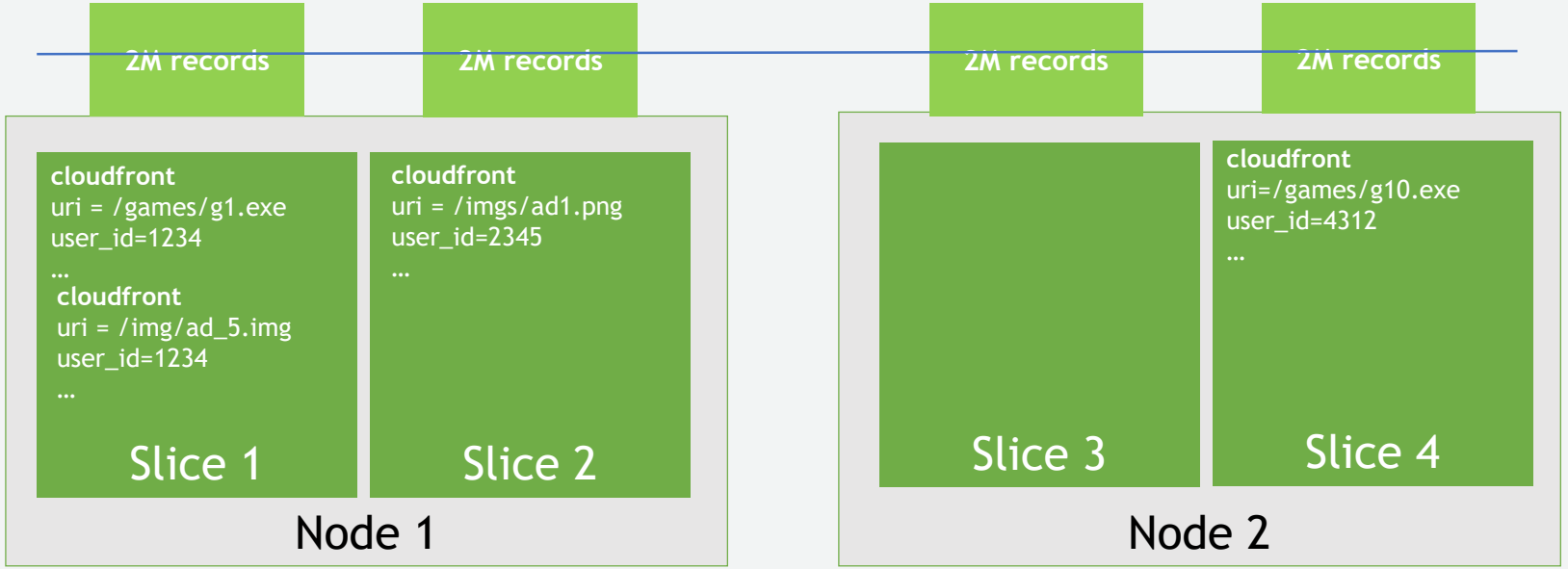
데이터 분산과 분산 키



적절치 못한 분산키 선택은 특정 슬라이스로의 데이터 집중 현상을 초래

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

데이터 분산과 분산 키



균등한 데이터 분산은 쿼리 성능 향상에 도움

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

정렬 키 (Sort Key)

데이터 로드 시 정렬 키 순서대로 디스크에 저장

정렬 키의 종류

- 정렬 키: 테이블의 단일 컬럼의 값을 기준으로 데이터를 정렬하여 저장
- 복합 정렬 키 (Compound Sort Key):
  - 다수의 컬럼을 활용 (최대 6개) 활용하여 정렬 키로 활용
  - 선언한 순서에 따라 순차적으로 정렬되어 저장됨
  - 조인 및 Group By, Order By에 효과적이며, 특히 사전에 순서대로 정렬이 되어 있기 때문에 Merge 조인에 효과적
- 인터리브 정렬 키 (Interleaved Sort Key)
  - 다수의 컬럼을 활용 (최대 8개) 활용하여 정렬 키로 활용
  - 정렬 키에서 각 열, 즉 열의 하위 집합에 똑같은 가중치를 부여
  - Ad-hoc 형태의 쿼리에서 높은 성능을 제공함

