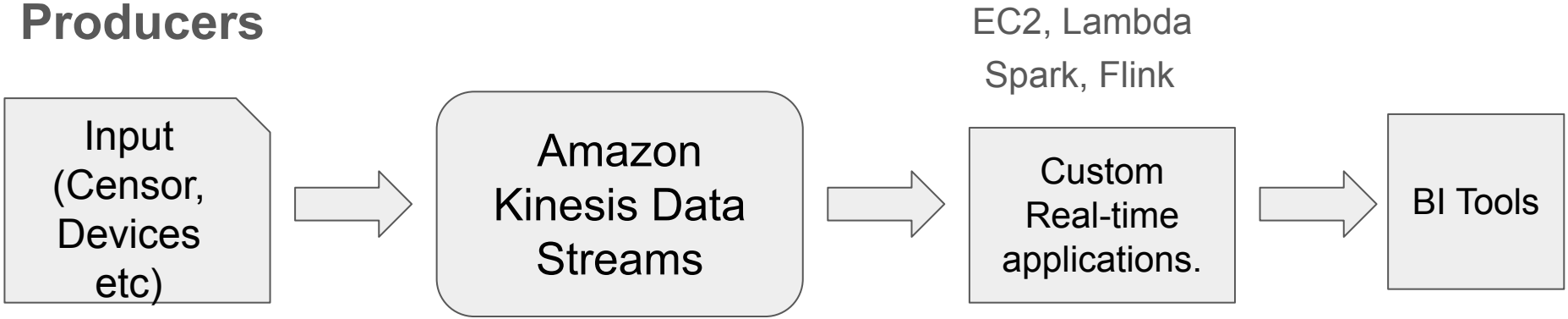
Amazon Kinesis Data Streams

Stream

Amazon Kinesis Data Stream can collect and process large streams

Kinesis Data Stream Applications

Producers



Capture and Send data to Amazon Kinesis Data Streams

Ingests and Store data streams for processing

Do the processing

Dashboarding

Scenarios for Kinesis Data Streams

Accelerated log and data feed intake

Real-time metrics and reporting

Real-time data analytics

Complex stream processing

Producers

Capture and Send data to Amazon Kinesis Data Streams

Ingests and Store data streams for processing

Questions

What is the primary purpose of Amazon Kinesis Data Streams?

- A) Store static data
- B) Collect and process real-time data streams
- C) Manage relational databases
- D) Generate static reports

Which of the following is a common use case for Kinesis Data Streams?

- A) Periodic batch processing
- B) Accelerated log and data feed intake
- C) Offline data warehousing
- D) Manual data entry

How does Kinesis Data Streams ensure data durability and elasticity?

- A) By batching data on servers
- B) Delaying data retrieval
- C) Putting data into streams
- D) Using static data storage

What is a key advantage of using the Kinesis Client Library?

- A) Real-time analytics
- B) Fault-tolerant data consumption
- C) Manual data archiving
- D) Static report generation

Which scenarios are typical for using Kinesis Data Streams? (Select all that apply)

- A) Batch processing
- B) Real-time metrics and reporting
- C) Offline data analysis
- D) Accelerated log and data feed intake

