Step 1: Start the cluster

**cd ~/graphlab/scripts/ec2**

**./gl-ec2 -i ~/.ssh/powergraph.pem -k powergraph -s [replace with number of slaves] launch [replace with cluster name]**

You can also add other parameters, just as follows:

**./gl-ec2 -i ~/.ssh/grapheast.pem -k grapheast -r us-east-1 -z us-east-1a -a ami-9eaa1cf6 -t m3.xlarge -s 1 launch cluster\_test**

Step 2: Update GraphLab

Install/update the GraphLab on cluster

**./gl-ec2 -i ~/.ssh/powergraph.pem -k powergraph update [replace with the cluster name]**

Step 3: Log into master and install vim, subversion

**update**

**apt-get install subversion**

Step 4: Configure sshd and reboot

special: in /etc/ssh/sshd\_config, add the following two lines in red.

MaxSessions Specifies the maximum number of open sessions permitted per network connection. The default is 10.

MaxStartups Specifies the maximum number of concurrent unauthenticated connections to the sshd daemon. Additional connections will be dropped until authentication succeeds or the LoginGraceTime expires for a connection. The default is 10.

MaxSessions 100

MaxStartups 100:30:600

**sudo reboot**

Step 5: On master, check out the source code folder (<https://com1333.eecs.utk.edu:8443/svn/source/Codes/DistPathFinding>)

**svn co** [**https://com1333.eecs.utk.edu:8443/svn/source/Codes/DistPathFinding**](https://com1333.eecs.utk.edu:8443/svn/source/Codes/DistPathFinding)

Step 6: On master, create links of distPathFinding.cpp and Cmakelist in graphlab/apps/DistPathFinding/

**cd ~/graphlab/apps/DistPathFinding**

**ln -s ~/DistPathFinding/codes/ distPathFinding.cpp .**

**ln -s ~/DistPathFinding/codes/ Cmakelist .**

Step 7: On master, replace files in graphlab/src/graphlab/engine/

**cp –f ~/DistPathFinding/codes/ engine/\* ~/ graphlab/src/graphlab/engine/**

Step 8: On On master, run make clean in graphlab/release/

**cd ~/ graphlab/release/**

**make clean**

Step 9: On master rerun configure in graphlab/ and make in graphlab/release/apps/DistPathFinding/

**cd ~/ graphlab/**

**./configure**

**cd ~/ graphlab/release/apps/DistPathFinding/**

**make**

Step 10: Don't do step 9 on slave because mpiexec requires exact same binaries. After compile on master use following commands to copy the binaries on master to all slavers.

**bash -x ~/graphlab/scripts/mpirsync**

Step 11: Run distPathFinding using the following commands:

**mpiexec -n 2 --hostfile ~/machines ./distPathFinding --graph ~/DistPathFinding/datasets/Livejournal/ --format tsv --ncpus 4 --max\_degree\_source 1 --saveprefix res --num\_exp 10000 &> log &**