27/02/2024

Kinesis

Amazon Kinesis Data Stream es un servicio para analizar grandes volúmenes de datos.

Un stream es un flujo de datos, es una secuencia de registros uno detrás del otro. Se procesan y luego se descartan. La arquitectura se compone:

* Productores: como dispositivos móviles, IoT
* Consumidores: aplicaciones de análisis de datos, data lakes, etc..

Dentro del flujo de Kinesis, los archivos se dividen en para distribuirlos en shards. Estos datos se quedan duplicados y replicados por AWS.

También se generan particiones para identificar los shards.

Los datos se almacenan durante un tiempo limitado, indicado por el usuario.

Se pueden utilizar servicios como lambda, o kinesis analytics para procesar los datos

Boto3

Librería de AWS para trabajar con AWS