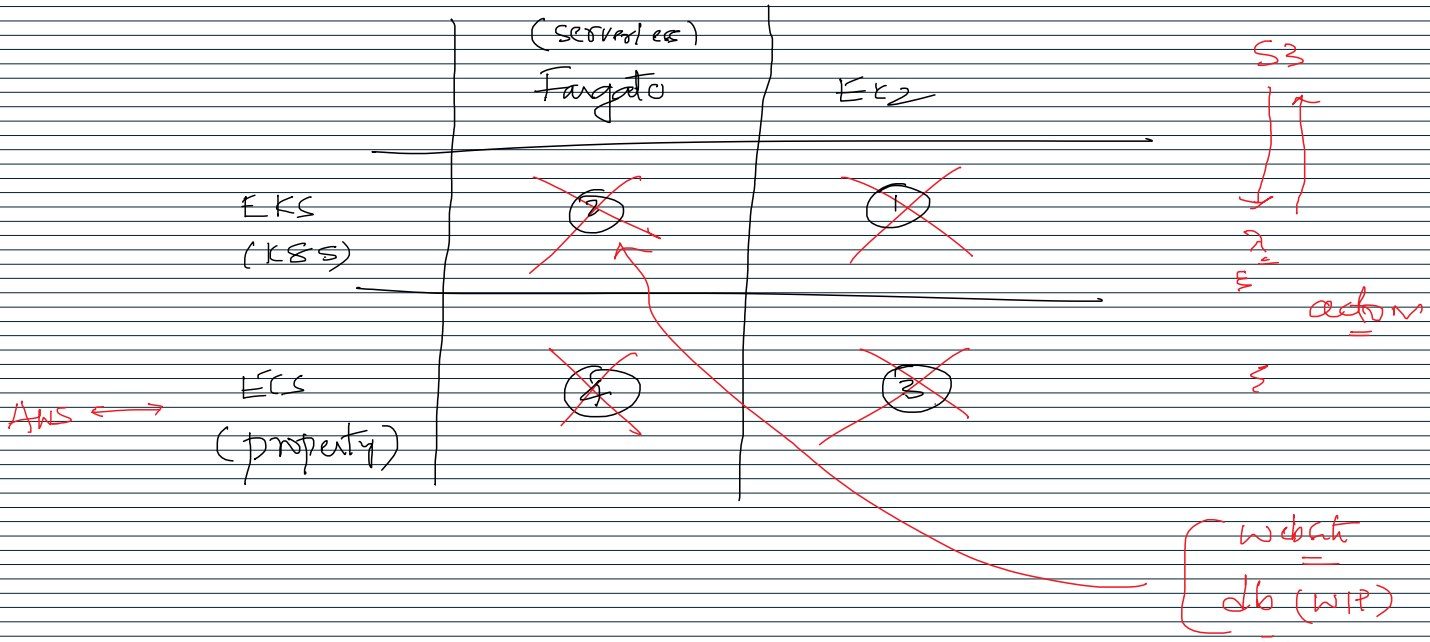
17 November 2020 11:28

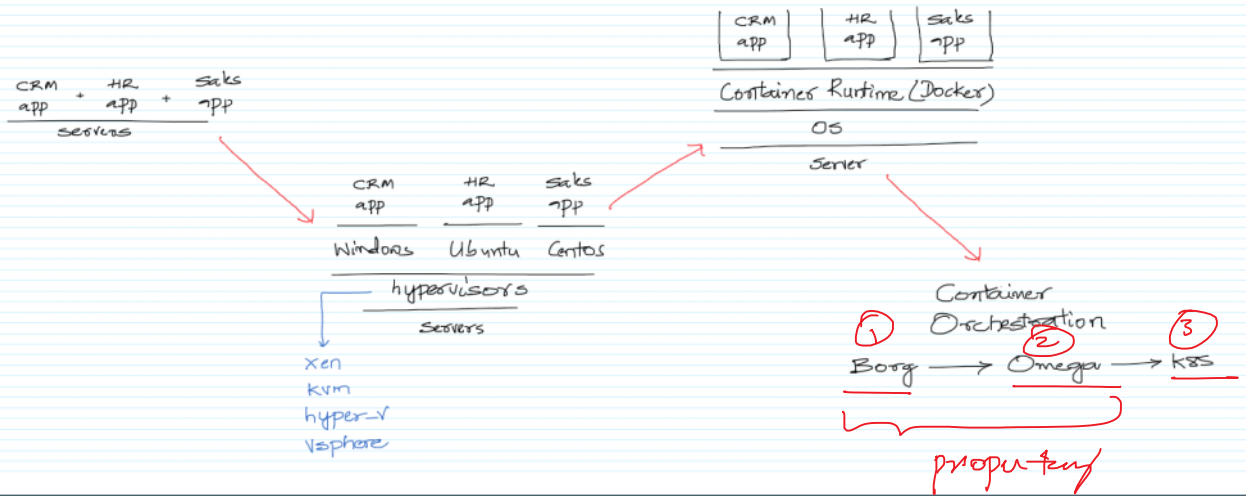
Agenda

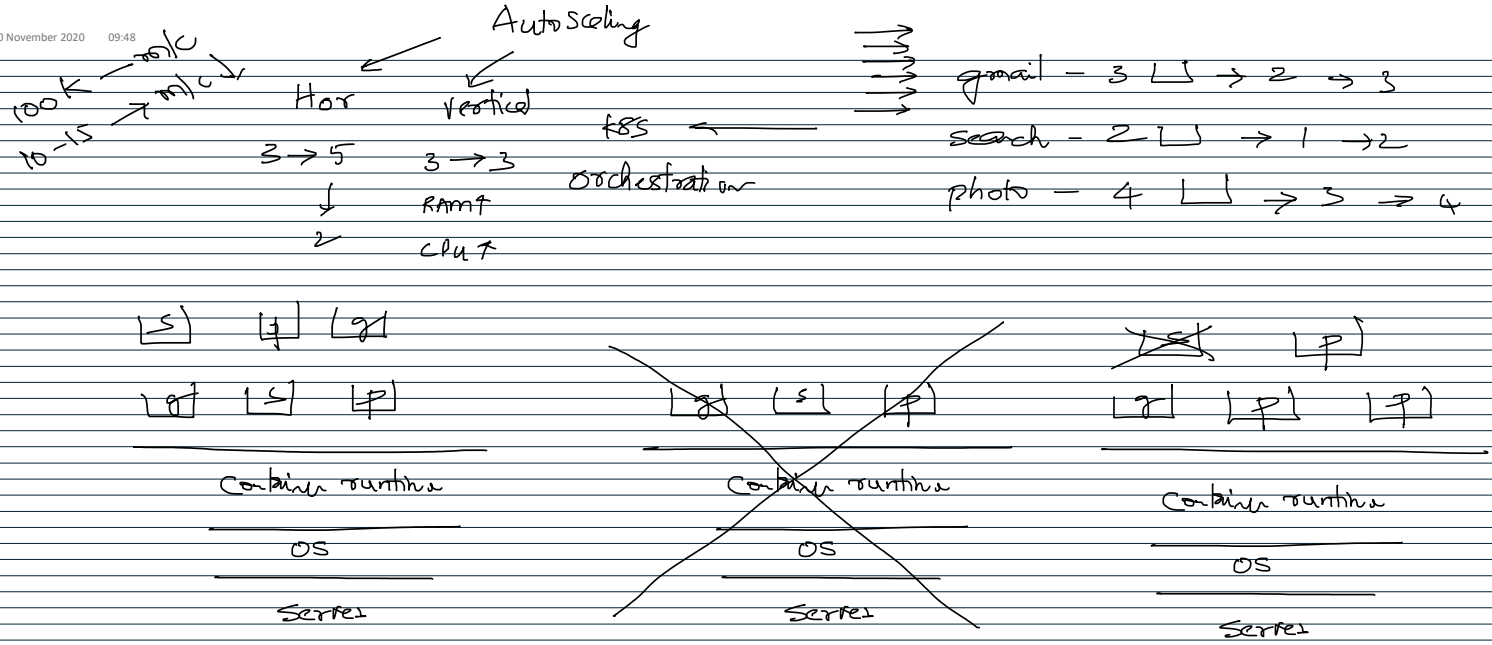
- Docker
- Kubernetes
- AWS ∩ Kubernetes

Different Ways

20 November 2020 12:19




CRM app + HR app + Saks app
Servers
CRM app
HR app
Saks app
Windows
Ubuntu
Centos
hypervisors
Servers
Xen
kvm
hyper-V
Vsphere
CRM app
HR app
Saks app
Container Runtime (Docker)
OS
Server
Container Orchestration
1
2
3
Borg
Omega
k8s
proprietary


Autoscaling
100K
10-15
Hor
Vertical
3 → 5
3 → 3
RAM
CPU
K8S
orchestration
gmail - 3 → 2 → 3
search - 2 → 1 → 2
photo - 4 → 3 → 4

Container runtime
OS
Server
Container runtime
OS
Server
Container runtime
OS
Server

