amazon web services | Pop-up Loft

Data warehousing in the era of Big Data: Intro to Amazon Redshift

Shree Kenghe, Solution Architect, AWS

Agenda

- Introduction
- Benefits
- Use cases
- Getting started
- Q&A

AWS big data portfolio

Collect

AWS Direct Connect

AWS Import/Export

Amazon Kinesis Streams

New

Amazon Kinesis Firehose

Store

Amazon S3

Amazon RDS, Amazon Aurora

Amazon Glacier

Amazon DynamoDB

Amazon CloudSearch

New

Amazon Elasticsearch Service

Analyze

Amazon EMR

Amazon EC2

Amazon Redshift

Amazon Machine Learning

New

Amazon QuickSight

New

AWS Database Migration Service

AWS Data Pipeline

Amazon
Redshift

a lot faster
a lot simpler
a lot cheaper

Relational data warehouse

Massively parallel; petabyte scale

Fully managed

HDD and SSD platforms

$1,000/TB/year; starts at $0.25/hour

The Amazon Redshift view of data warehousing

Enterprise

10x cheaper

Easy to provision

Higher DBA productivity

Big data

10x faster

No programming

Easily leverage BI tools, Hadoop, machine learning, streaming

SaaS

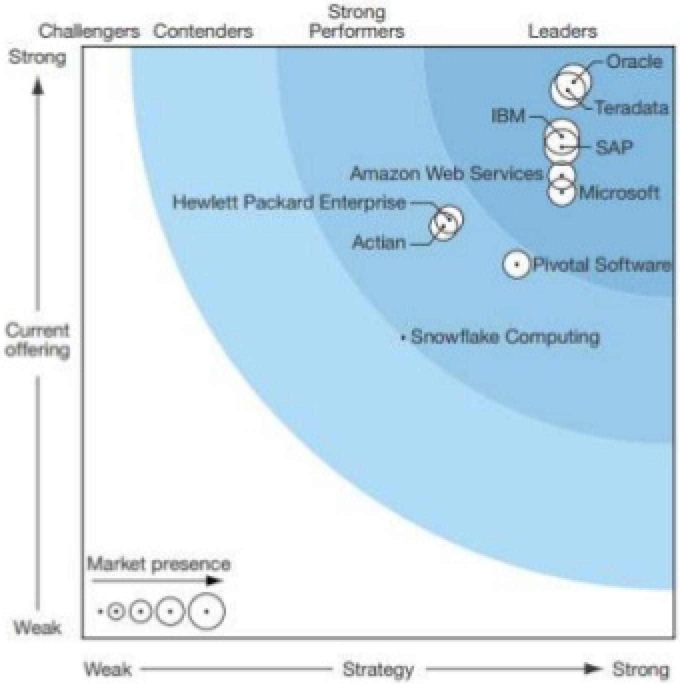
Analysis inline with process flows

Pay as you go, grow as you need

Managed availability and disaster recovery

Forrester Wave™ Enterprise Data Warehouse Q4 ’15

FIGURE 2 The Forrester Wave™: Enterprise Data Warehouse, Q4 '15

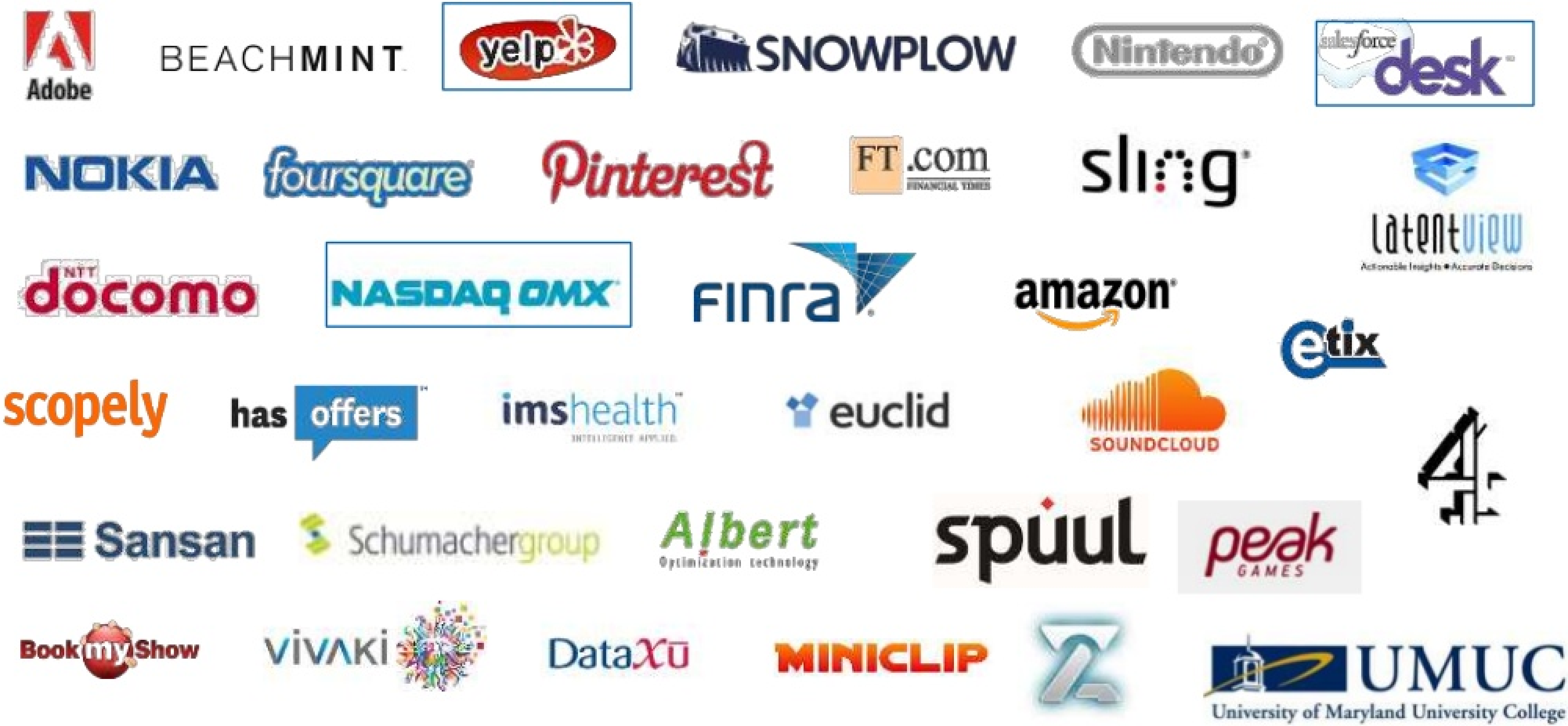


The Forrester Wave
Smart data for smart decisions

Go to Forrester.com to download the Forrester Wave tool for more detailed product evaluations, feature comparisons, and customizable rankings.

The Forrester Wave™ is a graphical representation of Forrester's call on a market and is plotted using a detailed spreadsheet with exposed scores, weightings, and comments. Forrester does not endorse any vendor, product, or service depicted in the Forrester Wave. Information is based on best available resources. Opinions reflect judgment at the time and are subject to change.

Selected Amazon Redshift customers



Amazon Redshift architecture

Leader node

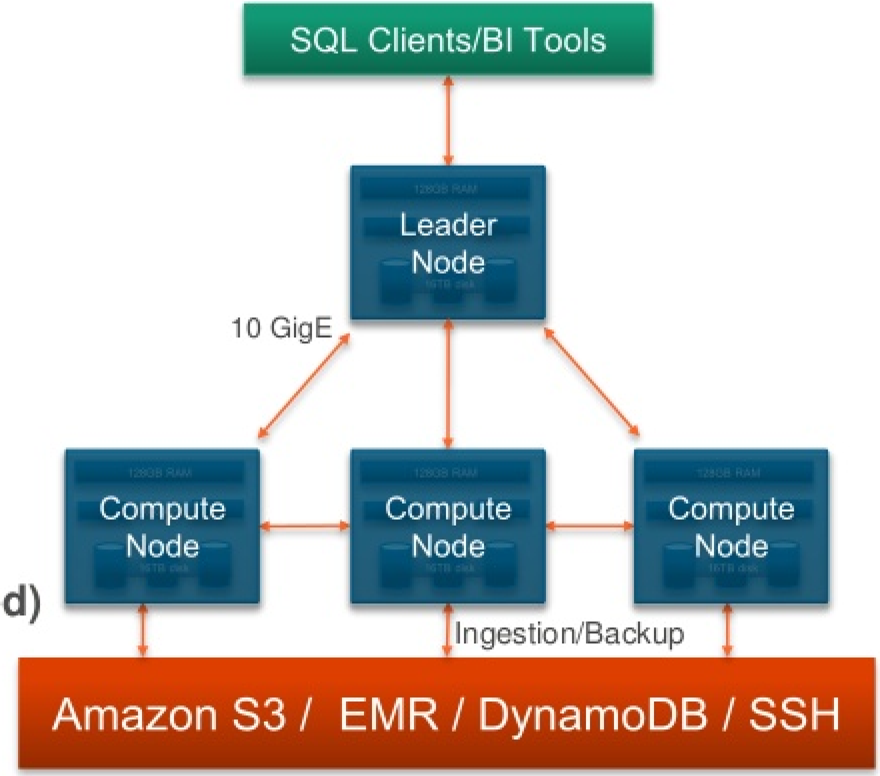
- Simple SQL endpoint
- Stores metadata
- Optimizes query plan
- Coordinates query execution

Compute nodes

- Local columnar storage
- Parallel/distributed execution of all queries, loads, backups, restores, resizes

Start at just $0.25/hour, grow to 2 PB (compressed)

- DC1: SSD; scale from 160 GB to 326 TB
- DS2: HDD; scale from 2 TB to 2 PB

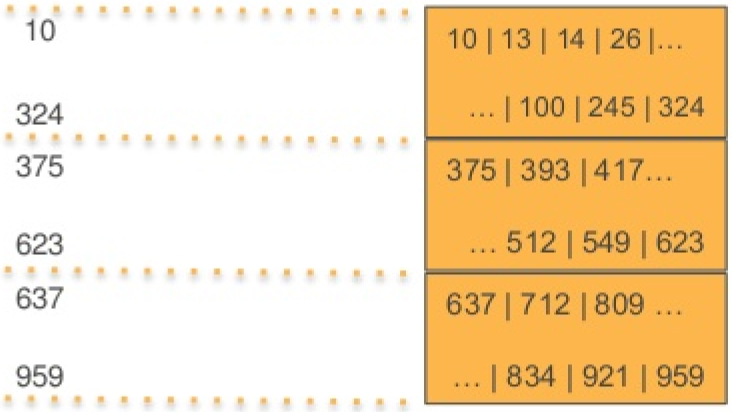


Benefit #1: Amazon Redshift is fast

Dramatically less I/O

- Column storage
- Data compression
- Zone maps
- Direct-attached storage
- Large data block sizes

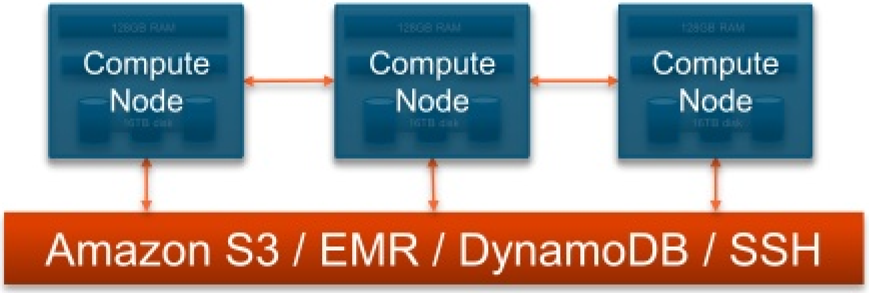
```
analyze compression listing;

  Table  |     Column     | Encoding
---------+----------------+----------
 listing | listid         | delta
 listing | sellerid       | delta32k
 listing | eventid        | delta32k
 listing | dateid         | bytedict
 listing | numtickets     | bytedict
 listing | priceperticket | delta32k
 listing | totalprice     | mostly32
 listing | listtime       | raw
```

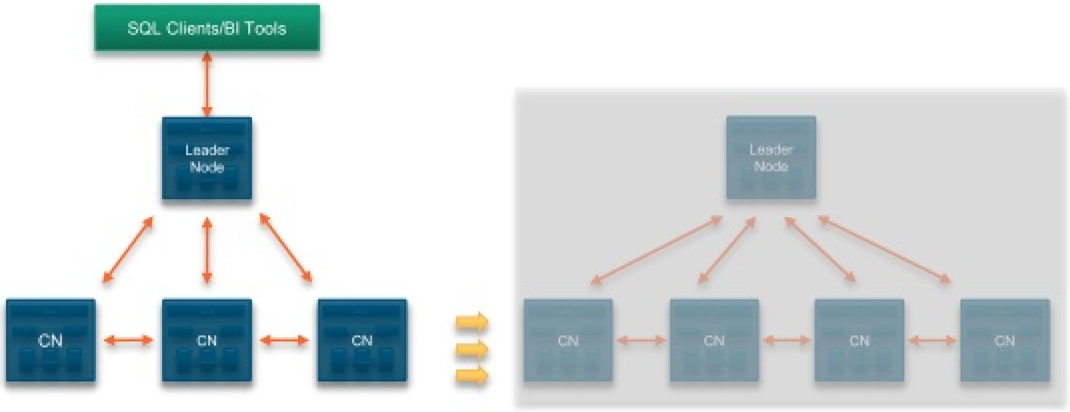


Benefit #1: Amazon Redshift is fast

Parallel and distributed

- Query
- Load
- Export
- Backup
- Restore
- Resize





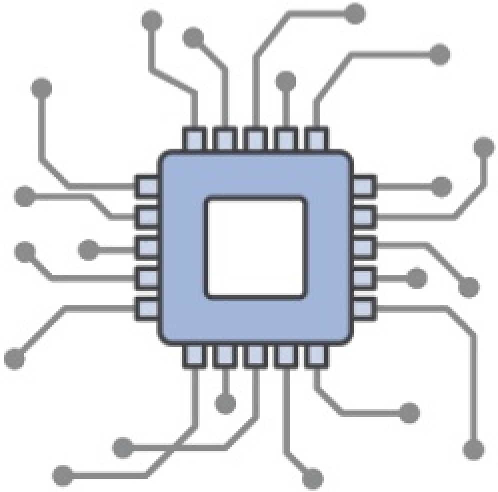
Benefit #1: Amazon Redshift is fast

Hardware optimized for I/O intensive workloads, 4 GB/sec/node

Enhanced networking, over 1 million packets/sec/node

Choice of storage type, instance size

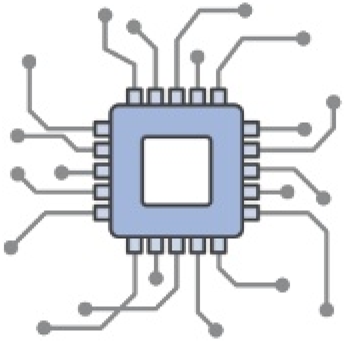
Regular cadence of auto-patched improvements



