

Part II

Start an EC2 Instance

The screenshot shows the AWS EC2 Management Console interface during the instance creation process, specifically Step 7: Review Instance Launch.

AMI Details:

- AMI: Amazon Linux AMI 2018.03.0 (HVM), SSD Volume Type - ami-0080e4c5bc078760e
- Free tier eligible
- Description: The Amazon Linux AMI is an EBS-backed, AWS-supported image. The default image includes AWS command line tools, Python, Ruby, Perl, and Java. The repositories include Docker, PHP, MySQL, PostgreSQL, and other packages.
- Root Device Type: ebs Virtualization type: hvm

Instance Type:

Instance Type	ECUs	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS-Optimized Available	Network Performance
t2.micro	Variable	1	1	EBS only	-	Low to Moderate

Security Groups:

Security Group ID	Name	Description
sg-0a29df4b263092e23	launch-wizard-1	launch-wiz

Inbound Rules (All selected security groups inbound rules):

Type	Protocol	Port Range
HTTP	TCP	80
HTTP	TCP	80
SSH	TCP	22

Terminal Output:

```
wansMacBookPro:Downloads phoebewu$ ssh -i "MyKeypair.pem" ec2-user@ec2-3-82-122-59.compute-1.amazonaws.com
Last login: Thu Dec 13 20:13:28 2018 from mobile-166-170-35-251.mycingular.net
[ec2-user@ip-172-31-94-229 ~]$
```

Bottom Navigation:

Cancel Previous Launch

Feedback English (US) © 2008 - 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

Firehose delivery streams

Kinesis Firehose - Create delivery stream

Step 1: Name and source

New delivery stream

Delivery streams load data, automatically and continuously, to the destinations that you specify. Kinesis Firehose resources are not covered under the [AWS Free Tier](#), and [usage-based charges apply](#). For more information, see [Kinesis Firehose pricing](#).

Delivery stream name* apache-fake-log

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

Choose source

Choose how you would prefer to send records to the delivery stream.

Firehose data flow overview

Source* Direct PUT or other sources

Choose this option to send records directly to the delivery stream, or to send records from

Step 4: Configure settings

Step 5: Review

Name and source

Delivery stream name apache-fake-log

Source Direct PUT or other sources

After creating the delivery stream, send records directly to the delivery stream, or send records from AWS IoT, CloudWatch Logs, or CloudWatch Events.