CPU → slow / costly

AI/ML

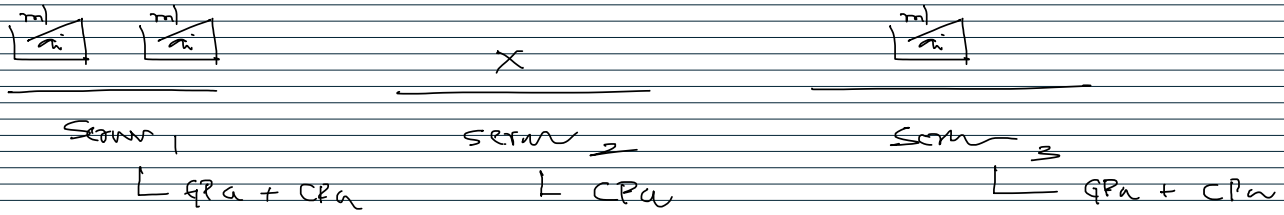
GPU → fast / cheap

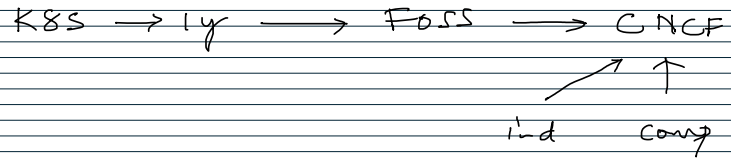
Nvidia T

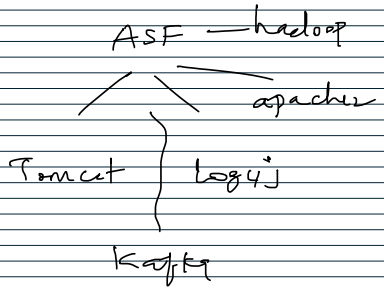
ECS

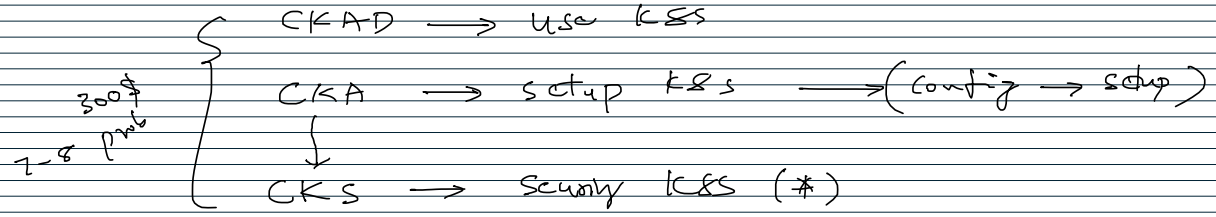
3 containers

Served with GPU




K8S
Ly
FOSS
CNCF
ind
Comp


ASF
hadoop
apache
Tomcat
log4j
Kafka



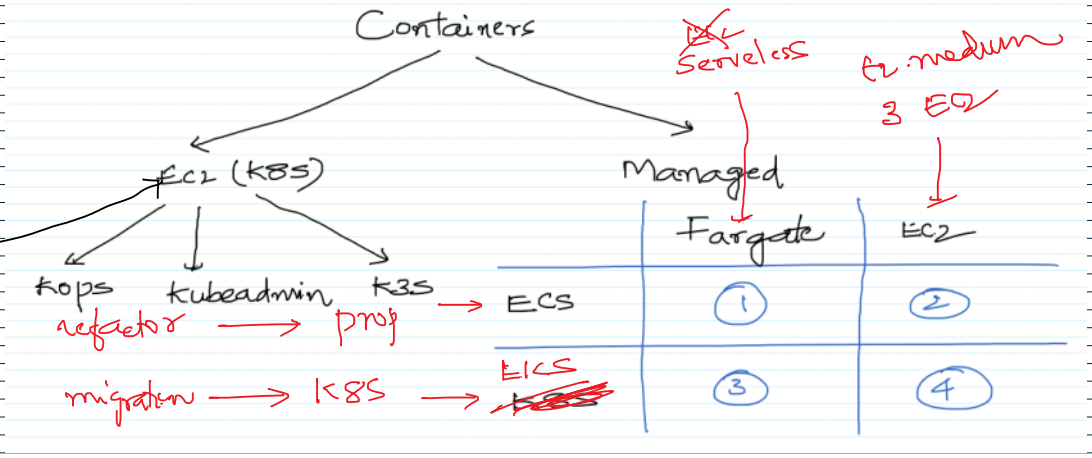
- open book
  - k8s doc ← exam
  - k8s blog

AWS → k8s

Linux → practical

- 2 attempts

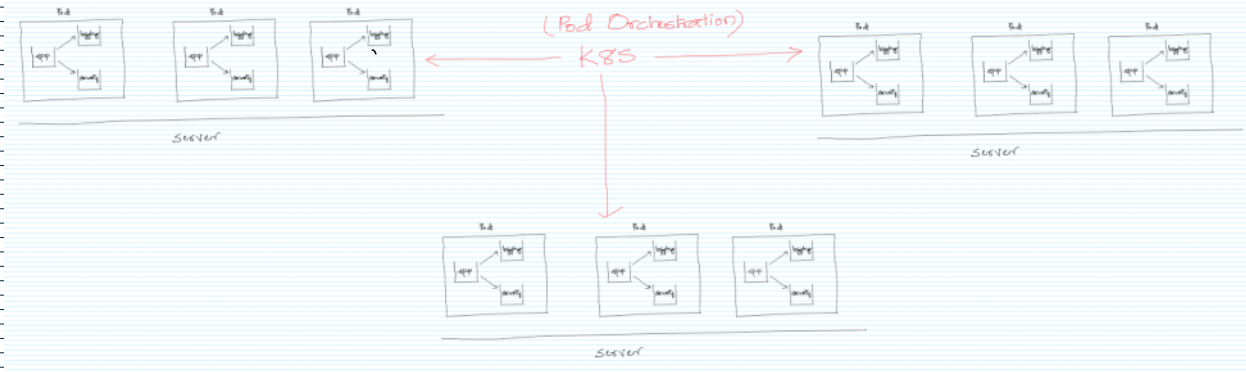
3 containers — 2 CPU, 2 GB RAM



- upgrade
- patching
- HA
- Scalability

managed services: SQS, RDS, SNS

Pod is basic unit of K8s

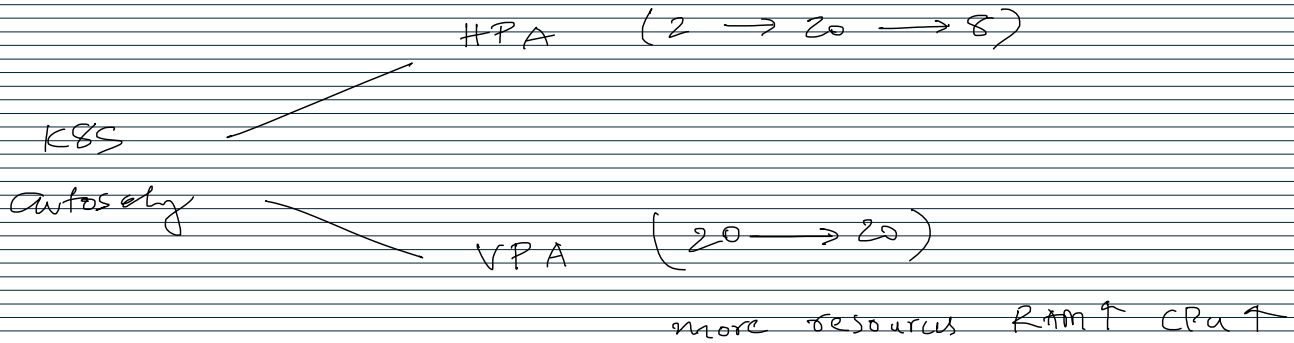


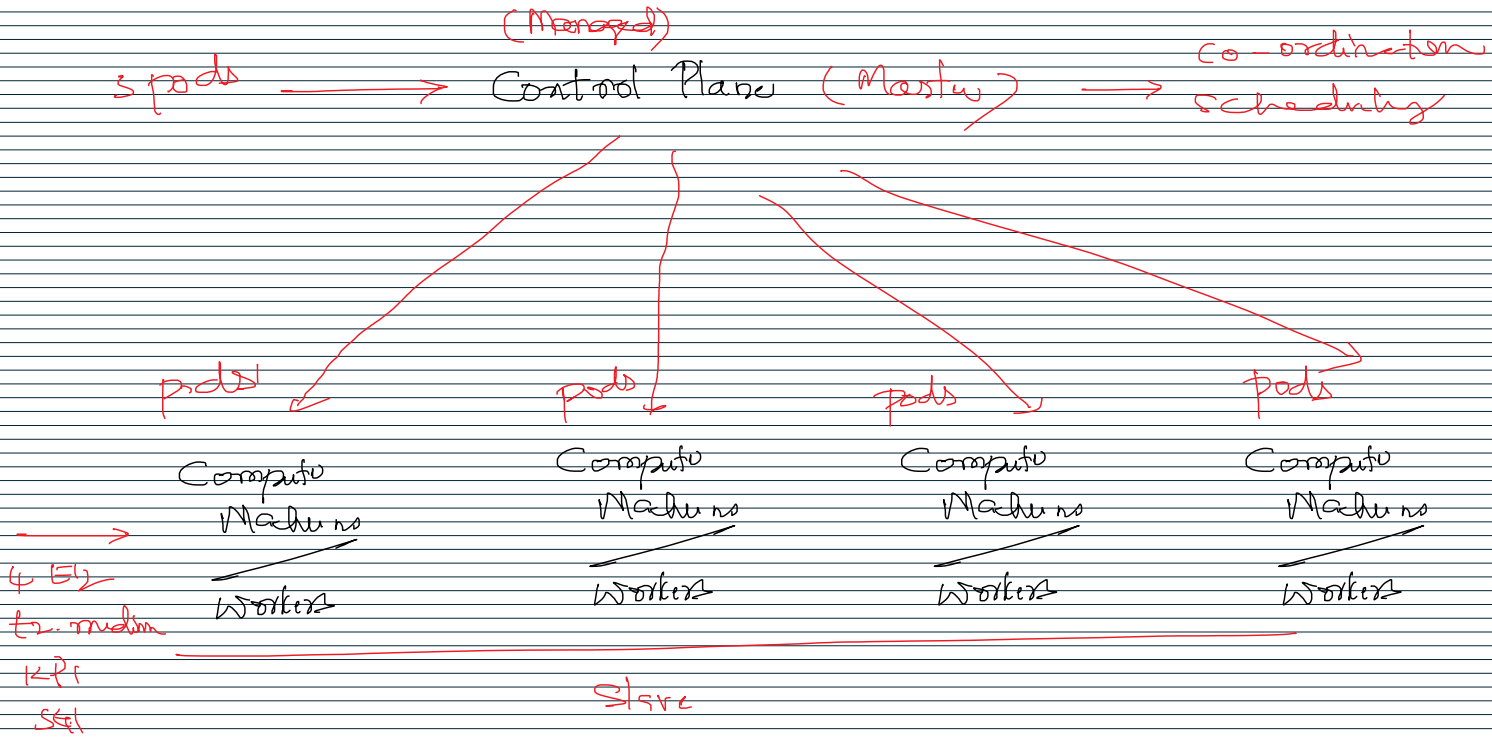
Pod can have single container

9 pods → K8S → Schedule

app log sent

Container


HPA (2 → 20 → 8)
K8S
Autoscaly
VPA (20 → 20)
more resources RAM↑ CPU↑


(Managed)
5 pods
Control Plane (Master)
Co-ordination
Scheduling
pods
pods
pods
pods
Compute
Machine
Worker
Compute
Machine
Worker
Compute
Machine
Worker
Compute
Machine
Worker
4 EC2
t2.medium
KPI
Skill
Store

20 November 2020 10:30

Stateless ⟶ stateful

(WIP)

Ecosystem

f() → GCP, AWS

Knative

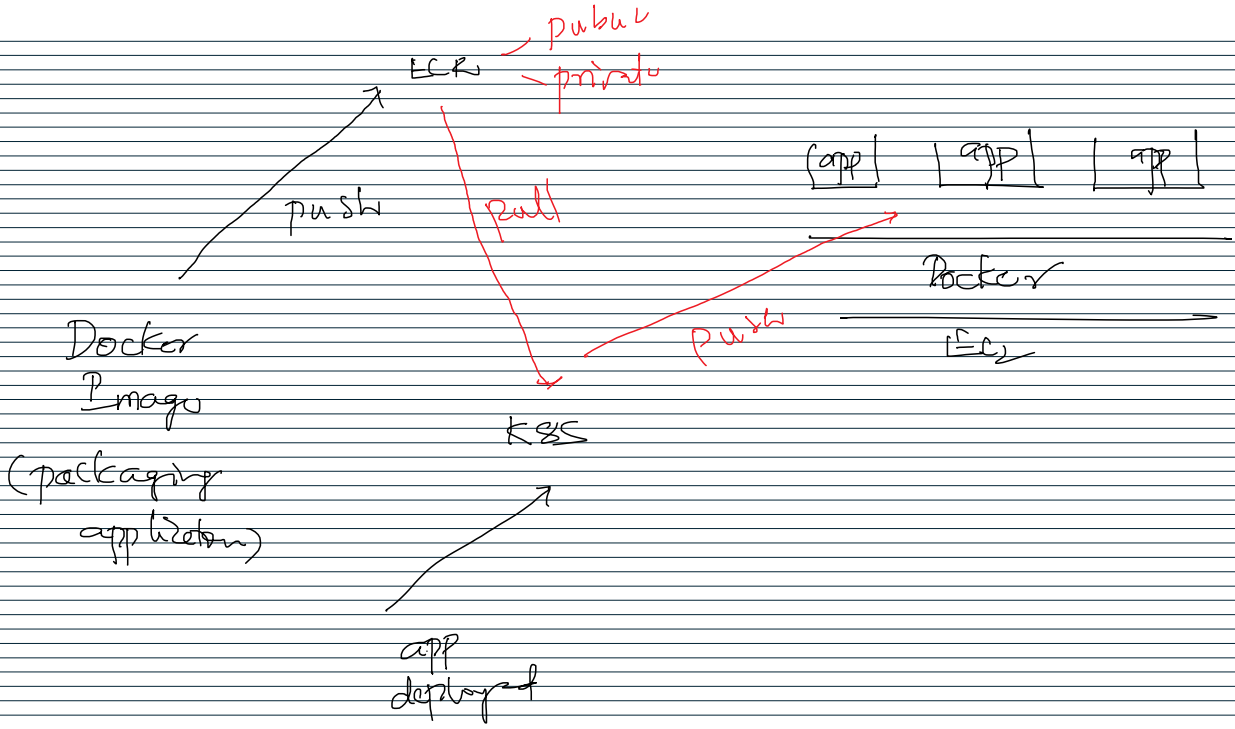
K8S

- Prometheus — monitoring
- Grafana — Visualization
- Helm — appl. dep. easier
- knative — serverless

GCP

λ → AWS

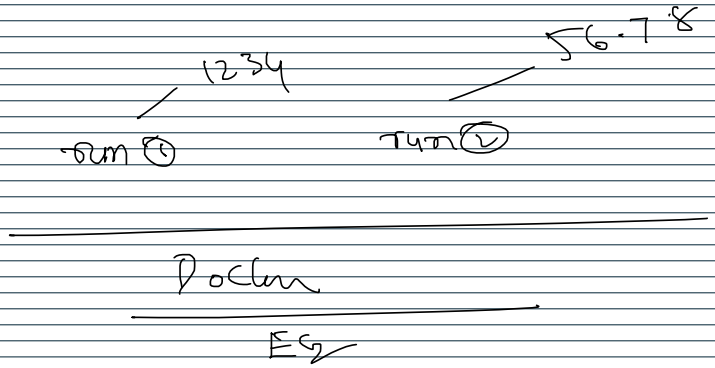
λ → Azure


ECR
public
private
push
pull
app
app
app
Docker
ECS
push
Docker
Image
(packaging
application)
K8S
app
deployed