Benefit #1: Amazon Redshift is fast

New Dense Storage (HDD) instance type (Jun 15)

- Improved memory 2x, compute 2x, disk throughput 1.5x
- Cost: Same as our prior generation!
- Performance improvement: 50%



Enhanced I/O and commit improvements (Jan 16)

- Reduce amount of time to commit data
- Throughput performance improvement: 35%

Improved memory allocation for query processing (May 16)

- Increased overall throughput by up to 60%

Benefit #2: Amazon Redshift is inexpensive

DS2 (HDD)	Price per hour for DS2.XL single node	Effective annual price per TB compressed
On-demand	$ 0.850	$ 3,725
1 year reservation	$ 0.500	$ 2,190
3 year reservation	$ 0.228	$ 999

DC1 (SSD)	Price per hour for DC1.L single node	Effective annual price per TB compressed
On-demand	$ 0.250	$ 13,690
1 year reservation	$ 0.161	$ 8,795
3 year reservation	$ 0.100	$ 5,500

Pricing is simple

Number of nodes x price/hour

No charge for leader node

No upfront costs

Pay as you go

Benefit #3: Amazon Redshift is fully managed

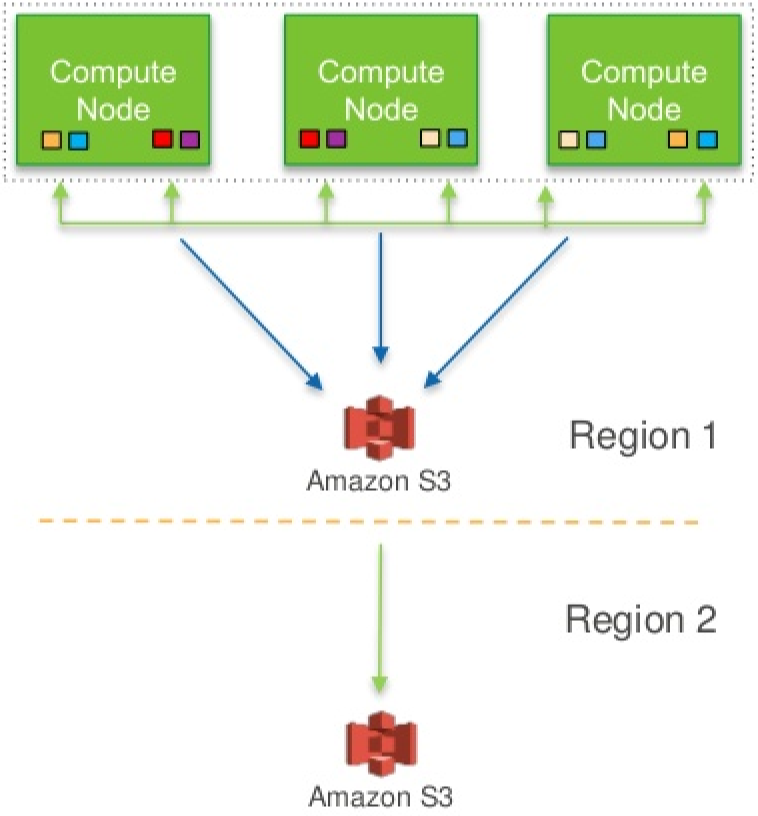
Continuous/incremental backups

Multiple copies within cluster

Continuous and incremental backups to Amazon S3

Continuous and incremental backups across regions

Streaming restore



Benefit #3: Amazon Redshift is fully managed

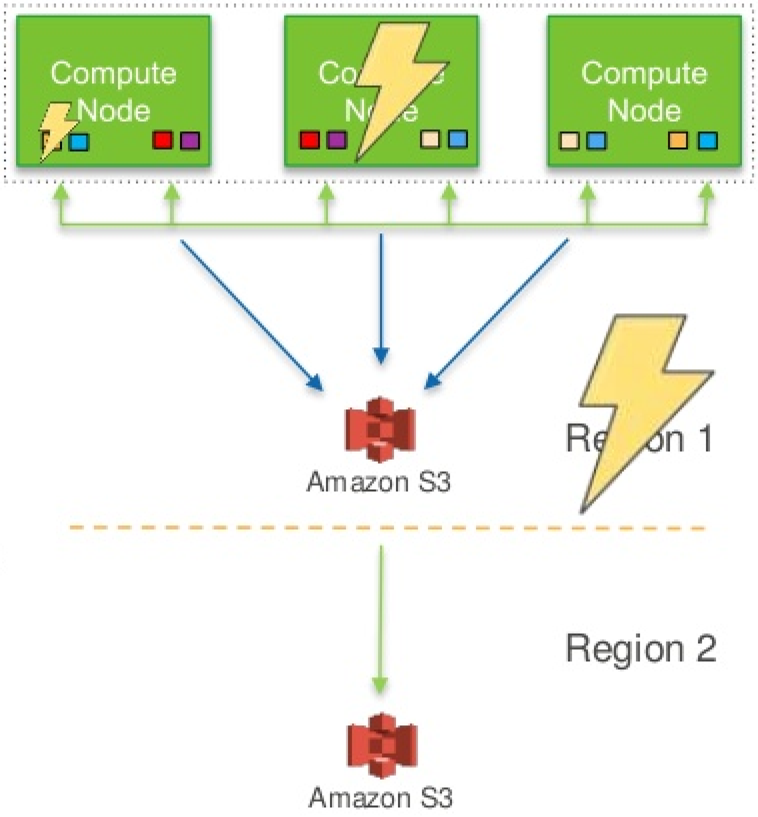
Fault tolerance

Disk failures

Node failures

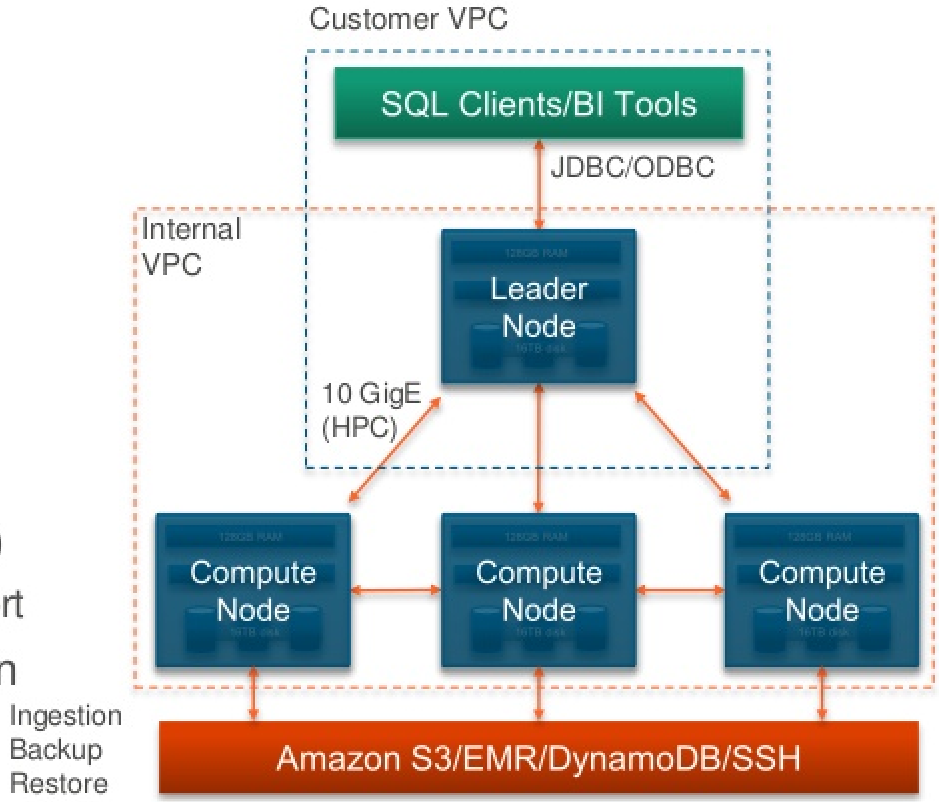
Network failures

Availability Zone/region level disasters



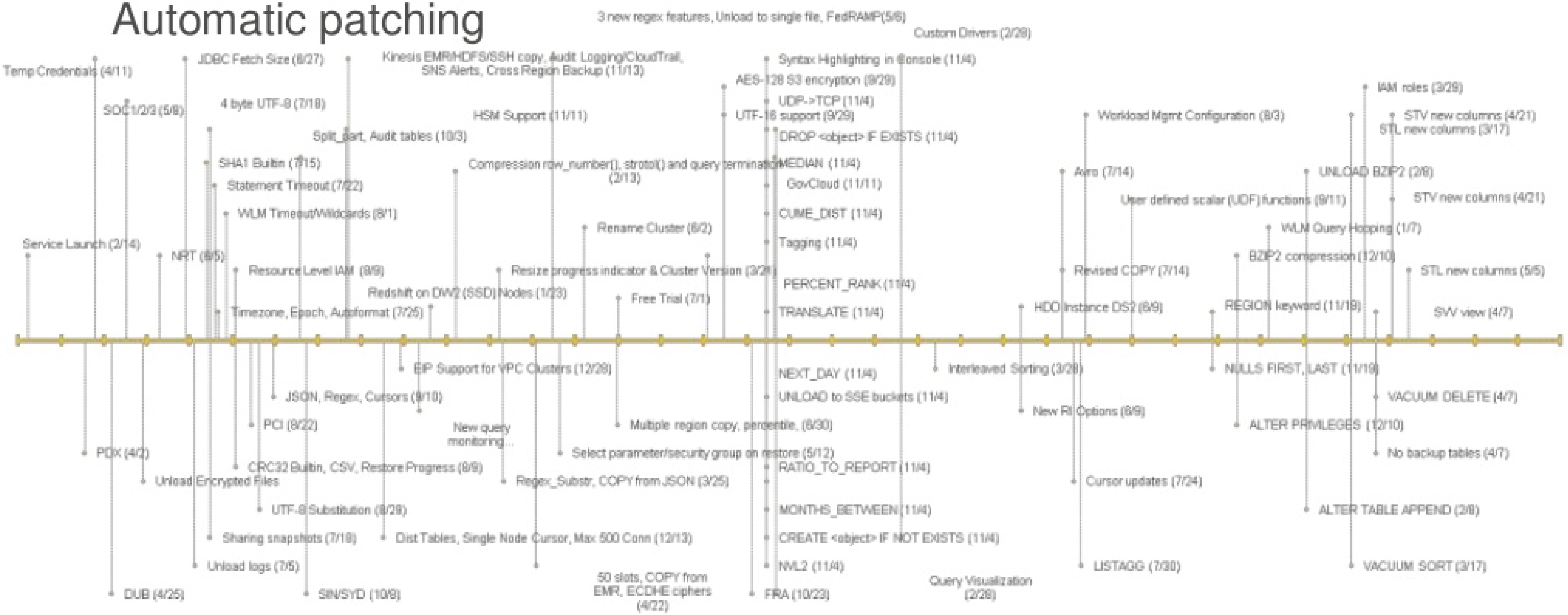
Benefit #4: Security is built-in

- Load encrypted from S3
- SSL to secure data in transit
    - ECDHE perfect forward security
- Amazon VPC for network isolation
- Encryption to secure data at rest
    - All blocks on disks and in S3 encrypted
    - Block key, cluster key, master key (AES-256)
    - On-premises HSM & AWS CloudHSM support
- Audit logging and AWS CloudTrail integration
- SOC 1/2/3, PCI-DSS, FedRAMP, BAA



Benefit #5: We innovate quickly

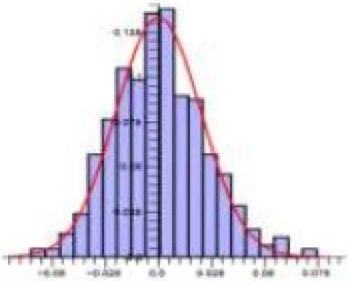
- Well over 125 new features added since launch
- Release every two weeks
- Automatic patching



Benefit #6: Redshift is powerful

- Approximate functions
- User defined functions
- Machine learning
- Data science

HyperLogLog: analysis of a near-optimal cardinality algorithm



Benefit #7: Amazon Redshift has a large ecosystem

Data integration

- ATTUNITY CloudBeam
- bryte Systems
- FlyData
- INFORMATICA
- Matillion
- pentaho
- Segment
- snapLogic
- talend integration at any scale
- Qubole
- xplenty