Process records

Source record transformation Disabled

Record format conversion Disabled

Destination

Destination Amazon S3

S3 bucket apache-fake-log-bucket1

S3 bucket Prefix None

Settings

S3 buffer conditions 5 MB or 300 seconds

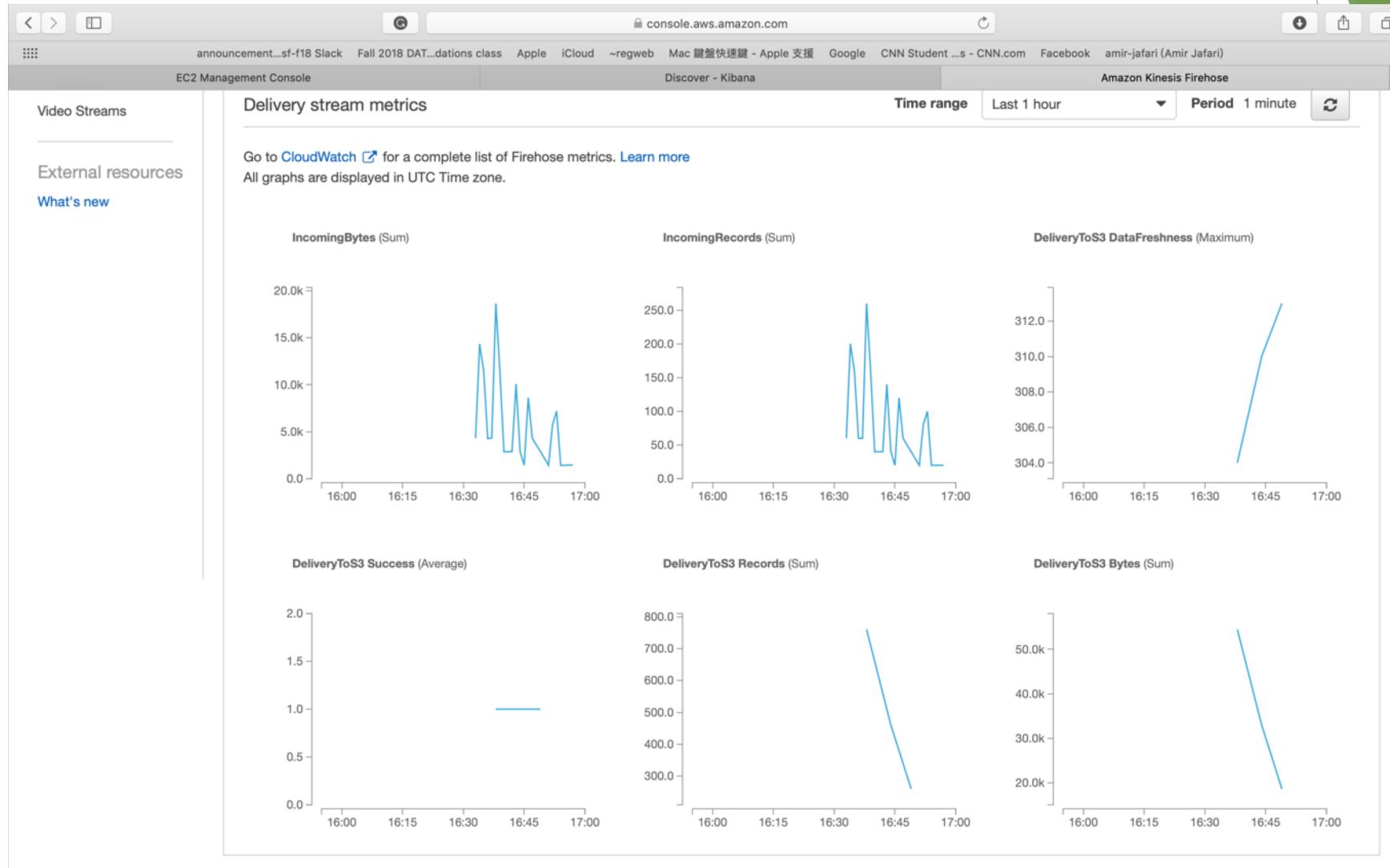
Compression Disabled

Encryption Disabled

Error logging Enabled

IAM role firehose_delivery_role

Test Firehose delivery streams with demo data



S3 bucket

The screenshot shows the AWS S3 Management Console interface. The URL in the browser is `console.aws.amazon.com`. The top navigation bar includes links for EC2 Management Console, Discover - Kibana, Amazon Kinesis Firehose, and S3 Management Console. The S3 Management Console tab is active. The sidebar shows navigation paths: Amazon S3 > apache-fake-log-bucket1 / 2018 / 12 / 13 / 21. The main area is titled "Overview". A search bar contains the placeholder text "Type a prefix and press Enter to search. Press ESC to clear.". Below the search bar are buttons for "Upload", "+ Create folder", "Download", and "Actions". To the right, it shows the region "US East (N. Virginia)" with a refresh icon. The table lists five log files:

<input type="checkbox"/>	Name	Last modified	Size	Storage class
<input type="checkbox"/>	apache-fake-log-1-2018-12-13-33-51-46d6ce05-dda0-48f6-9678-ef20a3c733cf	Dec 13, 2018 4:38:57 PM GMT-0500	53.2 KB	Standard
<input type="checkbox"/>	apache-fake-log-1-2018-12-13-21-38-56-aca9beae-ab38-4ff6-a3a3-b50e95deb7d4	Dec 13, 2018 4:44:08 PM GMT-0500	32.2 KB	Standard
<input type="checkbox"/>	apache-fake-log-1-2018-12-13-21-44-07-9fe0c194-e597-4904-b8d6-f13ee1c964bd	Dec 13, 2018 4:49:21 PM GMT-0500	18.2 KB	Standard
<input type="checkbox"/>	apache-fake-log-1-2018-12-13-21-51-43-17561df7-b141-489d-ad57-7625728ac7...	Dec 13, 2018 4:56:55 PM GMT-0500	25.3 KB	Standard
<input type="checkbox"/>	apache-fake-log-1-2018-12-13-21-56-58-6cf386da-c8d9-4c5f-b1b0-11bc6f9bdfaf	Dec 13, 2018 5:02:13 PM GMT-0500	18.2 KB	Standard

At the bottom, there are links for "Feedback", "English (US)", and "Feedback". The footer contains the copyright notice "© 2008 - 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved." and links for "Privacy Policy" and "Terms of Use".