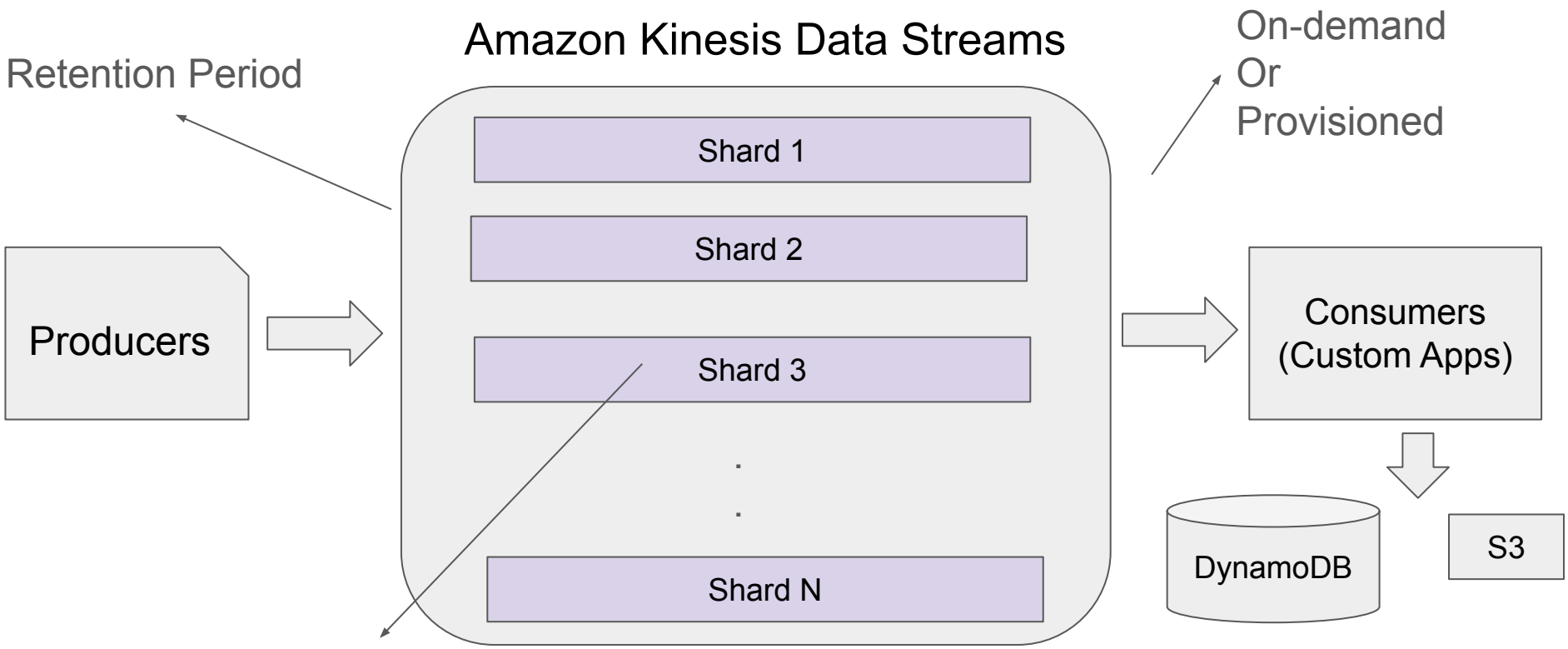
What are benefits of using Kinesis Data Streams? (Select all that apply)

- A) Operational burden relief
- B) Delayed data retrieval
- C) Static data storage
- D) Fault-tolerant data consumption

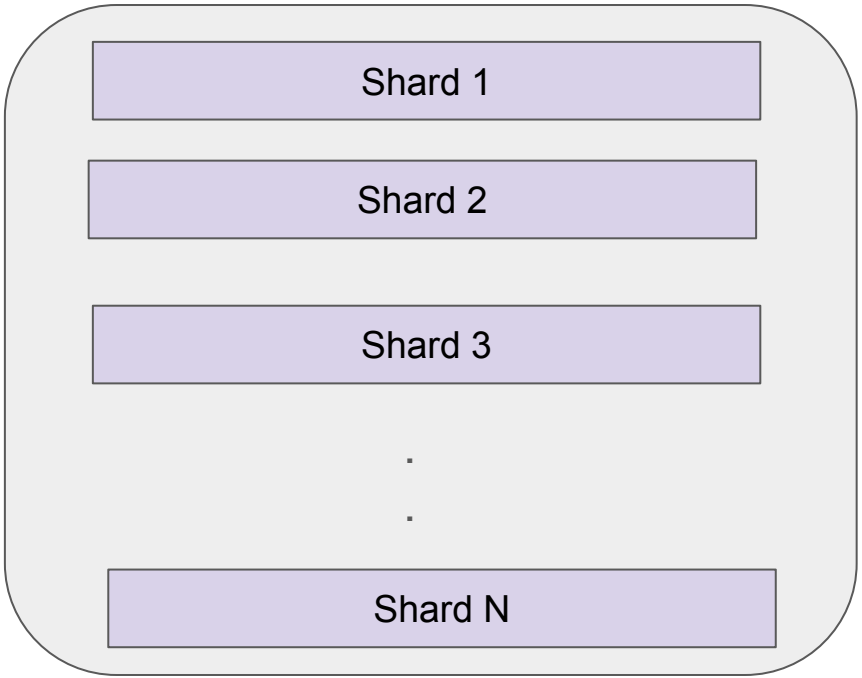
In what ways can multiple Kinesis Data Streams applications process data concurrently? (Select all that apply)

- A) Archiving data
- B) Calculating running aggregates
- C) Updating DynamoDB tables
- D) Generating static reports

Terminology and Concepts


Amazon Kinesis Data Streams
Retention Period
On-demand
Or
Provisioned
Shard 1
Shard 2
Shard 3
Shard N
Producers
Consumers
(Custom Apps)
DynamoDB
S3
Sequence of **data records**
Each record has a **sequence number**

Amazon Kinesis Data Streams



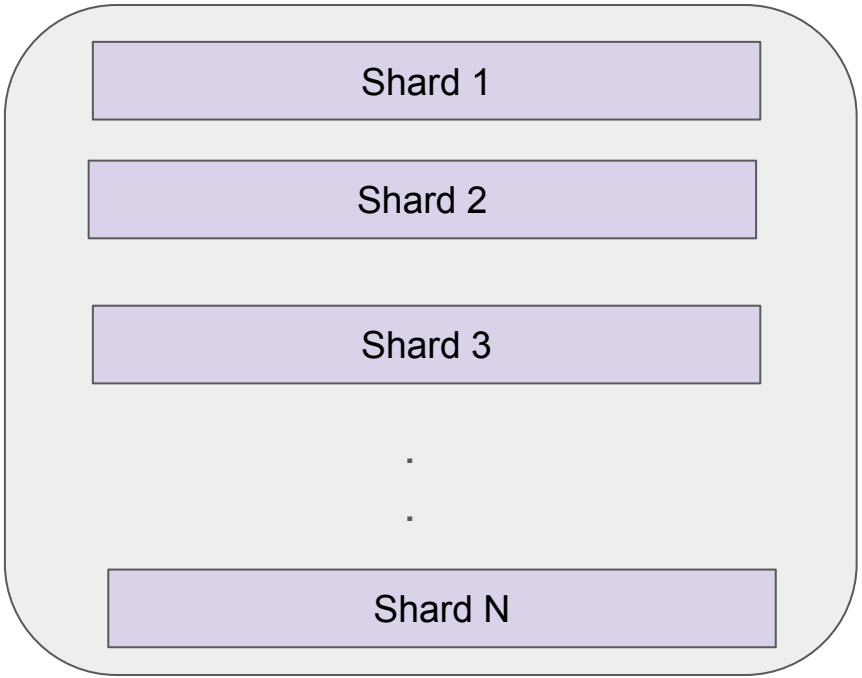
Sequence of **data records**

Each record has a **sequence number**

Each shard provide a fixed unit of capacity

- Shard read: Up to 5 transactions/second
- Maximum read rate: 2 MB/second
- Shard write: Up to 1,000 records/second
- Maximum write rate: 1 MB/second

