File Masking Run-through (Fixed File Masking)

The objectives of this run-through or taster is to allow you to try:

i) fixed-width file masking : the width or length of column(s) or field(s) in the files to be masked are defined

ii) on-the-fly masking: On-The-Fly (OTF) masking work with a source environment/connector and a target environment/connector, it is an 'Extract Transform Load' (etl) process. Please note that you can also do fixed file masking via in-place masking.

iii) using regex to identify files to be masked : allows you to filter for files to be masked according to certain search criteria

From Wikipedia:

> A **regular expression**, **regex** or **regexp** (sometimes called a rational expression) is a sequence of characters that define a search **pattern**. Usually such **patterns** are used by string searching algorithms for "find" or "find and replace" operations on strings, or for input validation.

iv) Defining a 'Header' record-type: Record types allow you to define different types of 'rows' or 'records'. In this example you will define a 'Header' record type. In some instances of file masking, the first record/row contains the Column/Field Names or titles and therefore should not be masked.

Please refer to the masking documentation at the site below, for more comprehensive information on masking and for masking terminology definitions:

maskingdocs.delphix.com

Please feel free to use your own masking engine:

To access your engine, use :

http://<engine address>/masking/

example:
http://md5350dc4.dc4.delphix.com/masking/

login: admin / Admin-12

- In this practice for on-the-fly masking, we will use the same host for file source and masking target, example:

  - 10.43.3.209   (md11204src.dcenter.delphix.com)

      o.s. logon = oracle/oracle

- source directory: /home/oracle/masking_source

- target directory: /home/oracle/<yourdirectory>

Note: the host/machine for the source, and the host/machine for the masking target, can be different.

Prerequisite actions:

i)   connect to the target host via ssh

ii)   create your target directory. Check and ensure that the masking user to be used , 'oracle' in this example, is able to create and write into files into this directory. For example, create a dummy file in that directory , using touch: touch dummyfile.txt

iii) Note that we haven't copied the source files across to the target directory. The masking process will create the target file(s) and load the mask data into the file(s).

The file that will get masked contains 3 records, and 1 header line which contains the string 'RECORD=3'

2

Note: To check what the data looked like before (in source directory) and after (in target directory) masking:

cat <filename>

```
[oracle@mdora11204:~/file_backups> cat s_gsm_ml_5_8_19.Sqlload.tmp.tozip
RECORD=3
02010211VMI21 00289402033D2E02033D2E27012017095329000240360111000  862993030014410 22201550000251639    335 8898118       14         PPRO39   335 8829995
    1139       TIMG2F0108201800000000000000009802HLRNAT 02033D2E                              002033D2ET  2220155000003615G     0
                                                        0000FAB0000142611                                                                     01F000111718
    N33735 515  5522201 22201 22201 8629930300144110
02010211VMI21 00293402033D2E02033D2E27012017100724000430036011200  862993030014410 22201550000251639    335 8898118       14         PPRO39   335 8829995
    1139       TIMG2F0108201800000000000000009833HLRNAT 02033D2E                              002033D2ET  2220155000003615G     0
                                                        0000FB70000142653                                                                     01F000111718
    N33747 528  5522201 22201 22201 8629930300144110
02012411VMI21 00294402033D2D02033D2D27012017101510004310360122000  356647090011430 22201550000361539    335 8829995       14         PRES
               TIMG2E0108201800000000000000009838HLRNAT 02033D2D39    335 8898660       1139       002033D2DT  2220155000002612G     0
                                                        0000FB90000142457                                                                     01F000111718
    N32575 528  5522201 22201 22201 3566470900114301
```

iv) Ensure that you have downloaded the file format file (in this case TIM_FF1.FF) on to your desktop/laptop.

v) Think up of your own names for the following and note these down or remember these for use later:

source environment name , example: MD_source_env

target environment name , example: MD_target_env

source connector name , example: MD_source_conn

target connector name , example: gMD_target_conn

Let's try Fixed file masking!

1) Login to the Masking Engine: http://md5350dc4.dc4.delphix.com/masking/

   login=admin/Admin-12

2) Define the file format for the file(s) being masked

For this exercise use the file containing the file format = TIM_FF1.FF

This file contains the column name and its length:

```
FILLER1,68
NUMBER,16
IMSI1,15
FILLER2,9
STRING1,14
FILLER3,25
STRING2,14
FILLER4,68
STRING3,14
FILLER5,26
IMSI2,15
FILLER6,18
STRING4,14
IMSI3,15
FILLER7,189
```
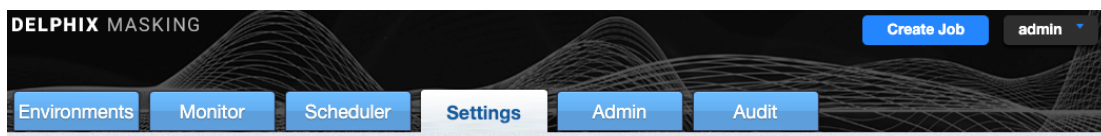
Import File Format

Import Format Type

Fixed Width File

Import Fields

Select...

