Automated Backup Process - User Manual V3.0

|  |  |  |
| --- | --- | --- |
| Version Number | Date | Details |
| 1.0 | 05th September 2023 | Initial Version |
| 2.0 | 28th September 2023 | Enhancements after feedback  -Creation of top-level entities automatically to avoid manual creation.  -Dealing case sensitive object names. SP is now handling case sensitive objects by applying double quotes.  -Enhanced logging information.  -Minor changes to audit table. Introduced default values for source schema and table. |
| 3.0 | 12th October 2023 | Added new column “exclusion\_tables” which is useful to add tables which need to be excluded from backup process. This is applicable for schema/database level backups and while taking PHYSICAL backup, But not CLONE  Also added another column “target\_backup\_type” which is defaulted to TRANSIENT. Now user can choose to create PERMANENT tables as well.  SP is limited to TABLE\_TYPEs BASE TABLE. Apart from Standard base tables rest of the objects are not in scope. |

Table Name: BACKUP\_TBL

|  |  |  |  |
| --- | --- | --- | --- |
| Column Name | Data Type | Mandatory | Details |
| BACKUP\_ID | INT | YES | Primary Key, Auto Increment value (IDENTITY) |
| SOURCE\_DB\_NM | VARCHAR | YES | Source database name which will be used in the backup |
| SOURCE\_SCHEMA\_NM | VARCHAR |  | Name of the Schema to be backed up. N.B. for schema level back up you will need to also specify the source database and the target in column TARGET\_DB\_NM as a pre-requisite |
| SOURCE\_TABLE\_NM | VARCHAR |  | If backup for specific table, then as above but include database and schema for both source and target. SOURCE\_SCHEMA\_NM should be created as pre-requisite |
| TARGET\_DB\_NM | VARCHAR | YES | Target database name where backup need to be created, Validation: TARGET\_DB\_NM should not be same as SOURCE\_DB\_NM |
| TARGET\_BACKUP\_TYPE | VARCHAR | YES | How target entities need to be created. Default is TRANSIENT. Any other value , SP creates as permanent type. |
| BACKUP\_MODE | VARCHAR | YES | Type of backup required. Permissible values CLONE/PHYSICAL |
| ENABLED\_IND | INT | YES | Active/In-active flag. Permissible values 1/0 |
| EXCLUSION\_TABLES | TEXT |  | Applicable for table level exclusions while doing PHYSICAL based backup. User can add comma delimited values which will be ignored while creating target tables. |
| WEEKLY\_DAY\_OF\_WEEK | INT |  | Day of the week, if backup is set for weekly. Target backup will be created with YYYYMMDD. Daily backup is default.  Note: If either (weekly/monthly) of the values are not set, default backup will be on daily. |
| MONTHLY\_DAY\_OF\_MONTH | INT |  | Day of the month, if backup is set for monthly. Target backup will be created with YYYYMMDD. Daily backup is default.  Note: If either (weekly/monthly) of the values are not set, default backup will be on daily. |
| RETAIN\_DAILY\_BACKUPS | INT |  | Number of daily backups to retain. Oldest will be dropped |
| RETAIN\_WEEKLY\_BACKUPS | INT |  | Number of weekly backups to retain. Oldest will be dropped |
| RETAIN\_MONTHLY\_BACKUPS | INT |  | Number of monthly backups to retain. Oldest will be dropped |
| EMAIL\_ACCOUNTS | VARCHAR | YES | Email accounts separated by comma. These email ID's need to be verified from snowflake. |

**Stored Procedure Name**: AUTO\_BACKUP\_PROC ()

**General**: Based on the values presented in SOURCE\_DB\_NM/SOURCE\_SCHEMA\_NM/ SOURCE\_TABLE\_NM, the respective Backup level will be derived.

If value is found in SOURCE\_DB\_NM and nothing under rest, then it will be considered as database level backup, if values found at SOURCE\_DB\_NM and SOURCE\_SCHEMA\_NM, then it is at schema level. If all three, then it is at table level.

