Amazon Web Services - Beyond EC2 and S3

DynamoDB, Redshift, Kinesis Streams and Lambda

Markus Dale, November 2015

AWS: Managed Services With Rich Environment

- EC2 - Elastic Load Balancer service, Auto-Scaling, Security groups...
- S3 - Server-Side Encryption, Versioning, Notification...
- RDS - Relational Database Service...

Developer Productivity

Right Tool for the Right Job

AWS DynamoDB

- Managed NoSQL database since 2012
    - Dynamo key-value paper 2007

- Cost: Provisioned read/write throughput and actual storage
- Scalable (up and down)
- Fault-tolerant
- SSD storage

DynamoDB Tables

	Name	Status	Partition key	Sort key	Indexes	Total rea
○	Forum	Active	Name (String)	-	0	1
○	ProductCatalog	Active	Id (Number)	-	0	1
○	Reply	Active	Id (String)	ReplyDateTime (String)	0	1
○	Thread	Active	ForumName (String)	Subject (String)	0	1

DynamoDB Provisioned Capacity

Overview	Items	Metrics	Alarms	**Capacity**	Indexes	Triggers	Access control

Provisioned capacity

Read capacity units   `1`   Table

Write capacity units   `1`   Table

Estimated cost $0.59 / month (Capacity calculator)

Cancel   Save

DynamoDB Sample Table - Reply

Reply

| Overview | Items | Metrics | Alarms | Capacity | Indexes | Triggers | Access control |

Create item   Actions ∨

Scan: [Table] Reply: Id, ReplyDateTime ∨

	Id	ReplyDateTime	Message	PostedBy
☐	Amazon DynamoDB#DynamoDB Thread 1	2015-11-10 05:02:08	DynamoDB Thread 1 Reply 2 text	User B
☐	Amazon DynamoDB#DynamoDB Thread 2	2015-11-03 05:02:08	DynamoDB Thread 2 Reply 3 text	User B
☐	Amazon DynamoDB#DynamoDB Thread 2	2015-11-17 05:02:08	DynamoDB Thread 2 Reply 2 text	User A
☐	Amazon DynamoDB#DynamoDB Thread 2	2015-11-23 05:02:08	DynamoDB Thread 2 Reply 1 text	User A

DynamoDB Query
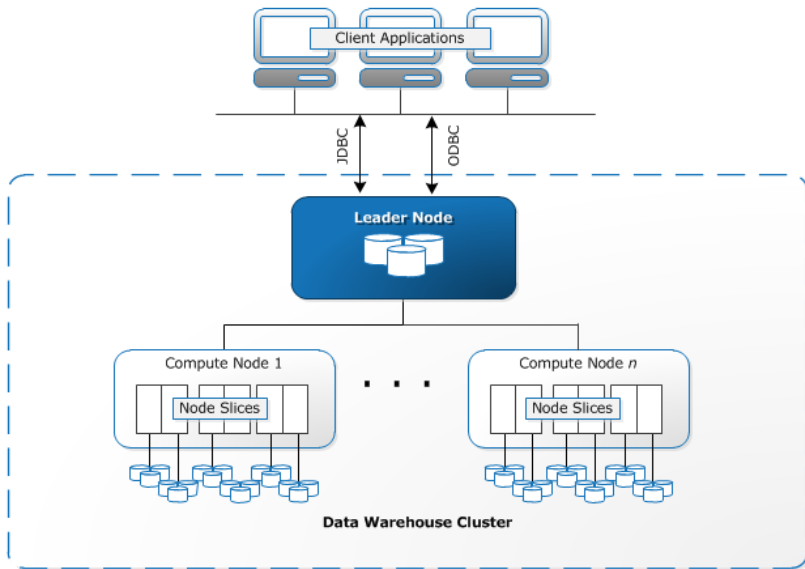
Amazon Redshift

- Fully managed, petabyte-scale data warehouse
- Online Analytic Processing (OLAP) and Business Intelligence
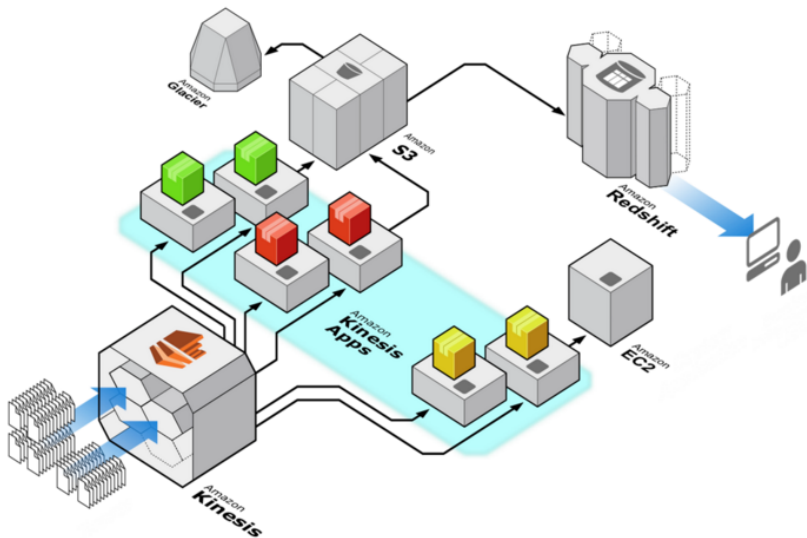- Single or multi-node cluster in a single AZ

Redshift Architecture

Redshift Integration

- Load/store in S3
- DynamoDB
- Elastic MapReduce (EMR)
- Kinesis Streams

AWS Kinesis Streams

Kinesis Streams Scaling and Data Records

- Shards: 1 shard = 1 MB/second write, 2MB/second reads, 1000 puts/second
- Sequence Number (assigned by Kinesis)
- Partition Key (user assigned)
- Data blob

AWS Lambda

- Event-driven compute
- Execute without provisioning/managing server
- Managed execution and scaling
- Python, Java, Node.js (JavaScript) - can call libraries, other languages

AWS Lambda Triggers Examples

- DynamoDB Table Update
- S3 Object modifications
- Amazon CloudWatch log entry
- Simple Email Service incoming email
- Kinesis Streams message
- Cron schedule

Lambda Execution

- Stateless
- Charged by 100ms increments, number of requests, GB of RAM allocated
- Default: 3sec execution limit (max 300sec)
- Default: 100 simultaneous executions (can increase)
- IAM policies/roles to manage permissions