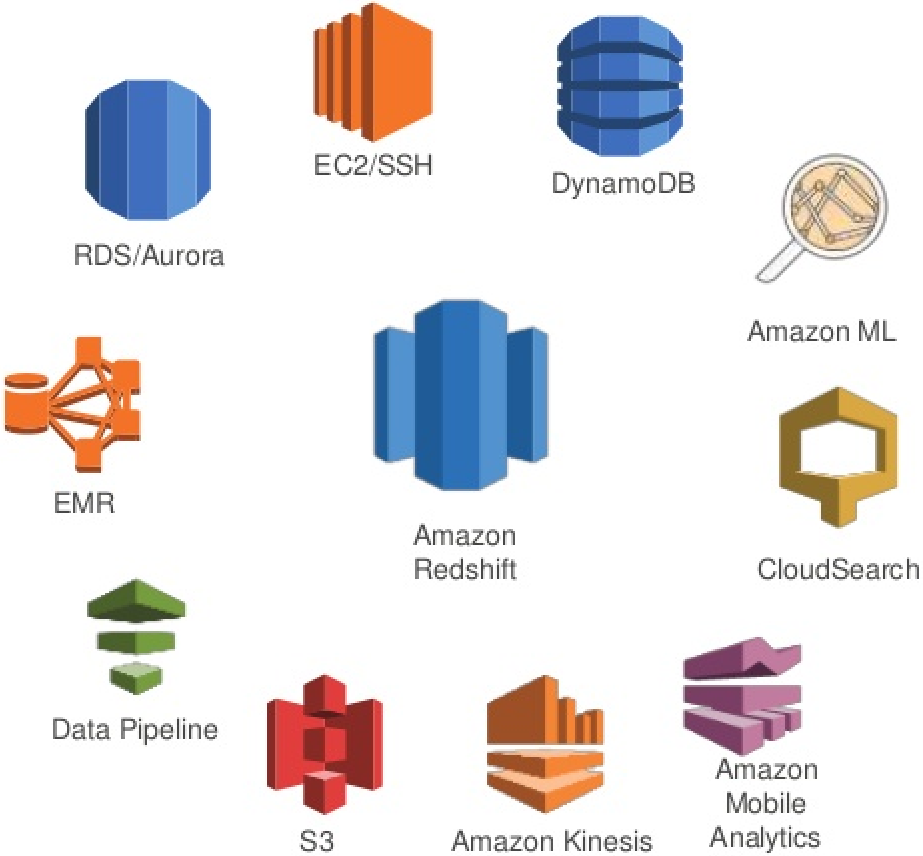
Business intelligence

- Actian
- alteryx
- ACTUATE The BIRT Company
- birst
- ClearStory DATA Now You See It
- Chart.io
- Dundas
- Business Partner IBM
- infor
- JReport JINFONET SOFTWARE
- Logi ANALYTICS
- looker
- MicroStrategy
- MODE
- PANORAMA
- pentaho A Hitachi Data Systems Company
- periscope
- Qlik
- SiSense
- REDROCK BI
- tableau software
- TIBCO
- Yellowfin

Systems integrators

- 2ND WATCH
- accenture
- BEEVA
- BRIGHTLIGHT
- Capgemini
- cloudpack
- Cognizant
- FULL 360
- iOLAP
- isobar
- MAGNUS DATA
- NorthBay solutions
- SLALOM CONSULTING

Benefit #8: Service oriented architecture



Use cases

NTT Docomo: Japan’s largest mobile service provider

NTT docomo

68 million customers

Tens of TBs per day of data across a mobile network

6 PB of total data (uncompressed)

Data science for marketing operations, logistics, and so on

Greenplum on-premises

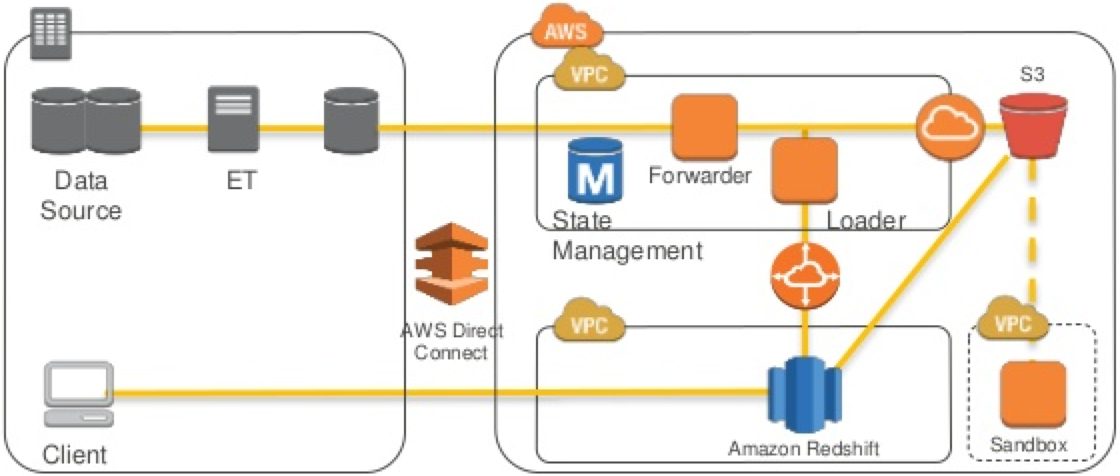
Scaling challenges

Performance issues

Need same level of security

Need for a hybrid environment

NTT Docomo: Japan's largest mobile service provider



125 node DS2.8XL cluster

4,500 vCPUs, 30 TB RAM

2 PB compressed

10x faster analytic queries

50% reduction in time for new BI application deployment

Significantly less operations overhead

Nasdaq: powering 100 marketplaces in 50 countries

Nasdaq

Orders, quotes, trade executions, market “tick” data from 7 exchanges

7 billion rows/day

Analyze market share, client activity, surveillance, billing, and so on

Microsoft SQL Server on-premises

Expensive legacy DW ($1.16 M/yr.)

Limited capacity (1 yr. of data online)

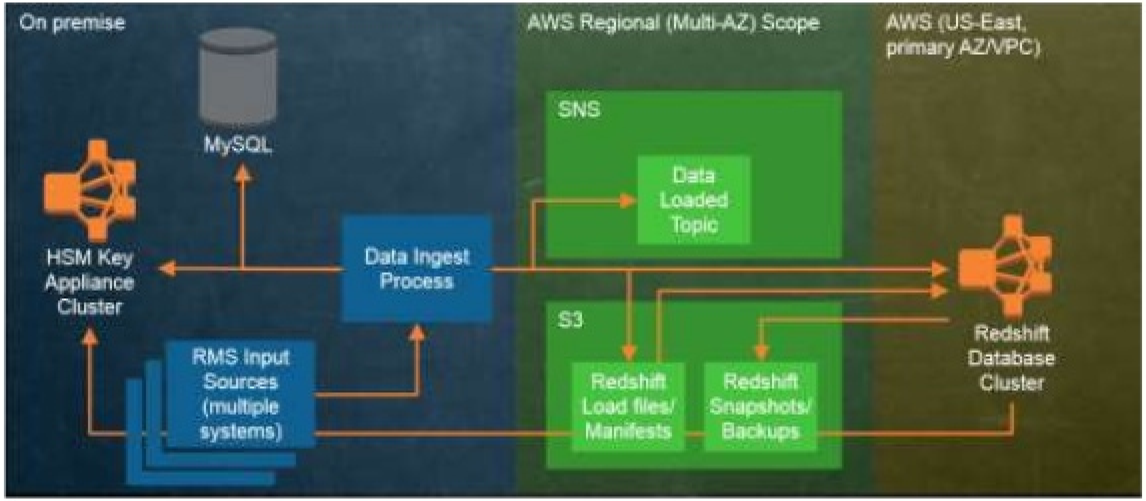
Needed lower TCO

Must satisfy multiple security and regulatory requirements

Similar performance

Nasdaq: powering 100 marketplaces in 50 countries

Nasdaq


On premise
AWS Regional (Multi-AZ) Scope
AWS (US-East, primary AZ/VPC)
MySQL
HSM Key Appliance Cluster
Data Ingest Process
RMS Input Sources (multiple systems)
SNS
Data Loaded Topic
S3
Redshift Load files/ Manifests
Redshift Snapshots/ Backups
Redshift Database Cluster

23 node DS2.8XL cluster

828 vCPUs, 5 TB RAM

368 TB compressed

2.7 T rows, 900 B derived

8 tables with 100 B rows

7 man-month migration

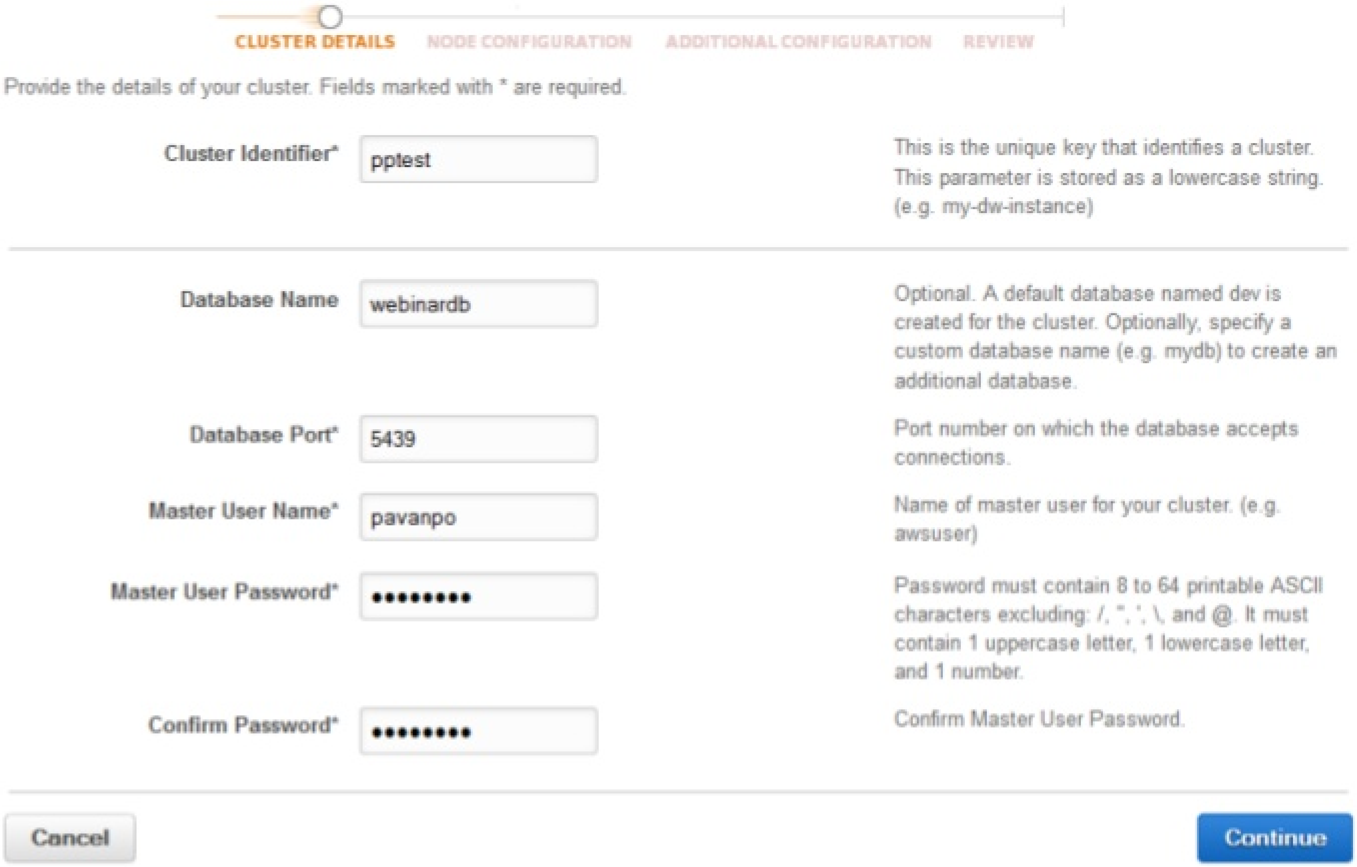
¼ the cost, 2x storage, room to grow

Faster performance, very secure

Getting started

Provisioning

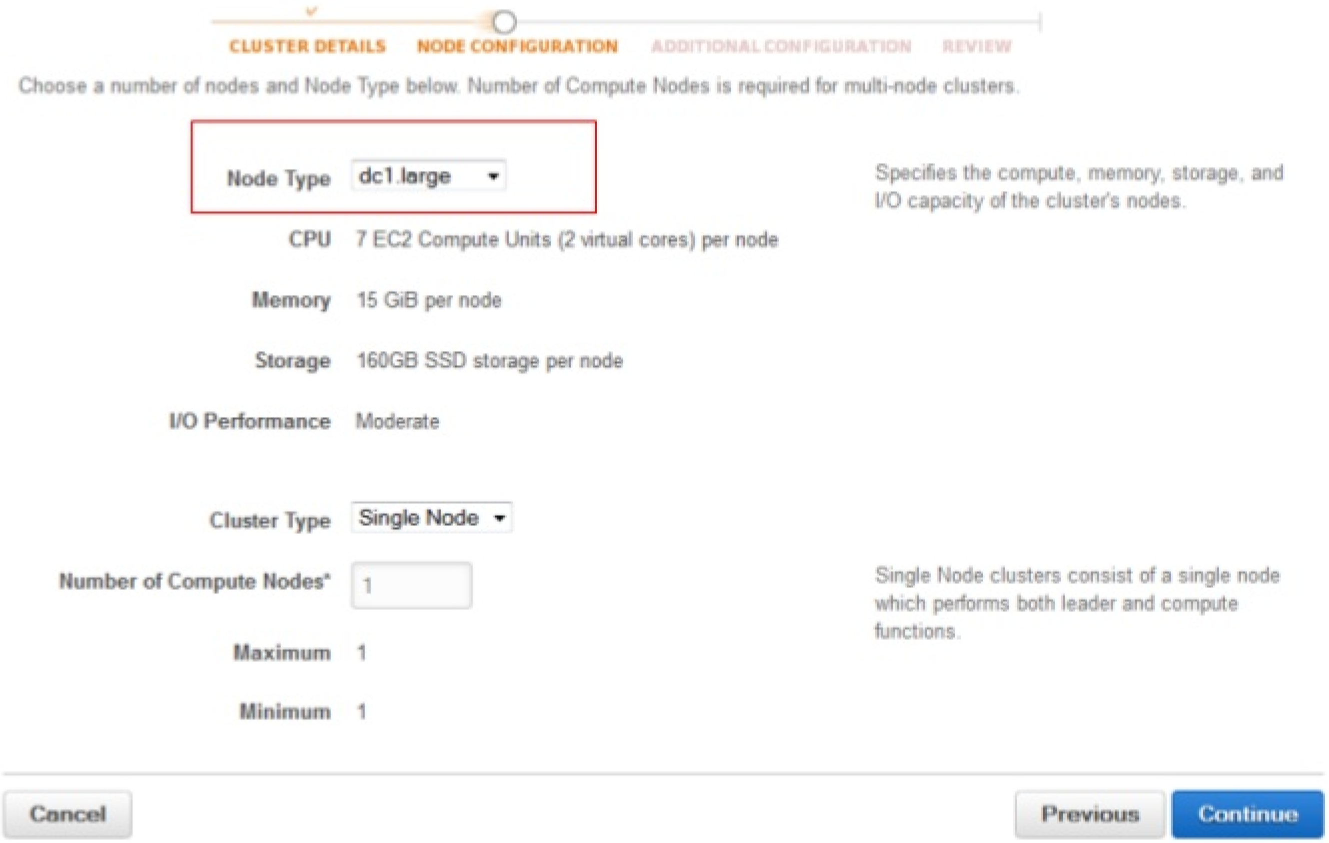
Enter cluster details



Provide the details of your cluster. Fields marked with * are required.

