AWS Certified Developer Associate

Lesson 6: AWS DynamoDB

simplilearn

What You'll Learn

- Describe the basic concepts of DynamoDB
- Identify the core components of DynamoDB
- Define Primary and Secondary Indexes
- Identify the purpose of Query and Scan Operations
- Describe Read/Write Operations
- Identify the usage of Atomic Counter
- Calculate provisioned throughput
- Identify the procedure for Data Import/Export using Data Pipeline

Basic Concepts of DynamoDB

DynamoDB Overview

Amazon DynamoDB

Fully Managed NOSQL Database Service

- Provides fast single-digit milliseconds latency performance
- Easy administration and cost effective
- Supports partitioning data over multiple instances
- Integrates with AWS IAM for fine grain access control
- Great fit for mobile, gaming, and Internet of Things

Components of DynamoDB

{
“Student Id”:”: 101,
“First Name “John”,
“Last Name”: “Smith”,
“Phone”: “703 323 1456”
}
	{
	“Student Id”: 102,
	“First Name”: “Lee”,
	“Last Name”: “Chang”,
	“Address”: {
	“Street”: “131 Park Street”,
	“City”: “Vienna”,
	“State”: “VA”,
	“Zip”: 22180
	}
	}
		{
		“Student Id”: 103,
		“First Name”: “Annie”,
		“Last Name”: “Carter”,
		“Address”: {
		“Street”: “222 M Street”,
		“City”: “Fairfax”,
		“State”: “VA”,
		“Zip”: 22030
		},
		“Degree”: “BS”
		}

Tables, Items, and Attributes are the three core components

Components of DynamoDB

```
{
    “Student Id”:”: 101,
    “First Name “John”,
    “Last Name”: “Smith”,
     “Phone”: “703 323 1456”
}
```

```
{
    “Student Id”: 102,
    “First Name”: “Lee”,
    “Last Name”: “Chang”,
     “Address”: {
                “Street”: “131 Park Street”,
                 “City”: “Vienna”,
                 “State”: “VA”,
                  “Zip”: 22180
                  }
}
```

```
{
    “Student Id”: 103,
    “First Name”: “Annie”,
    “Last Name”: “Carter”,
     “Address”: {
                “Street”: “222 M Street”,
                 “City”: “Fairfax”,
                 “State”: “VA”,
                  “Zip”: 22030
                  },
    “Degree”: “BS”
}
```

Tables are used to store information/collection of data

Components of DynamoDB

{
“Student Id”:”: 101,
“First Name “John”,
“Last Name”: “Smith”,
“Phone”: “703 323 1456”
}
	{
	“Student Id”: 102,
	“First Name”: “Lee”,
	“Last Name”: “Chang”,
	“Address”: {
	“Street”: “131 Park Street”,
	“City”: “Vienna”,
	“State”: “VA”,
	“Zip”: 22180
	}
	}
		{
		“Student Id”: 103,
		“First Name”: “Annie”,
		“Last Name”: “Carter”,
		“Address”: {
		“Street”: “222 M Street”,
		“City”: “Fairfax”,
		“State”: “VA”,
		“Zip”: 22030
		},
		“Degree”: “BS”
		}

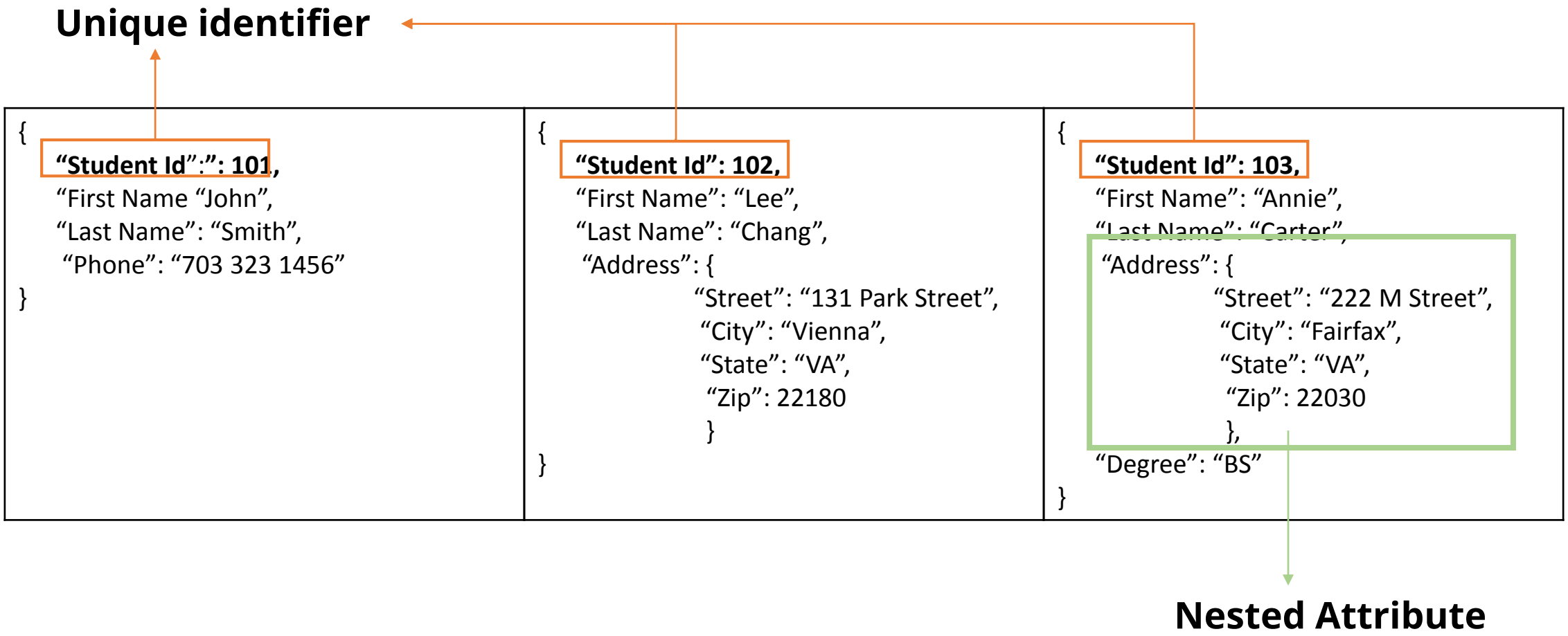
Each table is made of one or more items

Components of DynamoDB

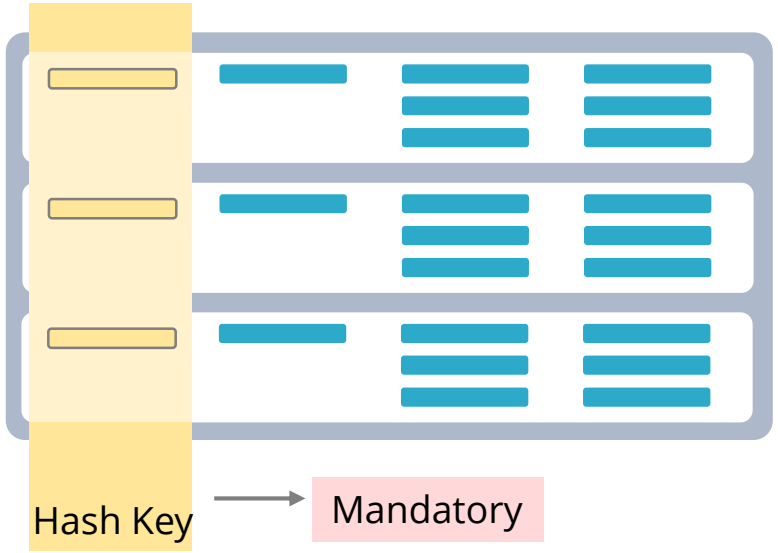
{
“Student Id”:”: 101,
“First Name “John”,
“Last Name”: “Smith”,
“Phone”: “703 323 1456”
}
	{
	“Student Id”: 102,
	“First Name”: “Lee”,
	“Last Name”: “Chang”,
	“Address”: {
	“Street”: “131 Park Street”,
	“City”: “Vienna”,
	“State”: “VA”,
	“Zip”: 22180
	}
	}
		{
		“Student Id”: 103,
		“First Name”: “Annie”,
		“Last Name” “Carter”,
		“Address”: {
		“Street”: “222 M Street”,
		“City”: “Fairfax”,
		“State”: “VA”,
		“Zip”: 22030
		},
		“Degree”: “BS”
		}

