Real-time Stream Processing using Kinesis

AWS Lambda and Amazon Kinesis can be used to process real-time streaming data for application activity tracking, transaction order processing, click stream analysis, data cleansing, metrics generation, log filtering, indexing, social media analysis, and IoT device data telemetry and metering.

- Create a Kinesis Stream
- Create a DynamoDB table named <stackname>-EventData
- Create Lambda Function 1 (<stackname>-DDBEventProcessor) which receives records from Kinesis and writes records to the DynamoDB table
- Create an IAM Role and Policy to allow the event processing Lambda function read from the Kinesis Stream and write to the DynamoDB table
- Create an IAM user with permission to put events in the Kinesis stream together with credentials for the user to use in an API client

Steps to achieve streaming data

Clone Github:

```
Microsoft Windows [Version 10.0.18363.1139]
(c) 2019 Microsoft Corporation. All rights reserved.

C:\Users\shive>git clone https://github.com/aws-samples/lambda-refarch-streamprocessing.git
Cloning into 'lambda-refarch-streamprocessing'...
remote: Enumerating objects: 150, done.
remote: Total 150 (delta 0), reused 0 (delta 0), pack-reused 150R
Receiving objects: 100% (150/150), 252.27 KiB | 3.28 MiB/s, done.
Resolving deltas: 100% (71/71), done.

C:\Users\shive>
```
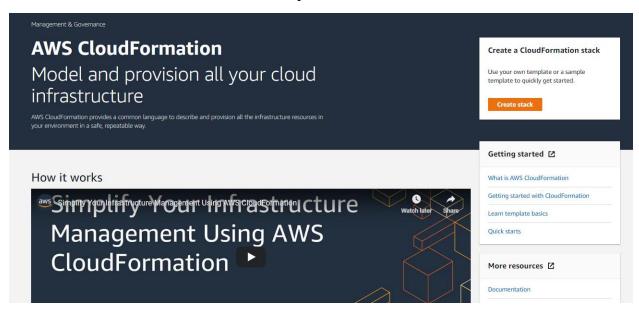
Open template.yml and change Runtime to nodejs12.x

```yaml
AWSTemplateFormatVersion: '2010-09-09'
Transform: AWS::Serverless-2016-10-31
Description: "Template to set up Kinesis stream, Lambda functions, S3 bucket, DynamoDB table and related IAM roles for AWS La
Parameters:
  LambdaS3Bucket:
    Type: String
    Default: awslambda-reference-architectures
    Description: Name of S3 bucket where Lambda function packages are stored.
  LambdaDDBEventProcessorS3Key:
    Type : String
    Default : stream-processing/ddb_eventprocessor.zip
    Description : Name of S3 key for Zip with Stream Processing DynamoDB Event Processor Lambda function package.
  LambdaDDBEventProcessorHandler:
    Type : String
    Default : ddb_eventprocessor.handler
    Description : Name of handler for Stream Processing DynamoDB Event Processor Lambda function.
Resources:
  EventStream:
    Type: 'AWS::Kinesis::Stream'
    Properties:
      ShardCount: 1
  DDBEventProcessor:
    Type: 'AWS::Serverless::Function'
    Properties:
      Description: Stream Processing DDB Event Processor
      Handler: !Ref LambdaDDBEventProcessorHandler
      MemorySize: 128
      Role: !GetAtt
        - EventProcessorExecutionRole
        - Arn
      Timeout: 10
      Runtime: nodejs12.x
      CodeUri:
        Bucket: !Ref LambdaS3Bucket
        Key: !Ref LambdaDDBEventProcessorS3Key
      Events:
```

Create AWS CloudFormation stack with template

Upload a template file

Give Stack Name

Specify stack details

Stack name

Stack name

kinesisTutorial

Stack name can include letters (A-Z and a-z), numbers (0-9), and dashes (-).

Parameters

Parameters are defined in your template and allow you to input custom values when you create or update a stack.