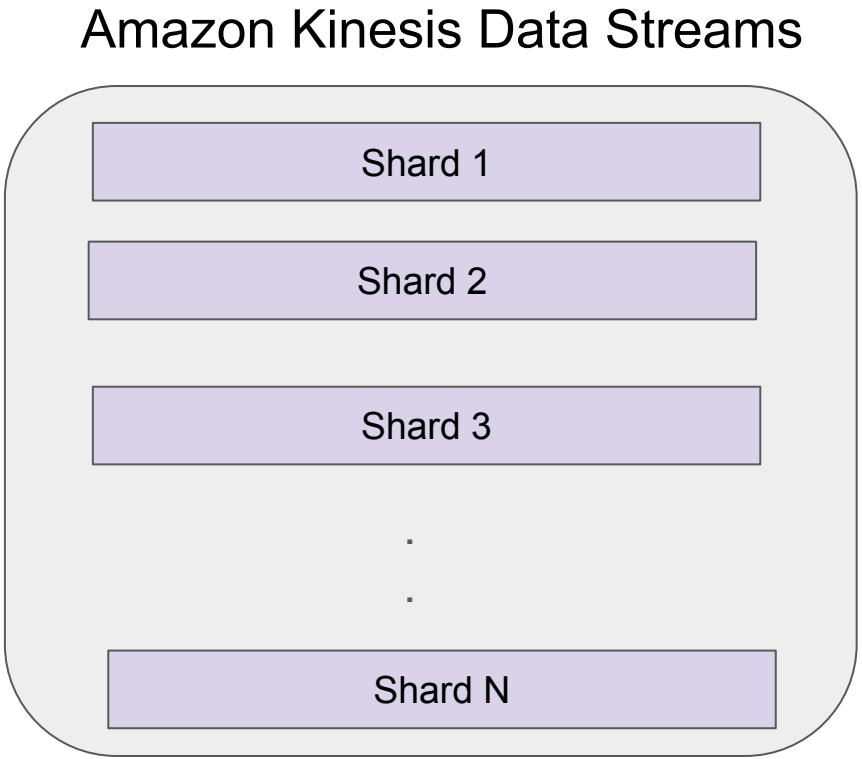
Amazon Kinesis Data Streams



Determines which shard

Partition key

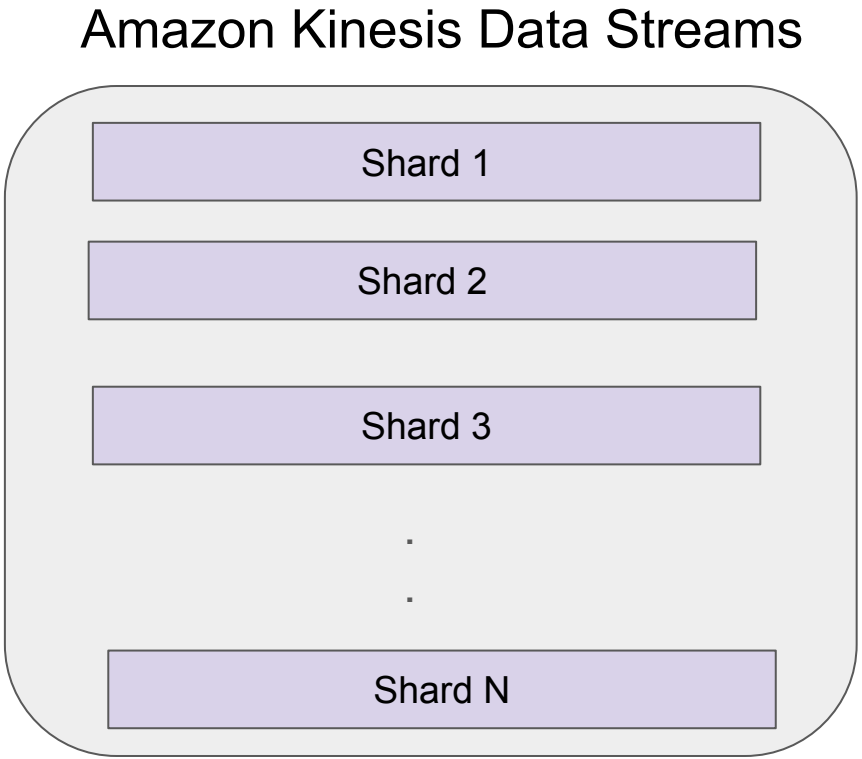
- **Shard Assignment**
- **Grouping by Partition Key**
- **Efficient Data Management**
- **Unicode strings**
- **Must specify partition key while storing data**



Sequence Number

Each data record has a unique **sequence number**

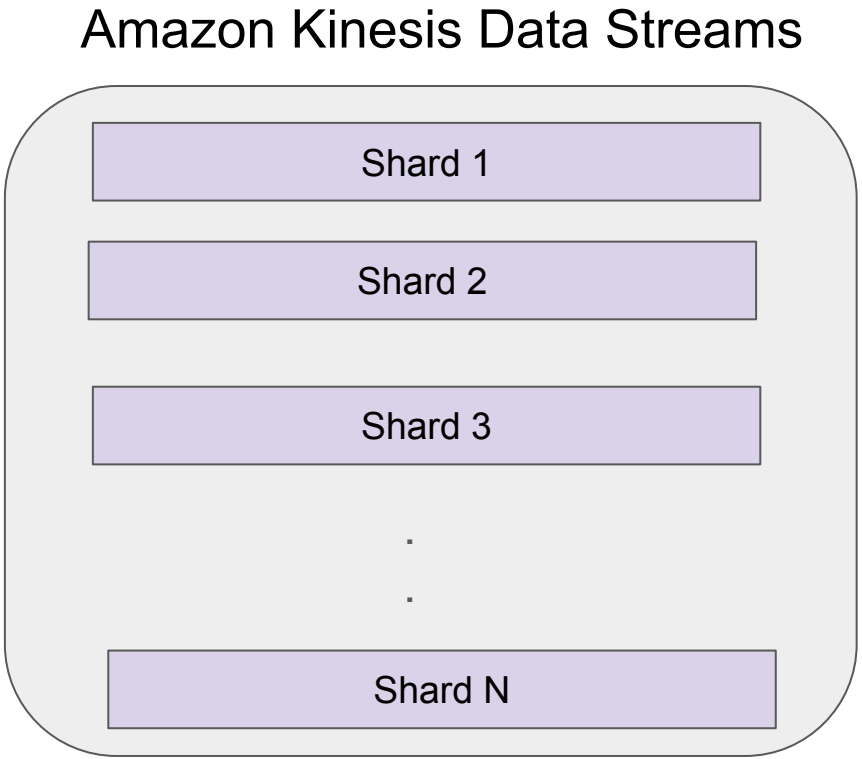
Kinesis Data Streams assigns the sequence number after writing to the stream using `client.putRecords` or `client.putRecord`.