Cluster Identifier*	pptest	This is the unique key that identifies a cluster. This parameter is stored as a lowercase string. (e.g. my-dw-instance)
Database Name	webinardb	Optional. A default database named dev is created for the cluster. Optionally, specify a custom database name (e.g. mydb) to create an additional database.
Database Port*	5439	Port number on which the database accepts connections.
Master User Name*	pavanpo	Name of master user for your cluster. (e.g. awsuser)
Master User Password*	•••••••••	Password must contain 8 to 64 printable ASCII characters excluding: /, ", ', \, and @. It must contain 1 uppercase letter, 1 lowercase letter, and 1 number.
Confirm Password*	•••••••••	Confirm Master User Password.

Select node configuration



Choose a number of nodes and Node Type below. Number of Compute Nodes is required for multi-node clusters.

Node Type: dc1.large

Specifies the compute, memory, storage, and I/O capacity of the cluster's nodes.

CPU: 7 EC2 Compute Units (2 virtual cores) per node

Memory: 15 GiB per node

Storage: 160GB SSD storage per node

I/O Performance: Moderate

Cluster Type: Single Node

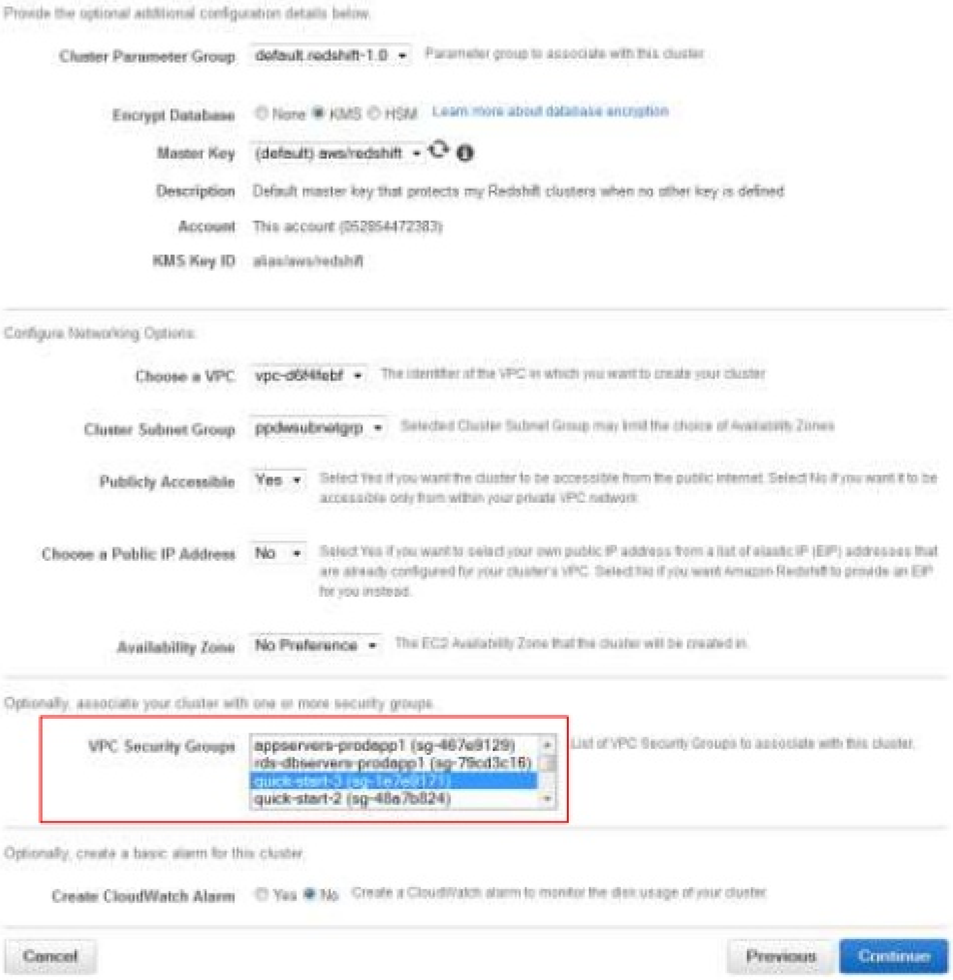
Number of Compute Nodes*: 1

Single Node clusters consist of a single node which performs both leader and compute functions.

Maximum: 1

Minimum: 1

Select security settings and provision



Provide the optional additional configuration details below.

Cluster Parameter Group: default.redshift-1.0 ▾ Parameter group to associate with this cluster

Encrypt Database: ○ None ● KMS ○ HSM Learn more about database encryption

Master Key: (default) aws/redshift ▾

Description: Default master key that protects my Redshift clusters when no other key is defined

Account: This account (052854472383)

KMS Key ID: alias/aws/redshift

Configure Networking Options:

Choose a VPC: vpc-d6f4febf ▾ The identifier of the VPC in which you want to create your cluster

Cluster Subnet Group: ppdwsubnetgrp ▾ Selected Cluster Subnet Group may limit the choice of Availability Zones

Publicly Accessible: Yes ▾ Select Yes if you want the cluster to be accessible from the public internet. Select No if you want it to be accessible only from within your private VPC network.

Choose a Public IP Address: No ▾ Select Yes if you want to select your own public IP address from a list of elastic IP (EIP) addresses that are already configured for your cluster's VPC. Select No if you want Amazon Redshift to provide an EIP for you instead.

Availability Zone: No Preference ▾ The EC2 Availability Zone that the cluster will be created in.

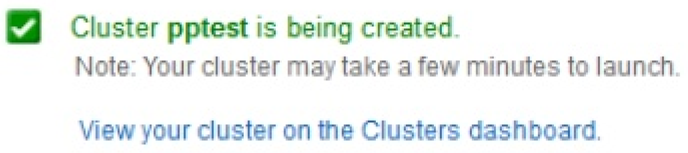
Optionally, associate your cluster with one or more security groups.

VPC Security Groups:
- appservers-prodapp1 (sg-467a9129)
- rds-dbservers-prodapp1 (sg-79cd3c16)
- quick-start-3 (sg-1e7e8171)
- quick-start-2 (sg-48a7b824)

List of VPC Security Groups to associate with this cluster.

Optionally, create a basic alarm for this cluster.

Create CloudWatch Alarm: ○ Yes ● No Create a CloudWatch alarm to monitor the disk usage of your cluster.

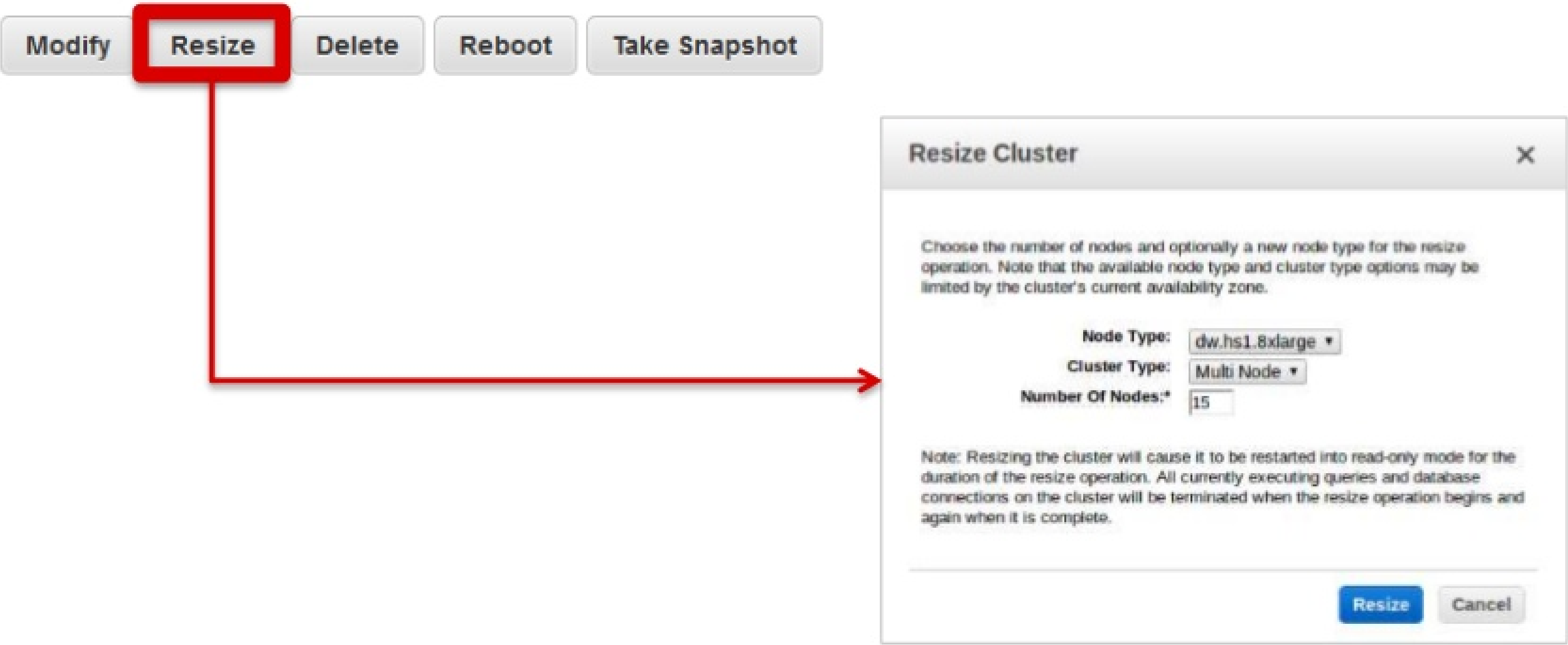


✓ Cluster **pptest** is being created.

Note: Your cluster may take a few minutes to launch.

View your cluster on the Clusters dashboard.

Point-and-click resize



Resize Cluster

Choose the number of nodes and optionally a new node type for the resize operation. Note that the available node type and cluster type options may be limited by the cluster's current availability zone.

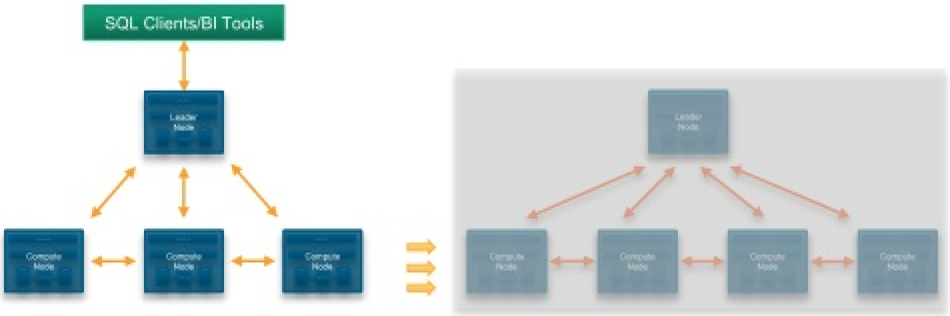
Node Type: dw.hs1.8xlarge ▾

Cluster Type: Multi Node ▾

Number Of Nodes:* 15

Note: Resizing the cluster will cause it to be restarted into read-only mode for the duration of the resize operation. All currently executing queries and database connections on the cluster will be terminated when the resize operation begins and again when it is complete.

