

Processing Engine and Resource Pools

Master Segments generate a workflow that unifies a Master Table, Attribute Tables, and Behavior Tables.

Because Workflow tends to be a computationally heavy process, Treasure Data provides a way to avoid the compute resource conflicts with other jobs and processes you're currently running.

In Master Segments, you can specify the type of Processing Engine to use and associated resource pools. For example:

Resource Pool

Presto and Hive Hive only

Presto Pool Name

Presto resource pool name

Hive Pool Name

Hive resource pool name

You can choose:

- Presto and Hive
- Hive only

Your selection determines which processing engine is used for the master segment creation.

We generally recommend Hive only, because it's designed for large-scale joins and known for robustness, while Presto and Hive might perform faster for smaller data sets.



When you select Presto and Hive, most of the jobs are issued with Hive, however, some Presto queries for the drop/create table operation are generated.

(Note: Hive only issues most of the jobs with Hive, however it generates some Presto queries for drop/ create table operation)

You can also specify which resource pools to use for each engine. Specifying resource pools gives you additional compute resource control.

- [Presto Resource Pools](#)
- [Hive Resource Pools](#)