

HPCC Data Tutorial

Equipe de documentação de Boca Raton



HPCC Data Tutorial

Equipe de documentação de Boca Raton

Copyright © 2021 HPCC Systems®. All rights reserved

Sua opinião e comentários sobre este documento são muito bem-vindos e podem ser enviados por e-mail para <docfeedback@hpccsystems.com>

Inclua a frase **Feedback sobre documentação** na linha de assunto e indique o nome do documento, o número das páginas e número da versão atual no corpo da mensagem.

LexisNexis e o logotipo Knowledge Burst são marcas comerciais registradas da Reed Elsevier Properties Inc., usadas sob licença.

HPCC Systems® é uma marca registrada da LexisNexis Risk Data Management Inc.

Os demais produtos e serviços podem ser marcas comerciais ou registradas de suas respectivas empresas.

Todos os nomes e dados de exemplo usados neste manual são fictícios. Qualquer semelhança com pessoas reais, vivas ou mortas, é mera coincidência.

2021 Version 8.2.80-1

Introdução	4
O processo de desenvolvimento do ECL	4
Trabalhando com Dados	5
Dados Originais	5
Programando	11
Publicando sua Consulta Thor	20
Compilar e Publicar uma Consulta Roxie	24
Resumo	27

Introdução

O processo de desenvolvimento do ECL

Este tutorial oferece uma orientação sobre o processo de desenvolvimento, do início até o fim, e foi projetado para servir como uma introdução para trabalhar com dados em qualquer um dos HPCC Systems.¹. Vamos programar em ECL² para processar nossos dados e consultá-los.

Este tutorial presume que:

- Você tem um HPCC em execução. Ele pode ser a VM Edition ou uma plataforma HPCC com um ou mais nós
- Ter o ECL IDE³ instalado e configurado

Neste tutorial, iremos:

- Baixar um arquivo de dados bruto

Há links para arquivos de dados disponíveis na página <http://hpccsystems.com/community/docs/data-tutorial-guide>

O download tem cerca de 30 MB (compactado) e está disponível nos formatos ZIP ou .tar.gz. Selecione o link correto.

- Realizar o spray do arquivo para um cluster da Refinaria de dados. Os clusters do HPCC realizam o processo de spray dos dados para partes de arquivo em cada nó.

Um *spray*, ou *importação*, é a transferência de um arquivo de dados de um local para um cluster do HPCC. O termo *spray* foi adotado devido à natureza da transferência dos arquivos – o arquivo é particionado entre todos os nós em um cluster.

- Examinar os dados e determinar o pré-processamento que precisamos realizar
- Realizar o pré-processamento dos dados para gerar um novo arquivo de dados
- Determinar os tipos de consultas desejadas
- Criar as consultas
- Testá-las
- Implementá-las para um cluster de motor de entrega rápida de dados (RDDE), também conhecido como um cluster Roxie.

¹High Performance Computing Cluster (HPCC) é uma plataforma de computação de processamento massivamente paralelo que soluciona problemas de big data. Acesse o link <http://www.hpccsystems.com/Why-HPCC/How-it-works> para obter mais detalhes

²Enterprise Control Language (ECL) é uma linguagem de programação declarativa e centrada em dados usada para gerenciar todos os aspectos da junção, classificação e compilação de dados massivos que realmente diferenciam o HPCC das demais tecnologias na sua capacidade de fornecer análise de dados flexíveis em escala massiva

³O ECL IDE (Ambiente de desenvolvimento integrado) é uma ferramenta usada para criar consultas em seus dados e arquivos ECL com os quais suas consultas serão compiladas.

Trabalhando com Dados

Dados Originais

Nesta situação, vamos receber um arquivo de dados estruturado que contenha registros com nomes e endereços de pessoas. O HPCC Systems também oferece suporte para dados não estruturados, mas este exemplo é mais simples. Este arquivo é documentado na tabela a seguir:

Nome do Campo	Tipo	Descrição
FirstName	Nome em string de 15 caracteres	Primeiro Nome
LastName	Nome em string de 25 caracteres	Sobrenome
MiddleName	Nome em string de 15 caracteres	Nome do Meio
ZIP	String de 5 caracteres	CEP
Street	String de 42 caracteres	Endereço
City	String de 20 caracteres	City
State	String de 2 caracteres	Estado

Isso totaliza um tamanho de registro de 124 (o total do tamanho de todos os campos). É necessário conhecer esse comprimento para o processo de **Spray do arquivo**.

Enviando os arquivos de dados de entrada para sua Zona de Entrada de Arquivos

Uma Zona de entrada (ou Drop Zone) é um local de armazenamento físico definido no ambiente do seu HPCC. Um daemon (DaFileSrv) precisa estar em execução no servidor para possibilitar sprays e desprays do arquivo.

Para arquivos de dados menores, é possível usar o utilitário de arquivos de envio/download no ECL Watch (uma interface baseada na Web para a sua plataforma HPCC). O arquivo de dados de amostra tem cerca de ~100 Mb.

1. Faça o download o arquivo de dados de amostra no portal do HPCC Systems.[®].

O arquivo de dados está disponível nos links na página <http://hpccsystems.com/community/docs/data-tutorial-guide>. O download tem cerca de 30 MB (compactado) e está disponível nos formatos ZIP ou .tar.gz. (**OriginalPerson.tar.gz** ou **OriginalPerson.zip**)

2. Extraia-o para uma pasta em sua máquina local.

3. Em seu navegador, acesse o URL do **ECL Watch**. Por exemplo, <http://nnn.nnn.nnn.nnn:8010>, onde nnn.nnn.nnn.nnn é o ESP.¹ Endereço IP do servidor.



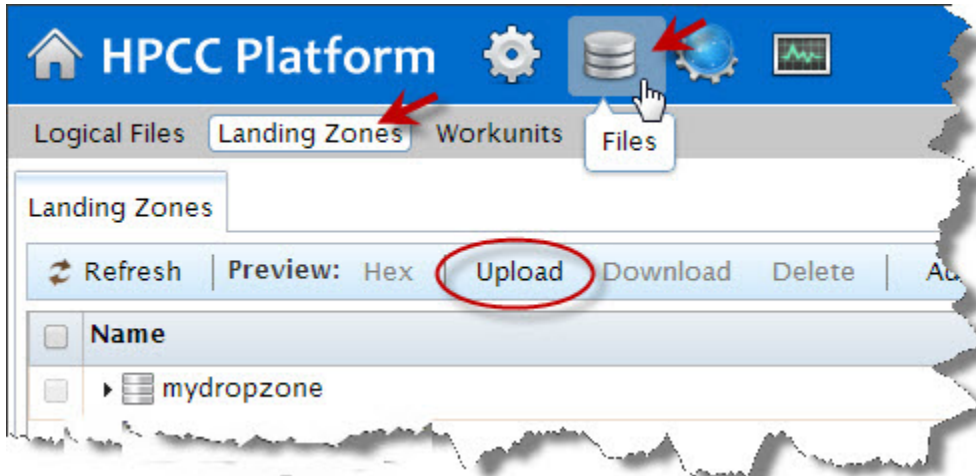
Seu endereço IP poderá ser diferente dos endereços fornecidos nas imagens de exemplo. Use o endereço IP fornecido pela **sua** instalação.

¹O ESP (Enterprise Services Platform) Server é o servidor de camada de comunicação no seu ambiente HPCC.

- Na página inicial do ECL Watch, clique no ícone **Files** e no link **Landing Zone (Zona de entrada de arquivos)** localizados no submenu de navegação.

Pressione o botão de ação **Upload** na guia Landing Zone (Zona de entrada de arquivos).