Resize



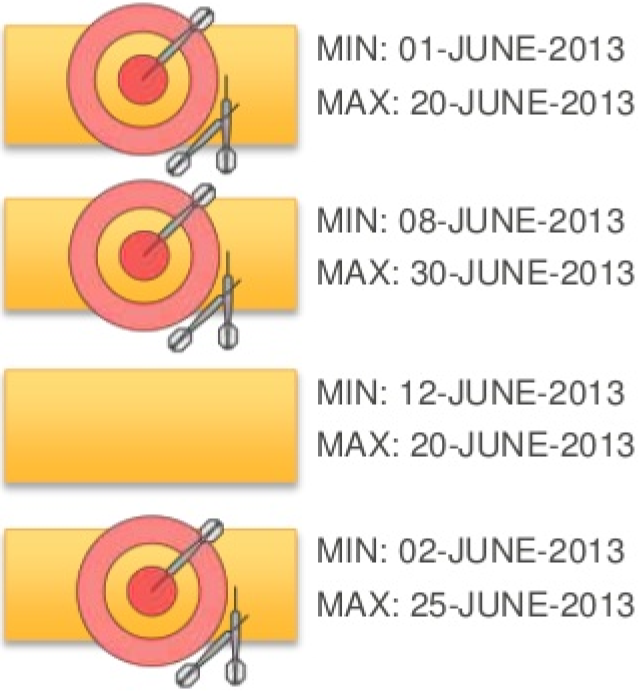
- Resize while remaining online
- Provision a new cluster in the background
- Copy data in parallel from node to node
- Only charged for source cluster

Data modeling

Zone maps

SELECT COUNT(*) FROM LOGS WHERE DATE = '09-JUNE-2013'

Unsorted table



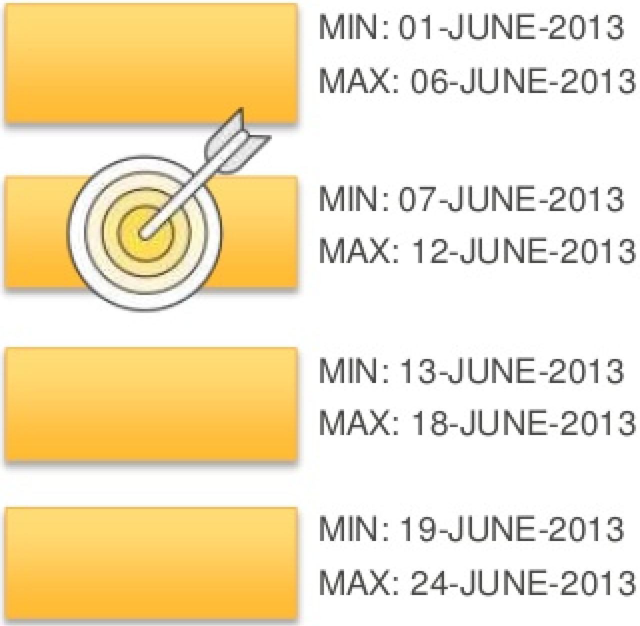
MIN: 01-JUNE-2013
MAX: 20-JUNE-2013

MIN: 08-JUNE-2013
MAX: 30-JUNE-2013

MIN: 12-JUNE-2013
MAX: 20-JUNE-2013

MIN: 02-JUNE-2013
MAX: 25-JUNE-2013

Sorted by date



MIN: 01-JUNE-2013
MAX: 06-JUNE-2013

MIN: 07-JUNE-2013
MAX: 12-JUNE-2013

MIN: 13-JUNE-2013
MAX: 18-JUNE-2013

MIN: 19-JUNE-2013
MAX: 24-JUNE-2013

Sort Keys

- Single column
- Compound
- Interleaved

Single Column

```
[ SORTKEY ( date ) ]
```

Date	Region	Country
2-JUN-2015	Oceania	New Zealand
2-JUN-2015	Asia	Singapore
2-JUN-2015	Africa	Zaire
2-JUN-2015	Asia	Hong Kong
3-JUN-2015	Europe	Germany
3-JUN-2015	Asia	Korea

- Best for:
  - Queries that use 1st column (i.e. *date)* as primary filter
  - Can speed up joins and group bys
  - Quickest to VACUUM

Sort Keys

Compound

- Table is sorted by 1st column , then 2nd column etc.

```
[ SORTKEY COMPOUND ( date, region, country) ]
```

Date	Region	Country
2-JUN-2015	Africa	Zaire
2-JUN-2015	Asia	Korea
2-JUN-2015	Asia	Singapore
2-JUN-2015	Europe	Germany
3-JUN-2015	Asia	Hong Kong
3-JUN-2015	Asia	Korea

- Best for:
  - Queries that use 1st column as primary filter, then other cols
  - Can speed up joins and group bys
  - Slower to VACUUM

Interleaved

- Equal weight is given to each column.

```
[ SORTKEY INTERLEAVED ( date, region, country) ]
```

Date	Region	Country
2-JUN-2015	Africa	Zaire
3-JUN-2015	Asia	Singapore
2-JUN-2015	Asia	Korea
2-JUN-2015	Europe	Germany
3-JUN-2015	Asia	Hong Kong
2-JUN-2015	Asia	Korea

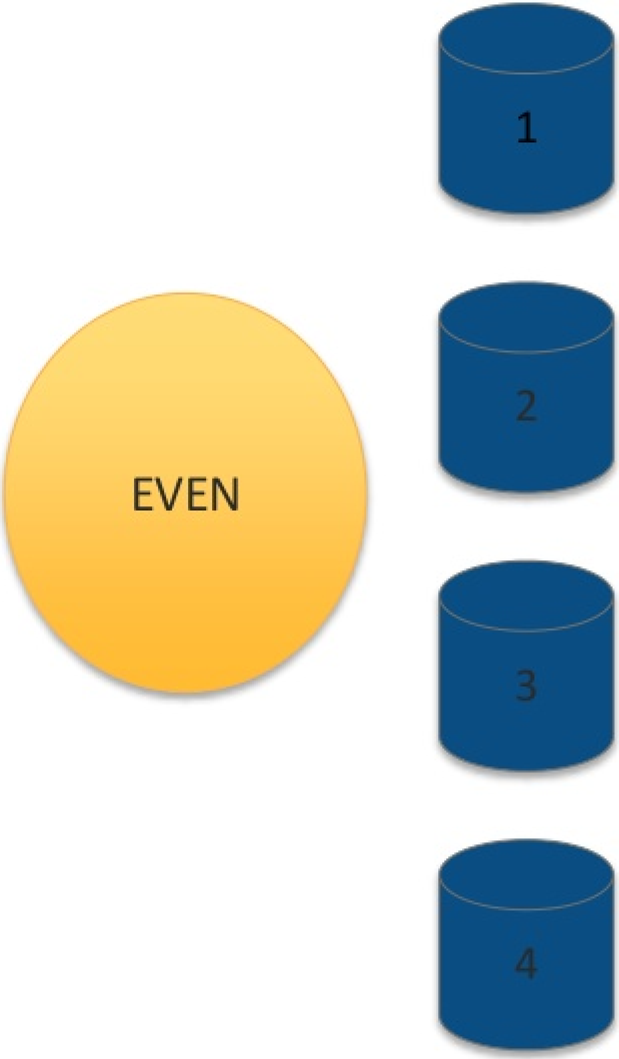
- Best for:
  - Queries that use different columns in filter
  - Queries get faster the more columns used in the filter
  - Slowest to VACUUM

Distribution

- EVEN
- KEY
- ALL

Distribution

ID	Gender	Name
101	M	John Smith
292	F	Jane Jones
139	M	Peter Black
446	M	Pat Partridge
658	F	Sarah Cyan
164	M	Brian Snail
209	M	James White
306	F	Lisa Green



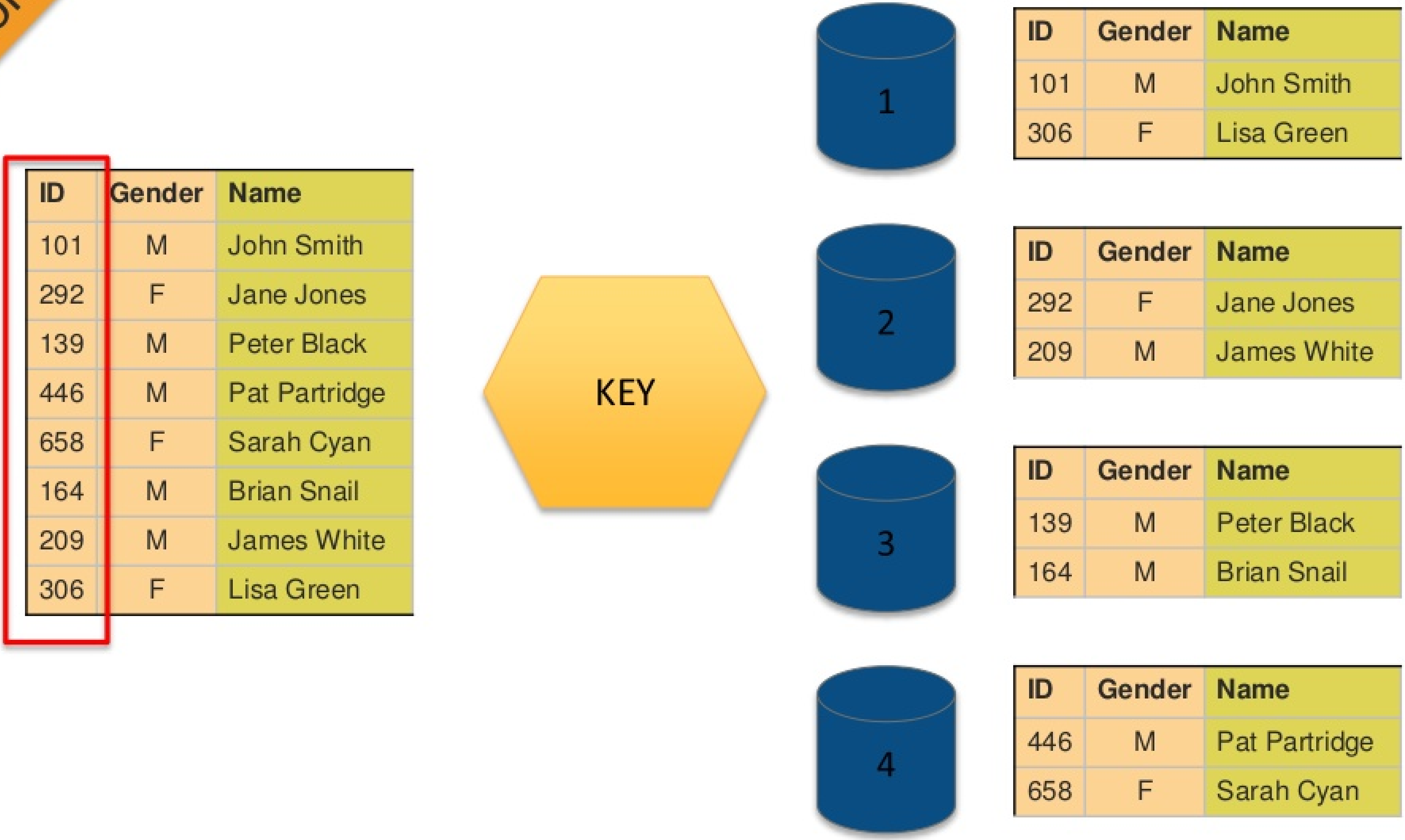
ID	Gender	Name
101	M	John Smith
306	F	Lisa Green

ID	Gender	Name
292	F	Jane Jones
209	M	James White

ID	Gender	Name
139	M	Peter Black
164	M	Brian Snail

ID	Gender	Name
446	M	Pat Partridge
658	F	Sarah Cyan

Distribution



ID	Gender	Name
101	M	John Smith
292	F	Jane Jones
139	M	Peter Black
446	M	Pat Partridge
658	F	Sarah Cyan
164	M	Brian Snail
209	M	James White
306	F	Lisa Green

ID	Gender	Name
101	M	John Smith
306	F	Lisa Green

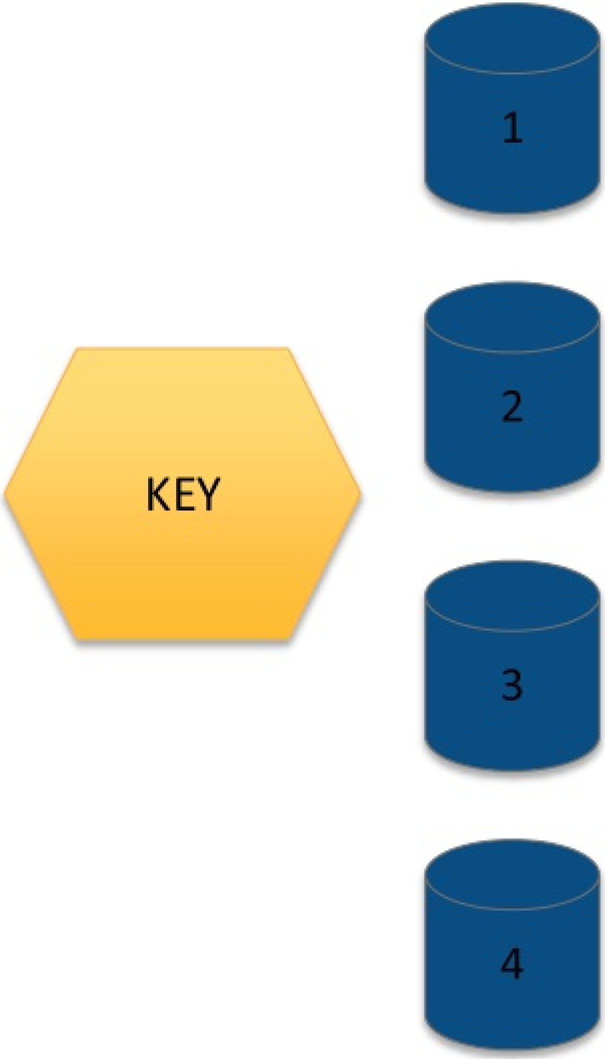
ID	Gender	Name
292	F	Jane Jones
209	M	James White

ID	Gender	Name
139	M	Peter Black
164	M	Brian Snail

ID	Gender	Name
446	M	Pat Partridge
658	F	Sarah Cyan

Distribution

ID	Gender	Name
101	M	John Smith
292	F	Jane Jones
139	M	Peter Black
446	M	Pat Partridge
658	F	Sarah Cyan
164	M	Brian Snail
209	M	James White
306	F	Lisa Green

