Carnegie Mellon University

14-848 Cloud Infrastructure: Design, Analysis and Implementation

Meeting Dates/Times: Mon/Wed 7 - 8:20 PM ET (4 – 5:20 PM PT)
Location: CIC 1201 for PIT Students & B23 118 for SV Students
Course Webpage: https://canvas.cmu.edu/courses/42627

Semester: Fall**, Year:** 2024

Units: 12**, Sections:** A (Pittsburgh), SV (Silicon Valley)

Instructor

Name	Mohamed Farag (he/his/him) – Addressed as “Mohamed” or “Dr. Farag”
Contact Info	farag@cmu.edu
Office Hours Location	CMU Remote
Office Hours	Monday 11am - 12PM ET (8 - 9am PT) and Thursday 3 - 4PM ET (12 – 1 pm PT). Conducted remotely via Zoom (URL is available on course home page)

Course Description

- Cloud computing has emerged as a new paradigm for efficient and highly elastic delivery of computing services over the Internet to achieve economies of scale. It focuses on the delivery of services via on-demand and fluidly scalable shared resources. This course focuses on the design and implementation of networked systems and software necessary to implement the infrastructure for elastic, global-scale computing, and storage clouds. Upon successful completion of the course, students will be able to design, implement, and analyze the infrastructure underlying cloud and edge-based services. Topics covered in this course include data center networking, cloud-scale storage, virtualization, computing programming models, large-scale distributed computing, edge computing, resource utilization and sharing, and cloud service frameworks. The course material will focus on recent and landmark research papers and existing tools and software systems. Students will have substantial programming project work in which they design, implement, and analyze aspects of cloud infrastructure and services. Students are expected to be proficient in object-oriented programming and Linux system programming and command-line tools.

- The premise of this course is to build a broad and solid foundation in cloud computing that will pay significant dividends throughout a student’s research and work career across cloud computing and DevOps related fields. In this class, we will focus on five main themes:

    - Cloud computing concepts and programming paradigms
    - Cloud data storage
    - Cloud Deployment and Orchestration
    - Big data processing frameworks.
    - Miscellaneous topics related to cloud computing.

- **Prerequisites**: project course in software systems (e.g., networking, OS, etc.) and object-oriented programming experience should be sufficient
- **Class Structure:** Lectures will be synchronous, and attendance is required. Lecture slides are delivered via TopHat, an online course delivery system. TopHat can be accessed on a smartphone or laptop.

For this purpose, I will ask you in advance to bring your laptop or mobile device, and we will review how to use the tool together in class. If you do not have the necessary equipment, please contact your HUB liaison who is available to help you tap into appropriate resources. Join code is posted on the course Canvas page. Recordings will be made available to students for asynchronous access. Class meeting times will include some lectures, quizzes, discussion of course materials and ongoing project tasks, and other activities as appropriate. Class sessions may be partitioned/rotated as needed to accommodate reduced classroom capacity due to COVID-19 restrictions.

- This class doesn't restrict students to use specific programming language but the lecture materials target Python programming language. In general, comparable object-oriented programming experience is required and you will be asked to do several tasks in your favorite programming language to gain the valuable perspective that this class offers. If you have questions about your programming experience, please let me know at the beginning of the semester, so we can decide how to proceed.

Learning Objectives

In taking and successfully completing this course, students will:

- Understand the current cloud computing technology and critical technology trends that are enabling cloud computing, including cloud Computing architecture along with the services and the applications that cloud computing offers;
- Gain hands-on experience in building and analyzing cloud infrastructure components and developing and testing services that operate within them;
- Develop big-data processing applications to handle large loads of data; and
- Simplify the deployment and portability of large complex applications to promote continuous integration and continuous delivery concepts using Docker.

The work that you will carry out in this course maps to some of the learning outcomes that INI has defined as vital to the full CMU graduate-level experience. Specifically:

- self-directed learning
- critical thinking; and
- complex problem solving.

Learning Resources

There is no primary textbook, as most reading material will come from research papers and other technical documentation. Additional background reading material can be suggested upon request.

Important Dates

- September 2nd, 2024: Labor Day (no class/office hours)
- October 14th – October 18th, 2024: Fall break (no classes/office hours).
- November 5th, 2024: Democracy Day (no classes/office hours).
- November 27th– November 29th, 2024: Thanksgiving Break (no class/office hours)
- December 4th, 2024: Final exam (during lecture time)

Assessments

Students are encouraged to attend class regularly, read the assigned reading material and participate in class discussions. The final grade will be based upon 1 exam, 1 project, 7 homework assignments, and in-class quizzes.