Create an Amazon Elasticsearch Service Domain

The screenshot shows the AWS Elasticsearch Service console with the domain 'web-log-summary' selected. The left sidebar shows 'My domains' with 'web-log-summary' highlighted. The main page displays the domain configuration with various tabs like 'Overview', 'Cluster health', 'Indices', 'Monitoring', 'Logs', and 'Upgrade history'. A message box通知 a new service software release R20180914-P2 is available. The domain status is Active, and it's running Elasticsearch version 6.3. The endpoint is https://search-web-log-summary-skeb4ai5kwvpayixck7hb2stfa.us-east-1.es.amazonaws.com. The domain ARN is arn:aws:es:us-east-1:146747544651:domain/web-log-summary, and the Kibana URL is https://search-web-log-summary-skeb4ai5kwvpayixck7hb2stfa.us-east-1.es.amazonaws.com/_plugin/kibana/. The instance type is m4.large.elasticsearch (default), and there is 1 instance. Storage is EBS, with General Purpose (SSD) volume type and 10 GB volume size. Both Encryption at rest and Node-to-node encryption are disabled. The upgrade status is listed as '-'.

Domain status: Active

Elasticsearch version: 6.3

Endpoint: <https://search-web-log-summary-skeb4ai5kwvpayixck7hb2stfa.us-east-1.es.amazonaws.com>

Domain ARN: arn:aws:es:us-east-1:146747544651:domain/web-log-summary

Kibana: https://search-web-log-summary-skeb4ai5kwvpayixck7hb2stfa.us-east-1.es.amazonaws.com/_plugin/kibana/

Availability zones: 1

Instance type: m4.large.elasticsearch (default)

Number of instances: 1

Storage type: EBS

EBS volume type: General Purpose (SSD)

EBS volume size: 10 GB

Encryption at rest: Disabled

Node-to-node encryption: Disabled

Upgrade status: -

Feedback English (US) © 2008 - 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

Create an Amazon Kinesis Analytics Application

The screenshot shows the AWS Kinesis Analytics console interface. On the left, a sidebar navigation includes 'Amazon Kinesis' (selected), 'Dashboard', 'Data Streams', 'Data Firehose', 'Data Analytics' (selected), 'Video Streams', 'External resources', and 'What's new'. The main content area displays a success message: 'Successfully created Application web-log-aggregation-tutorial. Next, choose Connect streaming data.' Below this, the 'Source' section shows a 'Streaming data' configuration with a 'Connect streaming data' button. The 'Real time analytics' section contains a placeholder for SQL queries. The right side of the screen shows the 'Access permissions*' section with two options: 'Create / update IAM role kinesis-analytics-web-log-aggregation-tutorial-us-east-1' (selected) and 'Choose from IAM roles that Kinesis Analytics can assume'. The 'Schema' section is titled 'Schema discovery can generate a schema using recent records from the source. Schema column names are the same as in the source, unless they contain special characters, repeated column names, or reserved keywords.' It includes tabs for 'Raw', 'Lambda output', and 'Formatted' (selected). A 'Refresh stream sample' button is at the top of the schema preview table. The schema table has columns: ROWTIME (TIMESTAMP), ticker_symbol (VARCHAR(4)), sector (VARCHAR(16)), change (REAL), price (REAL), and APPROXIMATE_ARRIVAL_TIME (TIMESTAMP). The table displays 10 rows of sample data.

