14-848 Cloud Infrastructure

FALL 2025

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.
- Skillsoft lists Infrastructure as a Code as the #2 skillet to learn in 2023
- The most in-demand hard skills of 2023 lists Cloud Computing among the top 10 hard skills.

GKE, EKS, and AKS are the managed Kubernetes services offered by
- Google Cloud Platform (GCP)
- Amazon Web Services (AWS)
- Microsoft Azure

Focus Areas of this Course

Cloud-related Technologies

Cloud Providers:

- Google Cloud Platform

Big Data Processing Platforms:

- Apache Hadoop
- Apache Spark

NoSQL Database:

- Neo4j
- BigTable

Large Language Models:

- 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

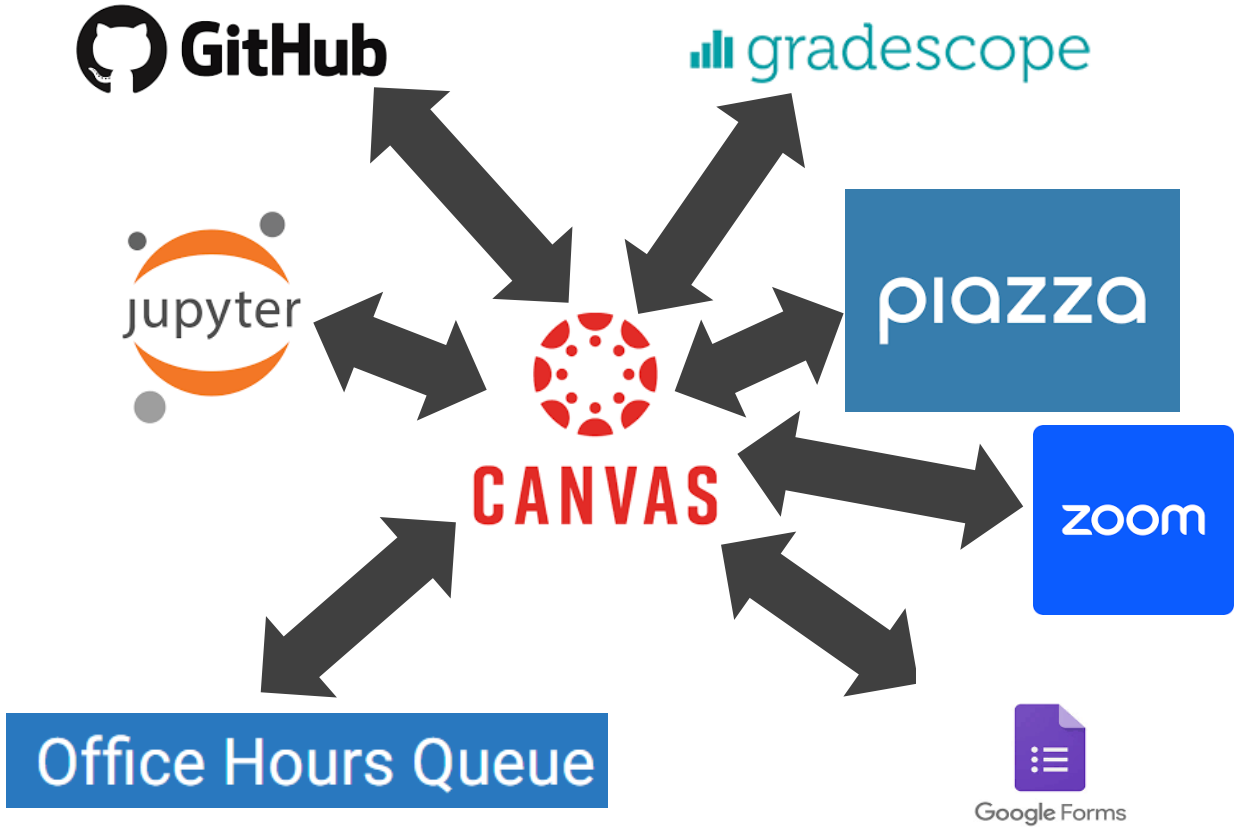
Instructor:

- Mohamed Farag: farag@cmu.edu

TAs:

- Ajay Sunandan asunanda@andrew.cmu.edu
- Fraol Dechasa fdechasa@andrew.cmu.edu

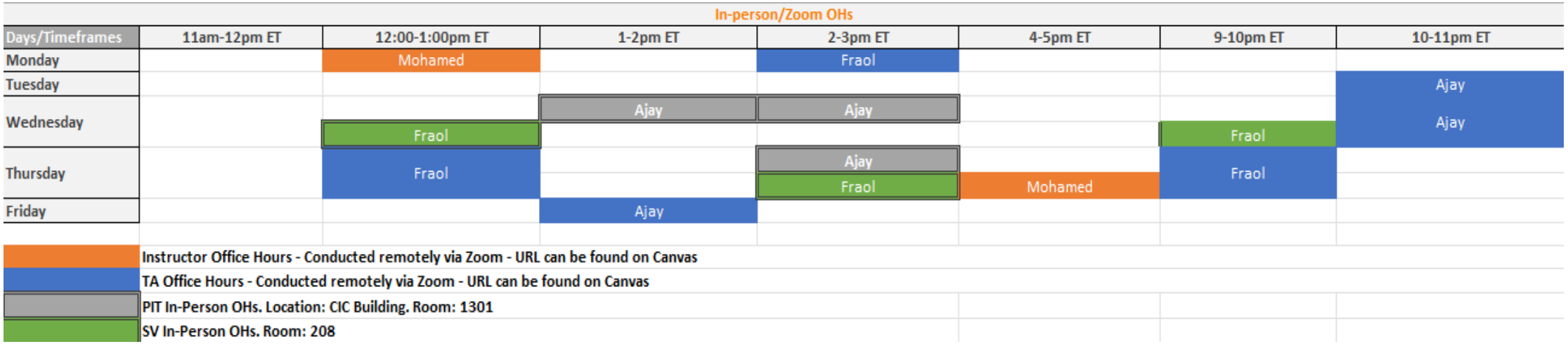
Course Delivery Technology Stack



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/In-Person)



In-person/Zoom OHs

Days/Timeframes	11am-12pm ET	12:00-1:00pm ET	1-2pm ET	2-3pm ET	4-5pm ET	9-10pm ET	10-11pm ET
Monday		Mohamed		Fraol			
Tuesday							Ajay
Wednesday		Fraol	Ajay	Ajay		Fraol	Ajay
Thursday		Fraol		Ajay Fraol	Mohamed	Fraol	
Friday			Ajay				

Instructor Office Hours - Conducted remotely via Zoom - URL can be found on Canvas

TA Office Hours - Conducted remotely via Zoom - URL can be found on Canvas

PIT In-Person OHs. Location: CIC Building. Room: 1301

SV In-Person OHs. Room: 208

- Instructor Office Hours will use this Zoom URL:
  - https://cmu.zoom.us/j/92231657955?pwd=mNyb0fCSLo1EZwcajFcLugNwiV6dBc.1
- TA Office Hours will use this Zoom URL:
  - https://cmu.zoom.us/j/92004227250?pwd=YnpiTW1qSXdscWRvbWYzd2t0elptQT09
  - In-person OHs won't have Zoom.

Course Logistics – Piazza Hours

Piazza OHs				
	10-11am ET	3-4pm ET	9-10pm ET	11pm-12am ET
Monday		Fraol		
Tuesday				Ajay
Wednesday				Ajay
Thursday	Ajay		Ajay	Fraol
Friday		Fraol		
Saturday		Fraol		

- 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.

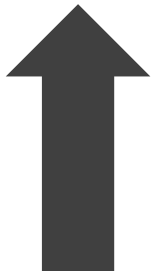
Office Hours Etiquette Reminder

- 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
15%	20%	40%	25%

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 December 3rd** during the lecture

Final Exam	Project	Assignments	Quizzes
15%	20%	40%	25%

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
15%	20%	40%	25%

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
15%	20%	40%	25%

- **Homework Assignments:** there will be 7 homework assignments provided throughout the semester covering the practical aspects of the class. There will be good learning curve that students will have to take on their own.
- 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
15%	20%	40%	25%

- **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 details are released in Week-3. Project submission deadline is **November 13th, 2025 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.
  - **Students will receive two excused absences for lectures (and their quizzes) for emergencies, sickness, etc.**
  - If you need to attend remotely for extended time period, please refer to the course homepage on Canvas.

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 (Aug. 25th)	- Introduction & Syllabus - Virtualization Basics	- System Setup homework released
Week-2 (Sep. 1st)	- Containerization	- System Setup homework deadline. - Docker homework released
Week-3 (Sep. 8th)	- Lab: Containerization - Deployment Orchestration - Pokémon Go Case Study	- Course Project released - Kubernetes homework is released
Week-4 (Sep. 15th)	- Lab: Deployment Orchestration - Kafka	- Docker homework deadline
Week-5 (Sep. 22nd)	- Lab: Confluent Kafka - Infrastructure-as-a-Code	- Kubernetes homework deadline - Kafka Homework released
Week-6 (Sep. 29th)	- Infrastructure-as-a-Code (Cont'd) Terraform - Lab: Terraform	
Week-7 (Oct. 6th)	- Cloud Data Storage Models - Lab: NoSQL Database - Neo4j AuraDB	- Course Project Checkpoint

Course Schedule – Cont'd

Week-8 (Oct. 20th)	- Cloud Infrastructure Concepts - Introduction to Hadoop	- Kafka Homework deadline - Terraform Homework released
Week-9 (Oct. 27th)	- Hadoop HDFS - Hadoop MapReduce	- Terraform Homework deadline
Week-10 (Nov. 3rd)	- Big Data Algorithms - Introduction to Spark	- Hadoop MapReduce homework released
Week-11 (Nov. 10th)	- Spark (Cont'd) - Lab: Spark Programming	- Course project submission deadline
Week-12 (Nov. 17th)	- Metaverse, Edge Computing and Fog Computing - Cloud Security & Privacy Concepts	- Hadoop MapReduce homework deadline - Apache Spark homework released
Week-13 (Nov. 24th)	- DevSecOps	- Apache Spark homework deadline
Week-14 (Dec. 1st)	- Introduction to GenAI on the Cloud using Vertex AI - 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 focuses on Environment Setup and GitHub Skills. It's released on Canvas, and you can submit it with no penalty until Thursday **September 4th, 11:59PM ET**.
- 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.
- Simple rules to follow:
  - Cite all the references you are using. Use APA citation style.
  - Cite ChatGPT (or other AI tools) for any code/info used in your answers.
  - Don’t use more than 30% of your solution/answer from external sources.
  - Collaborate and share ideas with your peers, and not code.
  - Don’t share code with your peers (including in-class group exercises). Don’t use your peer’s code even after changing variable names or statement order.
  - Don’t share quiz access codes with your peers.

create SCN

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

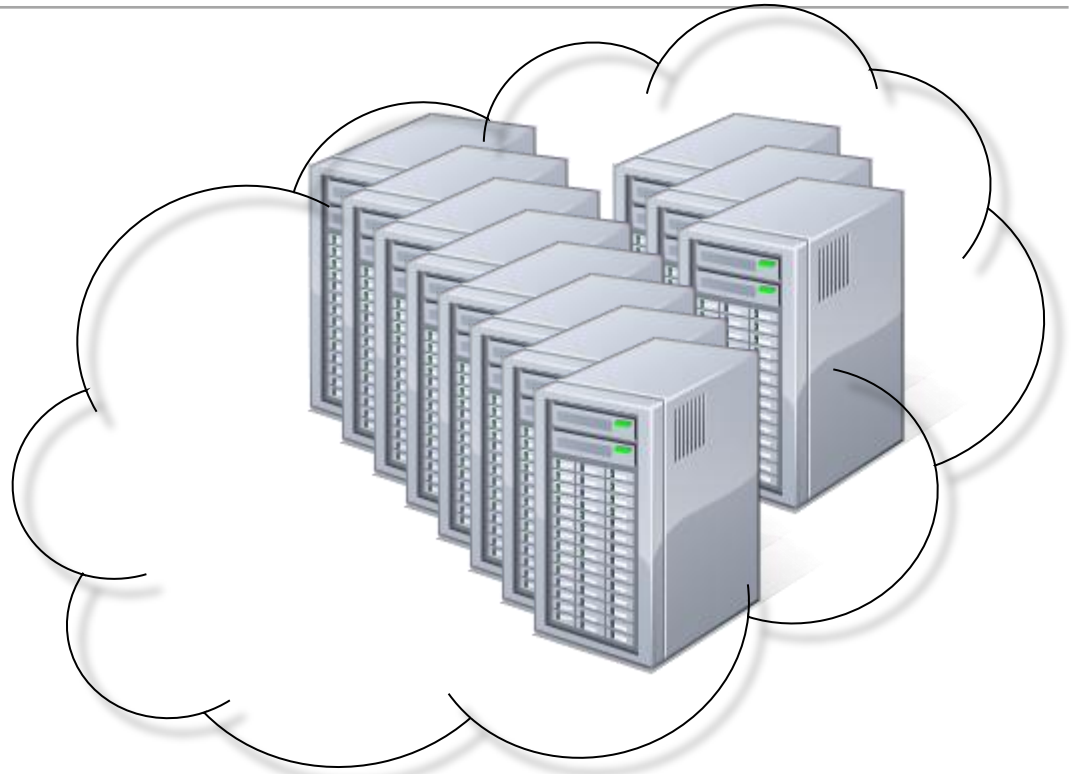
Bottom Line

- 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:
  - **Cloud = Lots of storage + Compute cycles nearby + Network bandwidth**

SCN: storage, computation, and networj

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

Scientifically, what is the Cloud?

Cloud is a model that offers the following characteristics:

Why Do We Care about Cloud Infrastructure?

Accessibility

Big Data

Cost

Cloud Cost Savings by Deployment Size

Consider the following scenarios for resource deployment:

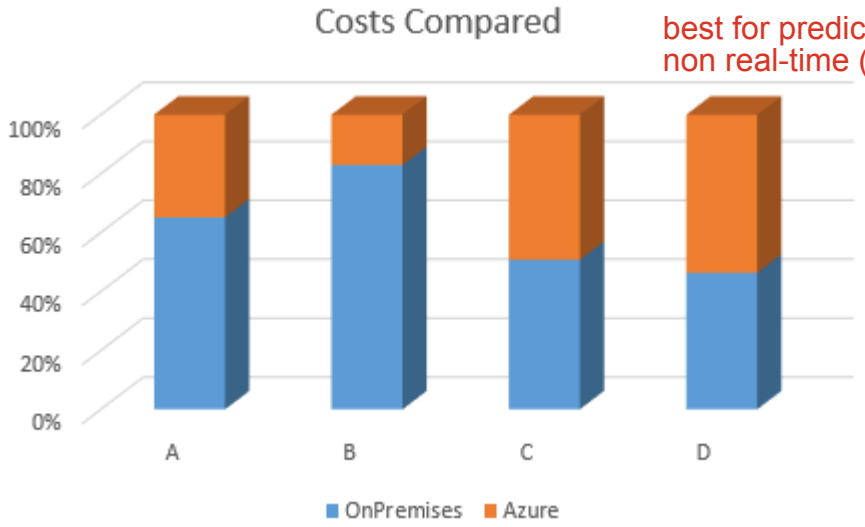
A. Modest Deployment

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



best for predictable workload + non real-time (fixed workers)

On premise (can just be at your basement) != private cloud (have cooling and surveilience systems)

Cloud costs vs. Traditional IT costs in Tiny Deployments

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

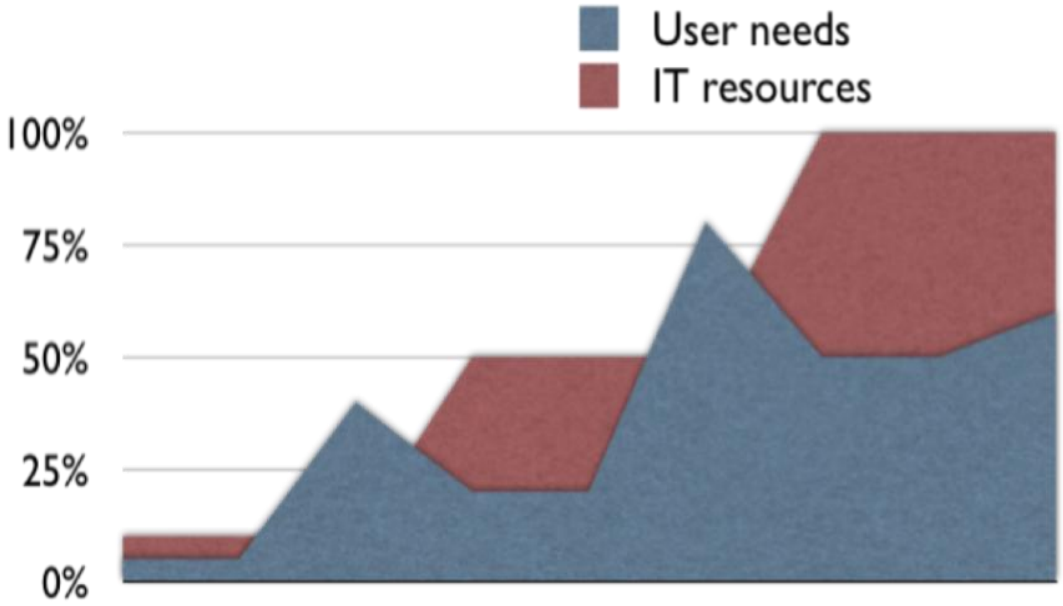
- **Dropbox**
- **Google Drive**
- **Microsoft OneDrive**
- **Apple iCloud**

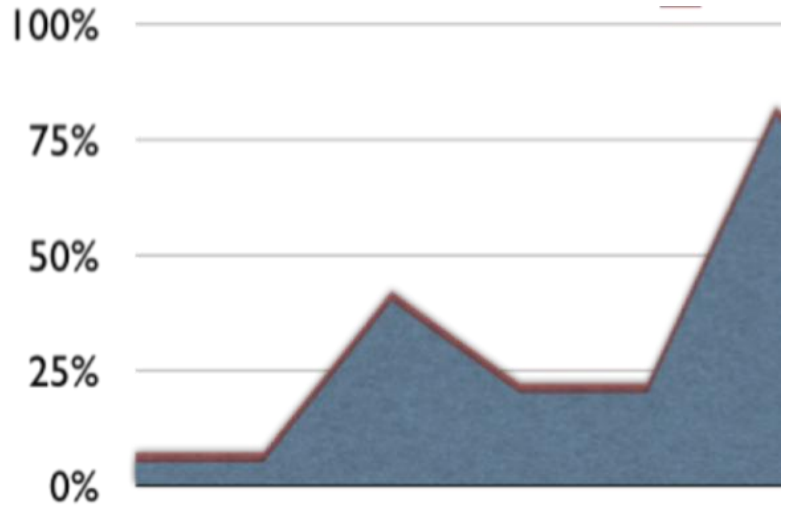
- Netflix - hosted on AWS
- Google search – Google Cloud
- Google Docs, Sheets, and Slides
- Facebook






Company Infrastructure and User Needs With vs. Without Cloud





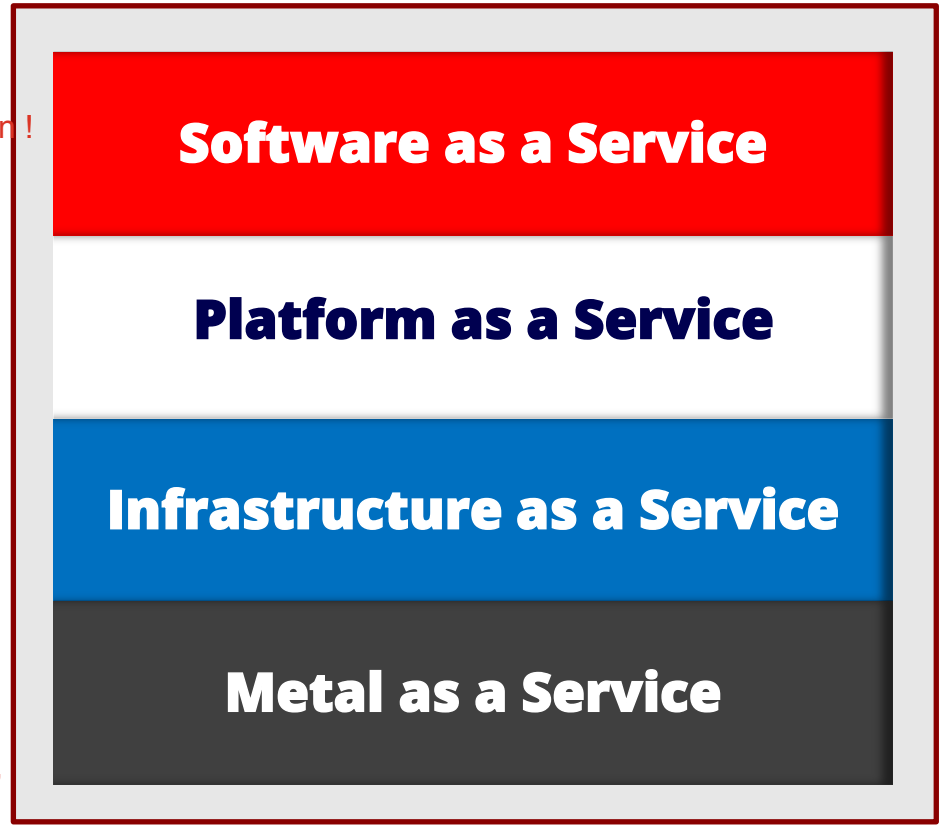
Which diagram reflects the IT resources and User needs for Infrastructure using the Cloud? And Why? right

Conventional Computing Infrastructure vs Cloud Computing Infrastructure

Conventional	Cloud
Manually Provisioned	Self-provisioned
Dedicated Hardware	Shared Hardware
Fixed Capacity	Elastic Capacity
Pay for Capacity	Pay per Use
Capital & Operational Expenses	Operational Expenses +risk mitigation expenses (this can be more expensive than the capital)
Managed via System administrators	Managed via APIs

Cloud Computing Service Model

most common !



cloud providers (AWS, GCP, Azure, IBM cloud)
---
third parties (netflix, ...)
---
user

CP

TP (salesforce)
---
user 1 (consumer of TP/ provider of the service)
they use PaaS to build SaaS
---
user 2

e.g. use GCP to create a virtual machine

CP (SCNs)
---
user

physical SCNs -> actual server
e.g. juju
users own the servers, it's like
they rent the storage.
This is not private cloud (on premise),
which needs to own the land, video
surveilience and cooling sys

SaaS: Software as a Service

- Provided with access to application software in the cloud
  - On-demand software
- Most applications can be run directly from web browser
- Largest cloud market
- **Examples:** Google Apps, Microsoft Office 365, Oracle's Netsuite, SAP's Concur, Cisco WebEx, GoToMeeting

  e.g. Netflix, dynamic CRM

PaaS is a middle man between end users and CP -> Users use PaaS to deploy their work on the cloud

PaaS: Platform as a Service

- Provides computing platforms which typically includes operating system, programming language, execution environment, database, web server, etc. to build cloud applications.
- Applications using PaaS inherit cloud characteristic such as scalability, high-availability, multi-tenancy, SaaS enablement, and more.

AMS lambda, google files, dynamic CRM

- **Examples:** Google App Engine, AWS Elastic Beanstalk, Salesforce.com, Amazon EMR, MS Azure HDInsight, GCP Dataproc

incl. Apache Spark

IaaS: Infrastructure as a Service

- Offers storage and computing resources that developers and IT organization use to deliver custom business solutions

- **Examples:** Amazon EC2, VMWare vCloud, GCP Compute Engine

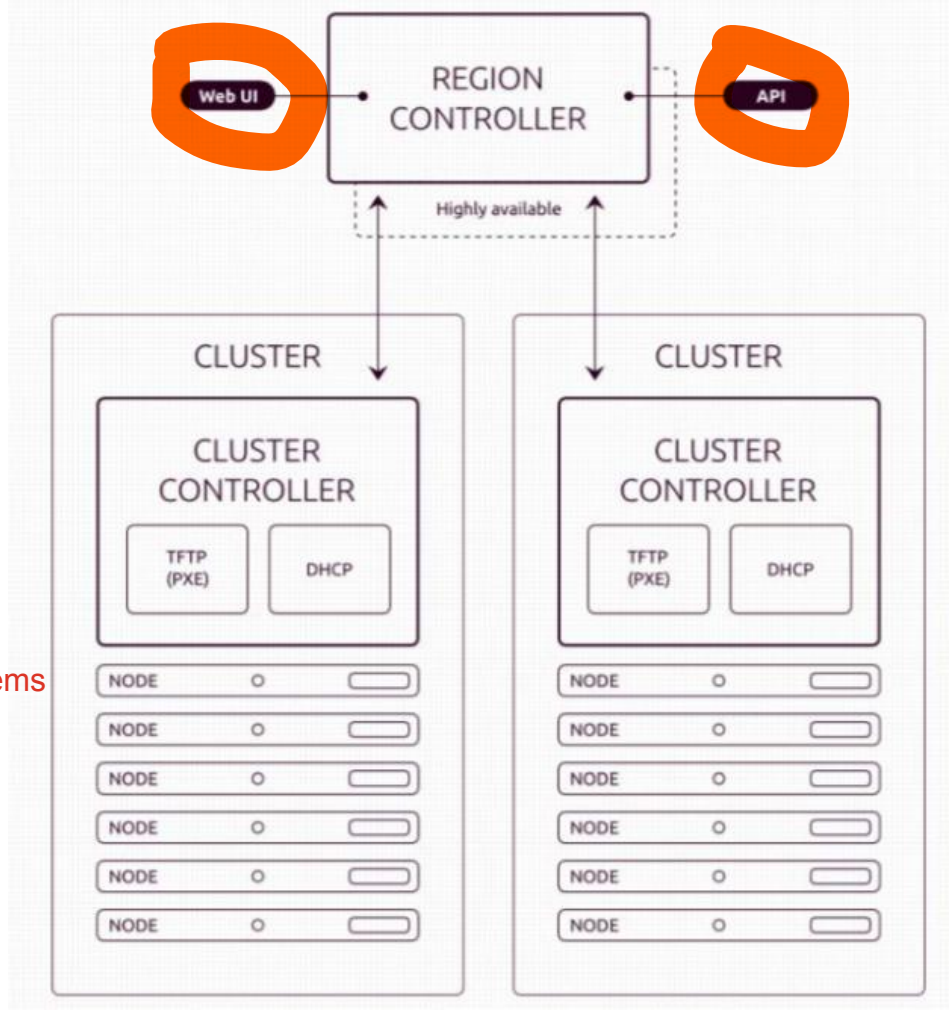
There are cloud security problems where in some cases it is not allowed to share the same hardware resources with other users