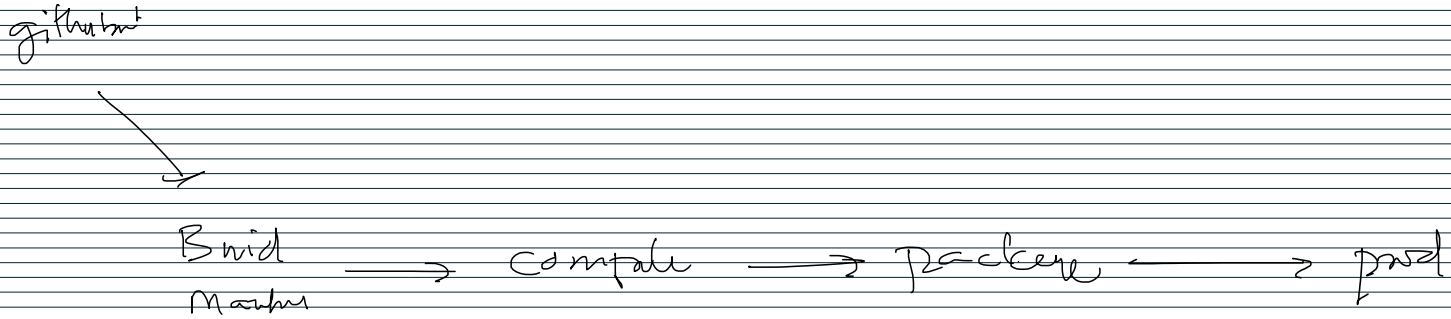
Docker Image

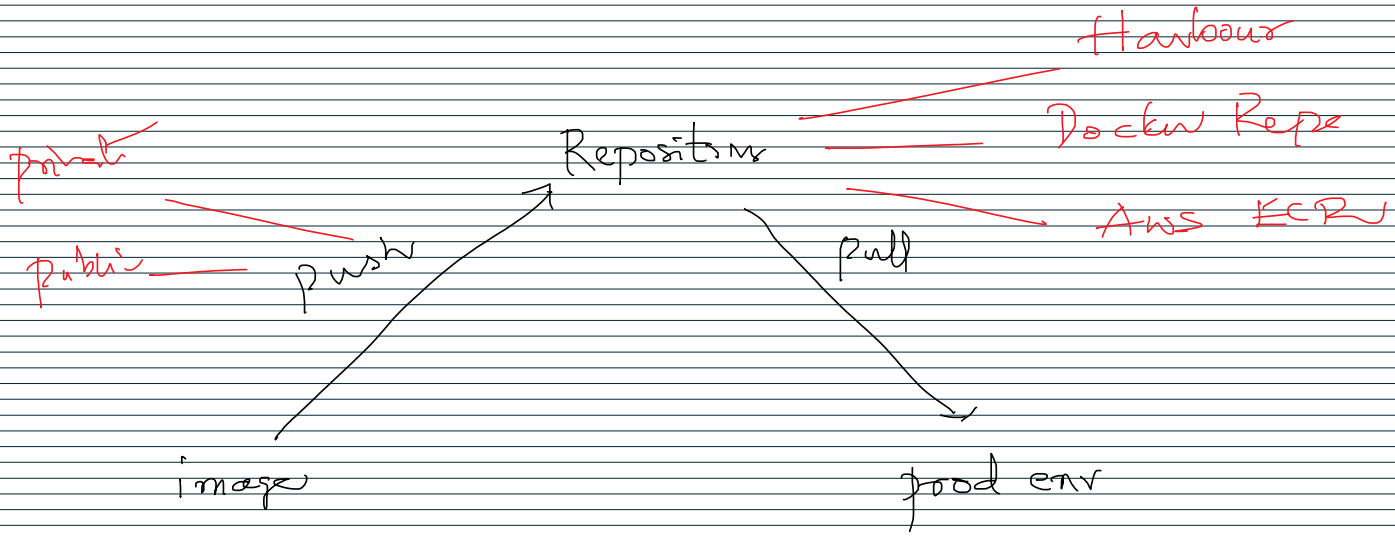
Docker

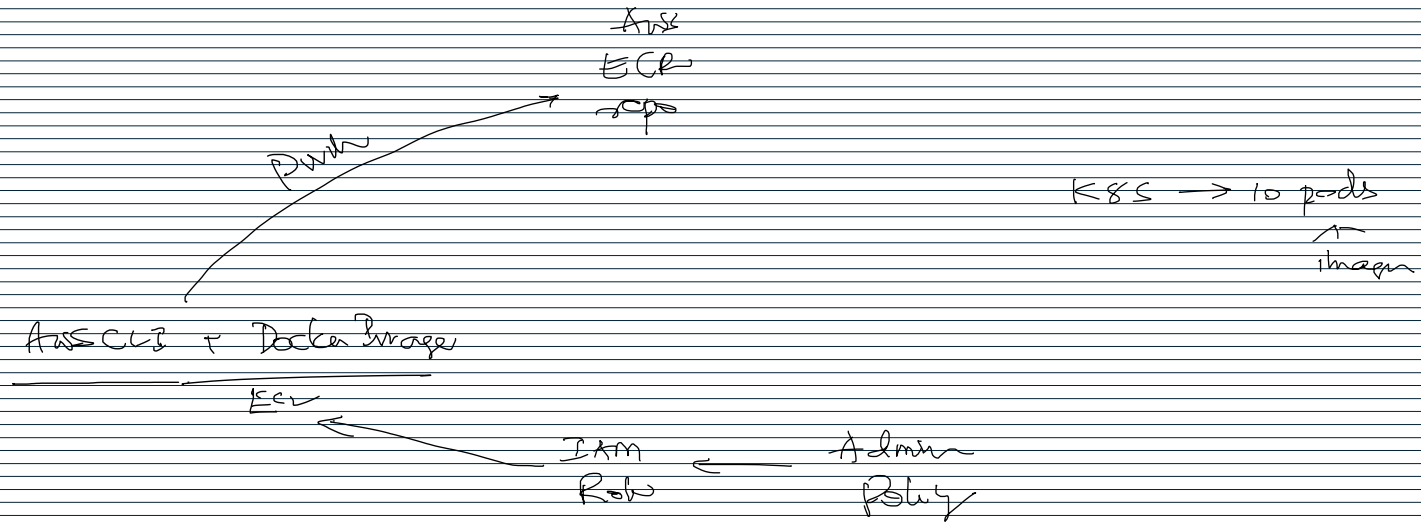
EC2



emphasize


github
Build
Maven
compile
package
post


Harbour
Docker Repo
AWS ECR
Repository
push
pull
image
prod env


Aws
ECR
repo
Push
K8S → 10 pods
image
AWS CLI + Docker Image
ECR
IAM
Role
Admin
Policy

Kubectl ⟶ interact with K.8S

↗

4 pods

AS

aws

Control Plane (Scheduling / Coordination)

└ ~~0.2 $/hr~~ → 0.1 $/hr

~~EKS~~

t2.medium

per second billing

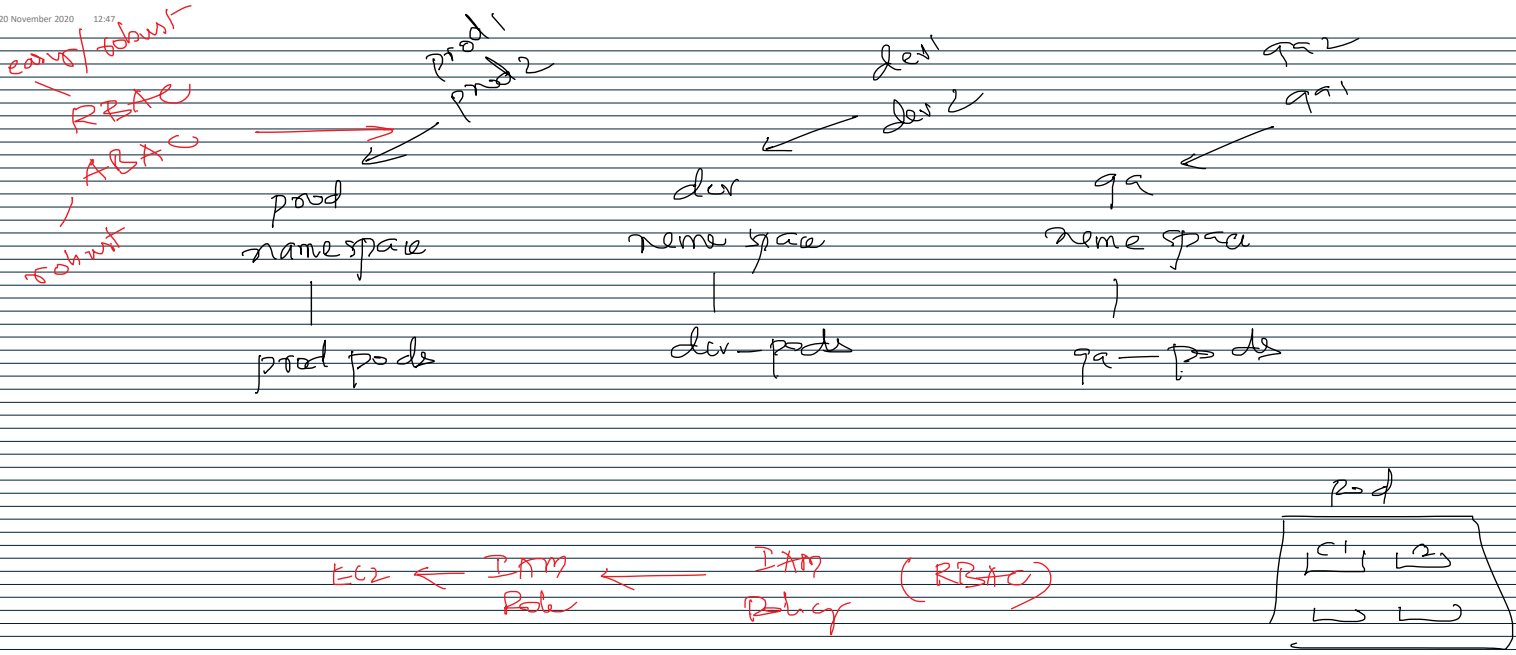
with min of 1 min

Worker nodes

min = 1

max = 3

Start = 1


prod1
prod2
dev1
dev2
qa2
qa1
prod
namespace
prod pods
dev
name space
dev-pods
qa
name space
qa-pods

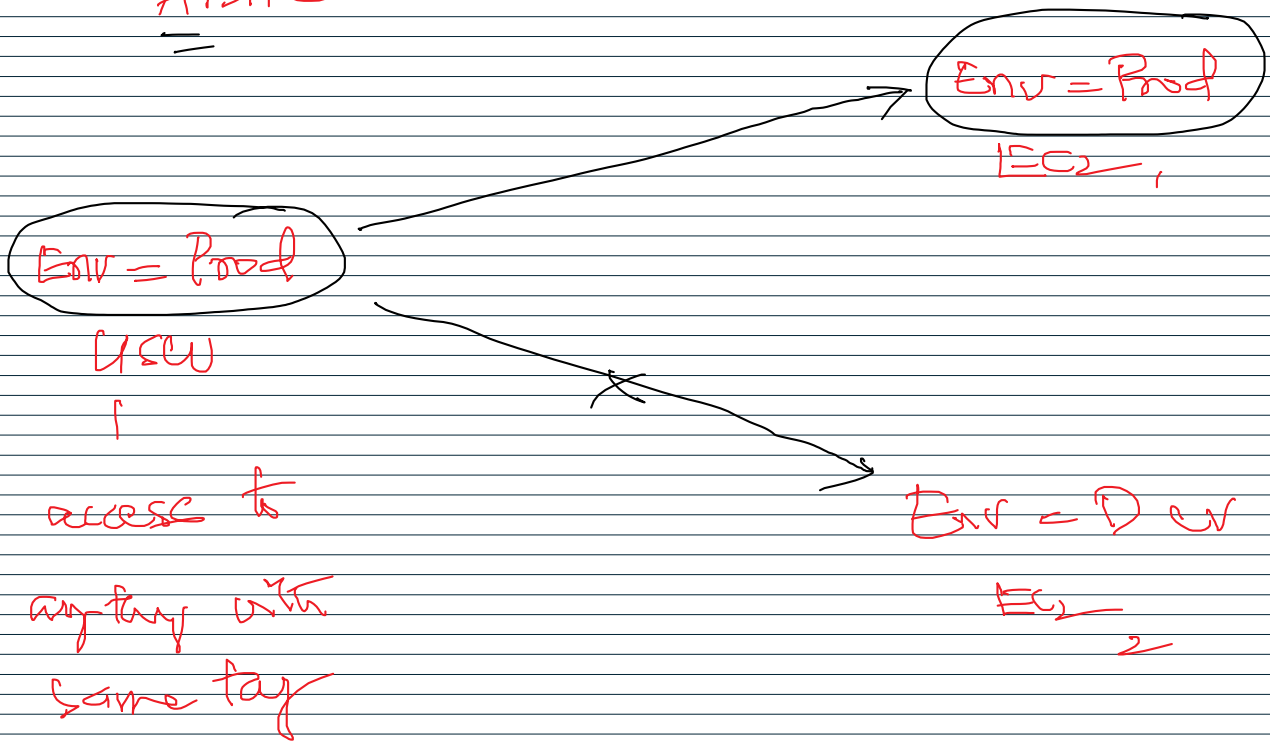
EC2
IAM
Role
IAM
Policy
(RBAC)

Pod
C1
C2

20 November 2020 12:50

Attribute Based Access Control

ABAC



(2) (1)