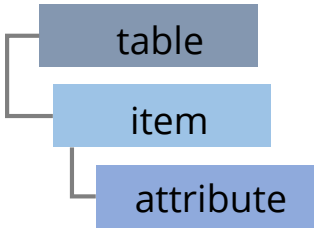
Attributes are basic data elements that provide value to items

How is DynamoDB different?



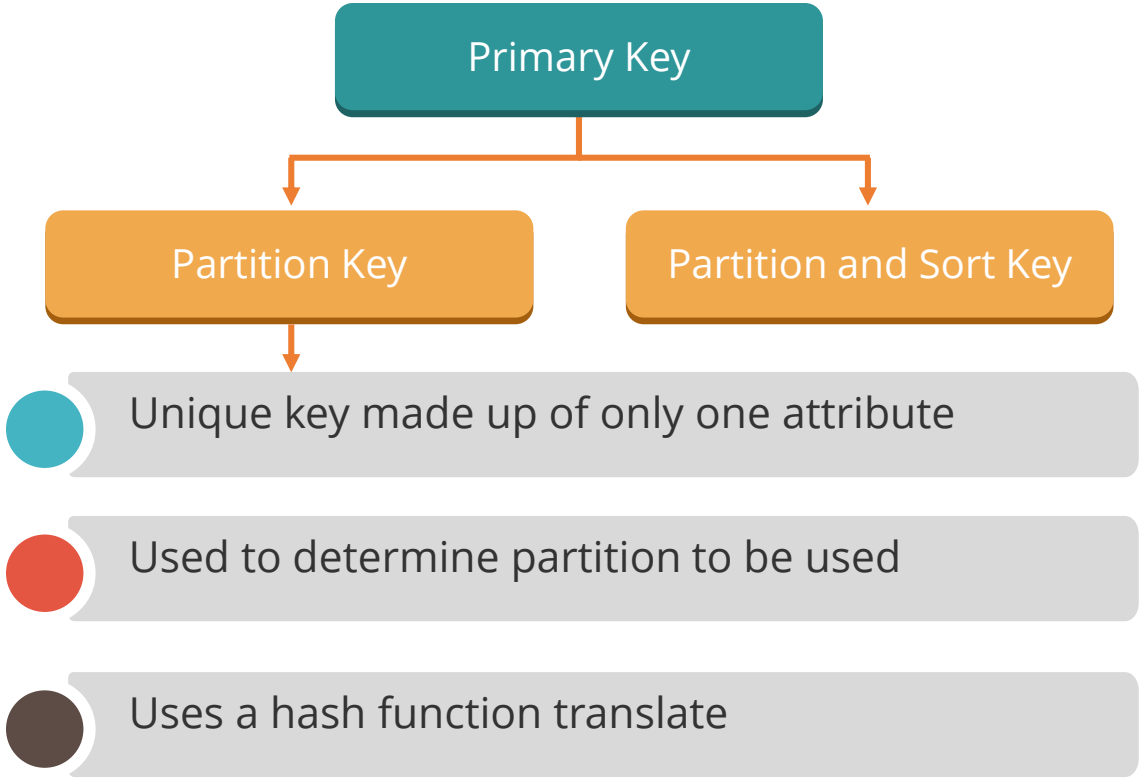
Primary Keys





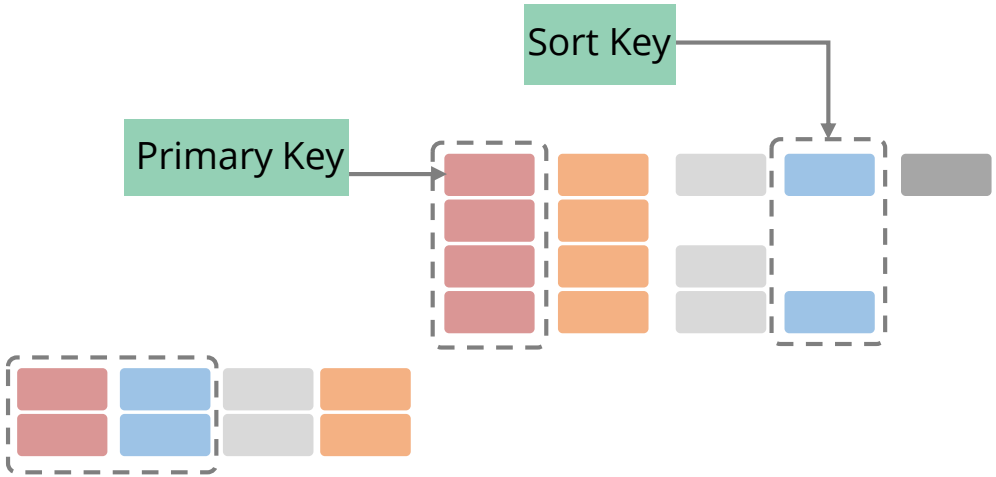
Primary Key is Unique
1.Hash only

amazon web services



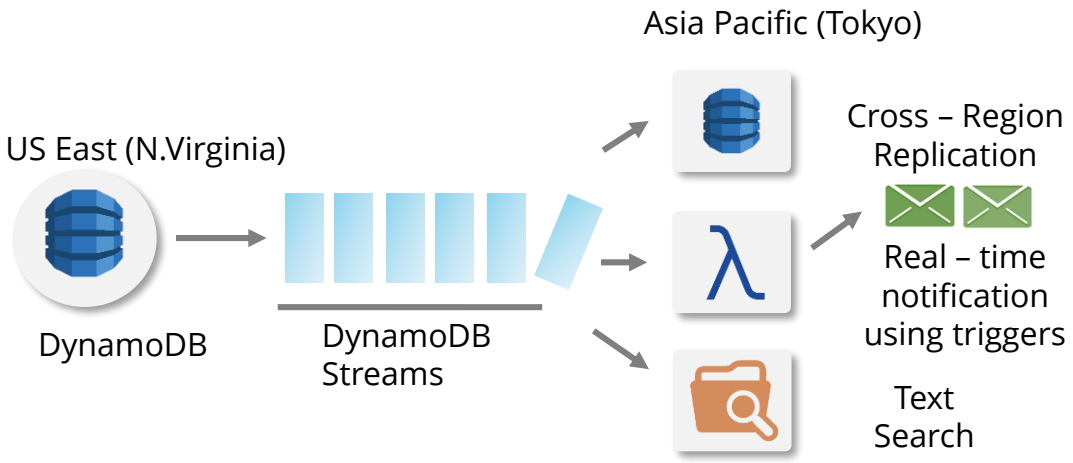
simplilearn

Secondary Indexes



- Can be used as an alternate key to query a table
- Two types: Local and global secondary index
- Local secondary index uses same primary key
- Global secondary index uses a different partition and sort key
- While creating, specify attributes to be maintained

