



14-848 Cloud Infrastructure

SPRING 2026

MOHAMED FARAG

FARAG@CMU.EDU



Agenda

- Welcome and Introductions
- Focus Areas of this Course
- Course Syllabus & Schedule
- Class Expectations
- Introduction to the Cloud
- Next Steps



Important Course Aspects

- Relevant course topics to today's market and research
- Practical course: 30% theory + 70% practice
- Course focuses on Cloud Infrastructure (compared to development)
- Assignments are implemented on public cloud environment
 - Google Cloud is our main cloud environment
- Three practical project options to choose from.
- **Note: Don't take this course if you took 15319/15619/15719.**



Cloud Infrastructure in Today's Market

In recent years, the demand for Cloud Computing and Infrastructure jumped significantly.

- The [US Bureau of Labor Statistics \(BLS\)](#) forecasts that cloud computing employment opportunities will grow by 15% between 2021 and 2031. This growth rate is considered faster than the average for other occupations.
- “The global cloud computing market size is projected to grow from \$677.95 billion in 2023 to \$2,432.87 billion by 2030”, [Fortune Business Insights](#).
- [Indeed.com](#) lists Cloud Computing as the #1 skillset to learn in 2024.
- [Gartner](#) identifies that by 2029, cloud computing will be absolutely essential if businesses hope to compete, and for government agencies to achieve their mission.

Focus Areas of this Course

Cloud-related Technologies

Cloud Providers:

- Google Cloud Platform

Big Data Processing Platforms:

- Apache Hadoop
- Apache Spark

NoSQL Database:

- BigTable

LLM Platforms:

- Google AI Studio

DevOps

Deployment Scaling and Orchestration:

- Docker
- Kubernetes

DevSecOps

- Falco

Cloud Infrastructure

- Infrastructure-as-Code
 - Terraform
- Kafka
- Metaverse
- Data Centers
- Edge Computing & Fog Computing



Expectations for Incoming Students

- ***You are expected to know Python or are willing to learn it.***
 - A Python recording is released week for members who need support with Python
- ***You are expected to have a basic understanding of Computer Virtualization.***
 - If you don't know about virtualization, check out this reading:
<https://www.vembu.com/blog/physical-server-vs-virtual-machine-choice-open/>

Instructor and TA Introductions

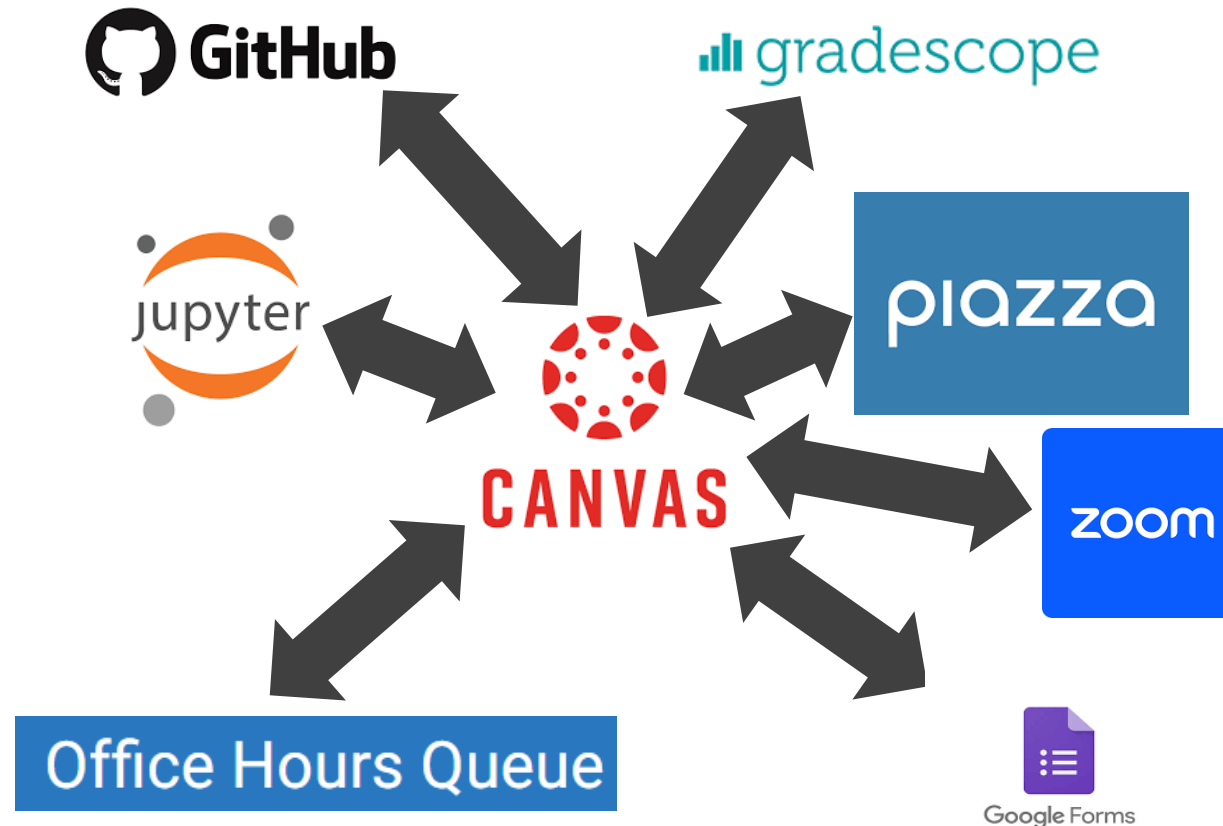
Instructor:

