

# AWS Certified Solutions Architect: Associate - 11.0 Additional Key Services

File Name: amazon-acsa-11-3-analytics\_services

Title: Analytics Services

Subtitle: AWS Certified Solutions Architect: Associate

## 11.3 Analytics Services

- Amazon Kinesis
  - Facilitates working with data streams
  - Stream can be analyzed and output fed into apps
  - Three services
    - Amazon Kinesis Firehose (Diagram)
      - Allows loading massive amounts of data into AWS
    - Amazon Kinesis Streams (Diagram)
      - Allows building custom apps to perform analysis in real-time
      - Custom apps run on EC2 instances and are called *consumers*
    - Amazon Kinesis Analytics
      - Not yet released
      - Allows analyzing the data via standard SQL
  - Tuned to handle near-unlimited amounts of data
  - Example uses
    - Website traffic data
    - Monitoring device output
    - Acting on real-time data
  - Not good for
    - Extracting/Transforming data (use Pipeline instead)
- Amazon Elastic MapReduce (EMR)
  - Fully managed Hadoop framework
  - Can spin up a full Hadoop cluster in a matter of clicks
  - Two file systems
    - Hadoop Distributed File System (HDFS)
      - Standard Hadoop default
      - Data is stored in EC2 or EBS storage
    - EMR File System (EMRFS)
      - Allows cluster to store data in S3
  - Possible uses
    - Log processing
    - Clickstream Analysis
    - Genomics and Life Sciences
- AWS Data Pipeline
  - Allows reliably processing data and moving it between services
  - Runs on a set schedule
  - Uses predefined *activities* to manipulate and move the data (Diagram)
  - Useful for Extract, Transform and Load (ETL) processes
- AWS Import/Export
  - Accelerates moving large amounts of data to/from AWS
  - Data is stored on physical devices and then shipped to AWS
  - Two primary features
    - AWS Snowball
      - Uses Amazon-provided shippable storage devices
        - Ruggedized so it is its own container
        - [Image](#)
        - Shipping label is e-ink
      - Mailed through UPS
      - Protected via AWS KMS (Encryption is required)
      - Comes in 50TB and 80TB models
      - Data can be imported into S3
    - AWS Import/Export Disk
      - Import/Export data directly from your own devices
      - Devices are shipped to Amazon and returned when completed
      - Can import to S3, Glacier or EBS
      - Limited to 16TB of data