DynamoDB Streams



- DynamoDB Streams are used to capture any modification events
- Stream captures an image of all items and attributes
- Stream captures before and after images of modified attributes
- Stream captures image of item prior to deletion

Query and Scan

Query
is used to search and find items from a table or secondary index

Scan
reads every item in a table or a secondary index

Query and Scan

Query

- Returns all data attributes for an item
- Uses sort key attribute name and value to refine results
- Is used with BatchGetItems API to return multiple items
- Is preferred to Scan operations

Scan

- Sequentially reads every item
- Scans entire table and then filters
- Not recommended for large tables
- Not a preferred form of searching and filtering

Knowledge Check

KNOWLEDGE CHECK

Which are the valid secondary index types? (Choose 2)

a. Hash key

b. Local secondary

c. Global secondary

d. Range key

KNOWLEDGE CHECK

Which are the valid secondary index types? (Choose 2)

a. Hash key

b. Local secondary

c. Global secondary

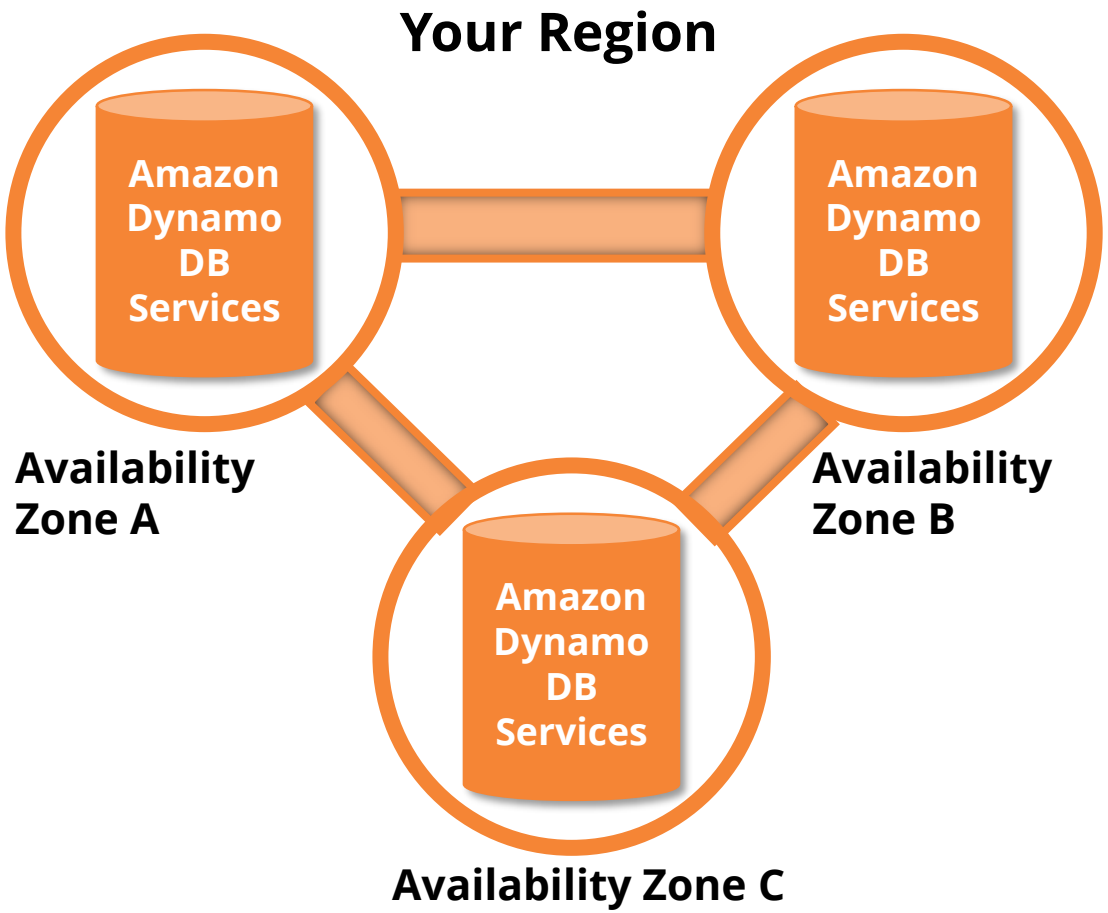
d. Range key

The correct answer is **Global secondary & Local secondary**

Explanation: DynamoDB provides two types of secondary indexes: Local secondary index and Global secondary index.

Capacity and Throughput Management

Capacity Management



- DynamoDB stores and manages any amount of data and traffic
- Provides high performance and less latency
- Automatically replicates data across three data centers

Read Operation and Consistency

GetItem Operation

- Helps to read an item from a DynamoDB table
- Whole key is needed while using composite key
- Performs eventual consistent reads by default
- Consumes additional read capacity units for strongly consistent reads
- Consumes half-read capacity for eventually consistent reads

