

Projects

Date given: August 31, 2015.

Return date: You must complete and return the complete project (report) on Friday, Oct. 30, 2015. After this date your project will not be accepted.

Title: Measurement of RTT (Round Trip Time) between any pair of active servers with error correction.

Objective

In this project you have to develop a RTT measurement system which will measure the RTT between any pair of server through exchange of files.

Work to do for the project:

1. Create a number of different size text files on your local computer (source computer).
2. Select at least 60 active nodes (as diverse as possible-spread over the entire country) all across the continental United States.
3. Store these files on destination servers that you selected in step 2.
4. Access these files, one at a time, from the source and measure the total time to get the file that you requested. This is RTT. The RTT value depends on two main parameters: scheduling delay at the destination node and network delay. Try to discover a way to measure the effect of these parameters on RTT. You need to come up with a way of identifying delays due to each of these parameters separately.
5. Measure RTT with each pair (source node being common in each pair), a total of 60 pairs.

Resource requirement and preparation needed for the project

1. Linux/Ubuntu installation
2. SCP/SSH
3. MySQL
4. Java/C

PlanetLab registration steps

1. Goto: <https://www.planet-lab.org/db/persons/register.php>
2. Register with your details and university email id, only one per team. Select site as: "University of Missouri Kansas City"
3. Email us stating completion of your registration with your email id
4. Select at least 60 nodes (as diverse as possible-spread over the entire country) all across the continental United States. You will find an inclination towards the concentration of nodes at east coast since it contains most of the PlanetLab nodes.

Steps: Data Collection

1. Goto: <https://www.planet-lab.org/doc/guides/user>
2. Read the documentation to create ssh/scp login key

3. Select at least 60 nodes in your slice all across the continental United States. Try to be as diverse as possible. You will find an inclination towards the concentration of nodes at east coast since it contains most of the planetlab nodes.
4. Create a database with the name IGOD. Create a table IGOD.nodes with the schema “nodename (PK), nodegeolocation”. Populate this database with your nodenames and their geolocation in latitude and longitude.
5. Once all your nodes are accessible (key has been propagated), create files of the size 32 byte, 1 KB, 512 KB, 1 MB, 1 GB and copy it to all your nodes using scp in loop linux shell script
6. Write a server/client socket program to download the files (32 byte, 1 KB, 512 KB, 1 MB, 1 GB) by each node from all other nodes. For each download measure the delay (time taken in download). Write these results in a log file. Use your own log file schema but it should contain “delay, source, target, timestamp, filesize”. Before you record the details of download make sure to tune the network. By tuning we mean download for 3 times without recording the details followed with 4th download but recording results this time. Do this for each file every time. Each node will run both client code and server code in parallel. Server providing files for download to other nodes and client downloading file from other nodes.
7. Write shell scripts for starting, stopping and downloading log from all the nodes. Stop all the nodes, download the log files and then start all the nodes, every 24 hours for 30 days.
8. Once the files are downloaded import them on MySQL. For this create a table “IGOD.downloaddetails” with the schema “id (PK), delay, source, target, timestamp, filesize, distanceinmiles”. You might have to write a program to insert all the data of files into the database.

Steps: Data Analysis

1. Create plots to show the cumulative effect of the measurements (Details will be shared with you in coming classes).
2. Use three different error correction techniques to reduce the error with high precision.
3. Compare these three techniques with each other and predict the best one to be used.
4. The RTT value depends on two main parameters: task scheduling at the destination node and network delay. Try to discover a way to measure the effect of these parameters on RTT. You need to come up with a way of identifying delays due to each of these parameters separately.

Learning Outcome

The objective of this project is to learn to measure and interpret RTT values and apply error correction to identify the actual RTT values. You will also learn to design and implement the database you design and use them to measure RTT.