ECL in HPCC Systems

Lily Xu

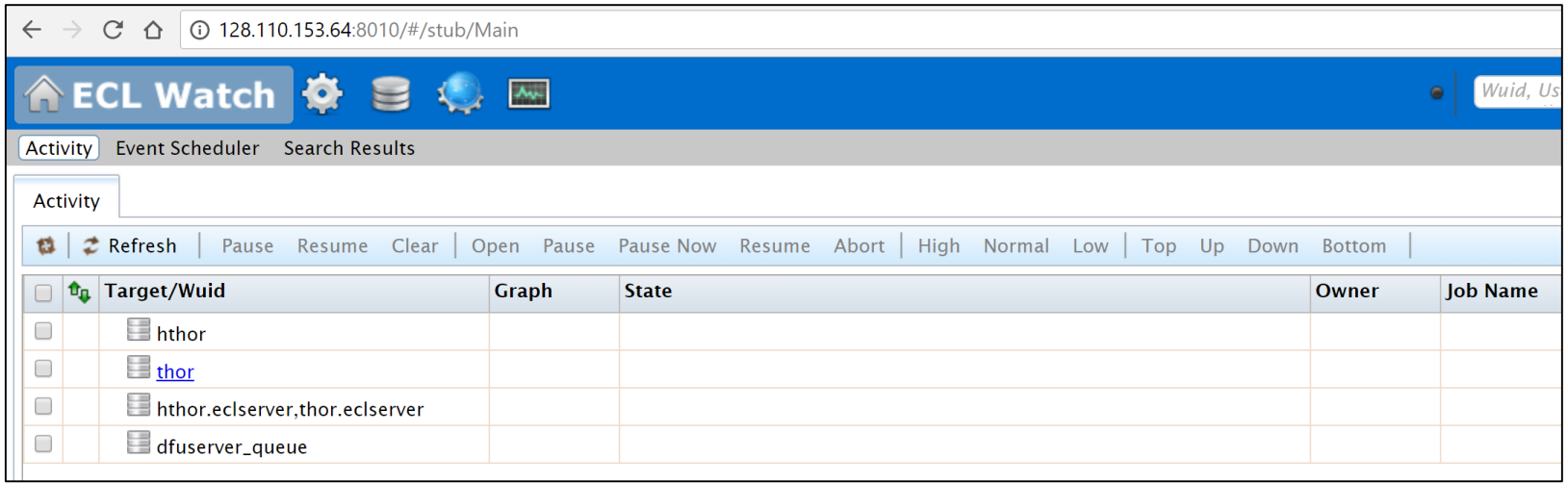
Step 1: Download dataset

The download is approximately 30 MB (compressed) and is available in either ZIP or .tar.gz format. Choose the appropriate link.