Write Operation and Conditional Writes

PutItem

Creates a new item

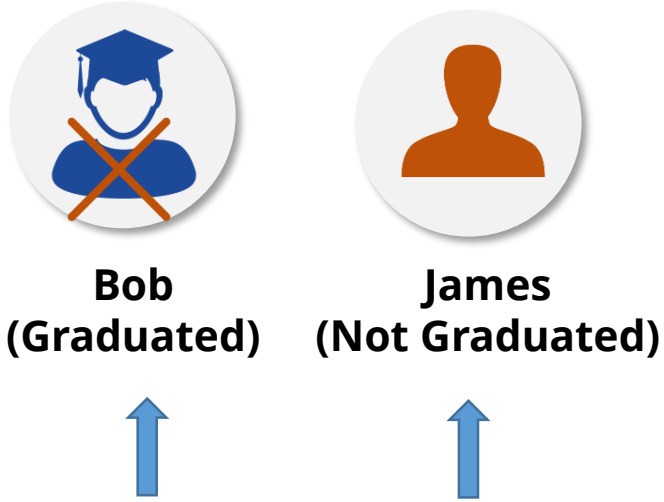
UpdateItem

Modifies existing item's attributes

DeleteItem

Deletes an existing item

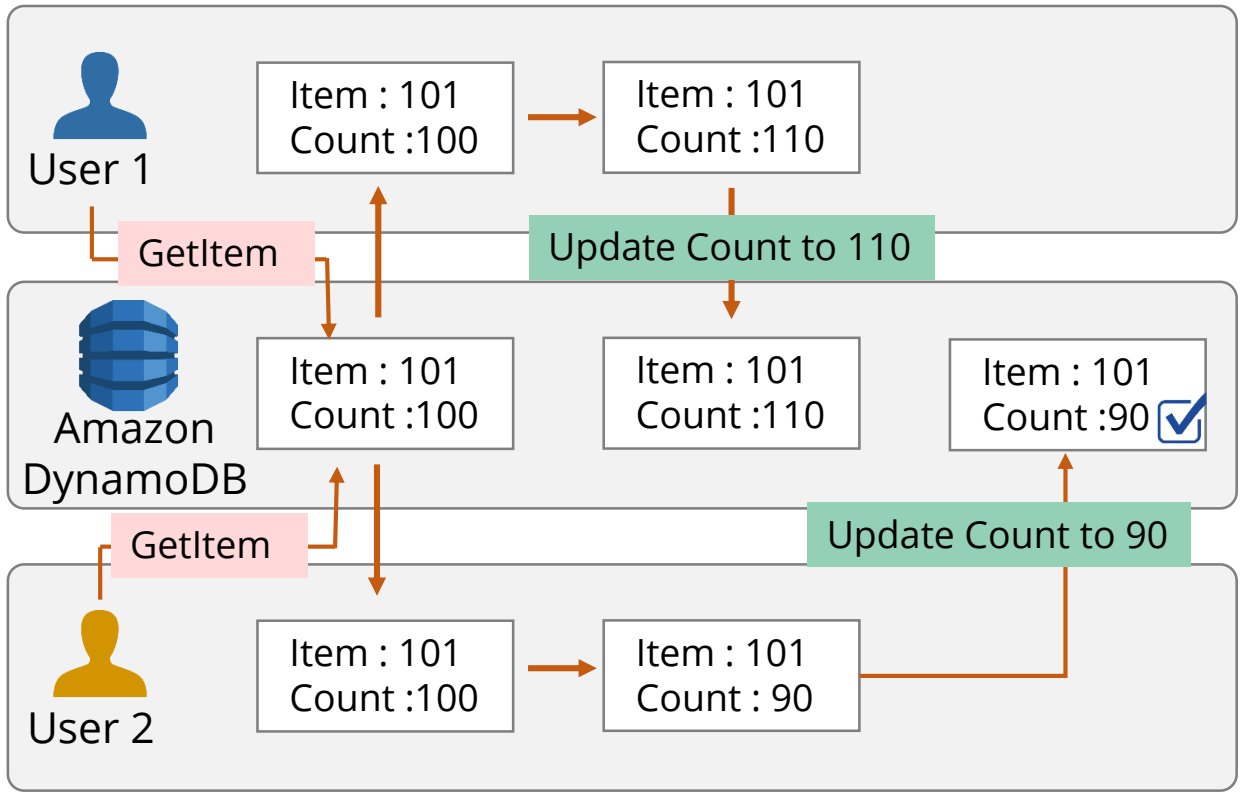
DeleteItem attribute_exists (Graduated)



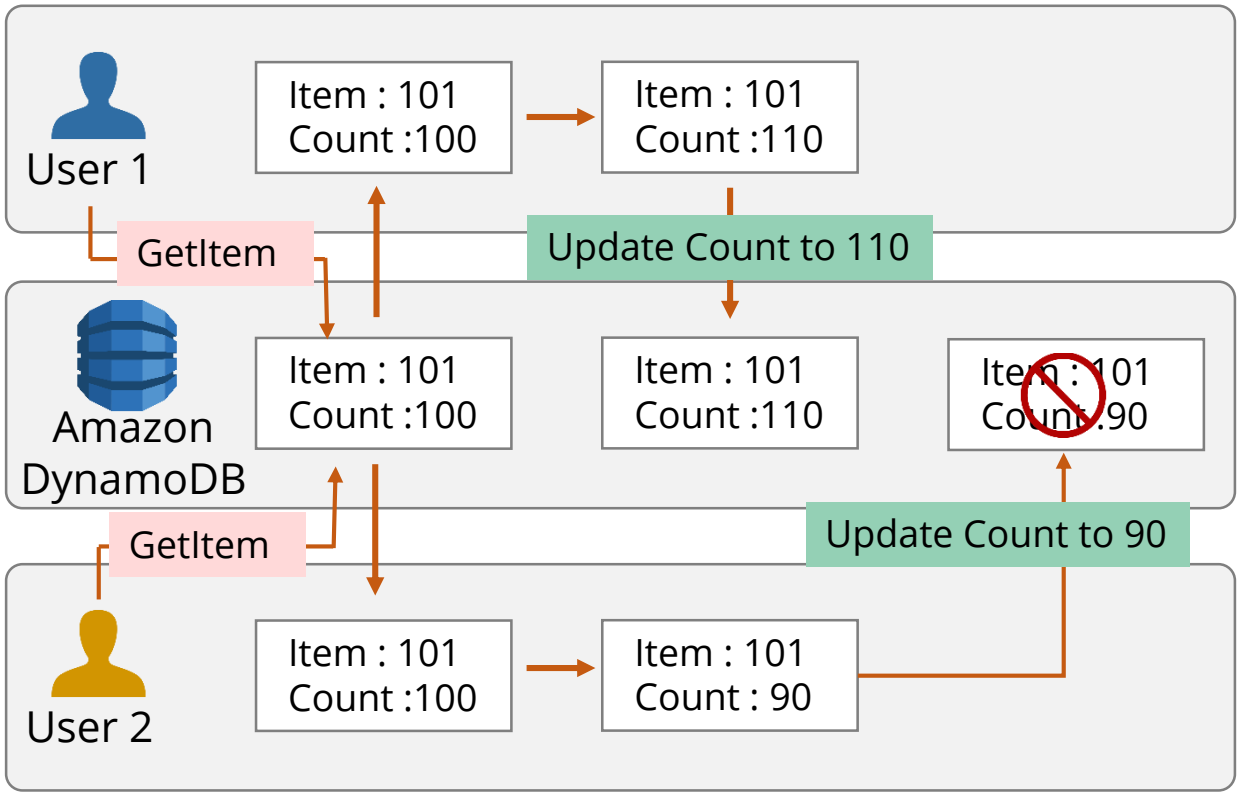
i Use ReturnConsumedCapacity to obtain information on units consumed

Using Conditional Writes

Without Conditional Writes



Without Conditional Writes



i Conditional writes are idempotent

Atomic Counter

Amazon DynamoDB Supports Atomic Counters

- Increment/decrement of existing data attribute value
- Uses UpdateItem operation
- **Atomic Counters are not idempotent**
- Not ideal for financial/mission-critical applications
- Use “set” action to update the counter

