A Journey to Production

Ts. Ammar Abdullah

About Me

Ts. Ammar Abdullah

Expertise area - Blockchain (Hyperledger & Ethereum), machine learning , IoT, full stack software development, cloud solution

Bachelor & Master in Engineering at Korea University
Currently pursuing part time PhD at UTM

Samsung Electronics, Korea - Asic Design Engineer - ASIC Chip Design
Funai R&D Malaysia - Assistance Chief Engineer - Embedded Software
OK Blockchain - R&D Manager - Blockchain, AI, Big data, IoT, Cloud, Full stack
Smart Tradzt - Innovation Lead - Blockchain, Machine Learning, Cloud, IoT, Full stack
Ever AI Technologies – CTO

https://www.linkedin.com/in/ammarabd/

Our Target Today

1. Understanding how the actual development flow is done in the industry from development to production

2. Understanding how to design proper software architecture

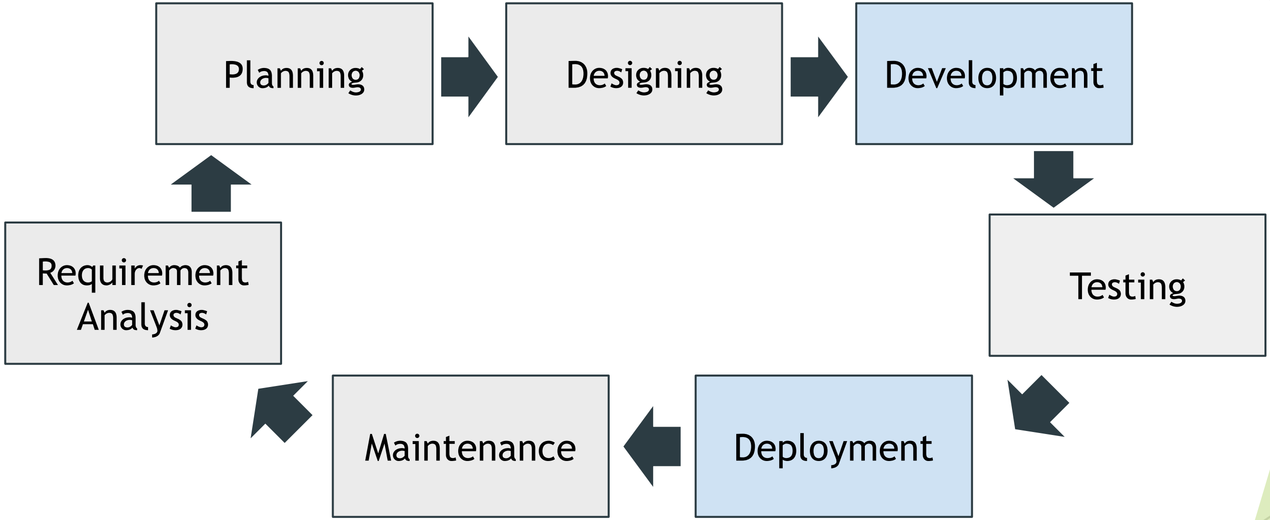
3. Understanding how to write proper code structure which cater most of the requirements

4. Understanding the proper setup for production and how continuous integration can be automated

So, to get solid understanding…

Let’s do more
hands on today !!

Software Development Life Cycle



Software Development Environment

Development

```
class DevConfig(BaseConfig):
   FLASK_ENV = 'development'
   DEBUG = True
   DBNAME = 'DBNAME_DEV'
   SQLALCHEMY_DATABASE_URI = DATABASE_URI
   SQLALCHEMY_TRACK_MODIFICATIONS = True
   SWAGGER = {
       'title': 'Development App',
   }
```

Staging

```
class StagingConfig(BaseConfig):
   FLASK_ENV = 'staging'
   DEBUG = False
   DBNAME = 'DBNAME_STAGE'
   SQLALCHEMY_DATABASE_URI = DATABASE_URI
   SQLALCHEMY_TRACK_MODIFICATIONS = True
   SWAGGER = {
       'title': Staging App',
   }
```

Production

```
class ProdConfig(BaseConfig):
   FLASK_ENV = 'production'
   DEBUG = False
   DBNAME = 'DBNAME'
   SQLALCHEMY_DATABASE_URI = DATABASE_URI
   SQLALCHEMY_TRACK_MODIFICATIONS = True
   SWAGGER = {
       'title': Production App',
   }
```

Important Elements In SW Management

Development Methodology

- Agile
- Waterfall
- etc

Task Management Board

- Jira
- Redmine
- Notion
- Clickup
- Gitscrum

CICD

- Jenkins
- Travis
- Circle
- Bamboo

Repository

- Git
- SVN
- TFS

IDE

- VS code
- Intellij IDEA
- Pycharm

Important Elements In SW Development

Frontend Framework

- ReactJS
- NextJS
- VueJS
- Angular
- Django
- etc

Backend Framework

- NodeJS
- Spring Boot
- Flask
- Django
- Symfony
- etc

Database

- MySQL
- Postgresql
- Mongodb
- Couchdb
- etc

Unit Test Framework

- Chai
- Mocha
- Jasmine
- Pytest
- etc

Code Structure

- MVC
- MVVP
- Flux
- Redux

Important Elements In Production

Deployment

- Kubernetes
- Managed Kubernetes
- AWS ECS
- Docker Swarm

Production Server

- AWS
- GCP
- Azure
- On Premise

Scalability

- Load balancer
- Response time
- Multi processing
- Auto scaling

Security

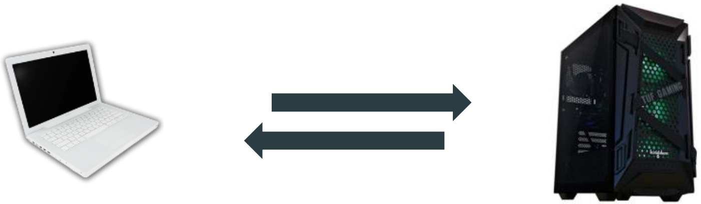
- AWS VPC
- Subnetting
- Access Control
- Open port

Let’s Setup Our GIT Repository

How to manage Git branch ?

