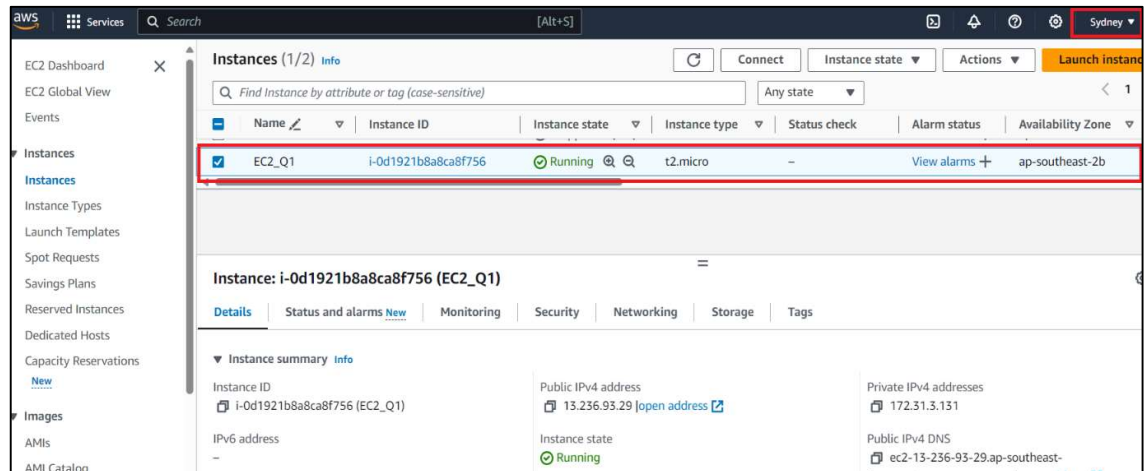
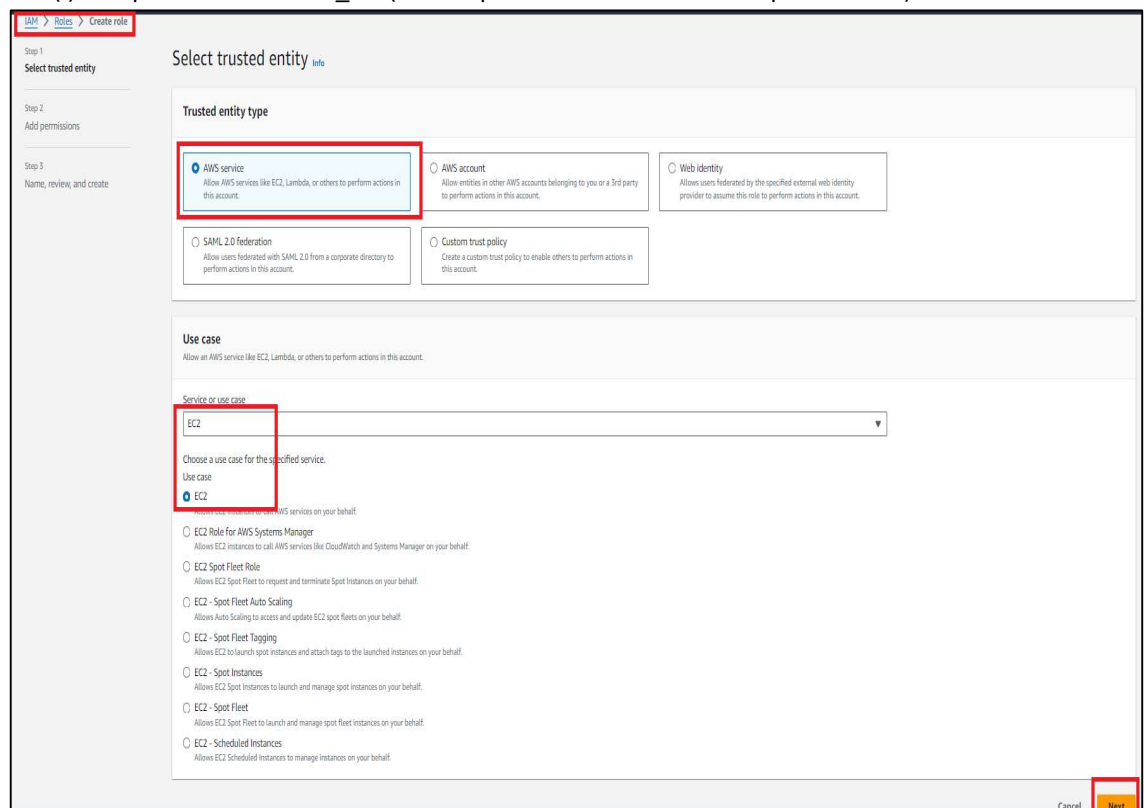


1) Create an EC2 instance: EC2_Q1



2) Granting permission to EC2_Q1 to access Kinesis:

(i) Setup IAM role for EC2_Q1 (Kinesis put record & CloudWatch permissions)



Step 1

Select trusted entity

Step 2

Add permissions

Step 3

Name, review, and create

Add permissions

Permissions policies (1/911)

Info

Choose one or more policies to attach to your new role.