Batch Operations

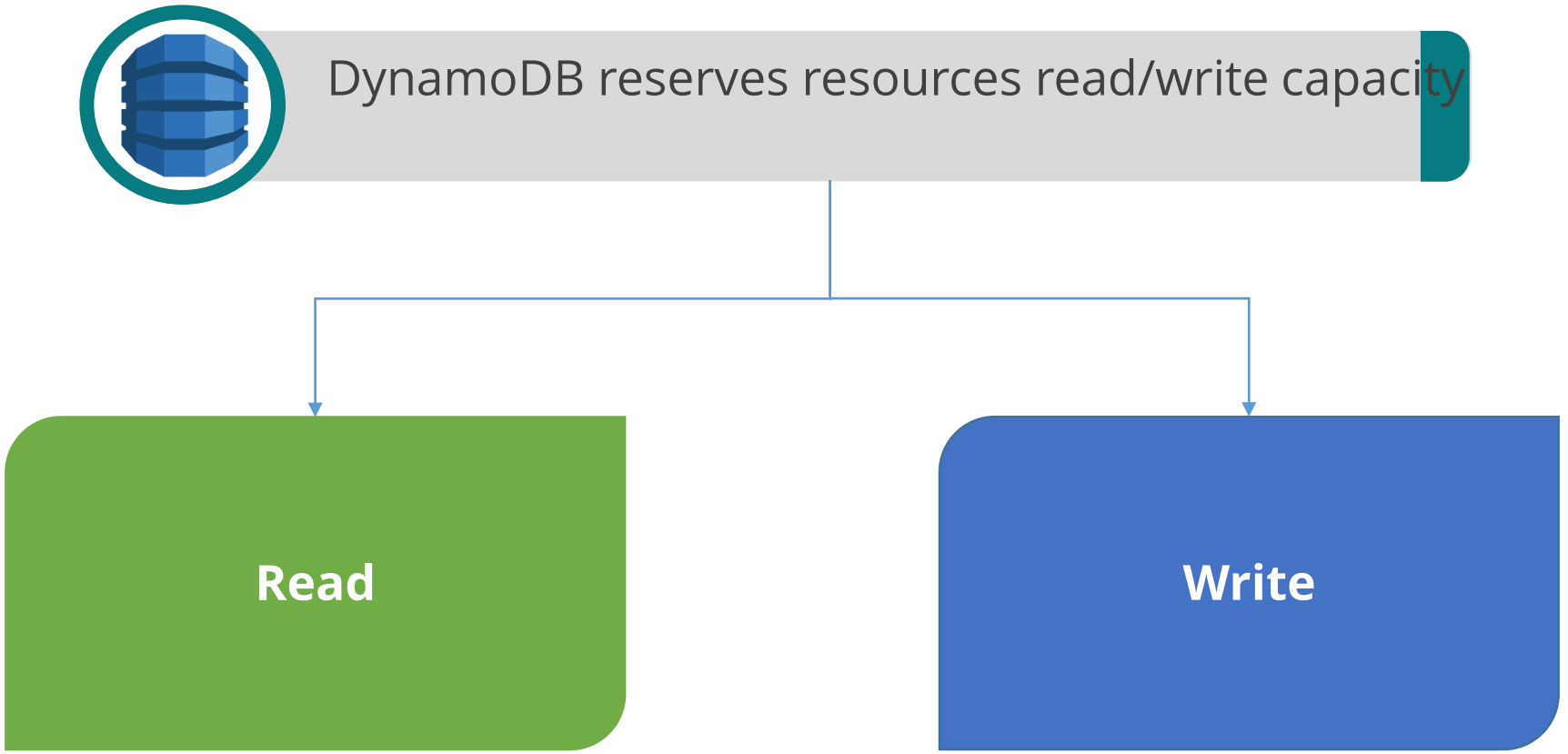
BatchGetItem

- Allows to read multiple items
- Can retrieve up to 16MB data
- Up to 100 items

BatchWriteItem

- Allows to create/delete multiple items
- Can add/delete up to 16MB data
- Up to 25 items

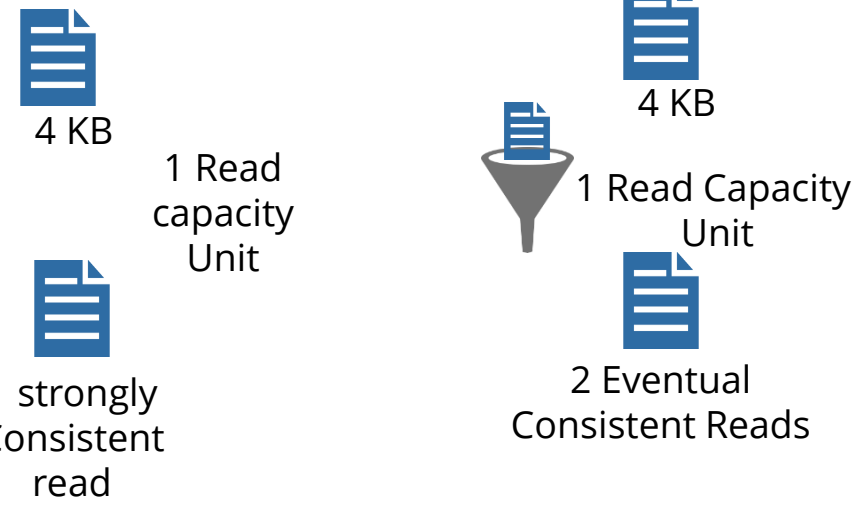
Provisioned Throughput



Provisioned Throughput

DynamoDB reserves resources read/write capacity

Read



One strongly/two eventual consistent per second

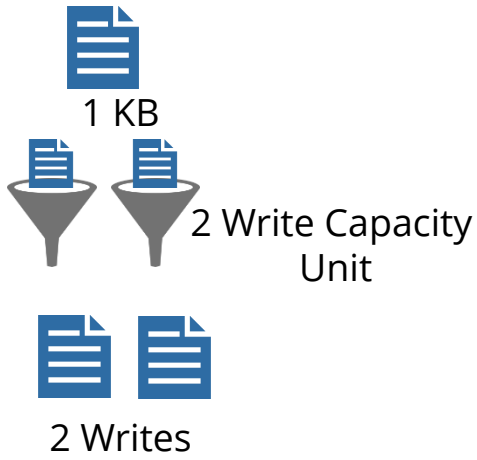
One read item supports up to 4kb

If more, DynamoDB consumes additional units

Provisioned Throughput

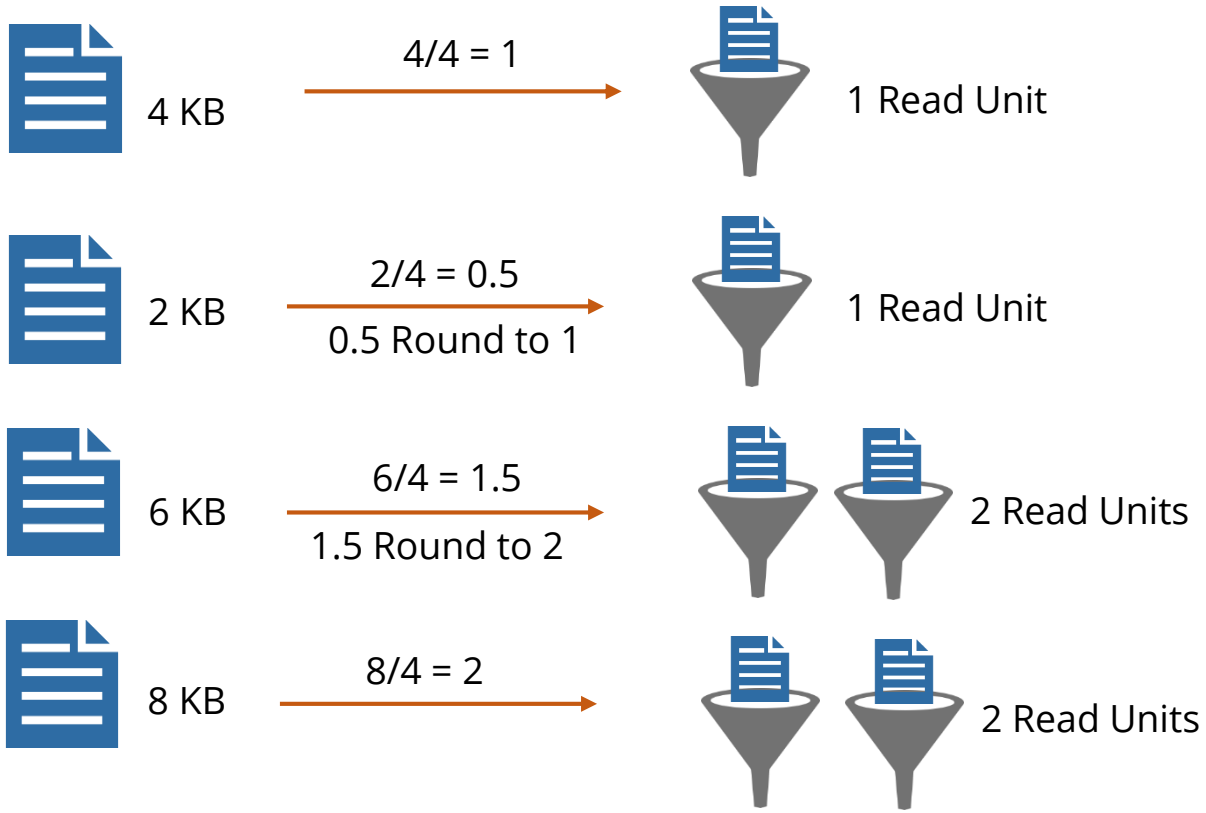
DynamoDB reserves resources read/write capacity

Write



- One write per second
- Supports up to 1KB per write
- Additional capacity for secondary indexes

Read Capacity Units



Read Capacity Units – Example 1

Example:

You want to perform 100 strongly consistent reads per second, where the items are 6 KB in size. How many read capacity units do you need?