MaaS: Metal as a Service

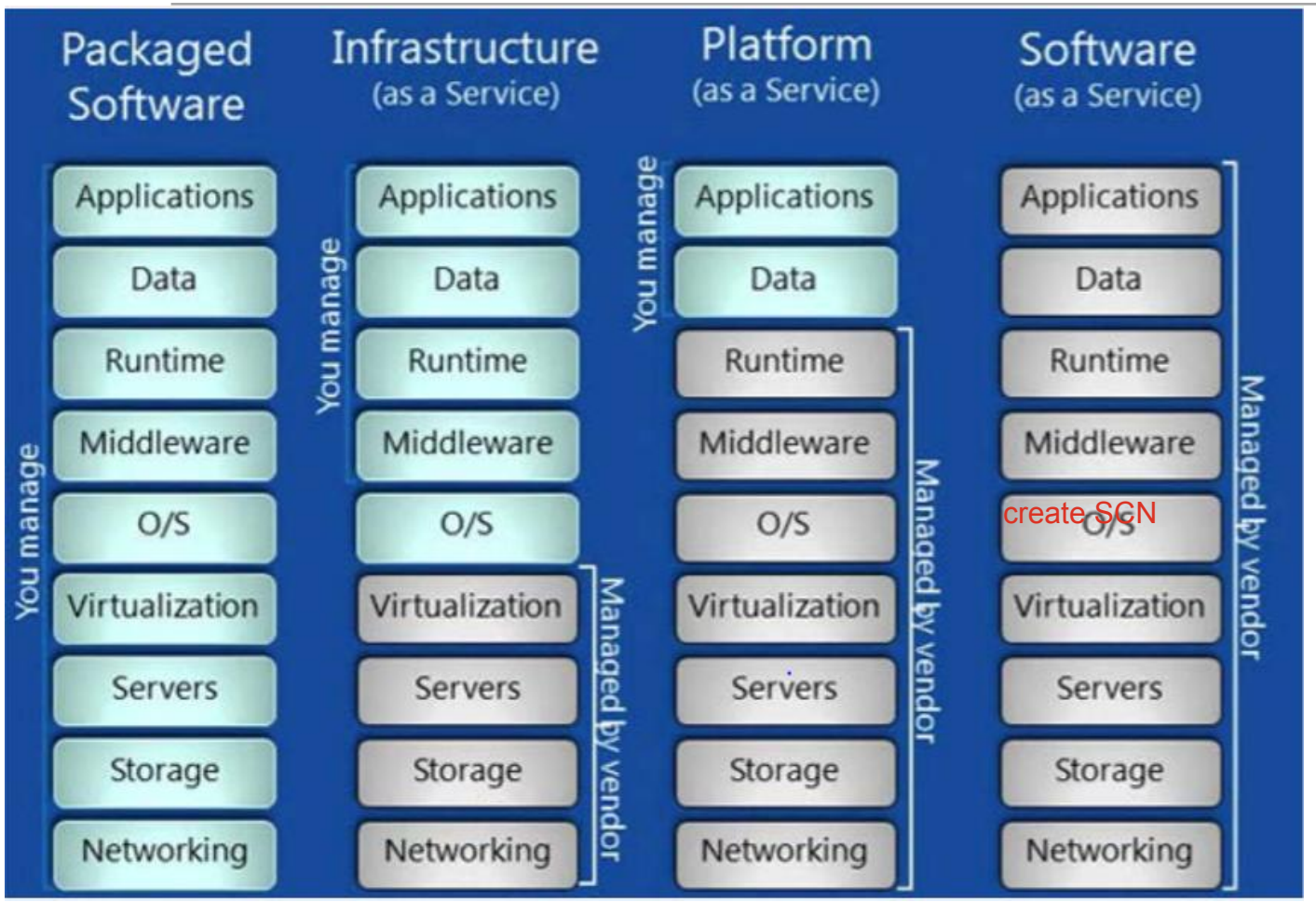
- Combines the flexibility and scalability of the cloud with the ability to harness the power of physical servers.
- **Example:** Juju

- For more information, watch this video (optional): https://www.youtube.com/watch?time_continue=280&v=FBCKCO45xIw

for users with security and privacy concerns, this gives more granularity with the systems



Cloud Computing Service Model



In MAAS

- You will have the option to control everything!

for quick production

PaaS or IaaS?

Vendor Lock-in: the ability to use “what you manage” in cloud environment with different cloud provider.

- PaaS may lock-in applications by requiring users to develop apps based on their specific APIs.
- If you are using PaaS, it might be difficult to switch to different vendor.

Development Tools

- PaaS providers usually allow a set of development tools for their users to shorten development time.
- Another trick for vendor lock-in!

Cloud Computing Enablers

Data Center + Virtualization

(Hardware) (Software)

create SCN that can be provisioned to users

Next Steps

- Complete Course Entry Survey
  - Link can be found on Canvas under “Modules” section.
- Join the course Piazza.
- Join the student Slack workspace.
- Familiarize yourself with the location of the in-person OHs
- Check Homework-1 PDF
  - Waitlisted students will receive a Google Form on September 3rd to submit their HW-1 solution.

Waitlisted Students

- All materials for first two weeks will be uploaded here