
APACHE HADOOP.

Apache Hadoop – (software framework)
Google, MapReduce, GFS BigTable,
Hadoop Hadoop Distributed File System
(HDFS), Hadoop. HDFS (
MapReduce,
Namenode TCP/IP.
ClientProtocol, Remote
Procedure Call (RPC). DatanodeProtocol,

Hadoop MapReduce

MapReduce –

Google

MapReduce

– map reduce,

MapReduce

list(k1,v1),

“ – ”.
– list(k2,v2).

- map reduce.

map(k1,v1) list(k2,v2)

MapReduce

v2,
reduce.

reduce(k2,list(v2)) list(v2)

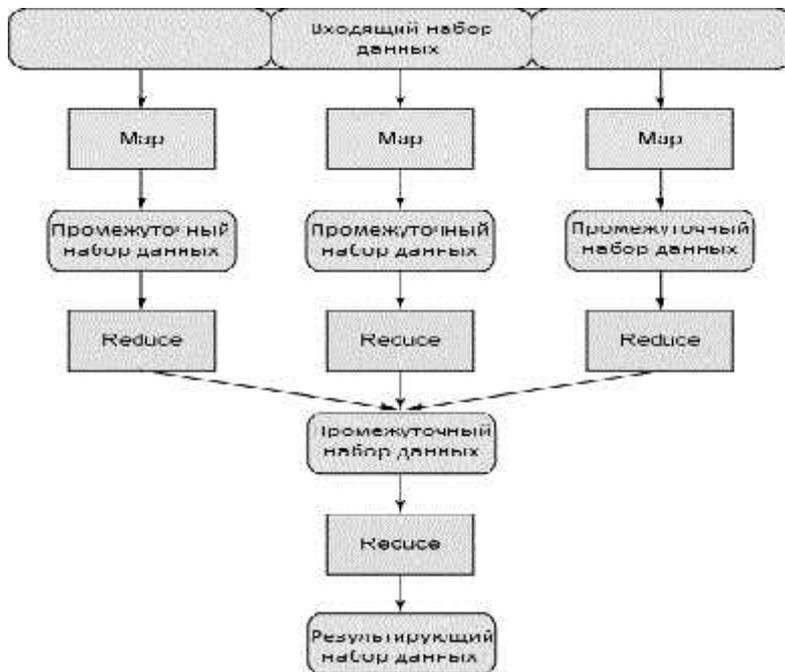
MapReduce

2.

Hadoop MapReduce

Java.

Java.



2.

MapReduce

MapReduce

JobTracker.

NameNode,

Hadoop,

