Addison Wurtz
CS 530
notebooks/Week10

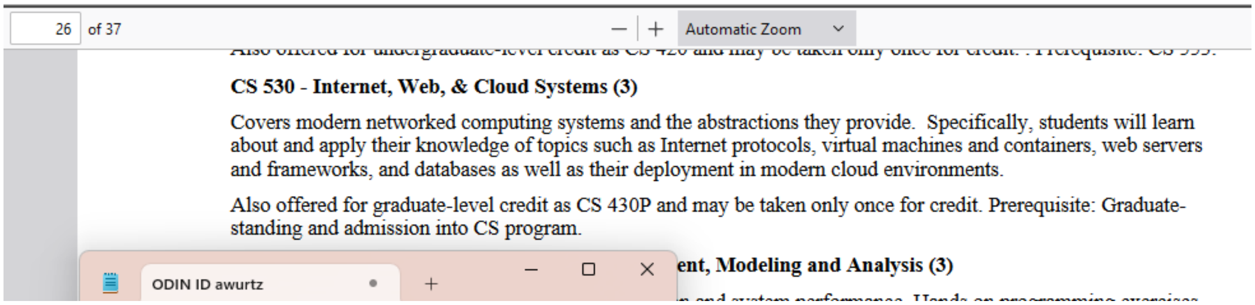
Lab Notebook – Week 10

Table of Contents:

Addison Wurtz
CS 530
notebooks/Week10

10.1g: LLMs

Walk through notebook



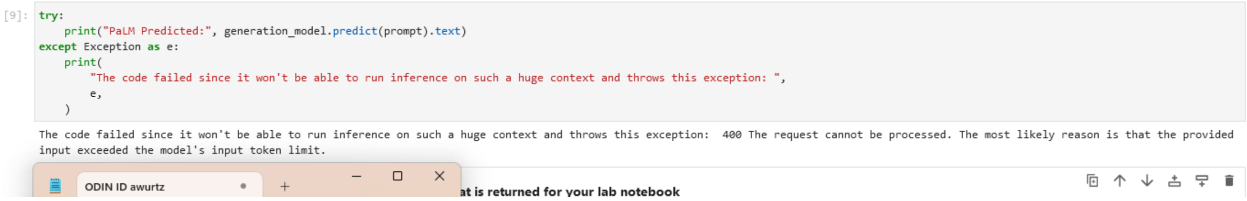
CS 530 - Internet, Web, & Cloud Systems (3)

Covers modern networked computing systems and the abstractions they provide. Specifically, students will learn about and apply their knowledge of topics such as Internet protocols, virtual machines and containers, web servers and frameworks, and databases as well as their deployment in modern cloud environments.

Also offered for graduate-level credit as CS 430P and may be taken only once for credit. Prerequisite: Graduate-standing and admission into CS program.

ent, Modeling and Analysis (3)

Method 1: Stuffing



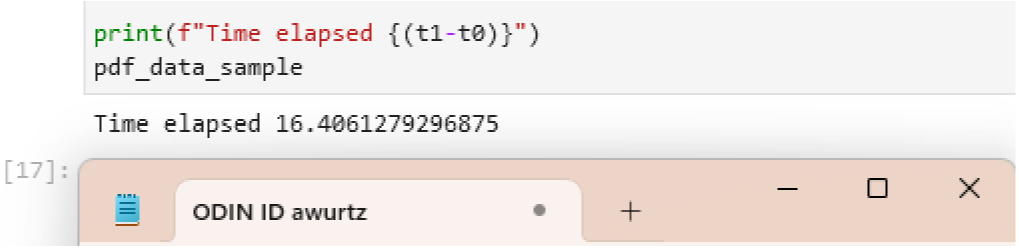
```
[9]: try:
         print("PaLM Predicted:", generation_model.predict(prompt).text)
     except Exception as e:
         print(
             "The code failed since it won't be able to run inference on such a huge context and throws this exception: ",
             e,
         )
```

```
The code failed since it won't be able to run inference on such a huge context and throws this exception:  400 The request cannot be processed. The most likely reason is that the provided
input exceeded the model's input token limit.
```

Provide an explanation as to why the description is not returned for your lab notebook.

The course description for CS530 is not in the context since it is limited to the first 5000 tokens.

Method 2: Map Reduce



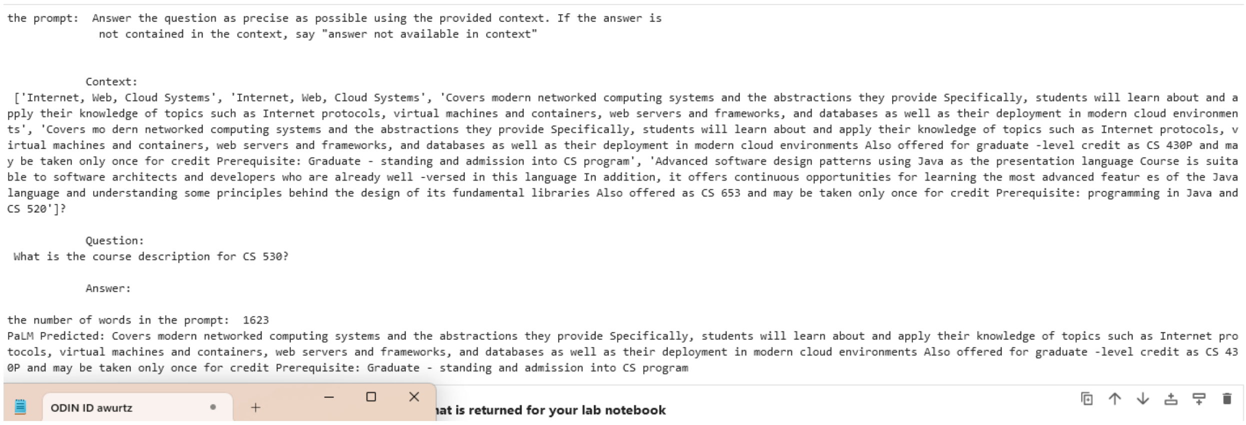
```
print(f"Time elapsed {(t1-t0)}")
pdf_data_sample
```

```
Time elapsed 16.4061279296875
```

