Data Pipeline

Data Pipeline can be broadly categorized into two

1. Batch data processing
2. Real Time data processing

- Batch processing of data can be delivered on Hourly, Daily, Weekly or Monthly

- Similarly Streaming data can be delivered as soon as the data comes in from IoT devices or other real time processing datasets.

DBMS

CSV Files

BODS/Informatica

Data Lake/ i.e. AWS S3

BI Reporting

IoT device

Message HUB/ ODS

Logs

Kafka/ Kinesis

Data Warehouse/ i.e. Teradata

Machine Learning

MDM

Streaming