

# 20120104 SLAC SEECS meeting

## Minutes for SLAC-SEECS Meeting January 4th, 2012.

### General

Sadia sent Ghulam new forms to get registered in SLAC directory. SLAC is shutdown until Jan 2nd 2012, so Ghulam had to wait for his SID to be renewed. He will also have to fill out a DoE form that Sadia is sending him and then apply for his computer account. **Ghulam will be sending new form tomorrow to Sadia.**

Anjum is working with Ghulam to try and improve Ghulam's Skype connectivity for calls with SLAC. **Ghulam and Bilal now attend the meeting from Lab to ensure reliable connectivity.**

**Ghulam will send Sadia username and passwords of maggie1.seecs or maggie2.seecs or atleast one of them. Amber and Sadia had a meeting with Joun after this meeting in which he shared the username and passwords of maggi1.seecs.edu.pk.**

Anjum, Umar and Les had a meeting over the holidays to go over future collaboration between NUST & SLAC in particular with respect to a student at SLAC (after Amber & Sadia leave), proposals to HEC etc.

### IPV6 - Anjum and Ghulam (this has been de-prioritized until new database PingER is working)

IPV6 machine is working fine. Ghulam installed pinger2 on it and tried to collect data. But it was unable to resolve the IPV6 address. Seems like we need to do some changes in the whole architecture to make it workable with IPV6. Les looked at pinger2.pl, it verifies the address is IPV4 4 octets. He made a suggestion to use valid\_ip.pl to verify both ipv4 and ipv6 addresses. Ghulam are there other parts of pinger2.pl that need modifying. It already can access ping6. In addition one will need a copy of pinger.xml with IPV6 hosts and their addresses.

A possible project would be to make traceroute.pl work on a dual stack IPV6 host (say to traceroute to ipv6.google.com). Will need to look at gethostbyname etc. Is there any interest?

Another is to make pingtable.pl and getdata.pl IPV6 capable, again this could use valid\_ip.pl. Since Ghulam is working on these it would be good for him to add this.

### pinger2.pl

There is a bug in pinger2.pl that results in the <BeaconList> being empty. I believe this is since before loading the new <BeaconList> it clears the old one. I have made a fix that may help. **The new version of just pinger2.p, is at wget<http://www-iepm.slac.stanford.edu/pinger/tools/pinger2.pl>. If you are deploying PingER2 make sure you are deploying the latest version. Someone needs to make a new tar file /afs/slac/g/www/www-iepm/pinger/tools/pinger-2.0.3.tar.gz, Done.**

**joun said he will install pinger2.pl at seeecs**

**Sadia updated the pinger.pern.edu.pk ip and beaon list at seeecs. The problem has to do with low memory issue at seeecs pinger machine. Low memory doesnot allow database updatation and no beacon list is updated as well. From pinger machine, unnecessary files should be deleted. It should be consider that Beacon list is updated automatically but Pinger.xml at seeecs has to be updated manually.**

Joun will install new pinger2.pl at seeecs.

Sadia updated the pinger.pern.edu.pk Ip and beaon list at seeecs. The problem has to do with low memory issue at seeecs pinger machine. Low memory doesnot allow database updatation and no beacon list is updated as well. From pinger machine, unnecessary files should be deleted. It should be consider that Beacon list is updated automatically but Pinger.xml at seeecs has to be updated manually

### ICFA Report - Les and Amber

- Fig 4 was MOS, it should be loss, **Amber has fixed**
- We need maps of 2008 and 2011 showing min RTT, it will require MapQuest, **Amber is working on**
- Normalized throughput had two errors, Les has fixed and re-run the data. Les has redone the SLAC normalized throughput. **Amber has re-done the CERN one.**
- Amber is working on the Pakistan case study, and the moves from satellite to terrestrial links.