- Mohamed Farag: farag@cmu.edu

TAs:

- Ali Maharramli: amaharra@andrew.cmu.edu
- Prabhat Pandey: prabhatp@andrew.cmu.edu

Course Delivery Technology Stack



Course Logistics



- Lectures recordings are made available after the lecture end time.
- You may ask questions interactively or via Google Form during the lecture.
Google Form Link is: <https://forms.gle/P1mQHN9z9JS7fdCJ9>
- Use the Student Space Slack Channel to find a teammate for your course project (No instructor or TA help is offered there)
- Students who have approved accommodation shall contact the course instructor to figure out how the instructor can meet their needs.
- You may contact the student affairs if you must miss a few classes due to illness.

Course Logistics – Office Hours (Zoom)

OH Schedule (Conducted over Zoom)				
Days/Timeframes	11:30am-12:30pm ET	1-2pm ET	2-3pm ET	9-10pm ET
Monday	Mohamed			Ali
Tuesday			Prabhat	
Wednesday			Ali	Prabhat
Thursday	Mohamed		Prabhat	Ali
Friday		Ali		
	Instructor Office Hours - Conducted remotely via Zoom - URL can be found on			
	TA Office Hours - Conducted remotely via Zoom - URL can be found on Canvas			

- Instructor Office Hours will use this Zoom URL:
 - <https://cmu.zoom.us/j/92807330462?pwd=Phnl3sZpsZsHWmktSnoC3JRSTCwAU9.1>
- TA Office Hours will use this Zoom URL:
 - <https://cmu.zoom.us/j/92004227250?pwd=YnpiTW1qSXdscWRvbWYzd2t0elptQT09>

Course Logistics – Piazza Hours

Piazza OHs							
	9:30-10:30am ET	11am-12pm ET	12-1pm ET	2-3pm ET	3-4pm ET	4-5pm ET	10:00-11:00pm ET
Monday				Ali			
Tuesday					Prabhat		
Wednesday			Prabhat				Ali
Thursday	Ali					Ali	Prabhat
Friday		Ali					
Saturday		Prabhat					

- Use Course Piazza to ask asynchronous questions that require instructor and/or TA help
- Please note that TAs will respond to inquiries/questions made ***before*** the Piazza OHs start time. Questions and inquiries that are made during the OHs time slot are not guaranteed to be answered during the same time slot.

Office Hours Etiquette Reminder

- Sign-up for a spot in the OH Queue: <https://ohq.eberly.cmu.edu/#/courses> and search for “14-848”
- Office Hours aim to help you find the path to maximize your learning experience.
- Getting the answers from the TA directly won't help you learn so **there won't be direct solutions provided during Office Hours.**
- The goal of the office hours is to **give you some ideas and pointers for you** to debug the issues.
- Please don't plan to spend **more than 15 minutes** in your conversation with the TA.
- Ask **good questions with due diligence**. Please research the issue and put an effort in implementing it before coming to Office hours.
 - **Example of a bad question:** I found this draft code online and I'm citing it but can't get it to work. Can you help?
 - **Example of a good question:** I'm getting a bug in my deployment to the cloud, I researched the issue and found these 3 different references (share the URLs). I implemented the first one and it didn't work. I'm trying the second one now and getting an error that I can't find enough references to it online. What could be the root cause of it?

