CS553 Homework #4

Understanding the Cost of Cloud Computing

Instructions:

- *Assigned date: Monday March 23rd, 2020*
- *Due date: 11:59PM on Monday April 10th, 2020*
- *Maximum Points: 100%*
- *This homework can be done in groups up to 3 students*
- *Please post your questions to the Piazza forum*
- *Only a softcopy submission is required; it will automatically be collected through GIT after the deadline; email confirmation will be sent to your HAWK email address*
- *Late submission will be penalized at 10% per day; an email to the TA with the subject "CS553: late homework submission" must be sent*

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. You can find many instance types 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 you have to estimate the cost of the hardware when building a private cloud, you can use hardware prices found at ThinkMate website (https://www.thinkmate.com) as good sources for server hardware. You must include a printout of your shoping cart in your final writeup report for this assignment; include this as an appendix at the end of your report.

You are to estimate the cost of different configurations for 3 different set of requirements:

- Configuration 1: Hadoop/Spark Cluster with 256K-cores, 2PB memory, 400PB HDD, and 25Gb/s Ethernet Fat-Tree network (each VM should be equivalent to the d2.8xlarge instance); in addition to the compute resources, a 800PB distributed storage shared across the entire cloud should be procured, with enough capacity for 800GB/sec throughput (for pricing comparison, see S3)
- Configuration 2: Support 1M virtual machines (VM) where each VM requires 2-core, 16GB RAM, 75GB NVMe storage, and 10Gb/s Fat-Tree network (each VM should be equivalent to the r5d.large instances); in addition to the compute resources, a 10PB distributed storage shared across the entire cloud should be procured, with enough capacity for 100GB/sec throughput (for pricing comparison, see S3)
- Configuration 3: Support deep learning with 1 exaflop of double precision performance (hint: each VM should be equivalent to p3.16xlarge instances; you will want to use the NVIDIA V100 GPUs (8 GPUs per node), and allocate 8-cores per GPU (64-cores per node) with 8GB of memory per core (512GB per node); the network to use is at least 10Gb/s per GPU (100Gb/s should work), and should be organized in a Fat-Tree network; in addition to the compute resources, a 1PB distributed storage shared across the entire cloud should be procured, with enough capacity for 10GB/sec throughput (for pricing comparison, see S3)

2. What you will submit?

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

- Report: A written document (typed, named hw4-report.pdf) describing your answers to the above questions.

- Compare the costs of the 3 different configurations between the public cloud (Amazon AWS) and the private cloud
  - 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); you can assume 1 system administrator is needed for every 1000 servers
  - show your data in three different tables with the costs of each of the 3 configurations, broken down by components (e.g. servers, network switches, cables, racks, cooling, power, administration, etc)
  - summarize your data in a 4th table, comparing the public cloud cost to the private cloud cost

- Explain in words if it is better to rent or buy. If it is better to buy, what utilization must you maintain over the 5 year lifetime of the private cloud in order to break even on the investment?

- Include your shopping cart of the 3 configurations. Your submission should be a single large PDF file, starting with your report, and followed by the shopping carts.

Table 1: Table summarizing each of the three configurations; please include 1 table per configuration, for a total of 3 tables

	Description	Price per Item	Quantity	Total Price
Compute Servers				
Network Switches				
Network Cables				
Racks				
Storage Servers				
Electric Power				
Cooling				
Administration				
TOTAL	N/A	N/A	N/A	

Table 2: Summary table comparing the 3 configurations between the public and private cloud; your cost of power, cooling, and administration should be to cover 5 years of costs

	Configuration 1	Configuration 2	Configuration 3
Public Cloud (including EC2 and S3) Cost over 5 years, 24/7 operation, with 100% usage			
Private Cloud cost over 5 years, 24/7 operation, with 100% usage			
What utilization must be achieved with the private cloud to make the private cloud option more attractive than the public cloud?			

Submit code/report through GIT. If you cannot access your repository contact the TAs. You can find a git cheat sheet here: https://www.git-tower.com/blog/git-cheat-sheet/
Grades for late programs will be lowered 10% per day late.