

20160803 SLAC SEECS and UNIMAS Meeting

Time & date

Wednesday Aug 3 9:00pm Pacific Standard Time, Thursday Aug 4, 2016 9:00am Pakistan time, Thursday Aug 4 2016 12:00noon Malaysian time, Thursday Aug 4 2016 02:00am Rio Time.

Coordinates of team members:

See: <http://pinger.unimas.my/pinger/contact.php>

Attendees

Invitees:

Hassaan Khaliq?, Muhammad Anas Abrar? (SEECS); Saqib+, Aqsa+ (UAF); Johari+, Adnan Khan (UNIMAS); Badrul, Ridzuan, Ibrahim (UM); (UTM); Adib+, Fatima (UUM); Fizi Jalil (MYREN); Les+, Bebo+, Joao (SLAC).

+ Confirmed attendance

- Responded but Unable to attend:

? Individual emails sent

Actual attendees:

Saqib, Johari. Aqsa (but we could not hear her), Saba, Les, Bebo, SLAC

Administration

- Membership of pinger-my in <https://groups.google.com>

Bebo

Topher is here for a few days. He and Bebo discussed whether it would be better to use native Android tools rather than loading the perl interpreter and using it.

1. Pro: easier for user, no need to load perl interpreter which may require missing skills, especially for a non technical user
2. Con: the android version has to mimic pinger2.pl with all its logs, careful checking and specific format

To a large extent it depend on how we plan to use this. If the phones are just MAs in a fixed location then 1 is easier and probably sufficient. If this is intended to grow into a mobile application for general use then it needs to be 2.

We agreed that Sara Masood at UUM should continue pursuing using the standard pinger2.pl version running under the perl interpreter while Amity will try and put together a native android version.

We also briefly discussed how to get the data from the phone MA to the archive. The current method ping_data.pl requires a public IP address for the phone which may not exist if its is mobile. Getting the MA to put the data to the archive may raise some security issue for the archiver. We did not come to a conclusion

UUM

From Adib:

- Adib asked a student (Jaafrau) to explore "Creative visualization of PingER data, including rich interaction;". He has come up with a good design, yet the development part needs a lot of work. Adib will ask him to upload his interface and share the link.
- Adib has a student from Thailand who is working with Adib at UUM who may be interested in installing a PingER MA in Thailand. Adib will talk next week after exams to see if can get permissions. 8/3/2016: No update yet, as they have a long holiday in Thailand. Will try to contact him.
- Adib has written a case study on the Internet performance in ASEAN countries and its implication on the Socio-Economic Indexes. It has been shared with Dr.Les and Prof. Bebo who have commented. We are looking for where to publish.

UNIMAS

pinger.unimas.my is working again as of ~ June 18th, also the traceroute server is working.

Student will be back in september and will start the ISO then.

Johari has given up on the RPi2. He now has an RPi3 with WiFi and Bluetooth which he will be testing in the coming weeks.

The postgrad who was working on anomaly detection quit. Johari will pick up what is left and submit a paper.

UAF (Saqib)

