

Cloud Computing Architecture

Google Cloud Platform (GCP) and Project Overview

Foteini Strati

foteini.strati@inf.ethz.ch

PhD Student, Systems Group



Agenda



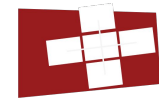
DINFK

1. Project Logistics
2. Intro to Cloud Providers and GCP
3. Cluster setup: Kops
4. Running applications in a cluster: Kubernetes



Goal: Schedule applications of different requirements (latency-sensitive, batch applications) in the cloud efficiently.

- **Part 1:** Measure the effect of resource **interference** (using **iBench**) in **memcached** application
- **Part 2:** Measure the effect of resource **interference** (using **iBench**) and **scalability** of **PARSEC** benchmarks.
- **Part 3:** Use insights from Part 1 & 2 to schedule memcached and PARSEC in **a cluster of heterogeneous VMs**.
- **Part 4:** Schedule memcached and PARSEC together (collocated) on **a single machine**.



- Deadline to submit group preferences for the project: **Thursday, March 9th, 2023.**
- Groups will be assigned on **Monday, March 13th, 2023.** You may then start working on the project.
- Project deadline for Part 1 and 2: **Thursday, April 6th, 2023 at 13:00**
- Project deadline for Part 3 and 4: **Thursday, May 25th, 2023 at 13:00**
- **You **MUST** signup on Moodle for the project (otherwise you cannot do the project):**
 - Sign up for a **specific** group
 - or**
 - Sign up for **general** group (and we will place you in a group)

Cloud Providers and GCP



- Many public cloud providers:
 - Google Cloud Platform (GCP)
 - Amazon Web Services (AWS)
 - Microsoft Azure
 - etc.
- They provide services such as:
 - **Platform as a Service (PaaS)**: app deployment, querying, ...
 - **Infrastructure as a Service (IaaS)**: VMs, storage, ...
 - **Function as a Service (FaaS)**: serverless computing
 - **Software as a Service (SaaS)**: cloud gaming, e-mail, ...
 - etc.



- Offers a set of cloud computing services:
 - Compute Engine
 - Cloud Storage/Filestore
 - BigQuery/DataFlow
 - Cloud Functions
 - AI Platform
 - etc.
- Same infrastructure that Google uses internally for its products (Gmail, Google Drive, etc.)



- Create a Project in GCP
- Configure GCP CLI
- Credits, Billing & Budget
- Create VMs
- Inspect/Access VMs
- Terminate VMs

GCP - Create Project (Get Credits)



DINFK

- You must have a Google Cloud account under your **ETH email address**
- **One group member** redeems the cloud credits (instructions in the project handout)
- **You should NOT redeem credits until the official start date of the project!**

Cloud Platform Education Grants

Use credits provided to you via the Google Cloud Platform Education Grants program to access Google Cloud Platform. Get what you need to build and run your apps, websites and services.

Thank you for your interest in Google Cloud Platform Education Grants. Please fill out the form below to receive a coupon code for credit to use on Google Cloud Platform.

First Name

Last Name

School Email @student.ethz.ch ▾

If you do not see your domain listed, please contact your course instructor: aklimovic@ethz.ch

By clicking "Submit" below, you agree that we may share the following information with your educational institution and course instructor (aklimovic@ethz.ch): (1) personal information that you provide to us on this form and (2) information regarding your use of the coupon and Google Cloud Platform products.

[Privacy Policy](#)



- Login and create the project via gcloud

```
$ gcloud init
```

```
[...]
```

```
You must log in to continue. Would you like to log in (Y/n)? Y
```

```
Go to the following link in your browser:
```

```
...
```

```
Enter verification code:
```

```
...
```



- Login and create the project via gcloud

```
$ gcloud init
```

```
[...]
```

```
You are logged in as: [ethzid@ethz.ch].
```

```
Pick cloud project to use:
```

```
[1] ...
```

```
...
```

```
[6] Enter a project ID
```

```
[7] Create a new project
```

```
Please enter numeric choice or text value (must exactly match list item): 7
```

```
Enter a Project ID. Note that a Project ID CANNOT be changed later.
```

```
Project IDs must be 6-30 characters (lowercase ASCII, digits, or  
hyphens) in length and start with a lowercase letter. cca-eth-2023-group-123
```

```
...
```

```
Your current project has been set to: [cca-eth-2023-group-123].
```



- Link the project with the **Billing Account for Education**

Set the billing account for project “cca-eth-2023-group-123”

Billing account *
Billing Account for Education ▼ ?