Course Assessment

Final Exam	Project	Assignments	Quizzes
20%	20%	40%	20%



You can boost your final exam score with up to 4 bonus points



Your lowest two quiz scores are dropped at the end of the semester
+
2 extra-credit quizzes

Course Assessment – Final Exam on April 23 during the lecture

Final Exam	Project	Assignments	Quizzes
20%	20%	40%	20%

Final Exam: is an open-note test.

- Students will have access to all the **PDFs** for lectures, readings and HW solutions. Students can **bring any hard-copied materials** with them.
- Students are required to follow the schedule of their registered section. **On the scheduled final lecture of each section, final exam will be released only to the registered students of the corresponding section.** Each section will have its final exam version(s).
- Exam will be offered via **Lockdown Browser** and no knowledge exchange is allowed among students during the exam.
- Students are expected to install and test Lockdown browser on their machines ahead of the exam. If students face an issue with Lockdown browser installation, students must reach out to the instructors **no later than 2 weeks** before the final exam date.
- **Sharing hard-copied notes is prohibited during the exam.**

Course Assessment – Cont'd – Final Exam Booster

Final Exam	Project	Assignments	Quizzes
20%	20%	40%	20%

You can get four bonus points on the final exam if you obtain TWO of the following certifications two weeks before the final exam

- Google Cloud: Associate Cloud Engineer
- AWS Certified Solutions Architect – Associate
- Microsoft Certified: Azure Developer Associate
- GCP Professional Data Engineer
- AWS Certified Data Engineer - Associate
- Microsoft Azure Data Engineer Associate
- Certified Kubernetes Application Developer
- NVIDIA's Fundamentals of Accelerated Data Science

Late submissions are not accepted.

Obtaining only 1 certificate will gain you two bonus points on the final exam score

Course Assessment – Cont'd

Final Exam	Project	Assignments	Quizzes
20%	20%	40%	20%

- **Homework Assignments:** there will be 7 homework assignments provided throughout the semester covering the practical aspects of the class and extending the exercises offered during the recitations.
- Students will have 3 days to submit the assignment after the due date with a late penalty. Late penalties are applied based on the timestamp of the last code commit on GitHub and it will follow this equation (no matter whether the delay is in minutes or in hours):
 - Total of 5 points for up to 24 hours delay
 - Total of 15 points for the next 24 hours delay
 - Total of 25 points for the next 24 hours delay
 - 100 points penalty (no grade) after this time.

After homework grades are released, **regrade requests can be made for 24 hours via Gradescope and CANNOT be submitted via email.**

Course Assessment – Cont'd

Final Exam	Project	Assignments	Quizzes
20%	20%	40%	20%

- **Course Project:** Each student will have the option to peer with a team member for the project and you will choose one of three project options to submit. This project leverages most of the topics and practices that are covered throughout the semester. Course project details are released in Week-3. Project submission deadline is **April 2nd, 2026, 11:59PM ET /8:59pm PT.**
- **Quizzes:** there will be 1 quiz published on Canvas during the lecture with a specific access code. The access code will be revealed during the lecture to the registered students of the corresponding section.
 - Quizzes will start from the second week of classes.
 - **Quizzes will cover lecture notes, readings, verbal explanations and recitations.**
 - **Students will receive two excused absences for lectures (and their quizzes) for emergencies, sickness, etc.**
 - If you need to miss classes for extended time, please refer to the course homepage on Canvas.



Course Recitations

- Recitations offer important supplementary information, exercises or practices sets that are not covered in class.
- There are 7 recitations held throughout the semester. Refer to the course schedule for the corresponding weeks.
- Typically, recitations are held on Fridays of select weeks from 3-5PM ET (12-2PM PT). Recitations will use the TA OH URL.
 - TAs may change the recitation date/time with a prior announcement on Piazza.
- Recitations are recorded and will be published on Canvas.
- There will be quiz questions covering notes presented in the recitations.