Figure 1. Upload/Download

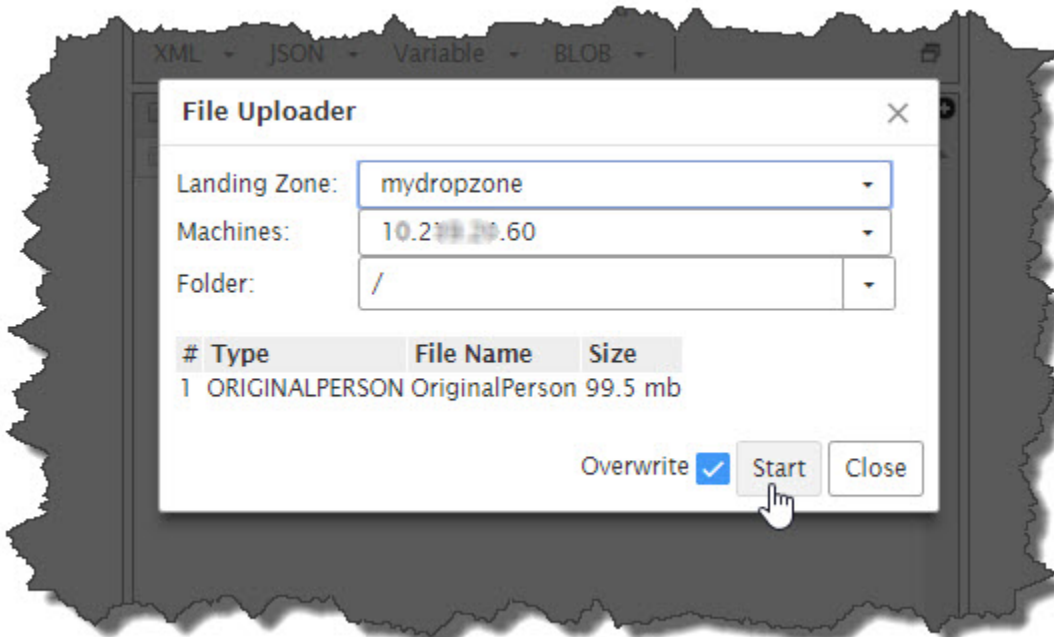


Depois de pressionar o botão Upload, é mostrada uma caixa de diálogo onde é possível selecionar um arquivo para envio.

- Navegue pelos arquivos na máquina local, selecione o arquivo a ser enviado e pressione o botão **Open**.

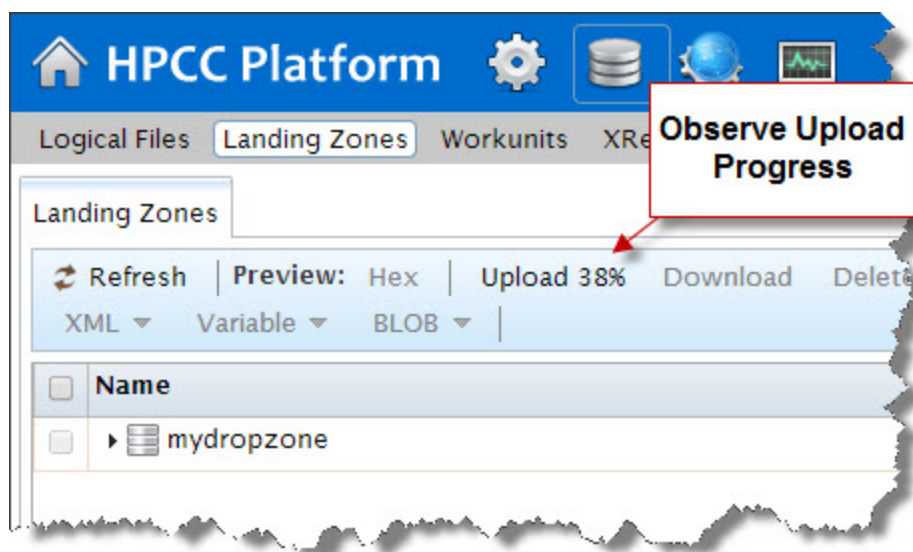
O arquivo selecionado é exibido na caixa de diálogo **File Uploader**.

Figure 2. Enviador de Arquivo



- Pressione o botão **Start** para concluir o envio do arquivo.

Figure 3. Progresso de envio



Spray do Arquivo de Dados para seu Cluster Thor

Para usar o arquivo de dados em nosso cluster HPCC System, precisamos primeiro realizar seu spray para um cluster Thor. Um *spray* ou *importação*, é a transferência de um arquivo de dados de um local para um cluster Thor. O termo spray foi adotado devido à natureza da transferência dos arquivos – o arquivo é particionado entre todos os nós em um cluster.

Neste exemplo, o arquivo está em sua zona de entrada de arquivos e é denominado **OriginalPerson**.

Vamos realizar seu spray para nosso cluster Thor e dar a ele um nome lógico de **tutorial::YN::OriginalPerson** onde **YN** são as suas iniciais. O Utilitário de arquivo distribuído mantém uma lista de arquivos lógicos e seus locais de arquivos físicos correspondentes.

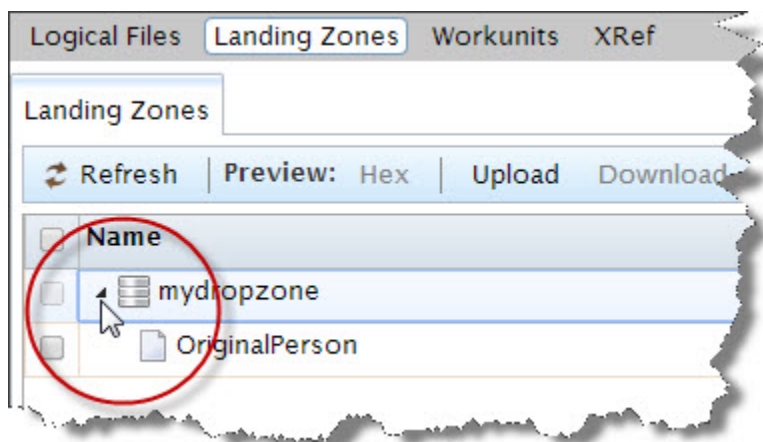
1. Abra o ECL Watch em seu navegador usando o seguinte URL:

http://nnn.nnn.nnn.nnn:pppp (onde **nnn.nnn.nnn.nnn** é o endereço IP do seu ESP Server e **pppp** é a porta. A porta padrão é 8010)

- Na página inicial do ECL Watch, clique no ícone **Files** e no link **Landing Zones** localizados no submenu de navegação.

Na guia landing Zone, clique na seta ao lado do contêiner mydropzone para ampliar a lista de arquivos enviados.

Figure 4. mydropzone

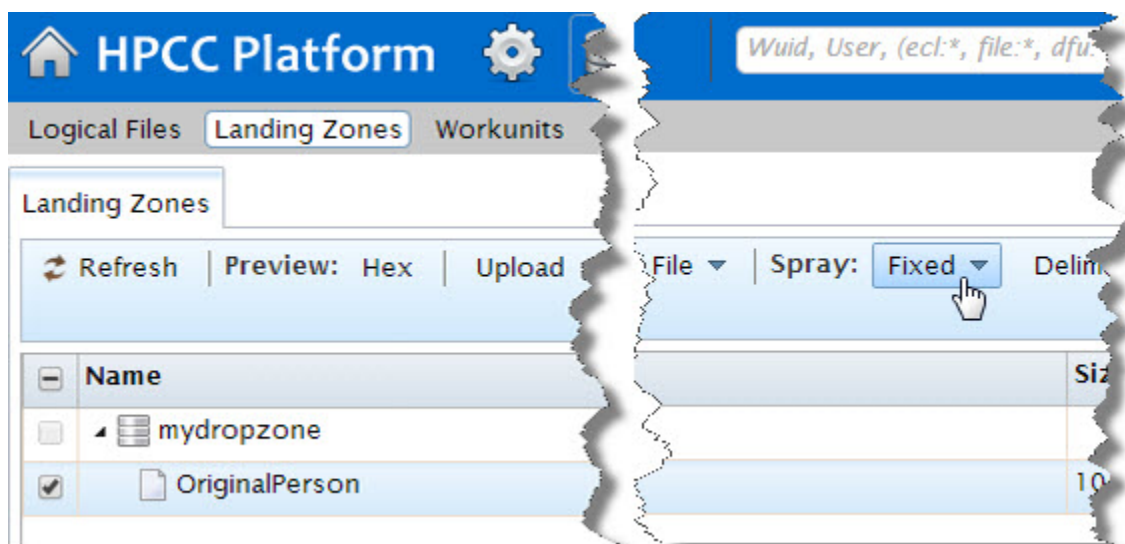


Encontre o arquivo que você deseja fazer spray na lista (OriginalPerson) e marque a caixa ao lado do nome do arquivo para selecioná-lo.

Depois de selecionar o arquivo na lista, os botões de ação **Spray** ficam ativados.

- Pressione o botão de ação **Fixed**. Isso indica que você está realizando o spray de um arquivo de largura fixa.

Figure 5. Spray: Botão de ação Fixed



A caixa de diálogo **Fixed Spray** é mostrada.

- O campo Target Name é preenchido automaticamente com o arquivo selecionado.

Figure 6. Caixa de diálogo Spray fixo



5. Selecione o cluster mythor na lista suspensa **Group** .
6. Preencher o **Tamanho do registro** (124).
7. Preencha o **Target Scope** usando a convenção de nomenclatura descrita anteriormente: **tutorial::YN** (lembre-se de que **YN** são as suas iniciais).

