Name: Kevin Strasberg NetID: strasbe1 HW#: 1

Question # 1:

For AWS EC2/S3:

Storage = \$0.05*(10PB/1GB) = \$0.05*(10,485,760) = \$524,288/month
CPUs = \$0.09/(CPU-hour)*8192Cores*24hrs*30days = \$530,841.60/month

Cost/M = \$1,055,129.60

Total = (Storage + CPU) 3years =\$1,055,129.60/M*36M = \$37,984,665.60

For Owning Owning Cluster:
Hardware = \$10,000,000
Hardware/M = \$10,000,000/36months = \$277,777.78/M
Power/M = \$277,777.78 * 0.4/0.45 = \$246,913.58/M
Network/M = \$277,777.78 * 0.15/0.45 = \$92,592.59/M
Admin/M = ((128 servers / 1024 CPU) * 8192CPU) / 100nodes * \$10,000 = \$100,000/M

Cost/M = \$100,000 + \$277,777.78/M + \$246,913.58/M + \$92,592.59/M = \$717,283.95

Total = 36months * \$717,283.95/M = \$25,822,222.20

Owning your own cluster is always cheaper/month and would be cheaper than renting from AWS.

Name: Kevin Strasberg NetID: strasbe1 HW#: 1

Question # 2:

The map function reverses the key and value so that all of the followers of the same user map to the same person

map (userA, userB)
        while(keyValuePairsExist)
                output.collect(userB, userA)

The reduce function adds all the followers of the same person to the same user in a hash bucket then iterates through each person who is followed and outputs their followers with who they follow.
reduce (key, value)

        while(keyValuePairsExist)
                hashTable[key].add(value)

        foreach key2 in hashTable
                foreach value2 in hashTable[key2]
                        output(key2, value2)

Name: Kevin Strasberg NetID: strasbe1 HW#: 1

Question # 3:

The computer with only a single core is going to cause a major bottleneck. It doesn't say how each machine is assigned a task, just that each has a maximum of 4 tasks. So if the company nievly evenly distributes the tasks among all the computers, the single core machine is going to take significantly longer to complete than the other machines. Also the single core machine only has 2GB of RAM, so if a task requires a lot of RAM it won't be able to complete it without swapping out the memory to its harddrive continuously.

I would suggest assigning the 7 machines with quadcores to have a maximum of 4 tasks and a configure the single core machine to be able to only run one task. Also If possible I would upgrade the single core machine to a quadcore.

Name: Kevin Strasberg NetID: strasbe1 HW#: 1

Question # 5:

The accuracy formula is given by +/-(RTT/2-min)

min = delay it takes server to send packet + delay it takes client to recieve packet
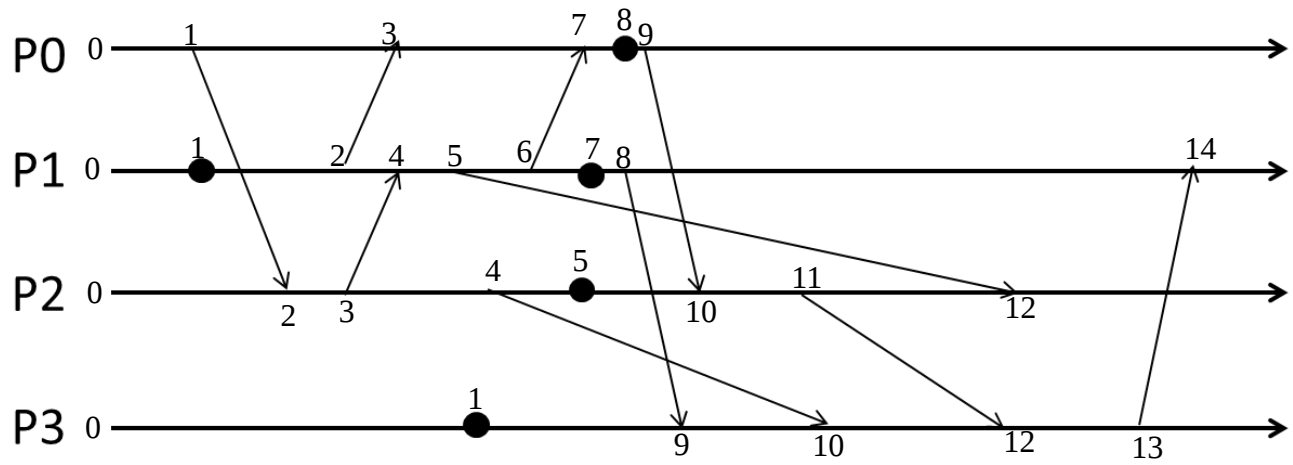min = 0.20 ms + 1.4 ms = 1.6 ms
RTT = 6.6

accuracy = +/-(3.3 ms – 1.6ms) = +/- 1.7 ms

Name: Kevin Strasberg NetID: strasbe1 HW#: 1

Question # 6:

Question # 7:

P0 <0,0,0,0><1,0,0,0>　　<2,2,0,0>　　<3,5,2,0>　<4,5,2,0>
　　　　　　　　　　　　　　　　　　　　　<5,5,2,0>

P1 <0,0,0,0><0,1,0,0>　<1,3,2,0>　<1,6,2,0>　<1,5,2,0>　　　　　　　<5,7,6,5>
　　　　　　　<0,2,0,0>　<1,4,2,0>　<1,7,2,0>

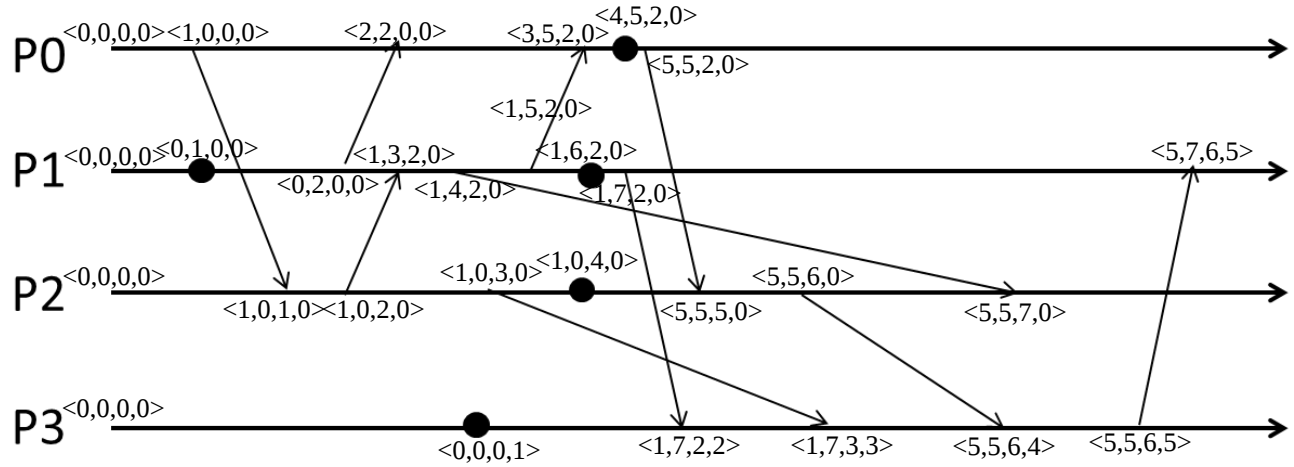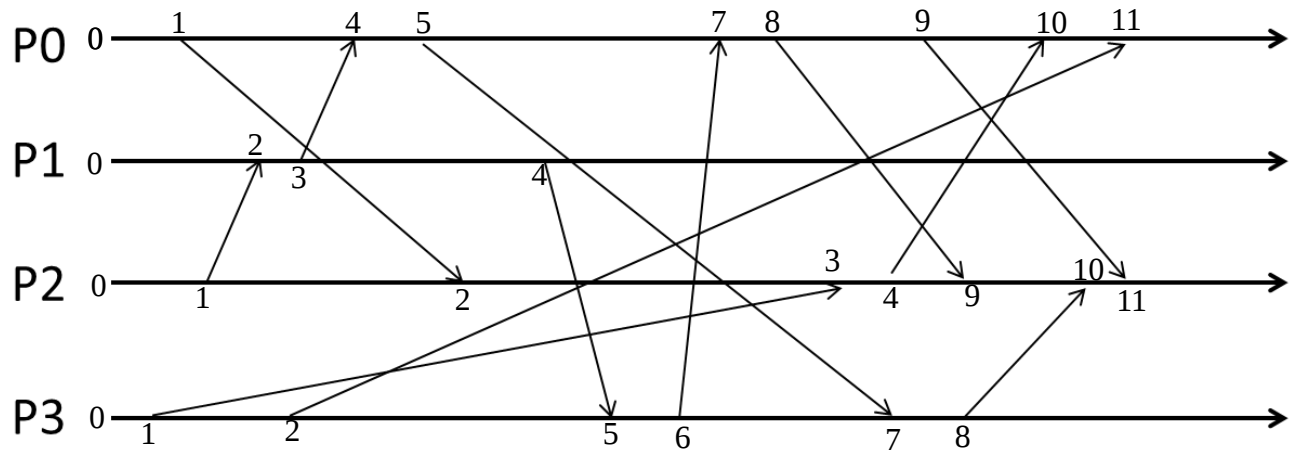**P2** <0,0,0,0>　　　　　<1,0,3,0><1,0,4,0>　　<5,5,6,0>
　　　　<1,0,1,0><1,0,2,0>　　　　<5,5,5,0>　　<5,5,7,0>

P3 <0,0,0,0>
　　　　　<0,0,0,1>　　<1,7,2,2>　<1,7,3,3>　<5,5,6,4>　<5,5,6,5>

Name: Kevin Strasberg NetID: strasbe1 HW#: 1

Question # 8:

Question # 9: