Addressbook Conversion Guide

Revision History

Version	*Date*	*Comments*	*Author*
0.1	07/23/03	Initial Draft	W. Ahmed
1	07/23/03	Incorportated feedback from Jeff	W. Ahmed

1. Overview

This guide outlines the procedures to follow for converting the current Captaris (CAP) webmail address books to Sun ONE Messaging Personal Address book (PAB) format. The main task of this conversion is done via running the custom perl script called cap2pab.pl. This reads the old address book(s) and list(s) per user and converts to ldif format which is directed to STDOUT.

The minimum requirement of the perl script is a list of active users. This list can either be provided in a file using the "-f" switch (recommended method) or via the "-u user1,user2,..., userN" switch. Note using the "-u" switch automatically implies that no suffix headers (such as o=pab, o=internet,o=pab etc) will be created. The generation of the headers can also be suppressed with the "-n" switch. The "-o" switch points to a log file and is highly recommended. So a typical invocation of the script is as follows.

```
# nohup ./cap2pab.pl -f active-users.txt -o conversion.log > pab.ldif
2> errors &
```

A sample of the log file output is given below. This can be used by helpdesk people and others for troubleshooting.

```
[aataylor] Found 189 address(es) in /webmail/ak/aataylor/ADDRBOOK.WM
[aataylor] 10 list(s) in /pab/webmailak/aataylor/LISTS
[aataylor]      List "water polo" with 16 addresses
[aataylor]      List "north penn 99" with 15 addresses
[aataylor]      List "basic forward" with 15 addresses
[aataylor]      List "s.seagulls" with 4 addresses
[aataylor]      List "DGfamily" with 14 addresses
[aataylor]      List "Delta Gamma 2000" with 67 addresses
[aataylor]      List "Water Polo 2002" with 16 addresses
[aataylor]      List "GW Water Polo 2003" with 11 addresses
```

[aataylor] List "warm fuzzy forward" with 27 addresses
[aataylor] List "Water Polo 2003" with 12 addresses
[aataylor] No /pab/webmailak/aataylor@gwu.edu ... Skipping
[aataylor] Found 183 address(es) in /pab/webmailfs/aataylor/ADDRBOOK.WM
[aataylor] 9 list(s) in /pab/webmailfs/aataylor/LISTS
[aataylor] List "water polo" with 16 addresses
[aataylor] List "north penn 99" with 15 addresses
[aataylor] List "basic forward" with 15 addresses
[aataylor] List "s.seagulls" with 4 addresses
[aataylor] List "DGfamily" with 14 addresses
[aataylor] List "Delta Gamma 2000" with 67 addresses
[aataylor] List "Water Polo 2002" with 16 addresses
[aataylor] List "GW Water Polo 2003" with 11 addresses
[aataylor] List "warm fuzzy forward" with 27 addresses
[aataylor] No /pab/webmailfs/aataylor@gwu.edu ... Skipping
[aataylor] No /pab/webmaillz/aataylor@gwu.edu ... Skipping
[aataylor] Summary: Merged to 206 unique address(es) and 10 unique list(s)

2. Procedure

2.1. List of active users

- Active users are defined as LDAP entries which have the attribute “inetsubscriberstatus” set to “active”.

- Edit the *get-active.sh* shell script to your production environment.
  `# cd <directory-server-root>/shared/bin; ./get-active.sh`
  - Contents of the shell script are as follows.

```
#!/bin/sh
./ldapsearch -p 19389 -h localhost -D "cn=directory manager" -w
dmanager -b "dc=gwu,dc=edu,o=internet" inetsubscriberstatus=active
uid | grep uid | awk '{print $2}' > active-users-unsorted.txt

# /usr/bin/sort active-users-unsorted.txt > active-users.txt
```

2.2. Preparation

- After several attempts to backup/restore the address book it was decided to use smbfs on a Linux machine as the approach to access the Captaris address books on various servers. Several successful tests have been run with this approach.

