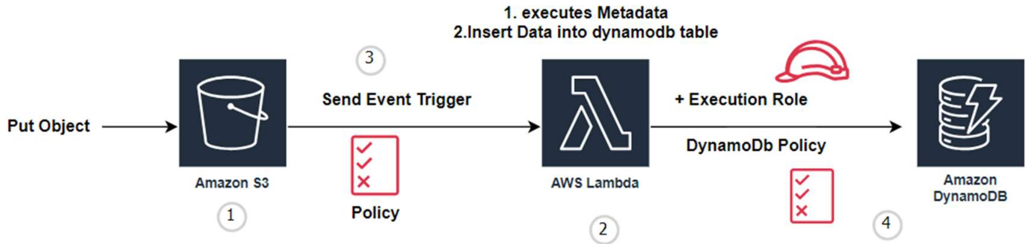
Problem statement:

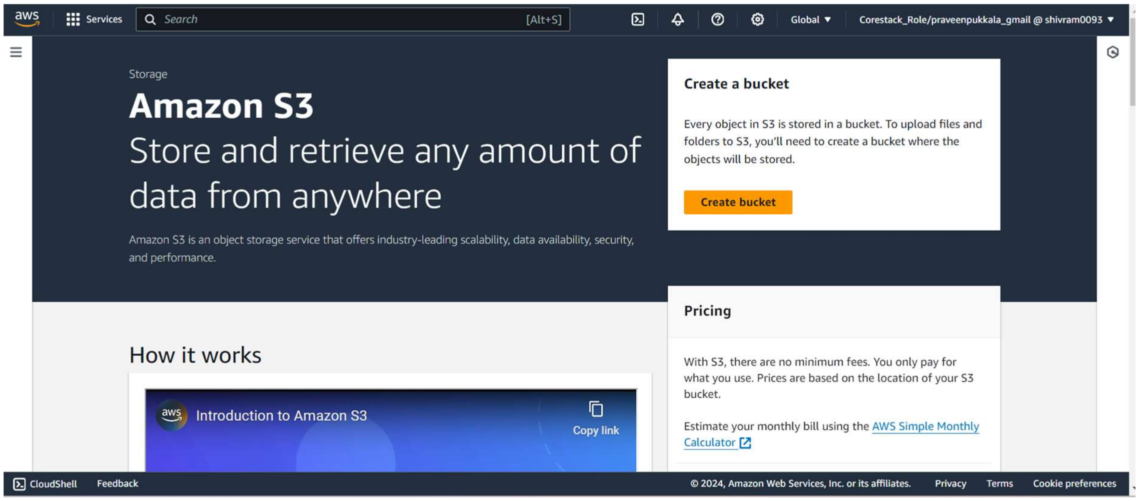
Data Processing Task- **When any file we receive**(S3), the **metadata** (lambda Code) of that file should **be store in a database (DynamoDB)**.

Architecture Diagram:



Implementation in AWS Console:

1. Create a bucket. And give name: mys3bucket20022024, and rest is default.


aws
Services
Search
[Alt+S]
Global
Corestack_Role/praveenpukkala_gmail @ shivram0093
Storage
Create bucket
How it works
Introduction to Amazon S3
Copy link
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

Amazon S3

Store and retrieve any amount of data from anywhere

Amazon S3 is an object storage service that offers industry-leading scalability, data availability, security, and performance.

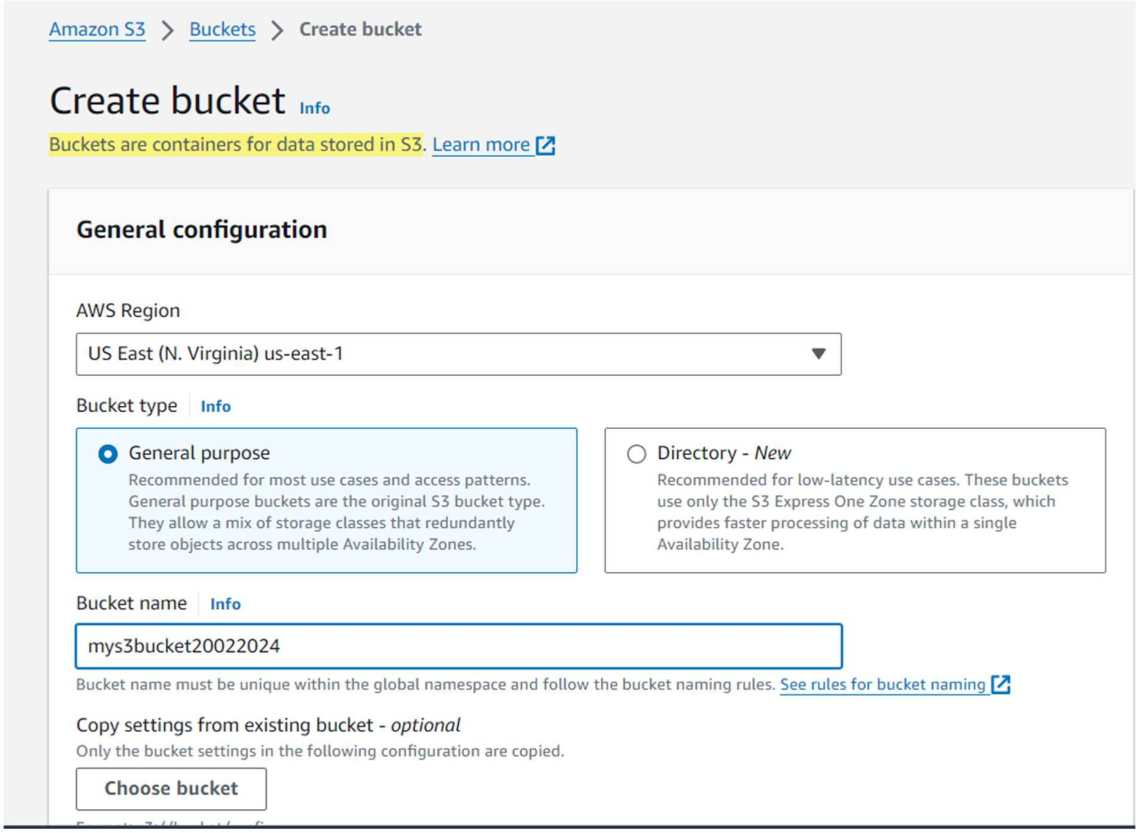
Create a bucket

Every object in S3 is stored in a bucket. To upload files and folders to S3, you'll need to create a bucket where the objects will be stored.

Pricing

With S3, there are no minimum fees. You only pay for what you use. Prices are based on the location of your S3 bucket.

Estimate your monthly bill using the AWS Simple Monthly Calculator



Create bucket Info

Buckets are containers for data stored in S3. Learn more

General configuration

AWS Region

US East (N. Virginia) us-east-1

Bucket type Info

● General purpose
Recommended for most use cases and access patterns. General purpose buckets are the original S3 bucket type. They allow a mix of storage classes that redundantly store objects across multiple Availability Zones.

○ Directory - *New*
Recommended for low-latency use cases. These buckets use only the S3 Express One Zone storage class, which provides faster processing of data within a single Availability Zone.

Bucket name Info

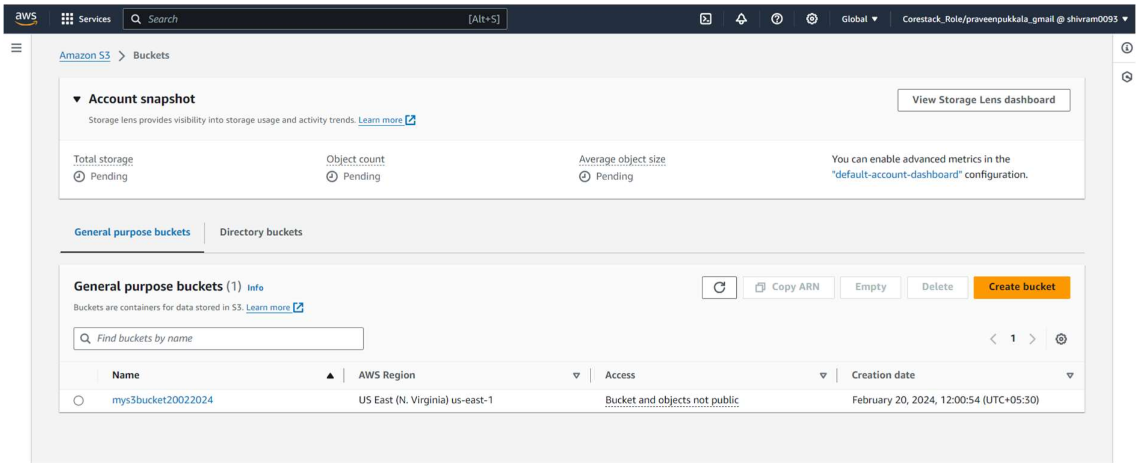
mys3bucket20022024

Bucket name must be unique within the global namespace and follow the bucket naming rules. See rules for bucket naming

Copy settings from existing bucket - *optional*

Only the bucket settings in the following configuration are copied.

After creating:



Account snapshot

Storage lens provides visibility into storage usage and activity trends. Learn more

Total storage: Pending

Object count: Pending

Average object size: Pending

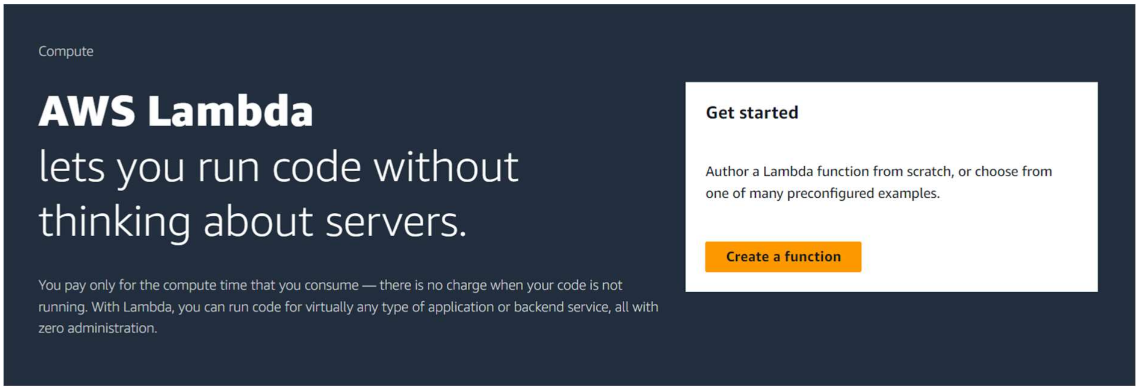
You can enable advanced metrics in the "default-account-dashboard" configuration.

General purpose buckets (1) Info

Buckets are containers for data stored in S3. Learn more

	Name	AWS Region	Access	Creation date
○	mys3bucket20022024	US East (N. Virginia) us-east-1	Bucket and objects not public	February 20, 2024, 12:00:54 (UTC+05:30)

2. Create a lambda function, give name and choose python 3.8 runtime



Compute

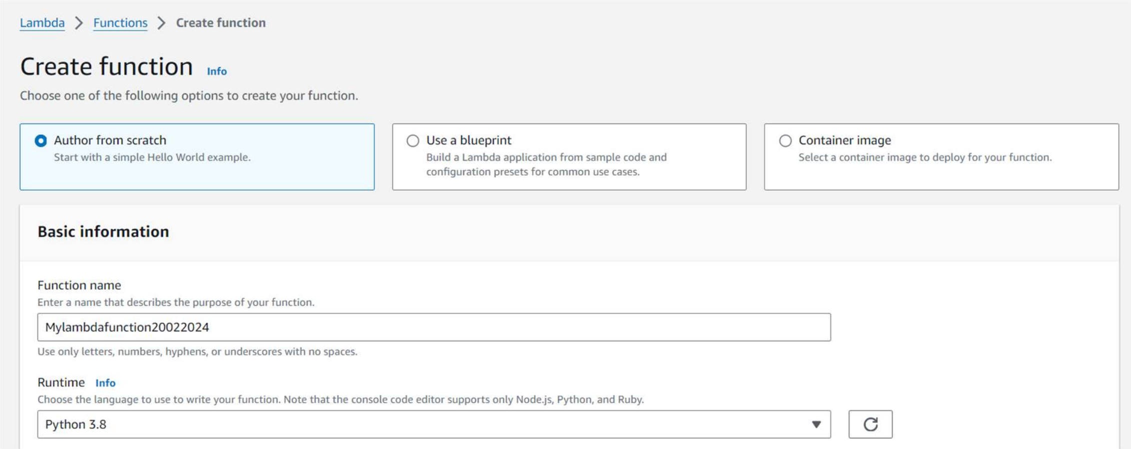
AWS Lambda

lets you run code without thinking about servers.

You pay only for the compute time that you consume — there is no charge when your code is not running. With Lambda, you can run code for virtually any type of application or backend service, all with zero administration.

Get started

Author a Lambda function from scratch, or choose from one of many preconfigured examples.



Create function Info

Choose one of the following options to create your function.

● Author from scratch
Start with a simple Hello World example.

○ Use a blueprint
Build a Lambda application from sample code and configuration presets for common use cases.

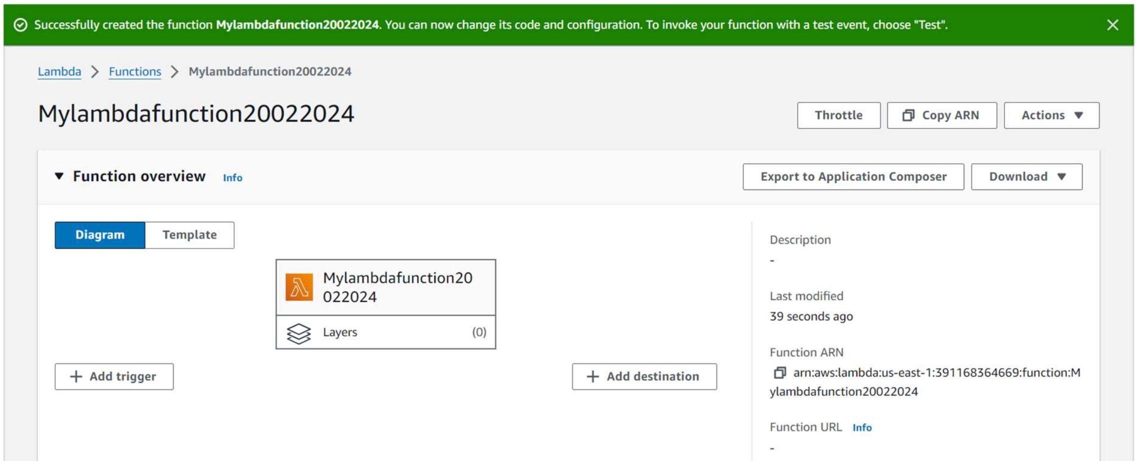
○ Container image
Select a container image to deploy for your function.

Basic information

Function name
Enter a name that describes the purpose of your function.
Mylambdafunction20022024
Use only letters, numbers, hyphens, or underscores with no spaces.