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

aws

Amazon Redshift 사용 예제: amazon.com

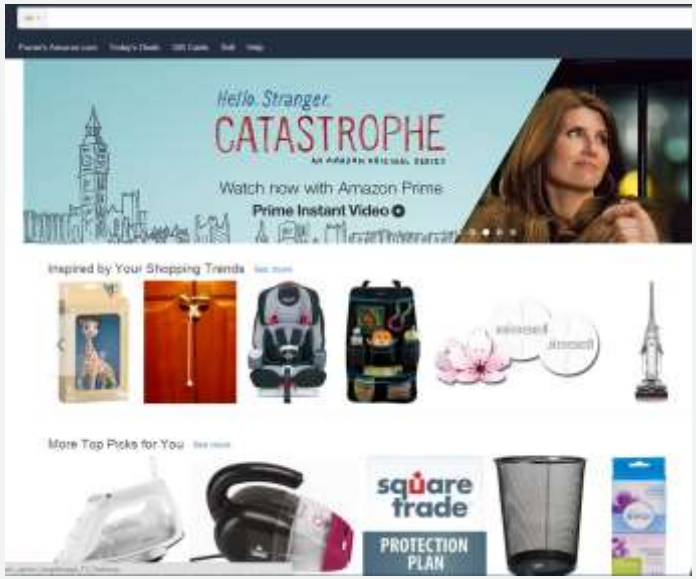
amazon.com의 웹 로그 분석

- 1PB+ 워크로드, 일간 2TB 축적, 매년 67% 증가
- 가장 큰 테이블: 400 TB

고객 행동 파악/분석 요건

AS_IS 시스템

- Data - 시간 당 일주일 치 분석
- Hadoop - 시간 당 한달 치 분석



© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

aws

Amazon Redshift 사용 예제: amazon.com

TO_BE 시스템

- 64 개 클러스터
- 800 노드
- 13 PB 데이터 크기
- 2명의 데이터베이스 관리자

수행 결과

- 15개월 치 1PB 처리 쿼리에 약 14분 소요
- 500억 건 데이터 로딩에 약 10분 소요
- 210억 건 데이터와 100억 건 데이터 Join 쿼리 약 2시간 소요
  (기존 Hive에서 3일)