Any charges for this project will be billed to the account you select here.

CANCEL

SET ACCOUNT



- Add your teammates to the project

Add principals

Principals are users, groups, domains, or service accounts. [Learn more about principals in IAM](#)

New principals

lcvetkovic@ethz.ch ✕ ?

Assign roles

Roles are composed of sets of permissions and determine what the principal can do with this resource. [Learn more](#)

Role *
Owner ▼

IAM condition (optional) ?
[+ ADD IAM CONDITION](#) ✕

Full access to most Google Cloud resources. See the list of included permissions.

[+ ADD ANOTHER ROLE](#)

[SAVE](#) [CANCEL](#)



- View available zones (usually have all VMs in the same availability zone)

```
$ gcloud compute zones list
```

```
API [compute.googleapis.com] not enabled on project [...]. Would you like to enable and retry (this will take a few minutes)? (y/N)? y
```

Shown
the first
time

```
Enabling service [compute.googleapis.com] on project [...]...
```

NAME	REGION	STATUS
us-east1-b	us-east1	UP
us-east1-c	us-east1	UP
us-east1-d	us-east1	UP
us-east4-c	us-east4	UP
us-east4-b	us-east4	UP
us-east4-a	us-east4	UP
us-central1-c	us-central1	UP
....		

Get available/active accounts:

```
$ gcloud auth list
```

```
Credentialed Accounts  
ACTIVE ACCOUNT  
*   username@ethz.ch
```

To set the active account, run:

```
$ gcloud config set account `ACCOUNT`
```

Set account

```
$ gcloud auth application-default login
```

Go to the following link in your browser:

....

Set/get current project

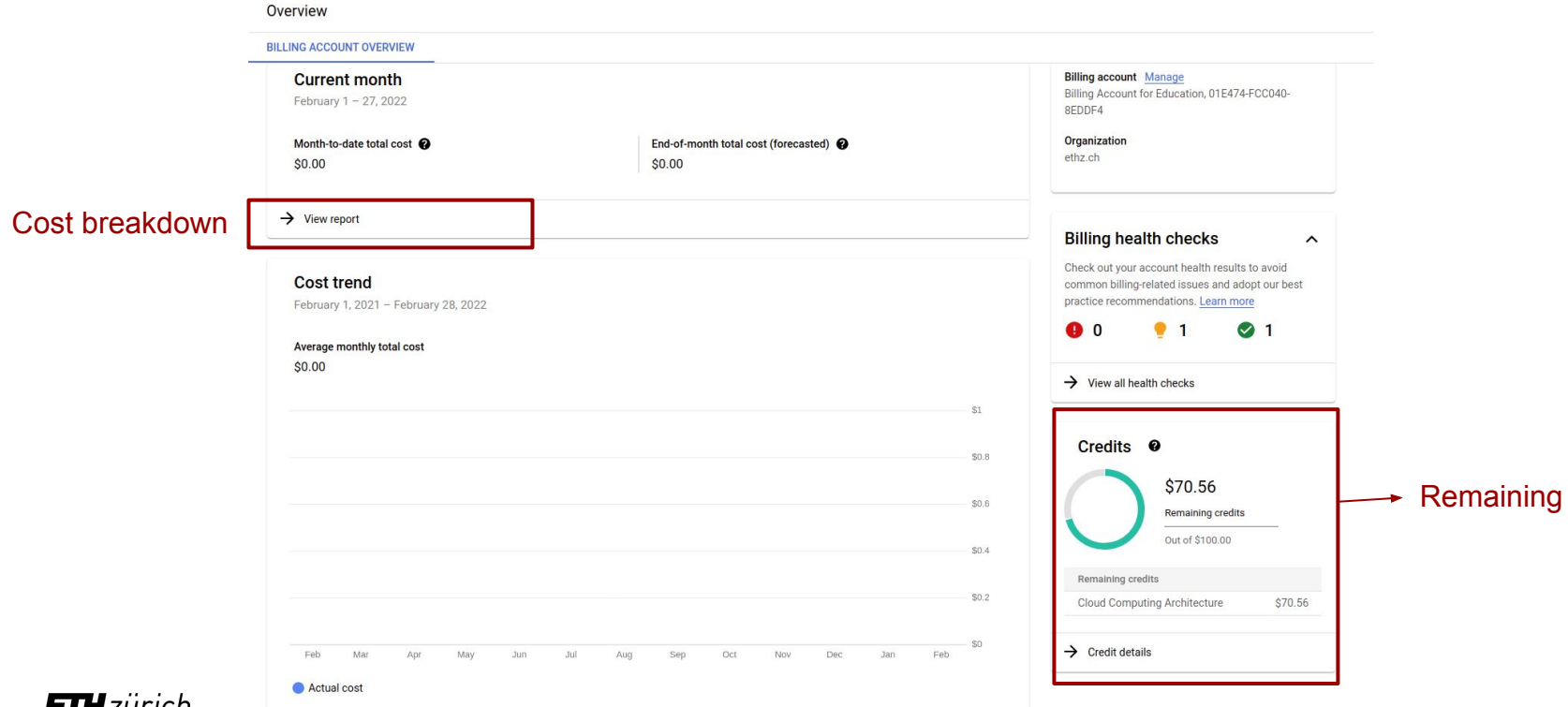
```
$ gcloud config set project myproject
```