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

- **Quizzes:** are offered during each lecture via Canvas. Each quiz will be accessible via a unique access code that will be provided to the students. Students will have 2-3 minutes to answer 1-2 multiple choice questions.
    - Please note that students are not allowed to share Quiz Access Code with their peers who didn't attend the class in-person without pre-approval from the instructor. Sharing the quiz access code, without instructor's approval, is considered an academic integrity violation.
- **Assignments:** will provide the opportunity to practice the concepts that are taught during the lectures. Students are expected to spend a good amount of time on their own to learn implementation details that are not provided during the lectures. Students will receive 7 assignments throughout the semester. Assignment schedule is shown on the last section of this syllabus. **While we will use Google Cloud Platform for most assignments, students will have to apply for trial version to use Azure and AWS. The trial version will require the students to enter their credit card information without being billed.**
- **Project**: details are released in week 3. Each student will have the option to choose another student for the project and you will choose one of two project options to submit. Students will be expected to record a video including a code-walkthrough of their work and functionality demo showing the running version of their application. Project submission deadline is **November 14th, 2023 11:59PM ET /8:59pm PT.** Course project will include a lot of self-learning that is needed from the students to complete the project. The course instructor will provide project-related hints, high-level directions, and clarifications during the lectures. However, students shouldn't expect any additional project support during office hours or via emails. Students are highly encouraged to give themselves enough time to learn the skills they need to complete the project. There will be a checkpoint to ensure that students are making good continuous progress (refer to the proposed course schedule in the last section of this syllabus). Project grading rubric and evaluation will be released along with project details.
- **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.**
    - **You can get full score 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
  - Late acquisition of these certificates beyond the deadline won't be accepted. Students who would like to waive their final exam must upload a proof of obtaining two certificates two weeks prior to the exam. Late submissions WILL NOT be accepted.
  - Obtaining only 1 certificate of the 3 won't provide any partial benefit on the final exam score

Students will be assigned the following final letter grades, based on rubric provided in the above table. +/- will be assigned in equal intervals to provide further granularity.

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

Homework & Grading Policies

- Students are expected to check the course webpage on canvas regularly for announcements, class schedules, lecture notes, homework assignments, reading assignments, and other related course material.
- Assignments will be submitted online via **GitHub Org URLs**. For this purpose, sign up for a (free) account at GitHub.
- Students are strongly encouraged to complete the assignments as early as they can. No homework extensions will be offered due to technical difficulties.
- **Students will have 3 days to submit an assignment after the due date** and a late penalty will be applied. Late penalties are applied based on the timestamp of the last code commit on GitHub and it will follow this equation:
  - 5 points for up to 24 hours delay
  - 15 points for the next 24 hours delay
  - 25 points for the next 24 hours delay
  - 100 points penalty (no grade) after this time.
- **Late submissions for the course project will receive no grade (0 points).**
- After homework and project grades are released, a Canvas announcement will be made with a link to submit regrade requests. **Regrade requests can be made for 24 hours via the URL** that is provided on the Canvas announcement and **CANNOT be submitted via email**.
- For grading questions, email the TA or ask your question during TA office hours. If you continue to have an issue with the grading, you may ask your question during instructor's office hours.
- Students are encouraged to read their privacy rights in Family Educational Rights and Privacy Act (FERPA).

Guidelines for Office Hours and Out-of-class Questions

- Students are encouraged to leverage office hours to get the support they need. Due to the current COVID-19 situations, we will continue to conduct office hours via Zoom. Refer to the course page for details.

- If you have materials-related questions, you can post them on **Piazza** or ask them during instructor's office hours. You should plan to receive the question responses on **Piazza** during office hours' allocated time.
- You may email personal inquiries and severe emergencies to the instructor(s). The email subject line should begin with **"14-848"**. Emails sent to the instructor should be 2-3 lines maximum and the instructor's response will not be more than few words. Generally, the course instructor is available to respond to personal inquiries during M-F 9am-8pm ET. Please don't expect responses over the weekend or late at night.

Expectations for Class Attendance

- Classes will be offered **synchronously and in-person.**
- Class attendance and participation are important parts of the learning in this course. To account for this, a portion of the final grade is based on quizzes that are offered during the lectures (see assessment section). That said, I also recognize that students may need to miss class for a variety of reasons (religious observance, job interview, university-sanctioned event, or illness). **For that reason, all students are permitted two class absences without any impact on the final grade.** If you encounter extenuating circumstances and must miss more than two classes, please email the course instructor for further guidance.
- When attending the class in-person, I expect that you will abide by all behaviors indicated in A Tartan's Responsibility, including any timely updates based on the current conditions.