Filter by Type

Q kinesis

X

All types

▼

10 matches

< 1 > ⚙

Policy name	Type
<input type="checkbox"/> AmazonKinesisAnalyticsFullAccess	AWS managed
<input type="checkbox"/> AmazonKinesisAnalyticsReadOnly	AWS managed
<input type="checkbox"/> AmazonKinesisFirehoseFullAccess	AWS managed
<input type="checkbox"/> AmazonKinesisFirehoseReadOnlyAccess	AWS managed
<input checked="" type="checkbox"/> AmazonKinesisFullAccess	AWS managed
<input type="checkbox"/> AmazonKinesisReadOnlyAccess	AWS managed

Step 1

Select trusted entity

Step 2

Add permissions

Step 3

Name, review, and create

Add permissions

Permissions policies (2/911)

Info

Choose one or more policies to attach to your new role.

Filter by Type

Q cloudwatch

X

All types

▼

29 matches

< 1 2 > ⚙

Policy name	Type
<input checked="" type="checkbox"/> CloudWatchFullAccess	AWS managed
<input type="checkbox"/> CloudWatchFullAccessV2	AWS managed
<input type="checkbox"/> CloudWatchLambdaInsightsExecutionRolePolicy	AWS managed

Step 1

Select trusted entity

Step 2

Add permissions

Step 3

Name, review, and create

Name, review, and create

Role details

Role name

role-awslogs-role-to-cloudwatch-role

role-awslogs-role-to-cloudwatch-role

Description

Add more information for this role.

Allows EC2 instances to call AWS services on your behalf.

Maximum 1024 characters. Use alphanumeric and '*' (asterisk) characters.

Step 1: Select trusted entities

Trust policy

```
1 {
2   "Version": "2012-10-17",
3   "Statement": [
4     {
5       "Effect": "Allow",
6       "Action": "iam:PassRole",
7       "Resource": "arn:aws:iam::123456789012:role/*",
8       "Condition": {
9         "StringEquals": {
10          "aws:PrincipalOrgID": "o-123456789012"
11        }
12      }
13    }
14  ]
15 }
```

Step 2: Add permissions

Permissions policy summary

Policy name	Type	Attached as
AmazonKinesisFullAccess	AWS managed	Permissions policy
CloudWatchFullAccess	AWS managed	Permissions policy

Step 3: Add tags

Add tags - optional

Tags are key-value pairs that you can add to AWS resources to help identify, organize, or search for resources.

No tags associated with the resource.

Add new tag

No more than 50 tags.

Cancel

Previous

Create role

Roles (5) [Info](#)

An IAM role is an identity you can create that has specific permissions with credentials that are valid for short durations. Roles can be assumed by entities that have permissions to do so.

Search

<input type="checkbox"/>	Role name	Trusted entities
<input type="checkbox"/>	AWSServiceRoleForSupport	AWS Service: support (Service-Linked Role)
<input type="checkbox"/>	AWSServiceRoleForTrustedAdvisor	AWS Service: trustedadvisor (Service-Linked Role)
<input checked="" type="checkbox"/>	EC2_Q1-Kinesis_role	AWS Service: ec2
<input type="checkbox"/>	Ec2Kinesisagentrole	AWS Service: ec2
<input type="checkbox"/>	KinesisFirehoseServiceRole-KDS-S3-Y-ap-southeast-2-1705209477854	AWS Service: firehose

(ii) Back to EC2 instances page to assign IAM role to EC2_Q1

EC2 Dashboard

EC2 Global View

Events

Instances

Instance Types

Launch Templates

Spot Requests

Savings Plans

Reserved Instances

Dedicated Hosts

Successfully started i-0d1921b8a8ca8f756

Instances (1/2) [Info](#)

Find Instance by attribute or tag (case-sensitive)

