Dag\_Batch

该表用于启动、完成、跟踪dag进程状态。

每次执行都会产生一个新的Dag\_Batch\_ID插入表中，记录批次开始时间，结束时间以及状态等。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dag\_Batch\_id | Serial | Not Null | Primary\_Key | Serial ID |
| Data\_Source\_ID | Integer | Not Null |  | Same as Data\_Source\_id |
| Dag\_Load\_Type | Varchar(30) |  |  | Historical Load or Increment Load |
| Batch\_Start\_Date | Timestamp |  |  |  |
| Batch\_End\_Date | Timestamp |  |  |  |
| Dag\_Batch\_Status | Char(1) |  |  |  |
| Insert\_Time | Timestamp |  |  | Record creation time |
| Update\_Time | Timestamp |  |  | Record modified time |

Data\_Source

用于存储数据源应用程序的相关信息。需要手工维护此表。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Data\_Source\_Name | Varchar(50) | Not Null | Primary\_Key | Defines data source batch |
| Data\_Source\_ID | Serial | Not Null |  | Serial ID |
| Data\_Source\_Status | Char(1) | Not Null |  | Active or Inactive(Y/N) |
| Data\_Source\_Description | Varchar(50) |  |  |  |
| Job\_Limit | Integer | Not Null |  | Number of Parallel Job |
| Insert\_Date | Timestamp |  |  | Creation Date |
| Last\_Update\_Date | Timestamp |  |  | Modified date |
| Version | Integer | Not Null |  |  |

Job

该表存储每一个Job的主要信息

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Job\_Name | Varchar(50) | Not Null | Primary\_Key | Name of the job |
| Job\_ID | Serial | Not Null |  | Serial ID |
| Schedule\_Type | Varchar(10) | Not Null |  | Type of schedule |
| Engine\_Type | Varchar(50) | Not Null |  | Where the job should be executed |
| Data\_Quality\_Check | Char(1) | Not Null |  | Y/N |
| Job\_Type | Varchar(10) | Not Null |  | Historical Load or Increment Load |
| Data\_Source\_Name | Varchar(50) | Not Null | Primary\_Key | Same as Data\_Source table |
| Is\_Deleted | Char(1) | Not Null |  | Soft delete (Y/N) |
| Insert\_Date | Timestamp |  |  | Creation date |
| Last\_Update\_Date | Timestamp |  |  | Modified date |
| Version | Integer | Not Null |  |  |

INDV\_Job\_Params

该表存储Job相关的参数信息, 可以用于调用Job。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Job\_ID | Integer | Not Null |  | Same as Job table |
| Param\_Name | Varchar(50) | Not Null |  |  |
| Param\_Value | Varchar(10000) | Not Null |  |  |
| Is\_Deleted | Char(1) | Not Null |  | Soft delete (Y/N) |
| Insert\_Date |  |  |  | Creation date |
| Last\_Update\_Date |  |  |  | Creation date |

Email\_Subscription—目前CCDH 生产该表为空

用于存储用户订阅信息。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Data\_Source\_Name |  |  |  |  |
| Success\_Topic\_Subscription |  |  |  |  |
| Failure\_Topic\_Subscription |  |  |  |  |
| Insert\_Date |  |  |  | Creation date |
| Last\_Update\_Date |  |  |  | Modified date |

Job\_Status

用于跟踪每个Job的状态。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Batch\_ID | Integer | Not Null | Primary\_Key | Batch id from Dag\_Batch table |
| Job\_ID | Integer | Not Null | Primary\_Key | Job id from Job table |
| Run\_ID | Bigint |  |  | Run information from execution end point |
| Engine\_Type | Varchar(50) |  |  | From Job table |
| Job\_Start\_Time | Timestamp |  |  | Job start time at end point |
| Job\_End\_Time | Timestamp |  |  | Job end time at end point |
| Job\_Status | Varchar(25) |  |  |  |
| Insert\_Date | Timestamp |  |  | Creation date |
| Last\_Update\_Date | Timestamp |  |  | Modified date |

Dag

用于存储Dag信息，该表需要手动维护。方便运维人员监控Airflow dag状态。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dag\_ID | Integer |  |  | Serial ID |
| Stream | Varchar(25) |  |  |  |
| DagID\_by\_Stream | Varchar(25) |  |  | Serial ID by stream |
| Subject | Varchar(25) |  |  |  |
| Dag\_Name | Varchar(50) |  |  |  |
| Data\_Source\_Name | Varchar(50) |  |  |  |