### HEC Report - Anjum and Amber

- POP to POP analysis: Analysis from Pern POP QAU Islamabad to main POP nodes of each region.
- MIN RTT findings and traceroute results from POP to NON-POP nodes
- Outliers
- For the future reports you can use September report as the border line for good and bad nodes. For example some metric threshold value would be: Jitter >0.5, Avg RTT >75ms, loss > x is bad.
- Imdad, a PhD student at SEECS will be helping Amber in HEC six monthly report. Amber and Imdad will have a meeting at 11pm Friday 16th December to discuss the report. They did not meet as Imdad had to fly to Riyadh for Honet conference. **He will be back on Januray 2nd 2012**

### PingER Explorer - Amber

- Amber has put together an email list of PingER contacts to send the video to. We will send email about Explorer to the list.

## Status of Pakistani PingER hosts - Amber

1. There is a discrepancy between the PERN monitors reported by Joun as not working and those SLAC is able to gather data from (reported in [chec kdata](#), [here](#)). Kashif and Joun are looking at. For example we are unable to gather data from [pinger.pern.edu.pk](#), there is something strange with the [wget](#) not working (see [here](#)). This maybe tied to some improvements to security installed on Dec 14th. **Kashif and Joun are looking at.**
  - a. Currently [pinger.perm.edu.pl](#) is not pingable so we are unable to test further.
  - b. All nodes deployed at PERN PoPs are being upgraded to enhance security. I had discussed with Umar as to what security features should be implemented on these nodes. **If anyone notices a change in working of PoP nodes within next 2 weeks, please highlight it so that we can fix the problem.**
2. In addition some PERN POP monitoring hosts ([pingerisl-fjwu.pern.edu.pk](#) and [pinger.pern.edu.pk](#) ) are only pingable from Pakistan and Jordan. This needs to be resolved. It is probably related to Amber's observation that nodes such as [nukhimain.seecs.edu.pk](#) ; [nuisb.seecs.edu.pk](#) and [pingerisl-fjwu.pern.edu.pk](#) are pingable from SEECS but not SLAC. Using [reflector.pl](#) to ping [nukhimain.seecs.edu.pk](#) and also [www.cern.ch](#), the number of landmarks able to ping nukhiman was 26, while for cern it was 106. It appears only landmarks in Pakistan, Algeria, India, Brazil, and Russia can ping nukhiman. Kashif and Joun are looking at. **Progress**
3. Also [http://pingermtn.pern.edu.pk/cgi-bin/traceroute.pl?function=ping&target=www-wanmon.slac.stanford.edu](#) is not responding, it appears the web server or CGI script may be down. **Progress**
4. Kashif reports we need a system for air university because they have a shortage of systems. **Progress**

As updated on 01/03/2011.

Responsible person: Joun Muhammad

**HEC is sending out letters to the contact persons (who are non-cooperative) after which the nodes will be more reliable. 2-3 weeks will show much more stable nodes.**

Node	Status	Description
<a href="#">pinger.ustb.edu.pk</a>	UP	Pinging by another IP. Data not collected. Issue will be resolved in a week
<a href="#">pinger.giki.edu.pk</a>	Down	On vacations, will be up after vacations.
<a href="#">hu.seecs.edu.pk</a>	Down	Pinging but not fetching data. Trouble shooting in progress.
<a href="#">pinger.uaar.edu.pk</a>	Down	Network issue, will be up soon.
<a href="#">airuniversity.seecs.edu.pk</a>	Down	Network issue, will be up soon.

## PingER Map

There were problems with it not displaying the pull down list of monitoring sites and also not displaying balloons for monitors and beacons. Les contacted Faisal and both have been fixed on Chrome, Safari, and Firefox (however the graphs do not work in Firefox). It does not work on the current version of MSIE.

## PingER Archive Site - Ghulam

Ghulam has rebuilt the database.