Solution

1. Number of read capacity units required per item = 6 KB/ 4 = 1.5

2. Round up 1.5 to the nearest whole number = 2 Read units per item

3. Total units needed for **strongly consistent reads** = 100 items * 2 = 200 Units

→ **Answer**

4. Total units needed for **eventual consistent reads** = 100 items/2 = 100 Units

Read Capacity Units – Example 2

Example:

An application requires to read 20 items of 1 KB per second using eventual consistency. What should you set the read throughput to be?

Solution

1. Number of read capacity units required per item = 1 KB/ 4 = 0.25

2. Round up 0.25 to the nearest whole number = 1 Read unit per item

3. Total units needed for **strongly consistent reads** = 20 items * 1 = 20 Units

4. Total units needed for **eventual consistent reads** = 20 items/2 = 10 Units

→ **Answer**

Write Capacity Units

One write capacity unit provides one write per second

Supports items up to 1KB in size per write

Example:

An application requires to write 10 items with each item being 20 KB in size per second. What should be set as the write throughput for the application?

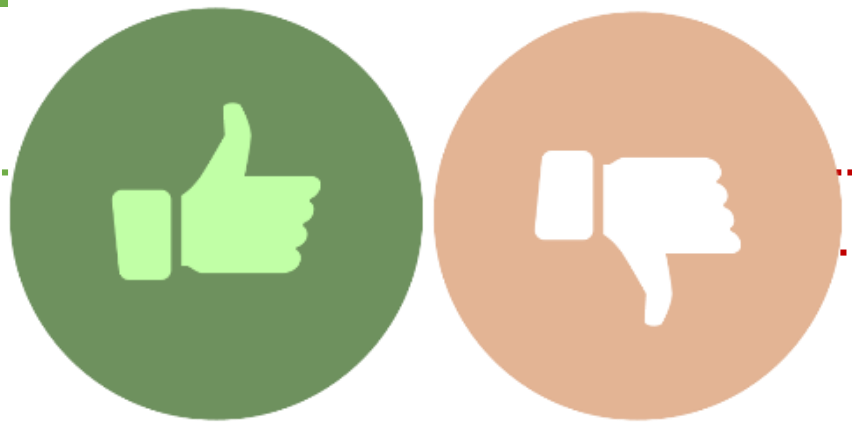
Solution

Number of write capacity units required = 20 KB * 10 = 200 → **Answer**

Best Practices

Use reduced page size scan or isolate scan operation

Use parallel scan instead of sequential scan



Sudden burst of usage may create capacity issues

May get exceptions for read requests while supporting scan operation

Knowledge Check

KNOWLEDGE CHECK

An application requires to read 10 items of 9 KB per second using eventual consistency. What should you set the read throughput to be?

a. 30

b. 90

c. 15

d. 10

KNOWLEDGE CHECK

An application requires to read 10 items of 9 KB per second using eventual consistency. What should you set the read throughput to be?

a. 30

b. 90

c. 15

d. 10

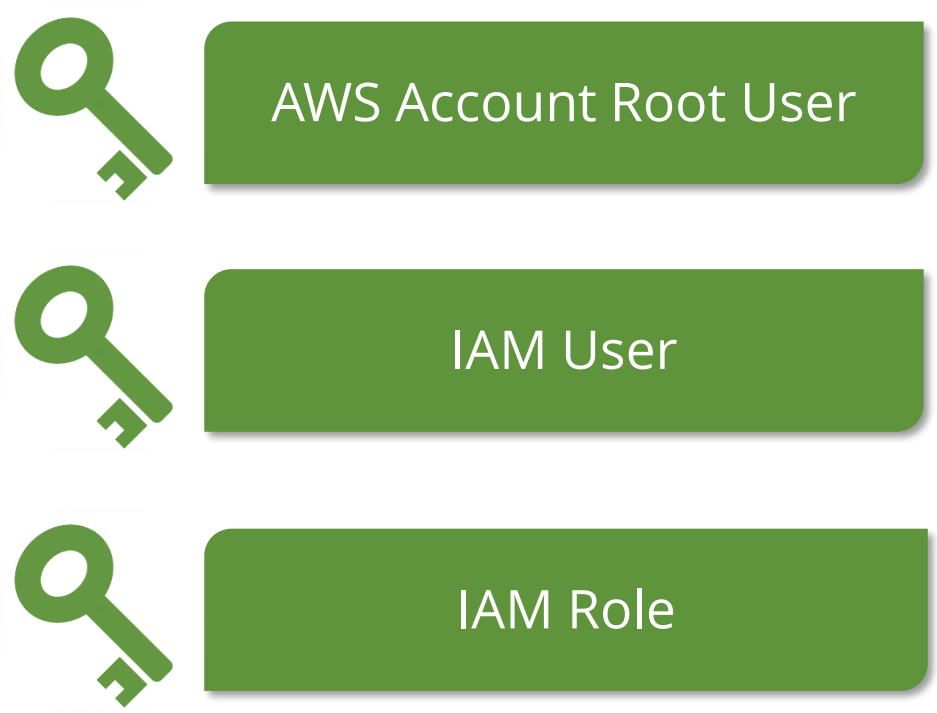
The correct answer is **15**

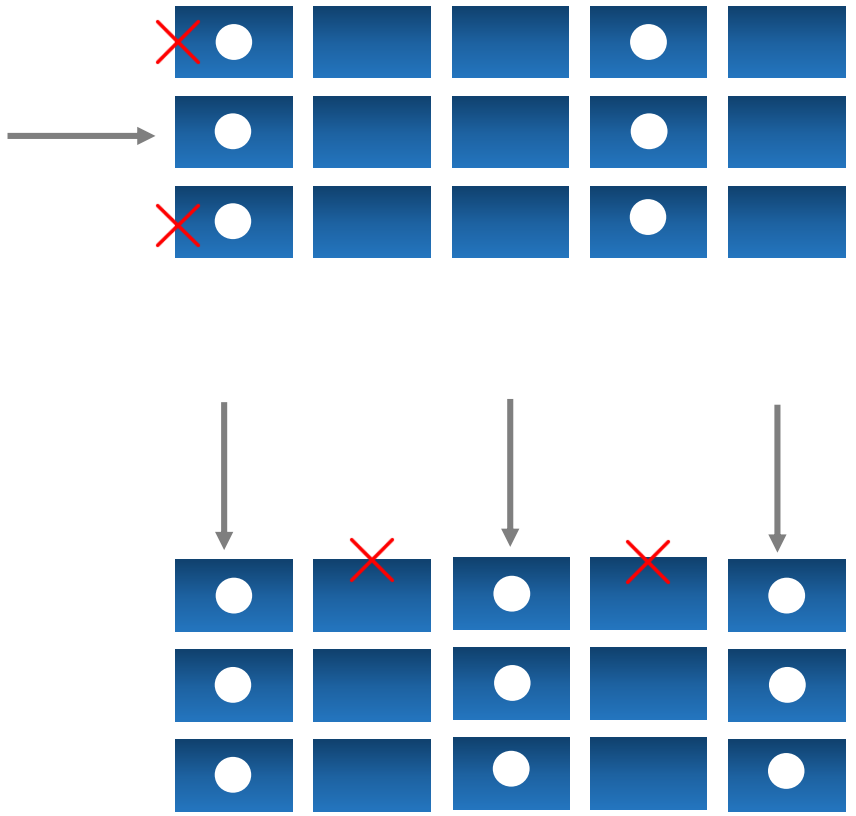
Explanation: Number of read capacity units required per item = 9 KB/4 = 2.25. Round up 2.25 to the nearest whole number = 3 Read unit per item.