Updated property [core/project].

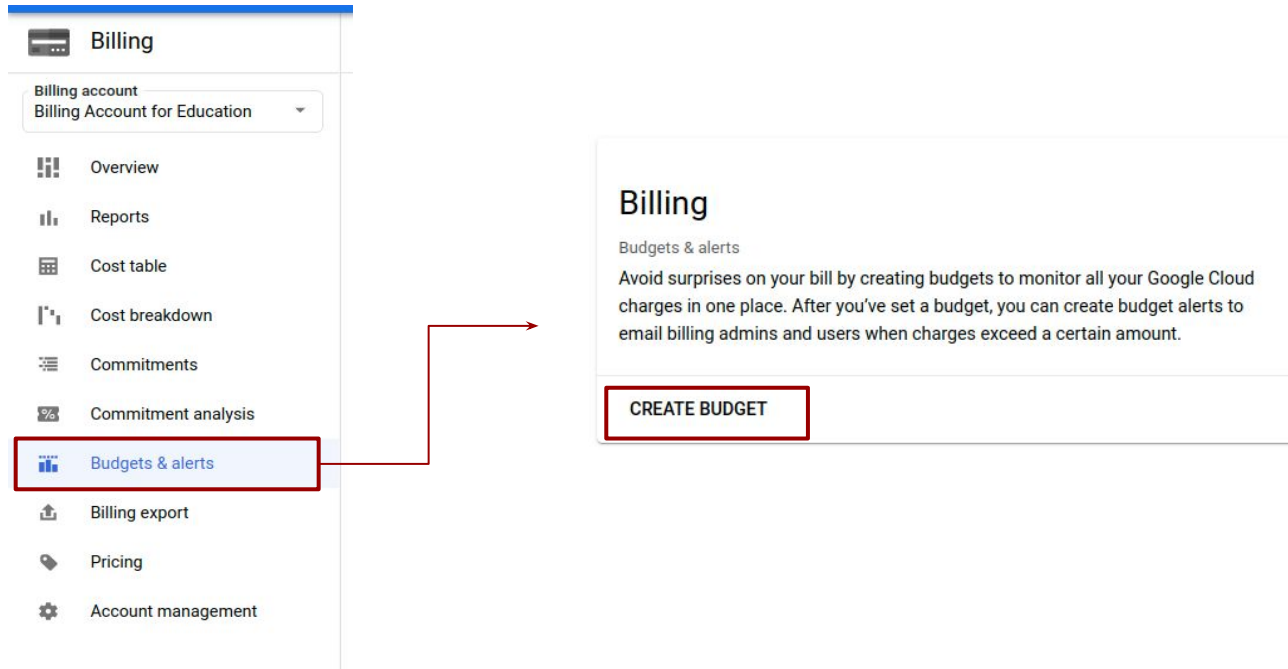
```
$ gcloud config get-value project
```