Master	Develop	Feature	Release	Hotfix

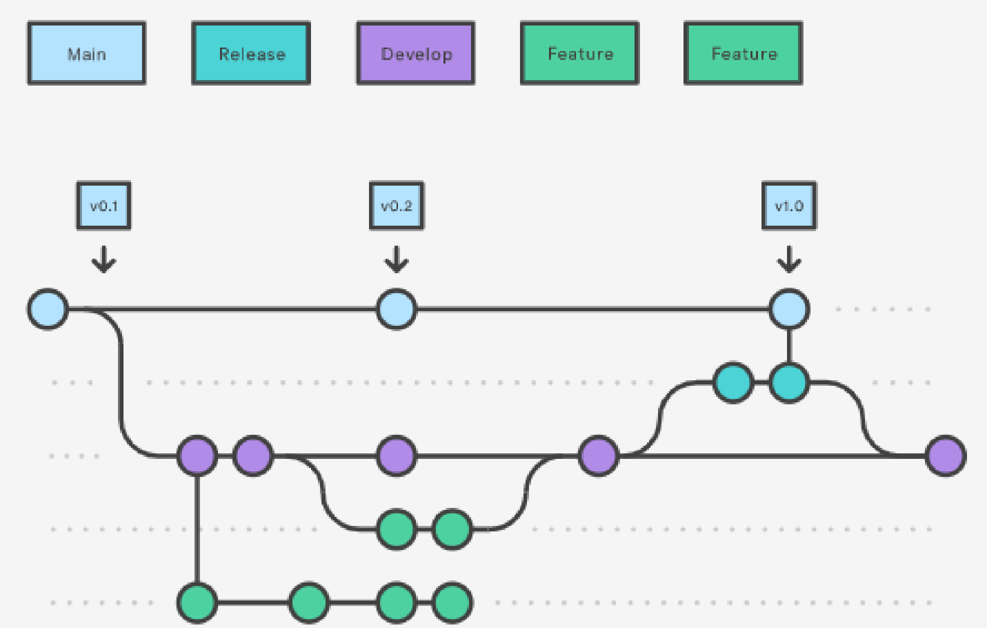
Local PC



Remote (Github / Gitlab / Bitbucket / etc)

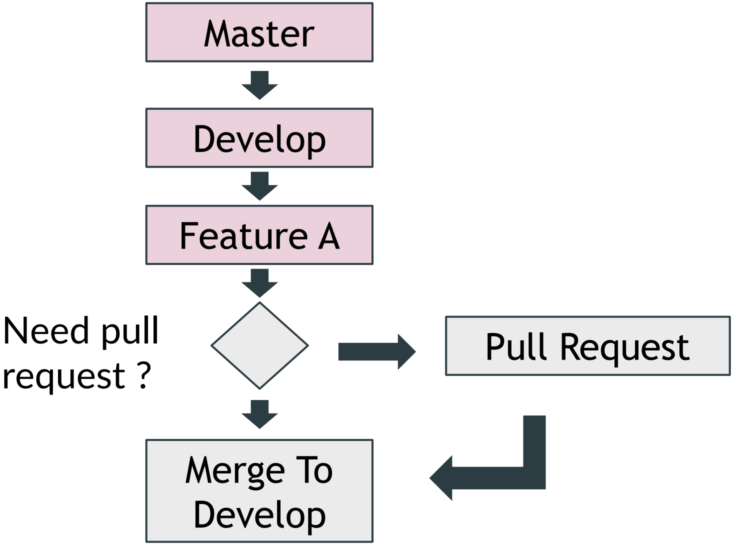
https://www.atlassian.com/git/tutorials/comparing-workflows/gitflow-workflow#:~:text=Gitflow%20is%20a%20legacy%20Git,software%20development%20and%20DevOps%20practices.

How to manage Git branch ?

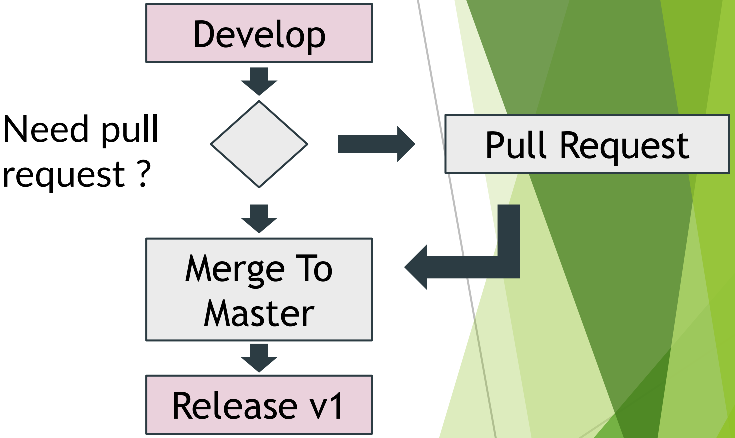


https://www.atlassian.com/git/tutorials/comparing-workflows/gitflow-workflow#:~:text=Gitflow%20is%20a%20legacy%20Git,software%20development%20and%20DevOps%20practices.

How to merge between Git branch ?



