CS553 Project:
Understanding the Cost of Computing in the Cloud

Instructions:

- ***Due date: 11:59PM on Tuesday, 11/25/14***
- ***Maximum Points: 100%***
- *You should work in teams of 3 for this assignment.*
- *Please post your questions to the Piazza forum.*
- *Please answer the questions below in an essay/report style; if you use any external sources (Internet, other books than the required textbook, other people, etc), please cite your work.*
- *Only a softcopy submission is required; it must be submitted to "Digital Drop Box" on Blackboard.*
- *For all programming assignments, please submit just the softcopy; please zip all files (report, source code, compilation scripts, and documentation) and submit it to BB.*
- *Name your file as this rule: "PROJ#_LASTNAME1_LASTNAME2_LASTNAME3.{zip|tar|pdf}". E.g. "Proj_Raicu_Li_Wang.tar".*
- *Late submission will be penalized at 10% per day (beyond the 7-day late pass).*

1. Introduction

You are hired by a startup company who is considering to use cloud computing instead of building its own infrastructure. There is concensus that a cloud computing software stack at the layer of IaaS will be used, but its not clear whether the computing resources should be rented from a public cloud on-demand, or whether a private cloud should be purchased. You are tasked to find the cost breakdown of a private cloud, and compare that to what Amazon would charge for the following instance types: t2.small, m3.large, c3.8xlarge, g2.2xlarge, r3.4xlarge, i2.8xlarge, and hs1.8xlarge. These instance types are defined at http://aws.amazon.com/ec2/instance-types/, and their prices are set at http://aws.amazon.com/ec2/pricing/. For pricing purposes, please stick to Linux on-demand pricing.

Since the instance types (see Figure 1) use a concept of EC2 Compute Unit to denote the compute capacity of the instance, its worthwhile to define it (http://aws.amazon.com/ec2/faqs/#What_is_an_EC2_Compute_Unit_and_why_did_you_introduce_it):

"The amount of CPU that is allocated to a particular instance is expressed in terms of these EC2 Compute Units. One EC2 Compute Unit provides the equivalent CPU capacity of a 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor."

Since you have to estimate the cost of the hardware when building a private cloud, you can use hardware prices found at Newegg.com (http://www.newegg.com/) as a good source for low cost and large variety of hardware. If you cannot find some particular hardware here, please cite whatever site you find where you obtained the pricing information.

Table 1: Instance Types Matrix

Instance Type	vCPU	Memory (GiB)	Storage (GB)	Networking Performance	Physical Processor	Clock Speed (GHz)	Intel® AES-NI	Intel® AVX†	Intel® Turbo	EBS OPT	Enhanced Networking	Price
t2.small	1	2	EBS Only	Low to Moderate	Intel Xeon family	2.5	Yes	Yes	Yes	-	-	$0.026
m3.large	2	7.5	1 x 32 SSD	Moderate	Intel Xeon E5-2670 v2*	2.5	Yes	Yes	Yes	-	-	$0.140
m3.2xlarge	8	30	2 x 80 SSD	High	Intel Xeon E5-2670 v2*	2.5	Yes	Yes	Yes	Yes	-	$0.560
c3.8xlarge	32	60	2 x 320 SSD	10 Gigabit	Intel Xeon E5-2680 v2	2.8	Yes	Yes	Yes	-	Yes	$1.680
g2.2xlarge	8	15	1 x 60 SSD	High	Intel Xeon E5-2670	2.6	Yes	-	-	Yes	-	$0.650
r3.4xlarge	16	122	1 x 320 SSD	High	Intel Xeon E5-2670 v2	2.5	Yes	Yes	Yes	Yes	Yes	$1.400
i2.8xlarge	32	244	8 x 800 SSD	10 Gigabit	Intel Xeon E5-2670 v2	2.5	Yes	Yes	Yes	-	Yes	$6.800
hs1.8xlarge	16	117	24 x 2,000	10 Gigabit	Intel Xeon Family	2	Yes	-	-	-	-	$4.600

Each vCPU is a hyperthread of an Intel Xeon core for M3, C3, R3, HS1, G2, and I2.

*M3 instances may also launch as an Intel Xeon E5-2670 (Sandy Bridge) Processor running at 2.6 GHz.

2. What you will submit?

Your deliverables for this project are to be written in a report, which will include the following:

- Assuming that 1 EC2 instance is 4.4GFlops, plot the cost of a compute cloud from 1GFlop to 1PFlop (1M GFlops)
- Compare the costs of 7 different instance types: t2.small, m3.large, c3.8xlarge, g2.2xlarge, r3.4xlarge, i2.8xlarge, and hs1.8xlarge
- Compute the private cloud cost equivalents to the 7 different Amazon EC2 Instance Types: t2.small.private, m3.large.private, c3.8xlarge.private, g2.2xlarge.private, r3.4xlarge.private, i2.8xlarge.private, and hs1.8xlarge.private
    - you may assume a 5 year amortization cost
    - you will have to factor in things other than hardware, such as cooling, power, administration costs, network infrastructure (e.g. switches)

    - show a separate table with the costs of each of the 7 different types, broken down by components (e.g. CPU, memory, motherboard, case, power supply, disk, network card, cooling, power, administration, etc)

- Plot (#1) the cost (in $) per instance type per hour, for the 7 Amazon EC2 instance types, vs. the 7 private cloud equivalent instance types, from 1GFlop to 1PFlop
- Plot (#2) the needed utilization of the private cloud from 1GFlop to 1PFlop for the different instance types in order to break even cost wise
- Explain in words the two plots, and anything interesting you might have found

Since your report will have 7 tables and 2 figures, its likely that your final report will be 3~6 pages long. A similar study looking to understand the cost of cloud storage was done previously; its writeup can be found at http://datasys.cs.iit.edu/projects/CloudStorage_summary12.pdf. This 1-page summary is a good example of what some of your graphs might look like, except that you are not scaling up the storage system in GB, but the compute capacity in GFlops. You are to make a PDF of your project writeup, and submit it on BB, with the following filename "Proj_LASTNAME1-LASTNAME2.pdf". Please ask questions on Piazza if there are any doubts about assumptions you need to make.