# Cloud Computing

DA 410/510

Spring 2026

# Last class!

- Assignment 5
- Course themes
- Next steps
- The final exam

# Assignment 5

- Deadline extended to Monday.
  - Meaningful extra credit for submitting by original deadline.
- Some people have already turned their work in
  - So far, it’s all been great!!!

# Course Themes

- Virtualization
  - Cloud systems use virtual machines and containers to enable renting resources at scale.
- Automation
  - Automating deployment of resources is strategically smart.
- System Architecture
  - We've seen how cloud technologies can support the design of useful and interesting systems.

# Where to go from here…

- More automation
- More system design
- More elasticity
- More security
- More monitoring

- Our new Linux Lab
- Also, other cloud systems.
  - Lambda Labs is a good one to look at for GPU machines.

# More Automation: Scripting

- We didn’t explore this in depth, but you can script every interaction you’ve had with AWS:

```
aws ec2 run-instances \
  --region us-east-1 \
  --image-id resolve:ssm:/aws/service/ami-amazon-linux-latest/al2023-ami-kernel-default-x86_64 \
  --instance-type t3.micro \
  --count 1 \
  --key-name my-key-pair \
  --security-group-ids sg-0123456789abcdef0 \
  --subnet-id subnet-0123456789abcdef0 \
  --tag-specifications 'ResourceType=instance,Tags=[{Key=Name,Value=small-demo-instance}]'
```

# More Automation: Scripting

- You can check status via the terminal:

```
aws ec2 describe-instances \
  --filters "Name=tag:Name,Values=small-demo-instance" \
  --query "Reservations[].Instances[].{ID:InstanceId,State:State.Name,PublicIP:PublicIpAddress}" \
  --output table
```

- And terminate instances as well:

```
aws ec2 terminate-instances --instance-ids i-0123456789abcdef0
```

- This means that it’s easy to script – for humans and AI
- **AI agents** will **almost certainly** be responsible for this work soon.

# More System Architecture

- Study systems:
  - https://aosabook.org/en/
- Designing Data-Intensive Applications
  - https://www.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/
  - This is a good book, especially if you want to get into AI-heavy development.

# More elasticity

- One of the big advantages of cloud is ***elasticity***.
  - Make resources start and stop automatically in response to load.
- We haven't focused on it much because of resource and practical limitations.
- If you're interested, these are things to explore:
  - EC2 Autoscaling
  - Elastic Load Balancing
  - Serverless

# More Security and Monitoring

- AWS provides for fine-grained identity management and access control
  - See here for details.
- As you scale a cloud system, monitoring becomes increasingly important and eventually essential for the operation to survive.
  - Prometheus is a good tool to look at.

# What can you build?

- You can build web frontends with nginx
- You can build APIs
  - Secure REST-based servers (FastAPI)
  - Secure API access with header-based tokens
- You can deploy a database-backed system that’s secure
- You can support users and logins
- You can support LLM-based servers
- This is pretty much everything you need to create a modern software-based product.
  - We didn’t cover mobile frontends, but that’s just an app that talks to the same REST API servers as web.

# What ***should*** you build?

- AI makes building things pretty easy.
- The hard question becomes ***what*** to build.
- You should think about this

# Final Exam Format

- Take-home exam: design a cloud system for a novel product.
  - Should take an hour.
  - Focus on analytical thinking, writing, and creativity.
- Released on Monday at noon.
- **Due by Wednesday, May 6 at MIDNIGHT**
- **NO LATE EXAMS WILL BE ACCEPTED!!!**