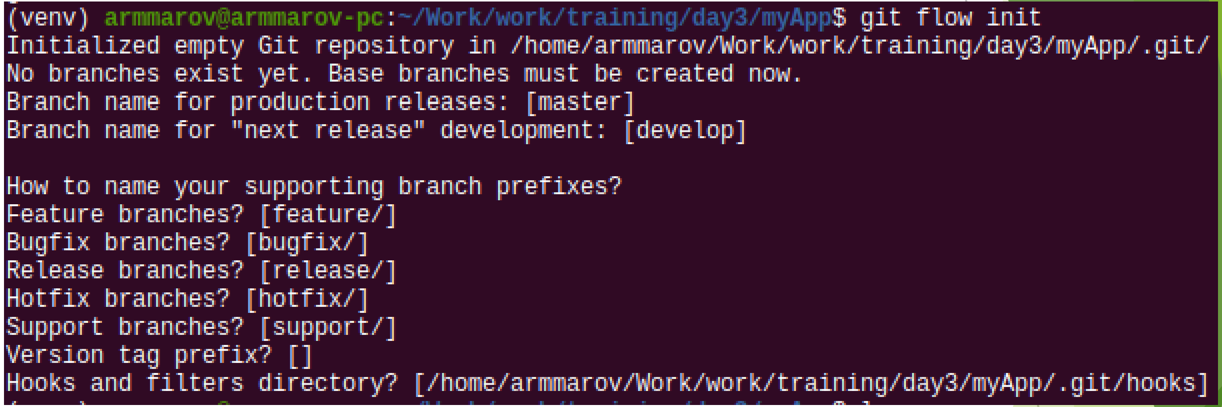
Use case for developer to merge feature branch



Use case for developer to merge to release branch

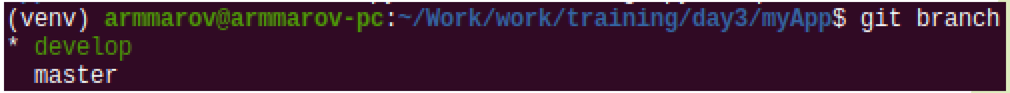
How to start with Git Flow

# git flow init



```
(venv) armmarov@armmarov-pc:~/Work/work/training/day3/myApp$ git flow init
Initialized empty Git repository in /home/armmarov/Work/work/training/day3/myApp/.git/
No branches exist yet. Base branches must be created now.
Branch name for production releases: [master]
Branch name for "next release" development: [develop]

How to name your supporting branch prefixes?
Feature branches? [feature/]
Bugfix branches? [bugfix/]
Release branches? [release/]
Hotfix branches? [hotfix/]
Support branches? [support/]
Version tag prefix? []
Hooks and filters directory? [/home/armmarov/Work/work/training/day3/myApp/.git/hooks]
```

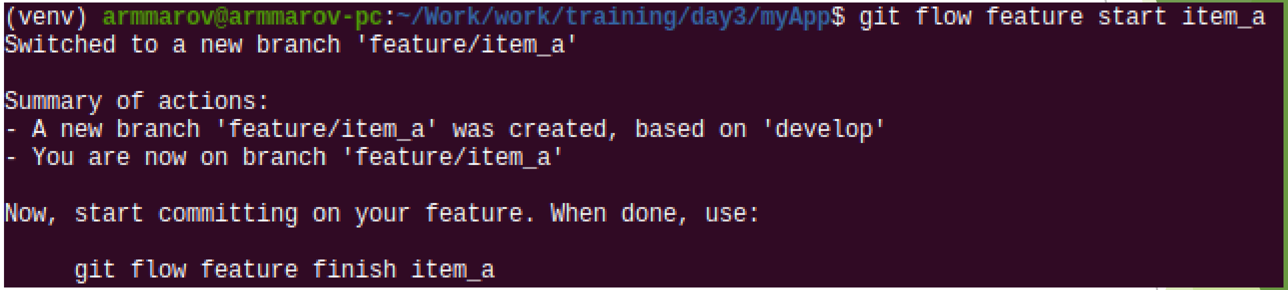
# git branch



```
(venv) armmarov@armmarov-pc:~/Work/work/training/day3/myApp$ git branch
* develop
  master
```

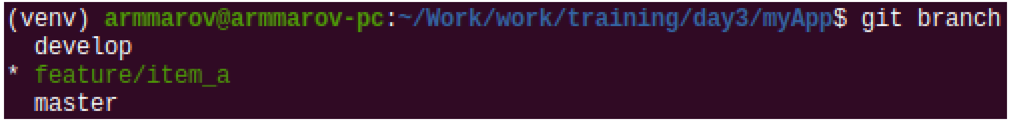
How to create feature branch ?

# git flow feature start <feature-name>



```
(venv) armmarov@armmarov-pc:~/Work/work/training/day3/myApp$ git flow feature start item_a
Switched to a new branch 'feature/item_a'

Summary of actions:
- A new branch 'feature/item_a' was created, based on 'develop'
- You are now on branch 'feature/item_a'

Now, start committing on your feature. When done, use:

     git flow feature finish item_a
```

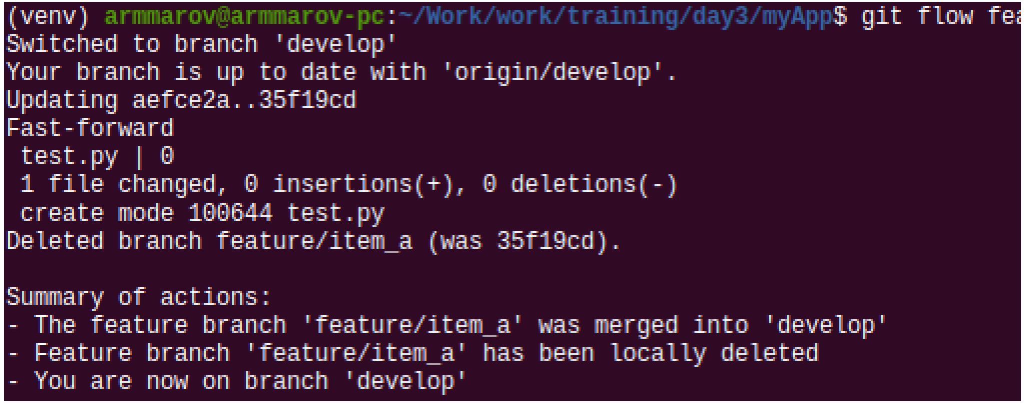
# git branch



```
(venv) armmarov@armmarov-pc:~/Work/work/training/day3/myApp$ git branch
  develop
* feature/item_a
  master
```