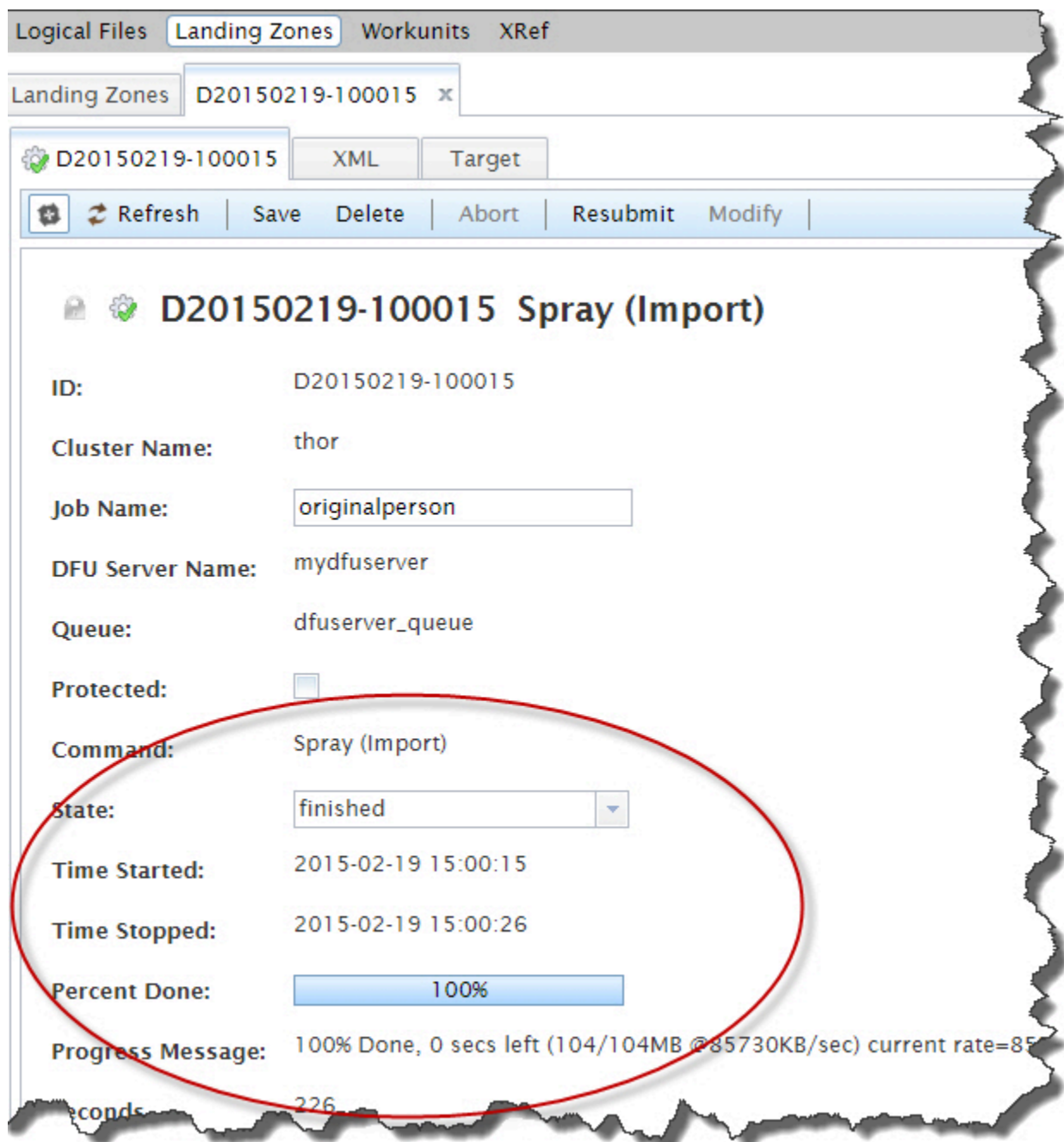
8. Não se esqueça de marcar a caixa **Replicate** .

Observação: Essa opção só está disponível em sistemas onde a replicação tiver sido ativada.

9. Pressione o botão **Spray** .

10.A página de detalhes da workunit é exibida. É possível ver o progresso do spray.

Figure 7. Visualizando Progresso



Logical Files **Landing Zones** Workunits XRef

Landing Zones D20150219-100015 x

D20150219-100015 XML Target

Refresh Save Delete Abort Resubmit Modify

D20150219-100015 Spray (Import)

ID: D20150219-100015

Cluster Name: thor

Job Name: originalperson

DFU Server Name: mydfuserver

Queue: dfuserver_queue

Protected: ☐

Command: Spray (Import)

State: finished

Time Started: 2015-02-19 15:00:15

Time Stopped: 2015-02-19 15:00:26

Percent Done: 100%

Progress Message: 100% Done, 0 secs left (104/104MB @85730KB/sec) current rate=85730KB/sec

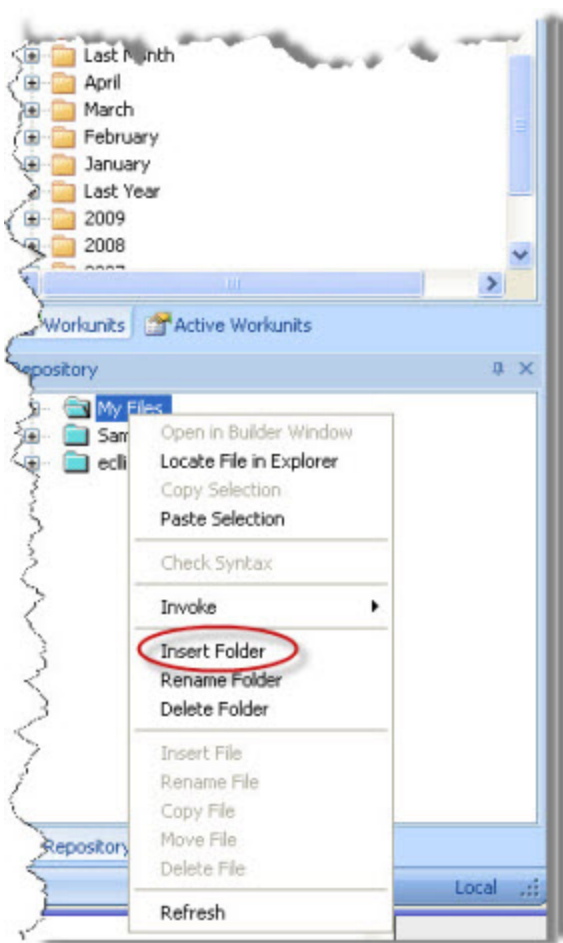
Depois de concluir o spray, podemos prosseguir.

Programando

Nesta parte do tutorial, vamos programar o código ECL para definir o arquivo de dados e executar consultas ECL simples nele, para que possamos avaliá-lo e determinar qualquer pré-processamento necessário.

1. Inicie o IDE ECL (Start >> All Programs >> HPCC Systems >> IDE ECL)
2. Faça login em seu ambiente
3. Clique com o botão direito na pasta **My Files** na janela Repository e selecione **Insert Folder** no menu pop-up.

Figure 8. Insert Folder



Para finalidades deste tutorial, vamos criar uma pasta chamada **Tutorial YourName** (onde *YourName* é o seu nome).

4. Insira **Tutorial YourName** (onde *YourName* é o seu nome) para o rótulo e depois pressione o botão OK.

Figure 9. Inserir o rótulo da pasta



5. Clique com o botão direito na pasta **Tutorial YourName** e selecione **Insert File** no menu pop-up.
6. Insira **Layout_People** para o rótulo e depois pressione o botão OK.

Figure 10. Inserir arquivo



Uma janela do compilador será aberta.

Figure 11. Layout de pessoas no compilador



Observe que parte do texto já foi escrita para você na janela. Isso ajuda a lembrar que o nome do arquivo (Layout_People) *sempre precisa corresponder exatamente* ao nome da definição EXPORT única (Layout_People) contida no arquivo. Isso é obrigatório – uma definição de EXPORT por arquivo e seu nome precisa corresponder ao nome do arquivo.

7. Escreva o código a seguir no espaço de trabalho do Compilador:

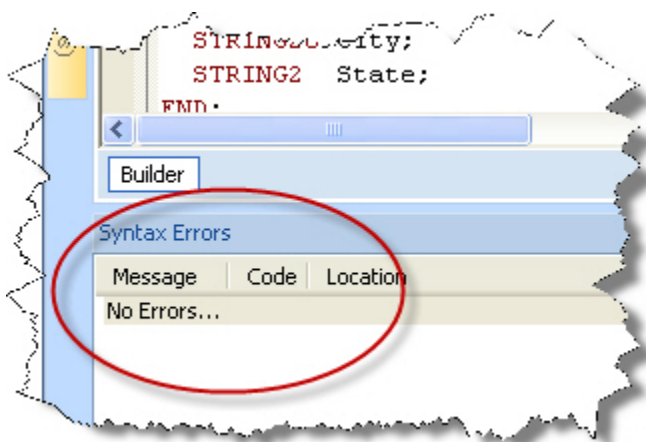
```
EXPORT Layout_People := RECORD
  STRING15 FirstName;
  STRING25 LastName;
  STRING15 MiddleName;
  STRING5 Zip;
  STRING42 Street;
  STRING20 City;
  STRING2 State;
END;
```

Figure 12. Programar na janela do Compilador



8. Pressione o botão de verificação de sintaxe localizado na barra de ferramentas principal (ou pressione F7).
É sempre recomendado verificar a sintaxe antes do envio.

