

20170921 SLAC SEECS and UNIMAS Meeting

Time & date

Thursday September 21st 9pm Pacific time, Friday September 22nd , 2016 9:00am Pakistan time, Friday September 22nd 2017 12:00noon Malaysian and Guangzhou time.

Format

New items and updates are in bold face.

Coordinates of team members:

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

Attendees

Invitees:

Wajahat Hussain+, ajmal.farooq@seecs.edu.pk, Muhammad Anas Abrar (SEECS); Saqib+, Aqsa (UAF); Johari?, Adnan Khan (UNIMAS); Badrul, Ridzuan, Ibrahim (UM); (UTM); Adib?, Fatima (UUM); Fizi Jalil (MYREN); Les+, Bebo?

+ Confirmed attendance

- Responded but Unable to attend:

? Individual emails sent

Actual Attendees

Saqib, Wajahat, Adib, Bebo, Les

Others

Administration

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

UUM

- The paper on Internet Performance and its Reflection on Socioeconomic Development in the ASEAN Countries to IEEE Access.
 - The paper was not accepted :(despite some positive reviews, two reviewers accept the paper and only one reject it. Anyway, I have been working on a more rigorous analysis to resolve the correlation/causation doubt (to answer David comments) and the new analysis is ready. Adib needs a few days to update our paper with the new analysis then will share with us to comment and advise.
 - He has been working with a statistician to identify correlations between Internet performance (IP) and ICT Development Index (IDI), Human Development Index (HDI), Fertility, GDP, and the Happy Planet Index (HPI).
 - They then look at the normality of the data. All except GDP pass the normality test. For GDP they use Spearman's non-parametric correlation, while Pearson Correlation is used for the rest.
 - There are strong positive correlations for IDI and HDI, a strong negative correlation for Fertility rate and a moderate correlation for HPI. GDP also has a moderate correlation
 - It then looks at the cause effect relationship using regression.
 - Finally it comes up with a model of the form:
$$IP = 2724.288 + 0.008GDP + 275.904 IDI - 3238.298 HDI - 220.264 - 8.084HPI$$
 - And they test its predictive behavior using various statistical tests. It looks very good.
 - **Adib is finalizing the paper and hopes to complete by the end of this week 9/25/2017.**
- "BIND: An Indexing Strategy for Big Data Processing" that uses PingER data. Submitted and accepted by the 2017 IEEE Region 10 Conference (TENCON) that takes place in November. In Penang Malaysia
- pinger.uum.edu.my they have a connection problem with UUM gateway. **The IP address had been re-assigned. Adib showed them the paperwork that the IP address belongs to the PingER MA. They apologized and hope to solve next week.**
- **Adib has sent PingER MA invitation to Nigerian person who used to be a student at UUM. He is awaiting a response.**
- **Bebo sent the information on the Web Observatory and the Internet archive to Adib. Bebo also sent an introduction to Adib and Ian. We are awaiting a response from Ian. Bebo will send a reminder message to Ian.**

NUST: Updates August 2017?

One student is working on: Data Mining doing research on PingER data. He will be using Hadoop on a supercomputer at NUST.

- The student is getting to know Hadoop. Once the student is comfortable with the tools we can start applying the tools to pinger data.

We were unable to gather data from:

- aup.seecs.edu.pk not pingable as of March 6th, 2017. Wajhat still trying, not given up. Still down 4/15/2017, 7/3/2017, 8/11/2017, **9/16/2017 (does not ping). Not responding.**
- cae.seecs.edu.pk working again as of 3/6/2017. Down since March 17, 2017. Down 7/3/2017, Down 8/11/2017 did not ping, **9/16/2017 pings but unable to gather any data. Reconfigured 9/19/2017.**
- kinnaird.seecs.edu.pl pings but unable to gather data 8/11/2017, **unable to ping 9/16/2017. Contacted.**
- maggie1.seecs.edu.pk Down 8/11/2017, **9/16/2017 pings unable to gather data. Reconfigured 9/19/2017.**
- maggie2.seecs.edu.pk Down 8/11/2017. **Working 9/16/2017**
- monitor.seecs.edu.pk down since 23 April, down 7/3/2017. **Usually working 9/16/2017.**
- namal.seecs.edu.pk (host is not pingable). Institution unreachable from start, we are trying to locate the person in charge, fingers crossed (Wajahat 1/10/2016). Working 2/1/2018, still working 3/20/2017. Down since April 1st 2017. Down 7/3/2017, 8/11/2017. **9/16/2016 pings unable to gather data. Contact person missing**
- nukimain.seecs.edu.pk down since March 24th, 2017. Down 7/3/2017, 8/11/2017, **9/16/2017 unable to ping**
- pinger.aiou.edu.pk down since April 15th, down 7/3/2017, 8/11/2017. **9/16/2017 unable to ping. Not Interested.**
- pinger-ncp.ncp.edu.pk pings but can't gather data 8/11/2017 and 9/16/2017. **Contacted.**
- pinger.vu.edu.pk down since May 22nd, down 7/3/2017. Problem with missing XML/Simple.pm missing. **Unable to gather data, pings 8/11/2017 and 9/16/2017. Reconfigured 9/19/2017.**

They are focusing on the hosts they can ping. In many cases the researcher has left and others do not see the purpose. Wajahat will keep trying pointing out the value for international research and papers.

UNIMAS (Johari unable to attend 9/21/2017 so no update)

- Johari attended a workshop organized by UNESCO about experience in China and Malaysia on analysis and visualization. He talked to several researchers from China. He will send a paper to a journal to see if there is interest.
- Johari is teaching this semester, he will enter the Raspberry Pi image into Dropbox or Github in the next few days. Bebo will follow up with Github information.
- He is looking at updating the PingER Malaysia web site.
- He has been unable to contact Hafiz to get MyREN monitor at UNIMAS working again.
- Writing a paper. Les will review and then Johari will submit. No progress 3/14/2017. Looking for a good journal 5/18/2017.

UAF/GHZU (Saqib)

The paper title: "Detecting Anomalies from End-to-end Internet Performance Measurements (PingER) using Cluster Based Local Outlier Factor" is submitted in ISPA 2017 (<http://trust.gzhu.edu.cn/conference/ISPA2017/>). **It has been accepted as of 9/17/2017.**

The paper title: "Internet Performance Analysis of South Asian Countries using End-to-end Internet Performance Measurements" is submitted in IEEE IUCC 2017 (<http://trust.gzhu.edu.cn/conference/IUCC2017/>). **Saqib can still update the submission. Therefore, he needs feedback.**

The paper title: "Missing Values Imputation in PingER Internet End-to-end Performance Measurements using k-nearest neighbors (k NN)" was not accepted in IMC 2017. I am updating the paper according to the review's comments. **Hopefully, Saqib will submit it at some other venue.**

The thesis of Aqsa Hameed title "Applying Data Mining and Visualization Techniques on Pinger Data" is published in ODBMs.org and is accessible through <http://www.odbms.org/2017/07/applying-data-mining-and-visualization-techniques-on-pinger-data/>

Pinger MA at GZHU is now accessible publicly and is successfully gathering data and incorporated in PingER. It would make a good beacon, however it was not pingable 7/16/2017, sent email. **Saqib is trying to convince the IT department to open the ping. No success yet. Hopefully in next 2 weeks.**

We now have about 230 educational sites in China that respond to pings. about 200 have been entered into the PingER Meta database. We still need to add about 48 municipality sites

Sara Masood

Currently, no data is available on PingER on Android due to unavailability of the live IP address. No update 4/19/2017, 7/6/2017. Email sent to Sara Masood. **No update yet.**

The 44th Asia-Pacific Advanced Network (APAN) Meeting was held in Dalian, China in collaboration with CERNET from August 26 to September 1, 2017. Saqib attended this meeting. He gave a presentation. He has a lot of interest from CERNET on using PingER for monitoring CERNET. CERNET backbone is currently at 10Gbps will upgrade to 100Gbps in a couple of years. There are about 20 core nodes that run IPv6 and they will be extending to 50 core nodes. Saqib is thinking of submitting a proposal to extend PingER to support IPv6 monitoring of the CERNET backbone. The submission deadline is 30 Sept 2017.