Runtime Info
Choose the language to use to write your function. Note that the console code editor supports only Node.js, Python, and Ruby.
Python 3.8

Click on create function



Successfully created the function **Mylambdafunction20022024**. You can now change its code and configuration. To invoke your function with a test event, choose "Test".

Mylambdafunction20022024

Description
-

Last modified
39 seconds ago

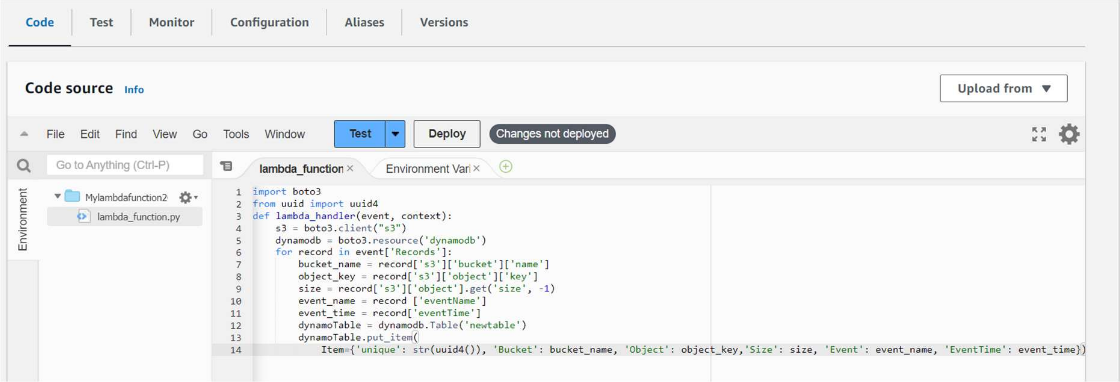
Function ARN
arn:aws:lambda:us-east-1:391168364669:function:Mylambdafunction20022024

Function URL Info
-

Now paste the code and click deploy.

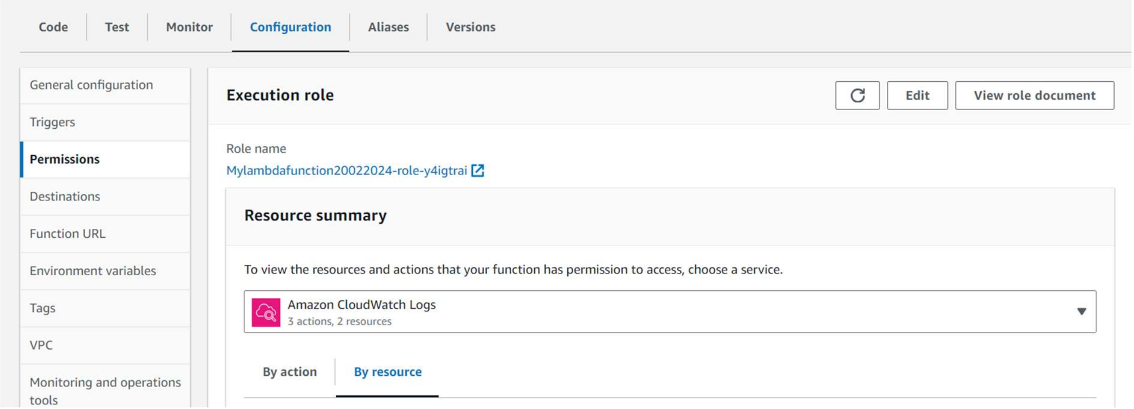
```
import boto3
from uuid import uuid4
def lambda_handler(event, context):
s3 = boto3.client("s3")
dynamodb = boto3.resource('dynamodb')
for record in event['Records']:
bucket_name = record['s3']['bucket']['name']
```

```
object_key = record['s3']['object']['key']
size = record['s3']['object'].get('size', -1)
event_name = record ['eventName']
event_time = record['eventTime']
dynamoTable = dynamodb.Table('newtable')
dynamoTable.put_item(
Item={'unique': str(uuid4()), 'Bucket': bucket_name, 'Object':
object_key,'Size': size, 'Event': event_name, 'EventTime':
event_time})
```



```
import boto3
from uuid import uuid4
def lambda_handler(event, context):
    s3 = boto3.client("s3")
    dynamodb = boto3.resource('dynamodb')
    for record in event['Records']:
        bucket_name = record['s3']['bucket']['name']
        object_key = record['s3']['object']['key']
        size = record['s3']['object'].get('size', -1)
        event_name = record ['eventName']
        event_time = record['eventTime']
        dynamoTable = dynamodb.Table('newtable')
        dynamoTable.put_item(
            Item={'unique': str(uuid4()), 'Bucket': bucket_name, 'Object': object_key,'Size': size, 'Event': event_name, 'EventTime': event_time})
```

lambda by default send logs to cloud watch and role will be created automatically



Execution role

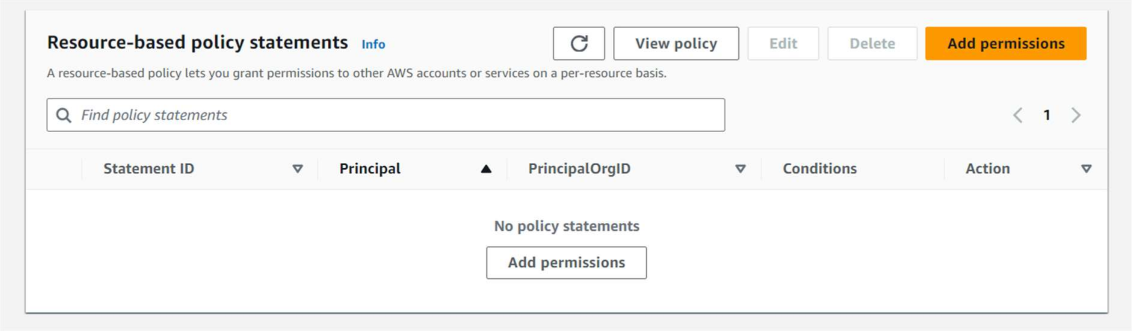
Role name
Mylambdafunction20022024-role-y4igtrai

Resource summary

To view the resources and actions that your function has permission to access, choose a service.

Amazon CloudWatch Logs
3 actions, 2 resources

before adding trigger observe there is no resource-based policy



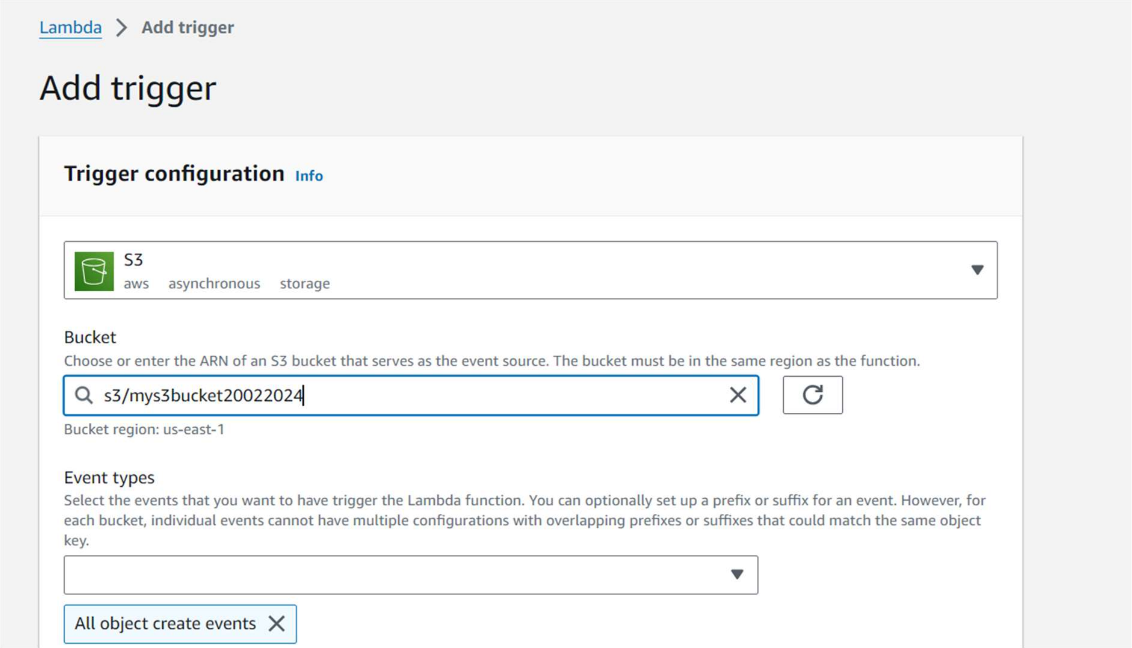
Resource-based policy statements Info

A resource-based policy lets you grant permissions to other AWS accounts or services on a per-resource basis.