Total units needed for strongly consistent reads = 10 items * 3 = 30 Units. Total units needed for eventual consistent reads = 30 items/2 = 15 Units.

Access Control and Data Migration

Authentication and Access Control





Identity Based Policies

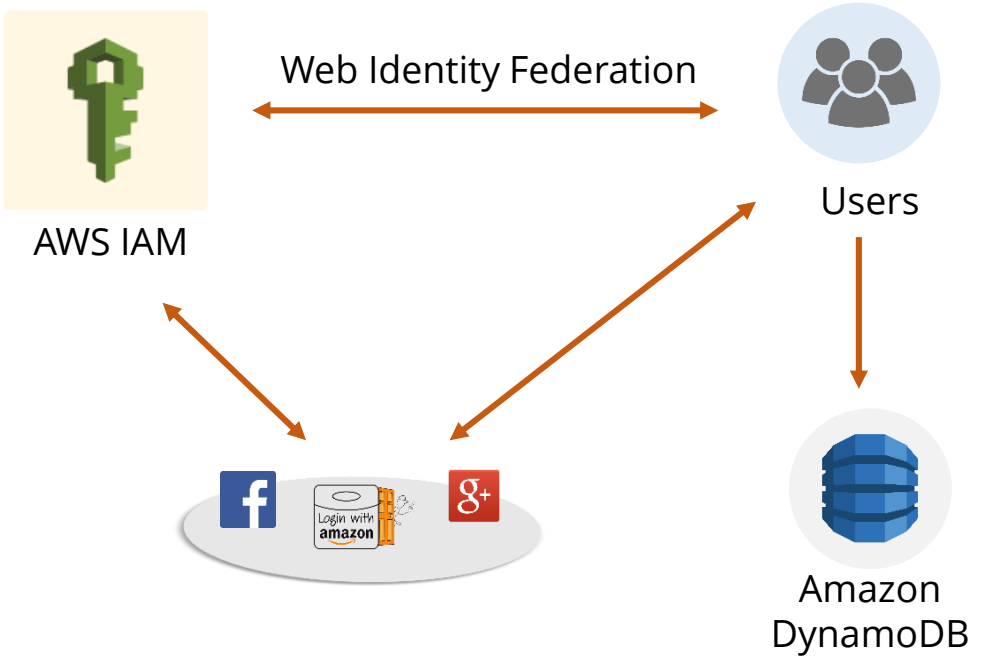
```
{ "Version": "2012-10-17",
      "Statement": [ {
            "Sid": "ReadOnlyAPIActionsOnBooks",
            "Effect": "Allow",
            "Action": [ "dynamodb:GetItem", "dynamodb:BatchGetItem" ],
            "Resource": "arn:aws:dynamodb:us-west-2:123456789012:table/Books" } ]
}
```

Grants access to GetItem and BatchGetItem

```
"Resource": "arn:aws:dynamodb:us-west-
2:123456789012:table/${aws:username}_ProductCatalog"
```

