Advanced Topics in Serverless and Edge Computing

CS 595 – Fall ‘22

Logistics

- Lecture
  - Tuesday/Thursday 3:35-4:50PM (JREC 241)
- There is no TA
- Office hours: Friday 2-3PM

Course Prerequisites

- This is an advanced graduate course
- I expect you to have taken or have experience with:
  - Computer Systems (350, 351)
  - Intro OS (450)
- Nice to have, but not necessary:
  - Distributed Systems (550)
  - Cloud Computing (553)
  - Computer Networking

It is my job to make sure that you

- Work on open-ended research projects
- Read (many) papers
- Learn how to think critically and dive deep into technical material we may know little about
- Learn how to present our research
- Learn how to critique research papers
- Become familiar with Edge computing and serverless

I will *not*

- Give exams
- Lecture
- Supply you with answers
- Hold your hand

Your job

- Think critically
- Work hard
- Manage your schedule (especially reading)
- Get excited about research
- Be welcoming and considerate

Course Philosophy

We're going to act like a research group:

- We are a group of peers
- No one is the expert
- Everyone's input is valuable
- We're trying to explore fun things and solve challenging problems
- Work together towards a common goal

Project

- Open-ended research
- Will likely involve a lot of coding
- Possibly also measurement and experimentation
- Current list of potential projects is up on course website, but you can also bring your own (upon approval)
- You'll pick projects this week, write a project proposal next week

Reading

- We'll be reading ~2 papers a week
- You will write a review for *one* of those papers every week (due Fridays, which paper is your choice)

Grades

- Research Project: 50%
- Project paper and presentation: 20%
- Discussion:
  - Participation: 10%
  - Paper reviews: 10%
  - Presentations: 10%

Daily Agenda

- 5 min project status updates from everyone
- 20-30 min paper presentation
- 30-45 min paper discussion
- Each of you will hopefully present ~2 papers throughout the semester
- Last class will run like a research workshop

Presentation Guidelines

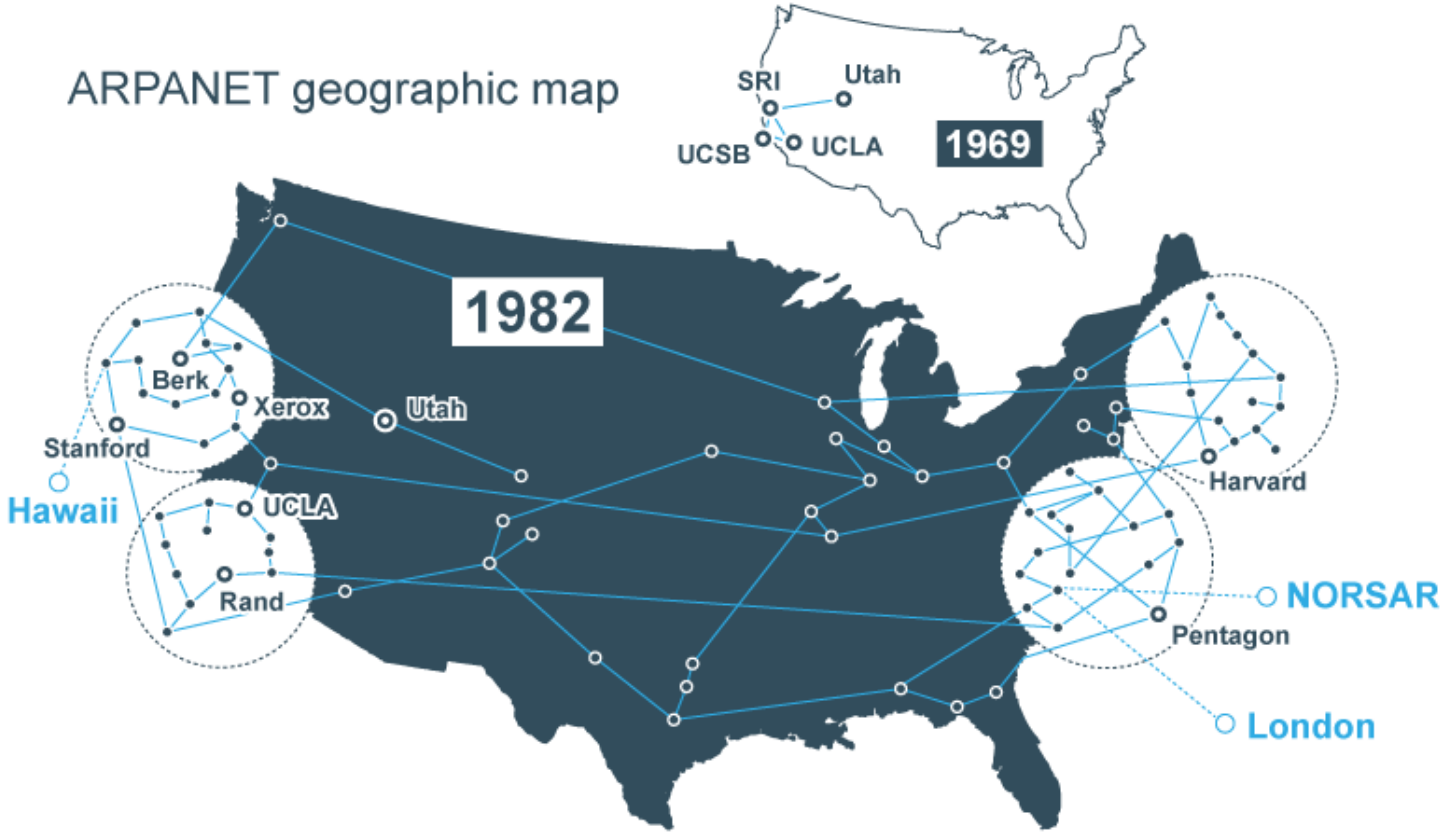
- ~20-30 min
- Can use existing author slides of available
- Summarize key points, contributions, and findings (no walls of text)
- Many papers have author presentation videos online: use them as a guide