Course Grade Scheme

+/- are used to provide granularity in equal intervals in B and C ranges

Grade	Percentage Interval
A/A-	[85-100%], A starts from 93
B	[70-85%)
C	[55-70%)
D	[40-55%)
R (F)	Below 40%

Course Schedule

Date	Topic	Notes
Week-1 (Jan. 12 th)	- Introduction & Syllabus - Computing Service Models.	- System Setup homework released
Week-2 (Jan. 19 th)	- Virtualization Basics - Containerization	- System Setup homework deadline. - Docker homework released - Recitation-1: Introduction to Google Cloud
Week-3 (Jan. 26 th)	- Lab: Containerization - Deployment Orchestration - Pokémon Go Case Study	- Course Project released. - Recitation-2: Docker Scripting and Debugging + Docker Compose Basics.
Week-4 (Feb. 2 nd)	- Lab: Deployment Orchestration - Kafka	- Docker homework deadline - Kubernetes homework released. - Recitation-3: YAML Scripting for Kubernetes.
Week-5 (Feb. 9 th)	- Lab: Confluent Kafka - Infrastructure-as-a-Code	- Recitation-4: More Kafka Examples.
Week-6 (Feb. 16 th)	- Infrastructure-as-a-Code (Cont'd) Terraform - Lab: Terraform	- Kubernetes homework deadline - Kafka Homework released - Recitation-5: HCL Basics and Terraform Practice Set.
Week-7 (Feb. 23 rd)	- Cloud Data Storage Models - Lab: NoSQL Database: Google BigTable	- Course Project Checkpoint
Spring Break (Mar. 2 nd - Mar. 6 th)		

Course Schedule – Cont'd

Week-8 (Mar. 9 th)	- Cloud Infrastructure Concepts - Introduction to Hadoop	- Kafka Homework deadline - Terraform Homework released
Week-9 (Mar. 16 th)	- Hadoop HDFS - Hadoop MapReduce	- Terraform Homework deadline
Week-10 (Mar. 23 rd)	- Big Data Algorithms - Introduction to Spark	- Hadoop MapReduce homework released - Recitation-6: Hadoop MapReduce Practice Set.
Week-11 (Mar. 30 th)	- Spark (Cont'd) - Lab: Spark Programming	- Course project submission deadline
Week-12 (Apr. 6 th)	- Metaverse, Edge Computing and Fog Computing	- Hadoop MapReduce homework deadline - Apache Spark homework released - Recitation-7: Spark Practice Set.
Week-13 (Apr. 13 th)	- Cloud Security & Privacy Concepts - DevSecOps	- Apache Spark homework deadline
Week-14 (Apr. 20 th)	- Student Vote Lecture. - Final Exam	



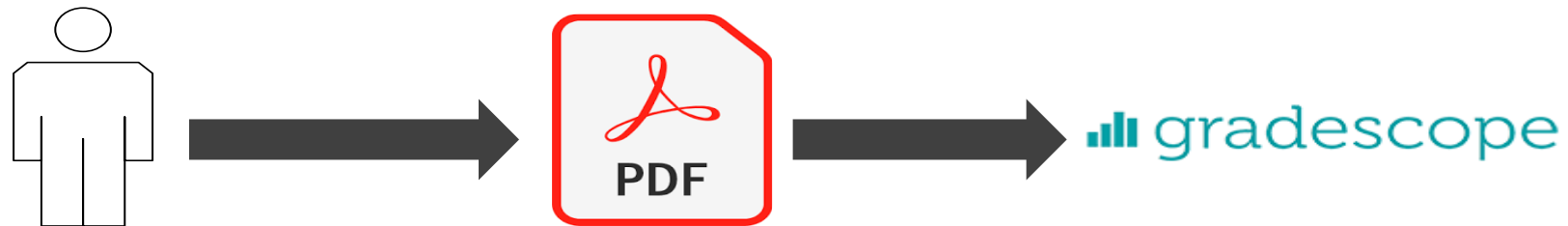
Expectations down the Road!

- For the final exam, we will use the Lockdown browser, and you will have access to the lecture PDFs via the browser. You will receive few training attempts on the final exam environment before the exam date.
- You are welcome to use other cloud platforms like Azure and AWS. That said, we don't have any credits/support that we can offer for these systems. You may need to use a Free Trial version.

HW Submission Guidelines

- HW-1 will focus on Environment Setup and essential GitHub Skills.

- HW-1 submission:



- HW-2 and later assignment submissions:





Academic Integrity Violations (AIVs)

- AIVs are serious and can have direct impact on your course grade, your scholarship -if any-, your graduation timeline, and/or your continuation in your degree program.
- You may use all the materials and readings that we explicitly provide you with in the assignment document. In this case, attribution is not required.
- You may use other course material, including lectures, Piazza posts by the instructors, and material from Canvas. For any such use, you must provide a proper attribution.
- You may not obtain code or other solution information from an unauthorized external source, including web pages, code repositories, blog posts, etc.
- You may not look at someone else's code (or other documents.) This includes one person looking at code and describing it to another. There is no notion of looking “too much,” since no looking is allowed at all.



Academic Integrity Violations (AIVs) – Cont'd

- Students may use Generative AI platforms (e.g., ChatGPT) to assist them with understanding the homework concepts.
- However, students can't use the direct homework question in their prompt to the AI tool.
- Students are expected to cite the text (or code) that was generated from ChatGPT carefully. This includes scenarios where generative-AI was used to generate base-code/scenario and minor (or significant) changes have been made to it.
- Beyond this, you may not obtain code or other solution information, either by copying, retyping, or looking at files or documents from any of the above-listed courses for this semester, or a previous semester.
- Don't share quiz access codes with your peers.



Teaching Expectations

- Not all material is captured on the slides; students are expected to take notes during lectures.
- Students are expected to attend recitations (or watch the recordings) prior to the following lecture.
- Students are expected to complete all assigned readings and **in-class exercises** on-time.
- Instruction is delivered in a passionate teaching style. Don't be scared! 😊

Other Syllabus Information

- **If you run out of Google Cloud credits, plan for 24-48-hour delay to get a new coupon.**
- Syllabus contains important information about student wellness, student academic success center, and food insecurity.
- The Syllabus can be found on Canvas under the Modules section

Waitlisted?

For enrollment questions and inquiries, please
email the INI Academics at

ini-academic@andrew.cmu.edu



Bottom Line!

- You are a graduate student at CMU, and we expect you to **pay close attention to the details** mentioned during lectures and in your homework assignments.
- Ask questions and **avoid making unreasonable assumptions**.
- Use your **intellectual abilities and problem-solving skills** to fill in any gaps beyond what the teaching team can share with you.
- We are **committed to your success** both inside and outside the classroom.
- Feel free to **reach out with any questions** related to research, internships, career advice, or related topics.

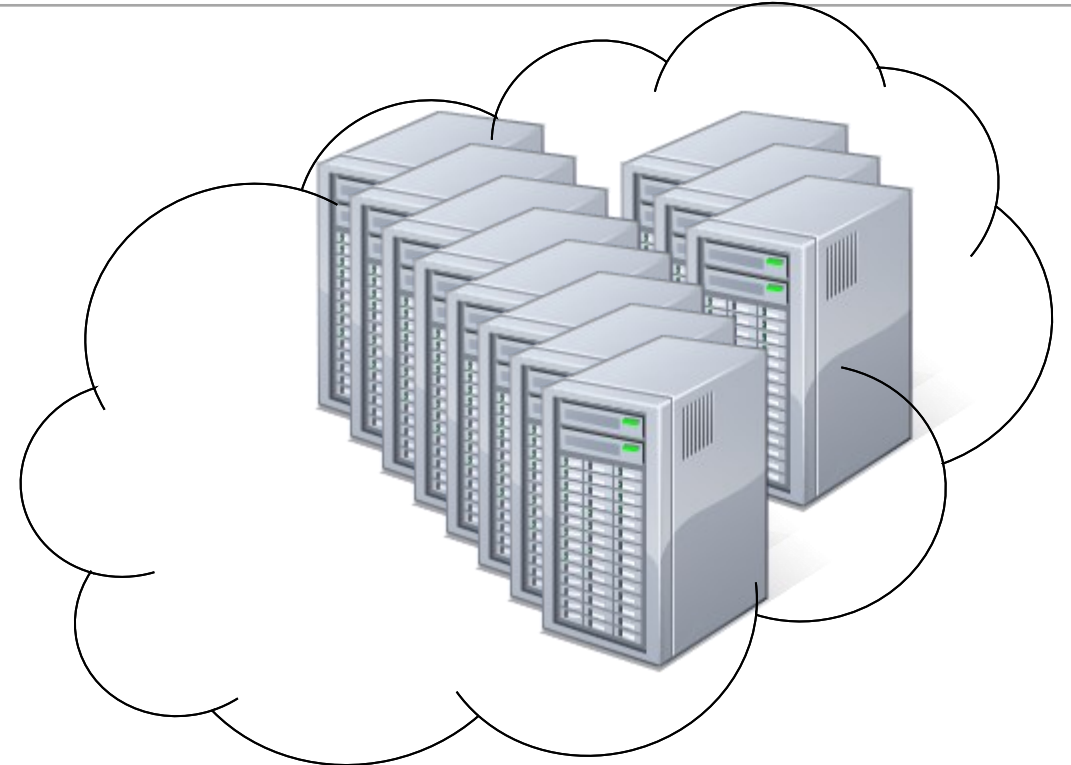


Now, let's start!
What is the Cloud?