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

aws

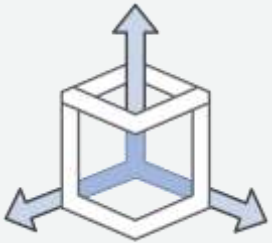
Redshift Spectrum

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

Amazon Redshift Spectrum

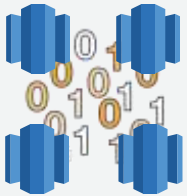
수천 대의 노드를 사용하여 S3에 저장된 데이터에 직접 쿼리를 실행

Fast @ exabyte scale

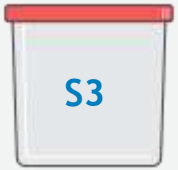


Elastic & highly available

On-demand, pay-per-query



High concurrency: Multiple clusters access same data



No ETL: Query data in-place using open file formats

Full Amazon Redshift SQL support

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

aws

1

```
SELECT COUNT(*)
FROM S3.EXT_TABLE
JOIN RS.TABLE …
GROUP BY…
```

JDBC/ODBC

Amazon Redshift

Redshift Nodes

Redshift Spectrum

Scale-out serverless compute

1 2 3 4 … N

Amazon S3

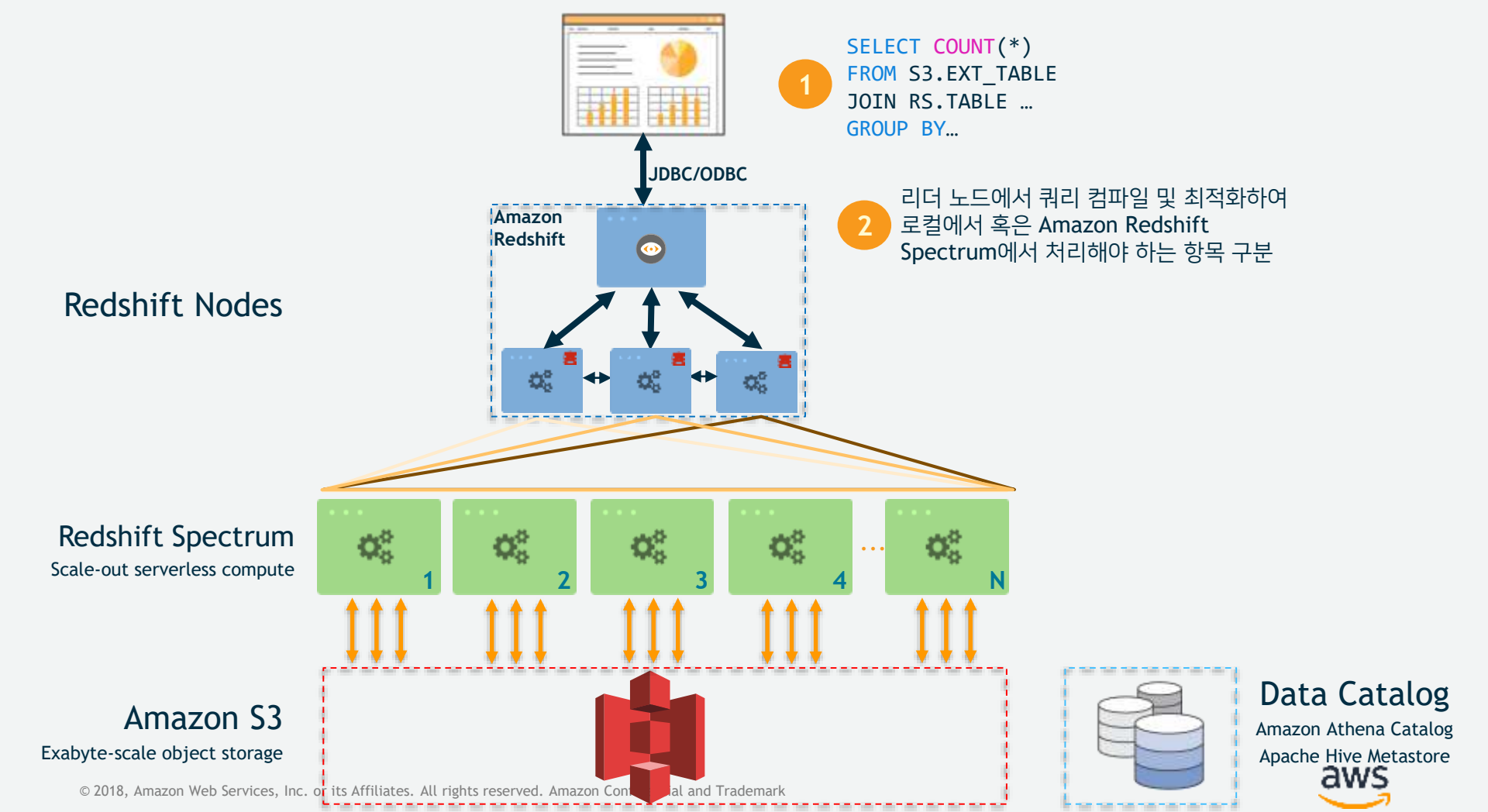
Exabyte-scale object storage

Data Catalog

Amazon Athena Catalog

Apache Hive Metastore

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark



© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Con al and Trademark

```
SELECT COUNT(*)
FROM S3.EXT_TABLE
JOIN RS.TABLE …
GROUP BY…
```

