

Batch Extractor Archival and Upload Utility

For

File Format Error Rectification

Version 1.9.15

Prepared by

Santosh Kulkarni

System Admin / Automation Enthusiast
WZ Nanded DO Cell 9960708564

Revision History

Date	Version	Description	Author
September 15, 2021	1.9.15		Santosh Kulkarni

Batch Extractor Archival and Upload Utility

Table of Contents

Batch Extractor Archival and Upload Utility	3
1. Introduction.....	4
a. Purpose	4
b. Objectives.....	6
2. How it works.....	6
a. Where it will work	6
b. Utility folder information and view.....	6
c. How Utility will work.	7
3. Utility usage / How to use.	9

1. Introduction

This Utility “Batch Extractor Archival and Upload Utility” is meant to reduce manual tedious tasks which often results delayed/repeated tasks. This tool is made just to serve the purpose. This section of the document describes the following.

- Purpose
- Objectives

a. Purpose

For rectifying the uploaded batches, we need Extract the password encrypted archival and we have to create new password less archive from the extracted batch. Later we have to upload the password less archive to application server in order to start IFF correction for the batch. It’s meant to perform 3-step process as mentioned below.

1. **Extracting Password Protected Batch (Step 1)**
2. **Creating Password less Archive (Step 2)**
3. **Uploading Password less Archive to Application Server (Step3)**

We will see above steps in details in order to get the purpose of the utility.

1. **Extracting Password Protected Batch: (Step1) Possible difficulty’s**

This is the 1st and most tedious task in the process and involves most possible failure in the process. The given DVD/CD as have been encrypted with password and those passwords are written in registers manually. That’s where our first failure cause in action. We will understand this with the following scenario.

Scenario 1: Incorrect password problem

Let’s say we have received a Batch 902-95h -0301 with given password “adv5da” There is possibility that password will not work. Cause can be since its written by hand mistakes and typo can happen the writer was writing u but it looks like v it’s also possible, he was writing s and its look like 5 there are N number of possible failure due to typo. It’s not possible to write all possible typos.

It’s quite difficult to check all possible password from command line manually. The given sample password is “adv5da” is small password some actual password can be big in size and can be difficult to deal with. Writing some actual passwords are as below.

fdr\$4c%fgh8n
ijd#jh5gt2%h
jgid%42\$jhk2
kdi#hg21%jh1
oijd\$g12#gh5
rdf*1#fty2bh
ytj\$4gt%1gyk
mgf\$hj52%jk1j

Above problem is handled in utility we will see this ahead

Scenario 2: No password received

This leads to strait rejection of the Batch. **This problem is handled in this utility we will check this ahead in the document**

2. Creating Password less Archive: (step 2) Possible difficulty's

The automation rule of IT says “**Anything we do twice or more has to be automated**” Just to increase efficiency and minimize manual mistakes we have to try automation for this. Let’s say the above batch we received “902-95h - 0301” is extracted successfully but while archiving the admin by mistake did typo and created zip file with name “902-95h-0301.zip” or “902-95H-0301.zip” or “902-95H-301.zip” or “902-95h -301.zip” it’s quite possible that this file will not be considered to IFF correction due to incorrect naming convention despite successful extraction, We cannot correct errors for this batch or admin has to create new zip file with correct naming convention in this case.

This kind of problems are well handled in this utility since this utility will create password less archive based on batch name which was extracted in step1. It will keep intact spaces in folder name as shown in above

Note: Step 3 Uploading Password less Archive to Application Server will be discussed later

b. Objectives

The objective is clearer in this case. Handle possible problem and find possible solutions. Since success of this activity is quite depend on system admins on site. But this is not the only work he have to complete. He has to parallelly monitor services/ Servers / Upload new Batches / Answer mails / Address any customer queries. It's quite possible that he will not dig deep if repeated problem arises. I have covered max possible problem (as per my understanding) with their solutions. We will understand this in details in how it works section. Since most of the tasks are automated there is vary less scope for manual mistake

2. How it works.

This section will give brief understanding of the utility

a. Where it will work

This utility is developed for the use on **Upload VM Machin (RHEL7)** which is used for day-to-day batches upload. This utility will not modify any Upload configuration on Upload VM. Even admin can do both jobs parallelly like Upload batches and using this utility in separate terminal.

b. Utility folder information and view

User need to copy this utility folder to upload VM in any preferred path. Utility Folder will look like below.

```
6988 root@p902vl ~ # ll (../Development/IFF_Extraction_Archival_and_Upload_utility)
total 32
drwxr-xr-x 2 root root  25 Jul 28 13:41 bin/
drwxr-xr-x 2 root root  34 Aug 26 12:05 conf/
drwxr-xr-x 3 root root  36 Aug 25 12:07 Extraction_Archival_Logs/
-rwxr-xr-x 1 root root 25270 Aug 26 12:23 IFF_Extraction_Archival_and_Upload.sh*
drwxr-xr-x 2 root root  78 Aug 25 12:13 Password_Inventory_Files/
drwxr-xr-x 2 root root  89 Jul 16 13:40 Password_Register_DIR/
drwxr-xr-x 4 root root 4096 Aug 26 12:17 Shell_TEMP_DIR/
```

```
qlmxl-xl-x 4 root root 4096 Aug 26 12:17 Shell_TEMP_DIR/
qlmxl-xl-x 5 root root 80 Jul 16 13:40 Password_Register_DIR/
```

Utility folder structure and file information:

bin: some useful files for utility

conf: contains some optional configuration file like `extract_archive.conf` useful for custom extraction and archival locations.

Extraction_Archival_Logs: This is where Extraction logs and passwordless archival(zip) creation logs are saved in batchwise folder in case we want refer it later.