Any state

Connect

Instance state

Actions

Launch instances

<input type="checkbox"/>	Name	Instance ID	Instance state	Instance type	Status check	Alarm
<input type="checkbox"/>	Ec21	i-09d6db7d9f3a6b506	Stopped	t2.micro	-	View
<input checked="" type="checkbox"/>	EC2_Q1	i-0d1921b8a8ca8f756	Running	t2.micro	-	View

Instance: i-0d1921b8a8ca8f756 (EC2_Q1)

Details

Status and alarms [New](#)

Monitoring

Security

Networking

Storage

Change security groups

Get Windows password

Modify IAM role

Connect

View details

Manage instance state

Instance settings

Networking

Security

Image and templates

Monitor and troubleshoot

EC2 > Instances > i-0d1921b8a8ca8f756 > Modify IAM role

Modify IAM role [Info](#)

Attach an IAM role to your instance.

Instance ID

i-0d1921b8a8ca8f756 (EC2_Q1)

IAM role

Select an IAM role to attach to your instance or create a new role if you haven't created any. The role you select replaces any roles that are currently attached to your instance.

Choose IAM role

Q |

No IAM Role

Choose this option to detach an IAM role

Ec2Kinesisagentrole

arn:aws:iam::851725364041:instance-profile/Ec2Kinesisagentrole

EC2_Q1-Kinesis_role

arn:aws:iam::851725364041:instance-profile/EC2_Q1-Kinesis_role

Create new IAM role

Cancel

Update IAM role

EC2_Q1-Kinesis_role

EC2_Q1

i-0d1921b8a8ca8f756

Running

t2.micro

2/2 checks passed

View alarms

3) Connect to EC2_Q1

Kinesis

Sydney

Analytics

Amazon Kinesis services

Collect, process, and analyze data streams in real time.

Get started

- ☒ **Kinesis Data Streams**
Collect streaming data with a data stream.
- ☐ **Kinesis Data Firehose**
Process and deliver streaming data with data delivery stream.
- ☐ **Managed Apache Flink**
Formerly Kinesis Data Analytics
Analyze streaming data with data analytics application.

Create data stream
Create data stream

Create data stream [Info](#)

Data stream configuration

Data stream name

stream-1

Acceptable characters are uppercase and lowercase letters, numbers, underscores, hyphens and periods.

Data stream capacity [Info](#)

☒ **On-demand**
Use this mode when your data stream's throughput requirements are unpredictable and variable. With on-demand mode, your data stream's capacity scales automatically.

☐ **Provisioned**
Use provisioned mode when you can reliably estimate throughput requirements of your data stream. With provisioned mode, your data stream's capacity is fixed.

Total data stream capacity
By default, data streams with on-demand mode scale throughput automatically to accommodate traffic of up to 200 MiB per second and 200,000 records per second for the write capacity. If traffic exceeds capacity, your data stream will throttle. To request capacity increase up to 2GB per second write and 4GB per second read, submit a [support ticket](#).

Write capacity
Maximum
200 MiB/second and 200,000 records/second

Read capacity
Maximum (per consumer)
400 MiB/second
Up to 2 default consumers. Use Enhanced Fan-Out (EFO) for more consumers. EFO supports adding upto 20 consumers, each having a dedicated throughput.

On-demand mode has a pay-per-throughput pricing model. See [Kinesis pricing for on-demand mode](#).

