

## Case STUDY – Hadoop



### Challenge

A global IT services company was extending into adjacent markets and new geographies and needed the back-office to keep pace with the expansion.

### Solution Profile

Created and implemented best practices for capacity planning, performance tuning and benchmarking to ensure cluster balancing.

### Services Provided:

- Implemented capacity schedulers on the Job tracker to share the resources of the Cluster for the Map Reduce jobs given by the users.
- Ensured Cluster balancing and performance tuning of Hadoop components like HDFS, Hive, Map Reduce, Oozie work flows.
- Commissioned and Decommissioning Hadoop Cluster nodes Including Balancing HDFS block data.
- Configured Kerberos and AD/LDAP for Hadoop cluster.
- Installed and setup Kafka cluster and monitored the cluster using Nagios and Ganglia.
- Used Hive to analyze the partitioned and bucketed data and compute various metrics for reporting.
- Set up automated processes for analyzing System and Hadoop log files for predefined errors and send alerts to appropriate groups.
- Used Flume and Sqoop for handling data exchange between HDFS & Web Applications and databases
- Performance Tuning, Client/Server Connectivity and Database Consistency Checks