

Project Description – Cluster Computing using Raspberry Pi

Project Goal – To set up multiple Raspberry Pis in a cluster that can utilize the Pis to share tasks and implement a database.

Project Implementation

Initially, we planned to implement our cluster computing project using ansible. However, Ansible presented too many problems for us and we weren't able to successfully set it up. Because of this, we switched to MPI (Message Passing Interface) and specifically we used MPICH and MPI4PY.

To physically implement our cluster we used 3 Raspberry Pis. We connected each Pi to a power source and to our switch via ethernet. Then we connected our switch to the router.

Components

The first component of the project is downloading MPI, MPICH and MPI4PY on our Raspberry Pi and link all three of our Pis together. This is done by putting each Pi's IP address in a file (titled machinefile) and then SSH'ing into each Pi in order to generate a unique RSA key for each Pi. We put these RSA keys into a file called authorized_keys which is placed on to each Pi.

The second component is to create and run a Python program where the Pis will divide the task and work together. We can run mpiexec and specify the number of Pis we're using along with the name of our Python program and the program will be handled by the Pi cluster. When we run the same program with just one Pi we see that it takes longer than when our cluster handles the task.

Summary

Overall, Ansible proved tougher to get working than we expected. This might have been the most challenging part of the project for us. We found MPI to be a great alternative and it offered less challenges to get installed than Ansible. One problem we did need to overcome with MPI is the fact that the IP addresses of all the Pis must be put in 'machinefile'. This isn't inherently difficult, but what sometimes happens with devices connected to a network is that their IP address will actually change. If the router or a device restarts, the router might change the IP address of that device. This would require us to update 'machinefile' every time. What we did instead was give each Pi a static IP address that won't change.

The one thing we didn't manage to do was set up a database working with our Pis but we were happy we had the Python program as a benchmark of our cluster's capabilities.

Group Members

Amitoz and Michael

Both of us worked to install Ansible and fix the problems we encountered before forgoing Ansible. Amitoz found the new method (MPI) for cluster computing and we both followed the tutorial and installed it on to our Pis for the cluster.

References/Resources

The tutorial we followed Part 1 and 2:

<http://www.tinkernut.com/2014/04/make-cluster-computer/>

<http://www.tinkernut.com/2014/05/make-cluster-computer-part-2/>

(A video of someone following tinkernut's tutorial):

<https://www.youtube.com/watch?v=x3Wlp3z218I>

Reference for diagrams

http://www.mcs.anl.gov/~balaji/permalinks/argonne14_mpi.php