What is the Cloud?



- Is it a cluster!
- Is it a supercomputer!
- Is it a datastore!
- Is it a superhero?!
- None of the above?
- All the above?



- In simple words: **SCN**
 - Cloud = Lots of storage + Compute cycles nearby + Network bandwidth



The cloud in the industry!

“Cloud” refers to large Internet services running on 10,000s of machines (Amazon, Google, Microsoft, etc.)

There services are offered to external customers for cycle renting and storage

- Amazon EC2: virtual machines at 8.5¢/hour
- Amazon S3: storage at 21¢/GB/month
- Google Cloud AppEngine
- Windows Azure

The 5 attributes of Cloud are defined
by National Institute of Standards and Technology (NIST).

Scientifically, what is the Cloud?

Cloud is a model that offers the following characteristics:



**On-demand
self-service**

No human
intervention
needed to get
resources



**Broad network
access**

Access
from
anywhere



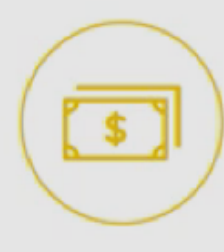
**Resource
pooling**

Provider
shares
resources
to
customers



**Rapid
elasticity**

Get more
resources
quickly as
needed



**Measured
service**

Pay only
for what
you
consume



Carnegie Mellon University

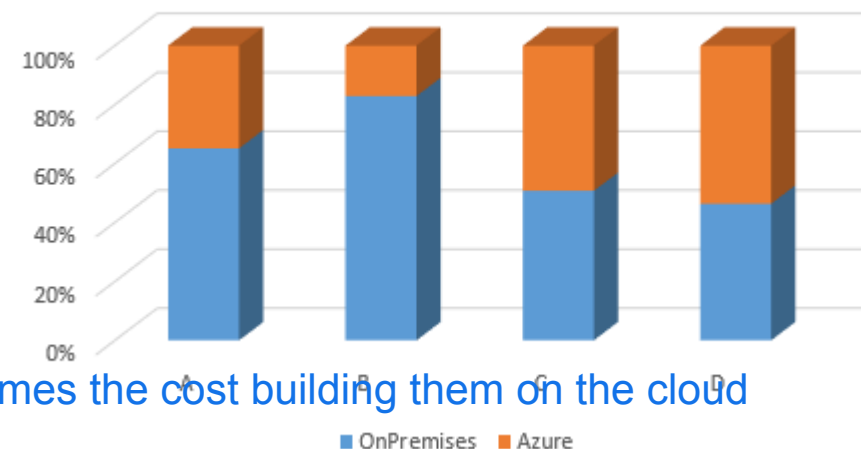
Cloud Cost Savings by Deployment Size

Consider the following scenarios for resource deployment:

- A. Modest Deployment
- B. Tiny Deployment
- C. Enterprise
- D. Very Large Enterprise

Scenario	A	B	C	D
Small VM	20	4	200	500
Medium VM	40	10	750	1750
Large VM	10	2	150	750
Storage (TB)	10.9	2.5	171.9	468.8

Costs Compared



A: you'd paying 60% Onpremises or 40% on Azure

B: the cost of building the services OnPremises is 4 times the cost building them on the cloud

Rental lock-in: it's costly for you to move your infrastructure from one to another compared with sticking to one

you don't see the upgrades since they are done automatically

Cloud costs vs. Traditional IT costs in Tiny Deployments

Cloud management cost is just a small portions of the cost compared with cloud operational cost