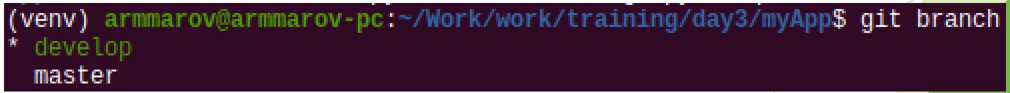
How to finish feature branch w/o pull request ?

# git add .
# git commit -m “Add new file”
# git flow feature finish <feature-name>



```
(venv) armmarov@armmarov-pc:~/Work/work/training/day3/myApp$ git flow fea
Switched to branch 'develop'
Your branch is up to date with 'origin/develop'.
Updating aefce2a..35f19cd
Fast-forward
 test.py | 0
 1 file changed, 0 insertions(+), 0 deletions(-)
 create mode 100644 test.py
Deleted branch feature/item_a (was 35f19cd).

Summary of actions:
- The feature branch 'feature/item_a' was merged into 'develop'
- Feature branch 'feature/item_a' has been locally deleted
- You are now on branch 'develop'
```

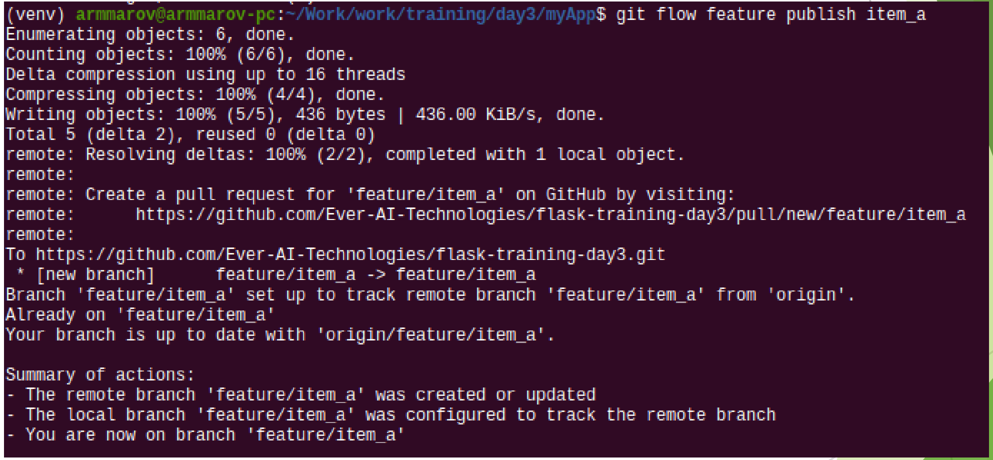
# git branch



```
(venv) armmarov@armmarov-pc:~/Work/work/training/day3/myApp$ git branch
* develop
  master
```

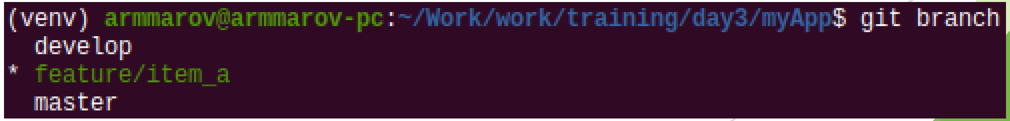
How to create pull request ?

# git add .
# git commit -m “Add new file”
# git flow feature publish <feature-name>



```
(venv) armmarov@armmarov-pc:~/Work/work/training/day3/myApp$ git flow feature publish item_a
Enumerating objects: 6, done.
Counting objects: 100% (6/6), done.
Delta compression using up to 16 threads
Compressing objects: 100% (4/4), done.
Writing objects: 100% (5/5), 436 bytes | 436.00 KiB/s, done.
Total 5 (delta 2), reused 0 (delta 0)
remote: Resolving deltas: 100% (2/2), completed with 1 local object.
remote:
remote: Create a pull request for 'feature/item_a' on GitHub by visiting:
remote:      https://github.com/Ever-AI-Technologies/flask-training-day3/pull/new/feature/item_a
remote:
To https://github.com/Ever-AI-Technologies/flask-training-day3.git
 * [new branch]      feature/item_a -> feature/item_a
Branch 'feature/item_a' set up to track remote branch 'feature/item_a' from 'origin'.
Already on 'feature/item_a'
Your branch is up to date with 'origin/feature/item_a'.

Summary of actions:
- The remote branch 'feature/item_a' was created or updated
- The local branch 'feature/item_a' was configured to track the remote branch
- You are now on branch 'feature/item_a'
```