Note: This is a fixed-width file format, that is, we have defined the width/length of the columns. In this case we are going to create the file format by importing from a file. The name of the file format will get derived from the name of the input file.

DELPHIX MASKING | Create Job | admin

Environments | Monitor | Scheduler | Settings | Admin | Audit

Settings

Click

File Formats

on : Settings tab , then 'File Formats', then 'Import Format'

Algorithms
Domains
Profiler
Roles
Custom Algorithms
File Formats

Import Format

Import File Format

Import Format Type

Fixed Width File

Import Fields

Select...

Click on 'Select' icon and choose TIM_FF1.FF

DRU_DFM_ML.FF   ⊗ Remove

from your local desktop/laptop directory, and Save.

Cancel   Save

3) If you have not defined an Application yet, create an 'Application' tag - under which Environments can be created.

   For example: "HR" or "Personnel" , the title often represents what type of application uses the data that is to be masked

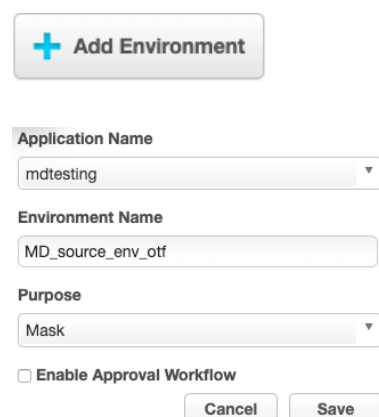Click on the Environments tab, then click on "Add Application".

4) In this run-through you are trying on-the-fly masking, and hence you will need

to define two environments and two connectors:

Source environment/File Connector

Target environment/File Connector

i) Create Environment 1: the source environment, example: MD_source_env_otf
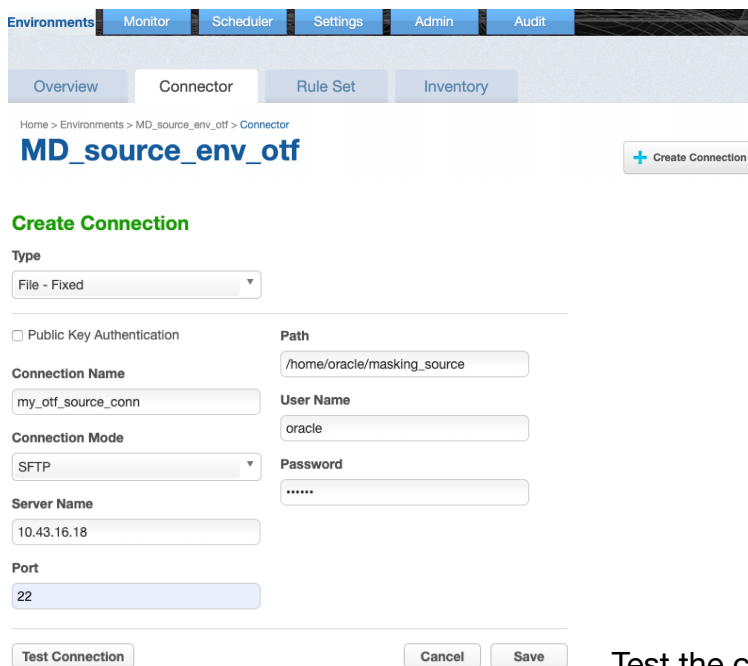
Within this environment define a Connector to the Source files.

After adding the environment, click on 'Environments' tab, select your 'source environment' from the list, then select 'Connector' from the sub-tab, as below:
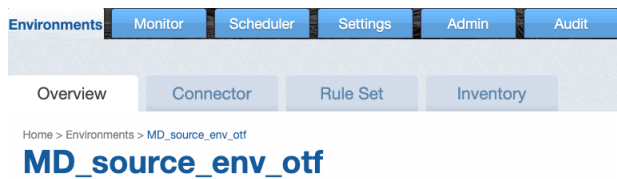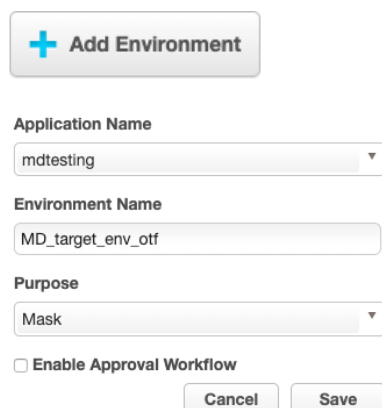

Click on 'Create Connection', per the screenshot below:

Test the connection before saving.

ii) Create Environment 2: the target environment, example: MD_target_env_otf

6

Within this environment define a Connector to the Target host/directory:

After adding the environment, click on 'Environments' tab, select your 'target environment' from the list, then select 'Connector' from the sub-tab.