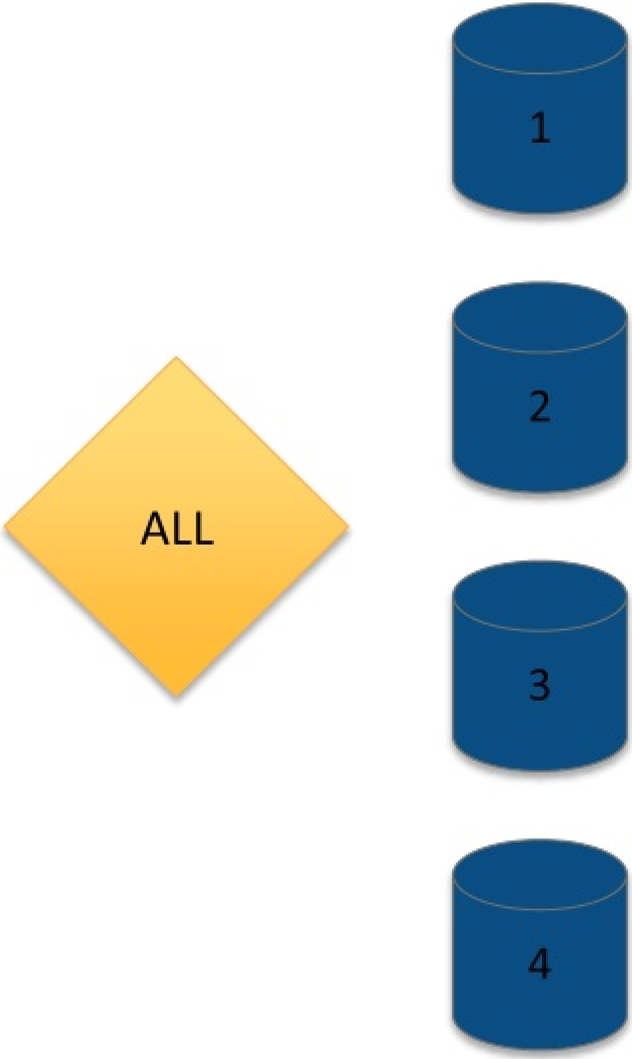


ID	Gender	Name
101	M	John Smith
139	M	Peter Black
446	M	Pat Partridge
164	M	Brian Snail
209	M	James White

ID	Gender	Name
292	F	Jane Jones
658	F	Sarah Cyan
306	F	Lisa Green

Distribution

ID	Gender	Name
101	M	John Smith
292	F	Jane Jones
139	M	Peter Black
446	M	Pat Partridge
658	F	Sarah Cyan
164	M	Brian Snail
209	M	James White
306	F	Lisa Green



101	M	John Smith
292	F	Jane Jones
139	M	Peter Black
446	M	Pat Partridge
658	F	Sarah Cyan
164	M	Brian Snail
209	M	Lisa Green
306	F	James White

101	M	John Smith
292	F	Jane Jones
139	M	Peter Black
446	M	Pat Partridge
658	F	Sarah Cyan
164	M	Brian Snail
209	M	Lisa Green
306	F	James White

101	M	John Smith
292	F	Jane Jones
139	M	Peter Black
446	M	Pat Partridge
658	F	Sarah Cyan
164	M	Brian Snail
209	M	Lisa Green
306	F	James White

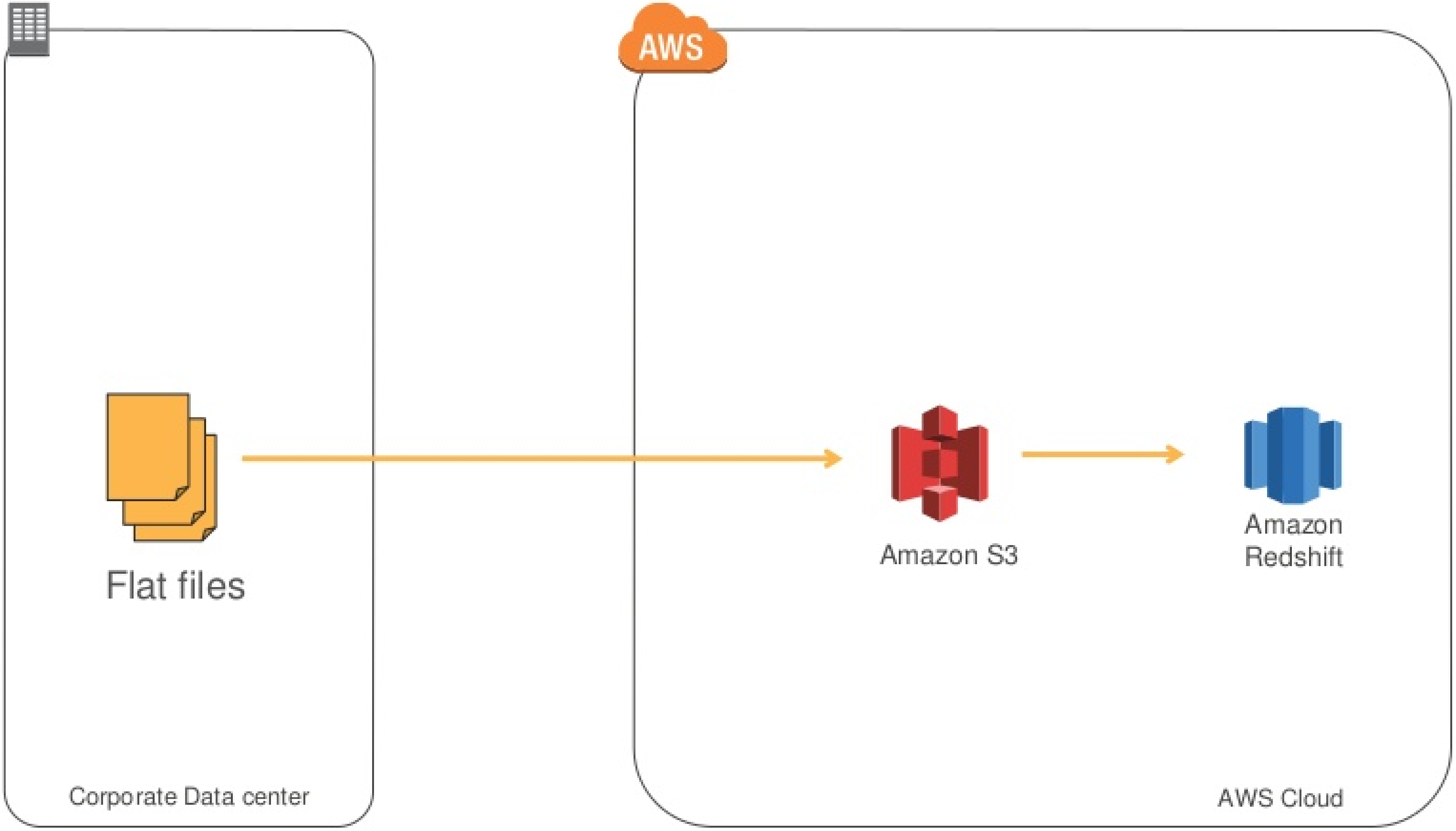
101	M	John Smith
292	F	Jane Jones
139	M	Peter Black
446	M	Pat Partridge
658	F	Sarah Cyan
164	M	Brian Snail
209	M	Lisa Green
306	F	James White

Distribution

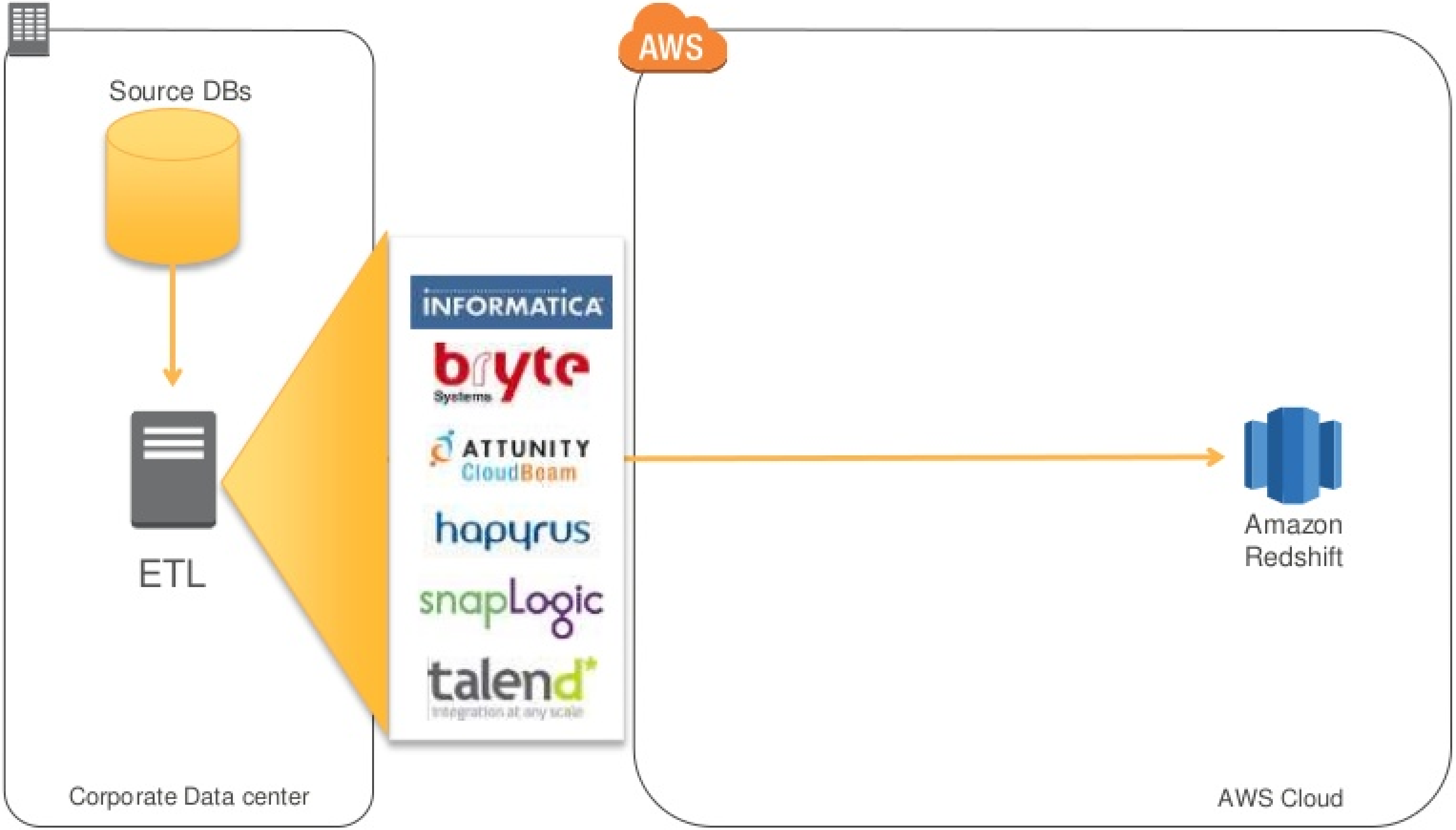
- EVEN
  - Tables with no joins or group by
- KEY
  - Large Fact tables
  - Large dimension tables
- ALL
  - Medium dimension tables (1K – 2M)
  - Small dimension tables

Loading data

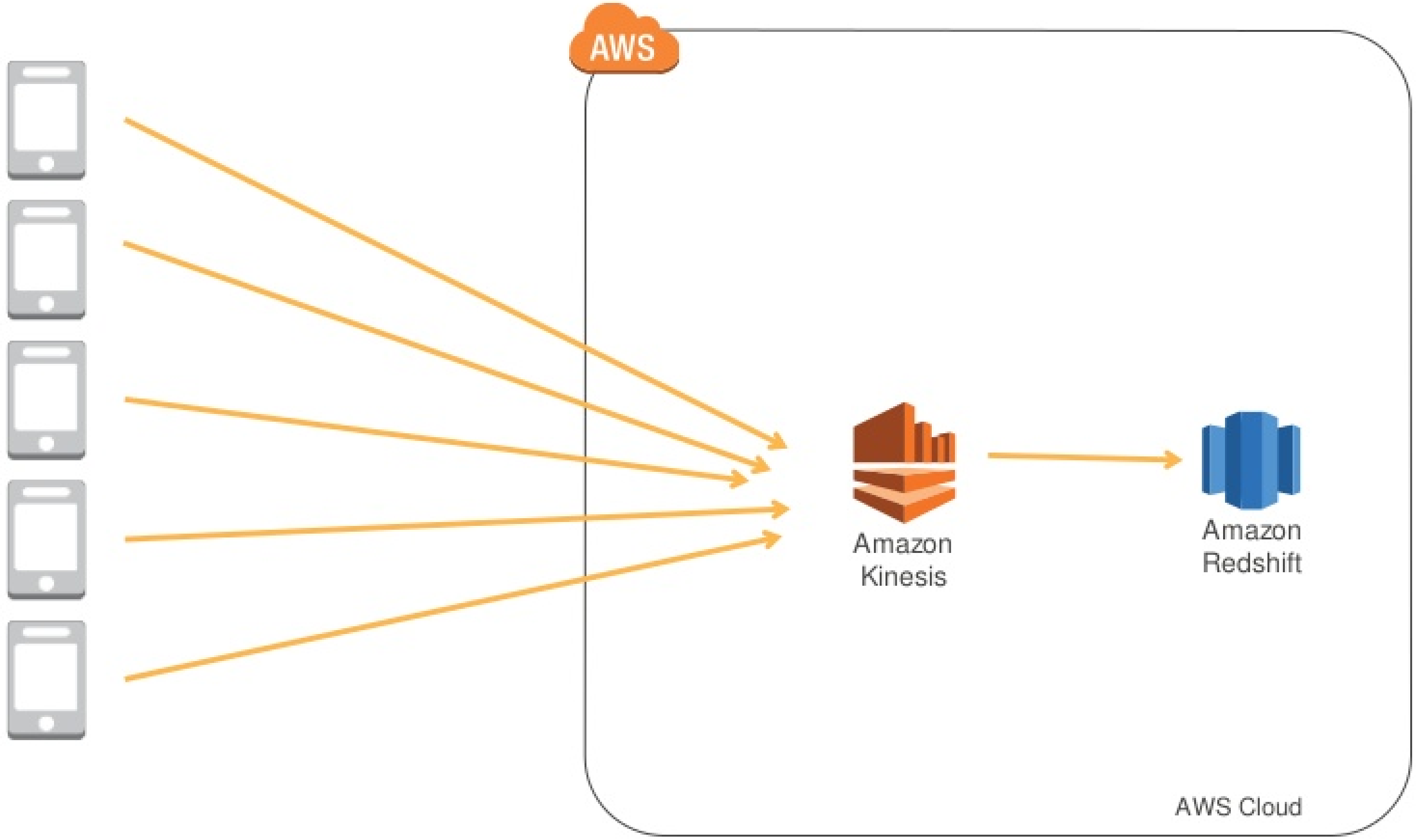
Data loading options



Data loading options



Data loading options



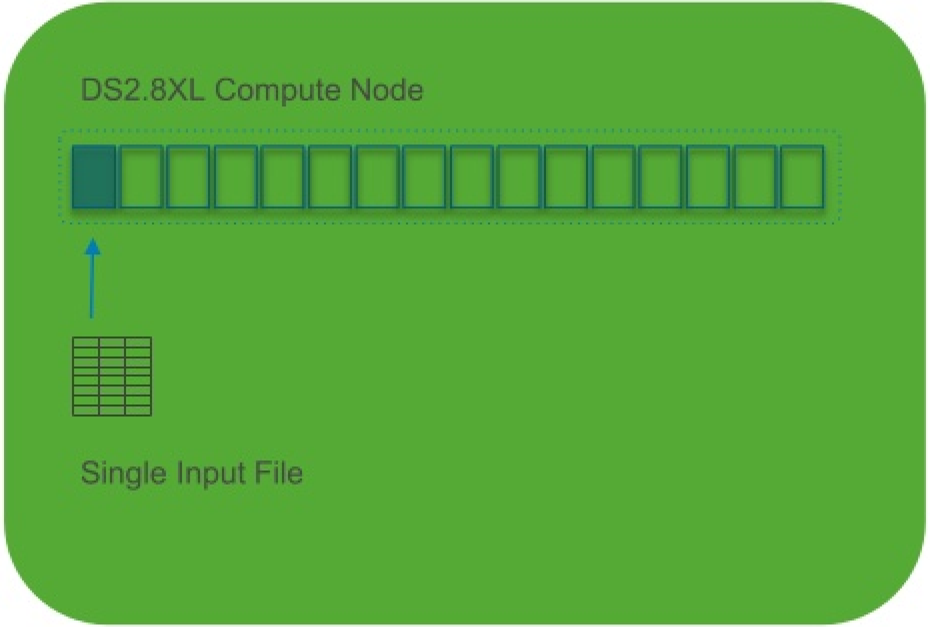
Use multiple input files to maximize throughput