Your active configuration is [xyz]
myproject

- View remaining credits/cost overview



- Budget alerts



The screenshot displays the Google Cloud Billing console interface. On the left, a sidebar menu lists various billing-related options: Billing, Billing account (set to 'Billing Account for Education'), Overview, Reports, Cost table, Cost breakdown, Commitments, Commitment analysis, **Budgets & alerts** (highlighted with a red box), Billing export, Pricing, and Account management. A red arrow points from the 'Budgets & alerts' menu item to the main content area. The main content area, titled 'Billing', includes a subtitle 'Budgets & alerts' and a descriptive paragraph: 'Avoid surprises on your bill by creating budgets to monitor all your Google Cloud charges in one place. After you've set a budget, you can create budget alerts to email billing admins and users when charges exceed a certain amount.' Below this text, a red-bordered button labeled 'CREATE BUDGET' is visible.



1 Scope

Name *

budget-alert-1

A budget enables you to track your actual spend against your planned spend.

Time range

Monthly

The month starts on the first of the month and reset at the beginning of each month.

A budget can be scoped to focus on a specific set of resources.

Projects

All projects (2)

Services

All services (1938)

Labels ?

Select the key and value of the label you want to filter.

Credits

Selected credits are applied to the total cost. Budget tracks the total cost minus any applicable selected credits.

☐

Discounts ?

☐

Promotions and others ?

NEXT

✓ These should be unchecked!



Scope



2 Amount

Set a monthly budget amount

Budget type

Specified amount

A fixed amount that your spend will be compared against.

Target amount *

\$ 30

NEXT



3 Actions

FINISH

CANCEL

- Get a list of available images

```
$ gcloud compute images list
```

NAME	PROJECT	FAMILY	DEPRECATED	STATUS
centos-7-v20220126	centos-cloud	centos-7		READY
centos-stream-8-v20220128	centos-cloud	centos-stream-8		READY
cos-85-13310-1416-5	cos-cloud	cos-85-lts		READY
cos-89-16108-604-11	cos-cloud	cos-89-lts		READY
cos-93-16623-102-12	cos-cloud	cos-93-lts		READY
cos-beta-93-16623-39-6	cos-cloud	cos-beta		READY
debian-10-buster-v20220118	debian-cloud	debian-10		READY
debian-11-bullseye-v20220120	debian-cloud	debian-11		READY
debian-9-stretch-v20220118	debian-cloud	debian-9		READY
fedora-cloud-base-gcp-33-1-2-x86-64	fedora-cloud	fedora-cloud-33		READY
fedora-cloud-base-gcp-34-1-2-x86-64	fedora-cloud	fedora-cloud-34		READY
fedora-cloud-base-gcp-35-1-2-x86-64	fedora-cloud	fedora-cloud-35		READY
....				



- Create the VM

```
$ gcloud compute instances create vm-gcloud --image-family=ubuntu-1804-lts --image-project=ubuntu-os-cloud  
--machine-type=e2-standard-2 --project=cca-eth-2023-group-fstrati --zone=europe-west3-a
```

Created

<https://www.googleapis.com/compute/v1/projects/cca-eth-2023-group-fstrati/zones/europe-west3-a/instances/vm-gcloud>

NAME	ZONE	MACHINE_TYPE	PREEMPTIBLE	INTERNAL_IP	EXTERNAL_IP	STATUS
vm-gcloud	europe-west3-a	e2-standard-2		10.156.0.2	35.159.171.70	RUNNING