Figure 13. Verificar a sintaxe



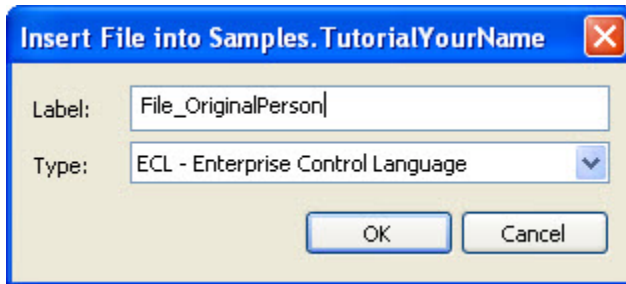
Esse arquivo define a estrutura de registro para o arquivo de dados. Em seguida, vamos examinar os dados.

Examinando os Dados

Nesta seção, vamos analisar os dados e determinar se é recomendado realizar algum pré-processamento. Essa é a etapa no processo de desenvolvimento na qual convertemos dados brutos para um formato utilizável.

1. Clique com o botão direito na pasta **Tutorial YourName** e selecione **Insert File** no menu pop-up.
2. Insira **File_OriginalPerson** para o rótulo e depois pressione o botão OK.

Figure 14. Inserir arquivo

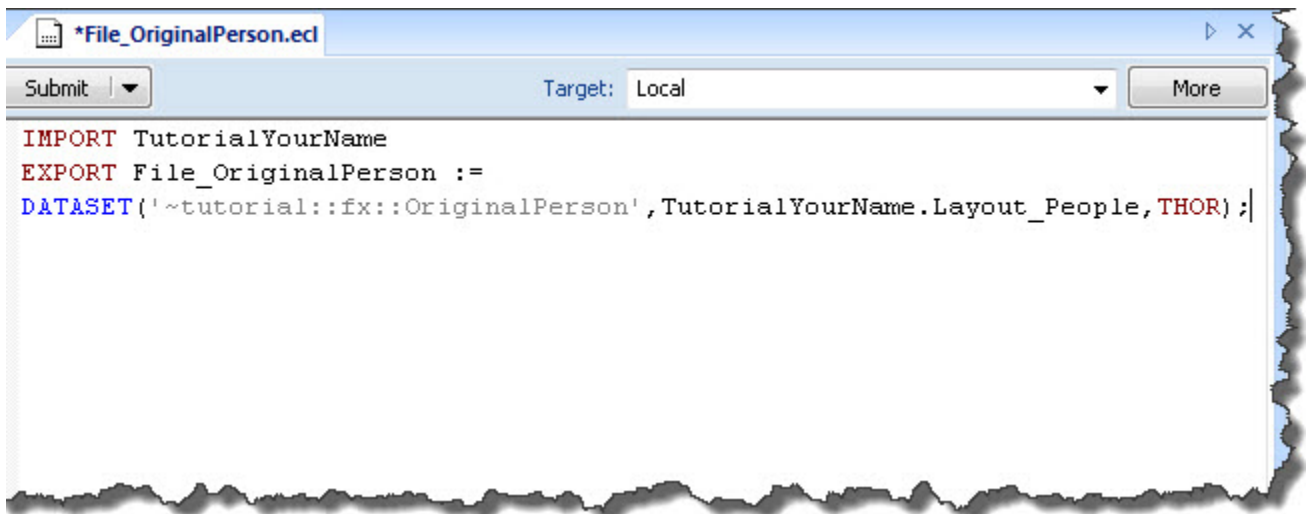


Uma janela do compilador será aberta.

3. Escreva o código a seguir (lembre-se de substituir *YN* pelas suas iniciais):

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

Figure 15. File_OriginalPerson.ecl



4. Pressione o botão de verificação de sintaxe localizado na barra de ferramentas principal (ou pressione F7) para verificar a sintaxe.

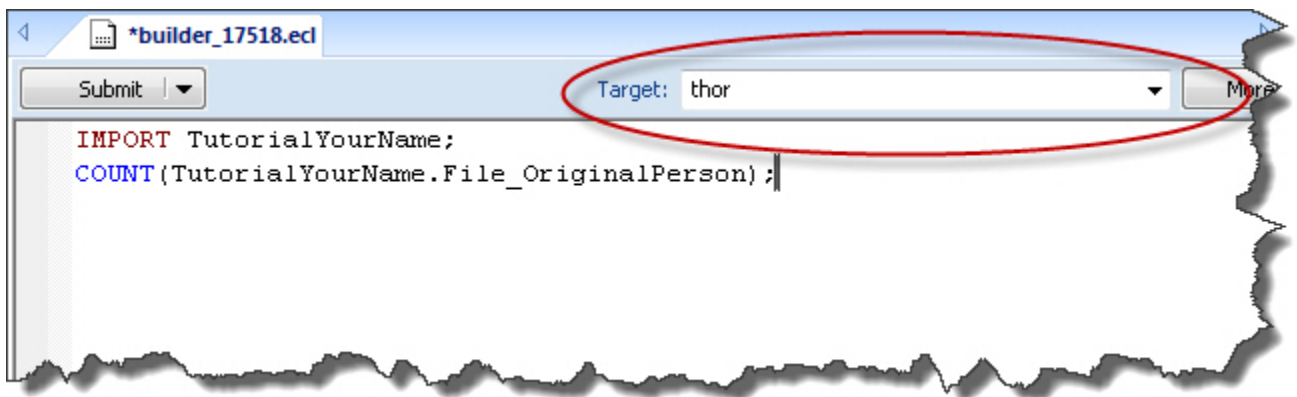
Isso define o dataset. Em seguida, vamos examinar os dados.

5. Abra uma nova janela no Compilador (CTRL+N) e escreva o código a seguir (lembre-se de substituir *YourName* pelo seu nome)

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

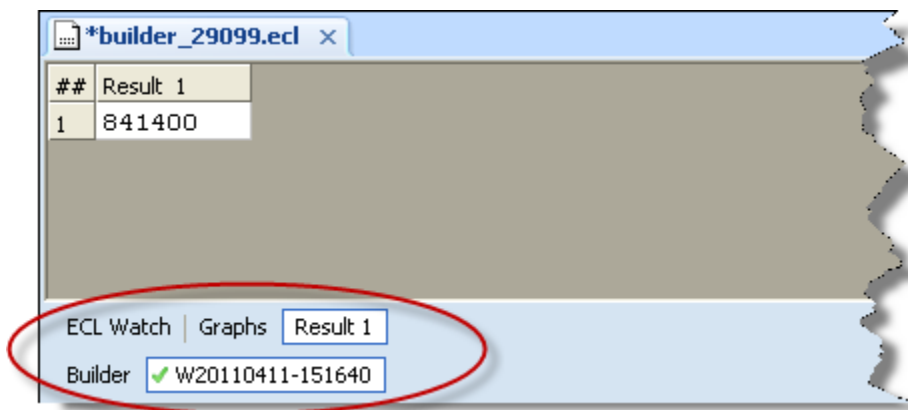
6. Pressione o botão de verificação de sintaxe localizado na barra de ferramentas principal (ou pressione F7) para verificar a sintaxe.
7. Verifique se o cluster selecionado é o Thor e pressione o botão **Submit** . Observe que o cluster de destino pode ter um nome diferente.

Figure 16. Thor de destino



8. Quando a tarefa for concluída, ela aparece com uma marca de verificação verde ✓.
9. Selecione a aba Workunit (aquela com o número ao lado da marca de verificação) e selecione a guia **Result 1** (ela pode já ter sido selecionada)

Figure 17. Aba Result



Essa guia nos mostra que há 841.400 registros no arquivo de dados.