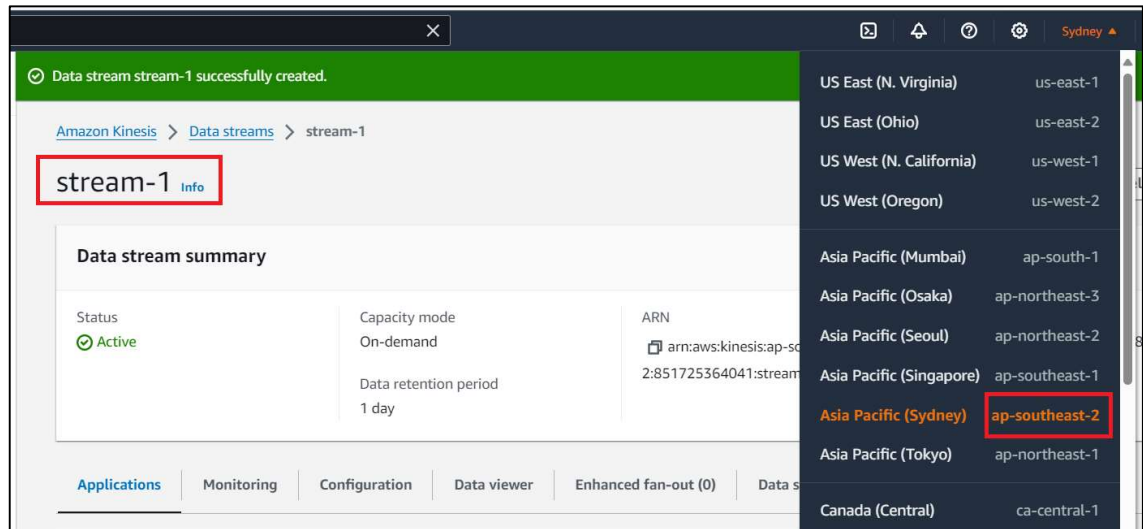
Data stream settings

You can edit the settings after the data stream has been created and is in the active status.

Setting	Value	Editable after creation
Capacity mode	On-demand	✔ Yes
Data retention period	1 day	✔ Yes
Server-side encryption	Disabled	✔ Yes
Monitoring enhanced metrics	Disabled	✔ Yes
Tags	-	✔ Yes
Data stream sharing policy	No policy	✔ Yes

Cancel

Create data stream



- 8) Run commands to configure Kinesis Agent:

clear

sudo nano /etc/aws-kinesis/agent.json

Copy/paste, and save:

```
{
  "cloudwatch.emitMetrics":true,
  "kinesis.endpoint":"https://kinesis.ap-southeast-2.amazonaws.com",
  "flows":[
    {
      "filePattern":"/opt/stream-1/logfile.log",
      "kinesisStream":"stream-1"
    }
  ]
}
```

- 9) Give ownership to Kinesis Agent user to access "stream-1" folder:

sudo chown aws-kinesis-agent-user:aws-kinesis-agent-user -R /opt/stream-1

- 10) Run Python program at the background to generate data, and test the logfile

sudo python3 /opt/stream-1/stream-1.py &

sudo tail -f /opt/stream-1/logfile.log

Ctrl+C to stop testing


```

2024-01-28 10:11:59,202Putting to stream: {'random': '72', 'timestamp': '1706436719', 'id': 'stream-1'}
2024-01-28 10:12:04,207Putting to stream: {'random': '51', 'timestamp': '1706436724', 'id': 'stream-1'}
2024-01-28 10:12:09,213Putting to stream: {'random': '48', 'timestamp': '1706436729', 'id': 'stream-1'}
2024-01-28 10:12:14,218Putting to stream: {'random': '45', 'timestamp': '1706436734', 'id': 'stream-1'}
2024-01-28 10:12:19,221Putting to stream: {'random': '22', 'timestamp': '1706436739', 'id': 'stream-1'}
2024-01-28 10:12:24,226Putting to stream: {'random': '16', 'timestamp': '1706436744', 'id': 'stream-1'}
2024-01-28 10:12:29,231Putting to stream: {'random': '63', 'timestamp': '1706436749', 'id': 'stream-1'}
2024-01-28 10:12:34,237Putting to stream: {'random': '75', 'timestamp': '1706436754', 'id': 'stream-1'}
2024-01-28 10:12:39,242Putting to stream: {'random': '58', 'timestamp': '1706436759', 'id': 'stream-1'}
2024-01-28 10:12:44,247Putting to stream: {'random': '54', 'timestamp': '1706436764', 'id': 'stream-1'}
2024-01-28 10:12:49,253Putting to stream: {'random': '7', 'timestamp': '1706436769', 'id': 'stream-1'}
2024-01-28 10:12:54,258Putting to stream: {'random': '79', 'timestamp': '1706436774', 'id': 'stream-1'}
2024-01-28 10:12:59,261Putting to stream: {'random': '7', 'timestamp': '1706436779', 'id': 'stream-1'}
2024-01-28 10:13:04,266Putting to stream: {'random': '2', 'timestamp': '1706436784', 'id': 'stream-1'}
2024-01-28 10:13:09,271Putting to stream: {'random': '17', 'timestamp': '1706436789', 'id': 'stream-1'}
2024-01-28 10:13:14,277Putting to stream: {'random': '86', 'timestamp': '1706436794', 'id': 'stream-1'}
2024-01-28 10:13:19,282Putting to stream: {'random': '97', 'timestamp': '1706436799', 'id': 'stream-1'}
2024-01-28 10:13:24,287Putting to stream: {'random': '74', 'timestamp': '1706436804', 'id': 'stream-1'}
2024-01-28 10:13:29,293Putting to stream: {'random': '28', 'timestamp': '1706436809', 'id': 'stream-1'}
2024-01-28 10:13:34,298Putting to stream: {'random': '45', 'timestamp': '1706436814', 'id': 'stream-1'}
2024-01-28 10:13:39,304Putting to stream: {'random': '10', 'timestamp': '1706436819', 'id': 'stream-1'}
2024-01-28 10:13:44,309Putting to stream: {'random': '74', 'timestamp': '1706436824', 'id': 'stream-1'}
2024-01-28 10:13:49,314Putting to stream: {'random': '28', 'timestamp': '1706436829', 'id': 'stream-1'}
2024-01-28 10:13:54,320Putting to stream: {'random': '33', 'timestamp': '1706436834', 'id': 'stream-1'}
2024-01-28 10:13:59,321Putting to stream: {'random': '94', 'timestamp': '1706436839', 'id': 'stream-1'}
2024-01-28 10:14:04,326Putting to stream: {'random': '20', 'timestamp': '1706436844', 'id': 'stream-1'}
2024-01-28 10:14:09,331Putting to stream: {'random': '23', 'timestamp': '1706436849', 'id': 'stream-1'}

