

# RFC-001: Supporting system databases and tables

## Authors

jeremydeng012@gmail.com

## Status

Accepted

## Problem

There is no place to store status data like jobs info in Task Manager other than ZooKeeper currently. But ZooKeeper is not suitable to store these data, which is a centralized service for maintaining configuration information, naming, providing distributed synchronization, and providing group services.

## Solution

We plan to store the jobs info in a system table called `JOB_INFO` which is auto created when OpenMLDB startup. Moreover, `tablet` need to be started before `nameserver`. All system tables will be stored in `__INTERNAL_DB` database which is not seen in `SHOW DATABASES` command.

Name	Value
DATABASE	__INTERNAL_DB
TABLE NAME	JOB_INFO
PARTITION_NUM	1
REPLICA NUM	2 if the number of tablet > 1
TTL	latest ttl 1

The schema of JOB\_INFO is shown as below:

Field Name	Type	Examples
id	bigint	1122, 1123
job	string	SparkBatchSql, ImportOfflineData, ImportOnlineData
state	string	Submitted, Running, Failed, Success, Pending
start_time	timestamp	145954243
end_time	timestamp	145954243
cluster	string	Local, Yarn
parameter	string	"{fileType: parquet, filePath: hdfs:///tmp/, db: db1, outputTable, t2}"...
application_id	string	application_1459542433815_0002
error	string	

All the relevant records will be selected from JOB\_INFO when executing SHOW JOB/SHOW JOBS command.

## Issues

- [Create system table when OpenMLDB startup #697](#)

## Changes and Additions to Public Interfaces

N.A

## Performance Impact

N.A

## Backwards Compatibility and Upgrade Path

N.A