Course Management

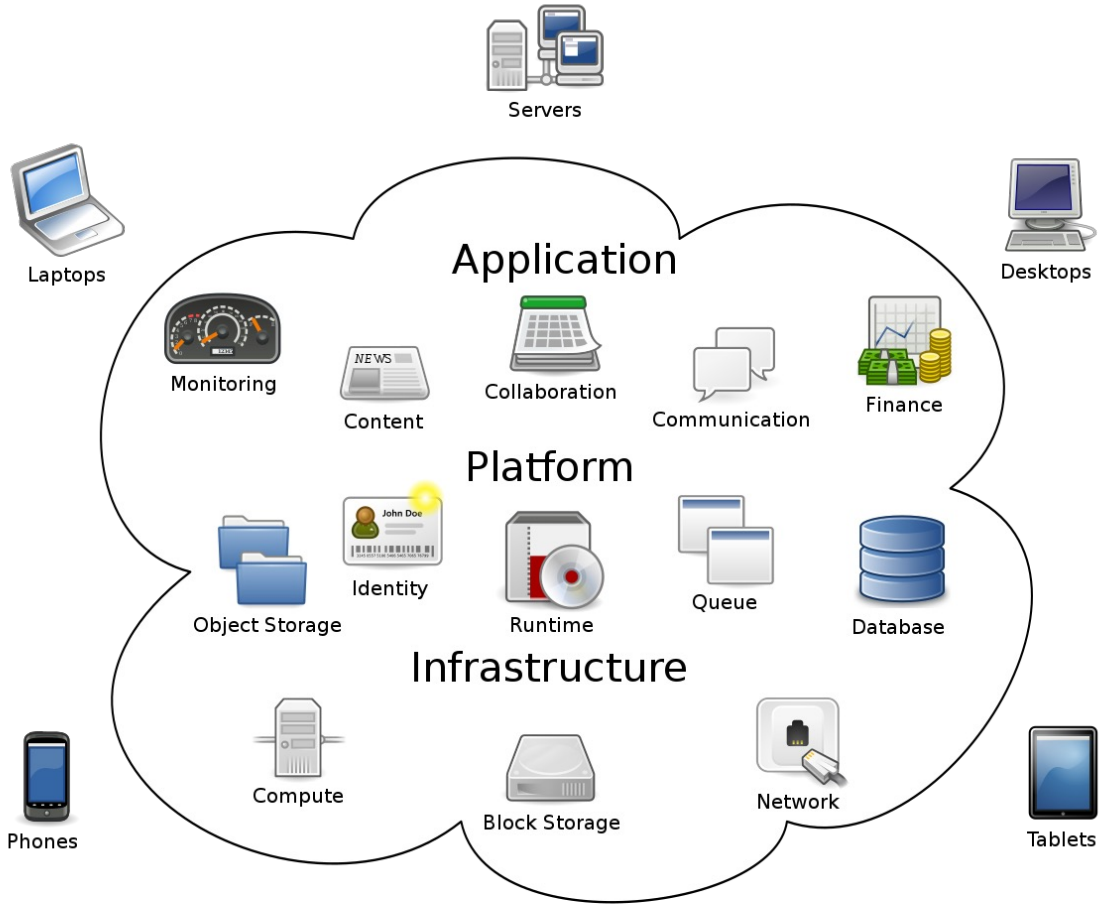
- Communication: Discord (link is on Blackboard)