JDBC/ODBC

Amazon Redshift

Redshift Nodes

2 리더 노드에서 쿼리 컴파일 및 최적화하여 로컬에서 혹은 Amazon Redshift Spectrum에서 처리해야 하는 항목 구분

3 모든 컴퓨팅 노드로 쿼리 플랜 전송

Redshift Spectrum

Scale-out serverless compute

1 2 3 4 … N

Amazon S3

Exabyte-scale object storage

Data Catalog

Amazon Athena Catalog

Apache Hive Metastore

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

JDBC/ODBC

Amazon Redshift

Redshift Nodes

```
SELECT COUNT(*)
FROM S3.EXT_TABLE
JOIN RS.TABLE …
GROUP BY…
```

1

2 리더 노드에서 쿼리 컴파일 및 최적화하여 로컬에서 혹은 Amazon Redshift Spectrum에서 처리해야 하는 항목 구분

3 모든 컴퓨팅 노드로 쿼리 플랜 전송

4 데이터 카탈로그에서 얻은 파티션 정보로 동적으로 파티션을 정리

Redshift Spectrum

Scale-out serverless compute

1 2 3 4 N

Amazon S3

Exabyte-scale object storage

Data Catalog

Amazon Athena Catalog

Apache Hive Metastore

aws

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

JDBC/ODBC

Amazon Redshift

Redshift Nodes

```
SELECT COUNT(*)
FROM S3.EXT_TABLE
JOIN RS.TABLE …
GROUP BY…
```

1

2 리더 노드에서 쿼리 컴파일 및 최적화하여 로컬에서 혹은 Amazon Redshift Spectrum에서 처리해야 하는 항목 구분

3 모든 컴퓨팅 노드로 쿼리 플랜 전송

4 데이터 카탈로그에서 얻은 파티션 정보로 동적으로 파티션을 정리

5 Amazon Redshift Spectrum 계층에 요청 전송

Redshift Spectrum

Scale-out serverless compute

1 2 3 4 … N

Amazon S3

Exabyte-scale object storage

Data Catalog

Amazon Athena Catalog

Apache Hive Metastore

aws

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

```
SELECT COUNT(*)
FROM S3.EXT_TABLE
JOIN RS.TABLE …
GROUP BY…
```