- Mount the export SMB shares via the following command on the Linux machine. Note the IP adddress might be incorrect.

```
# mount -t smbfs -o username=cking,ip=128.164.127.33,ro //
uncleleo/d$ /webmail/ak-mount
# mount -t smbfs -o username=cking,ip=128.164.127.34,ro
//newman/e$ /webmail/fs-mount
# mount -t smbfs -o username=cking,ip=128.164.127.35,ro
//puddy/e$ /webmail/lz-mount
```

- Create the following symbolic links to the Captaris user directories.

```
# ln -s /webmail/ak-mount/Program Files/InterChg/USERS
/webmail/ak
# ln -s /webmail/fs-mount/Program Files/InterChg/USERS
/webmail/fs
# ln -s /webmail/lz-mount/Program Files/InterChg/USERS
/webmail/lz
```

2.3. Conversion

- Before running the *cap2pab.pl* script make sure that the locale of the shell is set to C (acsii) i.e. export LANG=C. Otherwise Perl seg faults. Add this to your *.profile*.

- To save time, multiple instances of the script can be run. Recommended is 5 simultaneous instances. So split the *active-users.txt* file into 5 equal children.

```
# /usr/bin/split -l 6000 active-users.txt
```

- Edit the *cap2pab.pl* script and change the values (in bold) of the following variables to reflect your environment. The comments preceding each line explain what they are for.

```
# Your production dn to the people tree with o=pab added at the end
my $rrdn = "ou=People,dc=gwu,dc=edu,o=internet,o=pab";
# Mount point of your smbfs shares on the Linux machine.
my $base_dir = "/webmail";
# sub-directories under the above mount point which contain the
addressbooks.  Note this can be a symbolic link.
my @addrbooks = qw ( ak fs lz );
```

- Run the *convert.sh* shell script which is a wrapper. The contents of the shell script as as follows. Note only the first instance is without the “-n” (no headers) switch. Also the xa* files are created by split automatically.

```
#!/bin/sh
nohup /home/wahmed/cap2pab.pl -f xaa -o xaa.log > pab-xaa.ldif 2> xaa-
error &
nohup /home/wahmed/cap2pab.pl -f xab -n -o xab.log > pab-xab.ldif 2>
xab-error &
nohup /home/wahmed/cap2pab.pl -f xac -n -o xac.log > pab-xac.ldif 2>
```

```
xac-error &
nohup /home/wahmed/cap2pab.pl -f xad -n -o xad.log > pab-xad.ldif 2>
xad-error &
nohup /home/wahmed/cap2pab.pl -f xae -n -o xae.log > pab-xae.ldif 2>
xae-error &
```

- All scripts should finish in a cumulative 1 to 1½ hours. To make sure that each script has ended properly, tail -5 each log file and see if there is an ending time stamp there. For example.

  [wahmed@oldenglish wahmed]$ tail -5 xae.log
  [zzzak] 0 list(s) in /webmail/fs/zzzak/LISTS
  [zzzak] 0 list(s) in /webmail/fs/zzzak@gwu.edu/LISTS
  [zzzak] 0 list(s) in /webmail/lz/zzzak@gwu.edu/LISTS
  [zzzak] Summary: Merged to 0 unique address(es) and 0 unique list(s)
  [7/22/2003 18:13:43]:

- Concatenate all the output ldif files to one *pab.ldif* file.

```
#  cat pab-xaa.ldif pab-xab.ldif pab-xac.ldif pab-xad.ldif pab-
xae.ldif > pab.ldif
```