Replaces requester's username when policy is evaluated

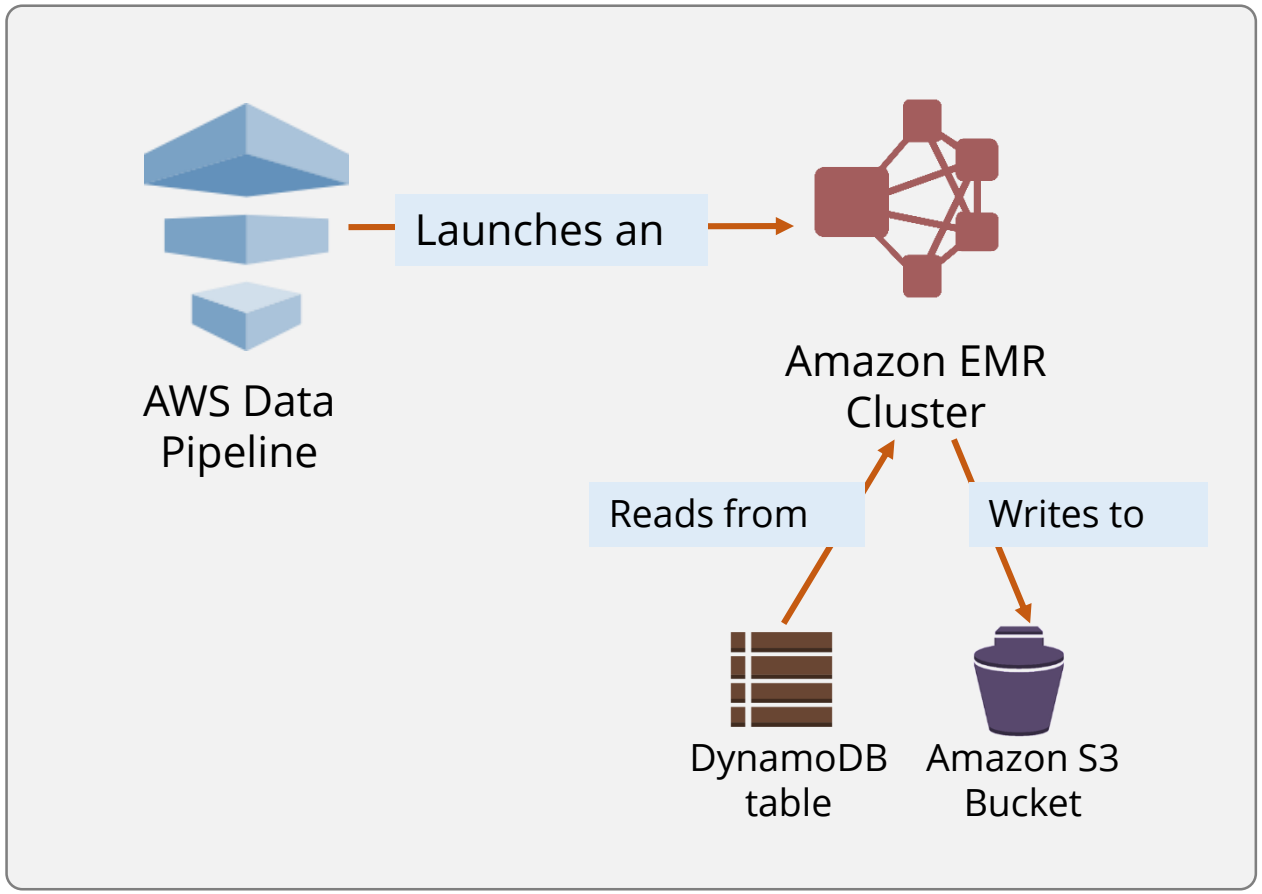
Web Identity Federation



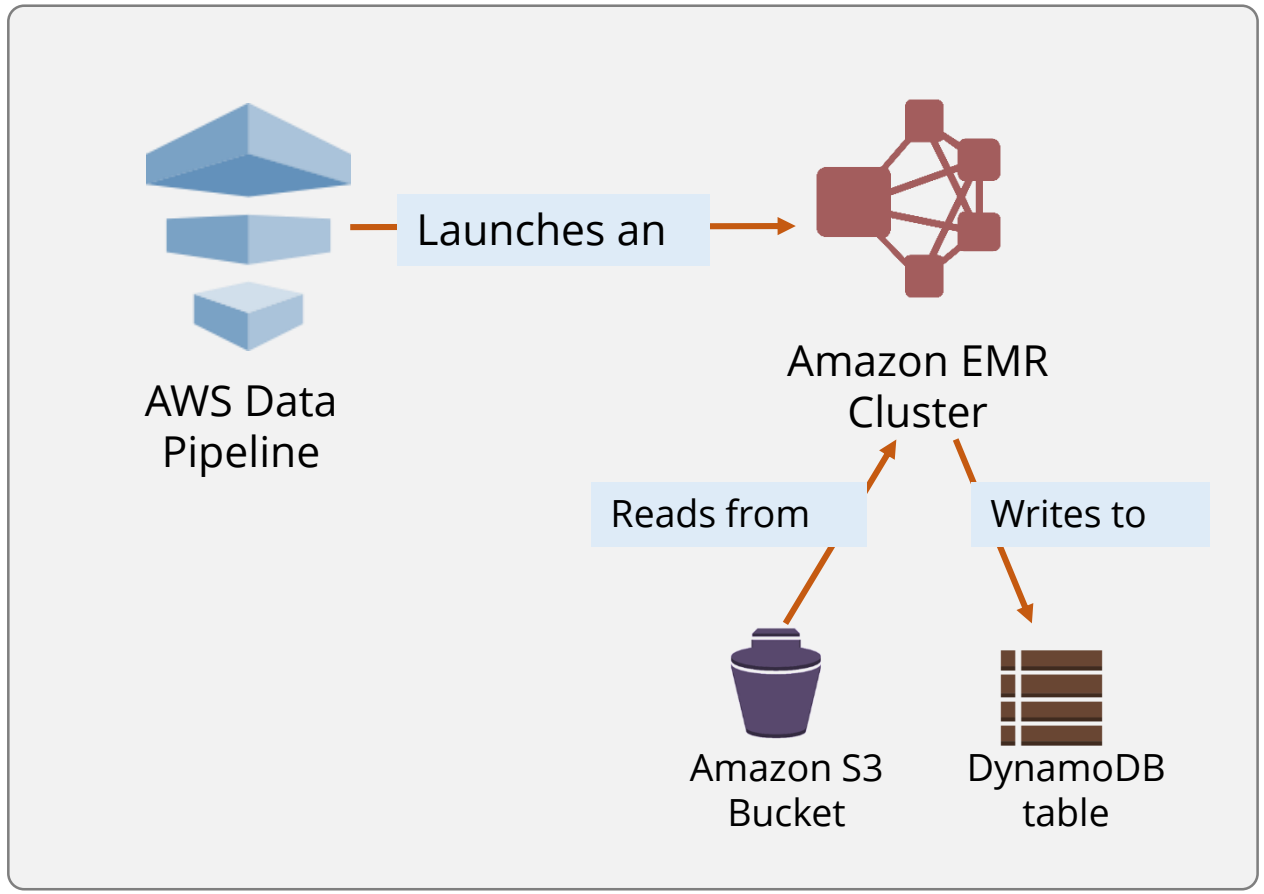
1. Signup and configure your application with IdP
2. Create an entity in IAM
3. Create an IAM role and define who can assume it

Web Identity Federation

Exporting Data From Dynamo to Amazon S3



Importing Data from Amazon S3 to DynamoDB



Knowledge Check

KNOWLEDGE CHECK

What API calls do you use to read data from the DynamoDB table?

a. ReadItem

b. GetItem

c. BatchReadItem

d. BatchGetItem

KNOWLEDGE CHECK

What API calls do you use to read data from the DynamoDB table?

a. ReadItem

b. GetItem

c. BatchReadItem

d. BatchGetItem

The correct answer is **GetItem & BatchGetItem**

Explanation: GetItem and BatchGetItem API calls provide read only access to a table.

Practice Assignment: Amazon DynomoDB

To create a DynamoDB table using the provided throughput

To create a DynamoDB table using the provided throughput

Create a DynamoDB table. When creating the DynamoDB table, calculate the provision a throughput requires to store a table with write capacity of 20 items with 4 KB size and read capacity of 80 items.

Prerequisites:
AWS Account

Task:
How to create a Dynamo DB table and use the read 80 and 20 items write capacity for provision throughput.

Quiz

QUIZ 1	In DynamoDB, do you need to define a table schema?

a. Yes, at the table level

b. Yes, at the item level

c. No, it is schema less

d. No, not for a table, but for a database

QUIZ 1	In DynamoDB, do you need to define a table schema?

a. Yes, at the table level

b. Yes, at the item level

c. No, it is schema less

d. No, not for a table, but for a database

The correct answer is **No, it is schema less**

Explanation: DynamDB tables are schema less. Except the primary key, you don't need to define any other attribute or their respective data types beforehand.

QUIZ 2	Which are the valid read consistency models in DynamoDB?

a. Eventual consistent

b. Softly consistent

c. Read after write consistent

d. None of the above

QUIZ 2	Which are the valid read consistency models in DynamoDB?

a. Eventual consistent

b. Softly consistent

c. Read after write consistent

d. None of the above

The correct answer is **Eventual consistent**

Explanation: By default, GetItem performs an eventual consistent reads. You can use GetItem with strongly consistent reads, but it consumes additional read capacity units.

QUIZ 3	How many read capacity units do you need to perform strongly consistent read on 120 items per minute?(Given that the items are 8 KB in size.)

a. 4

b. 240

c. 120

d. 2

QUIZ

3

How many read capacity units do you need to perform strongly consistent read on 120 items per minute?(Given that the items are 8 KB in size.)

a. 4

b. 240

c. 120

d. 2

The correct answer is **4**

Explanation: Number of read capacity units required per item = 8 KB/4 = 2. No of items to be read per second = 120/60 = 2.

Total units needed for strongly consistent reads = 2*2 = 4 Units. Total units needed for strongly consistent reads per second = 4 Units.

QUIZ

4

An application requires to write 30 items with each item being 10 KB in size per second. What should be set as the write throughput for the application?

a. 30

b. 90

c. 150

d. 300

QUIZ

4

An application requires to write 30 items with each item being 10 KB in size per second. What should be set as the write throughput for the application?

a. 30

b. 90

c. 150

d. 300

The correct answer is **300**

Explanation: Number of write capacity units required = 10 KB * 30 = 300.

QUIZ 5

When DynamoDB throttles your request, what exception do you receive?

a. ReadThroughputExceededException

b. ProvisionedThroughputExceededException

c. WriteThroughputExceededException

d. ThroughputExceededException

QUIZ

5

When DynamoDB throttles your request, what exception do you receive?

a. ReadThroughputExceededException

b. ProvisionedThroughputExceededException

c. WriteThroughputExceededException

d. ThroughputExceededException

The correct answer is **ProvisionedThroughputExceededException**

Explanation: If you exceed your provisioned throughput, DynamoDB might throttle your requests. In that case, your request fails with an HTTP 400 code, and is accompanied by a ProvisionedThroughputExceededException.

Key Takeaways

- Items are similar to records, each item size is limited to 64 Kb
- Primary key can be of two types: Partition key, and partition and sort key
- 5 global secondary indexes and 5 local secondary indexes are allowed
- Query operations are considered more efficient than scan operations
- DynamoDB supports conditional updates. Updates done by one user cannot be overwritten by another user. It supports atomic counters, increment or decrement of the value of an existing data attribute
- One read capacity unit provides one strongly consistent read per second or two eventual consistent reads per second. Supports items up to 4 KB in size per read
- One write capacity unit provides one write per second, supports items up to 1 KB in size per write

This concludes “AWS DynamoDB”

The next lesson is “AWS Application Services.”