Kinesis Client Library

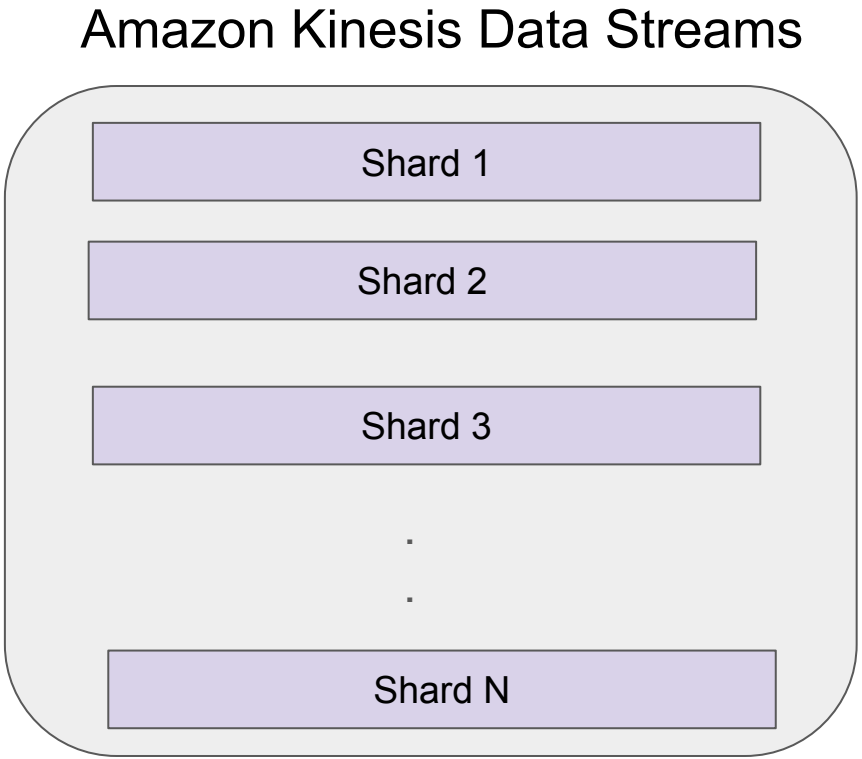
Ensures reliable consumption of data

Utilizes Amazon DynamoDB table



Application Name

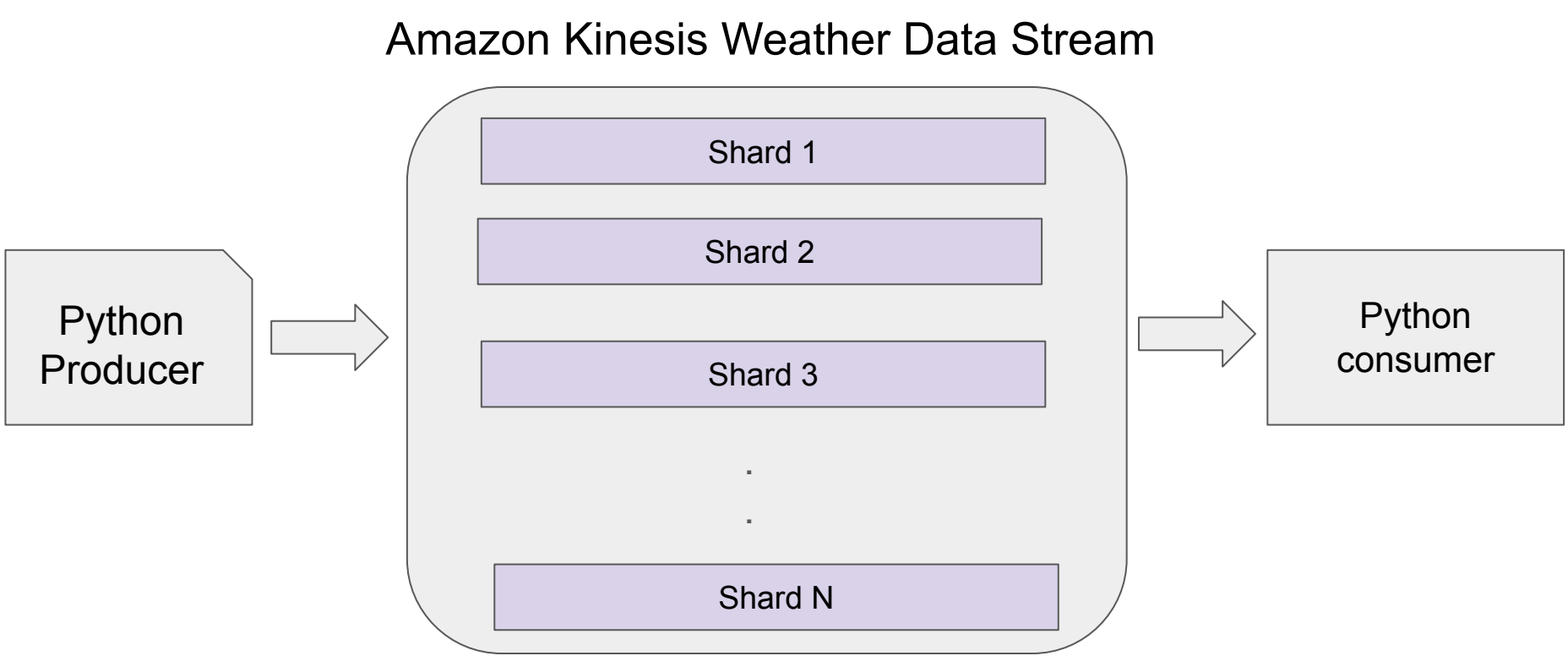
- Uniquely identifies an Amazon Kinesis Data Streams application.
- The name must be unique within the AWS account and Region.
- It serves as the identifier for the control table in Amazon DynamoDB.
- Also, it functions as the namespace for Amazon CloudWatch metrics.