Pinger has valuable historical data for the last 20 years. Many analysis and case studies have been carried using this data. A lot of information is available on the website. Saqib's idea is to publish the brief summary all these analysis through a survey paper covering the history and utilization of PingER data starting from 1998 to 2017. Saqib will start on it next month (August). **There has been correspondence on how to aggregate PingER data to look at losses by S. Asian countries over the last couple decades. See <http://www.iepm.slac.stanford.edu/pinger/prmout/>.**

Pinger at SLAC

Working with Saqib to document (add to PingER meta database) about 230 Chinese hosts and start monitoring them. Completed about 180.

Host	State	last seen	Status
pinger.vu.edu.pk	Unable to gather data missing XML/ Simple.pm , email sent 6/28/2017, no response .	May 22nd	
pinger.uum.edu.my	Unable to gather data since July 17 2017, also does not ping by IP address, by name gives Name does not resolve. Sent Email 8/9 /2017. Name does not resolve, address resolves to is.grid.uum.edu.my	July 2017	Does not ping
pinger.ictp.it	Pings but nothing else. Emails sent June 6, 2017, June 28, 2017, August 9, 2017 .	Mar 9th	Down
www.hep.ae.cn	Pings but nothing else. Email sent 8/9/2017. Fixed 8/10/2017	July 11, 2017	Back up
www.hepgrid.uerj.br	Due to a major issue with the cluster's electrical infrastructure we will have to shut down all machines, in order to preserve our hardware. The cause of the outage is still being investigated. Emails 11/21/2016, 8/9/2017. Given up. Disabled.	Sept 31, 2016	Does not ping

Next Meeting

Next meeting: Thursday October 26th 9pm Pacific time, Friday October 27th , 2016 9:00am Pakistan time, Friday October 27th 2017 12:00noon Malaysian and Guangzhou .time

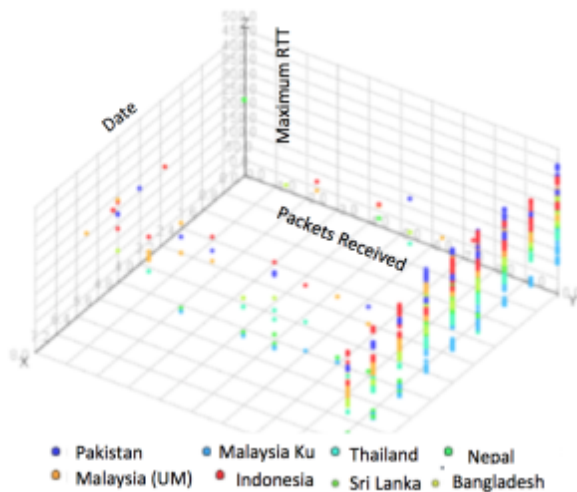
Old information

SEECs (moved here 9/19/2017.)

- Aqsa who was working with Saqib submitted "Applying Big Data Warehousing and Visualization Techniques on pingER Data", Aqsa Hameed, Dr. Saqib Ali, Dr. Les Cottrell and Bebo White, to BDSEA 2016.
- I see it is available from ACM online on the following link: <http://dl.acm.org/citation.cfm?doid=3006299.3006337> for \$15.
- This might be useful to Wajahat's student.

Amity (moved here 9/16/2017)

Preparing a paper on the impact of the cyclone Verdha that hit the Indian coast along with many countries like Thailand, Sri Lanka, Malaysia, Maldives on December 6th. They use K-Means clustering (see https://en.wikipedia.org/wiki/K-means_clustering) to identify anomalies in packets received (inverse of loss) and maximum RTT. Note that for December 7th the reduction in packets received.



Amity (moved here 5/18/2017)

From: Aayush Jain <aayush.2896@gmail.com>

Sent: 24 March 2017 12:31

To: A. Sai Sabitha; harysinha@gmail.com

Subject: PingER Android Team

Abstract for PingER on Android

Progress Made So Far

So far Shivnarayan Rajappa and Rohan Sampson's team have succeeded in making a bare-bones Android Application that can ping beacons, parse data, and generate a text file in a format specified by SLAC ready for uploading. The proposed model involved the application pulling the beacon list from SLAC's servers for pinging. However, the present application has a small percentage of the beacon list hard-coded into the application. As of now, the link between the application and proxy server has not been established.