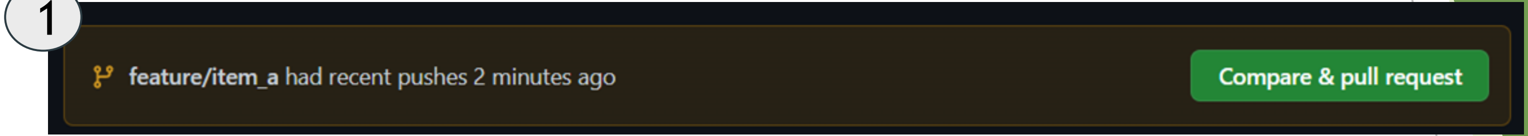
# git branch



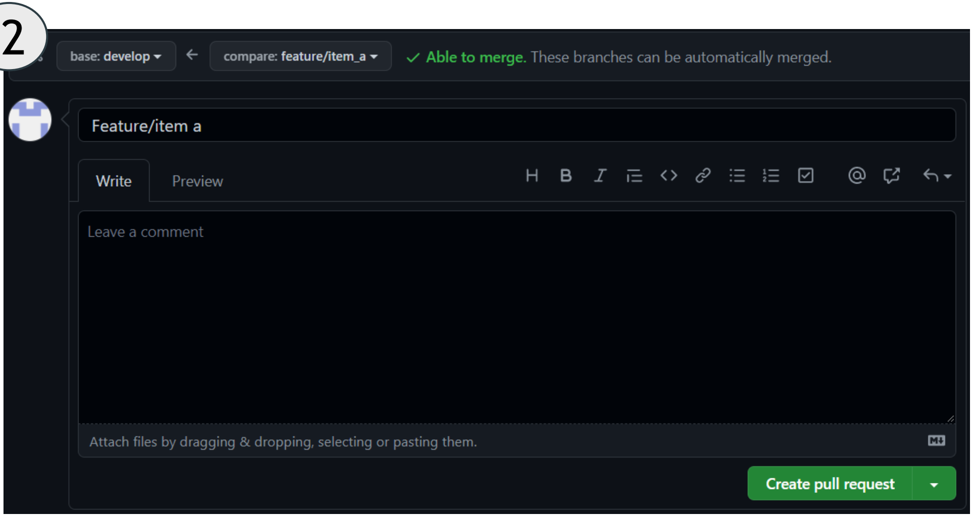
```
(venv) armmarov@armmarov-pc:~/Work/work/training/day3/myApp$ git branch
  develop
* feature/item_a
  master
```

How to create and merge pull request ?

1


feature/item_a had recent pushes 2 minutes ago

2

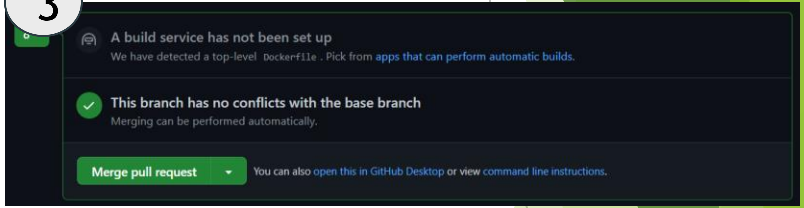

base: develop ← compare: feature/item_a ✓ Able to merge. These branches can be automatically merged.

Feature/item a

Leave a comment

Attach files by dragging & dropping, selecting or pasting them.

3


A build service has not been set up

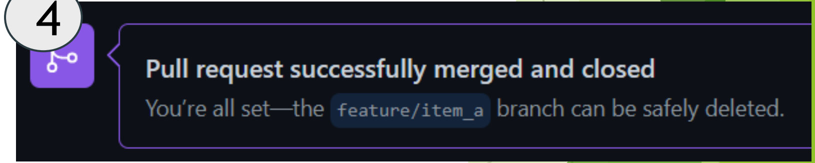
We have detected a top-level `Dockerfile` . Pick from apps that can perform automatic builds.

This branch has no conflicts with the base branch

Merging can be performed automatically.

You can also open this in GitHub Desktop or view command line instructions.

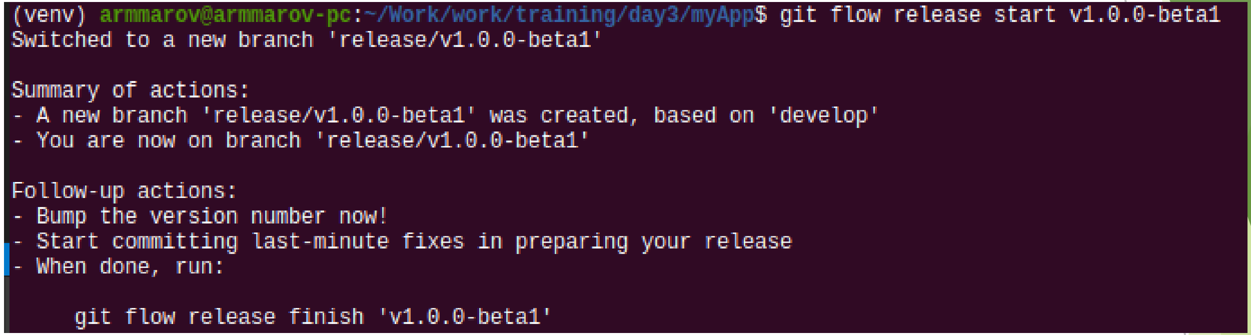
4


Pull request successfully merged and closed

You're all set—the `feature/item_a` branch can be safely deleted.

How to create and publish release branch ?

