日日是Oracle APEX

Oracle APEXを使った作業をしていて、気の付いたところを忘れないようにメモをとります。

2022年3月31日木曜日

データベースの内部表をオブジェクト・ストレージの外部表に 置き換える - 追記

こちらの記事の修正です。

DBMS CLOUD.EXPORT DATAによるエクスポートで上書きが発生しないようにする方法として、以 下がありました。

- 1. formatにmaxfilesize指定を含め、ファイルの最大サイズをあげる。
- 2. compressionにgizpを指定し、出力できるデータを増やす。

表BRICKSの内容をオブジェクト・ストレージにエクスポートするコードは以下になります。

```
declare
   C REGION constant varchar2(20) := 'us-ashburn-1';
    C_NAMESPACE constant varchar2(20) := 'ネームスペースに置き換える';
    C BUCKET constant varchar2(20) := 'bricks';
    C_FILENAME constant varchar2(20) := 'bricks';
   l_path varchar2(400);
begin
    l_path := 'https://objectstorage.' || C_REGION || '.oraclecloud.com/n/' || C_NAMESPACE
        || '/b/' || C_BUCKET || '/o/' || C_FILENAME;
    dbms_output.put_line(l_path);
    dbms_cloud.export_data
    (
       credential_name => 'DEF_CRED'
        , file_uri_list => l_path
        , format => json_object(
            'type' value 'csv'
           -- , 'maxfilesize' value '167772160'
            , 'compression' value 'gzip'
        , query => 'SELECT * FROM bricks'
    );
END;
/
                                                                                      view raw
```

export_data_fix.sql hosted with ♥ by GitHub

maxfilesizeやcompressionの指定が違いになります。

DBMS_CLOUD.EXPORT_DATAによる出力にはヘッダーが含まれません。そのため、外部表の作成スクリプトで指定されているskipheadersはコメントアウトします。

```
declare
   C_REGION constant varchar2(20) := 'us-ashburn-1';
    C_NAMESPACE constant varchar2(20) := 'ネームスペースに置き換える';
    C_BUCKET constant varchar2(20) := 'bricks';
    C_FILENAME constant varchar2(20) := 'bricks';
    l_path varchar2(400);
    l_filename varchar2(80);
    l_format clob;
begin
$IF true $THEN -- 圧縮されている場合
    l_filename := C_FILENAME || '*.csv.gz';
    l_format := json_object(
        'type'
                    value 'csv'
        , 'dateformat' value 'RR-MM-DD'
       -- , 'skipheaders' value '1'
        , 'compression' value 'auto'
    );
$ELSE
    l_filename := C_FILENAME || '*.csv';
    l_format := json_object(
        'type'
                    value 'csv'
        , 'dateformat' value 'RR-MM-DD'
        -- , 'skipheaders' value '1'
    );
$END
    l_path := 'https://objectstorage.' || C_REGION || '.oraclecloud.com/n/' || C_NAMESPACE
        || '/b/' || C_BUCKET || '/o/' || l_filename;
    dbms_cloud.create_external_table(
       credential_name => 'DEF_CRED'
        , table_name => 'bricks_ext'
        , file_uri_list => l_path
        , format
                     => l_format
        , column_list =>
            'brick_id number,
            colour varchar2(6),
            shape varchar2(8),
           weight number,
            insert_date date,
            junk varchar2(50)'
    );
end;
```

年ごとに分割した上で、オブジェクト・ストレージにエクスポートするスクリプトです。

```
declare
   -- 各種条件
   C_REGION constant varchar2(20) := 'us-ashburn-1';
    C_NAMESPACE constant varchar2(20) := 'ネームスペースに置き換える';
   C_BUCKET constant varchar2(20) := 'bricks';
   C_FILENAME constant varchar2(20) := 'bricks';
    -- エクスポートするデータを取り出すSELECT文
   l_path varchar2(400);
    l_sql varchar2(4000);
   C_SQL constant varchar2(4000) :=
q'~select
    brick_id
    , colour
    , shape
    , weight
    , insert_date
    , to_char(insert_date,'RRRR') year
    , junk
from bricks
where trunc(insert_date,'YEAR') = to_date('#NEW_YEAR_DAY#','RRRR-MM-DD')~';
begin
    /*
    * 年ごとにCSV形式でオブジェクト・ストレージに順次書き込む。
    */
    for c in
       -- 年単位でパーティションに分割する
       select to_char(year, 'RRRR') year, to_char(year, 'RRRR-MM-DD') new_year_day
           select distinct trunc(insert_date, 'YEAR') year from bricks
        ) order by year asc
    )
    loop
        l_path := 'https://objectstorage.' || C_REGION || '.oraclecloud.com/n/' || C_NAMESPACE
           || '/b/' || C_BUCKET || '/o/'
           || C_FILENAME || '/' || 'year=' || c.year || '/'
           || C_FILENAME || '-' || c.year;
        dbms_output.put_line(l_path);
        l_sql := replace(C_SQL, '#NEW_YEAR_DAY#', c.new_year_day);
        dbms_cloud.export_data
           credential_name => 'DEF_CRED'
```

パーティション分割された外部表を作成するスクリプトです。skipheadersの指定をコメントアウトしています。

```
declare
    C_REGION constant varchar2(20) := 'us-ashburn-1';
    C_NAMESPACE constant varchar2(20) := 'ネームスペースに置き換える';
    C_BUCKET constant varchar2(20) := 'bricks';
    C_FILENAME constant varchar2(20) := 'bricks';
    l_path varchar2(400);
    l_filename varchar2(80);
    l_format clob;
begin
$IF true $THEN -- 圧縮されている場合
    l_filename := C_FILENAME || '*.csv.gz';
    l_format := json_object(
        'type'
                    value 'csv'
        , 'dateformat' value 'RR-MM-DD'
        -- , 'skipheaders' value '1'
        , 'compression' value 'auto'
        , 'partition_columns' value '[{"name":"year","type":"varchar2(4)"}]' format json
    );
$ELSE
    l_filename := C_FILENAME || '*.csv';
    l_format := json_object(
        'type'
                    value 'csv'
        , 'dateformat' value 'RR-MM-DD'
       -- , 'skipheaders' value '1'
        , 'partition_columns' value '[{"name":"year","type":"varchar2(4)"}]' format json
    );
$END
    l_path := 'https://objectstorage.' || C_REGION || '.oraclecloud.com/n/' || C_NAMESPACE
        || '/b/' || C_BUCKET || '/o/' || l_filename;
    dbms_cloud.create_external_part_table(
        credential_name => 'DEF_CRED'
```

```
, table_name => 'bricks_part_ext'
        , file_uri_list => l_path
        , format
                        => l_format
        , column_list =>
            'brick_id number,
            colour varchar2(6),
            shape varchar2(8),
            weight number,
            insert_date date,
            year varchar2(4),
            junk varchar2(50)'
    );
end;
/
                                                                                           view raw
craete_external_table_part_fix.sql hosted with ♥ by GitHub
```

全体のエクスポートが30秒程度で終了するので、DBMS_CLOUD.EXPORT_DATAを使うと1/100くらいの処理時間になります。

完

Yuji N. 時刻: 19:40

共有

ホーム

ウェブ バージョンを表示

自己紹介

Yuji N.

日本オラクル株式会社に勤務していて、Oracle APEXのGroundbreaker Advocateを拝命しました。 こちらの記事につきましては、免責事項の参照をお願いいたします。

詳細プロフィールを表示

Powered by Blogger.