Statement ID	Principal	PrincipalOrgID	Conditions	Action

No policy statements

Now + add Trigger—Select s3→choose bucket



Lambda > Add trigger

Add trigger

Trigger configuration Info

S3
aws asynchronous storage

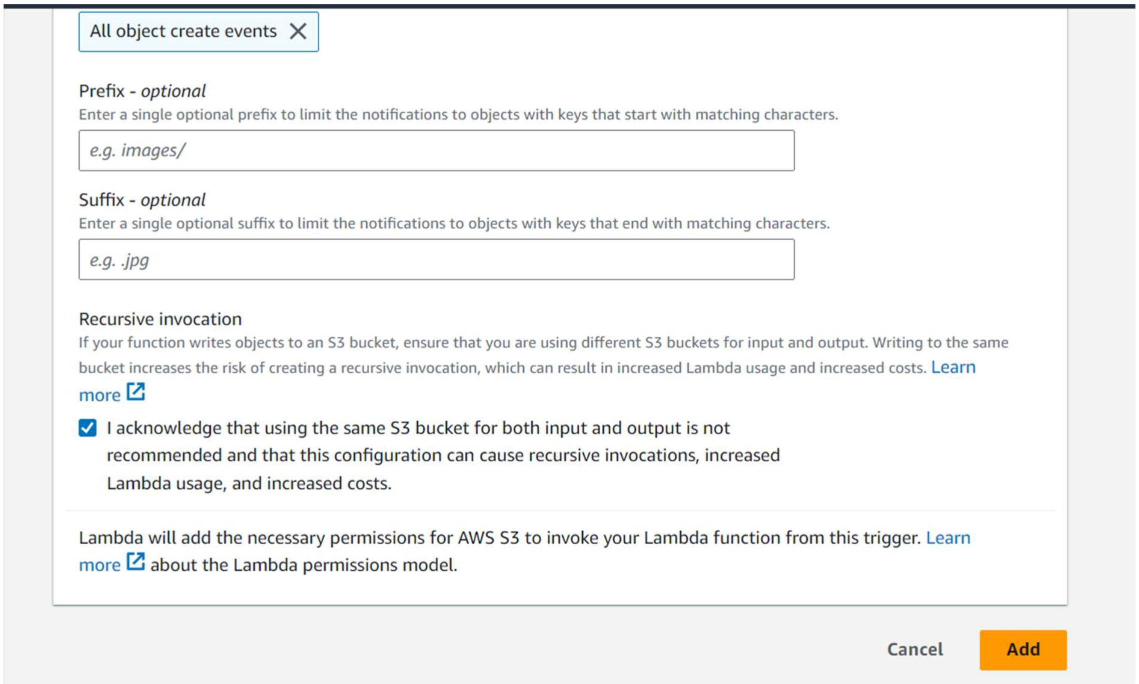
Bucket
Choose or enter the ARN of an S3 bucket that serves as the event source. The bucket must be in the same region as the function.

s3/mys3bucket20022024

Bucket region: us-east-1

Event types
Select the events that you want to have trigger the Lambda function. You can optionally set up a prefix or suffix for an event. However, for each bucket, individual events cannot have multiple configurations with overlapping prefixes or suffixes that could match the same object key.

All object create events



All object create events ✕

Prefix - *optional*

Enter a single optional prefix to limit the notifications to objects with keys that start with matching characters.

e.g. images/

Suffix - *optional*

Enter a single optional suffix to limit the notifications to objects with keys that end with matching characters.

e.g. .jpg

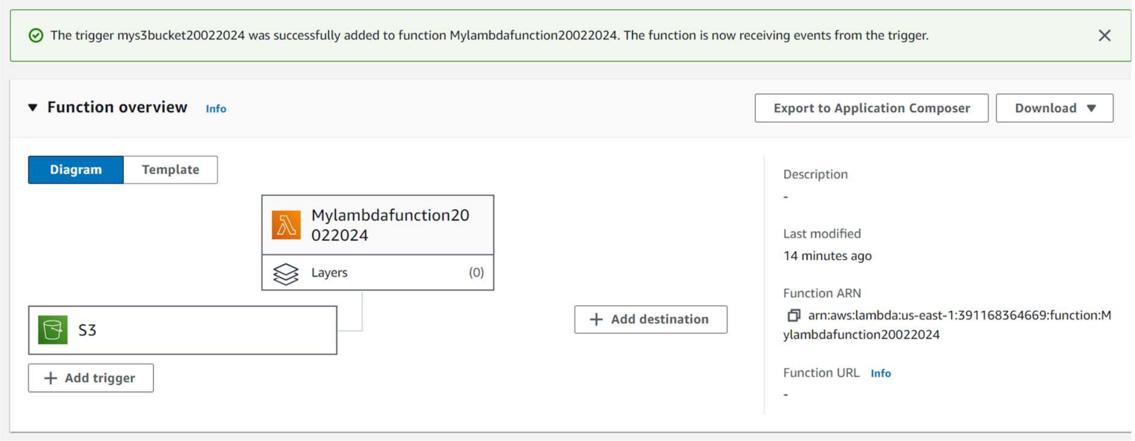
Recursive invocation

If your function writes objects to an S3 bucket, ensure that you are using different S3 buckets for input and output. Writing to the same bucket increases the risk of creating a recursive invocation, which can result in increased Lambda usage and increased costs. Learn more

☑ I acknowledge that using the same S3 bucket for both input and output is not recommended and that this configuration can cause recursive invocations, increased Lambda usage, and increased costs.

Lambda will add the necessary permissions for AWS S3 to invoke your Lambda function from this trigger. Learn more about the Lambda permissions model.

Click add



The trigger mys3bucket20022024 was successfully added to function Mylambdafunction20022024. The function is now receiving events from the trigger.

Description
-

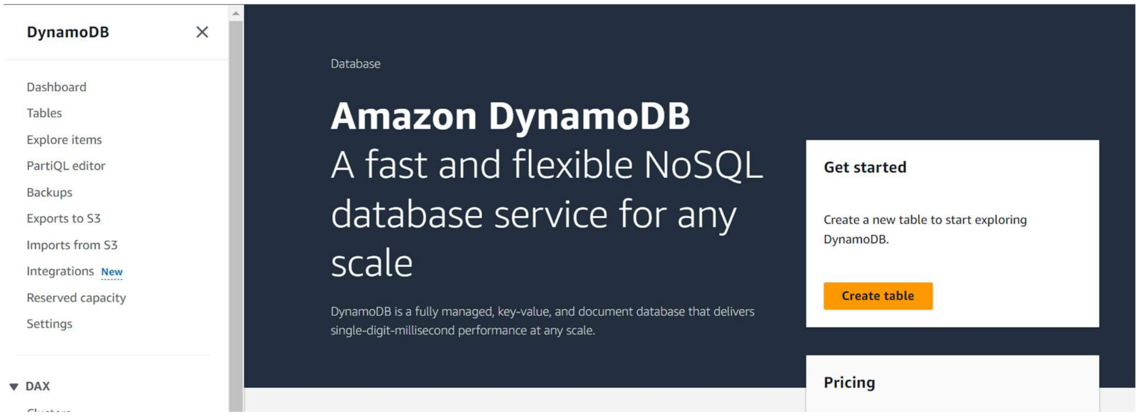
Last modified
14 minutes ago

Function ARN
arn:aws:lambda:us-east-1:391168364669:function:Mylambdafunction20022024

Function URL Info
-

Task3: Now go to dynamo db and crate a table

Here we will take the name of the table i.e "newtable" and partition key "unique". This value can be changed inside your code we used to integrate the dynamo DB.