# git flow release start <release-version>



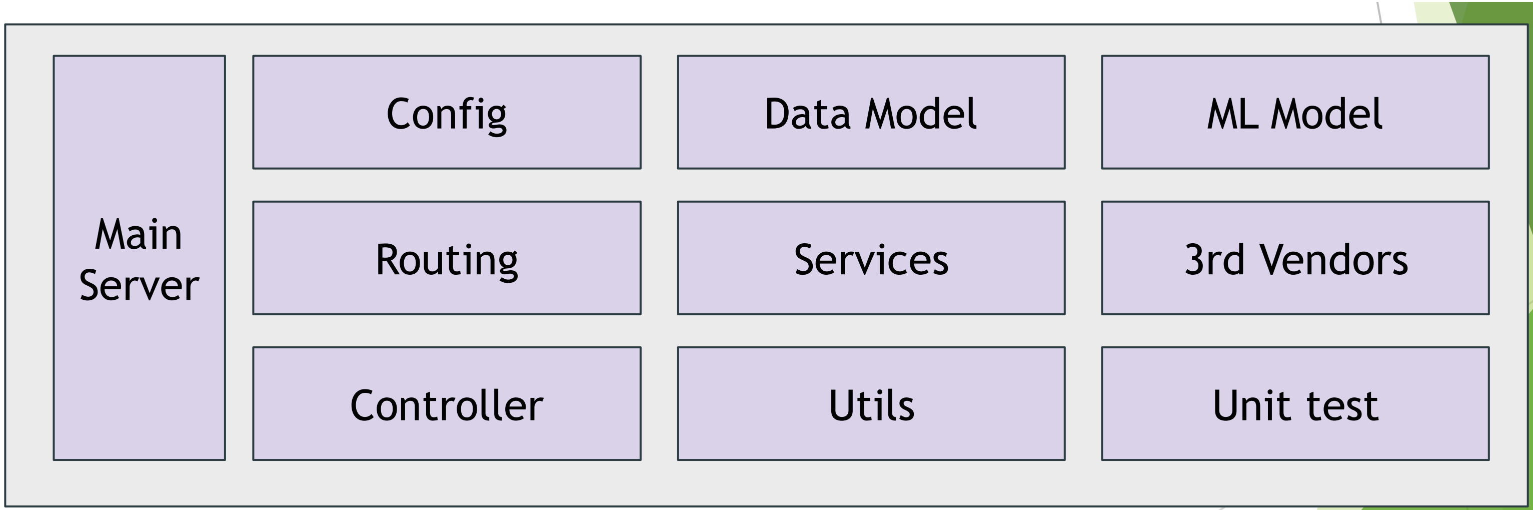
```
(venv) armmarov@armmarov-pc:~/Work/work/training/day3/myApp$ git flow release start v1.0.0-beta1
Switched to a new branch 'release/v1.0.0-beta1'

Summary of actions:
- A new branch 'release/v1.0.0-beta1' was created, based on 'develop'
- You are now on branch 'release/v1.0.0-beta1'

Follow-up actions:
- Bump the version number now!
- Start committing last-minute fixes in preparing your release
- When done, run:

     git flow release finish 'v1.0.0-beta1'
```

# git flow release publish <release-version>

Let’s restructure our codes

Designing Software Architecture

Flask



Main Server

Config

Data Model

ML Model

Routing

Services

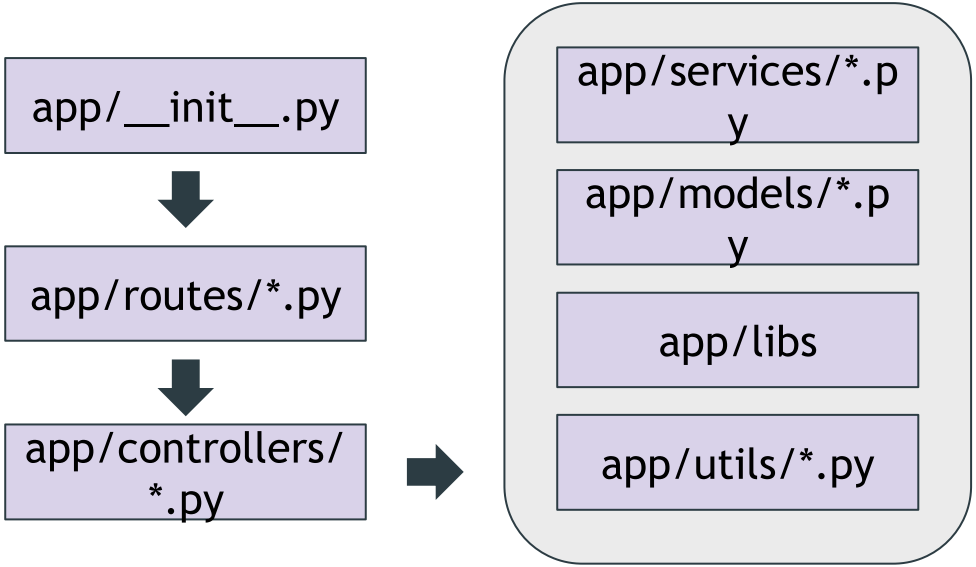
3rd Vendors

Controller

Utils

Unit test

How the structure looks like ?



app/__init__.py

app/routes/*.py

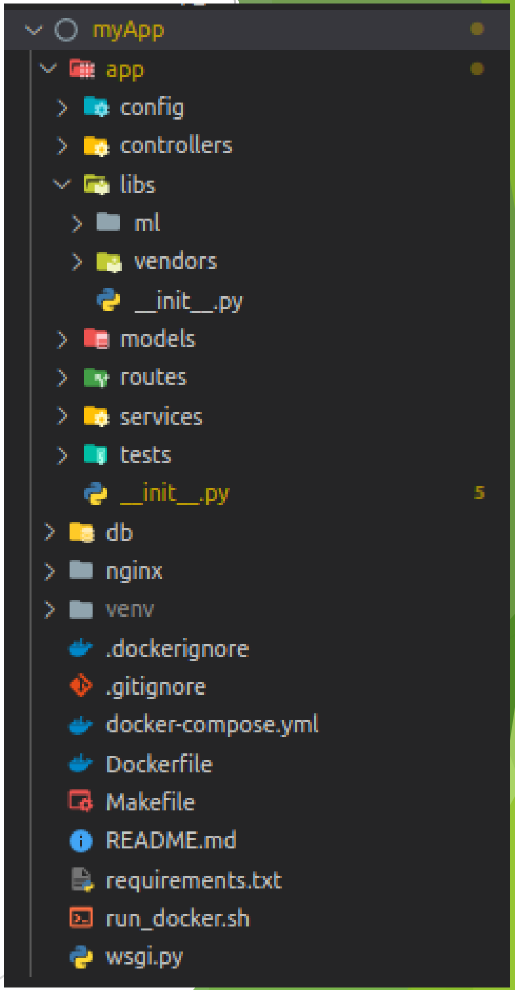
app/controllers/*.py

app/services/*.py

app/models/*.py

app/libs

app/utils/*.py



```
myApp
  app
    config
    controllers
    libs
      ml
      vendors
      __init__.py
    models
    routes
    services
    tests
    __init__.py                5
  db
  nginx
  venv
  .dockerignore
  .gitignore
  docker-compose.yml
  Dockerfile
  Makefile
  README.md
  requirements.txt
  run_docker.sh
  wsgi.py
```