GCP - Create VM - Console



DINFK

The screenshot shows the Google Cloud Platform console interface for creating a VM instance. The left sidebar contains navigation links: Home, View all products, and Compute Engine (highlighted with a red box). The main content area shows the 'VIRTUAL MACHINES' section with 'VM instances' highlighted (red box). An arrow points from 'VM instances' to the 'CREATE INSTANCE' button (red box). Below this, the 'Set name, zone, type' section is shown, with 'Name' set to 'vm-console', 'Region' set to 'europe-west3 (Frankfurt)', and 'Zone' set to 'europe-west3-a' (all highlighted with red boxes). The 'Boot disk' section shows 'Name' as 'instance-1', 'Type' as 'New balanced persistent disk', 'Size' as '10 GB', and 'Image' as 'Ubuntu 18.04 LTS' (highlighted with a red box). An arrow points from 'Ubuntu 18.04 LTS' to the 'CREATE' button (red box). The 'Machine configuration' section shows 'Machine family' as 'GENERAL-PURPOSE', 'Series' as 'E2', and 'Machine type' as 'e2-standard-4 (4 vCPU, 16 GB memory)' (all highlighted with red boxes). The bottom of the console shows 'vCPU' as 4 and 'Memory' as 16 GB.

Google Cloud Platform cca-eth-2022-group-123

Home > VIRTUAL MACHINES

VM instances

Instance templates

Sole-tenant nodes

Machine images

TPUs

Committed use discounts

PINNED

Compute Engine

VM instances

CREATE INSTANCE

IMPORT VM

REFRESH

START / RESUME

STOP

Set name, zone, type

Name *

vm-console

Labels

+ ADD LABELS

Region *

europe-west3 (Frankfurt)

Zone *

europe-west3-a

Region is permanent

Zone is permanent

Boot disk

Name

instance-1

Type

New balanced persistent disk

Size

10 GB

Image

Ubuntu 18.04 LTS

CHANGE

CREATE

CANCEL

EQUIVALENT COMMAND LINE

Machine configuration

Machine family

GENERAL-PURPOSE

COMPUTE-OPTIMIZED

MEMORY-OPTIMIZED

Machine types for common workloads, optimized for cost and flexibility

Series

E2

CPU platform selection based on availability

Machine type

e2-standard-4 (4 vCPU, 16 GB memory)

vCPU

4

Memory

16 GB

CPU PLATFORM AND GPU



- From CLI

```
$ gcloud compute instances list
```

NAME	ZONE	MACHINE_TYPE	PREEMPTIBLE	INTERNAL_IP	EXTERNAL_IP	STATUS
vm-console	europe-west3-a	e2-standard-4		10.156.0.3	35.198.83.193	RUNNING
vm-gcloud	europe-west3-a	e2-standard-2		10.156.0.2	35.159.171.70	RUNNING

- From Console

INSTANCES

OBSERVABILITY

NEW

INSTANCE SCHEDULES

VM instances

Filter

Enter property name or value

<input type="checkbox"/>	Status	Name ↑	Zone	Recommendations	In use by	Internal IP	External IP	Connect		
<input type="checkbox"/>	✓	vm-console	europe-west3-a			10.156.0.3 (nic0)	35.198.83.193 (nic0)	SSH	▼	⋮
<input type="checkbox"/>	✓	vm-gcloud	europe-west3-a			10.156.0.2 (nic0)	34.159.171.70 (nic0)	SSH	▼	⋮



- Connect via SSH

```
$ gcloud compute ssh vm-console --project cca-eth-2023-group-fstrati --zone europe-west3-a
```

```
[...]
```

```
Welcome to Ubuntu 18.04.6 LTS (GNU/Linux 5.4.0-1064-gcp x86_64)
```

```
user@vm-console:~$
```



- From CLI

```
$ gcloud compute instances delete vm-console --project cca-eth-2023-group-fstrati --zone europe-west3-a
```

The following instances will be deleted. Any attached disks configured to be auto-deleted will be deleted unless they are attached to any other instances or the `--keep-disks` flag is given and specifies them for keeping. Deleting a disk is irreversible and any data on the disk will be lost.