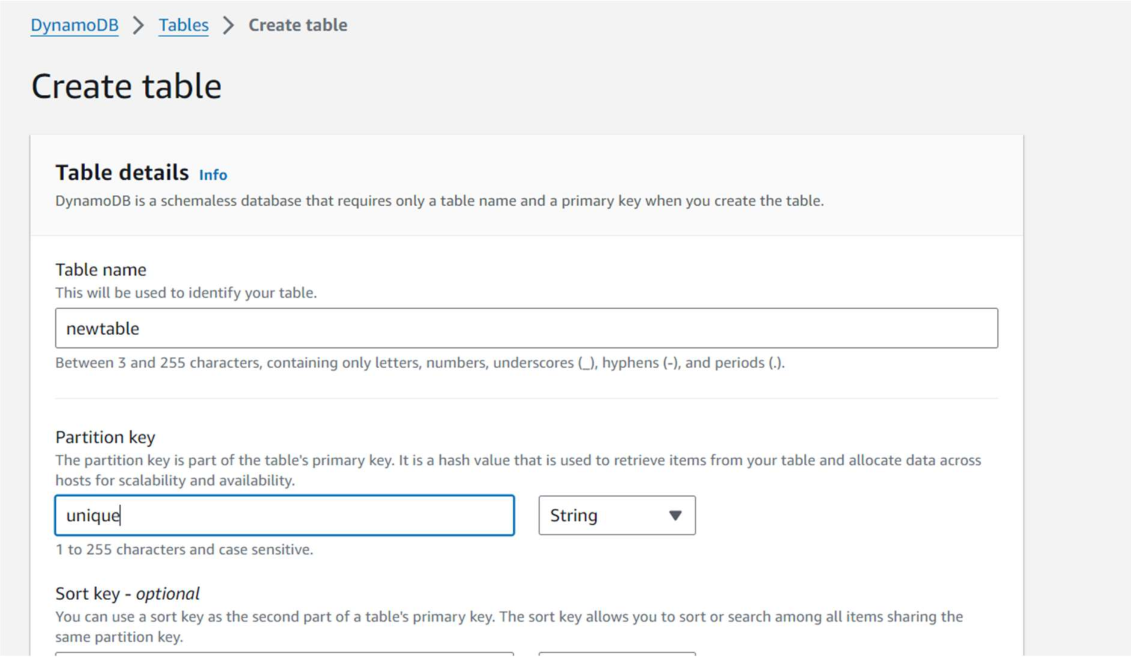
Database

Amazon DynamoDB
A fast and flexible NoSQL database service for any scale

DynamoDB is a fully managed, key-value, and document database that delivers single-digit-millisecond performance at any scale.

Get started

Create a new table to start exploring DynamoDB.



Create table

Table details

DynamoDB is a schemaless database that requires only a table name and a primary key when you create the table.

Table name

This will be used to identify your table.

newtable

Between 3 and 255 characters, containing only letters, numbers, underscores (_), hyphens (-), and periods (.).

Partition key

The partition key is part of the table's primary key. It is a hash value that is used to retrieve items from your table and allocate data across hosts for scalability and availability.

unique | String

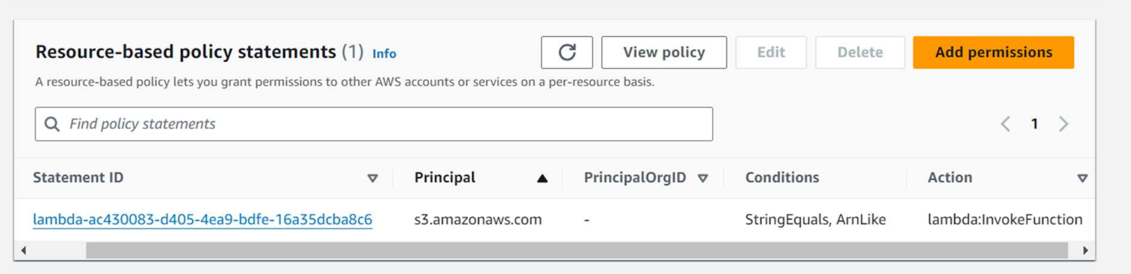
1 to 255 characters and case sensitive.

Sort key - *optional*

You can use a sort key as the second part of a table's primary key. The sort key allows you to sort or search among all items sharing the same partition key.

Create a table

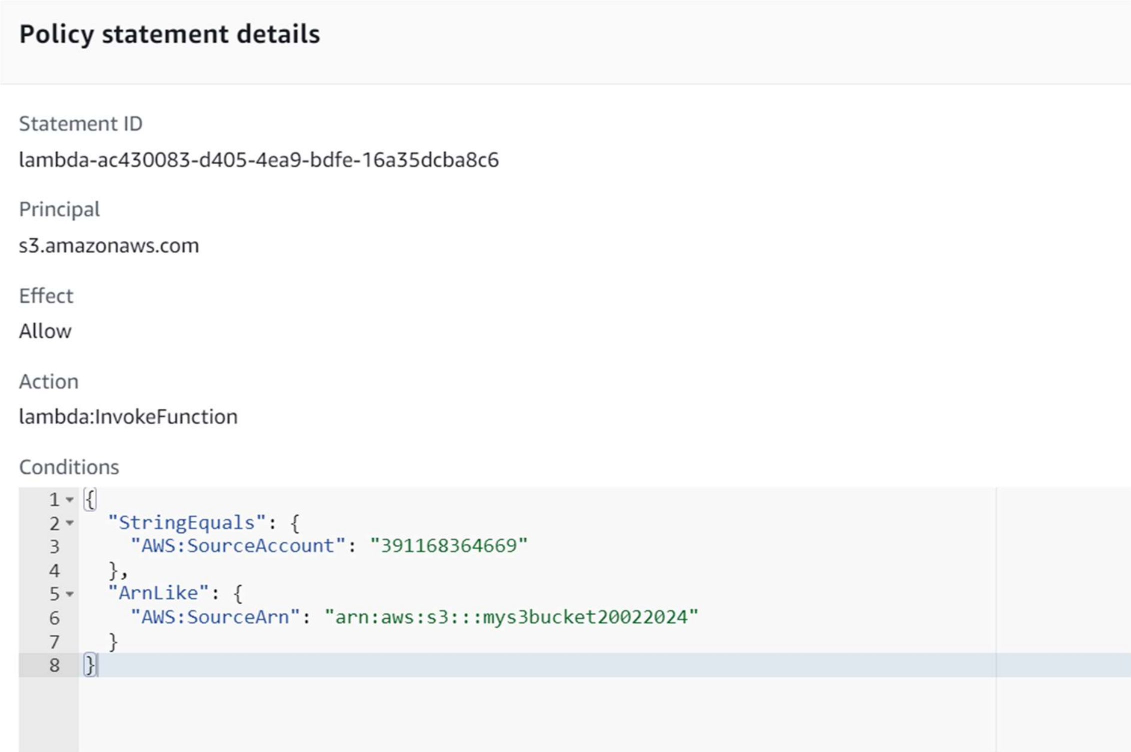
task4: now comes permission part
if you observe after adding s3 trigger,



Resource-based policy statements (1) Info

A resource-based policy lets you grant permissions to other AWS accounts or services on a per-resource basis.