Addison Wurtz
CS 530
notebooks/Week10

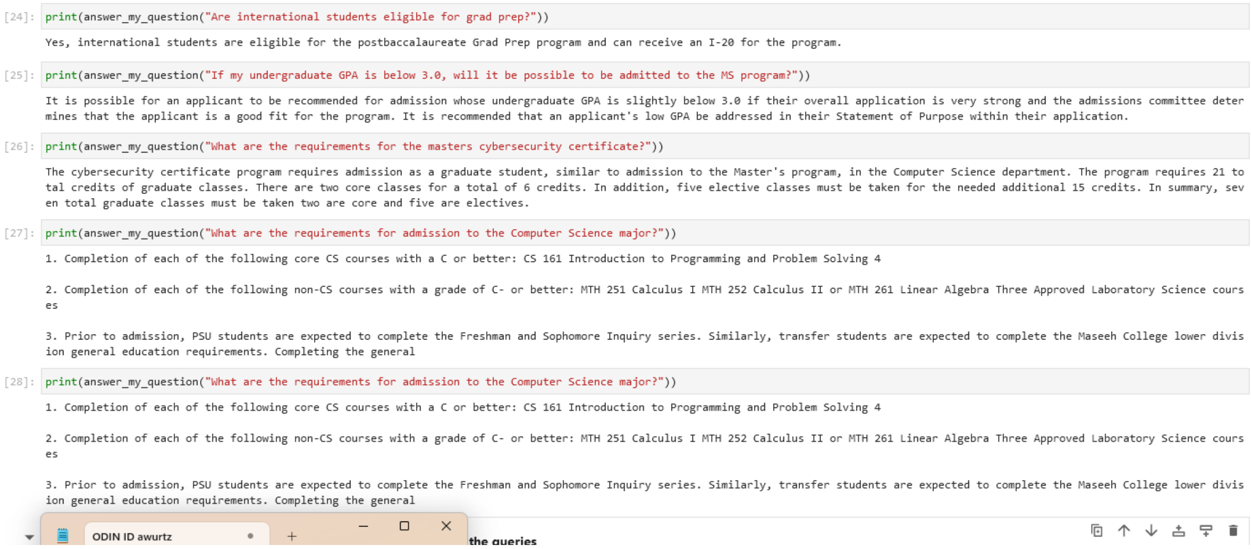
How many chunks returned predictions?

Five chunks returned predictions.



```
the prompt:  Answer the question as precise as possible using the provided context. If the answer is
              not contained in the context, say "answer not available in context"


            Context:
 ['Internet, Web, Cloud Systems', 'Internet, Web, Cloud Systems', 'Covers modern networked computing systems and the abstractions they provide Specifically, students will learn about and a
pply their knowledge of topics such as Internet protocols, virtual machines and containers, web servers and frameworks, and databases as well as their deployment in modern cloud environmen
ts', 'Covers mo dern networked computing systems and the abstractions they provide Specifically, students will learn about and apply their knowledge of topics such as Internet protocols, v
irtual machines and containers, web servers and frameworks, and databases as well as their deployment in modern cloud environments Also offered for graduate -level credit as CS 430P and ma
y be taken only once for credit Prerequisite: Graduate - standing and admission into CS program', 'Advanced software design patterns using Java as the presentation language Course is suita
ble to software architects and developers who are already well -versed in this language In addition, it offers continuous opportunities for learning the most advanced featur es of the Java
language and understanding some principles behind the design of its fundamental libraries Also offered as CS 653 and may be taken only once for credit Prerequisite: programming in Java and
CS 520']?

            Question:
 What is the course description for CS 530?

            Answer:

the number of words in the prompt:  1623
PaLM Predicted: Covers modern networked computing systems and the abstractions they provide Specifically, students will learn about and apply their knowledge of topics such as Internet pro
tocols, virtual machines and containers, web servers and frameworks, and databases as well as their deployment in modern cloud environments Also offered for graduate -level credit as CS 43
0P and may be taken only once for credit Prerequisite: Graduate - standing and admission into CS program
```

