**Cluster Resource Management**

**NoSQL**

**File Systems**

**Graph**

**High Level Abstractions for Data Processing**

**Data Processing**

**Giraph**

**Tez
(DAG)**

**MRQL
(SQL on Hadoopx, Hama, Spark)**

**HDFS**

**Cassandra
(DHT)**

**Hive
(SQL on Hadoop)**

**Mahout / MLlib / MLbase**

**Mesos / Yarn**

**Hadoop
(MapReduce)**

**Spark
(Iterative MR)**

**Accumulo
(Data on HDFS)**

**HBase
(Data on HDFS)**

**Storm**

**S4**

**Samza**

**Hama
(BSP**)

**Pig
(Procedural Language)**

**Shark
(SQL on Spark, NA)**

**Phoenix
(SQL on HBase)**

**GORA
(Object Mapping for NoSQL)**

**Data Analytics**

 **Thrift/Protobuf(NA)**

**ZooKeeper/JGroups/Hazelcast**

**Netty(NA)/ZeroMQ(NA)/ActiveMQ/QPid/Kafka**

**Distributed Coordination**

**Message** **Protocols**

**Communication**

**Cloud Deployment**

**Whir /JCloud**

**Tools**

**NA - Non Apache projects**

**Batch**

**Stream**

**Oozie**

**Orchestration/Workflow**

