Slide One – What is Amazon Kinesis part 1

What is Amazon Kinesis?

Amazon Kinesis is a real-time, fully managed, and scalable platform for streaming data on **Amazon Web Services.** It has multiple functionalities, allowing one to perform various tasks – such as ingesting and processing real-time data, and developing custom streaming applications for specific requirements.

Presentation Notes:

What is AWS Kinesis? AWS Kinesis – is core service for developing stream processing applications.

Why we need to ingest and process real time data? The area of BIG DATA is going to a massive change. The processing the vast of amount of data in real time is new norm and companies should processing incoming data in minutes or seconds just to stay competitive.

Slide Two – What is Amazon Kinesis part 2(Developing Stream Processing Applications with AWS Kinesis)

Architecture diagram with producers - > AWS Kinesis - > Consumers

Producers:
????
Consumers:
Dashboards
DataBase

Presentation Notes:

AWS Kinesis – is core service for developing stream processing applications. In a data stream processing application we have number of producer applications that are constantly writing data to data streaming service and we also have number of consumer that are constantly process that generated data in real time.

Here are some examples of producers applications that we can implement:

Kinesis can continuously capture gigabytes of data per second from hundreds of thousands of sources such as website clickstreams, database event streams, financial transactions, social media feeds, IT logs, and location-tracking events.

What consumers applications can we have:

We can have Dashboards that display information about data being generated

Database – we can store generated events to database for further quiring

Or we can have other application that react for coming data.

Slide Three – Kinesis Use-Cases

LOGS ICON(Log data collection and Processing)
MESSAGING ICON
REAL-TIME METRICS ICON
ACTIVITY TRACKER ICON

Presentation Notes:

Here are some common use cases for Amazon Kinesis:

Log Data collection and processing: we can collect and process application logs

Messaging: we can send messaging from one application to another

Real-time metrics: we can track metrics generating by applications

Activity tracking: we can analyze users activities in real time and react to it

All of these possible because of the following Kinesis features:

Slide Four – Kinesis Features

REAL TIME PERFORMANCE ICON
HIGH THROUGPUT ICON
ELASTIC ICON
AWS Integrations ICON
LOW COST ICON
EASY ADMINISTRATION ICON

Presentation Notes:

All of these use cases are possible because of the following Kinesis features:

Number one Kinesis can provide Real-time performance: which means there is low latency time between the moment when data written to AWS Kinesis stream and when the data can be read.

AWS Kinesis provide High Throughput: one of the example the companies can processing petabites of data every month

ELASTIC – which means it can scale up to process more data or it can scale down if demand is decreasing

AWS Integrations - also it's integrated with other AWS Services, so it's very convenient if you're already using AWS

LOW COST – AWS Kinesis has low cost, for example if you want to stream one megabyte of data every second, it will just cost you one cent per hour.

EASY ADMINISTRATION – and the last but not least Kinesis has very low administration version, it serverless, which means we don't need to take care about servers, upgrades, versions etc.

Slide Five – Batch processing vs Stream Processing

Producer - > S3 - > Consumers

Presentation Notes:

Why we do need new tool and why we can't use old tool to solve new problems.

So what was before stream processing? Before stream processing we had approach batch processing?

Batch Processing: We have number of producers to generate data. We aggregate this data and stored it in some distributed storage, like S3 for example and we have consumers that could read and process this data. And the biggest problem with this approach to have very high latency. It can take an hour or even a day from where the data can be generated and later consume.

More importantly recent data have a more values than old data. As example, if you want to analyze what are users are doing or fraud detection.

Old Architecture – History of kinesis

Unified log – new architecture: You have multiple number of producers but all of them are writing data to the same system, which are AWS Kinesis.

Slide Six - Streaming with Amazon Kinesis

Amazon Kinesis
Data Streams

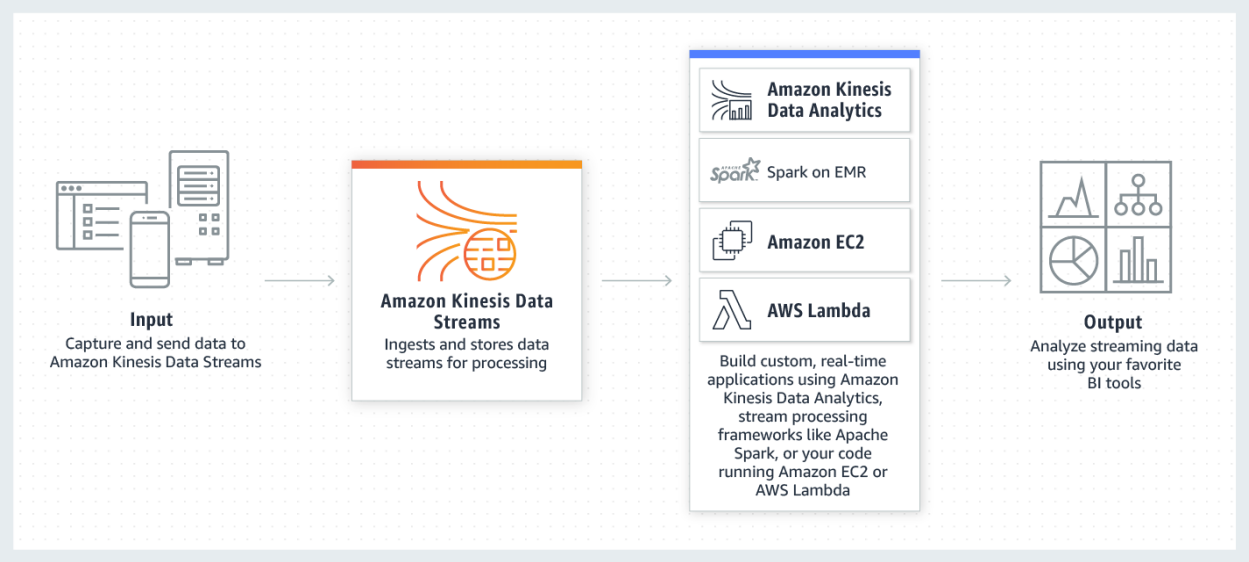
Amazon Kinesis
Data Firehose

Amazon Kinesis
Data Analytics

- **Kinesis Firehose** – to easily load streaming data into AWS;
- **Kinesis Analytics** – to easily process and analyze streaming data with standard SQL;
- **Kinesis Streams** – to build custom applications that process and analyze data.

Kinesis has multiple services under its name, like Data Streams, Firehose, Analytics, and Video Streams. We will only look at the Data Streams and Analytics service in this post.