Method 3: Map Reduce with embeddings



```
[24]: print(answer_my_question("Are international students eligible for grad prep?"))

      Yes, international students are eligible for the postbaccalaureate Grad Prep program and can receive an I-20 for the program.

[25]: print(answer_my_question("If my undergraduate GPA is below 3.0, will it be possible to be admitted to the MS program?"))

      It is possible for an applicant to be recommended for admission whose undergraduate GPA is slightly below 3.0 if their overall application is very strong and the admissions committee deter
      mines that the applicant is a good fit for the program. It is recommended that an applicant's low GPA be addressed in their Statement of Purpose within their application.

[26]: print(answer_my_question("What are the requirements for the masters cybersecurity certificate?"))

      The cybersecurity certificate program requires admission as a graduate student, similar to admission to the Master's program, in the Computer Science department. The program requires 21 to
      tal credits of graduate classes. There are two core classes for a total of 6 credits. In addition, five elective classes must be taken for the needed additional 15 credits. In summary, sev
      en total graduate classes must be taken two are core and five are electives.

[27]: print(answer_my_question("What are the requirements for admission to the Computer Science major?"))

      1. Completion of each of the following core CS courses with a C or better: CS 161 Introduction to Programming and Problem Solving 4

      2. Completion of each of the following non-CS courses with a grade of C- or better: MTH 251 Calculus I MTH 252 Calculus II or MTH 261 Linear Algebra Three Approved Laboratory Science cours
      es

      3. Prior to admission, PSU students are expected to complete the Freshman and Sophomore Inquiry series. Similarly, transfer students are expected to complete the Maseeh College lower divis
      ion general education requirements. Completing the general

[28]: print(answer_my_question("What are the requirements for admission to the Computer Science major?"))

      1. Completion of each of the following core CS courses with a C or better: CS 161 Introduction to Programming and Problem Solving 4

      2. Completion of each of the following non-CS courses with a grade of C- or better: MTH 251 Calculus I MTH 252 Calculus II or MTH 261 Linear Algebra Three Approved Laboratory Science cours
      es

      3. Prior to admission, PSU students are expected to complete the Freshman and Sophomore Inquiry series. Similarly, transfer students are expected to complete the Maseeh College lower divis
      ion general education requirements. Completing the general
```

Final questions and clean-up

Which of the approaches described would have issues with token limits on LLMs?

Stuffing

Addison Wurtz
CS 530
notebooks/Week10

Which of the approaches would result in the most queries for the LLM to handle? How many LLM requests are performed from a single user query in this approach?

Map Reduce. It performs a query for each of the N chunks plus a final query using the relevant chunks (so N+1).

Which of the approaches requires one to search a vector database for an appropriate context that is then sent to the LLM?

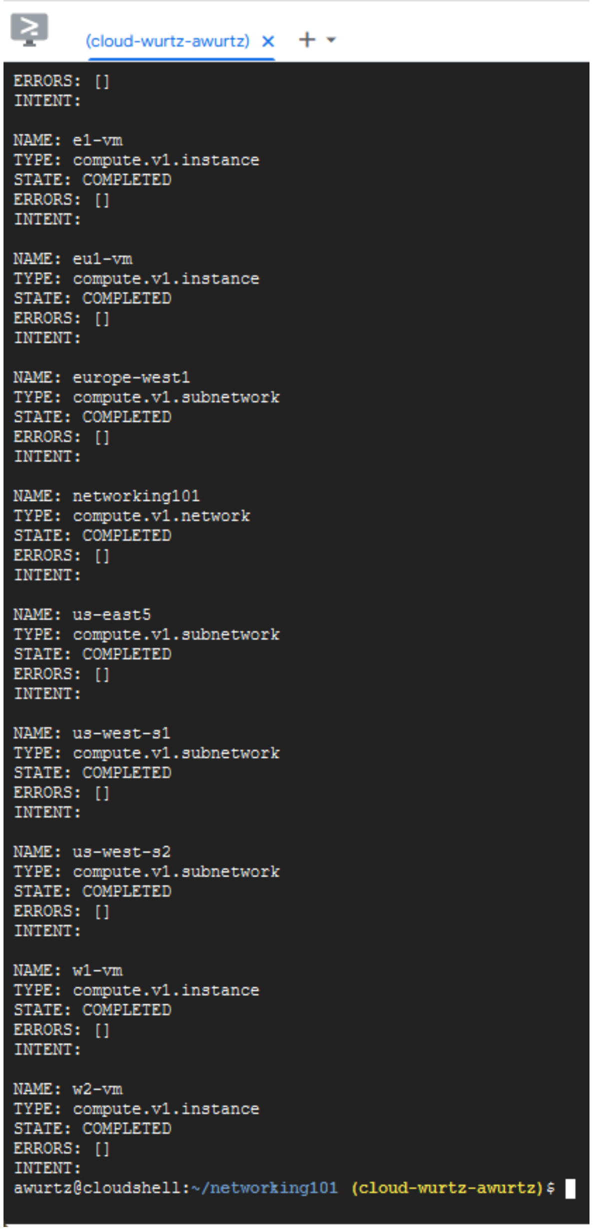
Map Reduce with embeddings.

Addison Wurtz
CS 530
notebooks/Week10

10.2g: CDN

Deployment

Take a screenshot of the output to include in your lab notebook. How many networks, subnetworks, and VM instances have been created?

