

**GOVERNANCE
INTEGRATION**

**Data Workflow,
Lifecycle &
Governance**

Falcon

WebHDFS

NFS

Flume

Sqoop

Kafka

DATA ACCESS

Script

Pig

Tez

SQL

Hive
HCatalog

Tez

Java/S...

Cascading

Tez

NoSQL

HBase
Accumulo
Phoenix

Slider

Stream

Storm

Slider

Search

Solr

In-Mem

Spark

Tez

Others...

Engines

S T

YARN: Data Operating System

HDFS

Hadoop Distributed File System

DATA MANAGEMENT

SECURITY

**Authentication,
Authorization, Audit &
Data Protection**

Storage: HDFS

Resources: YARN

Access: Hive

Pipeline: Falcon

Cluster: Knox

Cluster: Ranger

OPERATIONS

**Provision, Manage &
Monitor**

Ambari
ZooKeeper

Scheduling

Oozie