Deployment ⟶ 3 pods

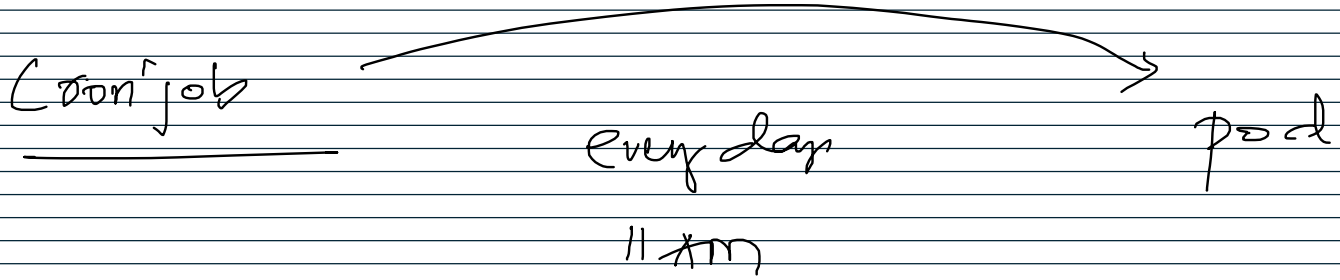
└ makes sure there are enough pods

pod

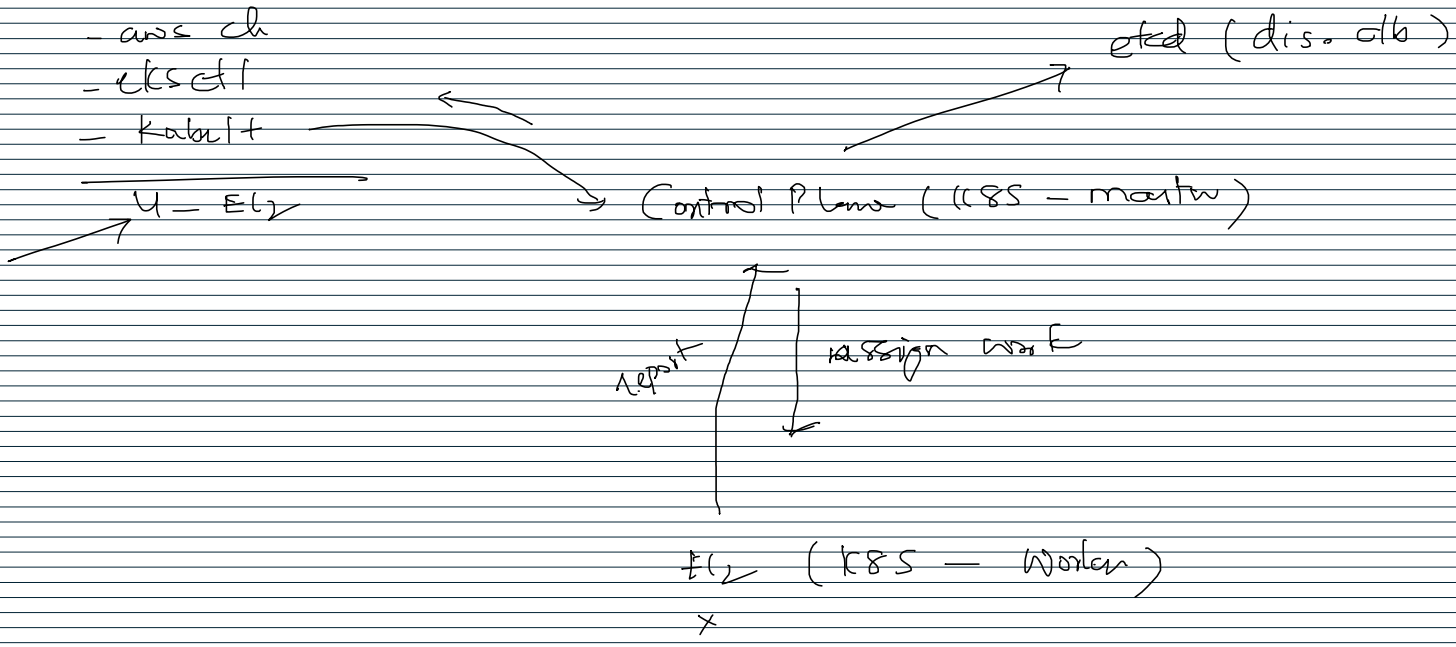
pods 1 ~~pods 2~~ pods 3

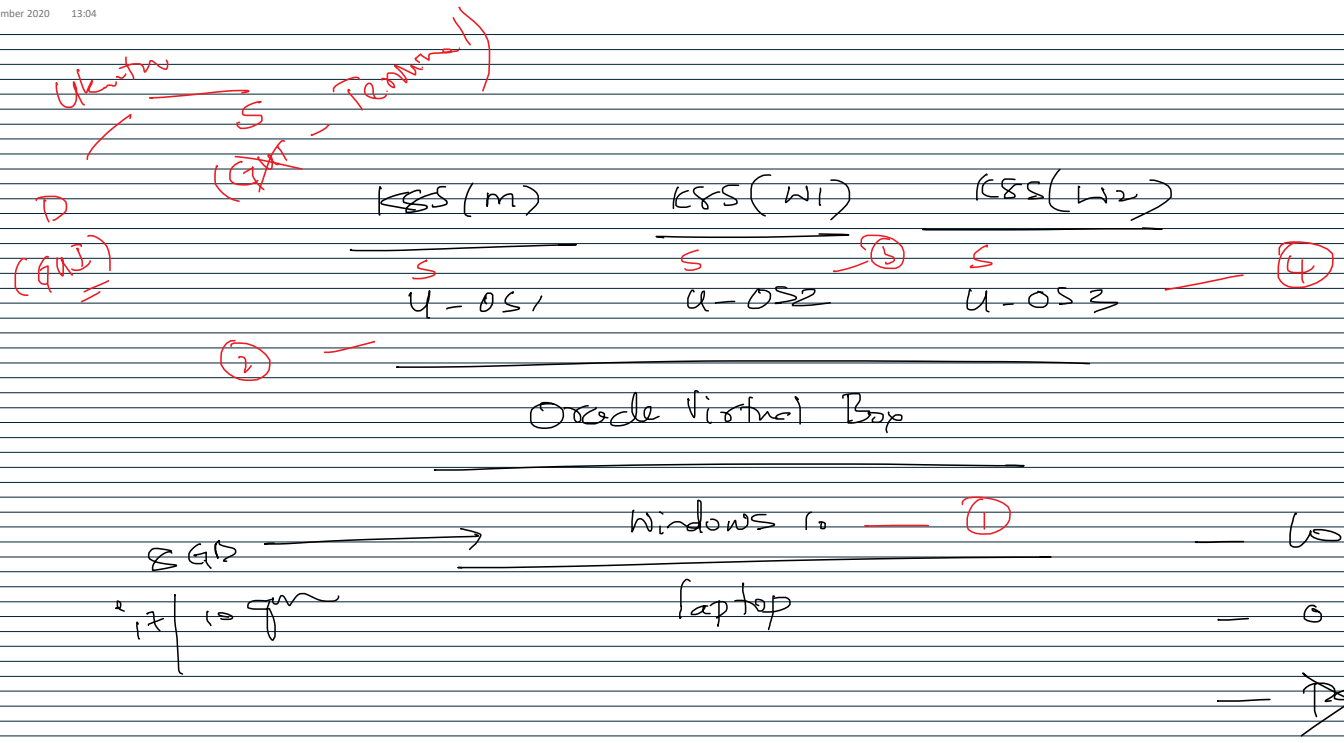
EC2 EC2 EC2

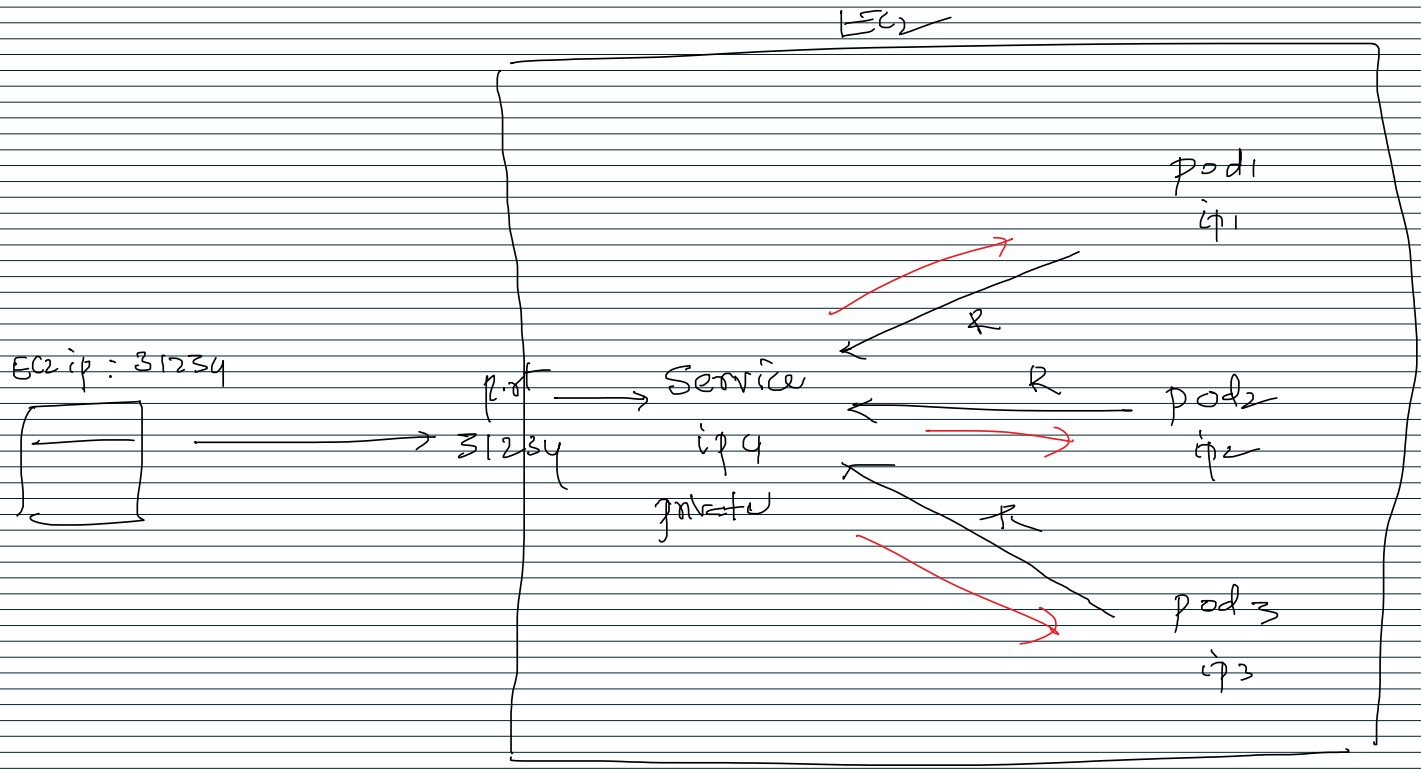
20 November 2020 12:53

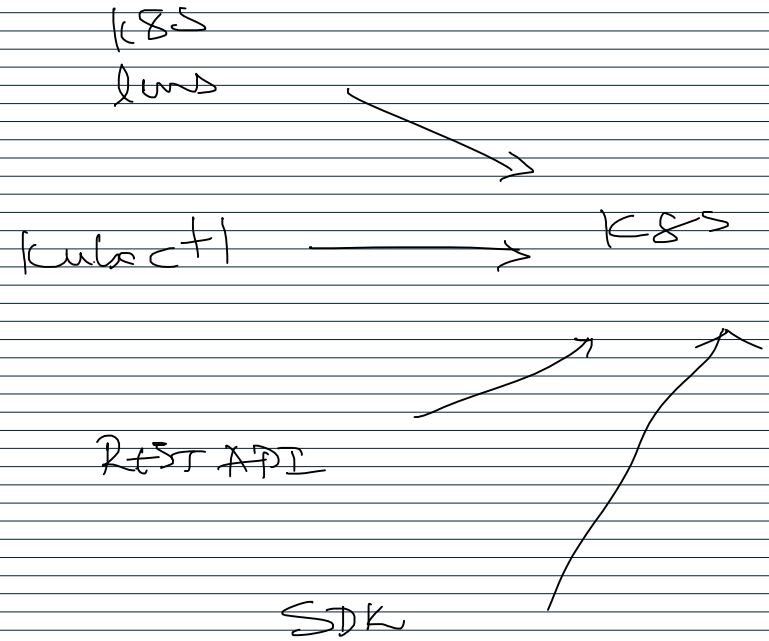