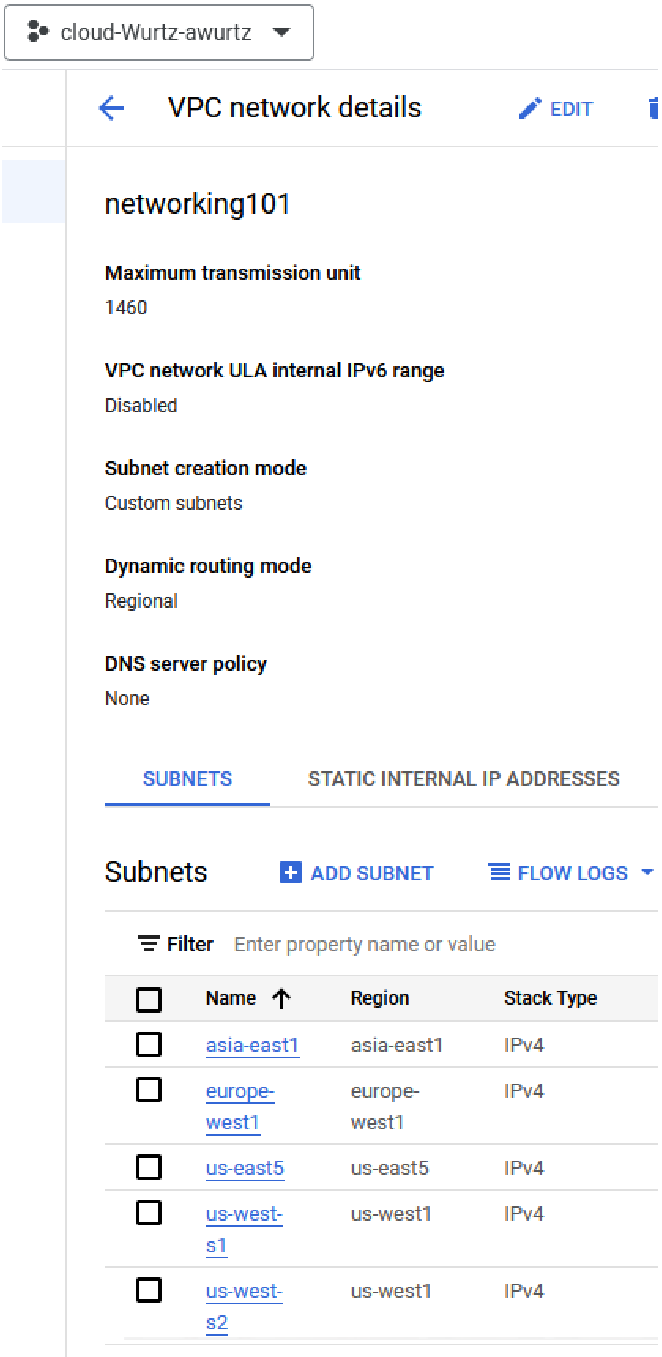


```
ERRORS: []
INTENT:

NAME: e1-vm
TYPE: compute.v1.instance
STATE: COMPLETED
ERRORS: []
INTENT:

NAME: eu1-vm
TYPE: compute.v1.instance
STATE: COMPLETED
ERRORS: []
INTENT:

NAME: europe-west1
TYPE: compute.v1.subnetwork
STATE: COMPLETED
ERRORS: []
INTENT:

NAME: networking101
TYPE: compute.v1.network
STATE: COMPLETED
ERRORS: []
INTENT:

NAME: us-east5
TYPE: compute.v1.subnetwork
STATE: COMPLETED
ERRORS: []
INTENT:

NAME: us-west-s1
TYPE: compute.v1.subnetwork
STATE: COMPLETED
ERRORS: []
INTENT:

NAME: us-west-s2
TYPE: compute.v1.subnetwork
STATE: COMPLETED
ERRORS: []
INTENT:

NAME: w1-vm
TYPE: compute.v1.instance
STATE: COMPLETED
ERRORS: []
INTENT:

NAME: w2-vm
TYPE: compute.v1.instance
STATE: COMPLETED
ERRORS: []
INTENT:
awurtz@cloudshell:~/networking101 (cloud-wurtz-awurtz)$
```

1 network, 4 subnetworks, and 5 VM instances were created.

Addison Wurtz
CS 530
notebooks/Week10

Visit the web console for VPC network and show the network and the subnetworks that have been created. Validate that it has created the infrastructure in the initial figure. Note the lack of firewall rules that have been created.



