

JobTracker

The [JobTracker](#) is the service within Hadoop that farms out [MapReduce](#) tasks to specific nodes in the cluster, ideally the nodes that have the data, or at least are in the same rack.

1. Client applications submit jobs to the Job tracker.
2. The [JobTracker](#) talks to the [NameNode](#) to determine the location of the data
3. The [JobTracker](#) locates [TaskTracker](#) nodes with available slots at or near the data
4. The [JobTracker](#) submits the work to the chosen [TaskTracker](#) nodes.
5. The [TaskTracker](#) nodes are monitored. If they do not submit heartbeat signals often enough, they are deemed to have failed and the work is scheduled on a different [TaskTracker](#).
6. A [TaskTracker](#) will notify the [JobTracker](#) when a task fails. The [JobTracker](#) decides what to do then: it may resubmit the job elsewhere, it may mark that specific record as something to avoid, and it may even blacklist the [TaskTracker](#) as unreliable.
7. When the work is completed, the [JobTracker](#) updates its status.
8. Client applications can poll the [JobTracker](#) for information.

The [JobTracker](#) is a point of failure for the Hadoop [MapReduce](#) service. If it goes down, all running jobs are halted.