Aspect	Traditional IT Costs	Cloud-Based Server Costs
Initial Setup Costs	High, includes hardware, software, and infrastructure setup	Low to moderate, primarily subscription-based or pay-as-you-go
Maintenance and upgrades	High, ongoing costs for hardware maintenance and software upgrades	Low, handled by cloud service providers
Scalability	Limited, requires significant investment for expansion	High, easily scalable based on demand
Operational flexibility	Low, fixed resources and limited flexibility	High resources can be adjusted as needed
Disaster recovery	High, requires investment in backup infrastructure	Low to moderate, often included in cloud services
Energy and cooling costs	High, significant energy consumption and cooling requirements	Low, managed by cloud providers
Staffing requirements	High need for IT staff to manage and maintain systems	Low, reduced need for in-house IT staff
Downtime costs	Potentially high, dependent on in-house capabilities	Low, high availability and reliability offered by providers
Security costs	High, requires investment in security measures	Variable, often included in cloud service plans
Time to deployment	Long, extensive setup and configuration time	Short, rapid deployment and provisioning of resources

Examples of Cloud Services

- Dropbox
- Google Drive
- Microsoft OneDrive
- Apple iCloud

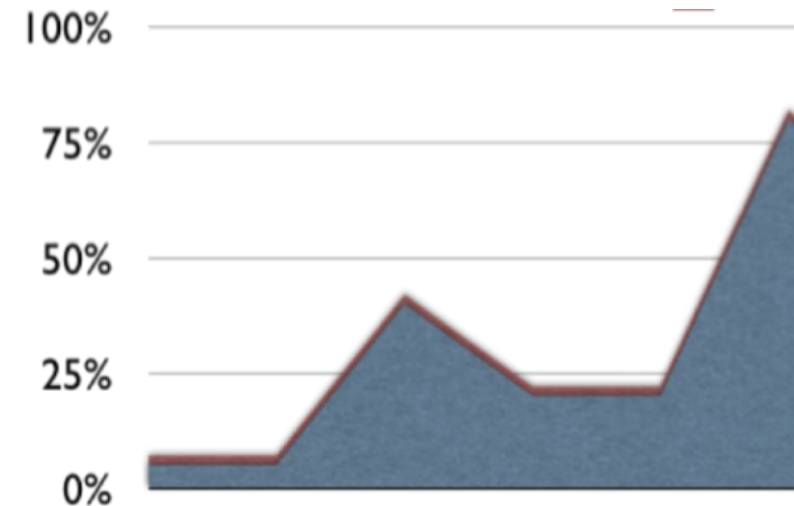
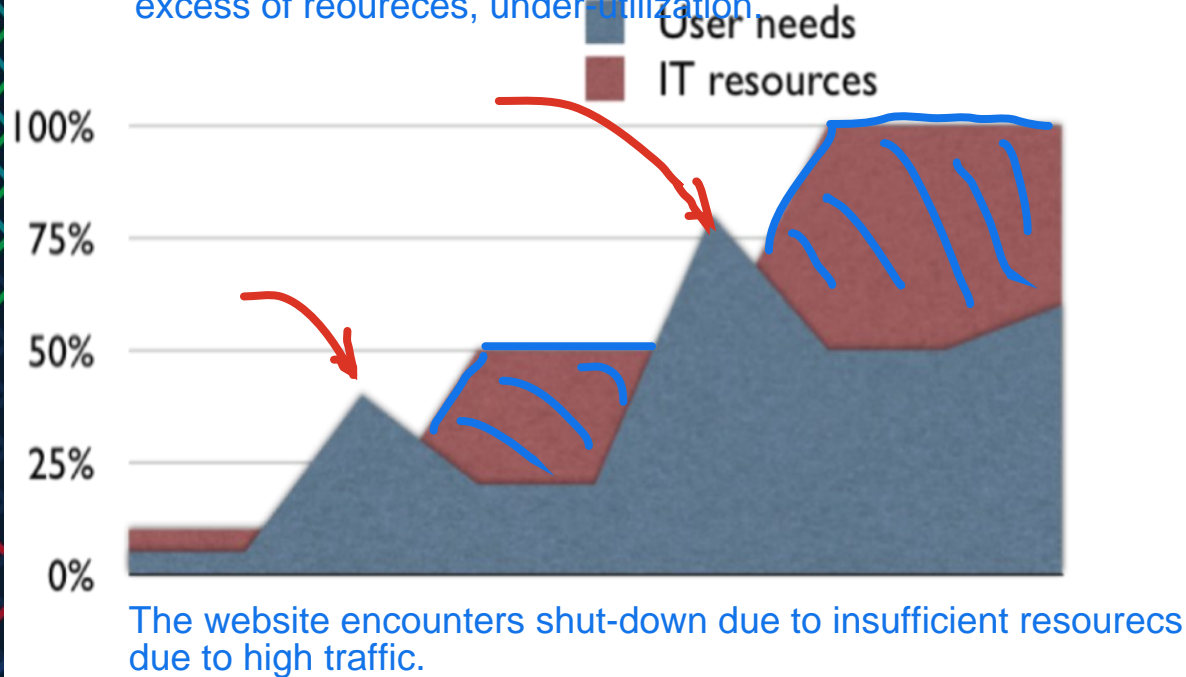


