

Agent Plugin for Manifest Files

Contents

- [Name](#)
- [Synopsis](#)
- [Description](#)
- [Agent Modes](#)
- [Options](#)
- [Example Actions](#)
- [Example Manifest File](#)

Name

hvrmanifestagent.py

Synopsis

hvrmanifestagent.py *mode chn loc* [*userargs*]

Description

The [agent plugin](#) **hvrmanifestagent** writes manifest file for every integrate cycle. A manifest file contains the summary of files or tables that have been changed during an integrate cycle so this information can be used for further downstream processing. This agent plugin should be defined in the HVR channel using action [AgentPlugin](#). The behaviour of this agent plugin depends on the agent mode and options supplied in parameter **/UserArgument**. The possible values for **/UserArgument** field in [AgentPlugin](#) screen are described in section [Options](#).

Agent Modes

Hvrmanifestagent supports only **integ_end** and **refr_write_end** [mode](#). This agent plugin should be executed using action [AgentPlugin](#) during either [Integrate](#) or [Refresh](#).

Parameter	Description
integ_end	Write manifest file implied by option -m <i>mani_fexpr</i> . Existing manifest files are not deleted by this. Value in manifest file for initial_load is false
refr_write_end	Write manifest file implied by option -m <i>mani_fexpr</i> . Existing manifest files are not deleted by this. Value in manifest file for initial_load is true

Options

This section describes the parameters that can be used with **Hvrmanifestagent**:

Parameter	Description
-i <i>integ_fexpr</i>	Integrate file rename expression. This is optional if there is only one table in an integrate cycle. If multiple tables are in a cycle this option is mandatory. It is used to correlate integrated files with corresponding table manifest files. Must be same as Integrate/RenameExpression parameter. Sub-directories are allowed. Example: {hvr_tbl_name}_{hvr_integ_tstamp}.csv

-m <i>mani_fexpr</i>	<p>Manifest file rename expression. This option is mandatory. Sub-directories are allowed. Example: manifest-{hvr_tbl_name}-{hvr_integ_tstamp}.json. It is recommended that table name is followed by a character that is not present in table name, such as:</p> <p>-m {hvr_tbl_name}-{hvr_integ_tstamp}.json or -m {hvr_tbl_name}/{hvr_integ_tstamp}.json or -m manifests/{hvr_tbl_name}/{hvr_integ_tstamp}.json</p>
-s <i>statedir</i>	<p>Use <i>statedir</i> for state files and manifest files instead of \$HVR_LOC_STATEDIR. This option is mandatory when \$HVR_LOC_STATEDIR points to a non-native file system (e.g. S3).</p>
-v <i>val</i>	<p>Set JSON path <i>a.b.c</i> to string value <i>val</i>/inside new manifest files. This option can be specified multiple times. Example: -v cap_locs.cen.dbname=mydb</p>

Example Actions

Group	Table	Action
SRC	*	Capture
TGT	*	Integrate /RenameExpression="{hvr_tbl_name}/{hvr_integ_tstamp}.xml"
TGT	*	AgentPlugin /Command="hvrmanifestagent.py" /UserArgument="-m {hvr_integ_tstamp}-{hvr_tbl_name}.m -s /hvr/hvr_config/files/work/manifests -i {hvr_tbl_name}/{hvr_integ_tstamp}.xml "

Example Manifest File

When using the above example, the manifest files are located in **/hvr/hvr_config/files/work/manifests** and are formatted as *{hvr_integ_tstamp}-{hvr_tbl_name}.m*. E.g. when source tables **aggr_order** and **aggr_product** are integrated in a cycle that ended at August 31st, 10:47:32, the manifest file names are **20170831104732-aggr_order.m** and **20170831104732-aggr_product.m**.

Example manifest file for table **aggr_product**

```
{
  "cap_rewind": "2017-08-31T08:36:12Z",
  "channel": "db2file",
  "cycle_begin": "2017-08-31T08:47:31Z",
  "cycle_end": "2017-08-31T08:47:32Z",
  "initial_load": false,
  "integ_files": [
    "aggr_product/20170831084731367.xml",
    "aggr_product/20170831084731369.xml",
    "aggr_product/20170831084731370.xml",
    "aggr_product/20170831084731372.xml",
    "aggr_product/20170831084731374.xml",
    "aggr_product/20170831084731376.xml"
  ],
  "integ_files_properties": {
    "aggr_product/20170831084731367.xml": {
      "hvr_tx_seq_min": "0000403227260001",
      "num_rows": 4
    }
  },
}
```

```
"aggr_product/20170831084731369.xml": {
  "hvr_tx_seq_min": "0000403227480001",
  "num_rows": 72
},

"aggr_product/20170831084731370.xml": {
  "hvr_tx_seq_min": "00004032280E0001",
  "num_rows": 60
},

"aggr_product/20170831084731372.xml": {
  "hvr_tx_seq_min": "0000403228B70001",
  "num_rows": 60
},

"aggr_product/20170831084731374.xml": {
  "hvr_tx_seq_min": "0000403229570001",
  "num_rows": 56
},

"aggr_product/20170831084731376.xml": {
  "hvr_tx_seq_min": "0000403229F50001",
  "num_rows": 56
},

"integ_loc": {
  "dir": "s3s://rs-bulk-load/",
  "name": "s3",
  "state_dir": "s3s://rs-bulk-load//_hvr_state"
},

"next": null,
"prev": "20170831104732-aggr_order.m",
"tables": {
  "aggr_product": {
    "basename": "aggr_product",
    "cap_tstamp": "2017-08-31T08:45:31Z",
    "num_rows": 308
  }
}
}
```