Click on 'Create Connection', per the screenshot below:
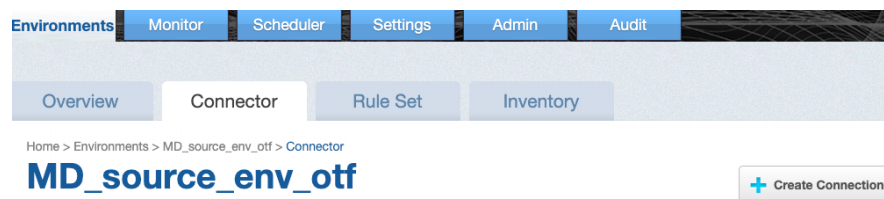
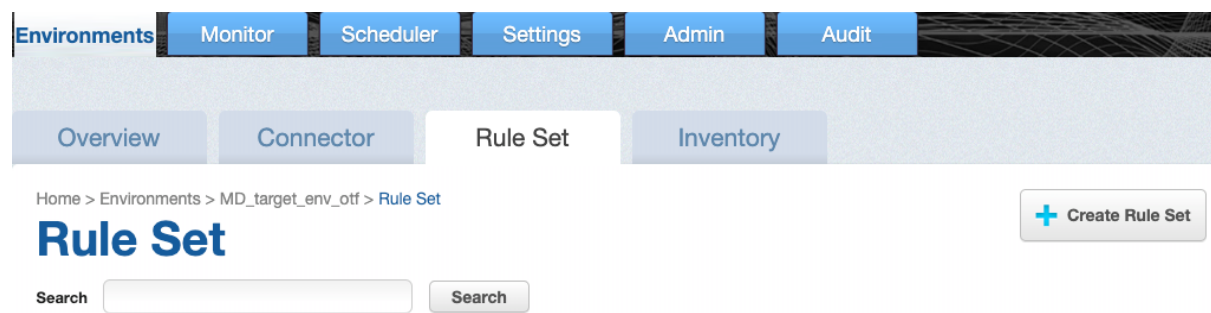then enter connection details:

Test the connection before saving.

5) Create a Ruleset

- this defines which files are to be masked and what file format to use for those files

i)   Click on 'Environments' tab, then select the your target environment. Click on the 'Ruleset' tab, and 'Create Ruleset'.

- Select the Target connector that you created earlier

- Specify that only files with the string 'ml', and 'tozip' in the filename, will be masked. To do this enter the following regex expression under 'File or Pattern' field, as below:

- .+_ml_.+tozip$

ii) Edit the Ruleset   to assign a file format to the Files to be masked. Click on the Green Pencil icon against the ruleset.

then click on the Edit (Green Pencil) Icon next to the 'File'/Pattern you had selected or added:

Rule Set Name md_otf_tgtenv_rs1

File or Pattern	Edit
.+_ml_.+tozip$	✏️

Edit File

Connector

my_otf_target_conn

File or Pattern

.+_ml_.+tozip$

File Format

File Format ▾

End Of Record

End Of Record ▾

Cancel   Save

Assign the file format you created, and specify that

each record is delineated or ended

with a Unix Line Feed

Edit File

Connector

my_otf_target_conn

File or Pattern

.+_ml_.+tozip$

File Format

DRU_DFM_ML.FF ▾
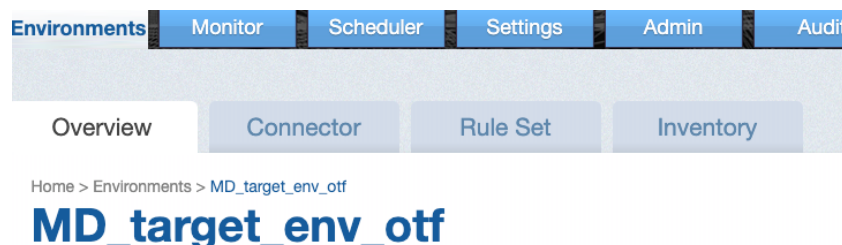
End Of Record

LF terminated (Unix) ▾

Cancel   Save

6)) Inventory

- this defines which files and columns (fields) are to be masked

- defines how they are to be masked

  - can define record types

  - which masking domain/algorithm to be applied

While within the context of your Target Environment, select Inventory

[Environments : Select the Target Environment from the list, click on the Inventory tab]