IFF_Extraction_Archival_and_Upload.sh: The shell which starts the utility

Password_Inventory_Files: This is where we can add possible password files in .csv extension

Password_Register_DIR: This is where utility will save information of extracted batches along with used password. In case of password not found it will add entry in password not found file with filename and date stamp.

Shell_TEMP_DIR: A temporary path for utility execution. Contains few caches files.

c. How Utility will work.

Let's say we want to extract batch "AI_902-98J -6089_01" we have received batch with incorrect password or no password but we have possible passwords with the batch no matter if they are 10000 or 100000 in count this utility will check all password and if it finds possible password then it will start extracting and creating password less archival process otherwise it will eject the batch with password not found message and wait for next batch.

The utility is superfast it can find possible password from around 500000 passwords in only 1 or 2 minutes.

How the password inventory file will look like?

Snap of the sample inventory

```
6992 root@p902vl λ ( ../../Developement/IFF Extraction Archival and Upload utility
# wc -l Password_Inventory_Files/*csv
41990 Password_Inventory_Files/New_Password_Inventory_2.csv
1869 Password_Inventory_Files/Old_Password_Inventory_1.csv
43859 total

6993 root@p902vl λ ( ../../Developement/IFF Extraction Archival and Upload utility
# head Password_Inventory_Files/Old_Password_Inventory_1.csv
2k*52kh$
abc*123#
ade%21g*
aj#22*jk
as$2*hy
aw*5op#
bc*97g$
bcf#41h%
bcy*5hg#4
bd$3h*g#

6994 root@p902vl λ ( ../../Developement/IFF Extraction Archival and Upload utility
# head Password_Inventory_Files/New_Password_Inventory_2.csv
AI9301LIC#1
AI9301LIC#10
AI9301LIC#100
AI9301LIC#101
AI9301LIC#102
AI9301LIC#103
AI9301LIC#104
AI9301LIC#105
AI9301LIC#106
AI9301LIC#107
VI030IFIC#J01
VI030IFIC#J02
VI030IFIC#J03
VI030IFIC#J04
VI030IFIC#J05
VI030IFIC#J06
VI030IFIC#J07
VI030IFIC#J08
```

The password inventory files will contain only single password field. The Utility will find password from these files. We can add /remove password as and when required. There is no restriction on password inventory count. The present password count inventory is 43859 as shown in above picture. User only need to add possible password in above inventory files.

What, if possible, passwords not available?

Most of the sites to have upload mail id configured in old EDMS app server which are still at site. We can get old password mail and get password file from there. Passwords were shared on mail. Even if we don't get from there we can combine available password file from all location and use that as most of the scanning partner have same password convention from same vendors so its quite possible that we have common passwords for batches and no harm in testing them. As checking is done via utility so no manual work load involved.

How extraction and password less archival creation is done?

This utility uses stranded binary's to extraction and archival it uses /usr/bin/unzip to extracted password protected archive and /usr/bin/zip to create password less archive which are provided with distribution. Which are used for day-to-day extraction by all system admin.

3. Utility usage / How to use.

Lets use the utility now user need to insert batch CD/DVD in disk drive. And run the utility as shown in picture below.

```
7001 root@p902vl λ ( ../Development/IFF_Extraction_Archival_and_Upload_utility
# ll
total 28
drwxr-xr-x 2 root root    25 Jul 28 13:41 bin/
drwxr-xr-x 2 root root    34 Aug 26 12:05 conf/
-rwxr-xr-x 1 root root 25439 Aug 26 15:46 IFF_Extraction_Archival_and_Upload.sh*
drwxr-xr-x 2 root root    78 Aug 25 12:13 Password_Inventory_Files/

7001 root@p902vl λ ( ../Development/IFF_Extraction_Archival_and_Upload_utility
# sh IFF_Extraction_Archival_and_Upload.sh
```

We have started the utility after inserting batch media in drive let's see how it will extract media.


```

NOW: Step 1 Extraction of Zip file with name AI_902-98J\ -6089_01.zip completed
-----
INFO: Option Create_zip_file is enabled. Creating Password less zip file AI_902-98J -6089_01.zip
-----
Step 2: Creating Password less zip file AI_902-98J\ -6089_01.zip
Archival Status (Password Less) : [ ████████████████████████████████████████ ] 100%
-----
INFO: Password less zip file is stored as /imagedata1/IFF_Persistent_ZIP_Dir>PasswordLess_archive/AI_902-98J -6089_01.z
ip
INFO: You can review zip Creation Logs at below path
Zip Creation log path : /imagedata1/MyCodes/Developement/IFF_Extraction_Archival_and_Upload_utility/Extraction_Archival
_Logs/26-Aug-21_16_02/Create_zip_Log_AI_902-98J-6089_01.zip.log
-----
NOW: Step 2 Archival password less (Zip) with name AI_902-98J\ -6089_01.zip creation completed.
```

As shown in above picture extraction and creating password less archive both steps performed by the utility itself along with extraction and archival logs. Let's explore Extraction and archival paths.

```

# ll /imagedata1/IFF_Extraction_Dir_TmpFS/UNZIP/
total 0
drwxr-xr-x 3 root root 80 Aug 26 16:03 AI_902-98J -6089_01/

6947 root@p902vl λ ( ../Development/IFF_Extraction_Archival_and_Upl
master U:1 ?:6$ | By: Santosh Kulkarni | )
# ll /imagedata1/IFF_Persistent_ZIP_Dir/PasswordLess_archive/
total 358380
-rw-r--r-- 1 root root 366977945 Aug 26 16:04 AI_902-98J -6089_01.zip

-lm-l--l-- j 100f 100f 300011042 vna 5e 1e:04 VI_005-08J -0080_0J's.b
1019f 328380

```

This concludes the utility information.