- Zafar has raised concerns about using the perfSONAR schema. As a result we propose to extend the data and meta table's columns (to eliminate the need for joins).
  - If we use the SEECS schema we have concerns over the number of rows exceeding 1 million and performance issues. SLAC proposes using monthly shards for the data. **We are unclear how SEECS would address the > 1 million rows. Have any tests been made?**
  - SLAC is unclear whether the data at SEECS is stored for each hour or each 30 minute measurement. If the former then how are the SEQs and RTTs stored (need for out of order and consecutive packet loss probability).
  - perfSONAR sets the interval between groups of pings to be 5 mins with a flat random distribution of 1 minute. The 5 minute setting is configured in Measurement Point Settings.
  - Ghulam will work on [getdata.pl](#) to modify the scripts to use the database rather than flat files and to use parallelization for speed (**do we need parallelization, since today [getdata.pl](#) at SLAC takes about only an hour and we thought most of the time goes on the wget so it should not get much worse using the database as opposed to flat files**).
  - If we do need parallelization then we will need to use parallel loops rather than parallel threads. Ghulam with the help of Zafar is removing threads and adding parallel loops. This would help in fixing the sorting issues of [pingtable-db.pl](#). **Ghulam is working on it.**
- 
- [Pingtable.pl](#) using database aggregates data on fly to show.
  - Question to be put was how long [pingtable.pl](#) takes to display the results for 12 months. Ghulam said that they did a testing few months back. Then [pingtable.pl](#) for 100 pakistani nodes to pakistani nodes took less then a minute to display result. It was suggested to do a testing proof for that.
  - Dr Les is inclined to the idea of PerfSonar database schema as it is well structured and a lot of support is present. It would not go worse than what is deployed at seeecs.
  - Sir Umer suggested to use the idea of views in database which are logical tables having fields from different tables.
  - Finally it is decided: we should go with PerfSonar schema using shards/views.
  - Brainstorm the answers of the following few of the questions and other possible questions if ghulam or any one finds interesting, kindly do share with all:
  - Following are the task list:
    1. What are typical queries for [pingtable.pl](#)
    2. Which query has highest frequency

3. Do we need to short the tables. e.g if we shard into monthly. it means there would be 1.8M rows for 1300 pairs with 48 pings a day in 30 days.It would be doubled if we use different packet size.
4. Which way is best to shard the data in terms of
  - Time
  - Region
  - Sites
  - Metrics
  - Months etc

5. Put a document having

- New schema
- Queries which are run
- Time queries took to run
- two/three lines of content regarding the query and results
- graph would be a good thing to add up

Sadia should

1. Document the database Schema
2. Data transfer from flat files to new database
3. See PerfSonar scripts to do the aggregation fo data like IPDV etc

### Adding MOS,max RTT and Alpha to pingtable.pl (awaits pingtable-db & getdata-db.pl working first)

- Schema has been updated by Sadia. The new schema had PerfSonar and pingtable required fields in separate tables. To remove joins we are looking to add extra columns to the perfSONAR table for metrics like throughput, MOS, alpha etc.
- Sadia will work on migrating data from flat files to the new databases. **Sadia is working on it.**
- Analysis scripts to add Mean Opinion Score and Alpha, some things need to be correctly configured. It has been deployed at <http://pinger.seecs.edu.pk/cgi-bin/pingtable.pl> for testing.
- Alpha, max RTT and MOS to be implemented at SLAC site. Sadia will be doing this with the help of Zafar.

## TULIP - Sadia and Bilal

Following targets in Europe are not plotted on maps.

Country Name	IP Address
Austria	62.218.39.47
Austria	212.33.36.188
Italy	193.206.84.12
Ukraine	193.29.220.3

For example the first target can be explored [here](#) . This can be compared to a [target which can be plotted on the map](#) .

Bilal looked into it and found that nodes are plotted using other GeoIP and IP tracking tools. Possibly there is some error in Tulip map code because of which these nodes are not plotted on Tulip map.