VPC network details

networking101

Maximum transmission unit
1460

VPC network ULA internal IPv6 range
Disabled

Subnet creation mode
Custom subnets

Dynamic routing mode
Regional

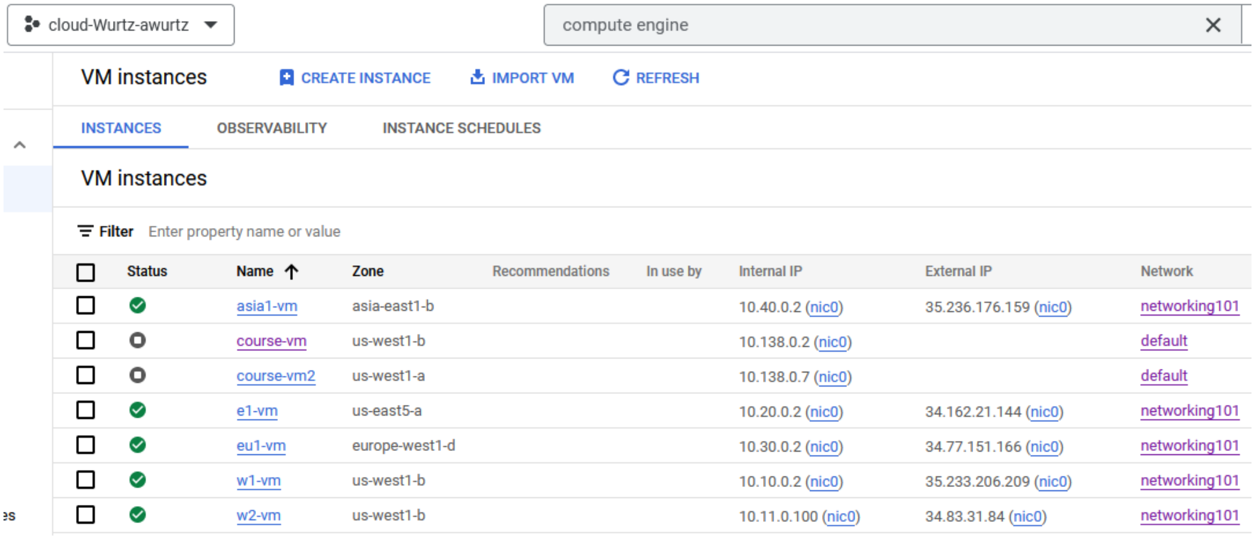
DNS server policy
None

Subnets

☐	Name ↑	Region	Stack Type
☐	asia-east1	asia-east1	IPv4
☐	europe-west1	europe-west1	IPv4
☐	us-east5	us-east5	IPv4
☐	us-west-s1	us-west1	IPv4
☐	us-west-s2	us-west1	IPv4

Addison Wurtz
CS 530
notebooks/Week10

Visit the web console for Compute Engine and show all VMs that have been created, their internal IP addresses and the subnetworks they have been instantiated on. Validate that it has created the infrastructure shown in the initial figure.



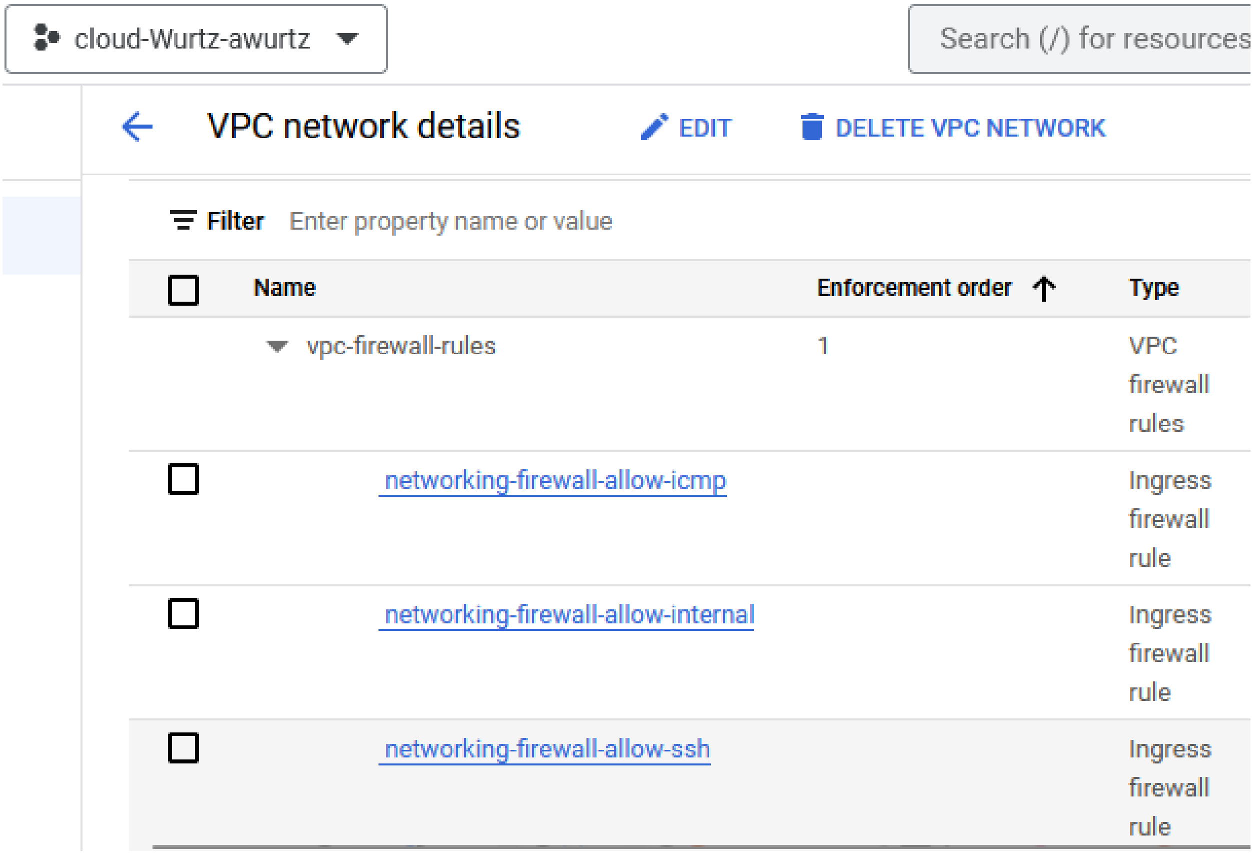
□	Status	Name ↑	Zone	Recommendations	In use by	Internal IP	External IP	Network
□	✓	asia1-vm	asia-east1-b			10.40.0.2 (nic0)	35.236.176.159 (nic0)	networking101
□	○	course-vm	us-west1-b			10.138.0.2 (nic0)		default
□	○	course-vm2	us-west1-a			10.138.0.7 (nic0)		default
□	✓	e1-vm	us-east5-a			10.20.0.2 (nic0)	34.162.21.144 (nic0)	networking101
□	✓	eu1-vm	europe-west1-d			10.30.0.2 (nic0)	34.77.151.166 (nic0)	networking101
□	✓	w1-vm	us-west1-b			10.10.0.2 (nic0)	35.233.206.209 (nic0)	networking101
□	✓	w2-vm	us-west1-b			10.11.0.100 (nic0)	34.83.31.84 (nic0)	networking101