In ZIP Format: OriginalPerson.zip
In tar.gz Format: OriginalPerson.tar.gz

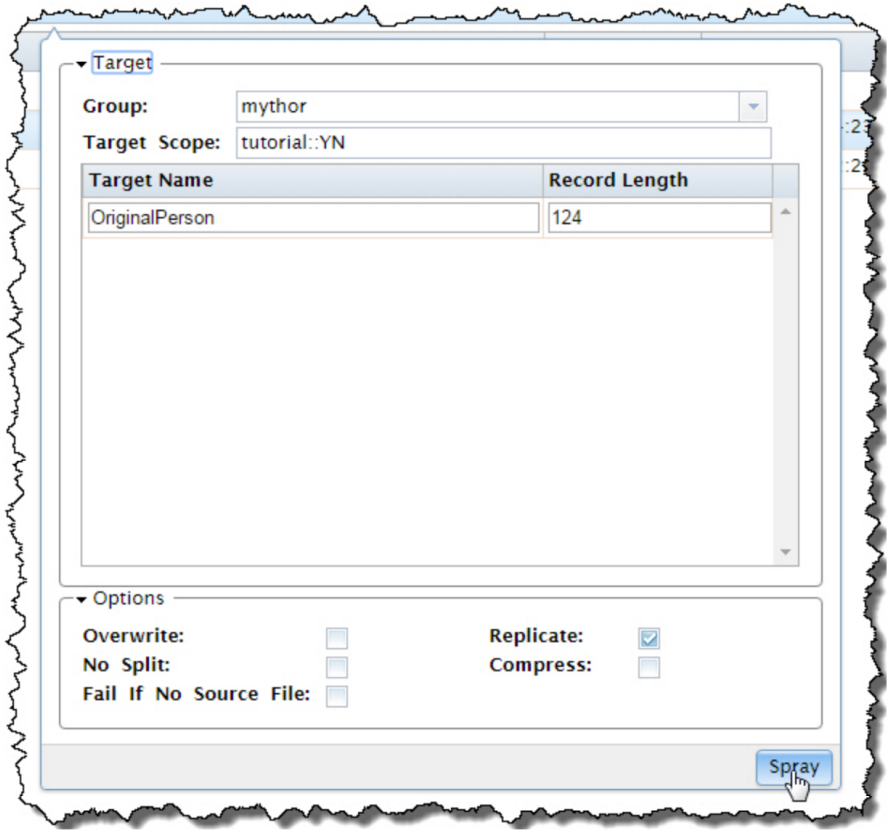
Get HPCC System Environment Ready

- Spray to Thor cluster: http://128.110.153.64:8010



Step 2: Spray the file to Thor

- The Target Scope should be ‘tutorial::yourname’
- The Target Name should be ‘OrginalPerson’



Step 3: Examine the data

- Begin Coding

1. Start the ECL IDE (Start >> All Programs >> HPCC Systems >> ECL IDE)

2. Log in to your environment

3. Right-click on the My Files folder in the Repository window, and select Insert Folder from the pop-up menu.

4. Enter TutorialYourName(where YourName is your name) for the label, then press the OK button.

5. Right-click on the TutorialYourName Folder, and select Insert File from the pop-up menu.

6. Enter Layout_People for the label, then press the OK button

7. Write the following code in the Builder workspace:

```
EXPORT Layout_People  := RECORD
       STRING15 FirstName;
       STRING25 LastName;
       STRING15 MiddleName;
       STRING5  Zip;
       STRING42 Street;
       STRING20 City;
       STRING2  State;
END;
```

1. Right-click on the TutorialYourName Folder, and select Insert File from the pop-up menu.
2. Enter File_OriginalPerson for the label, then press the OK button.
3. Write the following code

```
IMPORT TutorialYourName;
EXPORT File_OriginalPerson :=
DATASET('~tutorial::YN::OriginalPerson',TutorialYourName.Layout_People,THOR);
```

4. Press the syntax check button on the main toolbar (or press F7) to check the syntax.

This defines the Dataset. Next, we will examine the data.

5. Delete the EXPORT key word before the DATASET and add COUNT(TutorialYourName.File_OriginalPerson); to the end of the file.

6. Press the syntax check button on the main toolbar (or press F7) to check the syntax.

7. Make sure the selected cluster is your Thor cluster, then press the Submit button. Note that your target cluster might have a different name.

8. When the Workunit completes, it displays a green checkmark .

9. Select the Workunit tab (the one with the number next to the checkmark) and select the Result 1 tab (it may already

be selected).

10.Select the Builder tab and change COUNT to OUTPUT, as shown below:

```
IMPORT TutorialYourName;
OUTPUT(TutorialYourName.File_OriginalPerson);
```

11. Check the syntax, if no errors, press the Submit button.
12. When it completes, select the Workunit tab, then select the Result 1 tab.