Slide Seven - Amazon Kinesis Data Streams Overview

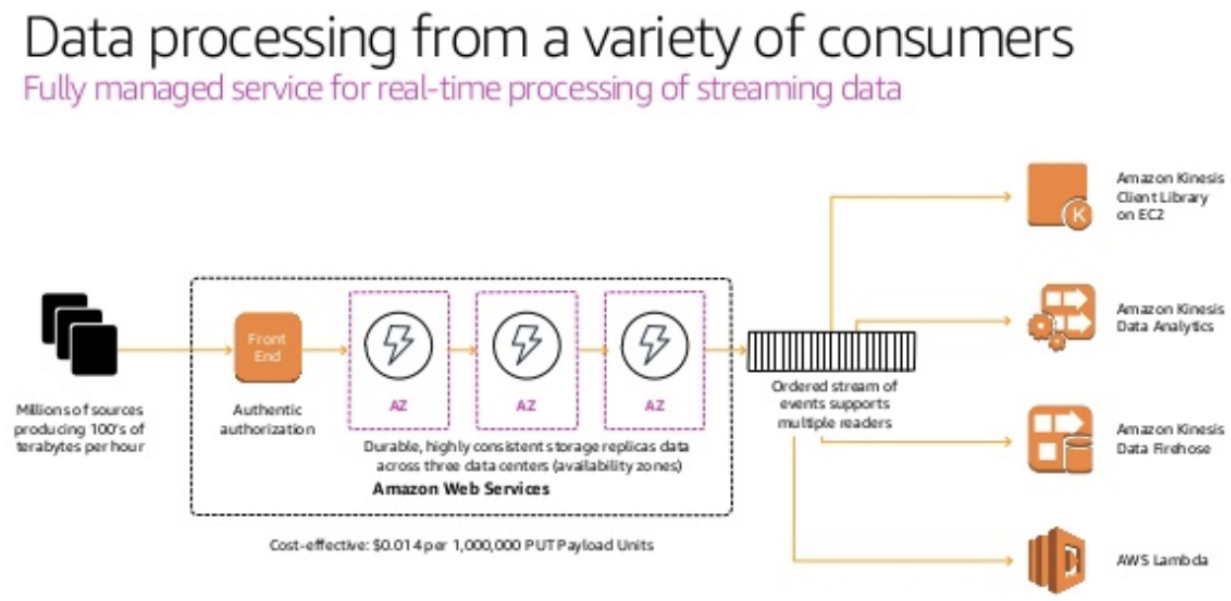


Demo notes:

Kinesis Data Streams (KDS) is a component of Kinesis. KDS is a service used to ingest and temporarily store data streams for processing by downstream consumers. Downstream consumers "get" data from streams (it is not pushed to applications by KDS).

and all other description is here: https://aws.amazon.com/kinesis/data-streams/faqs/

Maybe Slide:



Slide Four - Amazon Kinesis Firehose Overview

Slide Five - Amazon Kinesis Firehose Delivery Architecture

Slide Six- Amazon Kinesis Data Analytics Overview

**********DRAFT**********

Amazon Kinesis can collect and process hundreds of gigabytes of data per second from hundreds of thousands of sources, allowing you to easily write applications that process information in real-time, from sources such as web site click-streams, marketing and financial information, manufacturing instrumentation and social media, and operational logs and metering data.

Amazon Kinesis is a tool in the **Real-time Data Processing** category of a tech stack.

It provides key capabilities in the form of:

- **Kinesis Firehose** – to easily load streaming data into AWS;
- **Kinesis Analytics** – to easily process and analyze streaming data with standard SQL;
- **Kinesis Streams** – to build custom applications that process and analyze data.

Amazon Kinesis has attracted use from many well-known corporations. *Netflix*, for instance, monitors all communications between its applications using Kinesis, enabling it to detect and quickly solve any technical issues

The Amazon Kinesis suite of services is the central piece of the architecture that addresses the real-time, data ingestion needs integrating with managed services like Amazon EMR, Amazon DynamoDB, Amazon RDS Aurora, Amazon Redshift, AWS IoT and other AWS services (Amazon EC2, Amazon SNS, Amazon SQS) to offer optimal application experiences for analytics and machine learning. The Amazon Kinesis suite of services includes Amazon Kinesis Data Streams (Kinesis streams/KDS), Amazon Kinesis Data Firehose(Firehose), Amazon Kinesis Data Analytics (Kinesis data analytics) and Amazon Kinesis video streams.

Now let's review the technologies and patterns in the order of collection, processing, storage and analytics.

Presentation Notes:

Business today receive massive amount of data at a massive scale from various data sources that continuously generate streams of data. Enterprises are eyeing to extract precise insights on a real-time or near real-time from the unstructured data to deliver value and growth to their business.

Unstructured data can be from any available sources to the business, such as application logs, telemetry, social feeds, Internet Of Thing (IOT) etc. Business is enthusiastically looking forward to mine treasured information from the data that can help them more to learn about their customers, values so that they can delivery long term values to their customer.

Amazon Kinesis service is the central suite of the architecture to address the real-time , data ingestion needs.

Amazon Kinesis has attracted use from many well-known corporations. *Netflix*, for instance, monitors all communications between its applications using Kinesis, enabling it to detect and quickly solve any technical issues.