Statement ID	Principal	PrincipalOrgID	Conditions	Action
lambda-ac430083-d405-4ea9-bdfe-16a35dcba8c6	s3.amazonaws.com	-	StringEquals, ArnLike	lambda:InvokeFunction



Policy statement details

Statement ID
lambda-ac430083-d405-4ea9-bdfe-16a35dcba8c6

Principal
s3.amazonaws.com

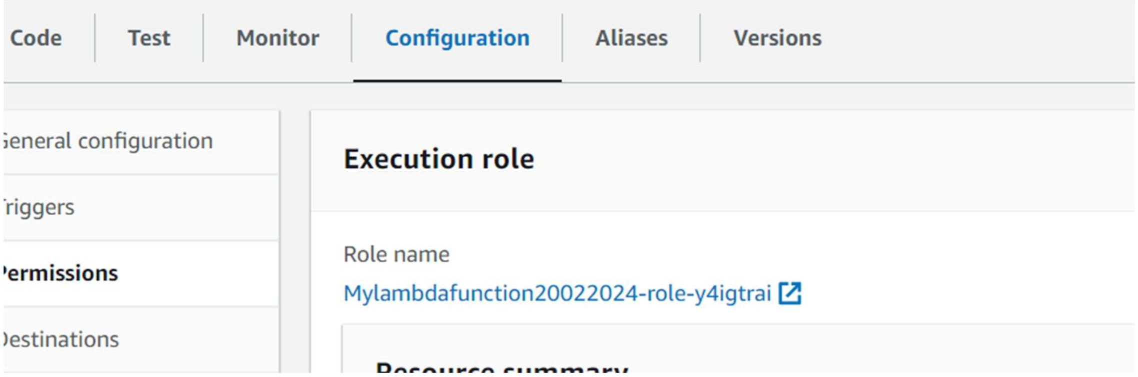
Effect
Allow

Action
lambda:InvokeFunction

Conditions

```
{
  "StringEquals": {
    "AWS:SourceAccount": "391168364669"
  },
  "ArnLike": {
    "AWS:SourceArn": "arn:aws:s3:::mys3bucket20022024"
  }
}
```

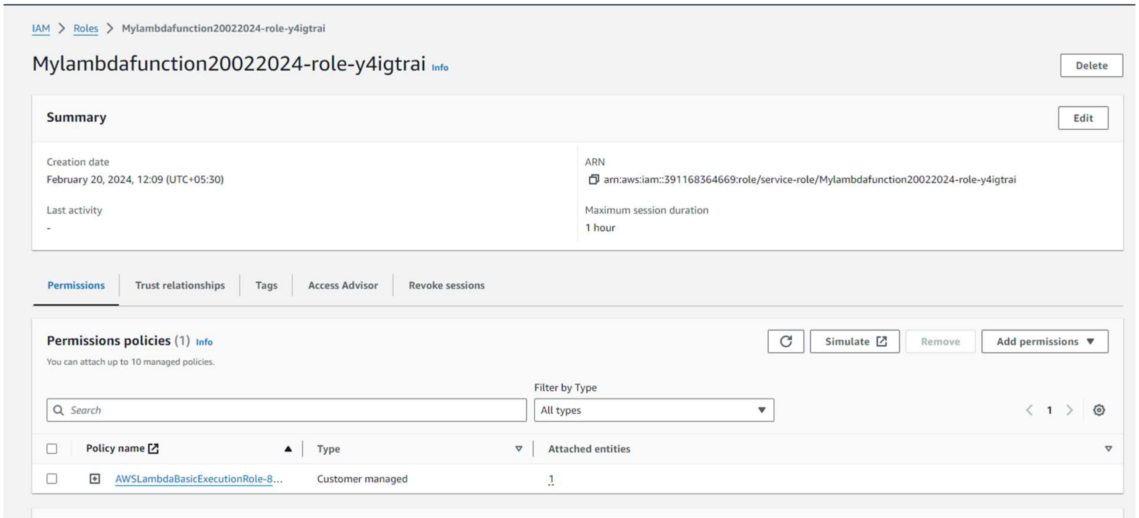
Now, if you observe lambda will have one execution role, so



Execution role

Role name
Mylambdafunction20022024-role-y4igtrai

Resource summary



IAM > Roles > Mylambdafunction20022024-role-y4igtrai

Mylambdafunction20022024-role-y4igtrai Info

Summary

Creation date
February 20, 2024, 12:09 (UTC+05:30)

Last activity
-

ARN
arn:aws:iam::391168364669:role/service-role/Mylambdafunction20022024-role-y4igtrai

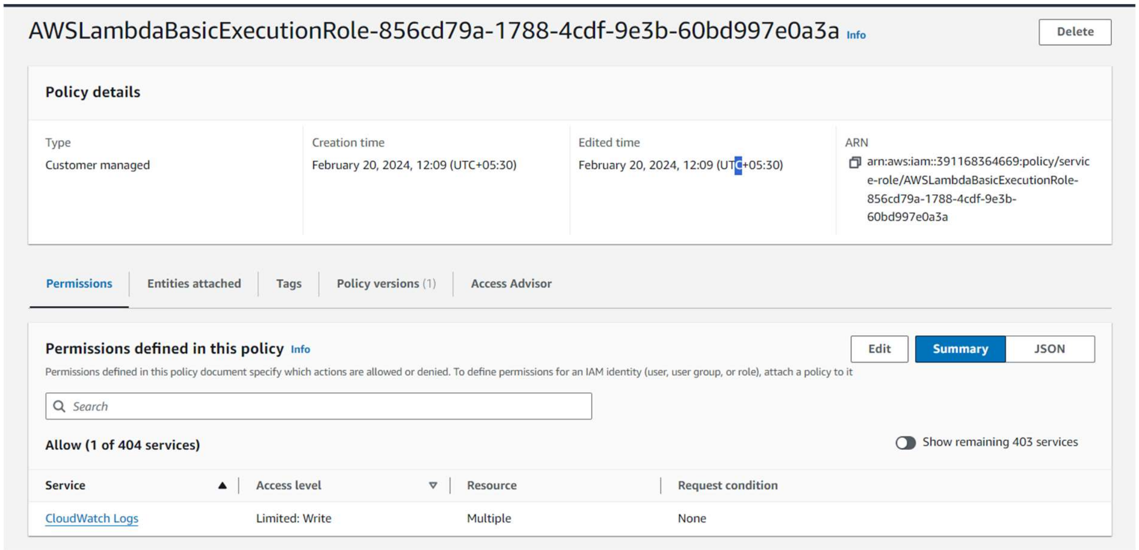
Maximum session duration
1 hour

Permissions policies (1) Info
You can attach up to 10 managed policies.