**He found that the problem is in the load function. He sent email to Faisal to try and understand what is going on. It is possible Sadia may also be able to help. Faisal responded Bilal mentioning that these issues might be the result of changes in the Java format; Bilal is looking into this, he will update through the mailing list.**

### CBG TULIP Integration -- FYP (Bilal)

- Bilal did some stress testing. The landmarks are 331 while the targets will be the ones generated by Sadia. He will compare the results with the 4 month old results with 59 hosts.
- **From the latest results it is apparent that if the landmark is also the target then we can get 0 error. Bilal has modified the tests to filter out such cases. Les has sent him the URL to the Landmarks file so he knows to filter out measurements from a landmark to itself. He will rerun the tests for N. America, Europe, S. Asia, E. Asia and Australia and will send the new results before Sunday.**

### PerfSONAR (Pakistan)

- **PerfSONAR at SEECS:** PerfSONAR throughput and latency nodes are now up and running at SEECS. Hostnames and corresponding IP addresses are:\*\* throughput measurement node: <http://psbw.seecs.edu.pk/> (http:115.186.132.154/toolkit/)
  - latency measurement node: <http://psla.seecs.edu.pk/> (http://115.186.132.155/toolkit/)
- Select options under "Service Graphs" to view throughput or latency graphs. Added 5 Stratum 1 NTP servers to cater for clock delay and everything seems to work fine.
- There are some interesting one-way latency graphs at 115.186.132.155 (SEECS PerfSONAR Latency node). **Dst to Src (e.g MIT to SEECS) latency is less than Src to Dst (e.g SEECS to MIT) latency.** This might uncover some trends in outbound network traffic from Pakistan.
- Bilal and Ghulam will have a meeting with Zafar to know about PerfSonar and to maintain it in future. **Update?**

## Possible projects

- There can be a paper kind of talking on Pinger if we could just find the right conference. MCN, ICC and Globecom do provide network monitoring topics. We can talk of GEO-Location experiences. For example within Pakistan it works fine, however as we go within regions or continents this gets worse. We can publish some stats on that for example. We are yet not ready for Tulip paper.
- See [<https://confluence.slac.stanford.edu/display/IEPM/Future+Projects>].
- Extend the NODEDETAILS data base to allow entry support for whether the host is currently pingable.
- Extend Checkdata to provide emails automatically, see [<https://confluence.slac.stanford.edu/display/IEPM/Extend+checkdata+to+make+it+more+useful>]. Many of the ideas in the script node-contacts.pl are a step in this direction.
- Improve the PingER2 installation procedures to make it more robust. This might be something for the person(s) in Pakistan who are responsible for installing PingER2 at the Pakistani monitoring sites. They probably have found where the failures occurs. Also look at the FAQ, and ping\_data.pl which has been improved to assist in debugging, could it be further improved (e.g. provide access to the httpd.conf file so one can see if it properly configured)? There are 2 students working on the PingER archive. Is this something they could work on?
- [[Fix PingER archiving/analysis package to be IPv6 conformant](#)]|IEPM:Make PingER IPV6 compliant]. Will build a proposal for an IPV6 testbed. They will try various transition techniques. A proposal has been prepared and that has been submitted to PTA. Adnan is a co PI. It is being evaluated today. A small testbed has been established in SEECS and the plan to shift some of the network to IPV6. Bilal is part of 3 students involved with PingER and they will be involved with IPV6. They are porting the PingER archive site to using a database. They have redeveloped the archive site using Umar's documentation. They have set up a small test archive site. They have gathering, archiving, analysis. They will design a new database. They will also try a port of PingER to IPV6.
- Look at RRD event detection based on thresholds and how to extend, maybe adding plateau algorithm. Umar's algorithm did not work in a predictable manner.
- Provide near realtime plots of current pinger data using getdata\_all.pl/wget. It will work as a CGI script with a form to select the host, the ping size, and the time frame to plot. It will use wget or getdata\_all.pl to get the relevant data and possibly RRD/smokeying to display the data.

## Future meeting time - Les

1. Next meeting on Wednesday 11th January, 2012 at 8:00 pm in US and Thursday 12th January, 2012 at 9:00am in Pakistan.