Click on the ssh button for one of the VMs and attempt to connect. Did it succeed?

No.

Addison Wurtz
CS 530
notebooks/Week10

Update deployment



VPC network details

□	Name	Enforcement order ↑	Type
	▾ vpc-firewall-rules	1	VPC firewall rules
□	networking-firewall-allow-icmp		Ingress firewall rule
□	networking-firewall-allow-internal		Ingress firewall rule
□	networking-firewall-allow-ssh		Ingress firewall rule

Latency measurements

Location pair	Ideal latency	Measured latency
us-west1 us-east5	~45 ms	51 ms
us-west1 europe-west1	~93 ms	135 ms
us-west1 asia-east1	~114 ms	116 ms
us-east5 europe-west1	~76 ms	89 ms
us-east5 asia-east1	~141 ms	186 ms
europe-west1 asia-east1	~110 ms	269 ms

Addison Wurtz
CS 530
notebooks/Week10

Test groups

Are the instances in the same availability zone or in different ones?

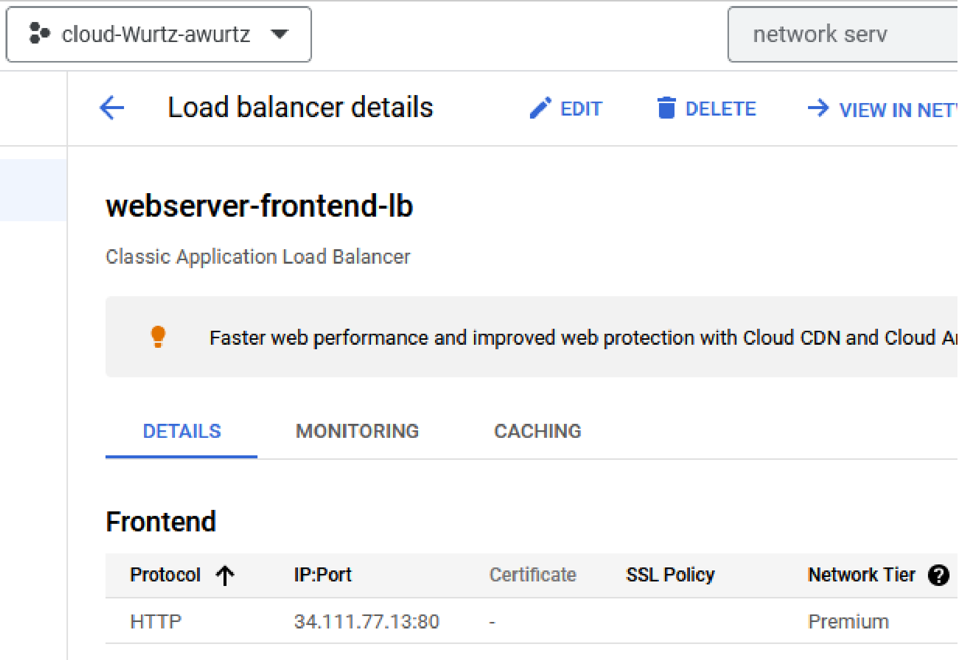
They are in different zones.

List all availability zones that your servers show up in for your lab notebook.

us-east5-c, us-east5-a, us-east5-b, europe-west1-d, europe-west1-c, europe-west1-b