- [vm-console] in [europe-west3-a]

Do you want to continue (Y/n)? Y

Deleted

[<https://www.googleapis.com/compute/v1/projects/cca-eth-2023-group-fstrati/zones/europe-west3-a/instances/vm-console>].



- From Console

VM instances [+ CREATE INSTANCE](#) [IMPORT VM](#) [REFRESH](#) [START / RESUME](#) [STOP](#) [⋮](#) [OPERATIONS](#)

[INSTANCES](#) [INSTANCE SCHEDULE](#)

VM instances are highly configurable virtual machines for running workloads on Google infrastructure. [Learn more](#)

[Filter](#) Enter property name or value

<input checked="" type="checkbox"/>	Status	Name ↑	Zone	Recommendations	In use by	Internal IP	External IP	
<input checked="" type="checkbox"/>	✓	vm-gcloud	europe-west3-a			10.156.0.2 (nic0)	35.246.201.130	SSH ⋮

[CREATE SCHEDULE](#)
[DELETE](#)
[RESET](#)
[SUSPEND](#)

Setting up a cluster with Kops



- Tool to create, maintain, update and delete Kubernetes-enabled clusters [1]
- Users define a *cluster specification* (e.g. a yaml file)
 - We will describe how our GCP VMs should look like (number of machines, machine type, etc.)
- Supports multiple cloud providers such as AWS, GCP

[1] [Kops guide](#)

```
apiVersion: kops.k8s.io/v1alpha2
kind: Cluster
metadata:
  creationTimestamp: null
  name: part1.k8s.local
spec:
  ...
  configBase:
  gs://cca-eth-2023-group-XXX-ethzid/part1.k8s.local
  ...
  project: cca-eth-2023-group-XXX
  ...
```

Change these!

```
apiVersion: kops.k8s.io/v1alpha2
kind: InstanceGroup
metadata:
  creationTimestamp: null
  labels:
    kops.k8s.io/cluster: part1.k8s.local
  name: memcache-server
spec:
  image: ubuntu-os-cloud/ubuntu-1804-bionic-v20210211
  machineType: t2d-standard-2
  maxSize: 1
  minSize: 1
  nodeLabels:
    cloud.google.com/metadata-proxy-ready: "true"
    kops.k8s.io/instancegroup: nodes-europe-west3-a
    cca-project-nodetype: "memcached"
  ...
```

Create a cluster with Kops



```
$ gsutil mb gs://cca-eth-2023-group-XXX-ethzid/  
$ export KOPS_STATE_STORE=gs://cca-eth-2023-group-XXX-ethzid/
```

→ For the cluster configuration

```
$ cd ~/.ssh & ssh-keygen -t rsa -b 4096 -f cloud-computing
```

→ Generate login keys

```
$ PROJECT=`gcloud config get-value project`  
$ export KOPS_FEATURE_FLAGS=AlphaAllowGCE
```

```
$ kops create -f part1.yaml
```

→ Set configuration

```
$ kops create secret --name part1.k8s.local sshpublickey admin -i ~/.ssh/cloud-computing.pub
```

→ Add login keys

```
$ kops update cluster --name part1.k8s.local --yes --admin
```

→ Deploy the cluster

```
$ kops validate cluster --wait 10m
```

→ Wait...

Inspect the cluster



DINFK

```
$ kubectl get nodes -o wide
```

NAME	STATUS	ROLES	AGE	VERSION	INTERNAL-IP	EXTERNAL-IP	OS-IMAGE	KERNEL-VERSION	CONTAINER-RUNTIME
client-agent-4v0l	Ready	node	2m38s	v1.23.16	10.0.16.4	34.159.225.244	Ubuntu 18.04.5 LTS	5.4.0-1036-gcp	docker://20.10.17
client-measure-0svs	Ready	node	2m41s	v1.23.16	10.0.16.3	34.141.87.51	Ubuntu 18.04.5 LTS	5.4.0-1036-gcp	docker://20.10.17
master-europe-west3-a-459v	Ready	control-plane,master	3m53s	v1.23.16	10.0.16.5	34.89.151.210	Ubuntu 18.04.5 LTS	5.4.0-1036-gcp	docker://20.10.17
memcache-server-z82c	Ready	node	2m45s	v1.23.16	10.0.16.2	34.159.171.70	Ubuntu 18.04.5 LTS	5.4.0-1036-gcp	docker://20.10.17

