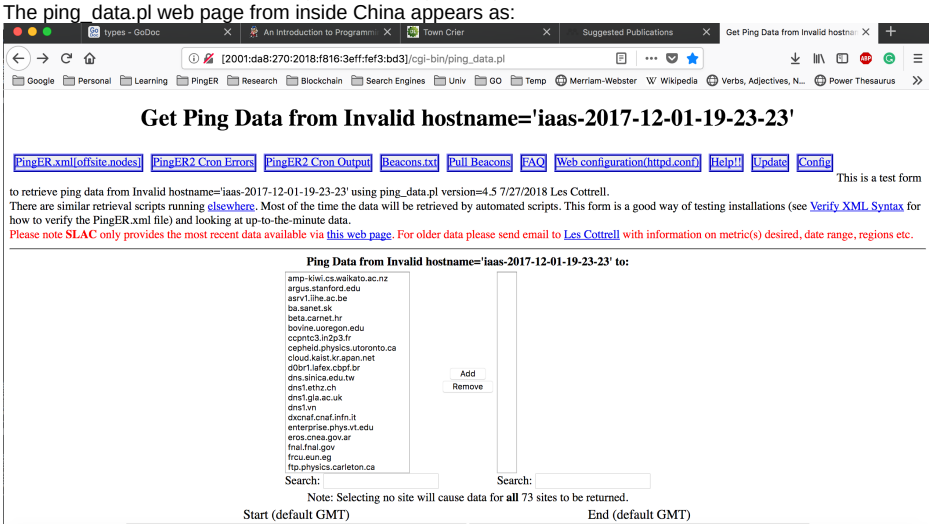
Proxy support for PingER

Table of Contents

Introduction

Originally (prior to June 2018), the gathering of data from the MAs was initiated by the SLAC end.

- However, access to a web server at the Beijing MA run by Guangzhou from outside China was not allowed.
  - The ping_data.pl web page from inside China appears as:



- Also since the MAs from the PingER Android project did not want to have to run a web server, they needed a proxy to copy their data to. This proxy, in turn, has to be accessible from SLAC.

Solution

Since we could not initiate the gathering from the SLAC end, we chose to use the anonymous ftp service at SLAC (see https://www.slac.stanford.edu/comp/unix/ftp.html).

- Be aware data is only kept for 4 days in anonymous incoming FTP and then is removed. The directories are preserved. One can delete files and directories by using rm and rmdir on /afs/slac/public/incoming from the Unix command line (one cannot delete files and directories within ftp).
- The anonymous FTP space is not accessible for reading from outside SLAC. This prevents bad actors from hosting illegal files here.

1. In this case, the MA (e.g. a script initiated at the Beijing MA and run at Beijing by say user Ann who does not have an account at SLAC) daily runs a script (as a cronjob) to use the incoming anonymous FTP at SLAC to store the day's data at SLAC. An example of an FTP session to copy a file is seen at the Proxy FTP script document.

1. At most MAs (with the exception of SLAC), the data is stored in /usr/local/share/pinger/data, e.g.

```
file: /usr/local/share/pinger/data found on dxmon3.cern.ch
total 66756
-rw-r--r-- 1 root root 12244303 Jun  7 18:56 ping-2018-06.txt
-rw-r--r-- 1 root root 56049692 Jun  1 01:55 ping-2018-05.txt
```

2. The data from the MA is copied to a directory associated with the MA and read-write permissions set for group sf. We chose to use the FTP anonymous incoming directory path /afs/slac/public/incoming/pinger/proxy/ and this was set up. The PingER MA then copies the data to the above directory with the PingER MA hostname subdirectory, e.g.

```
169cottrell@pinger:~$ls -l /afs/slac/public/incoming/pinger/proxy/2001:da8:270:2018:f816:3eff:fef3:bd3/
total 23277
-rw-rw-r-- 1 saqibali sf  9360738 Jun  6 20:20 ping-2018-04.txt
-rw-rw-r-- 1 saqibali sf 12394002 Jun  6 20:20 ping-2018-05.txt
-rw-rw-r-- 1 saqibali sf  2079565 Jun  6 20:01 ping-2018-06.txt
```

Since the user Ann does not have a SLAC user account this is done via anonymous ftp to ftp://ftp.slac.stanford.edu/incoming. If the directory does not exist then Ann will need to create the directory for her site, e.g. as seen above 2001:da8:270:2018:f816:3eff:fef3:bd3, i.e. having executed a cd /afs/slac/public/incoming/pinger/proxy/, then execute mkdir 2001:da8:270:2018:f816:3eff:fef3:bd3.

