ICSI516 Project 2 – Fall18
Internet path analysis using PlanetLab
(100 points)

Due November 28th at 11:59PM via Blackboard
(note the important PlanetLab account deadline below)

Note: While each PlanetLab slice will have two or three students associated with it, you need to complete this assignment individually.

PlanetLab account: Before you start your work, you need to create a PlanetLab account. ***Create your account within the first week of the assignment (October 31-November 7) using your albany.edu email***. **Late account requests will not be considered** (and you will get 0 points for this assignment).

Objective: PlanetLab (`https://www.planet-lab.org/`) is a global Internet testbed comprised of over 1000 machines distributed across the globe. Most of these machines are hosted by research institutions and all of them run a common software package on top of a Linux-based operating system. The objective of this assignment is to experiment with real, worldwide Internet testbed, to gather first-hand measurements of Internet performance from geographically distributed machines and evaluate path properties in today's Internet. The evaluation you will be performing here is inspired by a paper on Internet path stability by Vern Paxson ([Paxson96]).

Assignment: In this assignment you will implement a simple Internet measurement framework in which a minimum of ***10 PlanetLab nodes*** will perform periodic pairwise path measurements using `ping` (with 20 packets) and `traceroute`. Your measurement data should be collected over the course of at least ***two weeks***. Your measurement nodes should store *ping* and *traceroute* logs locally after each measurement and should move them periodically to your `VM` account. Once all of your measurements are completed you need to analyze your measurements and answer the questions below. Note, that you want to transfer your logs from the PlanetLab nodes to `VM` frequently, because PlanetLab nodes are prone to failures, which might result in permanent loss of your measurement data (and subsequent failure to submit your assignment). Also note that it is a good idea to begin partial analysis of your logs as soon as you have some (say one day). This will allow you to spot potential problems with your measurement framework early (and fix them before the deadline) and will distribute the load of implementing your analysis/evaluation tools. Finally, be sure to allow extra time for analysis before the deadline, in case you need to go back and run additional tests to answer some of the questions below. Please, read all the questions below before designing your measurement framework, since these questions will determine how to write your measurement daemons.

Questions:

1. Present results about your pair-wise path length in the form of a table. In this table include what was the min, max, average and standard deviation over all the measurements for any given path. Did you see any changes in path length over the different measurements? Why?
2. How often did you detect outages in pair-wise paths between your nodes (present this as a fraction of all monitored paths)?
    a. How many of them were temporary (as defined in section 6.7. from [Paxson96])? For each failure, can you determine the likely cause based on your measurements? If yes, what is it? If not, what additional information do you need to answer this question?

   b. How many failures were long-term outages? Can you determine the root-cause for these failures? (remember that PlanetLab nodes go down for repairs often).
3. Of the failures you detected, what percentage are in the core vs. in the edge (local ISP failures)?
4. Did you detect any significant differences in the reliability of continental links (e.g. US<->US, US<->Canada) vs inter-continental links crossing the Atlantic or the Pacific Ocean?
5. Did you detect any route fluttering (as defined in section 6.4 from [Paxson96])? Where and how often?
6. Did you see any inconsistencies (in terms of hop count and geographical location) between the forward (A→B) and reverse (B→A) path in your pair-wise measurements?
7. Did you detect any triangular routing, where the traceroute path traversed a network hop out of the expected general direction of the destination (e.g. a path from US to Germany going through Asia)? Describe these routes if any.
8. Plot your average pair-wise packet loss and latency as a CDF. You need to present two figures here: one containing the latencies of all five pairs (so five curves) and one containing the packet loss of all five pairs (again five curves).
9. Based on your tiny sample of measurements do you think that Internet routing instability has increased or decreased in comparison with the findings in [Paxson96]?

Choosing your PlanetLab nodes: You want to choose PlanetLab machines that are geographically spread. To this end, I suggest you choosing nodes such that you end up with the following subsets:

- At least 2-3 nodes from continents other than N. America
- At least one transatlantic and one transpacific link

Choosing measurement periodicity: You don't want to choose a periodicity that is too short or too long. Too long of a periodicity will not give you a fine-enough information to perform your measurements. Too short of a periodicity might not allow your measurements to complete in cases where the links are very long. For example, a link between New Zealand and Germany will have a long path, which will slow down the completion of traceroute. You need to account for this. I suggest picking a periodicity between 20min and 1hour.

Best practices on using PlanetLab: Remember that you will be sharing an account with one or two other students. To avoid any interference with your work, it is best to follow the practices below:

1. Pick nodes that are different than these of your slice buddy.
2. Don't ever delete nodes that were added to your slice. Adding a node to your slice creates a virtual machine on that node for you to work in. Deleting the node, deletes the virtual machine and all the data that lived in it (i.e. scripts, traces, etc.).
3. If you happen to work on the same node as your slice buddy, make sure to create your own directory and work only in that directory. This will avoid situations in which your files are overwritten by your buddy or you overwrite your buddy's files.

Submission: This assignment is due on November 28th at 11:59PM. **Late submission will not be accepted.** Submit your project as two files via blackboard.

- ***The first file*** should be a write up of your findings from this analysis in ***PDF*** format. Your write up should include the following sections:
   - *Your name and e-mail address.*
   - *Motivation* – why is Internet path stability important to measure?

- *Methodology* – include a description of your measurement framework (how many nodes, how often was probing performed, how was data retrieved, stored and processed, what hardware and software did you use).
- *Analysis*. You can begin this section with a paragraph describing your findings about path length and path length variability over the measurement period (Question 1). You can then divide the remainder of this section in two subsections. The first sub-section should be organized around the topic of path failure (so discussion of your findings while answering Questions 2-6). The second sub-section should cover the topic of path loss and latency (based on your answer to Question 7)
- *Conclusion* – summarize your findings. Answer Question 8 in your conclusion.

- ***The second file*** should be a ***zip archive*** that includes:
  - A text file with all the PlanetLab nodes you used.
  - All the scripts/daemons you used to collect, store, retrieve and analyze measurement data for this assignment.
  - All the data that you collected for this assignment.

Finally, when you are done with your measurements remember to KILL all daemons.

Grading: You will be graded based on the consistency of your measurements, the coherency of your analysis and the presentation of your findings.

Cheating: Cheating is not tolerated. See the syllabus on the class homepage for my cheating policy.