

Task Nodes

"Task Node" refers to a node that contributes only compute resources for TaskTrackers, and does not contribute any disk space to the cluster's storage pools. Generally, when permanently adding a node to a cluster, you want the node to contribute both compute and storage resources. However, there are cases for which it is preferable to prevent the cluster from storing data on a particular node. For example, Task Nodes are useful if you need the ability to add compute resources to the cluster at will, and later remove them spontaneously without provisioning for data on the nodes to safely replicate elsewhere.

Task Node Services and Topology

Task Nodes run the following services: TaskTracker and Fileserver. The Fileserver service is required, because TaskTrackers require local storage for intermediate data.

[Node Topology](#) settings prevent the cluster from storing unrelated data on the Task Node. You must assign a Task Node to the `/compute-only` topology, which has no storage volumes assigned to it. (The topology name is unimportant, so long as it has no storage assigned to it.) By contrast, nodes for data storage are generally assigned to the `/data` topology (or a sub-topology of `/data`).

Adding a Task Node to a cluster

To add a Task Node to a cluster, follow the steps outlined in [Adding Nodes to a Cluster](#), with the following modifications:

- Packages to install: `mapr-tasktracker`, `mapr-fileserver`
- *Before* you start the Warden service (which starts the Fileserver service), add the following line to `/opt/mapr/conf/mfs.conf`:

```
mfs.network.location=/compute-only
```

 After the Fileserver service is running, changing `mfs.conf` has no effect on the node's topology.

Converting an Existing Node to a Task Node

If the Fileserver is running on a node assigned to a topology with volume data assigned to it, you will need to use the `maprcli node move` command to move the node to the `/compute-only` topology (or some other topology with no volumes assigned to it). For example:

1. Find the `serverid` for the node, which you will use in the next step. (See [How to find a node's serverid](#).)
2. Issue the following command to re-assign the node's topology:

```
maprcli node move -serverids <serverid> -topology /compute-only
```

It will then take time for any data stored on the node to transition elsewhere in the cluster.