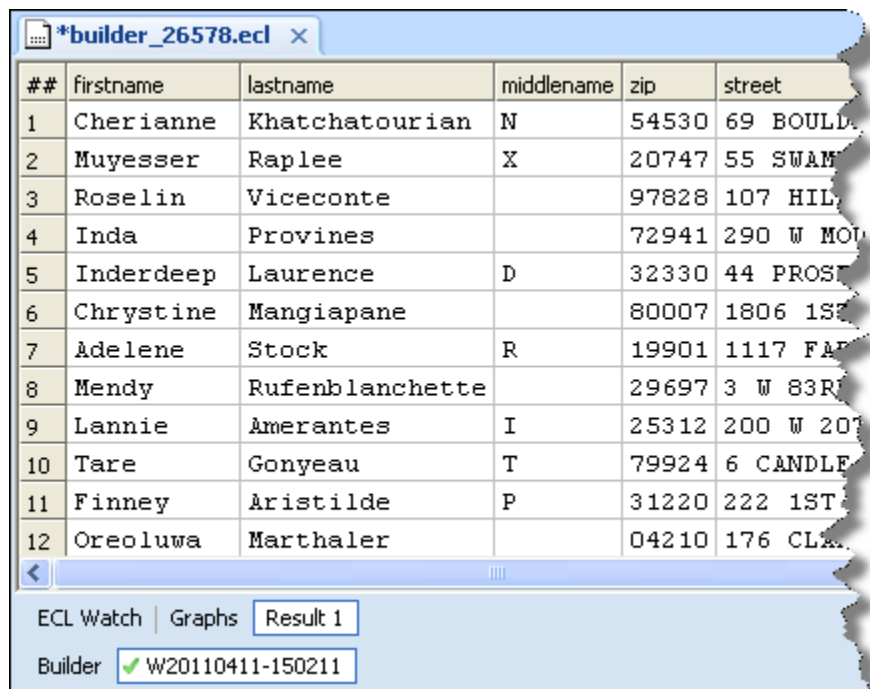
10. Selecione a aba Compilador e altere COUNT para OUTPUT como mostrado abaixo:

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

Observação: A parte modificada é mostrada em **negrito**.

11. Verifique a sintaxe e, se não houver erros, pressione o botão **Submit**.
12. Quando o envio estiver concluído, selecione a aba **Workunit** e em seguida a aba **Result 1**.

Figure 18. Exibição dos Resultados



##	firstname	lastname	middlename	zip	street
1	Cherianne	Khatchatourian	N	54530	69 BOULDA
2	Muyesser	Raplee	X	20747	55 SWAM
3	Roselin	Viceconte		97828	107 HIL
4	Inda	Provines		72941	290 W MO
5	Inderdeep	Laurence	D	32330	44 PROS
6	Chrystine	Mangiapane		80007	1806 1ST
7	Adelene	Stock	R	19901	1117 FAR
8	Mendy	Rufenblanchette		29697	3 W 83R
9	Lannie	Amerantes	I	25312	200 W 20
10	Tare	Gonyeau	T	79924	6 CANDLE
11	Finney	Aristilde	P	31220	222 1ST
12	Oreoluwa	Marthaler		04210	176 CLA

Observe que os nomes estão em caixa mista.

Para nossas finalidades, será mais fácil ter todos os nomes em caixa alta. Isso demonstra uma das etapas no processo básico de preparação de dados (extrair, transformar e carregar – ETL) usando ECL.

13. Feche a janela do Compilador.

Processando os Dados

Nesta seção, vamos programar o código para converter os dados originais para que todos os nomes fiquem em caixa alta. Vamos então gravar esse novo arquivo em nosso cluster Thor.