---

Test load balancer



webserver-frontend-lb

Classic Application Load Balancer

Faster web performance and improved web protection with Cloud CDN and Cloud A

Frontend

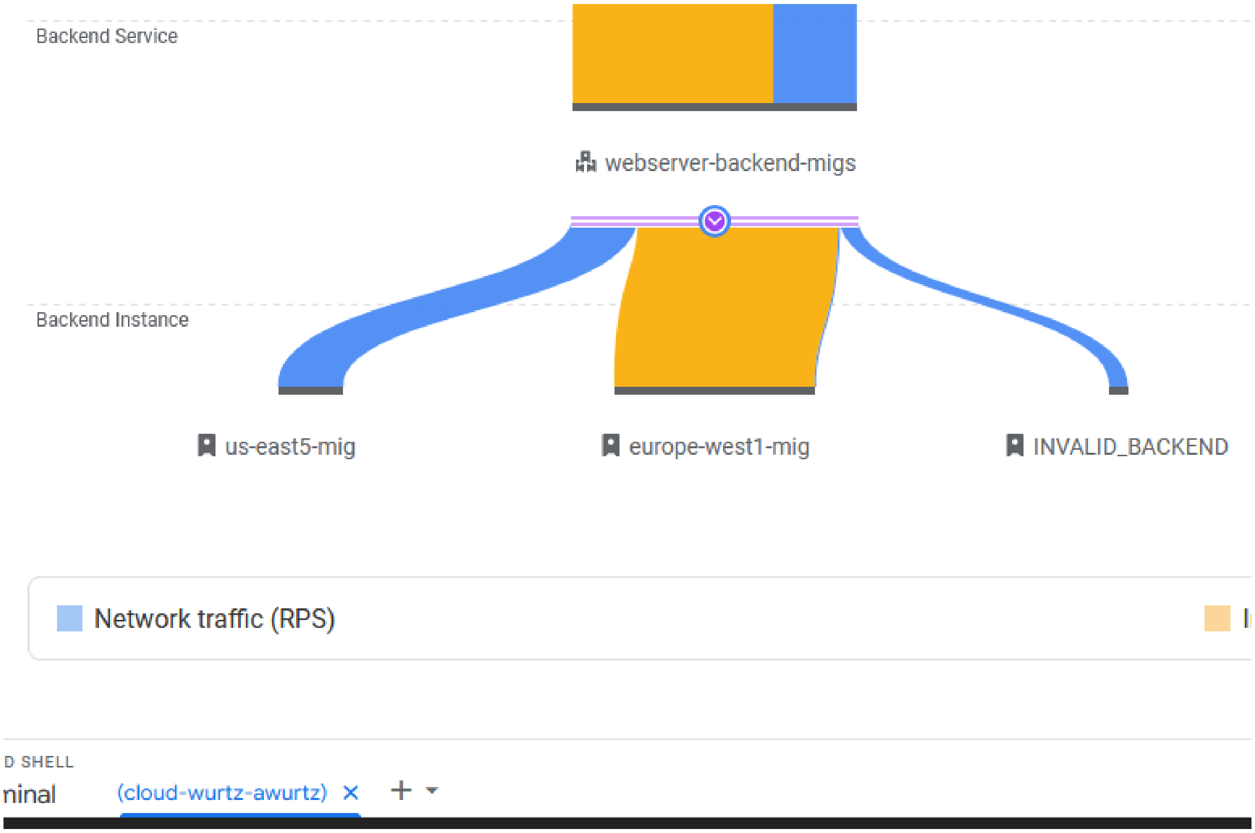
Protocol ↑	IP:Port	Certificate	SSL Policy	Network Tier
HTTP	34.111.77.13:80	-		Premium

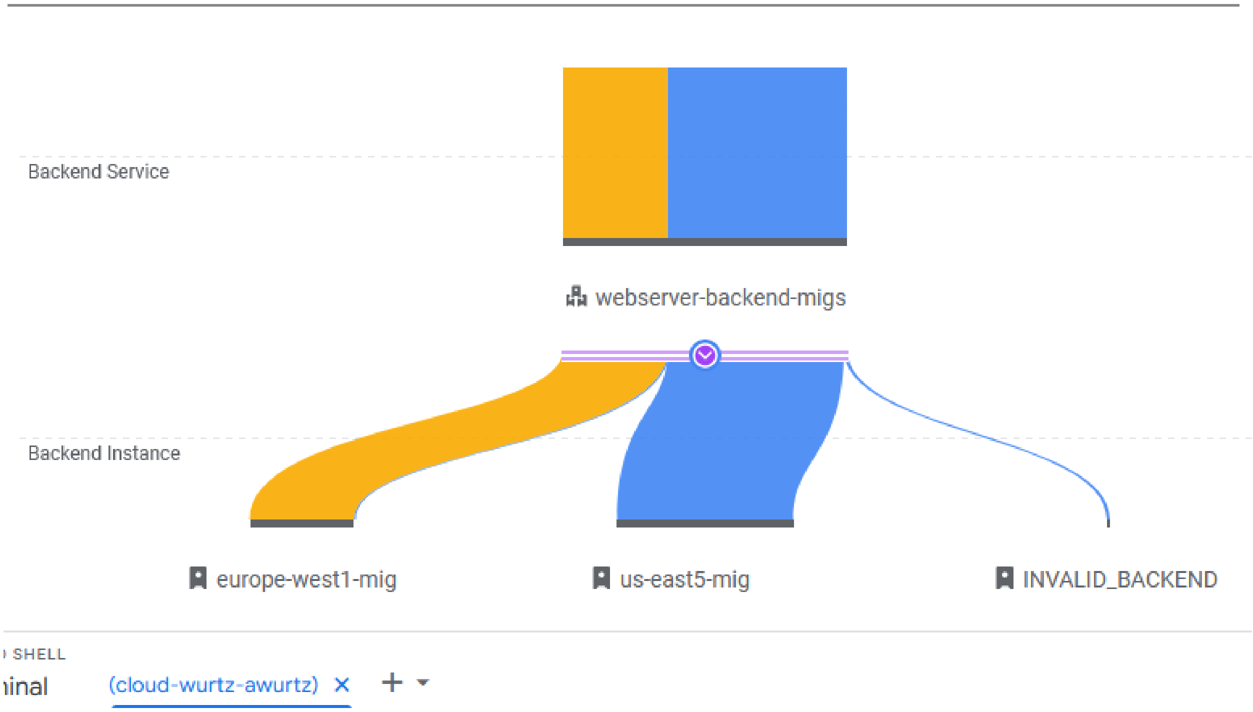
Which availability zone does the server handling your request reside in?