There are 4 students: Aqsa Hameed, Saba Muzamil, Tahseen and Sara Masood. T

- Aqsa has completed her research work on "visualization on pingER data" and now working on to publish a conference paper. Here are few details of the research work.
 - The query results can be exported as CSV file. I use the CSV file of Query results from Impala to draw Line and Bar charts by using Google API.
 - She has created a Data warehouse on pingER data. First we transform the pingER text files into CSV files. Then i upload these CSV files on HDFS and populate Impala Tables and queries.
 - Line and Bar charts are created on a webpage running or executed by localhost server. it can be updated as the query results varies.
- Aqsa and Saba are working together. Their goal is focusing on visualization of PingER historical data using warehouse. The idea is to develop a warehouse in UAF university and make it publicly available.
- Aqsa has put together an abstract of a conference paper and submitted to "The 3rd IEEE/ACM International Conference on Big Data Science, Engineering and Applications (BDSEA 2016)" (<http://computing.derby.ac.uk/bdcat2016/>). "Applying Big Data Warehousing and Visualization Techniques on pingER Data", Aqsa Hameed, Dr. Saqib Ali, Dr. Les Cottrell and Bebo White, submitted to BDSEA 2016. Authors can access it via: <https://easychair.org/conferences/?conf=bdsea2016>
- Sara Masood - pingER monitoring host on android .
 - She is using perl-android-apk Interpreter to run perl on android. It includes execution of 3 perl scripts:
 - getdata.pl, ping_data.pl, pinger2.pl
 - First step is to install pre-requisite i.e. lynx, perl, XML::Simple, make etc. However, she was having permission denied issue when she executes "make install" command at path usr/local/bin because it's a read only file system.
 - Perl-android-interpreter application was depreciated so install perl package using GNURoot Debian Application
 - "Make install" issue is resolved too it create directories and files required by pinger2 in usr/local/share/pinger
 - But now when i execute pinger2.pl <<http://pinger2.pl/>> (that will ping host and start getting data) using IP address its generating two errors image is attached. IP address is also updated in pinger.xml
 - First error is," cannot exec [DnsV4Cmd.pl](#), file not found", but [DnsV4Cmd.pl](#) is in PingER directory with execute permission.
 - Second error, "SrcIP could not be determined for DNS 203.124.57.137. Without the results produced by PINGER2 is not usable."
 - Les suspects the second error is result of the 1st error. maybe using the perl -d debugger would assist.
 - Step 2 Is to execute ping_data and cgi-bin, [traceroute.pl](#) <[http://traceroute.pl/](#)>
 - Step 3: To execute [pinger2.pl](#) <[http://pinger2.pl/](#)> that will start gathering data
- Saqib will proceed to install Pinger in UAF Pakistan 6/22/2016. There used to be an MA at UAF Les will send the information to Saqib. it is below:

```
NODENAME:      uaf.seecs.edu.pk
IPADDRESS:     121.52.152.22
SITENAME:      seecs.edu.pk
NICKNAME:      PK.UAF.SEECS.EDU.N1
FULLNAME:      University of Agriculture, Faisalabad
LOCATION:        Faisalabad
COUNTRY:       Pakistan
CONTINENT:     South Asia
LATANDLONG:    31.433741 73.070476
PROJECTTYPE:   D
PINGSERVER:    http://uaf.seecs.edu.pk/cgi-bin/traceroute.pl?function=ping
TRACESERVER:   http://uaf.seecs.edu.pk/cgi-bin/traceroute.pl?
DATASERVER:    http://uaf.seecs.edu.pk/cgi-bin/ping_data.pl?
URL:           http://www.uaf.edu.pk
CONTACTS:      Ahsan Raza Sattar ,
COMMENTS:      Add by Cottrell 11/5/2012.
Host not working as a monitor and not pingable. There is another Monitor in faisalabad, so Disabled by
Cottrell 6/26/2014.
```

UTM

Saqib's old supervisor agreed to appoint a master student to take of PingER in UTM. Saqib has emailed 3/9/2016, no progress 4/6/2016, no progress 5/18/2016.

1. Currently, Pinger server is located in the faculty server room. Recently, faculty has closed the Server room. Now we have to move it to the University Central Server Room. For this we need support from Prof Johari and Prof Abdul Hanan.
2. For the M.Sc. student in UTM we have to request again to Prof Hanan and Prof Asri and we need support from Prof Johari.

Given there is a working MYREN host at UTM, we will give up on this.

UM

MYREN

- pingersonar-usm.myren.net.my is down since June 18th, email sent to Fitzi 8/1/2016.

NUST

Asked for update on the following 7/30/2016:

pingerqta.pern.edu.pk (DNS resolves to 121.52.157.157) are asking to change the IP to 121.52.157.148,

- Currently it does not respond to ping pingerqta.pern.edu.pk
 - 271cottrell@pinger:~\$ping pingerqta.pern.edu.pk
ping: unknown host pingerqta.pern.edu.pk
 - [cottrell@pinger ~]\$ nslookup pingerqta.pern.edu.pk

272cottrell@pinger:~\$nslookup pingerqta.pern.edu.pk
Server: 134.79.111.111
Address: 134.79.111.111#53

** server can't find pingerqta.pern.edu.pk: NXDOMAIN
- 274cottrell@pinger:~\$ping 121.52.157.157
PING 121.52.157.157 (121.52.157.157) 56(84) bytes of data.
^C
--- 121.52.157.157 ping statistics ---
6 packets transmitted, 0 received, 100% packet loss, time 5625ms
- <http://121.52.157.148>, responds with the default Apache web page
- http://121.52.157.148/cgi-bin/ping_data.pl? responds (albeit very slowly) with Get Ping Data from sharedw228.b2bhosting.net Also there is data.
 - This is a funny hostname. It comes from the system response to the perl commands:
 - use Net::Domain qw(hostname hostfqdn hostdomain);
 - my \$hostname = hostfqdn();
- The following needs to reflect the new IP address at 121.52.157.148. It is very easy to change the IP address once the Measurement Agent is up and running and responding, in fact the PingER meta data at SLAC was changed in February 2015. However the DNS still points to:
 - Output from: /bin/cat /etc/hosts # Do not remove the following line, or various programs # that require network functionality will fail. 127.0.0.1 localhost.localdomain localhost 121.52.157.157 pingerqta.pern.edu.pk

Also asked for an update on the following 7/30/2016:

As of 7/30/2016 the following hosts are down and not responsive for >= 1 month. Anas Abrar reports the nodes were contacted many times for identifying their issues. i.e turn on their pinGER Computers but never responded well. Some have already been Disabled, (strikethroughs as proposed at the 6/22/2016 meeting and hearing nothing from Pakistan following the meeting), I will disable the others if I hear nothing.