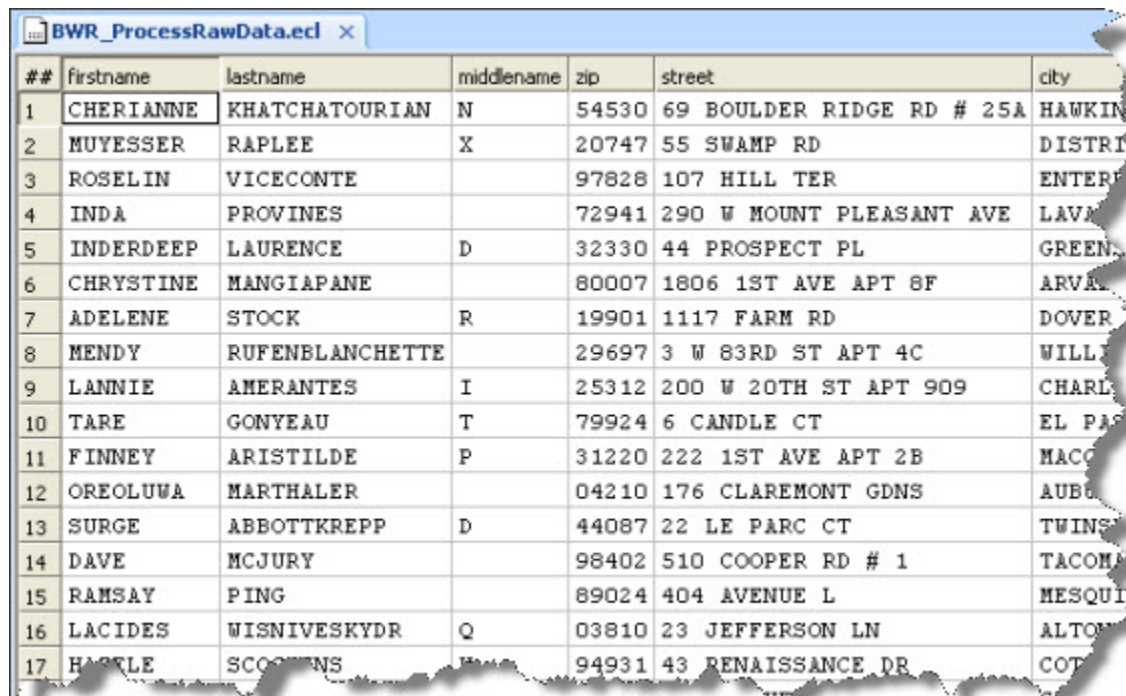
1. Clique com o botão direito na pasta **Tutorial YourName** e selecione Insert File no menu pop-up.
2. Nomeie esse arquivo como **BWR_ProcessRawData** e escreva o seguinte código (alterando YN e YourName como antes):

```
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. Verifique a sintaxe e, se não houver erros, pressione o botão **Submit**.
4. Quando o envio estiver concluído, selecione a aba Workunit e em seguida a aba Results 1.

Figure 19. Processar resultados



#	firstname	lastname	middlename	zip	street	city
1	CHERRIANNE	KHATCHATOURIAN	N	54530	69 BOULDER RIDGE RD # 25A	HAWKINS
2	MUYESSER	RAPLEE	X	20747	55 SWAMP RD	DISTRICT
3	ROSELIN	VICECONTE		97828	107 HILL TER	ENTERPRISE
4	INDA	PROVINES		72941	290 W MOUNT PLEASANT AVE	LAVAN
5	INDERDEEP	LAURENCE	D	32330	44 PROSPECT PL	GREEN
6	CHRYSTINE	MANGIAPANE		80007	1806 1ST AVE APT 8F	ARVA
7	ADELENE	STOCK	R	19901	1117 FARM RD	DOVER
8	MENDY	RUFENBLANCHETTE		29697	3 W 83RD ST APT 4C	WILL
9	LANNIE	AMERANTES	I	25312	200 W 20TH ST APT 909	CHARLE
10	TARE	GONYEAU	T	79924	6 CANDLE CT	EL PASO
11	FINNEY	ARISTILDE	P	31220	222 1ST AVE APT 2B	MACO
12	OREOLUWA	MARTHALER		04210	176 CLAREMONT GDNS	AUBU
13	SURGE	ABBOTTKREPP	D	44087	22 LE PARC CT	TWINS
14	DAVE	MCJURY		98402	510 COOPER RD # 1	TACOMA
15	RAMSAY	PING		89024	404 AVENUE L	MESQUIT
16	LACIDES	WISNIVESKYDR	Q	03810	23 JEFFERSON LN	ALTON
17	HARPLE	SCOTTENS	H	94931	43 RENAISSANCE DR	COT

Os resultados mostram que o processo converteu corretamente os campos de nome para caixa alta.

5. Depois de examinar os resultados, feche a janela do Compilador.


Utilizando nossos Novos Dados

Agora que temos nossos dados em um formato útil e o arquivo está salvo, podemos escrever mais códigos para usar o novo arquivo de dados. Vamos determinar os índices que precisamos e criá-los. Para esse tutorial, vamos supor que o campo que precisamos indexar é o do CEP.

Na definição DATASET, vamos adicionar um campo virtual à estrutura RECORD para a posição de arquivo. Isso é necessário para os índices.

1. Insira um arquivo na pasta **Tutorial YourName** . Nomeie-o como **File_TutorialPerson** e escreva esse código (alterando *YN* para as suas iniciais):

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

2. Verifique a sintaxe e, se não houver erros, pressione o botão **Submit** .
3. Depois de concluir, ela aparece com uma marca de verificação verde  .

Indexando os Dados

Em seguida, vamos definir o INDEX.

1. Insira um arquivo na pasta Tutorial. Nomeie-o como **IDX_PeopleByZip** e escreva esse código (alterando o *YN* e *YourName* como antes):

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

2. Verifique a sintaxe.

Em seguida, vamos compilar o arquivo de índice.

3. Insira um arquivo na pasta **Tutorial YourName** , nomeie-o como **BWR_BuildPeopleByZip** e escreva esse código (substituindo *YourName* pelo seu nome):

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

4. Verifique a sintaxe e, caso não haja erros, pressione o botão **Submit** .
5. Aguarde a conclusão da tarefa e feche a janela do Compilador.

Construa a Consulta

Agora que temos um arquivo de índice, vamos escrever uma consulta que o utilize.

1. Insira um arquivo na pasta Tutorial. Nomeie-o **BWR_FetchPeopleByZip** e escreva esse código (alterando *YourName* como antes):