```

i-0d1921b8a8ca8f756 (EC2_Q1)

PublicIPs: 13.55.1.50 PrivateIPs: 172.31.3.131

11) Start Kinesis Agent, and test status & log:

sudo service aws-kinesis-agent start

sudo service aws-kinesis-agent status

sudo tail -f /var/log/aws-kinesis-agent/aws-kinesis-agent.log

Ctrl+C to stop testing

```

^C
[ec2-user@ip-172-31-3-131 stream-1]$ sudo service aws-kinesis-agent start
Reloading systemd: [ OK ]
Starting aws-kinesis-agent (via systemctl): [ OK ]
[ec2-user@ip-172-31-3-131 stream-1]$ sudo service aws-kinesis-agent status
● aws-kinesis-agent.service - LSB: Daemon for Amazon Kinesis Agent.
   Loaded: loaded (/etc/rc.d/init.d/aws-kinesis-agent; generated)
   Active: active (running) since Sun 2024-01-28 10:19:58 UTC; 2min 9s ago
     Docs: man:systemd-sysv-generator(8)
  Process: 32433 ExecStart=/etc/rc.d/init.d/aws-kinesis-agent start (code=exited, status=0/SUCCESS)
    Tasks: 19 (limit: 1114)
   Memory: 100.2M
      CPU: 4.287s
   CGroup: /system.slice/aws-kinesis-agent.service
           └─32440 runuser aws-kinesis-agent-user -s /bin/sh -c "/usr/bin/start-aws-kinesis-agent "
             └─32442 /bin/java -server -Xms32m -Xmx512m -Dlog4j.configurationFile=file:///etc/aws-kinesis/lo
us>