**N.B. Values need to be added carefully, by avoiding spelling mistakes.**

**TARGET\_DB\_NM** is a key column. Backups will be created with this DB name. When a backup is set for a table, relevant target Database/Schema will be created automatically. **Please make sure TARGET\_DB\_NM is different than SOURCE\_DB\_NM. Although such scenario is taken care in SP.**

Target backups (Database/Schema/Table) will be concatenated with YYYYMMDD (e.g. 20230904 if run date is 4th September 2023).

By default, **ENABLED\_IND** = 1, means the Backup is active. If you want to deactivate the backup process, then make this flag as 0 (ZERO).

**WEEKLY\_DAY\_OF\_WEEK**: If day of the week matches with this, then backup will be considered as a weekly backup.

**MONTHLY\_DAY\_OF\_MONTH**: If day of the month matches with this, then backup will be considered as a monthly backup.

If either of the above is true, then it will set for daily backup.

**Clone**: A clone backup will be created when BACKUP\_MODE is mentioned as “CLONE”. It’s a ZERO copy clone which is a metadata operation.

**Physical Backup**: It will create a new physical copy of data. If backup level is at Database, a new database will be created as backup, same for schema and a table.

**Archival/Auto removal process:**

**RETAIN\_DAILY\_BACKUPS/ RETAIN\_WEEKLY\_BACKUPS/ RETAIN\_MONTHLY\_BACKUPS** are settings for how many recent backups required. Based on the settings in these columns, only recent number of backups will be retained, and older ones will be removed.

Ex: **RETAIN\_DAILY\_BACKUPS** – 5, then latest 5 daily backups will be retained and oldest will be removed.

**WEEKLY\_DAY\_OF\_WEEK** / **MONTHLY\_DAY\_OF\_MONTH** works together with **RETAIN\_WEEKLY\_BACKUPS/ RETAIN\_MONTHLY\_BACKUPS.** It means thatretain parameter get into action only when backup frequency is set accordingly.

NOTE: If RETAIN\_WEEKLY\_BACKUPS set as 5 and WEEKLY\_DAY\_OF\_WEEK is empty, then RETAIN\_WEEKLY\_BACKUPS won’t work. To work it, we need to keep valid values into RETAIN\_WEEKLY\_BACKUPS and WEEKLY\_DAY\_OF\_WEEK together.

**Logging**: Each entry into backup configuration table (BACKUP\_TBL) is assigned a primary key (auto increment). There is a log table BACKUP\_PROCESS\_LOG which stores list of statements executed in background and PK/FK keys are managed between these two tables (backup\_id).

**Restoring from backups**: At any point, if we need to restore data from backups (CLONE/PHYSICAL), naming of databases/schemas/tables need to be looked into. As stored procedure concatenates a date time to each backup, proper care needs to be taken while restoring. Restoring mechanism is not part of this component.

**Permissions**: Role which is executing this stored procedure should have grant of USAGE on source database and full access on target database (where backup is going to be created). You can find example script provided here.

**Email Alerts**: Upon successful completion or failure of stored procedure, it will send an automated email with relevant messages (success/fail) to the detailed email ID’s mentioned under **EMAIL\_ACCOUNTS** (please use comma separated lists for multiple accounts). These email ID’s may need to be validated with snowflake in advance.

**Sequence of Steps**:

Below are sequence of steps need to be followed for successful execution.

A screenshot of a computer

Description automatically generated

Please refer <https://docs.snowflake.com/en/user-guide/ui-preferences#verifying-your-email-address-in-the-classic-console>

**Scheduling**: Either use a scheduler of choice or you can use a snowflake task to execute this Stored procedure.

**Support:** Please contact Snowflake CoE in the event that you require help or have feedback.

[Csutherland@archgroup.com](mailto:Csutherland@archgroup.com)

[SivaM@archinsurance.com](mailto:SivaM@archinsurance.com)