LambdaDDBEventProcessorHandler
Name of handler for Stream Processing DynamoDB Event Processor Lambda function.

ddb_eventprocessor.handler

LambdaDDBEventProcessorS3Key
Name of S3 key for Zip with Stream Processing DynamoDB Event Processor Lambda function package.

stream-processing/ddb_eventprocessor.zip

LambdaS3Bucket
Name of S3 bucket where Lambda function packages are stored.

awslambda-reference-architectures

Cancel   Previous   Next

Create Stack

Capabilities and transforms

ⓘ **Transforms might require access capabilities**

A transform might add Identity and Access Management (IAM) resources that could provide entities access to make changes to your AWS account. If a transform adds IAM resources, you must acknowledge their capabilities to create or update them. Ensure that you want to create or update the IAM resources, and that they have the minimum required permissions. In addition, if they have custom names, check that the names are unique within your AWS account. Learn more

☑ I acknowledge that AWS CloudFormation might create IAM resources.

☑ I acknowledge that AWS CloudFormation might create IAM resources with custom names.

☑ I acknowledge that AWS CloudFormation might require the following capability: CAPABILITY_AUTO_EXPAND

Cancel   Previous   Create change set   Create stack

kinesisTutorial

<div>Delete Update Stack actions ▾ Create stack ▾</div>

Stack info | Events | Resources | Outputs | Parameters | Template | Change sets

Events (1)

Search events

Timestamp ▽	Logical ID	Status	Status reason
2020-11-04 20:39:46 UTC-0500	kinesisTutorial	ⓘ CREATE_IN_PROGRESS	User Initiated

Resource

kinesisTutorial

<div>Delete Update Stack actions ▾ Create stack ▾</div>

Stack info | Events | Resources | Outputs | Parameters | Template | Change sets

Resources (8)

Search resources

Logical ID ▲	Physical ID ▽	Type ▽	Status ▽	Status reason ▽
ClientKeys	AKIA5CNM2IXXY5VZMLM2	AWS::IAM::AccessKey	⊘ CREATE_COMPLETE	-
ClientPolicy	kines-Clie-198K9UETPRYUO	AWS::IAM::Policy	ⓘ CREATE_IN_PROGRESS	Resource creation Initiated
DDBEventProcessor	kinesisTutorial-DDBEventProcessor-I1P572NH7CUR ↗	AWS::Lambda::Function	⊘ CREATE_COMPLETE	-
DDBEventProcessorStream	5e621fbd-c8c6-405b-9ba6-895c411d7dc9	AWS::Lambda::EventSourceMapping	⊘ CREATE_COMPLETE	-
EventDataTable	kinesisTutorial-EventData ↗	AWS::DynamoDB::Table	⊘ CREATE_COMPLETE	-
EventProcessorExecutionRole	kinesisTutorial-EventProcessorExecutionRole-15RTLSFI9FN5W ↗	AWS::IAM::Role	⊘ CREATE_COMPLETE	-
EventStream	kinesisTutorial-EventStream-1RM28O3455USR ↗	AWS::Kinesis::Stream	⊘ CREATE_COMPLETE	-
streamprocessingclient	kinesisTutorial-streamprocessingclient-WI41ACRENN3M ↗	AWS::IAM::User	⊘ CREATE_COMPLETE	-

Output

Stacks (1)

Filter by stack name

Active ▾   ● View nested   ‹ 1 ›

kinesisTutorial
2020-11-04 20:39:46 UTC-0500
⊘ CREATE_COMPLETE

kinesisTutorial

<div>Delete Update Stack actions ▾ Cr</div>

Stack info | Events | Resources | Outputs | Parameters | Template | Change sets

Outputs (4)

Search outputs

Key ▲	Value ▽	Description ▽	Export name
AccessKeyId	AKIA5CNM2IXXY5VZMLM2	AWS Access Key Id of stream processing client user	-
KinesisStream	kinesisTutorial-EventStream-1RM28O3455USR	The Kinesis stream used for ingestion.	-
Region	us-east-1	The region this template was launched in.	-
SecretAccessKey	PIDXEpFhCMNqu1AvNrOCncPRB1DfQHsyIVtk7J2q	AWS Secret Key of stream processing client user	-

