

About Me

Vamsi Krishna

I am Vamsi Krishna.

working for a software

company in chennai.

India. I love to work

on Unix,Datastage

View my complete

profile

constantly keep trying things on Datastage.

and Database.I

This blog Contains Useful information collected from notes, documents, forums or blogs from the Internet

Showing posts with label Datastage-Scenarios. Show all posts

Thursday, April 18, 2013

Datastage scenario with small example

Consider Input file.txt as below

telephoneno 09700020075 919889110102 918571233668

Output.txt

telephoneno 09700020075 09889110102

08571233668 and length should be 11.If first two characters is "91" i need to replace as "0" $\,$

Ans:

In Transformer use following derivation

If (Len(InLink.Phone) = 11) Then InLink.Phone Else "0"
Right(InLink.Phone,10)

Common steps for Datastge Job Development:

1)Understand the requirement clearly and also exceptions

2)Form the algorithm in simple english and Do not go for Job development directly

Algorithm:

a)Check whether length of the string is 11 or not

b)If it 11 pass the input else extract right most 10 characters and append with zero $\,$

3)Now convert the Algorithm into Datastage stages

a) To implement first point in the Algorithm we need to use Len function and also If statement

b)To implement second point in the Algorithm we need to use Right function

4)In most of the cases requirement is Spilted into Jobs and we have to identity the stages used in each Job after forming the algorithm which makes you debugging easier instead of implementing entire requirement in one Job.

To identify when to use which stage the following link will help youwhen-to-use-which-stage-in-datastage

5)Performance tuning if you are facing any problem which is an iterative approach.

6)Finally connect all the Jobs through Job sequence depends on Dependency because one Datastage Job output may depends upon the Input of other Datstage Job

at 5:02 AM No comments: Posted by Vamsi Krishna

) **-**

Popular Posts

Difference between Normal lookup and Sparse lookup?

When to use which stage in Datastage?

What is dsjob in Datastage?

Difference between sequential file stage and Data set stage?

Datastage scenario with small example

Orchadmin

How to replace character by string in Datastage?

Double convert in Datastage

Difference between Hash and sort grouping methods in Aggregator stage

How To find list of jobs accessing a table in Datastage?

Subscribe To



Blog Archive

June (1)
April (16)
February (3)
January (4)
December (14)

Categories

- Abort (1)
- aggregator (1)
- command line (1)
- convert function (1)
- counter (1)
- dataset (1)
- Datastage-Admin (1)
- Datastage-Scenarios (12)
- Datastage-Stages (17)
- dscc (1)
- dsjob (1)

Recommend this on Google

Labels: Datastage-Scenarios

Sunday, April 14, 2013

How to Use different Configuration file for different Job activities in a Job sequence?

Define a APT_CONFIG_FILE Environmental variable at Job level and mention the path names of configuration files to the jobs in the Job activities in a Job sequence.

Tuesday, January 29, 2013

Double convert in Datastage

Convert Function

Replaces every instance of specified characters in a string with substitute characters.

Syntax-Convert (list, new.list, string)

The two lists of characters correspond. The first character of new.list replaces all instances of the first character of list, the second replaces the second, and so on. If the two lists do not contain the same number of characters:

1)Any characters in list with no corresponding characters in new.list are deleted from the result.

2)Any surplus characters in new.list are ignored.

3)We can not replace a string with string by using convert function we have to use parallel Ereplace c++ routine

Example-How to separate alphabets and numeric values?

Input col Abc123456 Pqrs698

output col_1,col_2 Abc,123456 Pgrs,698

 $\label{lowert} Convert (\convert(\con$

 $Convert(Convert('0123456789','',Upcase(Col)),'',Col) = col_2$

Here I am using second point as a reference

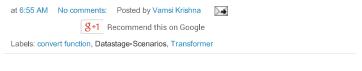
- dsxfiles (3)
- Job modified date (1)
- Job type (1)
- Lookup (1)
- Orchadmin (1)
- Other (3)
- sequence (1)
- Server Job (1)
- Server Routine (2)
- Transformer (2)
- Unix (2)

Search This Blog



Total Pageviews

36,255



Sunday, January 27, 2013

How to find the list of sequences that are restart able in a Datastage project?