Recording of Class Sessions

- Class recordings will be available after each lecture. Please note it may take a few hours for the recording to become available. The Class recordings will be published on Canvas.

Academic Integrity

- Discussing assignments with your classmates is allowed and encouraged, but it is important that every student get practice working on these problems. This means that all the work you turn in must be your own. You must devise and write your own solutions and carry out your own tests. The general policy on homework collaboration is:
    - You must first make a serious effort to solve the problem.
    - If you are stuck after doing so, you may ask for help from another student. You may discuss strategies to solve the problem, but you may not look at their code, nor may they spell out the solution to you step-by-step.
    - Once you have gotten help, you must write your own solution individually. You must disclose, in your GitHub pull request, the names of anyone you got help from.
    - This also applies in reverse: if someone approaches you for help, you must not provide it unless they have already attempted to solve the problem, and you may not share your code or spell out the solution step-by-step.
    - These rules also apply to getting help from other people: friends not in the course, homework help websites, Stack Overflow, and so on.
    - You can always, of course, ask for help from the course instructors.
- You may also use external sources (books, websites, papers, …) to:
    - Look up programming language documentation, find useful packages, find explanations for error messages, or remind yourself about the syntax for some feature,
    - Read about general approaches to solving specific problems (e.g. a guide to dynamic programming or a tutorial on unit testing in your programming language), or
    - Clarify material from the course notes or assignments.
    - But external sources must be used to support your solution, not to obtain your solution. You may not

use them to
    - Find solutions to the specific problems assigned as homework (in words or in code)—you must independently solve the problem assigned, not translate a solution presented online or elsewhere.
    - Find course materials or solutions from this or similar courses from previous years, or
    - Copy text or code to use in your submissions without attribution.
    - If you use code from online or other sources, you must include code comments identifying the source. It must be clear what code you wrote and what code is from other sources. This rule also applies to text, images, and any other material you submit.
- Students may use generative-AI platforms (e.g., ChatGPT) to assist them with a portion of the homework solution. However, 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.

Please talk to the instructor if you have any questions about this policy. Any form of plagiarism or cheating will result in sanctions to be determined by the instructors, including grade penalties (such as negative points for the assignment or reductions in letter grade) or course failure. Students taking the course pass/fail may have this status revoked. We are also obliged in report violations to your academic program and the appropriate University authorities. Please refer to the University Policy on Academic Integrity.

Student Wellness

- The last few years have been challenging. We are all under a lot of stress and uncertainty at this time. I encourage you to find ways to move regularly, eat well, and reach out to your support system or me (farag@cmu.edu) if you need to. We can all benefit from support in times of stress, and this semester is no exception.

Diversity Statement

- **We must treat every individual with respect**. We are diverse in many ways, and this diversity is fundamental to building and maintaining an equitable and inclusive campus community. Diversity can refer to multiple ways that we identify ourselves, including but not limited to race, color, national origin, language, sex, disability, age, sexual orientation, gender identity, religion, creed, ancestry, belief, veteran status, or genetic information. Each of these diverse identities, along with many others not mentioned here, shape the perspectives our students, faculty, and staff bring to our campus. We, at CMU, will work to promote diversity, equity and inclusion not only because diversity fuels excellence and innovation, but because we want to pursue justice. We acknowledge our imperfections while we also fully commit to the work, inside and outside of our classrooms, of building and sustaining a campus community that increasingly embraces these core values.
- Each of us is responsible for creating a safer, more inclusive environment.
- Unfortunately, incidents of bias or discrimination do occur, whether intentional or unintentional. They contribute to creating an unwelcoming environment for individuals and groups at the university. Therefore, the university encourages anyone who experiences or observes unfair or hostile treatment on the basis of identity to speak out for justice and support, within the moment of the incident or after the incident has passed. Anyone can share these experiences using the following resources:
    - **Center for Student Diversity and Inclusion**: csdi@andrew.cmu.edu, (412) 268-2150
    - Report-It **online anonymous reporting platform**: reportit.net username: tartans password: plaid
- All reports will be documented and deliberated to determine if there should be any following actions. Regardless of incident type, the university will use all shared experiences to transform our campus climate to be more equitable and just.

Food Insecurity

- If you are worried about affording food or feeling insecure about food, there are resources on campus that can help. Any undergraduate or graduate student can visit the CMU Pantry and receive food for free. Follow the directions on the CMU Pantry website to schedule your visit.