2. Due to security concerns*, we cannot access anonymous incoming FTP space from a web CGI script such as ping_data.pl (that is used by getdata.pl) thus we first move the data from the FTP anonymous space via a local (non-web job) to a space accessible via the web

- This is done by the script proxy.pl which **moves**** the files from /afs/slac.stanford.edu/public/incoming/pinger/proxy/2001:da8:270:2018:f816:3eff:fef3:bd3/ to /nfs/slac/g/net/pinger/pingerdata/hep/data/proxy/2001:da8:270:2018:f816:3eff:fef3:bd3/
  - the node name (in this case 2001:da8:270:2018:f816:3eff:fef3:bd3) is derived from NODEDETAILS.
  - we use a move since one cannot replace an existing file via the anonymous FTP server (one gets a message 'Overwrite permission denied'). Also it ensures we do not fill up the FTP data space.
- Then the getdata.pl script runs nightly at SLAC and calls ping_data.pl via wget to gather the data for all the active MAs (including the proxies). Getdata.pl saves the gathered data in:

```
/nfs/slac/g/net/pinger/pingerdata/hep/data/<host>/ping-<YYYY>-<MM>.txt
/nfs/slac/g/net/pinger/pingerdata/hep/data/2001:da8:270:2018:f816:3eff:fef3:bd3/ping-2018-06.txt
```

If this is done manually the SLAC account will need permissions to create a directory in /nfs/slac/g/net/pinger/pingerdata/hep/data/. This directory has permissions for user pinger and group iepm

```
176cottrell@pinger:~$ls -ld /nfs/slac/g/net/pinger/pingerdata/hep/data
drwxrwsr-x 220 pinger iepm 32768 May 16 00:45 /nfs/slac/g/net/pinger/pingerdata/hep/data/
```

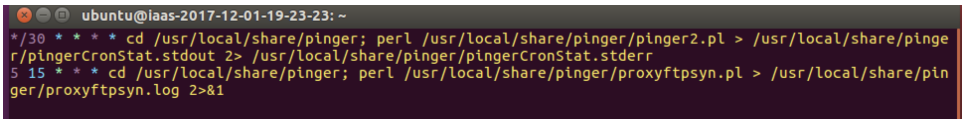
- To add the user to the iepm group

```
178cottrell@pinger:~$ypgroup adduser -group iepm -u saqibali
Jun  7 10:30:33 2018 User(s) 'saqibali' added to group 'iepm'
Jun  7 10:30:33 2018 User(s) must logout/login for this to take effect.
Jun  7 10:30:33 2018 Starting NIS post actions.
Jun  7 10:30:33 2018 [/usr/ccs/bin/make group]
copied group to /afs/slac/service/admin/NIS
updated group
pushed group
179cottrell@pinger:~$ypgroup exam -group iepm
Group 'iepm':
        GID:     2087
        Comment:
        Last modified at Jun  7 10:30:32 2018 by cottrell
        Owners:  cottrell
        Members: cottrell, iepm, pinger, saqibali, ytl
        This is a secondary group.
```

Automating

The steps are automated with synchronized cronjobs.

- A job at Beijing is required to do the anonymous ftp of the recent data from the Beijing MA to the incoming FTP server at SLAC. It runs at 15:05 Beijing time or 12:05am California summer time and 1:05am California standard (winter) time.



- The data then needs to be copied (by proxy.pl which takes a couple of seconds to run) from the anonymous FTP incoming space (/afs/slac/public/incoming/) to a directory accessible (see above) for reading from the ping_data.pl web CGI script that is called by wget from getdata.pl.
- getdata.pl is then scheduled to run at SLAC to copy the selected data from the accessible directory /nfs/slac/g/net/pinger/pingerdata/hep/data/proxy/2001:da8:270:2018:f816:3eff:fef3:bd3/

- We want to standardize the time of the various cronjobs. 1:05 am SLAC summer time is 15:05 in China, and 12:05 midnight SLAC standard (winter) localtime is also 15:05 in China. Thus to catch the Chinese data both winter and summertime California, we schedule the proxy.pl cron jobs at SLAC to run at just 01:20 am.