2)Identify restartability of Job by using dsjob -jobinfo roject> <job>command[Job Restartable].If the Job Restartable=0 means then that sequence is non-restartable.

3)The step 2 should be repeated for all the Jobs in a project[Use Unix scripting or Loop activity in Datastage]



Counter Using Transformer

In a parallel job a unique counter is generated by using special parallel macros.

- 1. Create a stage variable for the counter, eg. SVCounter.
- 2. At the Stage Properties form set the Initial Value of the Stage Variable to

"@PARTITIONNUM - @NUMPARTITIONS + 1".

- 3.Set the derivation of the stage variable to "svCounter + @NUMPARTITIONS".
- 4.Create a column(counter) in the output of the Transformer and set the derivation of the column(counter) to svCounter(stage variable)

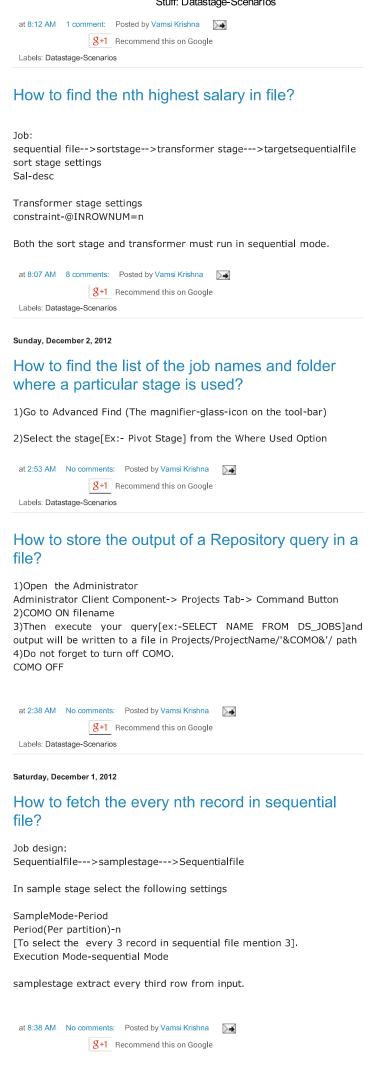
For 4 node configuration each instance will start at a different number, eg. -3, -2,-1,0 When the counter is incremented each instance is increment by the number of partitions, eg. 4. This gives us a sequence in instance 1 of 1, 5, 9, 13... Instance 2 is 2, 6, 10, 14... etc.In this process the counter will always start from 1.This method only works if the input data is evenly balanced i.e. equal number of rows going through each partition.

In 8.5 version developer can generate the unique counter(Surrogate Key) by using Surrogate Key tab in the transformer.User need to mention database sequence name and database crendentials where database sequence name is created.

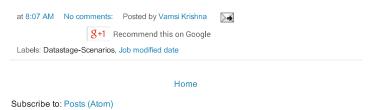
Wednesday, December 5, 2012

How to replace character by string in Datastage?

- 1)convert function replaces a single character with another single character .It is not possible to replace single character by string(group of characters).
- 2)Datastage does not have any inbuilt functionto implement this functionality. User has to write parallel routine which can called from transformer. The code for this routine can be found in the below link PXEReplace



Stuff: Datastage-Scenarios Labels: Datastage-Scenarios List of the stages that used in a jobs in a project 1)Get the list of Jobs in a project Using following command dsjob [options] -ljobs projectname 2)For each Job get the list of stages in a Job dsjob [options] -Istages projectname 3)Use above commands and loop concepts in unix shell scripting to achieve the required output at 8:22 AM No comments: Posted by Vamsi Krishna 8+1 Recommend this on Google Labels: Datastage-Scenarios How to find job modified date in 8.5 Method-1 Export the job as a DSX and look at the DateModified immediately lines below the BEGIN DSJOB and Identifier lines ex-**BEGIN DSJOB** Identifier "JobNameGoesHere" DateModified "2010-10-24" TimeModified "10.22.00" Method-2 2)a)Go to Advanced Find Option b)change view-->List c)Type Jobname in Name to find Option d)Click Find Option Datastage will show Job Type, Job modified date and who modified the Job etc.Method 2 is possible if the Datastage version is 8x,in such cases user can Use method 2



Simple template. Powered by Blogger.