1

JDBC/ODBC

Amazon Redshift

Redshift Nodes

2 리더 노드에서 쿼리 컴파일 및 최적화하여 로컬에서 혹은 Amazon Redshift Spectrum에서 처리해야 하는 항목 구분

3 모든 컴퓨팅 노드로 쿼리 플랜 전송

4 데이터 카탈로그에서 얻은 파티션 정보로 동적으로 파티션을 정리

5 Amazon Redshift Spectrum 계층에 요청 전송

Redshift Spectrum

Scale-out serverless compute

1 2 3 4 … N

6 Amazon Redshift Spectrum 노드에서 S3 데이터 스캔

Amazon S3

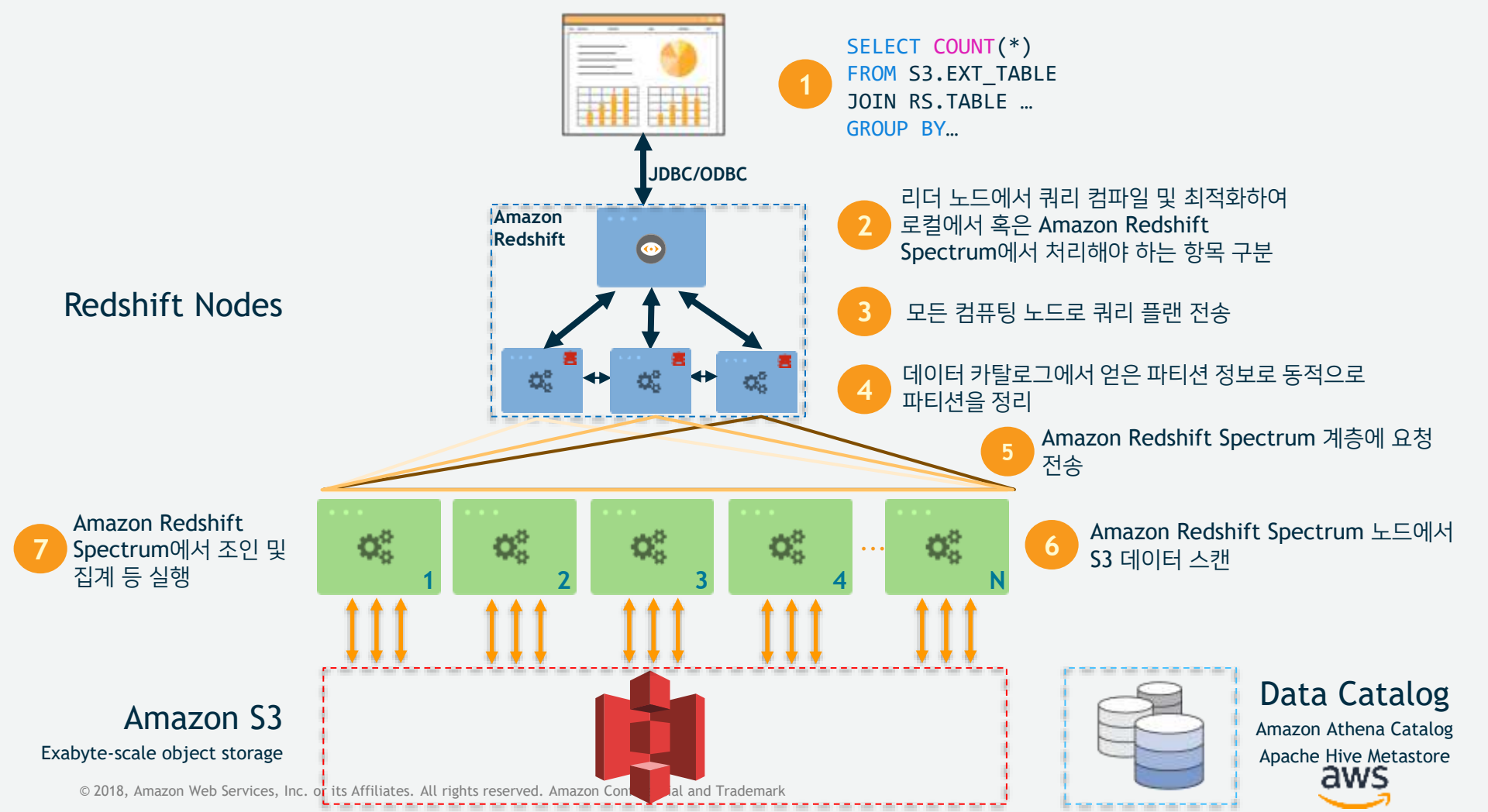
Exabyte-scale object storage

Data Catalog

Amazon Athena Catalog

Apache Hive Metastore

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark



© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

JDBC/ODBC

Amazon Redshift

Redshift Nodes

1
```
SELECT COUNT(*)
FROM S3.EXT_TABLE
JOIN RS.TABLE …
GROUP BY…
```

2 리더 노드에서 쿼리 컴파일 및 최적화하여 로컬에서 혹은 Amazon Redshift Spectrum에서 처리해야 하는 항목 구분

3 모든 컴퓨팅 노드로 쿼리 플랜 전송

4 데이터 카탈로그에서 얻은 파티션 정보로 동적으로 파티션을 정리

5 Amazon Redshift Spectrum 계층에 요청 전송

6 Amazon Redshift Spectrum 노드에서 S3 데이터 스캔

7 Amazon Redshift Spectrum에서 조인 및 집계 등 실행

8 로컬 클러스터 내에서 완료된 테이블과의 최종 집계 및 조인 수행

1 2 3 4 … N

Amazon S3

Exabyte-scale object storage

Data Catalog

Amazon Athena Catalog

Apache Hive Metastore

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

```
SELECT COUNT(*)
FROM S3.EXT_TABLE
JOIN RS.TABLE …
GROUP BY…
```