Let us look into the codes deeper

https://github.com/armmarov/journey-to-production.git

Please do me a favour ^^

Create an API with the endpoint “POST /add-me”, and the following json data

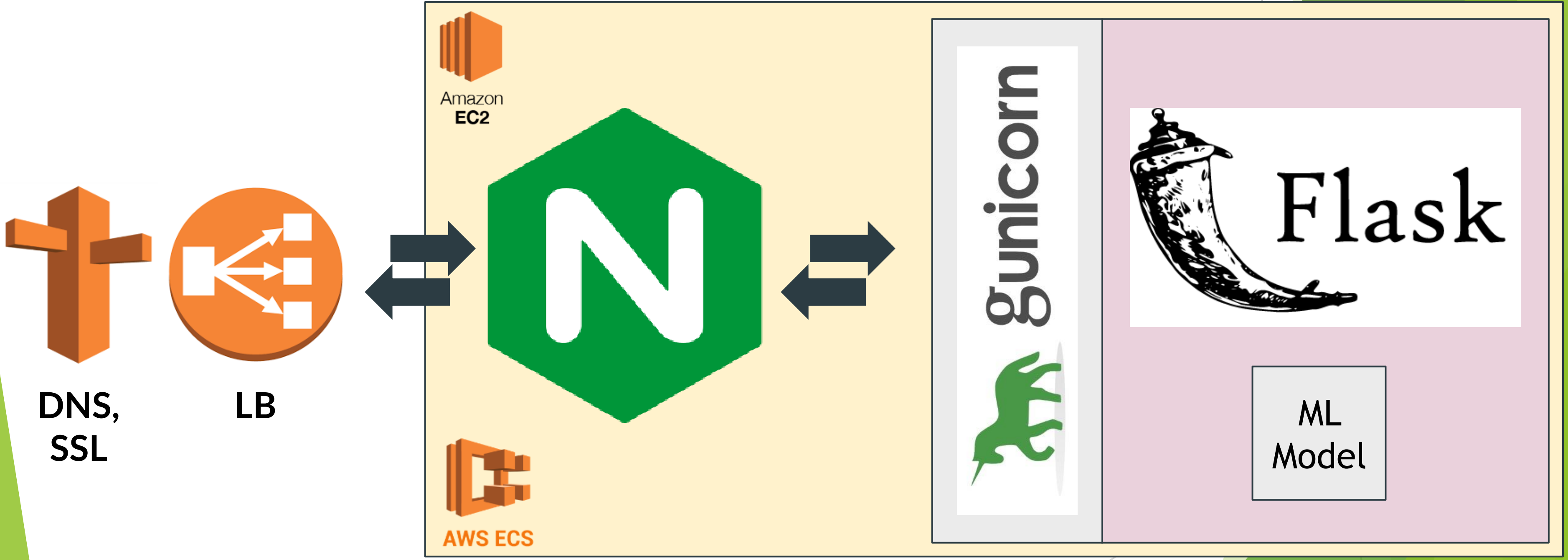
{ a: 4, b: 5 }

The function should be able to compute a + b and return me a value

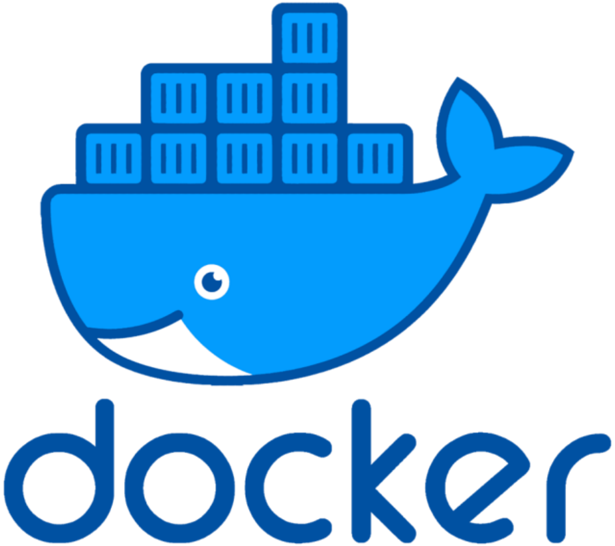
Please do pull request, and i will merge and let others test your function