2.4. Loading

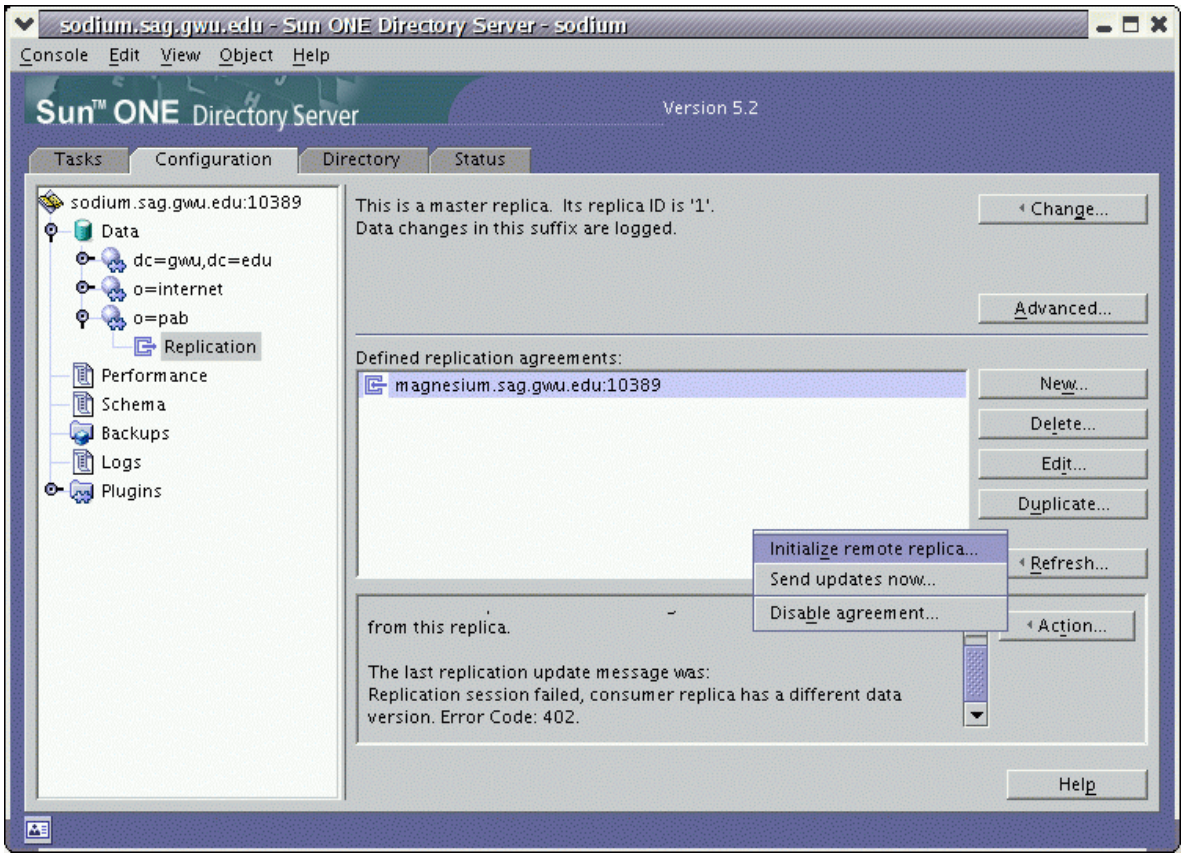
- On one of the PAB masters such as sodium.sag.gwu.edu, import the *pab.ldif* file.

```
# cd /app/ds52/slapd-sodium
# ./stop-slapd
# ./ldif2db -n pabdb -i /tmp/pab.ldif
```

- Once the import is done (which should not take more than 15 minutes), check the error logs of slapd under the logs directory for any discrepancies. There will always be some entries that are skipped or discarded because the had some binary data in the Captaris address book or the files were corrupt. Here is a typical output from a test run. Note that the number of entries for the actual production migration would be lower as the tests were done for all users (disregarding their active/inactive status).

  [23/Jul/2003:11:38:25 -0400] - import pabdb: Import complete. Processed 715890 entries (**1** bad entries were **skipped**, **213** entries were **skipped** because they don't belong to this database) in 745 seconds. (960.93 entries/sec)

- Once satisfied with the load, log into the directory server console and initialize the other master as illustrated by the following screen capture. You'll be warned that the data on the other consumer will be overwritten, which is the intention.


sodium.sag.gwu.edu - Sun ONE Directory Server - sodium
Console Edit View Object Help
Sun™ ONE Directory Server
Version 5.2
Tasks Configuration Directory Status
sodium.sag.gwu.edu:10389
Data
dc=gwu,dc=edu
o=internet
o=pab
Replication
Performance
Schema
Backups
Logs
Plugins
This is a master replica. Its replica ID is '1'.
Data changes in this suffix are logged.
Change...
Advanced...
Defined replication agreements:
magnesium.sag.gwu.edu:10389
New...
Delete...
Edit...
Duplicate...
Initialize remote replica...
Send updates now...
Disable agreement...
Refresh...
Action...
from this replica.
The last replication update message was:
Replication session failed, consumer replica has a different data version. Error Code: 402.
Help