us-east5-a

Addison Wurtz
CS 530
notebooks/Week10

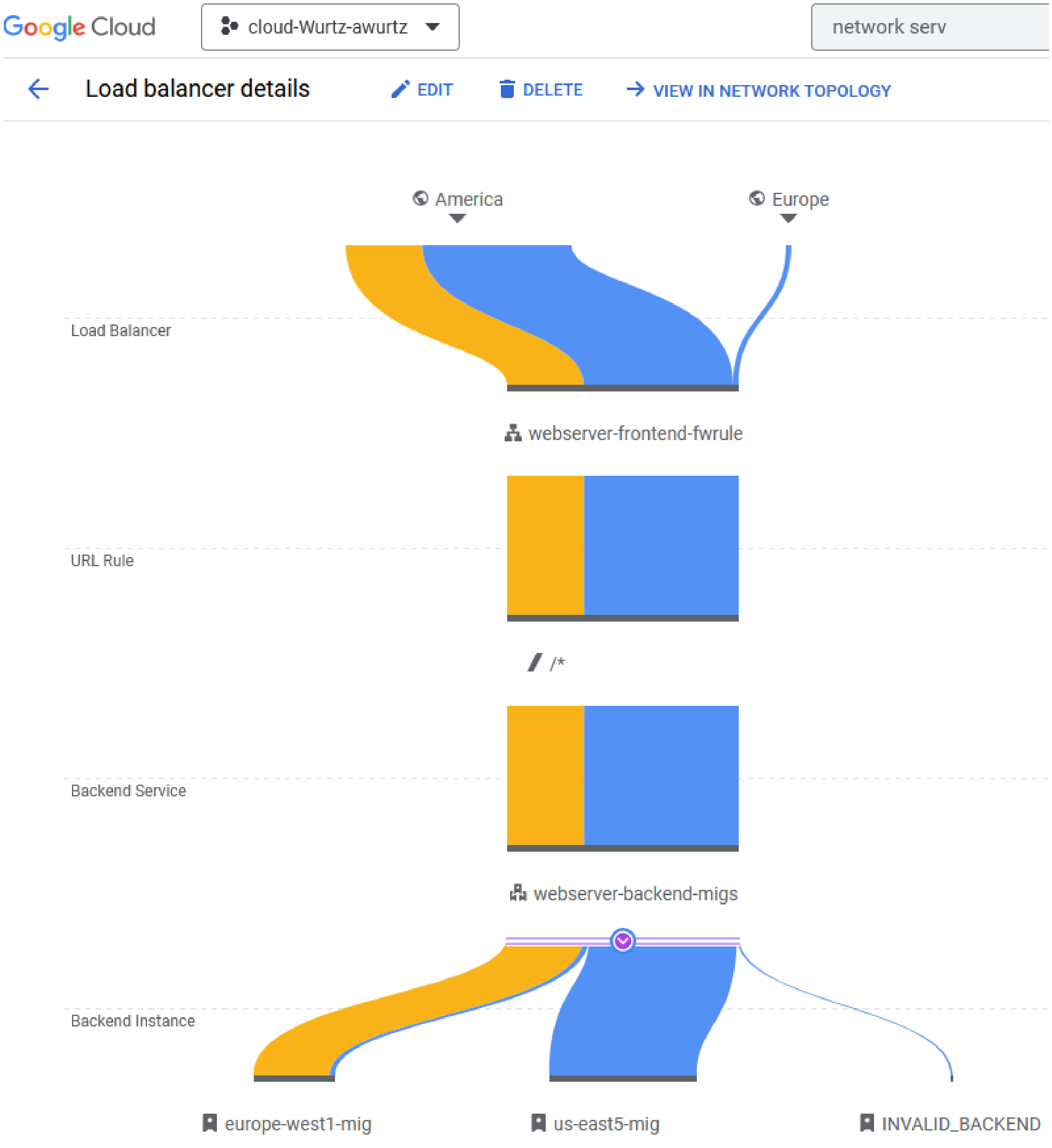
Siege! (Part 1)





Addison Wurtz
CS 530
notebooks/Week10

Siege! (Part 2)


Google Cloud
cloud-Wurtz-awurtz
network serv
Load balancer details
EDIT
DELETE
VIEW IN NETWORK TOPOLOGY
America
Europe
Load Balancer
webserver-frontend-fwrule
URL Rule
/*
Backend Service
webserver-backend-migs
Backend Instance
europe-west1-mig
us-east5-mig
INVALID_BACKEND