Pip install boto3

Pip install TwitterAPI

```
C:\Python\Python36>pip install boto3
Requirement already satisfied: boto3 in c:\python\python36\lib\site-packages (1.16.7)
Requirement already satisfied: jmespath<1.0.0,>=0.7.1 in c:\python\python36\lib\site-packages (from boto3) (0.10.0)
Requirement already satisfied: s3transfer<0.4.0,>=0.3.0 in c:\python\python36\lib\site-packages (from boto3) (0.3.3)
Requirement already satisfied: botocore<1.20.0,>=1.19.7 in c:\python\python36\lib\site-packages (from boto3) (1.19.7)
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in c:\python\python36\lib\site-packages (from botocore<1.20.
,>=1.19.7->boto3) (2.8.1)
Requirement already satisfied: urllib3<1.26,>=1.25.4; python_version != "3.4" in c:\python\python36\lib\site-packages (
rom botocore<1.20.0,>=1.19.7->boto3) (1.25.11)
Requirement already satisfied: six>=1.5 in c:\python\python36\lib\site-packages (from python-dateutil<3.0.0,>=2.1->boto
ore<1.20.0,>=1.19.7->boto3) (1.15.0)

C:\Python\Python36>pip install TwitterAPI
Collecting TwitterAPI
  Downloading TwitterAPI-2.6.2.1.tar.gz (10 kB)
Requirement already satisfied: requests in c:\python\python36\lib\site-packages (from TwitterAPI) (2.24.0)
Collecting requests_oauthlib
  Downloading requests_oauthlib-1.3.0-py2.py3-none-any.whl (23 kB)
Requirement already satisfied: chardet<4,>=3.0.2 in c:\python\python36\lib\site-packages (from requests->TwitterAPI)
0.4)
Requirement already satisfied: certifi>=2017.4.17 in c:\python\python36\lib\site-packages (from requests->TwitterAPI)
020.6.20)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in c:\python\python36\lib\site-packages (from
uests->TwitterAPI) (1.25.11)
Requirement already satisfied: idna<3,>=2.5 in c:\python\python36\lib\site-packages (from requests->TwitterAPI) (2.10
Collecting oauthlib>=3.0.0
  Downloading oauthlib-3.1.0-py2.py3-none-any.whl (147 kB)
     |████████████████████████████████| 147 kB 3.3 MB/s
Using legacy 'setup.py install' for TwitterAPI, since package 'wheel' is not installed.
Installing collected packages: oauthlib, requests-oauthlib, TwitterAPI
    Running setup.py install for TwitterAPI ... done
Successfully installed TwitterAPI-2.6.2.1 oauthlib-3.1.0 requests-oauthlib-1.3.0

C:\Python\Python36>
```

Run python twitter2kinesis.py

```
*Dems do not want to abide by U S Constitution…
Why do you think Obama &amp; Fauci gave them 3.5 million? This has been in the works a long time.
I'm Hannah Strong and #IHaveWonPennsylvania because I'm from Australia and just felt like it
RT @GregRubini: Trump won Michigan:
TRUMP 54.5%
BIDEN   43.5%

what Fake News are saying is FAKE (including CROOKED Fox News).
wait and see…
RT @RL9631: "I heard the voice of God say, 'The enemy has intended to strike out Trump at a critical hour in history. B
t supernatural help…
RT @_burnett22_: Boiiiii if Biden win this I can't wait to start telling trump supports "respect him that's your presid
nt, if you don't li…
RT @Jali_Cat: ‼️FYI @RudyGiuliani @pnjaban @MarkMeadows @DHS_Wolf @TomFitton @TheJusticeDept @tedcruz

▣ ▣ ▣▣▣▣38,000 ballots mysteriously show…
RT @joncoopertweets: BREAKING: @JoeBiden has now pulled within just 37,000 votes of Trump in Georgia! ▣ ▣▣
#CountEveryVote
RT @jasithedon: it's the fact that trump supporters are getting escorted and protected by police but BLM activists are
etting killed by th…
RT @mattgaetz: Last hour the Fox News decision desk said Trump would get 45% of the outstanding Arizona vote.

He just got 57% of a batch…
RT @sethweathers: Trump has won the State of Georgia

All he needs now is AZ and he's won re-election
RT @JasonMillerinDC: .@realDonaldTrump is going to be the winner of Pennsylvania.
RT @jennycohn1: 6/ Former CBS News executive producer:
RT @andybg_andy: I am not a conspiracy theorist. However Trump brings back 35 new car factories and revitalized Michiga
```