Step 4: Process the Data

1. Right-click on the TutorialYourName Folder, and select Insert File from the pop-up menu.

2. Name this one BWR_ProcessRawData and write the following code (changing YN and YourName as before):

```
IMPORT TutorialYourName, Std;
    TutorialYourName.Layout_People toUpperPlease(TutorialYourName.Layout_People pInput)
      := TRANSFORM
     SELF.FirstName := Std.Str.ToUpperCase(pInput.FirstName);
     SELF.LastName := Std.Str.ToUpperCase(pInput.LastName);
     SELF.MiddleName := Std.Str.ToUpperCase(pInput.MiddleName);
     SELF.Zip := pInput.Zip;
     SELF.Street := pInput.Street;
     SELF.City := pInput.City;
     SELF.State := pInput.State;
    END ;

OrigDataset := TutorialYourName.File_OriginalPerson;
UpperedDataset := PROJECT(OrigDataset,toUpperPlease(LEFT));
OUTPUT(UpperedDataset,,'~tutorial::YN::TutorialPerson',OVERWRITE);
```

- 3. Check the syntax, if no errors press the Submit button.
- 4. When it completes, select the Workunit tab, then select the Result 1 tab.

Using our New Data

- 1. Insert a File into the TutorialYourName Folder. Name it File_TutorialPerson and write this code (changing YN to your initials):

```
IMPORT TutorialYourName;
EXPORT File_TutorialPerson :=
DATASET('~tutorial::YN::TutorialPerson',
       {TutorialYourName.Layout_People,
        UNSIGNED8 fpos {virtual(fileposition)}},THOR);
```

Index the Data

1. Insert a File into your Tutorial Folder. Name it IDX_PeopleByZipand write this code (changing YN and YourName as before):

```
IMPORT TutorialYourName;
EXPORT IDX_PeopleByZIP :=
INDEX(TutorialYourName.File_TutorialPerson,{zip,fpos},'~tutorial::YN::PeopleByZipINDEX');
```

2. Insert a File into the TutorialYourName Folder and name it BWR_BuildPeopleByZip and write this code (replacing YourName with your name):

```
IMPORT TutorialYourName;
BUILDINDEX(TutorialYourName.IDX_PeopleByZIP,OVERWRITE);
```

4. Check the syntax and if there are no errors, press the Submit button.
5. Wait for the Workunit to complete, then close the Builder Window

Step 5: Build a Query

- 1. Insert a File into your Tutorial Folder. Name it BWR_FetchPeopleByZip and write this code (changing YourName as before):

```
IMPORT TutorialYourName;
ZipFilter :='33024';
FetchPeopleByZip :=
FETCH(TutorialYourName.File_TutorialPerson,
      TutorialYourName.IDX_PeopleByZIP(zip=ZipFilter),
      RIGHT.fpos);
OUTPUT(FetchPeopleByZip);
```

- 2. Check the syntax and if there are no errors, press the Submit button.
- 3. When it completes, select the Workunit tab, then select the Result tab.
- 4. Examine the result, then close the Builder window and resubmit the code.

Step 6: Publishing your Thor Query

- 1. Insert a File into the TutorialYourName Folder and name it FetchPeopleByZipService
- 2. Write this code (changing YourName as before):

```
IMPORT TutorialYourName;
STRING10 ZipFilter := '' :STORED('ZIPValue');
resultSet :=
            FETCH(TutorialYourName.File_TutorialPerson,
            TutorialYourName.IDX_PeopleByZIP(zip=ZipFilter),
            RIGHT.fpos);
OUTPUT(resultset);
```

3. Check the syntax, and save the file.

4. Press the Submit button.

5. When the workunit completes, select the Workunit tab, then select the ECL Watch tab.

6. Press the Publish button, on the ECL Watch tab.

7. If there are no error messages, the workunit is published. Leave the builder window open, you will need it again later

Execute using WsECL

- Using the following URL:
- http://128.110.153.66 :8002
- (where nnn.nnn.nnn.nnn is your ESP Server's IP address and pppp is the port. Default port is 8002)

Step 7 : Test the queries

1. Click on the + sign next to thor to expand the tree.

2. Click on the fetchpeoplebyzipservice hyperlink. The form for the service displays

3. Provide a zip code (e.g., 33024) in the zipvalue field. Select Output Tables from the drop list, then press the Submit button. The results display.