ROWTIME TIMESTAMP	ticker_symbol VARCHAR(4)	sector VARCHAR(16)	change REAL	price REAL	APPROXIMATE_ARRIVAL_TIME TIMESTAMP
2018-12-14 20:17:25.455	WAS	RETAIL	0.35	12.8	2018-12-14 20:17:24.63
2018-12-14 20:17:25.455	MJN	RETAIL	7.01	178.41	2018-12-14 20:17:24.63
2018-12-14 20:17:25.455	BAC	FINANCIAL	-0.47	15.97	2018-12-14 20:17:24.63
2018-12-14 20:17:25.455	TGT	RETAIL	-2.57	56.61	2018-12-14 20:17:24.63
2018-12-14 20:17:25.455	ABC	RETAIL	1.39	29.94	2018-12-14 20:17:24.63
2018-12-14 20:17:25.455	WSB	RETAIL	-0.3	8.41	2018-12-14 20:17:24.63
2018-12-14 20:17:25.455	WAS	RETAIL	-0.22	12.58	2018-12-14 20:17:24.63
2018-12-14 20:17:25.455	UHN	FINANCIAL	16.44	573.66	2018-12-14 20:17:24.63
2018-12-14 20:17:25.455	UHN	FINANCIAL	-9.28	564.38	2018-12-14 20:17:24.63
2018-12-14 20:17:25.455	PJN	RETAIL	-0.72	31.67	2018-12-14 20:17:24.63
2018-12-14 20:17:25.455	DFT	RETAIL	-2.8	85.34	2018-12-14 20:17:24.63

View the Aggregated Streaming Data

console.aws.amazon.com

EC2 Management Console Amazon Kinesis Firehose CloudWatch Management Console Amazon Kinesis Data Stream

Save and run SQL Add SQL from templates Download SQL SQL Reference guide Kinesis Data generator tool

```
15 "APPROXIMATE_ARRIVAL_TIME" timestamp,  
16 "ticker_symbol" VARCHAR(12),  
17 "symbol_count" integer;  
18  
19 CREATE OR REPLACE PUMP "STREAM_PUMP" AS  
20 INSERT INTO "DESTINATION_SQL_STREAM"  
21 SELECT STREAM STEPC("SOURCE_SQL_STREAM_001".ROWTIME BY INTERVAL '60' SECOND) AS "ingest_time",  
22 STEPC("SOURCE_SQL_STREAM_001".APPROXIMATE_ARRIVAL_TIME BY INTERVAL '60' SECOND) AS "APPROXIMATE_ARRIVAL_TIME",  
23 "ticker_symbol",  
24 COUNT(*) AS "symbol_count"  
25 FROM "SOURCE_SQL_STREAM_001"  
26 GROUP BY "ticker_symbol",  
27 STEPC("SOURCE_SQL_STREAM_001".ROWTIME BY INTERVAL '60' SECOND),  
28 STEPC("SOURCE_SQL_STREAM_001".APPROXIMATE_ARRIVAL_TIME BY INTERVAL '60' SECOND);
```

Application status: RUNNING

Source data Real-time analytics Destination

In-application streams:

- DESTINATION_SQL_STREAM
- error_stream

Pause results New results are added every 2-10 seconds. The results below are sampled.

Scroll to bottom when new results arrive.

Filter by column name

2018-12-14 21:15:00.0	2018-12-14 21:14:00.0	2018-12-14 21:14:00.0	CRM	2
2018-12-14 21:15:00.0	2018-12-14 21:14:00.0	2018-12-14 21:14:00.0	WMT	1
2018-12-14 21:15:00.0	2018-12-14 21:14:00.0	2018-12-14 21:14:00.0	VVS	2
2018-12-14 21:15:00.0	2018-12-14 21:14:00.0	2018-12-14 21:14:00.0	BNM	1
2018-12-14 21:15:00.0	2018-12-14 21:14:00.0	2018-12-14 21:14:00.0	QWE	2
2018-12-14 21:15:00.0	2018-12-14 21:14:00.0	2018-12-14 21:14:00.0	AZL	1
2018-12-14 21:15:00.0	2018-12-14 21:14:00.0	2018-12-14 21:14:00.0	HJV	2
2018-12-14 21:15:00.0	2018-12-14 21:14:00.0	2018-12-14 21:14:00.0	AMZN	1
2018-12-14 21:15:00.0	2018-12-14 21:14:00.0	2018-12-14 21:14:00.0	IVD	4