Filter by Type: All types

	Policy name	Type	Attached entities
□	AWSLambdaBasicExecutionRole-8...	Customer managed	1

By default, lambda will send logs to CloudWatch



AWSLambdaBasicExecutionRole-856cd79a-1788-4cdf-9e3b-60bd997e0a3a Info

Policy details

Type
Customer managed

Creation time
February 20, 2024, 12:09 (UTC+05:30)

Edited time
February 20, 2024, 12:09 (UTC+05:30)

ARN
arn:aws:iam::391168364669:policy/service-role/AWSLambdaBasicExecutionRole-856cd79a-1788-4cdf-9e3b-60bd997e0a3a

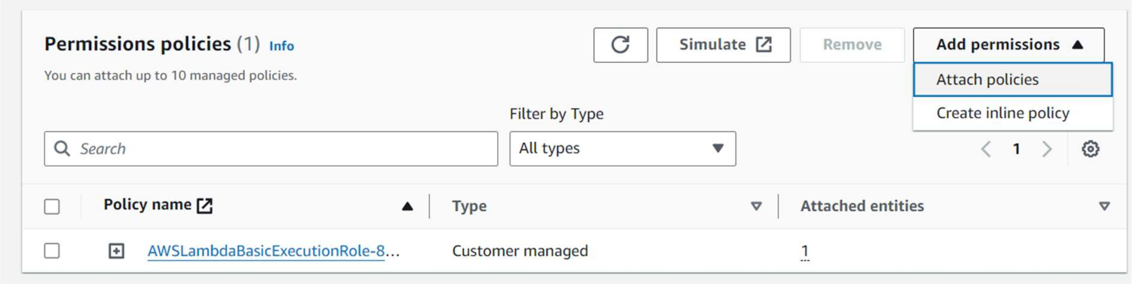
Permissions defined in this policy Info
Permissions defined in this policy document specify which actions are allowed or denied. To define permissions for an IAM identity (user, user group, or role), attach a policy to it

Allow (1 of 404 services)

Show remaining 403 services

Service	Access level	Resource	Request condition
CloudWatch Logs	Limited: Write	Multiple	None

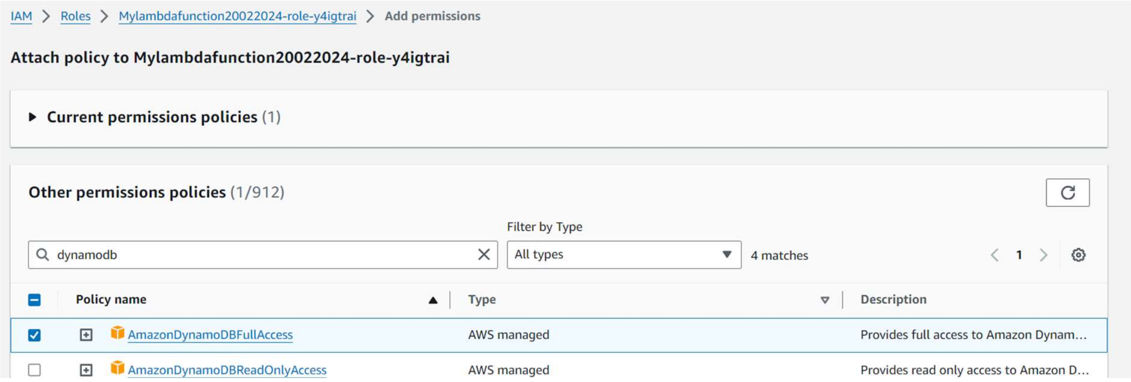
Now to get access to dynamo db., we must attach a DynamoDB policy to lambda (execution role
)



Permissions policies (1) Info
You can attach up to 10 managed policies.

Filter by Type: All types

	Policy name	Type	Attached entities
□	AWSLambdaBasicExecutionRole-8...	Customer managed	1



IAM > Roles > Mylambdafunction20022024-role-y4igtrai > Add permissions

Attach policy to Mylambdafunction20022024-role-y4igtrai

▸ Current permissions policies (1)

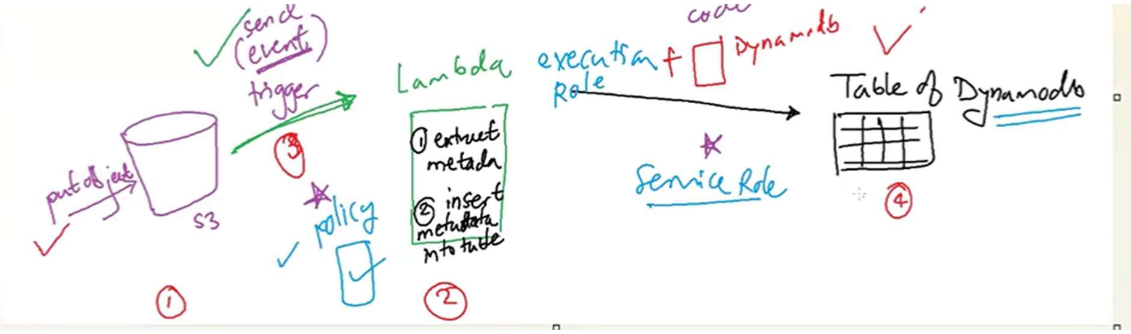
Other permissions policies (1/912)

Filter by Type

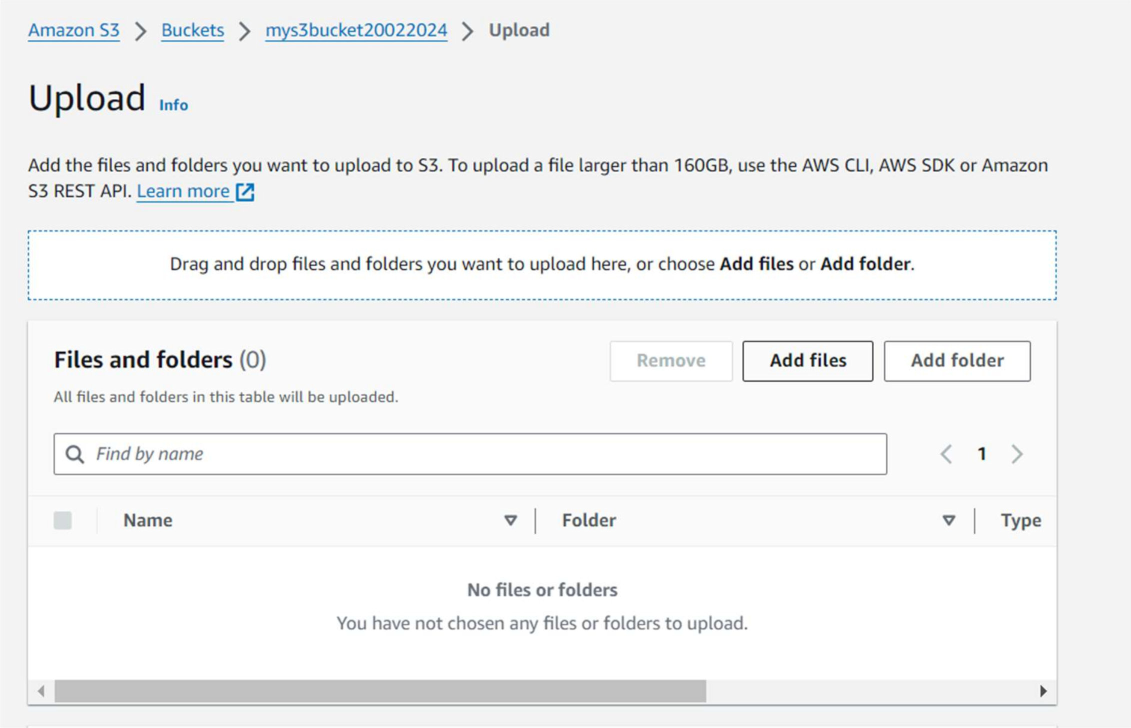
dynamodb | All types | 4 matches

☐	Policy name	Type	Description
☑	AmazonDynamoDBFullAccess	AWS managed	Provides full access to Amazon Dynam…
☐	AmazonDynamoDBReadOnlyAccess	AWS managed	Provides read only access to Amazon D…

architecture diagram:



Now I WILL UPLOAD files in s3,



Amazon S3 > Buckets > mys3bucket20022024 > Upload

Upload Info

Add the files and folders you want to upload to S3. To upload a file larger than 160GB, use the AWS CLI, AWS SDK or Amazon S3 REST API. Learn more

Drag and drop files and folders you want to upload here, or choose **Add files** or **Add folder**.