- aup.seecs.edu.pk^ #Anas Abrar reports: AUP said that they will look into the issue in a while everytime I ask them or give them the reminder. Does not ping
- cae.seecs.edu.pk^ #Does not ping
- maggie1.seecs.edu.pk^ # Ping responds with Time to live exceeded
- maggie2.seecs.edu.pk^ #Does not ping
- namal.seecs.edu.pk^ #Does not ping
- ~~npm.neduct.edu.pk^ # Not pingable I no longer gather data from it so state unknown 7/30/2016,~~
- pinger.comsats.edu.pl^ #Does not ping
- pinger.iba-suk.edu.pk(30)^ # Pings, ping_data.pl gives page not found from Sukkar Institute of Business Administration
- ~~pinger.isra.edu.pk^ # Not pingable, No longer gather data from it so state is unknown 7/30/2016~~
- ~~pinger.uettaxila.edu.pk^ #Anas Abrar reports: UET taxila are not receiving the provided Number, Unknown host. No longer gather data from it so state is unknown 7/30/2016~~
- ~~pingerisl-qau.pern.edu.pk^ #Anas Abrar reports: number is not responding for QUA. Not pingable. No longer gather data from it so state is unknown 7/30/2016~~
- ~~pingerkhi-uok.pern.edu.pk^ #Unknown host. No longer gather data from it since state is unknown 7/30/2016~~
- ~~pingerjms.pern.edu.pk^ #Unknown host, No longer gather data from it so state is unknown 7/30/2016~~
- ~~pingerlhr-pu.pern.edu.pk^ #Unknown host. No longer gather data from it so state is unknown~~
- ~~quest.seecs.edu.pk^ #Anas Abrar reports: Quest are not receiving my Call now. Not pingable. No longer gather data from it so state is unknown 7/30/2016 7/30/2016~~

The following are close to working:

- pinger.isra.edu.pk # http://pinger.isra.edu.pk/cgi-bin/ping_data.pl works but the xml config file contains no beacons. See FAQ at <http://www-iepm.slac.stanford.edu/pinger/faq.html#nodata> on how to fix

The following 8 .pk hosts are still working:

- 121.52.146.180 (pinger.kohat.edu.pk))

- comsatsswl.seecs.edu.pk
- monitor.seecs.edu.pk
- nukhimain.seecs.edu.pk
- pinger.aiou.edu.pk
- pinger.nwfpuet.edu.pk
- pinger-ncp.ncp.edu.pk
- pinger.vu.edu.pk

* This host was noted at the 3/9/2016 meeting as being down for the last 30 days.

+ Not seen in June 2016

^ Not seen in July 2016

Les proposes 7/30/2016):

- Give up (i.e. remove from Monitoring node table by setting Projecttype to D (for Disabled)) on non-responsive hosts not working for a month.
- I would not oppose extending this to all non-responsive, non-working hosts.
- Focus on the close to working hosts

PingER at SLAC

Joao has provided a cronjob to automatically update the PingER hourly data on. He has also provided for a client to get new data via the web. Needs documenting.

Also has a 4 node cluster with Cloudera manager for Hadoop, Next step are to get impala working for queries.

Working on the following hosts to be able to gather data

Host	State	last seen	Status
pinger.unimas.my	email 5/12/2016, fixed 6/16/2016, fixed June 16th	March 2016	Does not ping
pingeronar-utm.myron.net.my	email to Fitzi 5/12/2016, fixed May 18, 2016		Does not ping
pinger.unesp.br	email 6/15/2016, fixed 6/20/2016	Feb 2016	Looked like cronjob not running, cannot ping
pinger.fsksm.utm.my	email 6/15/2016, no response	May 17	Does not ping
www.univ-ouaga.bf	Unreliable, No data since June 20th.	June 2016	Does not ping
www.ihep.ac.cn	Email 8/3/2016, ping_data.pl works but there is no data	July 28th, 2016	ping_data.pl works but there is no data

Next Meeting

Next meeting: Wednesday Aug 24 9:00pm Pacific Standard Time, Thursday Aug 25, 2016 9:00am Pakistan time, Thursday Aug 25 2016 12:00noon Malaysian time, Thursday Aug 25 2016 02:00am Rio Standard Time.