Use the COPY command

Each slice can load one file at a time

A single input file means only one slice is ingesting data

Instead of 100MB/s, you're only getting 6.25MB/s



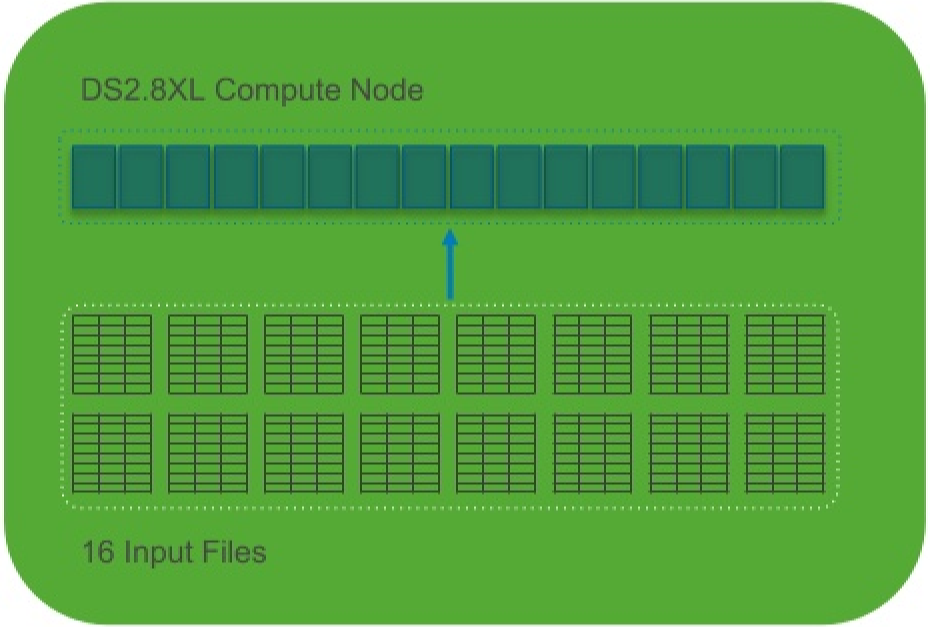
Use multiple input files to maximize throughput

Use the COPY command

You need at least as many input files as you have slices

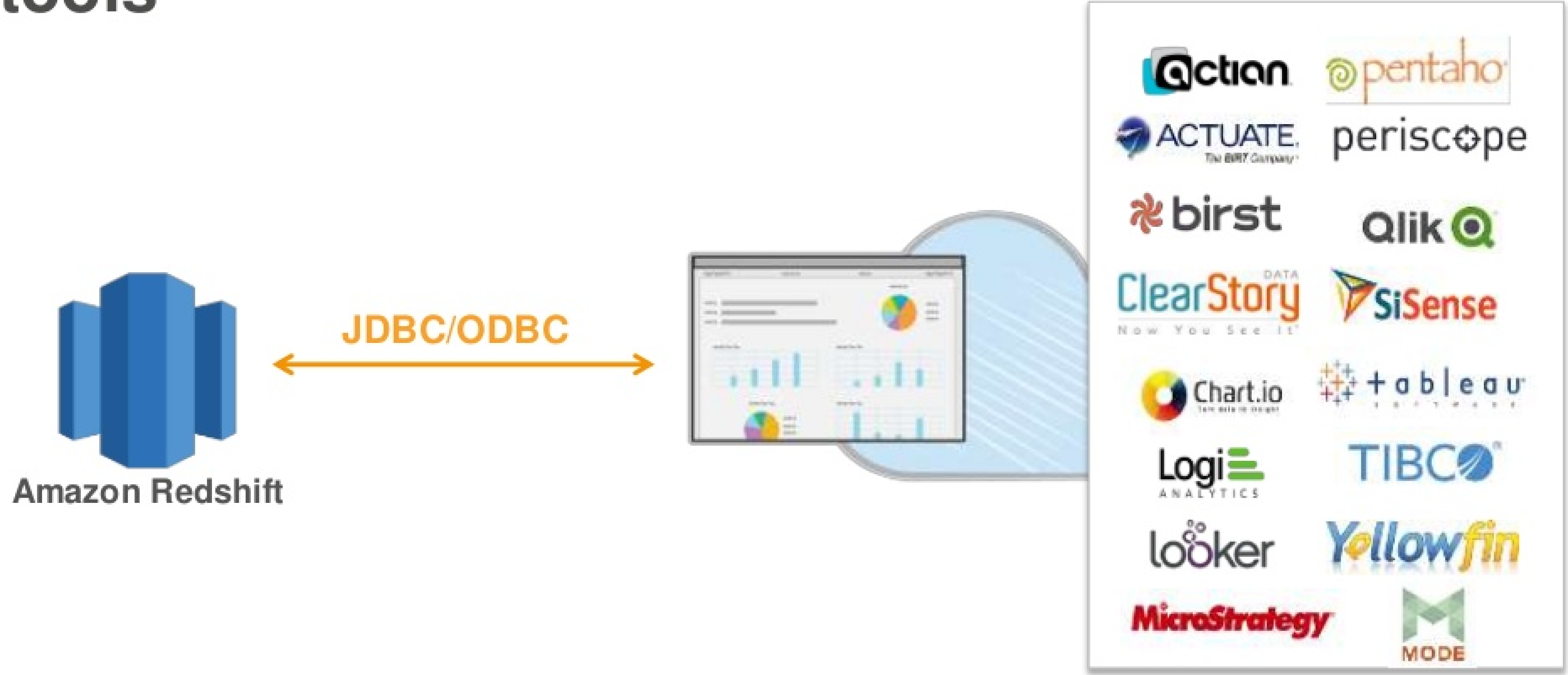
With 16 input files, all slices are working so you maximize throughput

Get 100MB/s per node; scale linearly as you add nodes



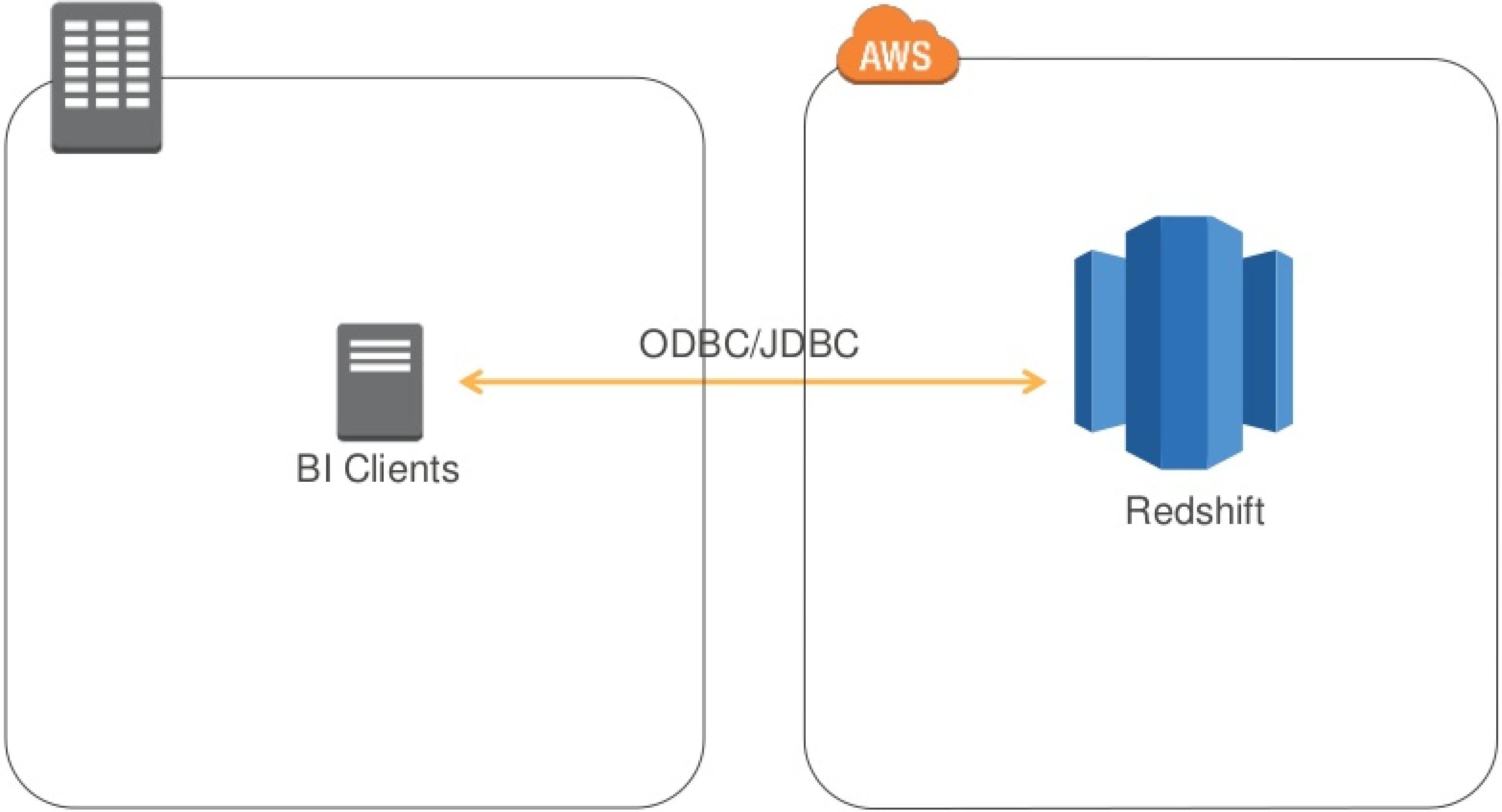
Querying

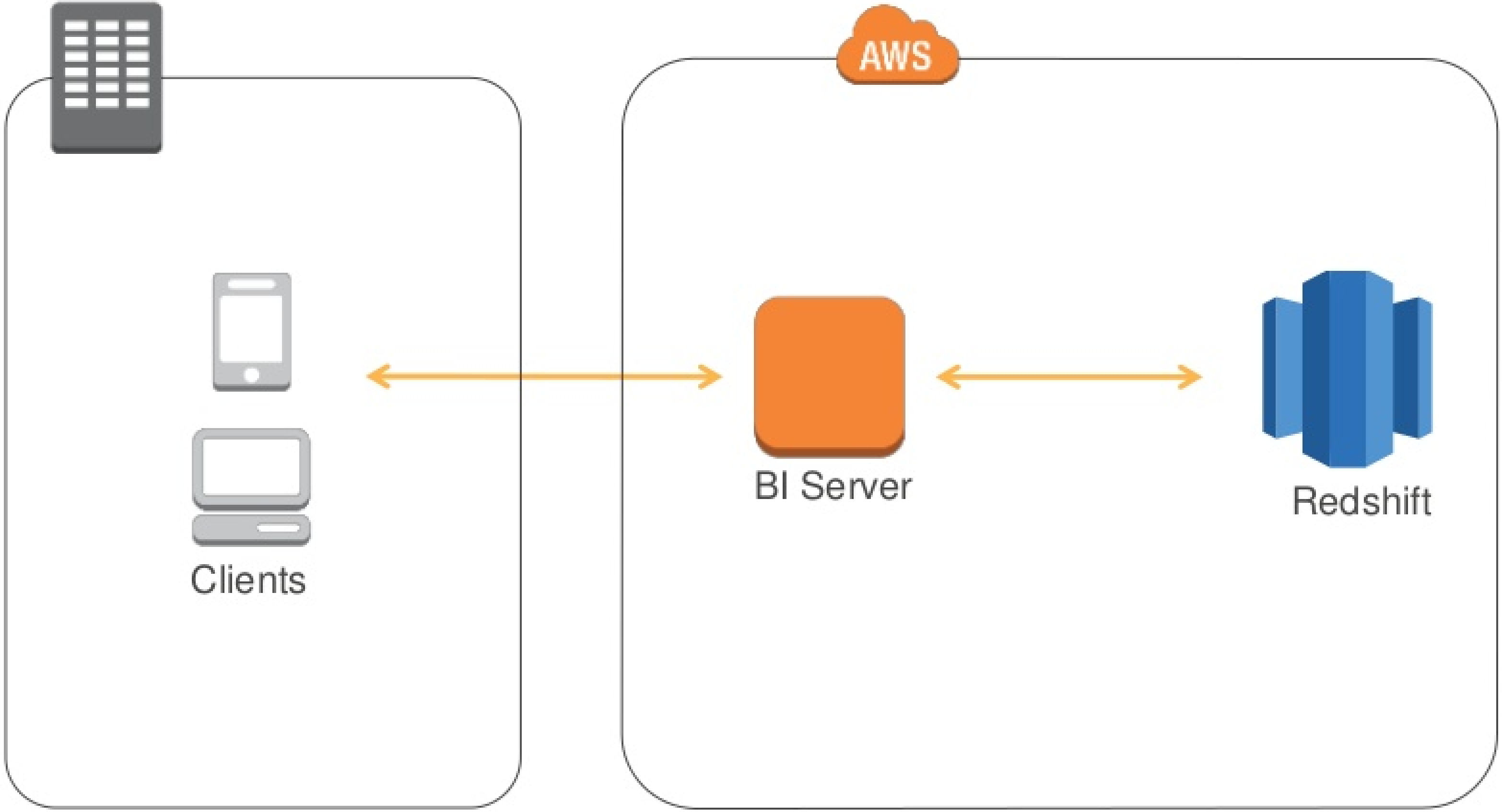
Amazon Redshift works with your existing BI tools