1

JDBC/ODBC

Amazon Redshift

2 리더 노드에서 쿼리 컴파일 및 최적화하여 로컬에서 혹은 Amazon Redshift Spectrum에서 처리해야 하는 항목 구분

3 모든 컴퓨팅 노드로 쿼리 플랜 전송

4 데이터 카탈로그에서 얻은 파티션 정보로 동적으로 파티션을 정리

5 Amazon Redshift Spectrum 계층에 요청 전송

6 Amazon Redshift Spectrum 노드에서 S3 데이터 스캔

7 Amazon Redshift Spectrum에서 조인 및 집계 등 실행

8 로컬 클러스터 내에서 완료된 테이블과의 최종 집계 및 조인 수행

9 클라이언트로 결과 전송

Redshift Nodes

1 2 3 4 … N

Amazon S3

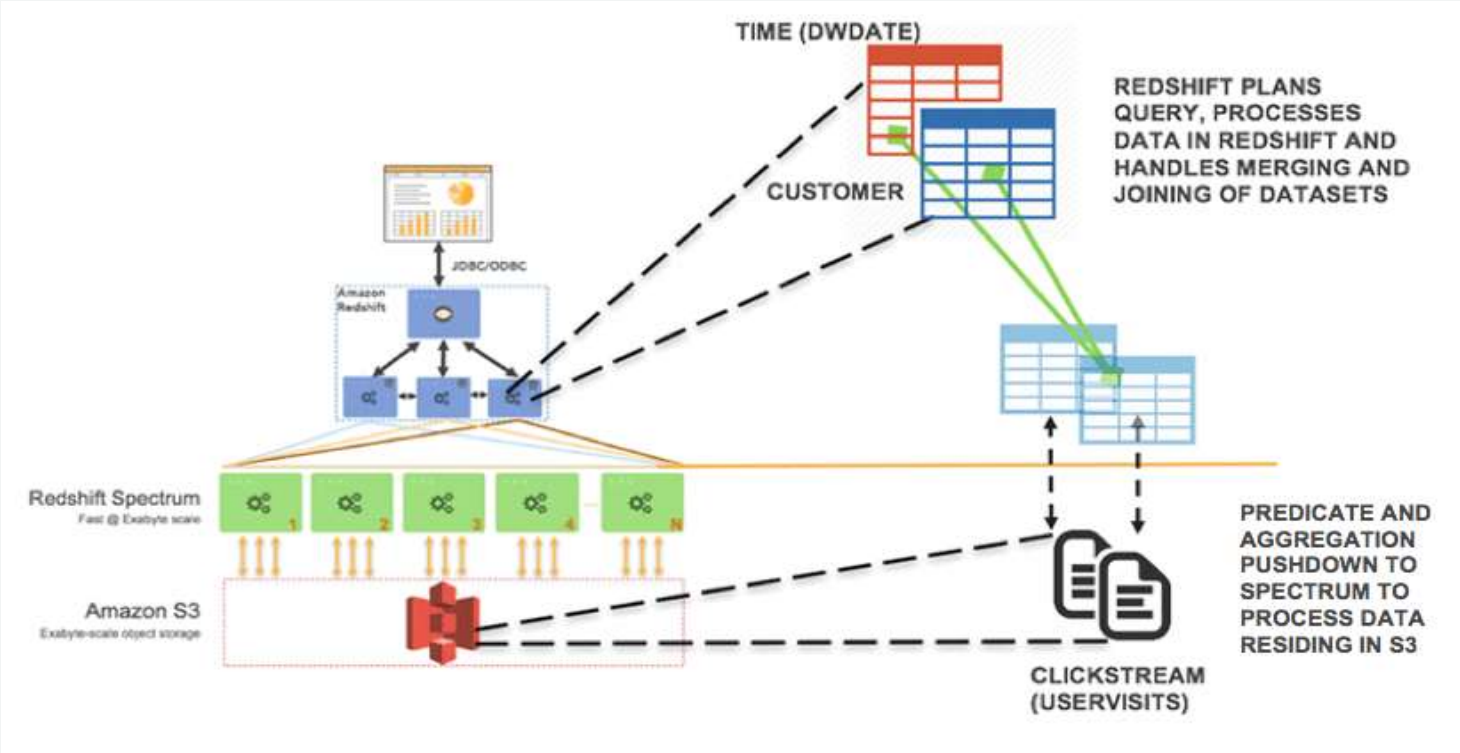
Exabyte-scale object storage

Data Catalog

Amazon Athena Catalog

Apache Hive Metastore

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark



© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

Amazon Redshift Spectrum Best Practice

- 파티셔닝 및 Columnar 파일 포맷 (ORC, Parquet) 사용
    - 파티셔닝 컬럼의 조건
        - 필터 및 조인의 조건이 되는 컬럼
        - 비니지스 유닛
        - 비니지스 그룹
        - 날짜 및 시간
- 파일의 갯수는 Redshift의 Slice의 수량 이상
- 파일의 크기는 64MB 이상을 권고
- 각 파일은 동일한 크기를 권고

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

Spectrum 사용 예제

```
SELECT
        P.ASIN,
        P.TITLE,
        R.POSTAL_CODE,
        P.RELEASE_DATE,
        SUM(D.QUANTITY * D.OUR_PRICE) AS SALES_sum
FROM
        s3.d_customer_order_item_details D,
        asin_attributes A,