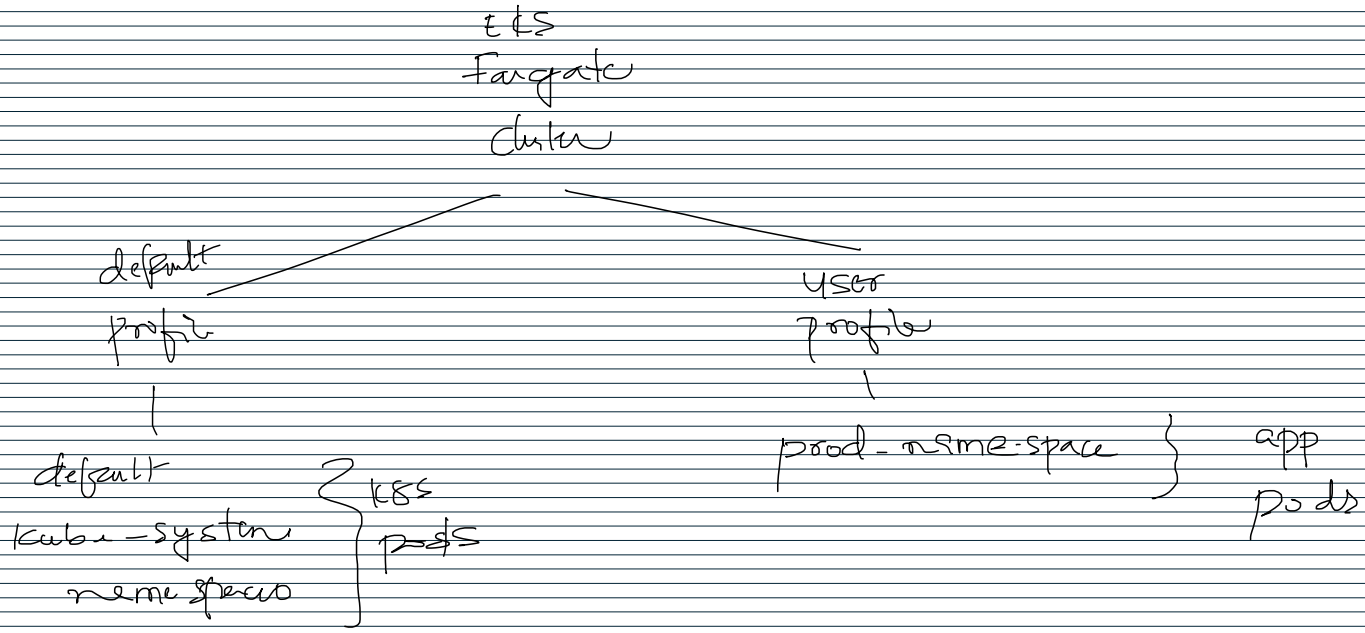
KV

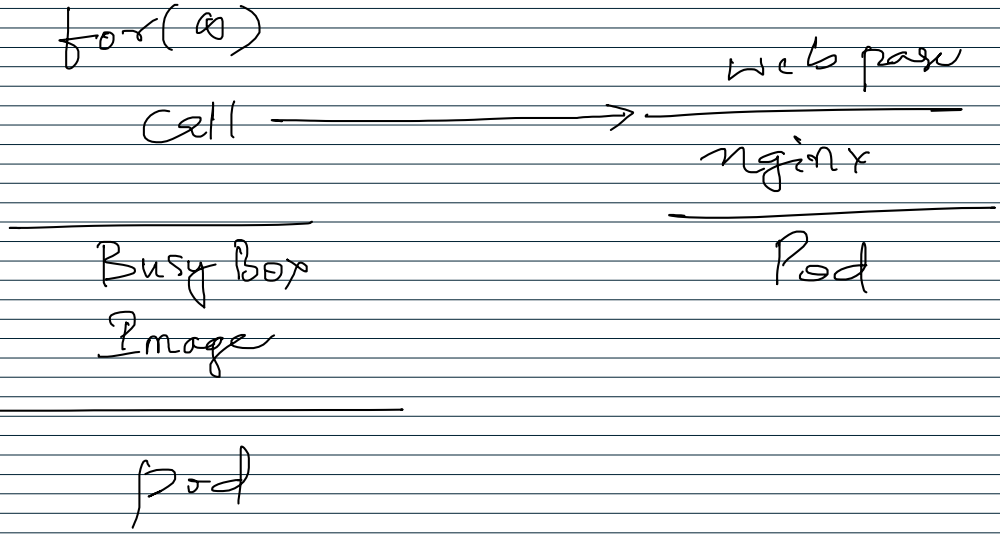

- aws cli
- eksctl
- kubectl
Control Plane (EKS - master)
etcd (dis. db)
report
assign work
EC2 (EKS - worker)

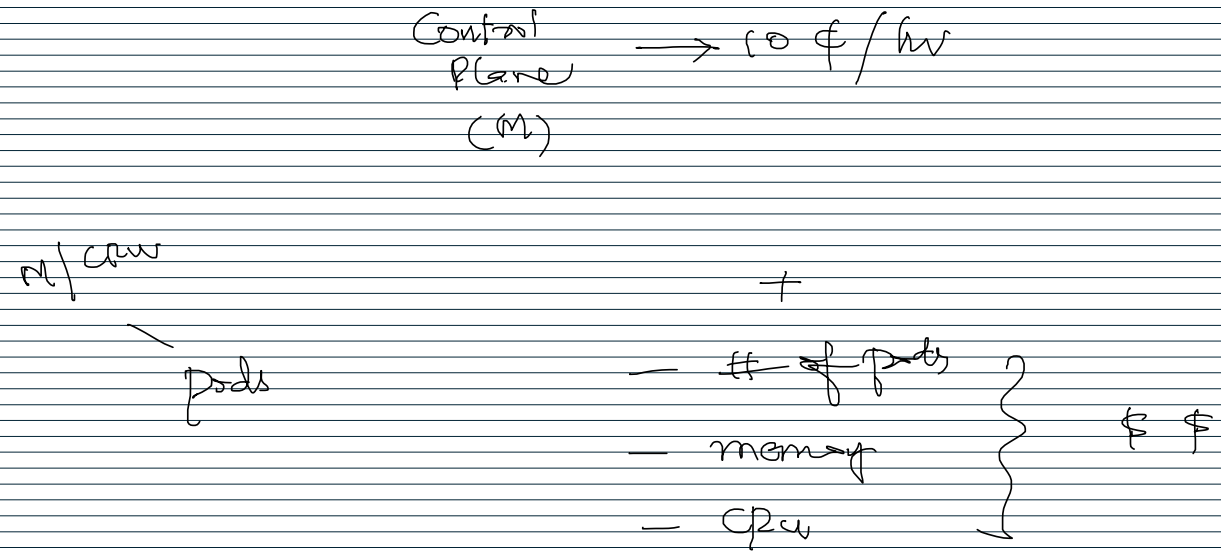

K8S(m)
K8S(W1)
K8S(W2)
U-051
U-052
U-053
Oracle Virtual Box
Windows 10
laptop
locally
offline