Old Items

Visualization ideas for PingER moved here 5/20/2016

- One way of drilling down is via a map of all sites. Faisal Zahid put something together a couple of years ago. It is at <http://www.slac.stanford.edu/comp/net/wan-mon/viper/pinger-coverage-gmap.html>. It uses Google maps, allows pan and zoom, plus selection of type (measurement agent, beacon, target) of host to display. It also allows one to select a host (Search location), mouse click on the balloon to find out more about the host, from there click on the sites web site, click on the country or region to get the relevant [pingtable.pl](#) information. The Visualize allows one to choose a measurement agent and a country and it draws coloured lines between the sites to represent the metric choosen. This was very useful but parts of it are broken. It would be great to fix or replace. It has some documentation.
- For plotting time series, Grafana looks very promising
- Updating the motion charts, adding newer data etc. would be most useful. See <http://www-iepm.slac.stanford.edu/pinger/pinger-metrics-motion-chart.html> and <http://www-iepm.slac.stanford.edu/pinger/explorer.html>
- Enhance the selection Metrics, Ticks (see [pingtable.pl](#)) to allow multiple metrics (min, avg, max RTT, loss, throughput, , MOS ..) to be plotted as a time series.
- Enhanced selection remote hosts for selected tick, selected metric selected mutiple targets ([pingtable.pl](#) currently only allows one target), e.g. Once you have plot enable fitting the data to timelines (linear, exponential, power series)
- Map pf countries the world colored by metric allow selection of metric and for a particular year. with chosen MA,

Huffington Post moved here 3/8/2016

- article on PingER and African Internet performance, see http://www.huffingtonpost.com/david-tereshchuk/a-giant-leap-in-2016-africa_b_8901556.html
- Anjum said it would be good for Hassaan to attach this to his proposal. Any action?
- Les is working with the SLAC CIO to try and put something together for publicity concerning the life of PingER.

UOA (Saqib) placed here 2/3/2016.

Saqib has a 5 MS students from the Database team

- - Sara Massoud came up with ideas to improve the info put together by Amity
 - Sabah Massoumil working on Linked Open Data (please excuse spelling)
 - Aqsa Hameed working on big data/analysis of PingER data. The initial idea was to set up a big data/big analysis
 - Aqsa Hameed has been working with Anjum to look at a project to create a Hadoop/Cloudera PingER warehouse to enable easier, more powerful access to PingER historical data.
 - Thinking a bit on Aqsa her work appears to be related to the work done by Thiago on PingER warehouse using a cluster /Cloudera/HDFS/Impala earlier this year. In particular see the presentation at NETAPPS2015 (see https://confluence.slac.stanford.edu/download/attachments/123309267/NETAPPS_PRESENTATIONrevLuiza.pptx). Once the paper is published we can also provide her with that (Adib will let us know when this is OK).
 - Les emailed the relevant people to put them in contact with one another. It appears there is a lot of overlap between what Aqsa proposed as a Masters project and what Thiago has already done. However Thiago's system was mainly a proof of concept and not in production. We need to look at the next steps: Internet access, auto-updating of information in near realtime, production service, maintainable etc. With this in place one could really mine the data looking for all kinds of interesting correlations, clustering, event impacts, comparisons etc.
 - Given that such a warehouse is available, then the next step would be to automatically create queries that would produce in near real-time the plots we produce manually for the PingER annual reports.
 - Following this publish the data in RDF to tie in with the Southampton RDF web observatory repository (similar to what Renan did as a proof of concept).
 - Since Thiago's warehouse is only available at SLAC, it may assist to get Aqsa an account at SLAC, alternatively Aqsa will need to set up a repository at here home base.
 - There is some documentation written by Les on the usage of the warehouse at SLAC. It is at: [PingER Data Warehouse using Big Data with Cloudera on Nebula](#).
 - Les will contact Thiago to see if there is other documentation or where to find the programs etc.
 - We need to put together all we know:
 - Paper from Thiago on PingER warehouse presented in Malaysia at NETAPP2015 Thiago M. Da S. Barbosa, Renan F. Souza, Sergio M. S. Da Cruz, Maria L. Campos and R. Les Cottrell. [Applying Data Warehousing and Big Data Techniques to Analyze Internet Performance](#)
 - Document on Warehouse
 - [PingER Data Warehouse using Big Data with Cloudera on Nebula](#)
 - Documents on PingERLOD
 - [Possible follow on to PingERLOD project, PingER#PingERLOD](#)
 - [Linked Open Data Publication Strategies: Application in Networking PPerformance Measurement Data](#), Renan Souza, Les Cottrell, Bebo White, Maria Campos, Marta Mattoso, poster presented at the [BigDataScience - Stanford](#) conference, CA, USA May 27-31, 2014.
 - Reviews/proposals/ from Aqsa/Fatima
 - [Survey on Big Data Indexing strategies](#), Fatima Bintu Adama, Adib Habbal, Suhaidi Hassan, R. Les Cottrell, Bebo White, Ibrahim Abdullahi.
 - Saqib will provide Aqsa's
 - Tehseen is working on missing PingER data.
 - Saqib has submitted a project with title "A Fundamental Active Internet Performance Monitoring Framework for Pakistan Education & Research Network (PERN) in University of Agriculture, Faisalabad"
 - Project is accepted.
 - working on a project to develop a node using Raspberry Pi 2 and IoT to measure the air and soil quality.
 - Have 2 RPis and setting up air quality with sensors from market, idea is to distribute on campus.
 - Currently, students are analyzing the project to develop a problem statement for their research project.
 - For the GeoLocation Saqib should contact Anjum
 - Project on Internet performance has been accepted, but no funding yet.