Server-side Encryption

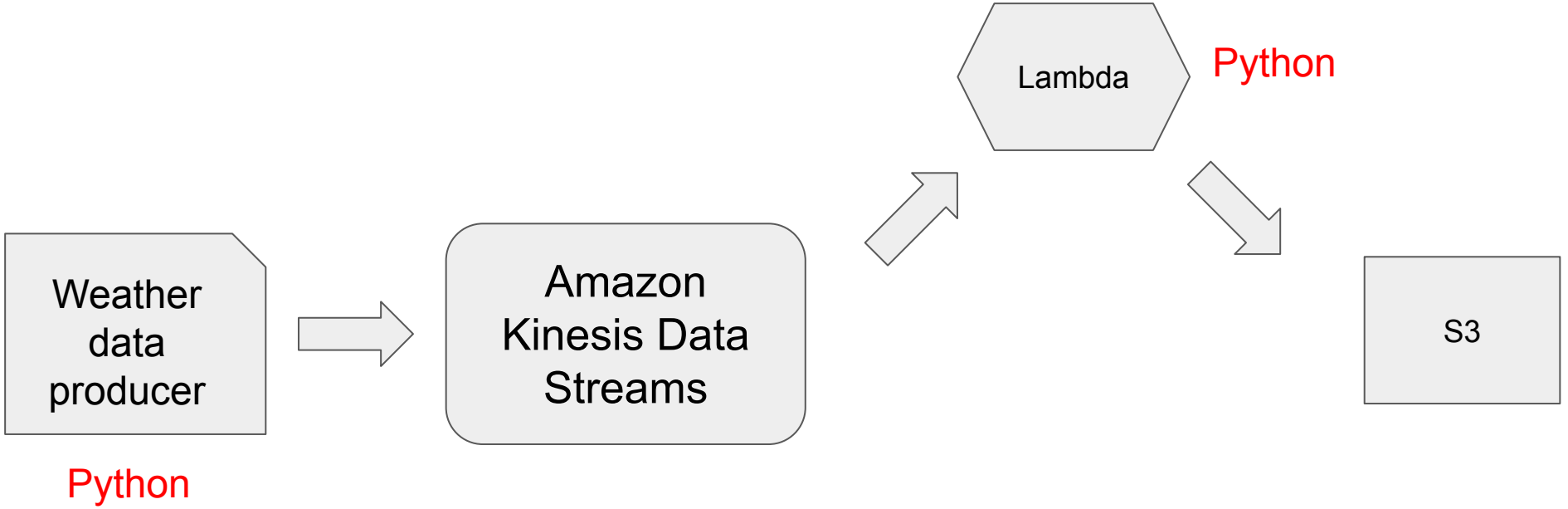
AWS KMS master keys for encryption

Basic Lab


Amazon Kinesis Weather Data Stream
Shard 1
Shard 2
Shard 3
Shard N
Python Producer
Python consumer

Kinesis Lambda

Kinesis Lambda


Lambda
Python
Weather data producer
Amazon Kinesis Data Streams
S3
Python

Questions

What is a Kinesis Data Stream?

- A. A set of shards
- B. A data record
- C. A Kinesis data stream
- D. A sequence number

What is the function of a partition key?

- A. Encrypts data
- B. Manages stream capacity
- C. Groups data by shard
- D. Stores control data

How is the capacity managed in on-demand mode?

- A. Automatically by Kinesis Data Streams
- B. Specified by the user
- C. Using DynamoDB
- D. Through CloudWatch metrics

What is a shard in Kinesis Data Streams?

- A. A set of data records
- B. A sequence number for a partition key
- C. A unit of capacity
- D. A Kinesis application

Which AWS services can consumers use to store results from Kinesis Data Streams?

- A. Amazon DynamoDB
- B. Amazon Redshift
- C. Amazon S3
- D. Amazon EC2

What are characteristics of a data record in Kinesis Data Streams?

a. A. Immutable sequence of bytes
b. B. Maximum size of 1 MB
c. C. Always inspected and modified by Kinesis
d. D. Contains a partition key

What is a role of the Kinesis Client Library?

A. Simplifies reading data from the stream

B. Manages shards in on-demand mode

C. Provides server-side encryption

D. Creates DynamoDB tables for consumers