Check DynamoDB for data produced

kinesisTutorial-EventData Close

Overview | Items | Metrics | Alarms | Capacity | Indexes | Global Tables | Backups | Contributor Insights | Triggers | Access control | Tags

Create item | Actions

Scan: [Table] kinesisTutorial-EventData: Username, Id

Viewing 1 to 100 items

Scan | [Table] kinesisTutorial-EventData: Username, Id

Add filter

Start search

Username	Id	Message	Timestamp
jorge ivan	1324176631871668224	RT @777jorgeivan: if Biden wins i want the cameras to follow trump aroun…	2020-11-05T02:28:05.000Z
josh_tre4	1324176640075878401	RT @lexcorinne: If trump loses and people start throwing a fit & going …	2020-11-05T02:28:07.000Z
lolo	1324176645008379904	Ik dis not real	2020-11-05T02:28:08.000Z
marleneyctoday Parler@MYCTrump2020	1324176631662071815	RT @TheBadDadder: PENNSYLVANIA CORRUPTION MACHINE!! @real…	2020-11-05T02:28:05.000Z
nae	1324176648489586690	So, Trump isn't concerned about if the states that he won cheated, right?	2020-11-05T02:28:09.000Z
rae	1324176628243726337	RT @BLACKGIRLMANI: Trump told his supporters not to vote by mail and …	2020-11-05T02:28:04.000Z
readsector	1324176645016690690	Trump files Georgia lawsuit over late ballots after suing PA and NV https://t…	2020-11-05T02:28:08.000Z
save the oceans	1324176632593051648	RT @jdtitan: The only way Trump gets to 270 now is if he loses 50 lbs	2020-11-05T02:28:05.000Z
sherry perry	1324176636116406274	@BretBaier Me too ! Bunch of #fakenews Weasels! Trump is winning this …	2020-11-05T02:28:06.000Z
xDrive	1324176628235165696	RT @kayleighmcenany: UNACCEPTABLE	2020-11-05T02:28:04.000Z
Nara	1324176640105238528	@yeriinroses Amg eu tô falando que só uma pessoa que não sabe interpre…	2020-11-05T02:28:07.000Z
Medusa	1324176645020917763	RT @rafaelkjls: o trump vai perder a eleição pra um tiktoker… o karma nun…	2020-11-05T02:28:08.000Z
michelle	1324176635973742593	RT @JewishWorker: The NYPD's unions endorsed Donald Trump and are …	2020-11-05T02:28:06.000Z
Arbiter	1324176649043193857	RT @zerohedge: TRUMP CAMPAIGN SAYS DECLARING VICTORY IN P…	2020-11-05T02:28:09.000Z
Chef Rocky #RESIST#BLM#ClimateActionNow!!	1324176641082462213	TOO LATE DR SCARF!… THOUSANDS DIED WHILE YOU SAT ON #D…	2020-11-05T02:28:07.000Z
Muscalex AUGBOM	1324176632744013825	RT @KatMeanJean: @TomiLahren The only way Trump gets to 270 now is…	2020-11-05T02:28:05.000Z
usJustCallMeSparkly us	1324176649336836096	RT @mattgaetz: Last hour the Fox News decision desk said Trump would …	2020-11-05T02:28:09.000Z
TheBullDroppa	1324176625488056321	RT @NeesaNicki: I'm seeing tons of signs in oro valley saying "CONGRAT…	2020-11-05T02:28:04.000Z

Create table | Delete table

Filter by table name

Choose a table … | Actions

Name
- kinesisTutorial-EventData
- NASA_Turbofan_Jet_Engine_Datase
- Temperatures