Future Plans

The new team members are:

1. Rohit Raj
2. Shagun Seth
3. Savy Gupta
4. Aayush Jain
5. Tanuj Saraf

Owing to the advancements in Android technologies during the time of development of the project, our team believes that we can create a more capable and robust application for this project. This involves rewriting the entire application from scratch.

We also propose to create a proxy server that can act as an intermediary between the Android application and SLAC's servers. The proxy server would thus allow handling multiple hosts for greater data collection.

Approach

Our team plans to start off by completing the work on the Android app within 20 days. We will recreate the entire app, with an improved workflow for greater stability. The app will parse the beacon list from SLAC's servers and save as an XML on the device. The data generated after every ping will be appended to a file after cleaning it up with RegEx matches. We first plan to test the app with only a few members of the ping list (which will later be expanded to auto-update in its entirety).

Once we accomplish our work with the app, we will move on to the task of establishing a proxy server. Our entire team will focus on the components of networking, host management, host authentication, file synchronization, and security.

By the end of the project, the server will be able to handle multiple hosts which would all forward it data, and it would in turn reorganise it again for SLAC's servers to pull.

Amity (moved here 4/13/2017)

- The paper on Implementation of PingER on Android has been accepted by IEEE Section. The paper to be online will take 5 months.
- Students are very interested in working with different projects. They have divided the students into three batches (each batch has min of 4 students). The projects currently they are working are:
 - android,
 - data analysis(vardha cyclone)
 - and bigdata

Amity (moved here 3/12/2017)

The students successfully presented the paper on the PingER implementation on [Android.at](#) the confluence 2017 conference.. The paper is submitted to IEEE section.

[Tropical cyclone Vardah](#) hit Chennai in India on the Dec 13th. It impacted the Internet, in particular one of Airtel's undersea cables. Les sent email to A. Sai Sabitha to see if PingER from Amity could see any effect.

- During the next 6 months their research will study the impact Vardha cyclone that hit the Indian coast(South India/Chennai) and a few other neighboring countries in December 2016 as seen by PingER.
 - The idea is to study and analyze the PingER data during the corresponding time frame and deduce significant trends and patterns from the data using
 - 1. Clustering techniques
 - 2. Time series
 - 3. Correlation and Regression concepts

Amity - Java approach (A. Sai Sabitha and Shivnarayan Rajappa)

1. They are using the native java tools, they are not running the `pinger2.pl` <http://pinger2.pl> script on android since the native java tools have the following advantages
 - a. easier for user,
 - b. no need for prior installation of any software, e.g. load perl interpreter which may require missing skills, especially for a non technical user
 - c. doesn't need a rooted phone
 - d. only the apk needs to be installed to run
2. They have fixed the final sequence number change by using regex, and pushed these changes to github repository.

3. They have installed apache tomcat in the server and plan to use a java file on the server which would connect to the phones that send the request. This java file will then take the input stream received from the phone and write the output stream to a file that would be stored on the server. We are facing some problems regarding a blocked port that is not allowing the phone to connect to the server we are currently working on resolving the issue.
4. SLAC can then regularly pull these files which would be stored based on the month they are received.
5. The Android students have started writing a paper on " implementation of pinger on android " .
6. Next steps:
 - a. Extend the target list by getting the Beacon list from SLAC. It is at <http://www-iepm.slac.stanford.edu/pinger/pinger.xml> on a regular basis and updating the <BeaconList> section at their site. This was part of [pinger2.pl](#).
 - b. Also they will need a utility to clean out old recorded data (say older than 3 months), since it will be gathered from SLAC (via the proxy) and eventually they may run out memory on the Android.

Discussion

To a large extent it depends on how we plan to use this.

- If the phones are just MAs in a fixed location then simply porting [pinger2.pl](#) is easier and probably sufficient.
- If this is intended to grow into a mobile application for general use then it needs to be the Java implementation.

A next step is to get the data from the phone MA to the archive at SLAC. 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.
