

CS553 Homework #4

Understanding the Cost of Cloud Computing

Instructions:

- Assigned date: Thursday October 6th, 2022
- Due date: 11:59PM on Friday October 14th, 2022
- Maximum Points: 100%
- This homework can be done in teams of up to 3 students
- Please post your questions to BB
- Only a softcopy submission is required; submission is a 2-step process: 1) push changes to GIT repository, and email confirmation will be sent to your HAWK email address at the deadline; a confirmation document with all team member names and A# must be submitted through BlackBoard for your submission to be graded; only 1 student must submit the assignment, and only the submitting student will receive the confirmation email
- Late submission will be penalized at 10% per day

1. Introduction

You are hired by a startup company who is considering to use cloud computing instead of building its own infrastructure. There is consensus that a cloud computing software stack at the layer of IaaS will be used, but it's 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. There are a variety of Amazon calculators for S3 (<https://calculator.aws/#/createCalculator/S3>) and EC2 (<https://calculator.aws/#/createCalculator/EC2>), please use them if you find them useful.

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 (for configuration #1 and #3). For configuration #2, you will need to use the Apple website (<https://www.apple.com/mac-mini/>). You must include a printout of your shopping 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; compute prices for a 5-year period:

- Configuration 1: Hadoop/Spark Cluster with 160K-cores, 128TB memory, 24PB HDD, and 100Gb/s Ethernet Fat-Tree network (each VM should be equivalent to the d3.8xlarge instance); in addition to the compute resources, a 48PB distributed storage shared across the entire cloud should be procured, with the expectation that 48PB of data will be read and written to S3 every year from outside of Amazon with enough capacity for 1GB/sec throughput (for pricing comparison, see S3 Standard). For EC2, you must use the reserved instance pricing with a standard 5-year term.
- Configuration 2: Support 1K application developers who are designing MacOS and iPad OS applications. They require a MacOS system with 6-cores (3GHz), 32GB RAM, 1TB storage, and 10Gb/s network (Amazon has mac1.metal instances that have everything you need except the 1TB storage, which you can provision through EBS). The developers work 40 hours/week, 48 weeks/year (they get 4 weeks of vacation per year). You must use on-demand EC2 pricing as developers are expected to provision their systems at the beginning of each working day, and release their systems at the end of each working day.
- Configuration 3: Ethereum crypto currency mining; you have an investor who has \$10M to buy hardware

to mine Raven Coin RVN (and pay for maintenance / sys admin, power, and cooling), or rent resources from Amazon EC2 to mine Raven Coin. Configure the best hardware you can from ThinkMate. For buying hardware solution, make sure to leave funds to pay for power, cooling, and system administrator. Raven Coin mining can be done on any compute hardware (CPUs or GPUs), but you will likely find that its most profitable to mine using GPUs. Since Ethereum mining is compute intensive, your processor, memory, hard drive, and network requirements are minimal (4-cores, 8GB RAM, 100GB HDD, and 1Gb/sec network). Identify the best Amazon instance (you must use Spot Instances to make sure you get the best hardware for the cheapest price); although spot pricing fluctuates over time, you can assume the spot price will remain fixed for the duration of your evaluation. For the purchase of the hardware scenario, you are free to locate the hardware in any state in the USA (for a full list of average electricity cost by state, see <https://www.chooseenergy.com/electricity-rates-by-state/>); since this will be a business venture, use the business electricity rates. If electricity is too expensive to make a profit, invest part of the \$1M in solar power (solar panels), and estimate the amount of energy you can extract. For an overview of various GPUs and their respective hashrates (the higher the hashrates, the more Raven Coin that can be mined), see <https://whattomine.com> (KawPow); this online resource has an even more exhaustive list of GPUs and their hashrate; <https://www.betterhash.net/mining/gpu/?page=1>. Once you have a hashrate, you can estimate how much money can be made mining Ethereum by using an online calculator such as <https://www.cryptocompare.com/mining/calculator/eth?HashingPower=0&HashingUnit=MH%2Fs&PowerConsumption=0&CostPerkWh=0&MiningPoolFee=1>. The mining calculator gives an instantaneous mining number, although in reality the amount of coin that can be mined would vary based on many factors (hash rate, hash difficulty, fees, etc). The profit similarly can vary based on the Raven Coin pricing, which can vary wildly. When computing the mining coins and expected profit, you can use the calculator above to compute it for a 5-year period, assuming the mining continues at the same rate, and the price remains at the same level. Your task is to compute the amount of profit that is expected after \$10M is invested in buying hardware and running it for 5-years, vs. renting the hardware from Amazon. Its possible that the profits you make will be less than the original investment (especially with the Amazon scenario).

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 500 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, and by how much.
- 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 2 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
Public Cloud (including EC2 and S3) Cost over 5 years		
Private Cloud cost over 5 years		

Table 3: Summary table comparing the profits with configuration #3 between the public and private cloud over a 5-year period

	Configuration 3
Public Cloud Mining Profit over 5 years	
Private Cloud Mining Profit over 5 years	

You will have to submit your solution to a private git repository created for you at [git@github.com:datasys-classrooms/cs553-fall2022-hw4-<team name>.git](https://github.com/datasys-classrooms/cs553-fall2022-hw4-<team name>.git). Make sure you have the following file in your repo: [hw4-report.pdf](#). The repository is created through GitHub Classroom and you will need to accept the assignment before you can clone it. The invitation link for HW4 is <https://classroom.github.com/a/36GEvn29>. You will also need to create a new team or join an existing team. Your submission will not be graded unless you submit a confirmation document through BlackBoard (BB) that clearly shows the pushing of your final homework to your GIT repository. This confirmation document can simply be a screen shot of your final commands to push your repository to GIT. The timestamp on the BB submission will be used to determine if the submission is on-time. **You must also include the names and A# of all your team members in this confirmation document.** 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.