```

```

60] FileTailer[kinesis:stream-1:/opt/stream-1/logfile.log]: Tailer Progress: Tailer has parsed 138 records (51960 bytes), transformed 0 records, skipped 0 records, and
has successfully sent 129 records to destination.
2024-01-28 10:31:28.124+0000 (Agent.MetricsEmitter RUNNING) com.amazon.kinesis.streaming.agent.Agent [INFO] Agent: Progress: 138 records parsed (51960 bytes), and 129
records sent successfully to destinations. Uptime: 690057ms
2024-01-28 10:31:58.121+0000 (FileTailer[kinesis:stream-1:/opt/stream-1/logfile.log].MetricsEmitter RUNNING) com.amazon.kinesis.streaming.agent.tailing.FileTailer [IN
60] FileTailer[kinesis:stream-1:/opt/stream-1/logfile.log]: Tailer Progress: Tailer has parsed 144 records (52582 bytes), transformed 0 records, skipped 0 records, and
has successfully sent 142 records to destination.
2024-01-28 10:31:58.124+0000 (Agent.MetricsEmitter RUNNING) com.amazon.kinesis.streaming.agent.Agent [INFO] Agent: Progress: 144 records parsed (52582 bytes), and 142
records sent successfully to destinations. Uptime: 720057ms
2024-01-28 10:32:08.500+0000 (cw-metrics-publisher) com.amazon.kinesis.streaming.agent.metrics.CWPublisherRunnable [INFO] Successfully published 5 datums.
2024-01-28 10:32:28.121+0000 (FileTailer[kinesis:stream-1:/opt/stream-1/logfile.log].MetricsEmitter RUNNING) com.amazon.kinesis.streaming.agent.tailing.FileTailer [IN
60] FileTailer[kinesis:stream-1:/opt/stream-1/logfile.log]: Tailer Progress: Tailer has parsed 150 records (53206 bytes), transformed 0 records, skipped 0 records, and
has successfully sent 142 records to destination.
2024-01-28 10:32:28.124+0000 (Agent.MetricsEmitter RUNNING) com.amazon.kinesis.streaming.agent.Agent [INFO] Agent: Progress: 150 records parsed (53206 bytes), and 142
records sent successfully to destinations. Uptime: 750057ms
2024-01-28 10:32:58.121+0000 (FileTailer[kinesis:stream-1:/opt/stream-1/logfile.log].MetricsEmitter RUNNING) com.amazon.kinesis.streaming.agent.tailing.FileTailer [IN
60] FileTailer[kinesis:stream-1:/opt/stream-1/logfile.log]: Tailer Progress: Tailer has parsed 156 records (53830 bytes), transformed 0 records, skipped 0 records, and
has successfully sent 155 records to destination.
2024-01-28 10:32:58.124+0000 (Agent.MetricsEmitter RUNNING) com.amazon.kinesis.streaming.agent.Agent [INFO] Agent: Progress: 156 records parsed (53830 bytes), and 155
records sent successfully to destinations. Uptime: 780057ms
2024-01-28 10:33:09.119+0000 (cw-metrics-publisher) com.amazon.kinesis.streaming.agent.metrics.CWPublisherRunnable [INFO] Successfully published 5 datums.
2024-01-28 10:33:28.121+0000 (FileTailer[kinesis:stream-1:/opt/stream-1/logfile.log].MetricsEmitter RUNNING) com.amazon.kinesis.streaming.agent.tailing.FileTailer [IN
60] FileTailer[kinesis:stream-1:/opt/stream-1/logfile.log]: Tailer Progress: Tailer has parsed 162 records (54453 bytes), transformed 0 records, skipped 0 records, and
has successfully sent 155 records to destination.
2024-01-28 10:33:28.124+0000 (Agent.MetricsEmitter RUNNING) com.amazon.kinesis.streaming.agent.Agent [INFO] Agent: Progress: 162 records parsed (54453 bytes), and 155
records sent successfully to destinations. Uptime: 810057ms
^C
lec2-user@ip-172-31-3-131 stream-1$

```

12) Check incoming data at Kinesis data stream AWS console: **data is incoming to Kinesis data stream successfully**

Amazon Kinesis > Data streams > stream-1

stream-1 Info Delete

Data stream summary

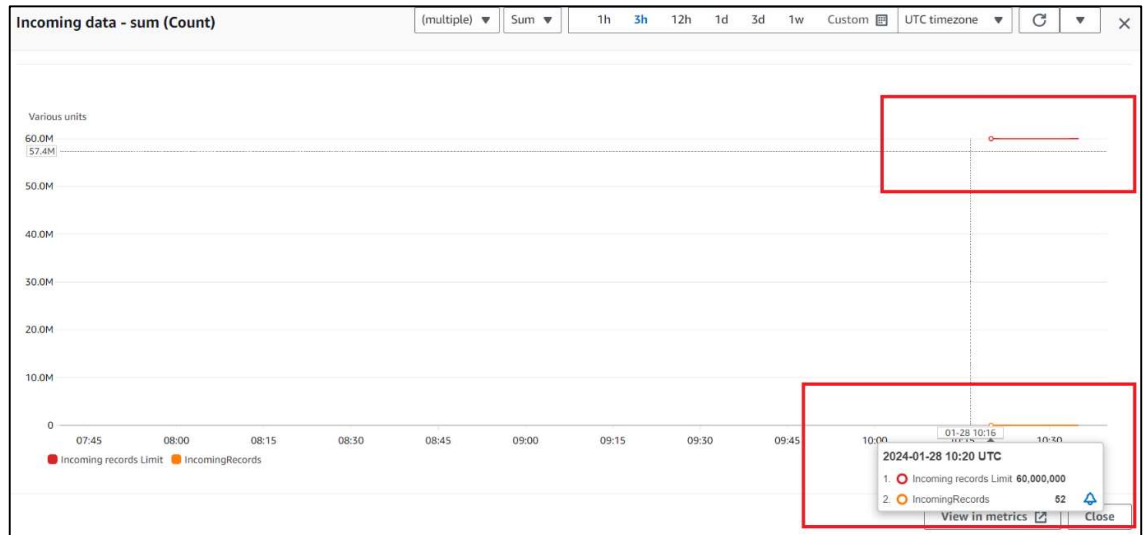
Status Active	Capacity mode On-demand	ARN arn:aws:kinesis:ap-southeast-2:851725364041:stream/stream-1	Creation time January 28, 2024 at 16:45 GMT+8
Data retention period 1 day			

Applications
Monitoring
Configuration
Data viewer
Enhanced fan-out (0)
Data stream sharing - new
EventBridge Pipes

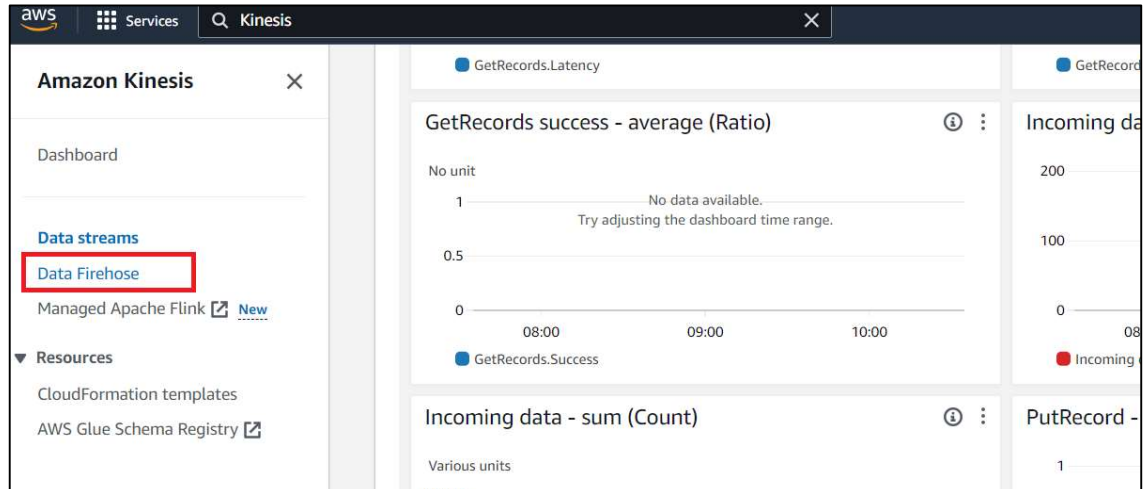
Stream metrics Info

☐ Alarm recommendations

3h
1d
1w
UTC timezone
Add to dashboard



13) Configure Kinesis Firehose



Amazon Kinesis Data Firehose

Real-time streaming delivery for any data, at any scale, and at low-cost.

Amazon Kinesis Data Firehose provides the easiest way to reliably ingest, transform, and deliver streaming data into data lakes, data warehouses, and analytics services.

Getting started

Create a delivery stream that processes and delivers streaming data to destinations.

[Create delivery stream](#)

How it works

Pricing (Asia Pacific (Sydney))

► Amazon Kinesis Data Firehose: How it works

Choose source and destination

Specify the source and the destination for your delivery stream. You cannot change the source and destination of your delivery stream once it has been created.

Source [Info](#)

Amazon Kinesis Data Streams

Destination [Info](#)

Amazon S3

Source settings

Kinesis data stream

arn:aws:kinesis:ap-southeast-2:851725364041:stream/stream-1

Browse

Create [↗](#)

Format: arn:aws:kinesis:[Region]:[AccountId]:stream/[StreamName]

Destination settings [Info](#)

Specify the destination settings for your delivery stream.

S3 bucket

s3://s3jzm

Browse

Create [↗](#)

Format: s3://bucket

New line delimiter

You can configure your delivery stream to add a new line delimiter between records in objects that are delivered to Amazon S3.

☒ Not enabled

☐ Enabled

Dynamic partitioning [Info](#)

Dynamic partitioning enables you to create targeted data sets by partitioning streaming S3 data based on partitioning keys. You can partition your source data with inline parsing and/or the specified AWS Lambda function. You can enable dynamic partitioning only when you create a new delivery stream. You cannot enable dynamic partitioning for an existing delivery stream. Enabling dynamic partitioning incurs additional costs per GiB of partitioned data. For more information, see [Kinesis Data Firehose pricing](#). [↗](#)

☒ Not enabled

☐ Enabled

S3 bucket prefix - optional

By default, Kinesis Data Firehose appends the prefix "YYYY/MM/dd/HH" (in UTC) to the data it delivers to Amazon S3. You can override this default by specifying a custom prefix that includes expressions that are evaluated at runtime.