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

2. Verifique a sintaxe e, caso não haja erros, pressione o botão **Submit** .
3. Quando o envio estiver concluído, selecione a aba **Workunit** e em seguida a aba **Results** .
4. Examine o resultado, feche a janela do Compilador e reenvie o código.

Observação: É possível alterar o valor do campo **ZipValue** para obter resultados de diferentes CEPs.

Publicando sua Consulta Thor

Agora que criamos uma consulta indexada, a próxima etapa é possibilitar o acesso a ela através de uma interface da Web.

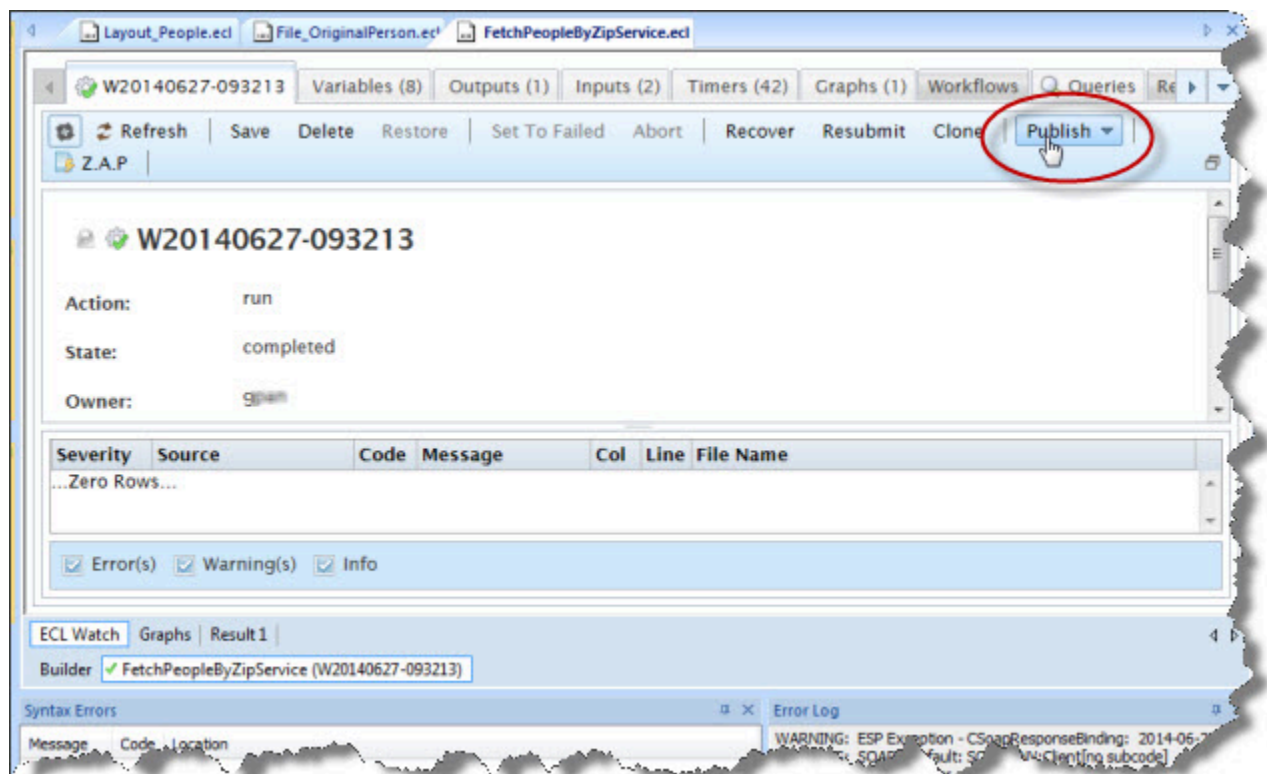
Nossas variáveis STORED oferecem um meio para passar valores como parâmetros de consulta. Neste exemplo, o usuário pode fornecer o ZIP (CEP) para que os resultados sejam pessoas na região do ZIP (CEP).

1. Insira um arquivo na pasta **TutorialYourName** e nomeie-o como **FetchPeopleByZipService**
2. Escreva esse código (alterando *YourName* como antes):

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

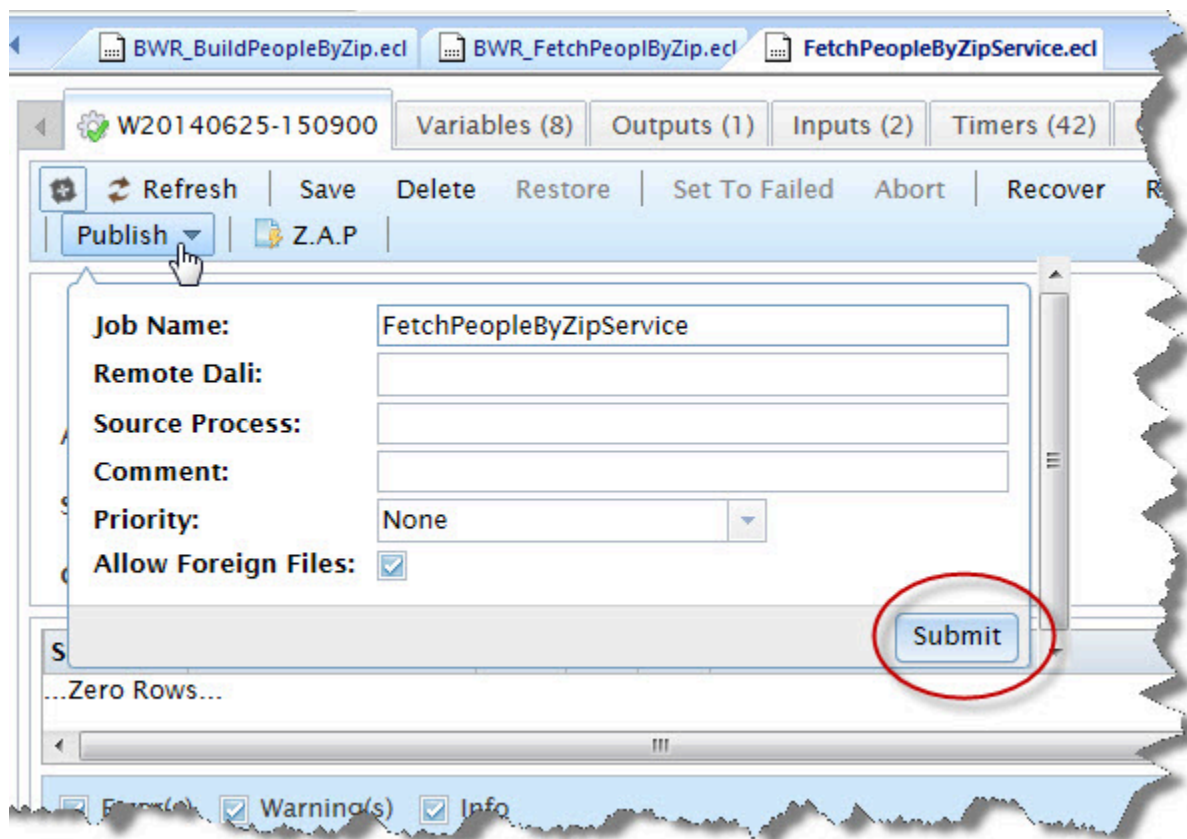
3. Verifique a sintaxe e salve o arquivo.
4. Pressione o botão **Submit**.
5. Quando a tarefa estiver concluída, selecione a aba Workunit e em seguida a aba ECL Watch.
6. Pressione o botão **Publish** na guia ECL Watch.

Figure 20. Publicar Workunit



A caixa de diálogo Publish é exibida, com o campo Job Name automaticamente preenchido. Você pode adicionar um comentário no campo Comment se quiser, e depois pressione Submit.

Figure 21. Caixa de diálogo Publish



7. Se não houver mensagens de erro, a workunit é publicada. Mantenha a janela do Compilador aberta – ela será necessária posteriormente.

Executar utilizando WsECL

Agora que a consulta foi publicada, podemos executá-la usando o serviço da Web WsECL. O WsECL oferece uma interface baseada na Web para sua consulta publicada. Ele também cria automaticamente um formulário de entrada para executar a consulta.

Use a seguinte URL:

http://nnn.nnn.nnn.nnn:pppp (onde nnn.nnn.nnn.nnn é o endereço IP do seu ESP Server e pppp é a porta. Porta padrão é 8002)

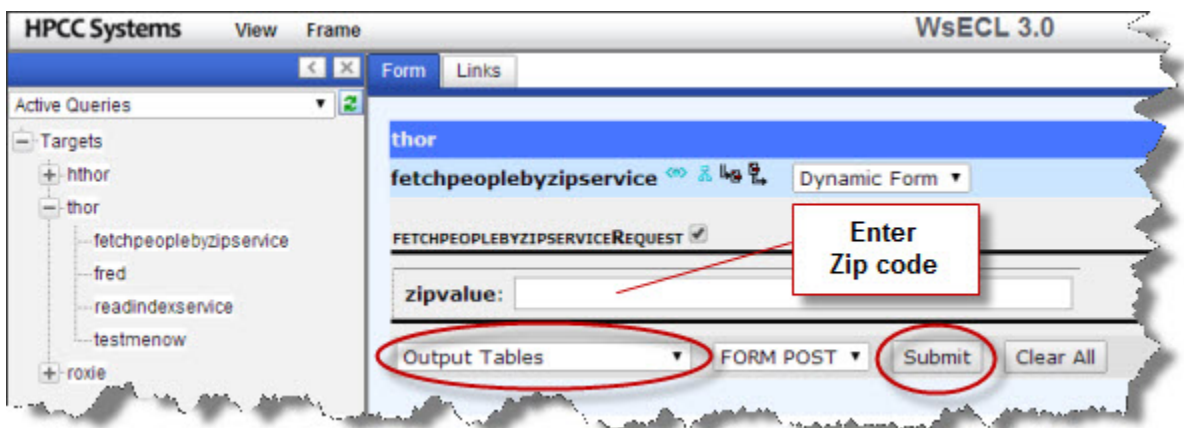
Figure 22. WsECL



1. Clique no sinal + ao lado de **thor** para expandir a árvore.
2. Clique no hiperlink **fetchpeoplebyzipservice**.

O formulário de serviço será exibido.

Figure 23. Formulário de serviço



3. Informe um CEP (por exemplo, 33024) no campo **zipvalue**. Selecione **Output Tables** na lista suspensa e pressione o botão **Submit**.

Os resultados serão exibidos.

Figure 24. Resultados



HPCC Systems View Frame

Active Queries

Targets

- + hthor
- thor
- ... fetchpeoplebyzipservice
- + roxie

fetchpeoplebyzipservice Response

Dataset: Result 1

	firstname	lastname	middlename	zip	
1	ZHIAGN	ARCHIE	W	33024	3
2	ASHANTI	TSAMES		33024	200
3	KIRAKOS	ROBSHAW		33024	1630
4	LEESHA	MONTALVO		33024	1 AST
5	REYNOLD	BELLERO		33024	489 WOL
6	KAROLYN	SCATENA		33024	2 NEW

Compilar e Publicar uma Consulta Roxie

A última etapa no processo é publicar a consulta indexada para um cluster de motor de entrega rápida de dados (Roxie).

Vamos recompilar o código com o Roxie como cluster de destino e depois publicá-lo em um cluster Roxie.

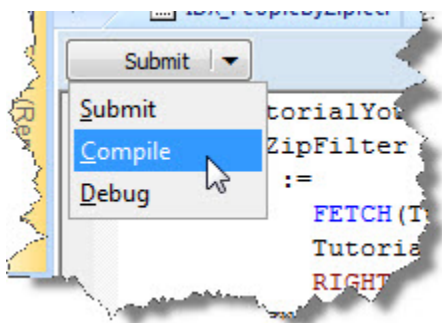
1. No ECL IDE, selecione a guia Compilador na janela do compilador de arquivos FetchPeopleByZipService.
