

# Implementation and Support for Cassandra platform for a leading American multinational Bank

## Challenges

- Build a Operation Data store for Digital consumption.
- Build a scalable architecture that would allow real-time processing, data lineage and faster data Ingestion.
- Data will be consumed by various systems like Statement generator, recommendation engine, regulatory reporting.
- Build a Data ingestion Pipeline to Ingest the Data from various LOBs like DDA, Card and Investment, etc.

## Solution

- Built Hortonworks Hadoop cluster[6 nodes | 10 TB]
- Built DSE Cassandra [20 nodes | 270 TB]. Multi datacenter setup.
- Kafka with Spark real time data ingestion into Cassandra.
- Cassandra Multidata center replication setup to allow work load isolation.
- Cassandra Solr to allow free text search option for Support team.
- Developed Enrichment engine (A Wrapper) around apache spark and Data Frame API.
- Setup a support function using our ASM (Application support and Maintenance) framework.

## Value delivered

- Provides consistent and complete view to consumers.
- Scalable & flexible architecture which can cater to future growth requirements of bank
- Operational efficiencies providing end to end automation
- Reduce HW, SW and Operation cost.