Others

- Jan 5, 2015 Hassaan reports "I have received revisions on my proposal and these days I am revising my proposal. In the meanwhile, I have also added another student (Anas Abrar) on this project. He is in learning phase and will follow the nodes which are not working. I shall give you an update very soon."
 - Hassaan is very hopeful that if the proposal is accepted then we can easily have a full time RA for the project.

Hassaan has re-submitted the proposal after revisions. He would like to get Anas Abrar more trained on monitoring operation and then will inform us to add him to the mailing list at <http://pinger.unimas.my/pinger/contact.php>.

- Oct 2015. Following the last meeting, Anjum, Hassaan and Les met to discuss the way forward.
 - "Adnan currently is unable to find resources for handling the project. Similarly, there is no progress on hiring of a full time RA by NUST HQ.
 - However, I (Hassaan) checked from HEC about the proposal that I submitted last year. They have informed me that 2 reviewers have asked for revisions while they are waiting for the third review. I am very hopeful about it. If the proposal is accepted then we can easily have a full time RA for the project. I have plans to talk to Dr. Zaidi about hiring an RA on assuming that our proposal will be accepted by HEC. We can then get his salary deducted later from the HEC project. I shall update you very soon in this regard."
 - Hassaan is waiting to hear from HEC about the comments on the proposal.
 - Moreover, he has asked a student to work on the project for the time being. His name is Mian Anas however he will need few weeks to understand the project.
- Thiago completed setting up the PingER data SQL Impala warehouse running on a Nebula/Cloudera cluster using the Hadoop File System (HDFS). Unfortunately it is not currently accessible from outside SLAC. There have been several attempts to provide outside access, but no success yet, we need to engage the subject matter experts. Thiago is now a SLAC associate so he still has an account at SLAC. There was a cyber security alert on the version of java installed with Cloudera. Les has replaced the cloudera version of java which should fix the vulnerability. However the new version has not been tested.

Geolocation

Anjum believes the TULIP Geolocation application can be improved significantly. At least there are few ideas that we can try. For this, either a group of undergraduate students or an active masters student is required. The resultant work can easily be the thesis of masters level. Who is interested?

- Saqib at Faisalabad has an MS student interested to work on Geolocation project. He requests an initial paper on the project. Les has responded to Saqib. He also has some other students. Anjum will contact him. Potential projects/asks include: take over management of PingER monitoring in Pakistan (say 5 monitors/student; case study of how Pakistan's network performance/connectivity has improved over the years especially as function of funding etc; geolocation with variable alpha; indoor geolocation
- Johari will contact Anjum to learn more of the requirements. Update Johari/Adnan
- See <http://www.slac.stanford.edu/comp/net/tulip/>. Basically TULIP uses pings to a target from landmarks at known locations and converts the minimum RTTs to estimate the distances. Then uses the distances with multilateration to estimate the location of the target
- To improve TULIP one needs the right selection of landmarks, i.e. good (working landmarks) at the right locations (not too far from the target), straddling the target, and with a reasonable estimate of the indirectness (directivity or alpha) of the path from the landmark to the target (so we can reasonably accurately estimate the distance). One also needs a reasonable density of landmarks (e.g. number of targets/100,000sq km)
- The landmarks come from PingER and perSONAR sites. We have a reasonable density in the US, Pakistan and Europe. Currently Anjum is getting better than 20km accuracy for Pakistani targets
- As the number of landmarks goes up so does the accuracy, but so does the time to make the measurements (pings).
- One needs to find the optimal density
- Anjum proposes to speed up the measurements using a cluster for parallelization and also proposes to improve the adaptation of alpha based region. He regards the adaptive geolocation and parallelization as MS projects.
- He is also interested in geolocation in small proximity (e.g. indoors), e.g. using cell tower signals. This is a new area of research. It is possible that the port of PingER to an Android could be related to this. This is a PhD project
- Anjum reports he can supervise the students on Geolocation. He will need to know when the students are ready. We can start with a joint meeting involving Les and the students. Later on, Anjum can have the meeting with students every week while Les can join if he has time.