Disability Resources

- If you have a disability and have an accommodations letter from the Disability Resources office, we encourage you to discuss your accommodations and needs with us as early in the semester as possible. We will work with you to ensure that accommodations are provided as appropriate. If you suspect that you may have a disability and would benefit from accommodations but are not yet registered with the Office of Disability Resources, we encourage you to contact them at access@andrew.cmu.edu.

Religious Observances

- In order to accommodate the observance of religious holidays, students should inform the instructor by email, within the first two weeks of the term, of any such days which conflict with scheduled class activities.

Student Academic Success Center

SASC programs to support student learning include the following (program titles link to webpages):

- Academic Coaching: This program provides holistic, one-on-one peer support and group workshops to help undergraduate and graduate students implement habits for success. Academic Coaching assists students with time management, productive learning and study habits, organization, stress management, and other skills. Request
- an initial consultation here.
- Peer Tutoring: Peer Tutoring is offered in two formats for students seeking support related to their coursework. Drop-In tutoring targets our highest demand courses through regularly scheduled open tutoring sessions during the fall and spring semesters. Tutoring by appointment consists of ongoing individualized and small group sessions. You can utilize tutoring to discuss course related content, clarify and ask questions, and work through practice problems.
- Communication Support: Communication Support offers free one-on-one communication consulting as well as group workshops to support strong written, oral, and visual communication in texts including IMRaD and thesis-driven essays, data-driven reports, oral presentations, posters and visual design, advanced research, application materials, grant proposals, business and public policy documents, data visualisation, and team projects. Appointments are available to undergraduate and graduate students from any discipline at CMU. Schedule an appointment (in-person or video), attend a workshop, or consult handouts or videos to strengthen communication skills.
- Language and Cross-Cultural Support: This program supports students seeking help with language and cross-cultural skills for academic and professional success through individual and group sessions. Students can get assistance with writing academic emails, learning expectations and strategies for clear academic writing, pronunciation, grammar, fluency, and more. Make an appointment with a Language Development Specialist to get individualized coaching.
- Supplemental Instruction (SI): This program offers a non-remedial approach to learning in historically difficult courses at CMU. It utilizes a peer-led group study approach to help students succeed and is facilitated by an SI leader, a CMU student who has successfully completed the course. SI offers a way to connect with other students studying the same course, a guaranteed weekly study time that reinforces learning and retention of information, as well as a place to learn and integrate study tools and exam techniques specific to a course.

Mental Health Resources

- CaPS has partnered with TimelyCare for provision of virtual well-being services, including immediate emotional support 24/7 as frequently as needed, scheduled appointments with therapists that can be chosen by identity group and other features, health coaching (e.g., sleep issues, weight management, etc.), and group sessions for things like yoga, meditation, etc.

Preliminary Course Schedule (Subject to Change)

Date	Topic	Notes
Week-1 (Aug. 26th)	- Introduction & Syllabus - Virtualization Basics	- System Setup homework released
Week-2 (Sep. 2nd)	- Containerization	- System Setup homework deadline. - Docker homework released
Week-3 (Sep. 9th)	- Lab: Containerization - Deployment Orchestration - Pokémon Go Case Study	- Course Project released
Week-4 (Sep. 16th)	- Lab: Deployment Orchestration - Cloud Computing Concepts	- Docker homework deadline - Kubernetes homework is released
Week-5 (Sep. 23rd)	- Cloud Computing Paradigms - Kafka	
Week-6 (Sep. 30th)	- Lab: Confluent Kafka - Infrastructure-as-a-Code	- Kubernetes homework deadline - Kafka Homework released
Week-7 (Oct. 7th)	- Infrastructure-as-a-Code (Cont'd) Terraform - Lab: Terraform	- Kafka Homework deadline - Course Project Checkpoint
Fall Break (Oct. 14th - Oct. 18th)		
Week-8 (Oct. 21st)	- Introduction to Hadoop - Hadoop HDFS	- Terraform Homework released
Week-9 (Oct. 28th)	- Hadoop MapReduce	- Terraform Homework deadline - Hadoop MapReduce homework is released.
Week-10 (Nov. 4th)	- Big Data Algorithms - Introduction to Spark	- Hadoop MapReduce homework deadline
Week-11 (Nov. 11th)	- Spark (Cont'd) - Lab: Spark Programming	- Course project submission deadline
Week-12 (Nov. 18th)	- Cloud Data Storage Models - Lab: NoSQL Database - Neo4j AuraDB	- Apache Spark homework released
Week-13 (Nov. 25th)	- Cloud Security & Privacy - Metaverse, Edge Computing and Fog Computing	- Apache Spark homework deadline
Week-14 (Dec. 2nd)	- Introduction to GenAI on the Cloud using Vertex AI - Final Exam	