Computing 1.0: Centralized Processing

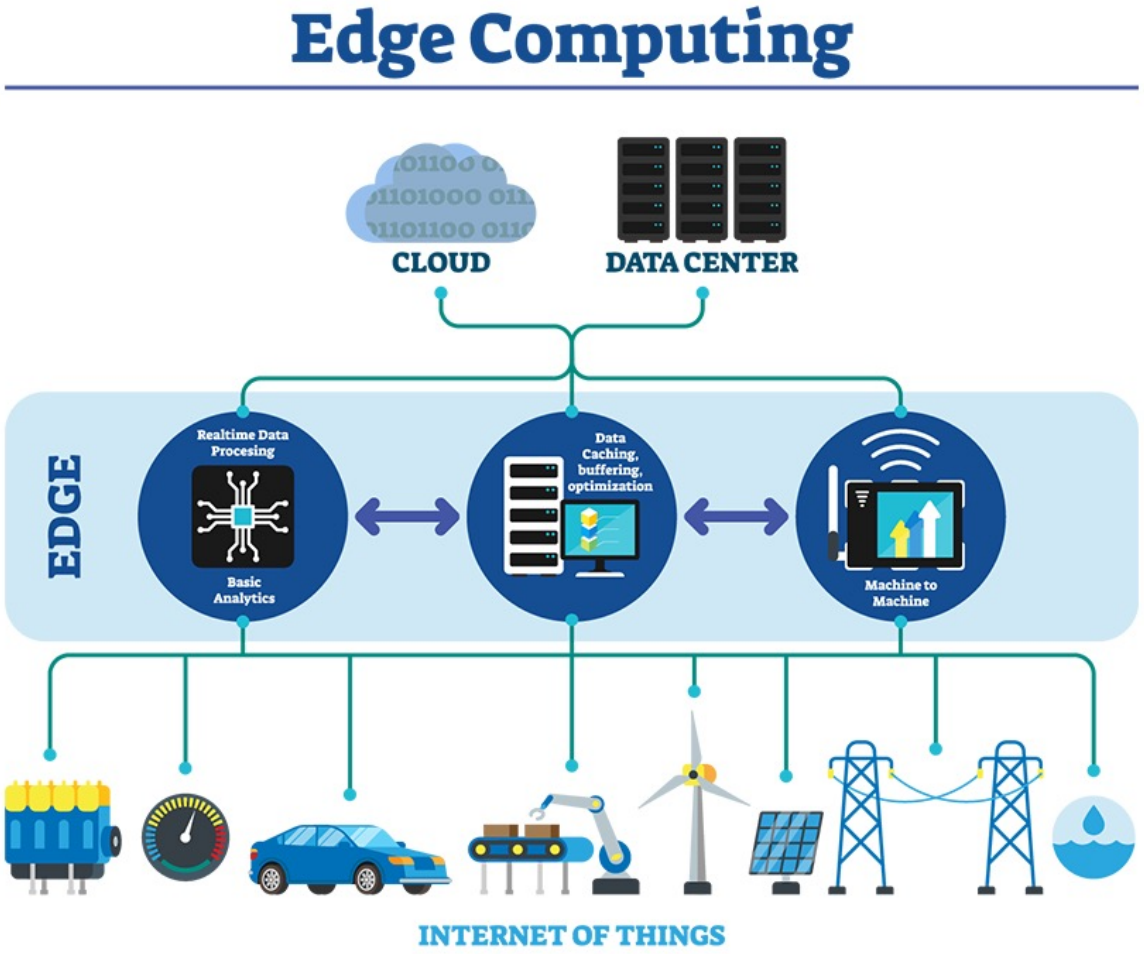
Computing 2.0: Distributed Computing



Computing 3.0: Centralized Cloud



Computing 4.0? Decentralized Edge

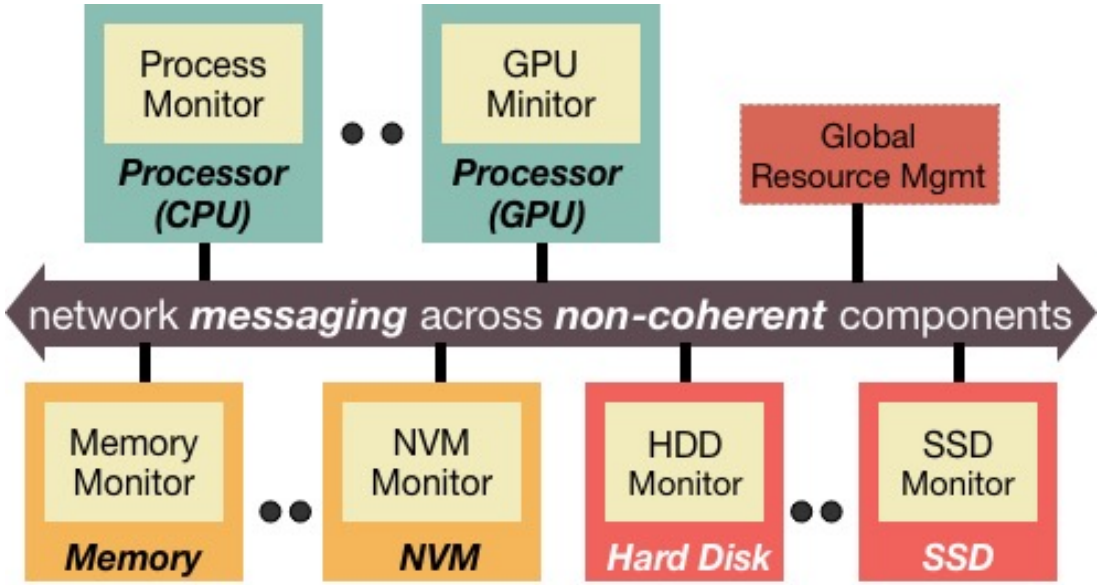


Cloud Provisioning → Serverless

- Trend has been towards:
  - Finer granularity
  - More control
  - Decreased cost
- First we had VMs tied to physical machines
- Then to subsets of resources (cores, memory) *within* those machines
- Moving more towards *what to run* and away from *what to run it on*
- *Function as a Service (FaaS):*
  - `on event X: run function Y`
- Infrastructure is invisible, software stack is implicit

Resource Disaggregation

- Networks are faster
- A lot of unutilized resources in cloud
- Can we consolidate resources by using them remotely?



First Tasks

- Form a group (no more than 2)
- The project list will be out tomorrow: think about ones you'd like
- Read *how to read*
- Volunteer for paper presentation slots