2. Usando a lista suspensa **Target** , selecione Roxie como o cluster de destino.

Figure 25. Roxie de destino



3. Na janela do compilador, no canto superior esquerdo, há uma seta suspensa ao lado do botão **Submit** . Selecione a seta para exibir a opção **Compile** (compilar.)

Figure 26. Compilar



4. Selecione **Compile (Compilar)**

5. Quando a workunit for concluída, ela exibirá um círculo verde indicando que foi compilada.

Figure 27. Compilada

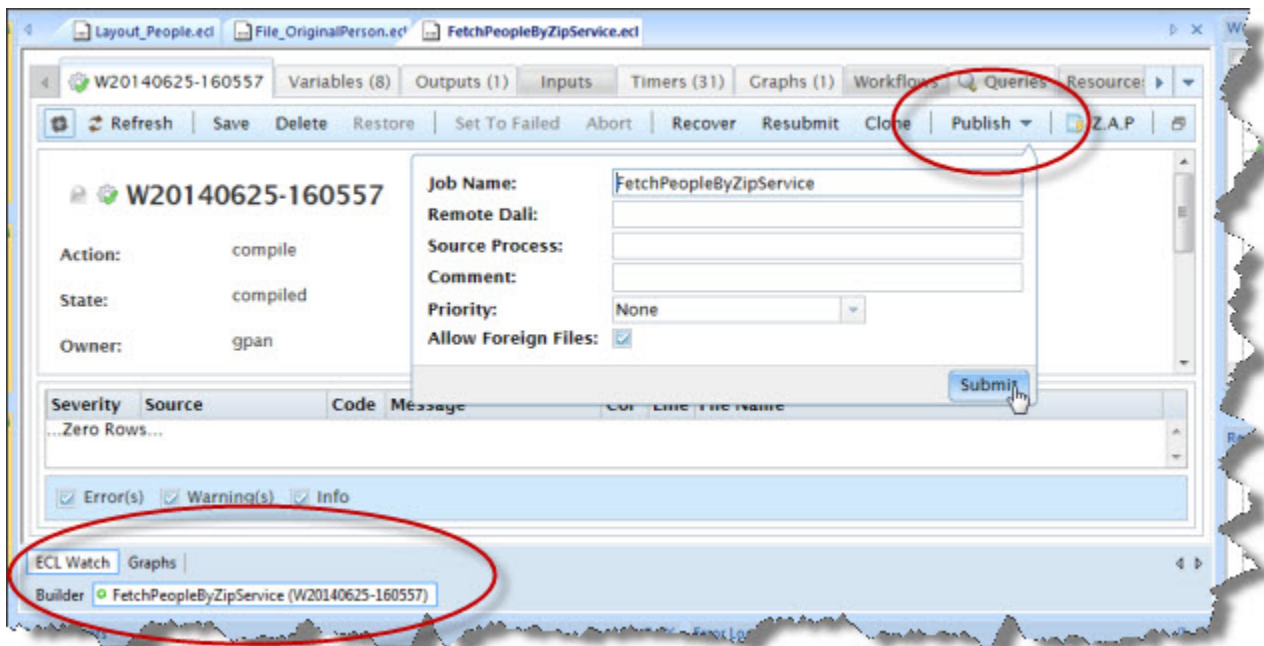


Publicar uma consulta Roxie

Agora vamos publicar a consulta em um cluster Roxie.

1. Selecione a aba Workunit para a workunit FetchPeopleByZipService que você acabou de compilar.
Isso abre a tarefa em uma aba no ECL Watch.
2. Pressione o botão de ação **Publish**, e depois verifique as informações na caixa de diálogo e pressione **Submit**.

Figure 28. Publicar consulta



Isso publica a consulta.

Execute uma consulta Roxie Query no WsECL

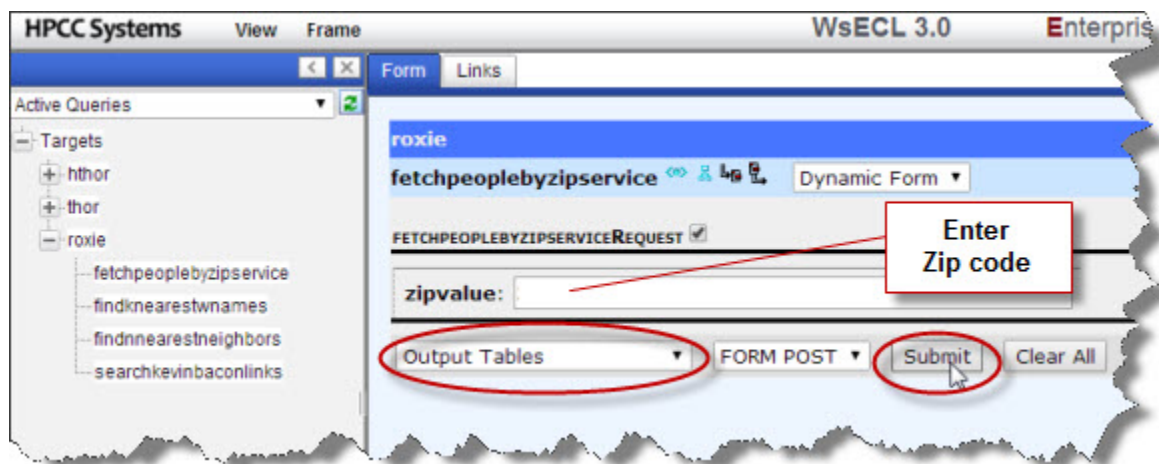
Agora que a consulta foi implementada a um cluster Roxie, podemos executá-la usando o serviço WS-ECL no URL a seguir:

http://nnn.nnn.nnn.nnn:pppp (onde **nnn.nnn.nnn.nnn** é o endereço IP do seu ESP Server e **pppp** é a porta. A porta padrão é 8002)

1. Clique no sinal + ao lado de **myroxie** para expandir a árvore.
2. Clique no hiperlink **fetchpeoplebyzipservice**.

O formulário de serviço será exibido.

Figure 29. RoxieECL

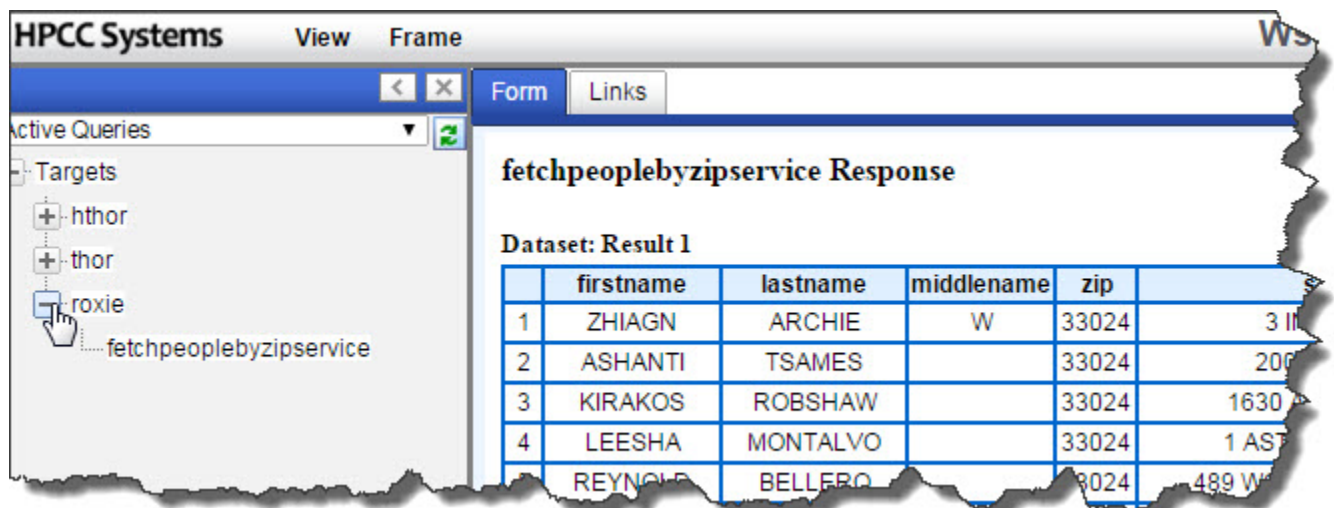


The screenshot shows the HPCC Systems WsECL 3.0 interface. On the left, the 'Active Queries' tree is expanded to show 'roxie' and its sub-items, including 'fetchpeoplebyzipservice'. The main panel displays the 'fetchpeoplebyzipservice' form. It includes a 'zipvalue:' input field, a 'Dynamic Form' dropdown, a 'FORM POST' dropdown, and a 'Submit' button. A red box highlights the 'zipvalue:' field with the text 'Enter Zip code'. The 'Output Tables' dropdown menu is circled in red, and the 'Submit' button is also circled in red.

3. Informe um CEP (por exemplo, 33024), selecione **Output Tables** na lista suspensa e pressione o botão Submit.

Os resultados serão exibidos.

Figure 30. RoxieResults



The screenshot shows the HPCC Systems WsECL 3.0 interface displaying the 'fetchpeoplebyzipservice Response'. The 'Dataset: Result 1' is shown as a table with the following data:

	firstname	lastname	middlename	zip	
1	ZHIAGN	ARCHIE	W	33024	3 II
2	ASHANTI	TSAMES		33024	200
3	KIRAKOS	ROBSHAW		33024	1630
4	LEESHA	MONTALVO		33024	1 AS
	REYNOL	BELLER		3024	489 W

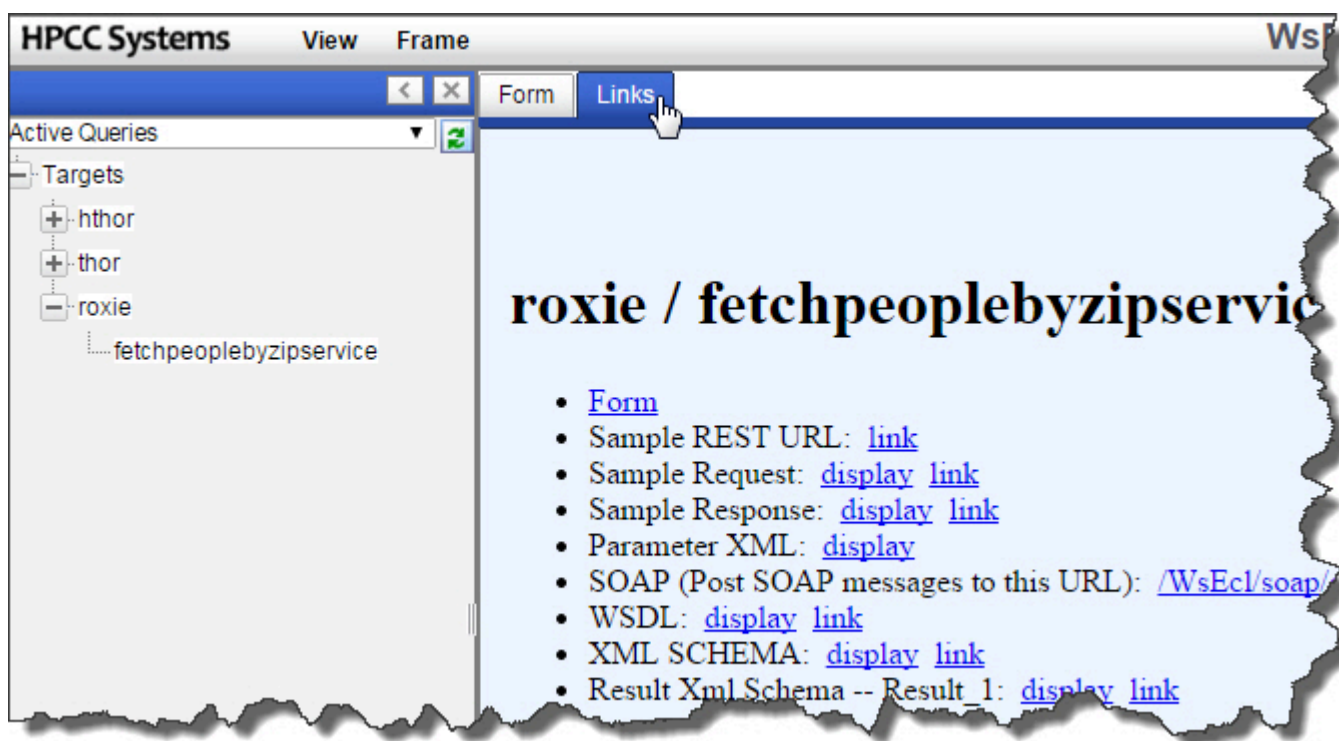
Resumo

Agora que os dados brutos foram processados corretamente, o spray para um cluster foi feito e os dados foram implementados em um cluster RDDE, qual é a próxima etapa?

Aqui está uma pequena lista de sugestões do que você pode fazer a seguir:

- Criar índices em outros campos e criar consultas usando esses índices.
- Programar aplicações client para acessar suas consultas usando interfaces
- Conferir os recursos disponíveis na aba Links

Figure 31. Links



A aba Links oferece fácil acesso a um formulário, uma amostra da busca, uma amostra de resposta, o WSDL, o esquema XML (XSD) e muito mais.

- Siga os procedimentos neste tutorial usando seus próprios dados!