Enter a prefix

You can repeat the same keys in your S3 bucket prefix. Maximum S3 bucket prefix characters: 1024.

S3 bucket error output prefix - optional

You can specify an S3 bucket error output prefix to be used in error conditions. This prefix can include expressions for Kinesis Data Firehose to evaluate at runtime.

Enter a prefix

► Buffer hints, compression and encryption

The fields below are pre-populated with the recommended default values for S3. Pricing may vary depending on storage and request costs.

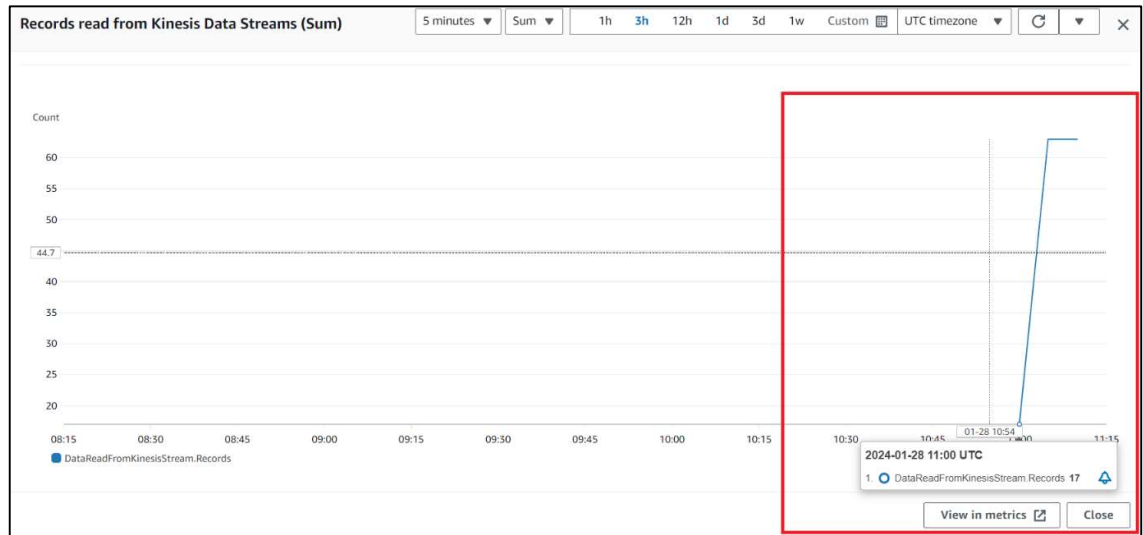
► Advanced settings

Server-side encryption not enabled; error logging enabled; IAM role KinesisFirehoseServiceRole-KDS-S3-S-ap-southeast-2-1706439300041; no tags.

Cancel

Create delivery stream

- 14) Check if data is picking up by Kinesis Firehose under “Monitoring” tab: **data is successfully picking up by Kinesis Firehose.**



15) Check if data stream is delivered into S3 bucket: **data stream is successfully delivered into S3**

Amazon S3

Amazon S3 > Buckets > s3jzm > 2024/ > 01/ > 28/ > 11/

11/

Objects (12) Info

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly

Find objects by prefix

	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	KDS-S3-SyMa5-1-2024-01-28-11-04-13-08c9f008-9fcf-4de7-b1e2-953d92222166	-	January 28, 2024, 19:09:14 (UTC+08:00)	1.8 KB	Standard
<input type="checkbox"/>	KDS-S3-SyMa5-1-2024-01-28-11-04-13-4fe95ea4-b38c-4e6f-90d3-5c0c0d7773de	-	January 28, 2024, 19:09:14 (UTC+08:00)	1.9 KB	Standard
<input type="checkbox"/>	KDS-S3-SyMa5-1-2024-01-28-11-04-13-a81f5316-b2bf-4865-8e07-05abab4cb5fb	-	January 28, 2024, 19:09:14 (UTC+08:00)	1.3 KB	Standard
<input type="checkbox"/>	KDS-S3-SyMa5-1-2024-01-28-11-04-13-e33352ca-5d39-408c-b6c7-42ff2b92691e	-	January 28, 2024, 19:09:14 (UTC+08:00)	1.5 KB	Standard
<input type="checkbox"/>	KDS-S3-SyMa5-1-2024-01-28-11-09-38-117ce904-6336-4bef-	-	January 28, 2024, 19:14:39 (UTC+08:00)	1.4 KB	Standard