- Netflix - hosted on AWS
- Google search – Google Cloud
- Google Docs, Sheets, and Slides
- Facebook



Company Infrastructure and User Needs With vs. Without Cloud

It takes time to set up, to configure whenever the infrastructure is scaling up. And after the increase of resources, there is an excess of resources, under-utilization.



**Which diagram reflects the IT resources and User needs for Infrastructure using the Cloud?
And Why?**



Conventional Computing Infrastructure vs Cloud Computing Infrastructure

Conventional

On-premises (本地部署)

Manually Provisioned

Dedicated Hardware

Fixed Capacity

Pay for Capacity

Capital & Operational Expenses

Managed via System administrators

Cloud

Self-provisioned

Shared Hardware

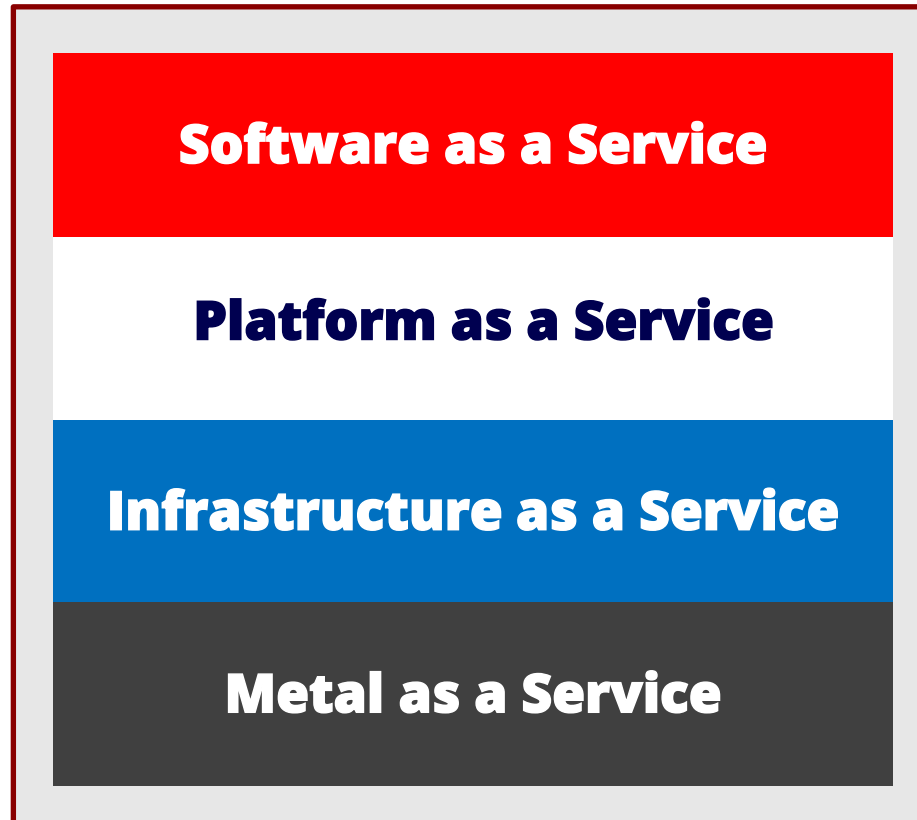
Elastic Capacity

Pay per Use

Operational Expenses

Managed via APIs

Cloud Computing Service Model





Cloud Computing Enablers

Data Center + Virtualization

(Hardware) (Software)

SCN



Next Steps

- Complete Course Entry Survey
 - Link can be found on Canvas under “Modules” section.
- Join the Course Piazza.
- Join the student Slack workspace.
- Read the Course Syllabus.

Waitlisted Students

- All materials for first two weeks will be uploaded here