VM instances

Filter Enter property name or value										
<input type="checkbox"/>	Status	Name ↑	Zone	Recommendations	In use by	Internal IP	External IP	Connect		
<input type="checkbox"/>	✓	client-agent-4v0l	europe-west3-a		a-client-agent-part1-k8s-local	10.0.16.4 (nic0)	34.159.225.244 (nic0)	SSH ▾	⋮	
<input type="checkbox"/>	✓	client-measure-0svs	europe-west3-a		a-client-measure-part1-k8s-local	10.0.16.3 (nic0)	34.141.87.51 (nic0)	SSH ▾	⋮	
<input type="checkbox"/>	✓	master-europe-west3-a-459v	europe-west3-a		a-master-europe-west3-a-part1-k8s-local , api-part1-k8s-local	10.0.16.5 (nic0)	34.89.151.210 (nic0)	SSH ▾	⋮	
<input type="checkbox"/>	✓	memcache-server-z82c	europe-west3-a		a-memcache-server-part1-k8s-local	10.0.16.2 (nic0)	34.159.171.70 (nic0)	SSH ▾	⋮	

Instance Groups



DINFK

Instance groups

CREATE INSTANCE GROUP

REFRESH

DELETE

HELP ASSISTANT

LEARN

Instance groups are collections of VM instances that use load balancing and automated services, like autoscaling and autohealing. [Learn more](#)

Filter Enter property name or value											
<input type="checkbox"/> Status	Name	Instances	Template	Group type	Creation time	Recommendation	Autoscaling	Zone	In Use By		
<input type="checkbox"/>	a-client-agent-part1-k8s-local	1	client-agent-part1-k8s-local-1677426005	Managed	Feb 26, 2023, 4:40:26 PM UTC+01:00		No configuration	europe-west3-a			
<input type="checkbox"/>	a-client-measure-part1-k8s-local	1	client-measure-part1-k8s-local-1677426006	Managed	Feb 26, 2023, 4:40:26 PM UTC+01:00		No configuration	europe-west3-a			
<input type="checkbox"/>	a-master-europe-west3-a-part1-k8s-local	1	master-europe-west3-a-par-ee08bm-1677426006	Managed	Feb 26, 2023, 4:40:29 PM UTC+01:00		No configuration	europe-west3-a	api-part1-k8s-local		
<input type="checkbox"/>	a-memcache-server-part1-k8s-local	1	memcache-server-part1-k8s-local-1677426006	Managed	Feb 26, 2023, 4:40:25 PM UTC+01:00		No configuration	europe-west3-a			



Load balancing					
+ CREATE LOAD BALANCER REFRESH DELETE					
LOAD BALANCERS BACKENDS FRONTENDS					
Filter Enter property name or value					
<input type="checkbox"/>	Name	Load balancer type ↑	Protocols	Region	Backends
<input type="checkbox"/>	api-part1-k8s-local	Network (target pool-based)	TCP	europe-west3	✓ 1 target pool (1 instance) ⋮

To view or delete load balancing resources like forwarding rules and target proxies, go to the [load balancing components view](#).