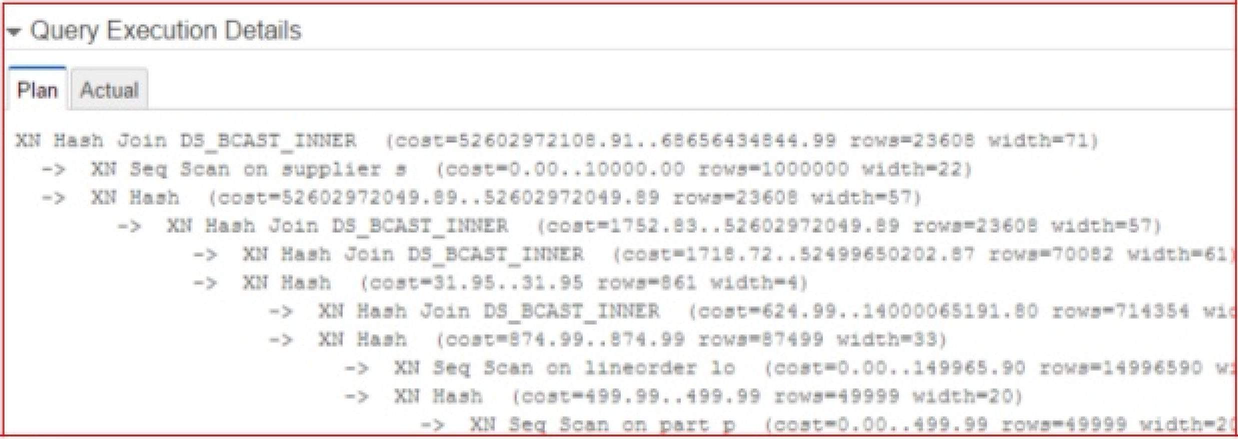
Amazon Redshift

JDBC/ODBC

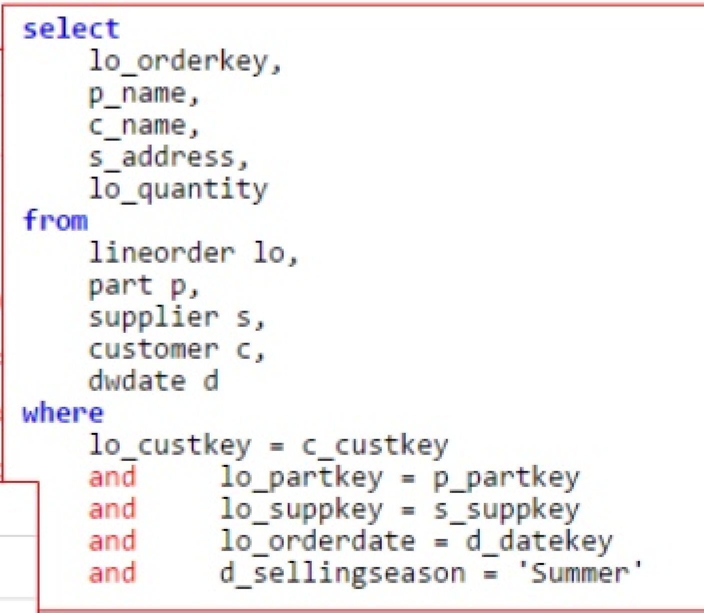

AWS
ODBC/JDBC
BI Clients
Redshift


AWS
Clients
BI Server
Redshift

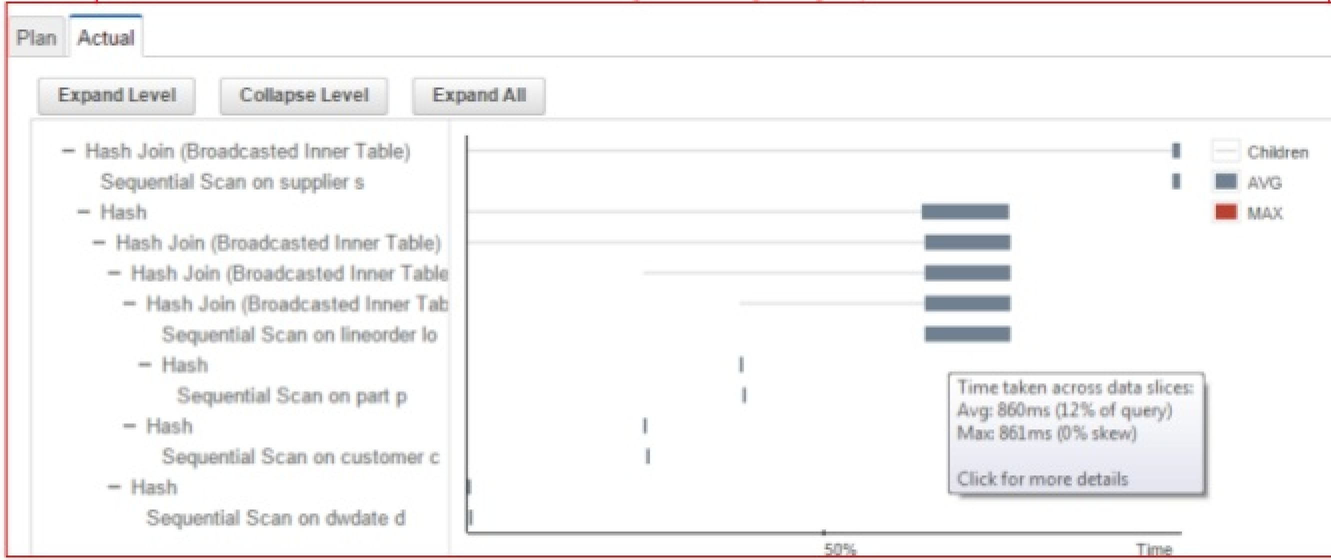
View explain plans



```
XN Hash Join DS_BCAST_INNER  (cost=52602972108.91..68656434844.99 rows=23608 width=71)
  ->  XN Seq Scan on supplier s  (cost=0.00..10000.00 rows=1000000 width=22)
  ->  XN Hash  (cost=52602972049.89..52602972049.89 rows=23608 width=57)
        ->  XN Hash Join DS_BCAST_INNER  (cost=1752.83..52602972049.89 rows=23608 width=57)
              ->  XN Hash Join DS_BCAST_INNER  (cost=1718.72..52499650202.87 rows=70082 width=61)
              ->  XN Hash  (cost=31.95..31.95 rows=861 width=4)
                    ->  XN Hash Join DS_BCAST_INNER  (cost=624.99..14000065191.80 rows=714354 wi
                    ->  XN Hash  (cost=874.99..874.99 rows=87499 width=33)
                          ->  XN Seq Scan on lineorder lo  (cost=0.00..149965.90 rows=14996590 wi
                          ->  XN Hash  (cost=499.99..499.99 rows=49999 width=20)
                                ->  XN Seq Scan on part p  (cost=0.00..499.99 rows=49999 width=2
```



```
select
    lo_orderkey,
    p_name,
    c_name,
    s_address,
    lo_quantity
from
    lineorder lo,
    part p,
    supplier s,
    customer c,
    dwdate d
where
    lo_custkey = c_custkey
    and     lo_partkey = p_partkey
    and     lo_suppkey = s_suppkey
    and     lo_orderdate = d_datekey
    and     d_sellingseason = 'Summer'
```

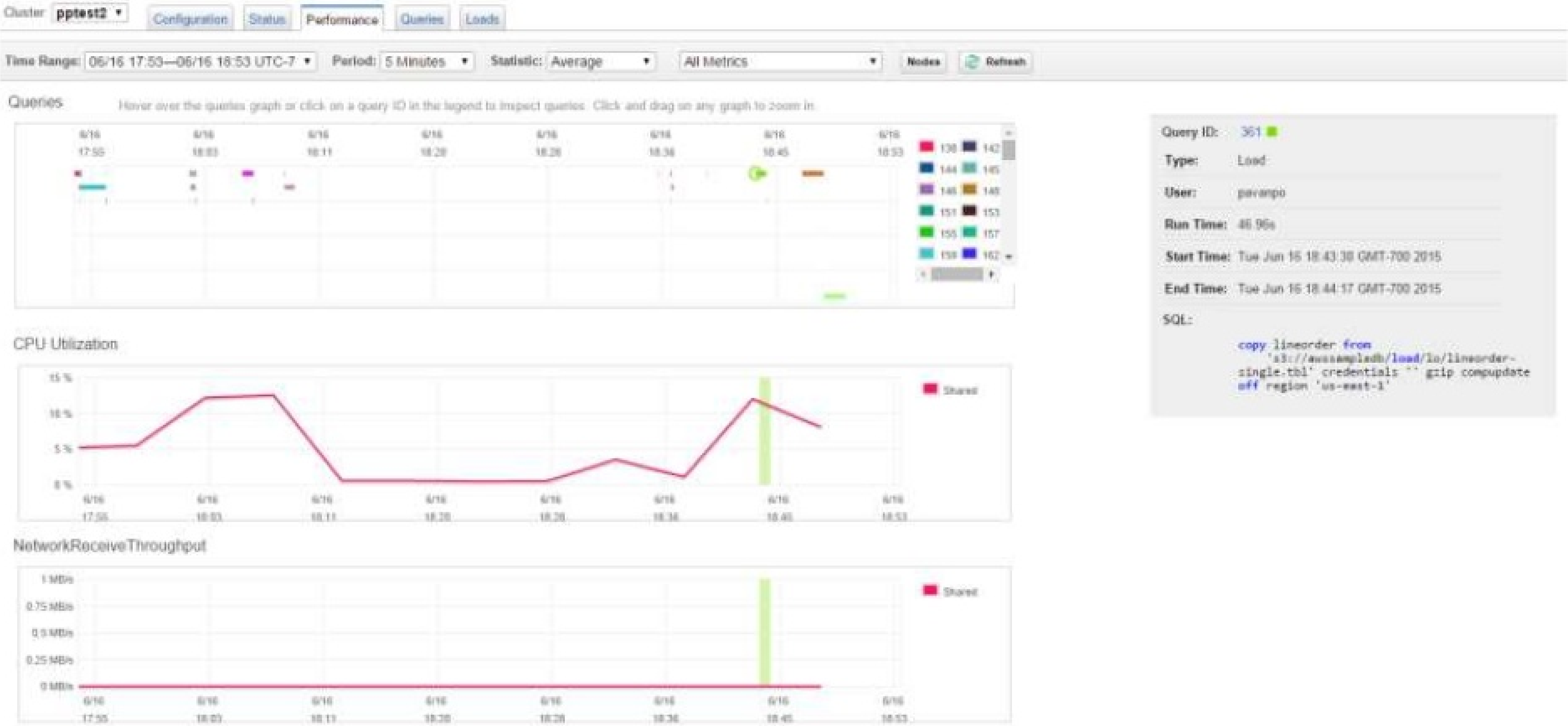


Plan
Actual
Expand Level
Collapse Level
Expand All
− Hash Join (Broadcasted Inner Table)
Sequential Scan on supplier s
− Hash
− Hash Join (Broadcasted Inner Table)
− Hash Join (Broadcasted Inner Table
− Hash Join (Broadcasted Inner Tab
Sequential Scan on lineorder lo
− Hash
Sequential Scan on part p
− Hash
Sequential Scan on customer c
− Hash
Sequential Scan on dwdate d
Children
AVG
MAX
50%
Time

Time taken across data slices:
Avg: 860ms (12% of query)
Max: 861ms (0% skew)

Click for more details

Monitor query performance



Cluster: pptest2

Time Range: 06/16 17:53—06/16 18:53 UTC-7

Period: 5 Minutes

Statistic: Average

All Metrics

Queries

Hover over the queries graph or click on a query ID in the legend to inspect queries. Click and drag on any graph to zoom in.

Query ID: 361

Type: Load

User: pavanpo

Run Time: 46.96s

Start Time: Tue Jun 16 18:43:30 GMT-700 2015

End Time: Tue Jun 16 18:44:17 GMT-700 2015

SQL:

```
copy lineorder from
    's3://awssampledb/load/lo/lineorder-
single.tbl' credentials '' gzip compupdate
off region 'us-east-1'
```

Resources

Detail Pages

- http://aws.amazon.com/redshift
- https://aws.amazon.com/marketplace/redshift/
- Amazon Redshift Utilities - GitHub

Best Practices

- http://docs.aws.amazon.com/redshift/latest/dg/c_loading-data-best-practices.html
- http://docs.aws.amazon.com/redshift/latest/dg/c_designing-tables-best-practices.html
- http://docs.aws.amazon.com/redshift/latest/dg/c-optimizing-query-performance.html

amazon web services | Pop-up Loft

Thank you!