Environments	Monitor	Scheduler	Settings	Admin	Audit

Overview	Connector	Rule Set	Inventory

Home > Environments > MD_target_env_otf

MD_target_env_otf

As seen under the Inventory, a default record type 'ALL RECORDS' , has been automatically created. Assign a masking domain/Algorithm to the first field. In this exercise use an out-of-the-box domain/algorithm as below:

Edit

Field Name	Record Type
FILLER1	All Records ▼

Domain	Position
ACCOUNT_NO ▼	1

Algorithm	Length
ACCOUNT SL ▼	68

ID Method	Notes
Auto ▼	

Cancel Save

7) Define an on-the-fly masking job to mask the files

[Environments -> select the Target Environment]

Within the context of the Target Environment, click on the [🔲 Mask] Mask Icon.

Environments	Monitor	Scheduler	Settings	Admin	Audit

Overview	Connector	Rule Set	Inventory

Home > Environments > MD_target_env_otf

MD_target_env_otf

🔍 Profile   🔲 Mask

Provide similar details as seen below:

8) Run the masking job by clicking on the icon against the job in the jobs list.

A successful job would return with a 'Succeeded' status as below:

11

9) Return to the target host and cat or view the masked file(s) in the target directory:

```
oracle@mdora11204:~/marisa> cat s_ml_5_8_20.Sqlload.tmp.tozip
45125                                          nullnullnullnullnullnullnullnullnullnullnullnullnull
78452                              862993030014410 22201550000251639   335 8898118        14            PPRO39    335 8829995
     1139        TIMG2F010820180000000000000009802HLRNAT 02033D2E                002033D2ET  222015500003615G     0
                                   0000FAB0000142611                                                              01F000111718
  N33735 515  5522201 22201 22201 8629930300144110
34345                              862993030014410 22201550000251639   335 8898118        14            PPRO39    335 8829995
     1139        TIMG2F010820180000000000000009833HLRNAT 02033D2E                002033D2ET  222015500003615G     0
                                   0000FB70000142653                                                              01F000111718
  N33747 528  5522201 22201 22201 8629930300144110
78452                              356647090011430 22201550000361539   335 8829995        14            PRES
                 TIMG2E010820180000000000000009838HLRNAT 02033D2D39   335 8898660     1139       002033D2DT  222015500002612G     0
                                   0000FB90000142457                                                              01F000111718
  N33575 528  5522201 22201 22201 356647090011430
```

You will notice that the first field/column has been masked. However also notice that the first row/record, which was a Header row with entry 'RECORDS=3', has also been masked.

10) Define a new record type in order that the Header record/row is not masked.

Within the context of the Rule set, click on Inventory, then click on the

'Record Types' icon   [≣] Record Types   then

+ Add a Record Type

Add Record Type

Record Name   Header/Body/Trailer

rec_header   Header   ▾

Number of lines

1|

                                            Cancel   Save

11) Re-run the masking job , check the masked file and notice that the Header row/record has been left unmasked:

```
oracle@mdora11204:~/marisa> cat s_ml_5_8_20.Sqlload.tmp.tozip
RECORD=3
78452                                                        862993030014410 22201550000251639    335 8898118        14         PPRO39   335 8829995
    1139       TIMG2F010820180000000000000009802HLRNAT 02033D2E                           002033D2ET  222015500003615G      0
                                              0000FAB0000142611                                                                    01F000111718
   N33735 515  5522201 22201 22201 8629930300144110
34345                                                        862993030014410 22201550000251639    335 8898118        14         PPRO39   335 8829995
    1139       TIMG2F010820180000000000000009833HLRNAT 02033D2E                           002033D2ET  222015500003615G      0
                                              0000FB70000142653                                                                    01F000111718
   N33747 528  5522201 22201 22201 8629930300144110
78452                                                        356647090011430 22201550000361539    335 8829995        14         PRES
             TIMG2E010820180000000000000009838HLRNAT 02033D2D39    335 8898660        1139        002033D2DT  222015500002612G      0
                                              0000FB90000142457                                                                    01F000111718
   N32E7E E28  EE22201 22201 22201 3E66470000114301
```

This is the end of Fixed-file Masking run-through.

13