The various jobs have to be synchronized:

- The copying of data to the anonymous FTP server and moving from there to the PingER raw data archive needs to complete beforegetdata.pl starts at 01:32 local time each night at SLAC
- The data is copied from the MA to anonymous FTP incoming space at then proxy.pl, needs to be scheduled to copy the data from anonymous FTP directory to the directory accessible by the ping_data.pl CGI script.
  - This also has to complete before getdata.pl is scheduled,
  - proxy.pl takes < 5 seconds to execute.
  - proxy.pl is therefore currently scheduled to run at 01:20amCalifornialocaltime
- Once the move is completed by proxy.pl then getdata.pl can be scheduled to gather and save the selected data from the MAs in the PingER raw data archive at:
  `/nfs/slac/g/net/pinger/pingerdata/hep/data/<host>/ping-<YYYY>-<MM>-<DD>.txt.gz`
  - This (getdata.pl) takes about 15 minutes.
  - getdata.pl is scheduled to run at 01:32 am local time.
- The analysis of the hourly data by analyze-all.pl needs to start after getdata.pl has completed. Currentlyanalyze-all.pl starts as a cron job at 55 minutes past 2 am local time at SLAC each morning. The scheduling of the jobs at SLAC is shown below:

```
20 1 * * * /afs/slac/package/pinger/proxy.pl
32 1 * * * /afs/slac/package/pinger/getdata.pl > /afs/slac/g/www/www-iepm/pinger/slaconly/getdata.err
55 2 * * * /afs/slac/package/pinger/analysis/analyze-all.pl --date 1days #Takes 25:1410/20/2011(55mins 9/21
/2011, 70minutes 5/11/2018)
```

The analyzed data from analyze-all.pl is saved in files of the form below, the contents are described in PingER data flow at SLAC:

```
/nfs/slac/g/net/pinger/pingerreports/hep/<metric>/<metric>-<len>-<by>-<year>-<month>-<day>.txt.
gz#len=100|1000, by=by-node|by-site.
/nfs/slac/g/net/pinger/pingerreports/hep/minimum_rtt/minimum_rtt-100-by-node-2011-05-01.txt.gz
```

- See https://confluence.slac.stanford.edu/display/IEPM/PingER+data+flow+at+SLAC for the data flow. The use of the proxy is triggered by the NODEDETAIL Data Server entry (e.g. proxy=1) to tell ping_data.pl to gather the data from the SLAC anonymous ftp server (via the copy)that is basically acting as a proxy.

This whole mechanism is also interesting since it could be extended and a step to providing support for the Android PingER project at Amity in Delhi, India which also needs a proxy.

Support for ping_data.pl for debugging etc.

The CGI script ping_data.pl provides remote access to PingER information at the MA. Besides providing access to the measured data, this also provides validation and debugging information (e.g. logging information from execution of pinger2.pl, the pinger.xml configuration file, the version of pinger2.pl running, the pinger.pl environment at the MA, etc). However, it is not possible to access the Beijing website (or a PingER Android) from outside. Thus we need a solution to provide access to the various information that ping_data.pl provides for the MA. This includes (where for example:

1. The pinger.xml file (e.g. http://2001:da8:270:2018:f816:3eff:fef3:bd3/cgi-bin/ping_data.pl?admin_func=pinger_xml) - item 5 below
2. Error log from pinger2.pl (e.g. http://2001:da8:270:2018:f816:3eff:fef3:bd3/cgi-bin/ping_data.pl?admin_func=std_cron_err) - missing below
3. Progress log from pinger2.pl (e.g. http://2001:da8:270:2018:f816:3eff:fef3:bd3/cgi-bin/ping_data.pl?admin_func=std_cron_out) - item 6 below
4. Beacon.txt (e.g. http://2001:da8:270:2018:f816:3eff:fef3:bd3/cgi-bin/ping_data.pl?admin_func=beacons) - item 1 below
5. FAQ (e.g. http://www-iepm.slac.stanford.edu/pinger/faq.html) - available locally at SLAC
6. Web server configuration (e.g. http://2001:da8:270:2018:f816:3eff:fef3:bd3/cgi-bin/ping_data.pl?admin_func=httpd) - missing below
7. Help for the current version of pinger2.pl being run at MA (e.g. http://2001:da8:270:2018:f816:3eff:fef3:bd3?help=1) - item 4 below
8. PingER configuration (e.g. http://2001:da8:270:2018:f816:3eff:fef3:bd3/cgi-bin/ping_data.pl?admin_func=fileLog) - missing below
9. Plus access to the actual data for a specified time range.