Delete the cluster



DINFK

```
$ kops delete cluster part1.k8s.local --yes
```

```
[...]
```

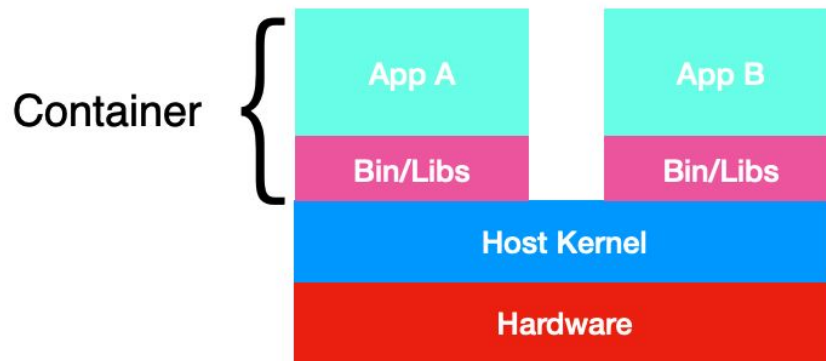
```
Deleted kubectl config for part1.k8s.local
```

```
Deleted cluster: "part1.k8s.local"
```

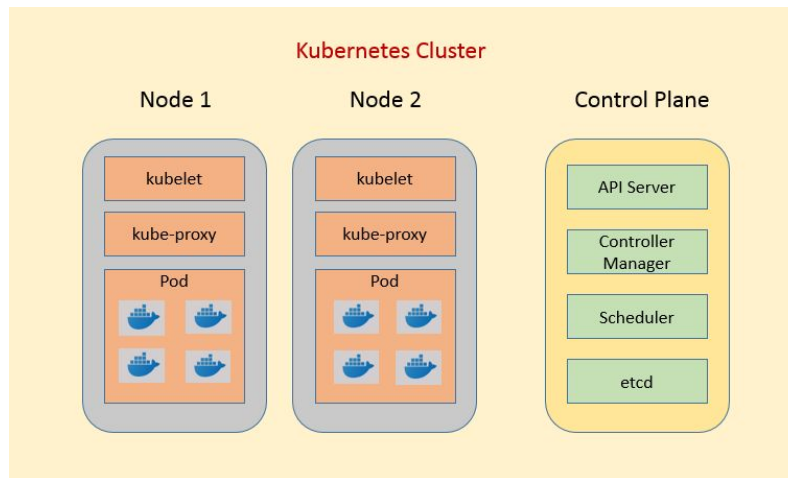
**It is very important to delete your cluster,
otherwise you may spend your credits quite quickly!**

Running applications with Kubernetes

- Packages of software (application code and dependencies)
- Allow applications to be “packed” and run in any environment
- In this project, we are going to use **Docker** containers



- Runs, monitors, and manages containers
- Schedules containers on the cluster (unit of scheduling: **pod**) and migrates them
- Load balancing across containers
- ... and many more (we will see more later during the semester)





- [Memcached](#): distributed, in memory, key-value store - latency-sensitive
 - [mcperf](#): load generator for memcache
- [PARSEC](#): benchmark suite of multi-threaded applications [1]
- [iBench](#): suite of microbenchmarks that cause interference to various resources (CPU, caches, memory bandwidth)

[1] [The PARSEC benchmark suite](#)

Run workloads with kubectl



DINFK

```
$ kubectl create -f memcache-t1-cpuset.yaml  
$ kubectl expose pod some-memcached --name some-memcached-11211 \  
  --type LoadBalancer --port 11211 \  
  --protocol TCP
```

→ Launch memcached

```
$ sleep 60
```

```
$ kubectl get service some-memcached-11211
```

```
$ kubectl get pods -o wide
```

NAME	READY	STATUS	RESTARTS	AGE	IP	NODE
some-memcached	1/1	Running	0	4m40s	100.96.2.3	memcache-server-z82c

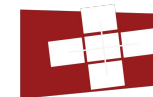
↓
You will need
this for mcperf

Yaml file for workload description



DINFK

```
apiVersion: v1
kind: Pod
metadata:
  name: some-memcached
  labels:
    name: some-memcached
spec:
  containers:
    - image: anakli/memcached:t1
      name: memcached
      imagePullPolicy: Always
      command: ["/bin/sh"]
      args: ["-c", "taskset -c 0 ./memcached -t 1 -u memcache"]
  nodeSelector:
    cca-project-nodetype: "memcached"
```



Thank you!
Questions?