Files and folders (0)

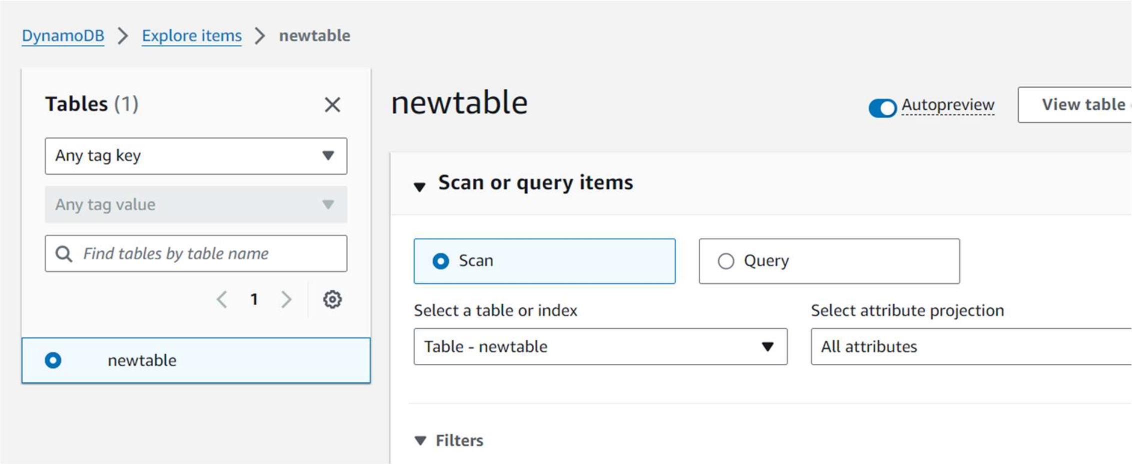
All files and folders in this table will be uploaded.

☐	Name	Folder	Type

No files or folders

You have not chosen any files or folders to upload.

After uploading, observe the dynamodb table , explore items, and click run(scan)



newtable

Scan or query items

● Scan ○ Query

Select a table or index: Table - newtable

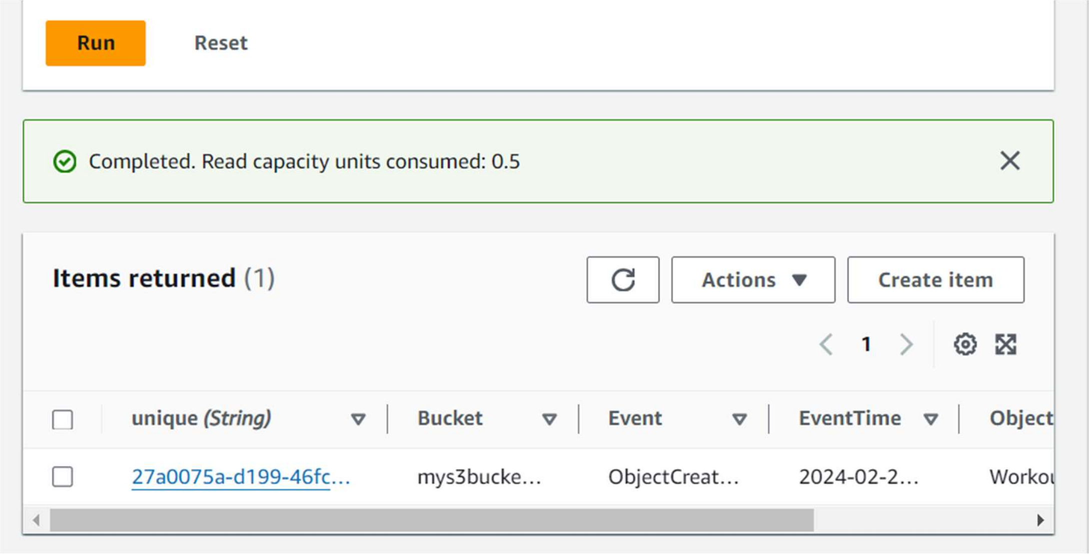
Select attribute projection: All attributes

Any tag key

Any tag value

Find tables by table name

● newtable

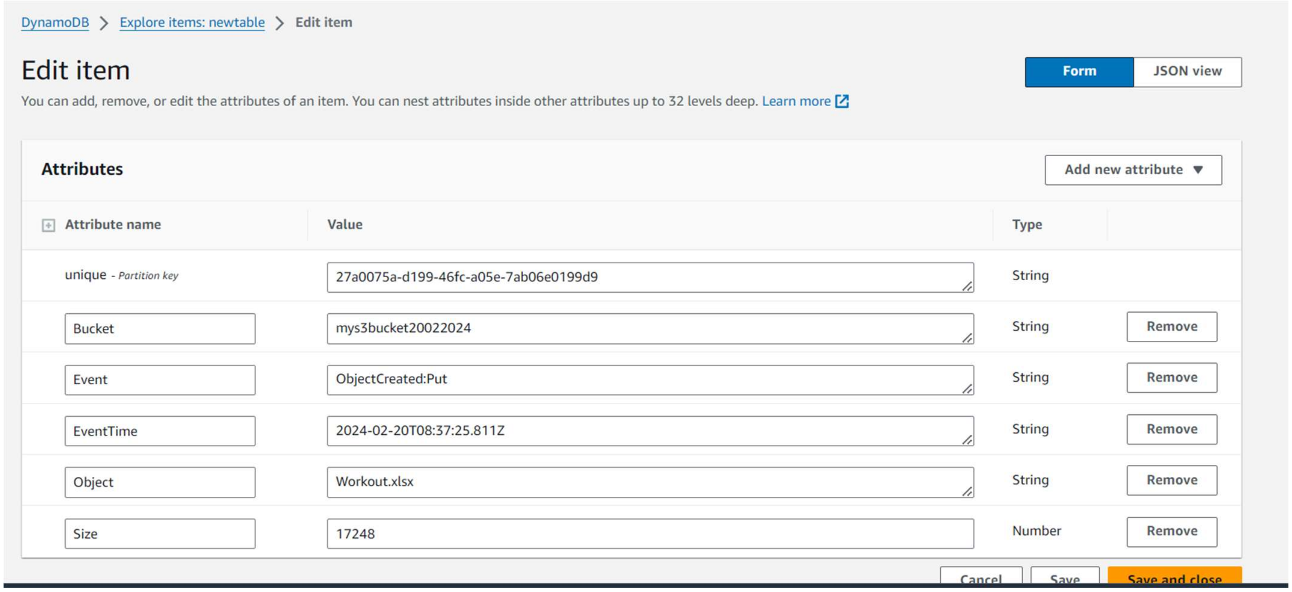


Completed. Read capacity units consumed: 0.5

Items returned (1)

□	unique (String)	Bucket	Event	EventTime	Object
□	27a0075a-d199-46fc…	mys3bucke…	ObjectCreat…	2024-02-2…	Workou

Output:



Edit item

You can add, remove, or edit the attributes of an item. You can nest attributes inside other attributes up to 32 levels deep. Learn more

Attributes

Attribute name	Value	Type
unique - *Partition key*	27a0075a-d199-46fc-a05e-7ab06e0199d9	String
Bucket	mys3bucket20022024	String
Event	ObjectCreated:Put	String
EventTime	2024-02-20T08:37:25.811Z	String
Object	Workout.xlsx	String
Size	17248	Number