NUST/SEECs Pakistani PingER nodes status

Pink Background indicates host was bad last month, strike through says it is fixed, yellow is a new bad host.

Current status of Pakistani Hosts 7/1/2015

1. airuniversity.seecs.edu.pk	Down	Called (Person Not Responding).
2. comsatscw1.seecs.edu.pk	Down	Called (Link Issue)
ns3.pieas.edu.pk	Pingable	
3. nuisb.seecs.edu.pk	Down	Called (Not Responding)
4. nukhimain.seecs.edu.pk	Down	Called (Will be up within two days)
5. pinger.cemb.edu.pk	Pingable	Called (Need Access)
6. pinger.kohat.edu.pk	Down	Email sent to the concern Person (DNS Entry issue)
7. pinger.lhr.nu.edu.pk	Down	Called (Person Not Responding)
8. pinger.lcwu.edu.pk	Down	Working now?
9. pinger.nca.edu.pk	Down	Called (Will be up within two days)
10. pinger.numl.edu.pk	Pingable	Need Visit
11. pinger.pern.edu.pk	Down	Need Visit
12. pinger.usindh.edu.pk	Down	Called (Person Not Responding)
13. pingerisl-fjwu.pern.edu.pk	Down	Need Visit
pingerisl-qau.pern.edu.pk	Down	
pingerkhi.pern.edu.pk	Down	
pingerlhr.pern.edu.pk	Down	
14. pingerqta.pern.edu.pk	Pingable	Email sent to the concern person (DNS Entry Issue)
15. www.upesh.edu.pk	Pingable	Called (Person not cooperating)
sau.seecs.edu.pk	Down	

Is it time to start paring down the list of PingER monitor hosts in Pakistan, starting with those that have been down for a while and despite your efforts they are not cooperating. One might also look at the coverage by region in Pakistan and try and keep good coverage for all regions.

Traceroute at UTM 5/9/2015

The traceroute problem regarding maximum reachable hops (i.e. 11 hops) may be since the Unix/Linux/OSX traceroute uses UDP to send the requests. The first request is sent to a particular port (33434), with a ttl to tell it how many hops to go to. The ttl starts at 1 is incremented as it tries the next hop, also the port is incremented (up to 33465). It looks like the first few UDP ports are enabled and then they are blocked. The Windows traceroute uses ICMP to send the probes so does not see the problem.

Linked Open Data

Cristiane reports (7/1/2015): "I am trying to automatize the triplification of PingER data on Kettle. For now, part of the transformation is made on Kettle and another is made by a Java code. Although this solution works for a data sample, is important to have the entire process on Kettle because it facilitates to understand, modify and control the triplification process."

Feb 2015

The plan is still the one seen before (see [project proposal](#)), experimenting those alternatives. Right now, they managed to triplify the data according to a new ontology that takes advantage of a combination of a current standard for multidimensional data (called data cube vocabulary) and a revised version of Renan's Moment ontology adaptation. With this we expect to have a better data organization than the previous solution.

They are now preparing a test plan (like a small benchmark) to be used on all alternatives so that we can compare the results accordingly.

Aug 2014

Renan finished the new pingerlod web site. The new thing is that it should be much easier now to modify the info texts. What Renan did was to put the texts into a separate [file](#). The new version has been loaded on the server and some text added to describe how to use the map. However there is a bug that prevents it from executing the map. Renan reports that the bugs should be easy to fix. He has talked to his professor who suggested trying RDF OwlLink, it should have faster responses to queries. Renan will research this. It will probably mean reloading the PingER data so is a lot of work, hopefully this will improve performance. Before the rebuild he will make the fixes and provide a new WAR for us to load on pingerlod.slac.stanford.edu. He is also working on documentation (he has finished the ontology and has a nice interactive tool for visualizing it, since the ontology is the core of the data model of our semantic solution, this will be very helpful for anyone who uses our system, both a developer of the system and a possible user) and his thesis. Bebo pointed out that to get publicity and for people to know about the data, we will need to add pingerlod to lod.org.