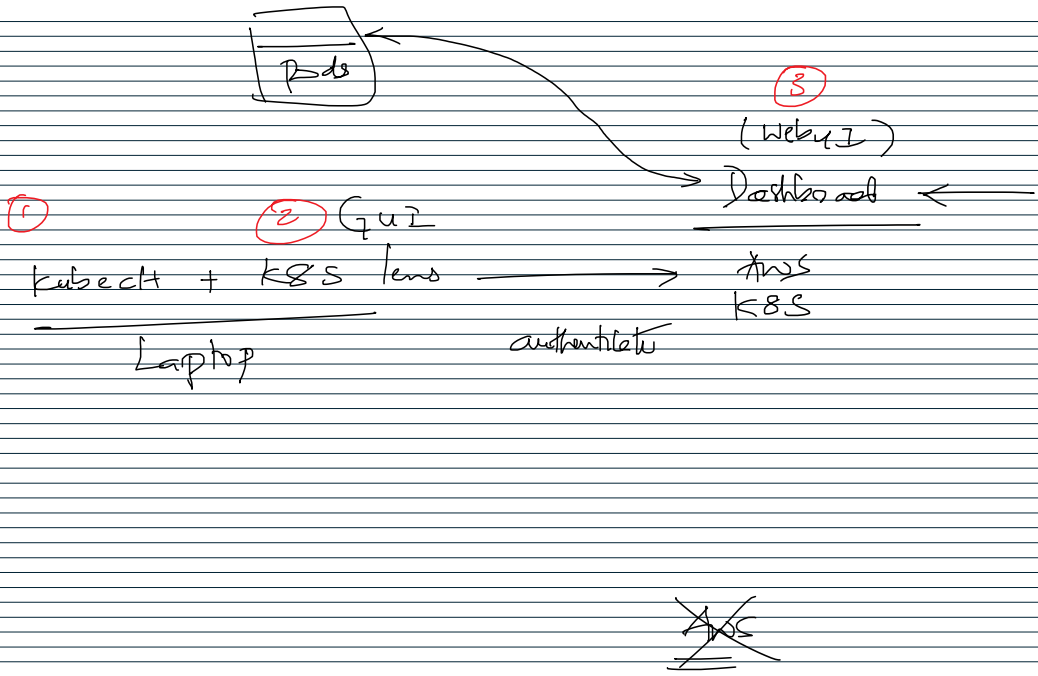

EC2
pod1
ip1
R
EC2 ip : 31234
port
31234
Service
ip4
protocol
R
pod2
ip2
R
pod3
ip3


K8S
kubectl
K8S
REST API
SDK

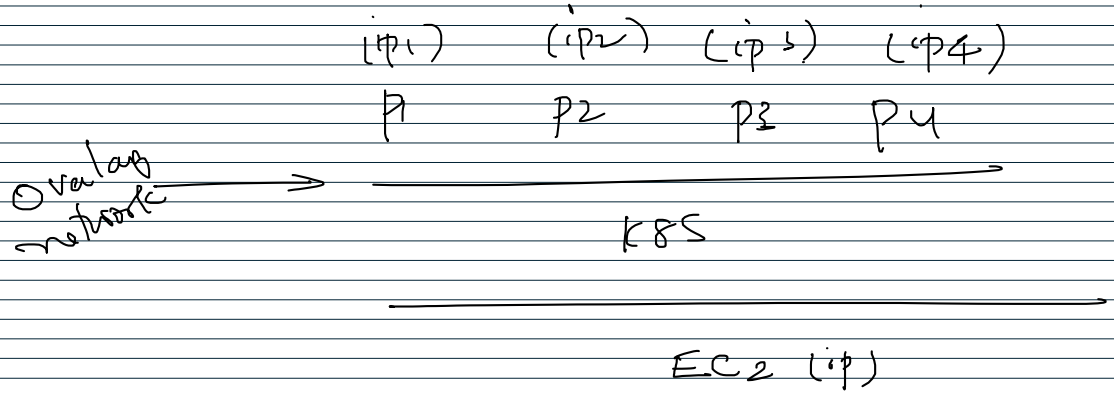

EKS
Fargate
Cluster
default
profile
default
kube-system
namespace
K8S
pods
user
profile
prod-name-space
app
pods