        products P,
        regions R
WHERE
        D.ASIN = P.ASIN AND
        P.ASIN = A.ASIN AND
        D.REGION_ID = R.REGION_ID AND
        A.EDITION LIKE '%FIRST%' AND
        P.TITLE LIKE '%Potter%' AND
        P.AUTHOR = 'J. K. Rowling' AND
        R.COUNTRY_CODE = 'US' AND
        R.CITY = 'Seattle' AND
        R.STATE = 'WA' AND
        D.ORDER_DAY :: DATE >= P.RELEASE_DATE AND
        D.ORDER_DAY :: DATE < dateadd(day, 3, P.RELEASE_DATE)
GROUP BY P.ASIN, P.TITLE, R.POSTAL_CODE, P.RELEASE_DATE
ORDER BY SALES_sum DESC
LIMIT 20;
```

Hive (1000 nodes)	Redshift Spectrum
5 years	155 seconds

- 지난 20 년간 매일 약 140TB의 고객 품목 주문 내역 기록
- S3에서 15,000 개의 파티션에 걸쳐 1 억 9 천만 개의 파일
- 미국 및 기타 국가에서 하루에 하나의 파티션 추가
- 총 데이터 크기가 EB 이상

Optimization:

- Compression ……..…….5X
- Columnar file format……..……..10X
- Scanning with 2500 nodes……2500X
- Static partition elimination……..……2X
- Dynamic partition elimination……350X
- Amazon Redshift query optimizer..40X

- 20 노드 Hive 클러스터 및 1.4TB를 사용하여 추정
- 쿼리는 20 DC1.8xLarge Redshift Cluster 사용
- 실제 판매 데이터 아닌 Amazon Retail에서 사용하는 데이터 형식을 기반으로 생성한 데모 데이터

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark

aws

Thank you!