We address items 1 thru 8 by Beijing copying the files to anonymous incoming FTP at SLAC on a daily basis at the same time the actual data is copied. Thus the files (updated as of the previous night) would be available for review. Since Bejing has a copy of ping_data.pl is used to access the files so they can be copied.

- The FTP files are located in: /afs/slac/public/incoming/pinger/proxy/2001:da8:270:2018:f816:3eff:fef3:bd3/web
- The NFS files are located in: /nfs/slac/g/net/pinger/pingerdata/hep/data/proxy/2001:da8:270:2018:f816:3eff:fef3:bd3/web/

```
[cottrell@pinger ~]$ ls -l /nfs/slac/g/net/pinger/pingerdata/hep/data/proxy/2001:da8:270:2018:f816:3eff:
fef3:bd3/web
total 288
-rw-rw-r-- 1 cottrell iepm 16844 Sep 21 01:05 beacons.txt            #item 4 above
-rw-rw-r-- 1 cottrell iepm 15101 Sep 21 01:05 cgi-lib.pl             #missing above, not needed
-rw-rw-r-- 1 cottrell iepm 51968 Sep 21 01:05 ping_data.pl           #missing above, not needed
```

```
-rw-rw-r-- 1 cottrell iepm 40625 Sep 21 01:05 pinger.xml             #item 1 above
-rw-rw-r-- 1 cottrell iepm 63779 Sep 21 01:05 pinger2.pl             #item 7 above
-rw-rw-r-- 1 cottrell iepm   924 Sep 21 01:05 pingerCronStat.stdout   #item 3 above
```

Extension to Android PingER

The Androids do not have a web server so using ping_data.pl to remotely look at the information in the MNA is not going to work. Thus we cannit use ping_data.pl in the process to copy the relevant information to the proxy, so some new code would be needed to copy the files to the proxy. To simplify this we note below the locations of various files in the table below.

- NFS web/directory = /nfs/slac/g/net/pinger/pingerdata/hep/data/proxy/2001:da8:270:2018:f816:3eff:fef3:bd3/web/
- The ping_data.pl URl is prefixed by http://2001:da8:270:2018:f816:3eff:fef3:bd3/cgi-bin/ping_data.pl?

Purpose		Primary location at MA	Alternative location	ping_data.pl URL	NFS web/location at SLAC
List of beacons		/usr/local/share/pinger/beacons.txt		admin_func=beacons	beacons
PingER configuration		Generated by ping_data.pl on demand		admin_func=fileLog	fileLog
httpd		/etc/httpd/conf/httpd.conf		admin_func=httpd	httpd
ping_data.pl		/usr/local/share/pinger/ping_data.pl		N/A	ping_data.pl
pinger2.pl		/usr/local/share/pinger/pinger2.pl		N/A	pinger2.pl
pinger.xml		/usr/local/share/pinger/pinger.xml		admin_func=pinger.xml	pinger_xml
pinger2.pl error log		/usr/local/share/pinger/pingerCronStat.stderr		admin_func=std_cron_err	std_cron_err
pinger2.pl execution log		/usr/local/share/pinger/pingerCronStat.stdout		admin_func=std_cron_out	std_cron_out

---

'* If an anonymous ftp server allows anonymous upload, and also world read access, it quickly becomes a haven for stolen software distribution

'** It moves all the files (except ./ and ../) from for example /afs/slac/public/incoming/pinger/proxy/2001:da8:270:2018:f816:3eff:fef3:bd3/. Hence after a successful move the files will no longer be in /afs/slac/public/incoming/pinger/proxy/2001:da8:270:2018:f816:3eff:fef3:bd3/ but should be in /nfs/slac/g/net/pinger/pingerdata/hep/data/proxy/2001:da8:270:2018:f816:3eff:fef3:bd3/