Things he will soon do regarding documentation:

1. A task/process flow writing all java classes involved on all those batch jobs;
2. A Javadoc <<http://www.oracle.com/technetwork/java/javase/documentation/index-jsp-135444.html>> which will explain all classes and how they are used.

For the Linked Open Data / RDF which is in pre-alpha days, you can go to <http://pingerlod.slac.stanford.edu>. As can be seen this page is not ready for prime time. However the demos work as long as one carefully elects what to look at:

- Click on Visualizations, there are two choices:
 - Multiple Network Metrics: Click on the image: gives a form, choose from Node pinger.slac.stanford.edu pinging to www.ihep.ac.cn, time parameters yearly, 2006 2012, metrics throughput, Average RTT Packet loss and display format Plot graph, then click on submit. In a few seconds time series graph should come up. Mouse over to see details of values at each x value (year).
 - A mashup of network metrics x university metrics Click on image: gives another form, pinging from pinger.slac.stanford.edu, School metric number of students, time metric years 2006 2012, display format plot graph, click on submit. Longer wait, after about 35 seconds a google map should show up. Click on "Click for help." Area of dots = number of students, darkness of dots = throughput (lighter is better), inscribing circle color gives university type (public, private etc.) Click on circle for information on university etc.
- Renan will be working on providing documentation on the programs, in particular the install guide for the repository and web site etc. This will assist the person who takes this over.

Renan is using [OWLIM](#) as RDF Repository. He is using an evaluation version right now. Renan looked into the price for OWLIM (that excellent RDF Database Management System he told us about). It would cost 1200EUR minimum (~ 1620 USD, according to Google's rate for today) for a one time eternal license. It seems too expensive. No wonder it is so good. Anyhow, he heard about a different free alternative. Just not sure how good it would be for our PingER data. He will try it out and evaluate. He will also get a new evaluation of the free OWLIM lite.

He has also made some modifications on the ontology of the project (under supervision of his professor in Rio) hence he will have to modify the code to load the data accordingly.

Maria and Renan are advancing in some approaches to deal with PingER data, making it easier to be analyzed and integrated. In particular they have been busy studying and evaluating alternatives, analyzing results from the latest benchmarks on NoSQL (including RDF and graph based storages) database management, distributed processing and mediated solutions over relational databases, and also other experiments with multidimensional analyses on Linked Data. The new students involved are now understanding better the scenario and they have been interacting with Renan regularly.

Cristiane has studied the PinGER data and how to cast it into Linked Open Data form. The size of the [PingER hourly data for 1998-Sep 2014 archived via FTP in text form](#) amounts to ~ 5.12GB and this corresponds to 15.66×10^9 (billion) triples. Then using 5 triples for each measurement and using Turtle without compression gives us 685 Gbytes or an inflation factor of ~ 200.

When Christiane made the estimation of PingER triples, she wrote two documents that explain the process but they were in Portuguese. She has written the new versions in English.

- Counting PingER Measurements: https://www.dropbox.com/s/35itp7v6yasy3rb/Counting%20PingER%20Measurements%20_EnglishVersion.docx?dl=0
- PingER LOD Triples: https://www.dropbox.com/s/4oj5jqupwbujja5/PingERLOD%20Triples%20_EnglishVersion.docx?dl=0

Christiane's report is at: [Size Inflation of PingER Data for use in PingER LOD](#)

UM

Moved here 3/4/2015:

Ibrahim has setup distributed hadoop clusters. He has 2TB of disk space. Les has provided information on getting a subset of PingER data by anonymous ftp via <ftp://ftp.slac.stanford.edu/users/cottrell>. It was put there last September. Information on how the data was put together is at <https://confluence.slac.stanford.edu/display/IEPM/Archiving+PingER+data+by+tar+for+retrieval+by+anonymous+ftp>. There is information on formatting etc at <http://www-iepm.slac.stanford.edu/pinger/tools/retrievedata.html> and some on the dataflows at <https://confluence.slac.stanford.edu/display/IEPM/PingER+data+flow+at+SLAC>. Renan at UFRJ has successfully used this data, he has also characterized the data in terms of bytes/metric per year etc.

Ibrahim has started downloading all zip files in the local machines. 6 weeks ago he downloaded 2 GB of Weather data to test his nodes cluster, he wrote a simple Java program (Map, Reduce) to find the Average and it was working fine.

Anjum reported that UM had experienced a TCP syn DOS attack prior to Mar 12th (when an IDS was put in place). It occurred mainly for several days before between the hours on noon- 2pm and 7-7 in the evening (Malaysia time). He suggested looking to see if PingER could spit the effect. Ibrahim, Les and Anjum will look at. Les analyzed the data and sent it to Anjum

NUST

The following is from Samad 2/24/2015.