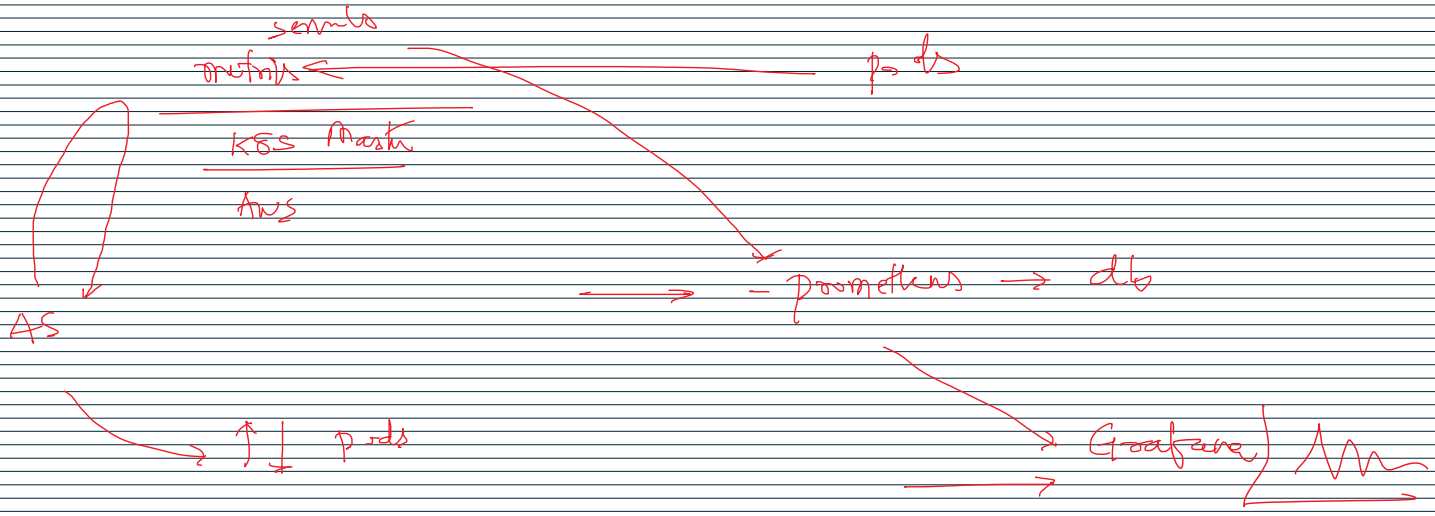

for(∞)
Call
web pod
nginx
Pod
Busy Box
Image
Pod


Control Plane
(M)
→ 10 ¢/hr
M/c
Pods
+
— # of pods
— memory
— CPU
$ $


Pods
3
(Web UI)
Dashboard
1
2
GUI
kubectl + K8S lens
Laptop
authenticate
AWS
K8S

dns




servers
metrics
pods
K8S Master
AWS
AS
Prometheus
db
pods
Grafana

ECS Cluster → Free

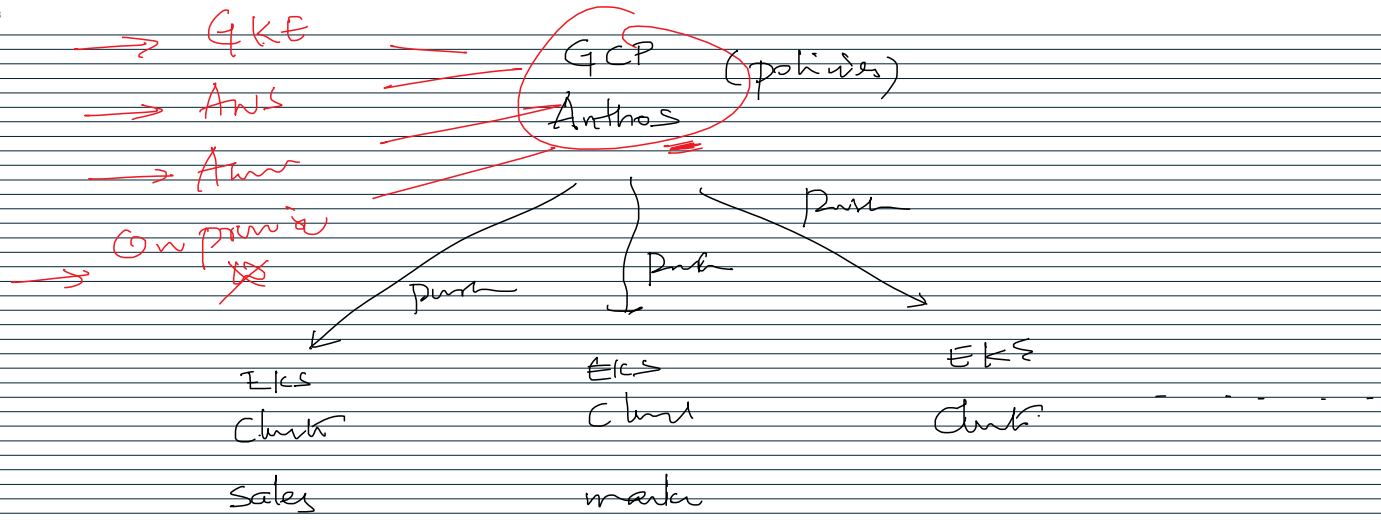
↳ ECS Service

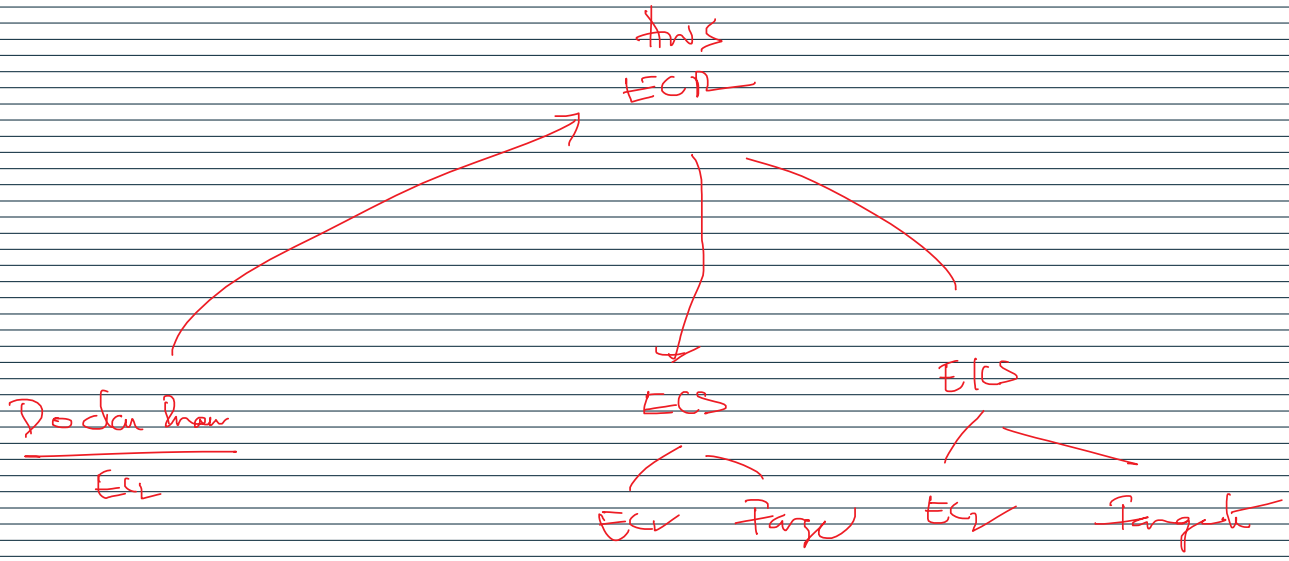
↳ ECS Task

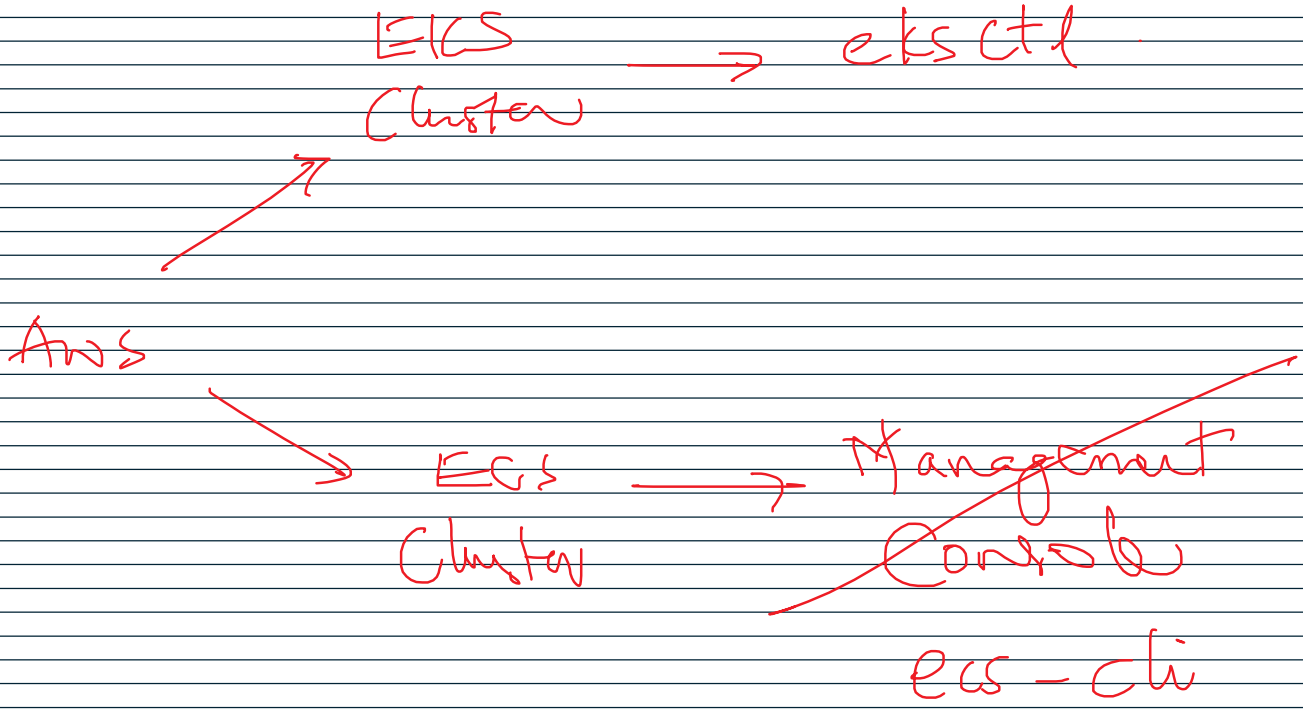
EC2 ↑ $

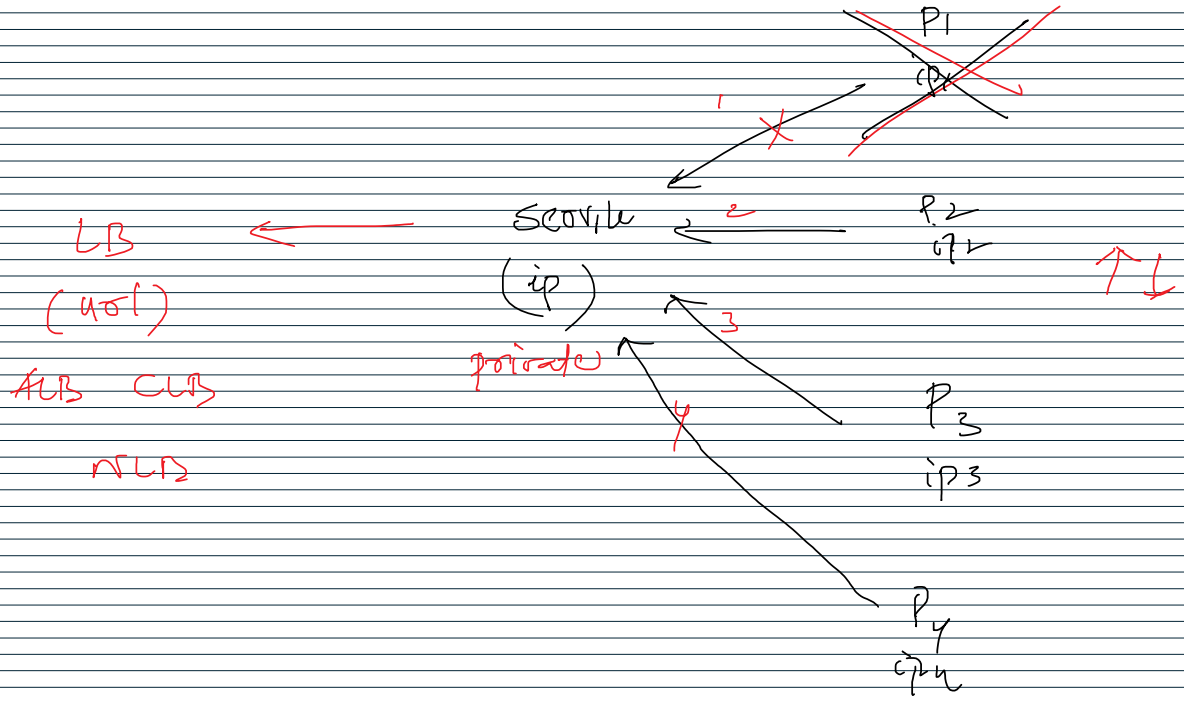
Fargate ↑ 4vu/Pod

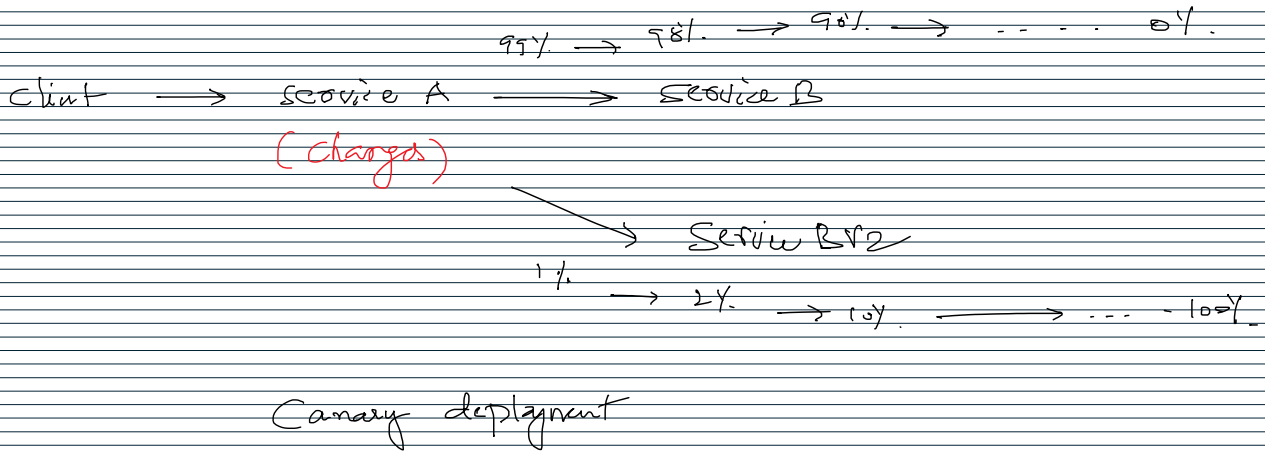
Task def (JSON) (Image, CPU, Memory, ...)


GKE
AWS
Azure
On prem
GCP
Anthos
(policies)
push
push
push
EKS
EKS
Cluster
Cluster
sales
marketing


Aws
ECR
Docker Image
ECS
ECS
ECS
Target
ECS
Target


EKS
Cluster
eksctl
AWS
ECS
Cluster
Management
Console
ecs-cli


P1
1
2
Scarcity
(ip)
P2
3
4
P3
ip3
P4
ip4
LB
(UOI)
ALB CLB
NLB



Canary deployment