- [buitms.seecs.edu.pk](#) #We have to disable gathering data from this host because the person still don't want to continue with us as i have tried once again to convince him but the answer is same. Les has disabled from SLAC.
- [nukhimain.seecs.edu.pk](#) # We were unable to gather data since 20th November, 2014 and now the Node is working fine and collecting data as well.
- [pinger.uettaxila.edu.pk](#) #The node is working fine from last two weeks.
- [sau.seecs.edu.pk](#) #This Node is working fine now.
- [pingerjms.pern.edu.pk](#) #This node is working now.
- [pinger.uet.edu.pk](#) # this was also not working from so many days. and now its working fine and collecting data as well.
- [pinger.isra.edu.pk](#) # This node is also working fine now.
- [pingerlhr-pu.pern.edu.pk](#) # This is also working fine now.
- [pinger.kohat.edu.pk](#) # Collecting data now.

- The IP of "pingerqta.pern.edu.pk" has been changed, Les has updated the databas at SLAC with the following

DNS: pingerqta.pern.edu.pk

Old IP: 121.52.157.157

New IP: 121.52.157.148

Follow up from workshop

- Hossein Javedani of UTM is interested in anomalous event detection with PingER data. Information on this is available at <https://confluence.slac.stanford.edu/display/IEPM/Event+Detection>. We have sent him a couple of papers and how to access the PingER data. Hossein and Badrul have been put in contact. Is there an update Badrul?

The Next step in funding is to go for bigger research funding, such as LRGS or eScience. Such proposals must lead to publications in high quality journals. They will need an infrastructure such as the one we are building. We can use the upcoming workshop (1 specific session) to brainstorm and come up with such proposal. We need to do some groundwork before that as well. Johari will take the lead in putting together 1/2 page descriptions of the potential research projects.

1. Need to identify a few key areas of research related to PingER Malaysia Initiative and this can be shared/publicized through the website. These might include using the infrastructure and data for: anomaly detection; correlation of performance across multiple routes; and for GeoLocation. Future projects as Les listed in Confluence here <https://confluence.slac.stanford.edu/display/IEPM/Future+Projects> can also be a good start and also Bebo's suggestion.
2. Need to synchronize and share research proposals so as not to duplicate research works. how to share? Maybe not through the website, or maybe can create a member only section of the website to share sensitive data such as research proposal?

Anjum suggested Saqib, Badrul and Johari put together a paper on user experiences with using the Internet in Malaysia as seen from Malaysian universities. In particular round trip time, losses, jitter, reliability, routing/peering, in particular anomalies, and the impact on VoIP, throughput etc. It would be good to engage someone from MYREN.

Ibrahim

Ibrahim Abaker is planning to work on a topic initially entitled "leveraging pingER big data with a modified pingtable for event-correlation and clustering". Ibrahim has a proposal, see <https://confluence.slac.stanford.edu/download/attachments/17162/leveraging+pingER+big+data+with+a+modified+pingtable+for+event-correlation+and+clustering.docx>. Ibrahim reports 7/15/2014 "I have spent the last few months trying to understand the concept of big data storage and its retrieval as well as the traditional approach of storing RDF data. I have integrated a single hadoop cluster in our cloud. but for this project we need multiple clusters, which I have already discussed with Dr. Badrul and he will provide me with big storage for the experiment." No Update 8/20/2014.

"I have come up with initial proposed solution model. This model consists of several parts. The upper parts of the Figure below shows the data source, in which PingER data will be convert into RDF format. Then the data pre-processor will take care of converting RDF/XML into N-triples serialization formats using N-triples convertor module. This N-triple file of an RDF graph will be as an input and stores the triples in storage as a key value pair using MapReduce jobs"

? Unknown Attachment

Potential projects

[See list of Projects](#)

1) pingER monitoring host on android .

I am using perl-android-apk Interpreter to run pearl on android.

it include execution of 3 pearl scripts

* PingER2.pl

* ping_data.pl

* getdata.pl

2) Resources Required

i need ping_data.pl script file and get data.pl script file .

3) Research Status

Perl-android-interpreter is not executing .pl file mainly because old Perl version (Perl 5.10) which is not comparable with android lollipop 5.0 version, i am compiling Perl .5.22.0 for android ,it may solve issue.

Host	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
121.52.146.180	

```
| timechart span=1d count by application limit=300
| rename comment AS "The search is for subnet VPN-CLIENTS-GW2 and VPN-CLIENTS-GW1 inbound"
```