Let’s Design Our Production Architecture

Designing Production Architecture

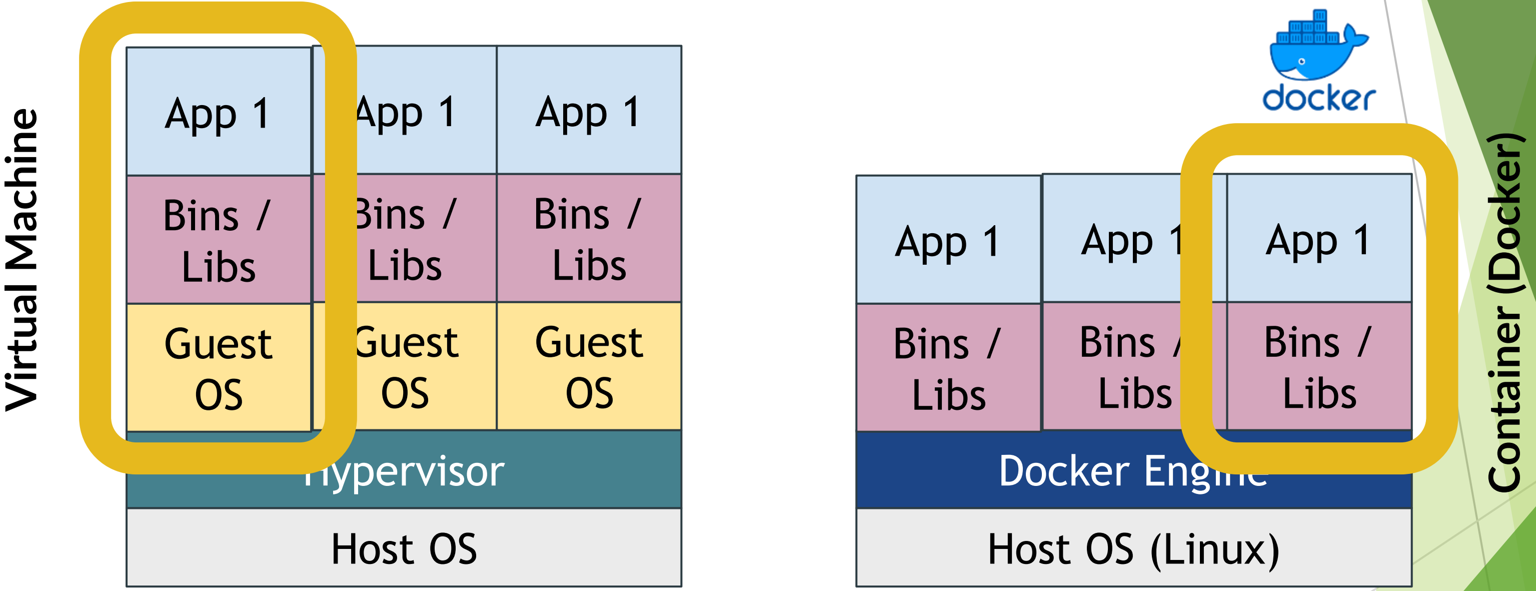


How to set the whole architecture up ?



Thanks to this
WHALE !!!

But wait, what is the Containerization in the first place ?



Virtual Machine

App 1 | App 1 | App 1

Bins / Libs | Bins / Libs | Bins / Libs

Guest OS | Guest OS | Guest OS

Hypervisor

Host OS

docker

Container (Docker)

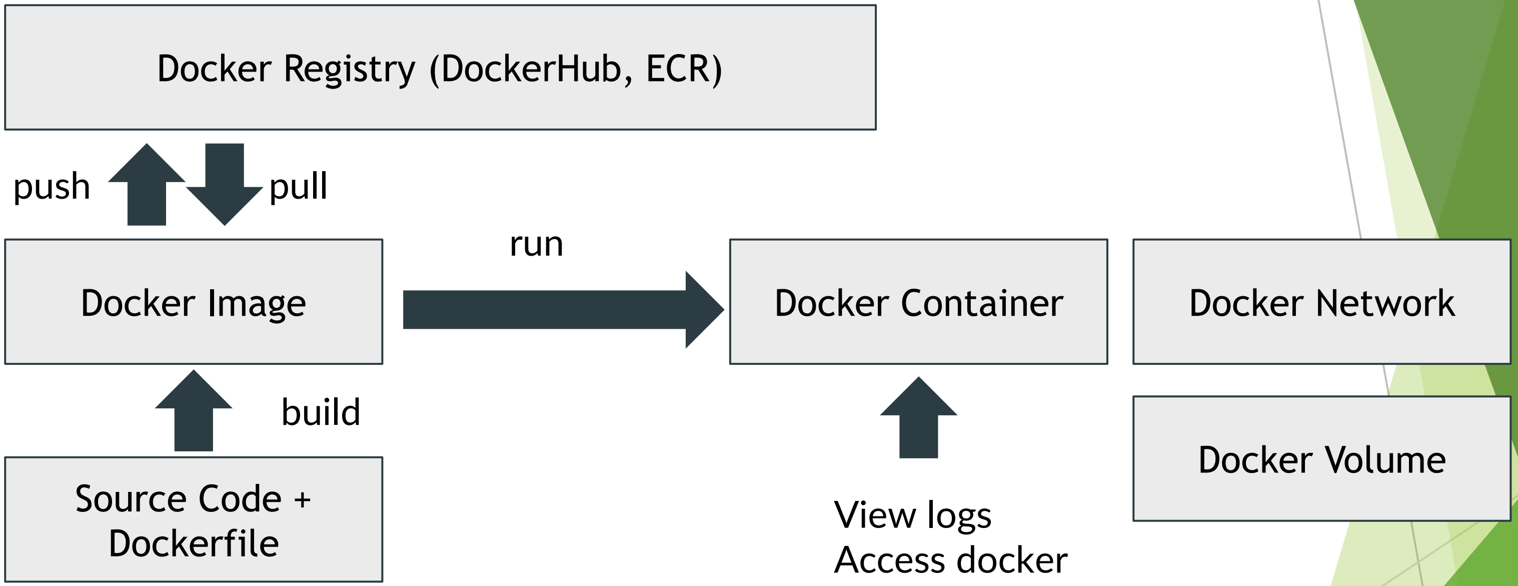
App 1 | App 1 | App 1

Bins / Libs | Bins / Libs | Bins / Libs

Docker Engine

Host OS (Linux)

What is the important components for Docker ?



Docker Registry (DockerHub, ECR)

push

pull

Docker Image

run

Docker Container

Docker Network

build

Docker Volume

Source Code + Dockerfile

View logs
Access docker

How to use Docker command line ?

- **Docker Image**

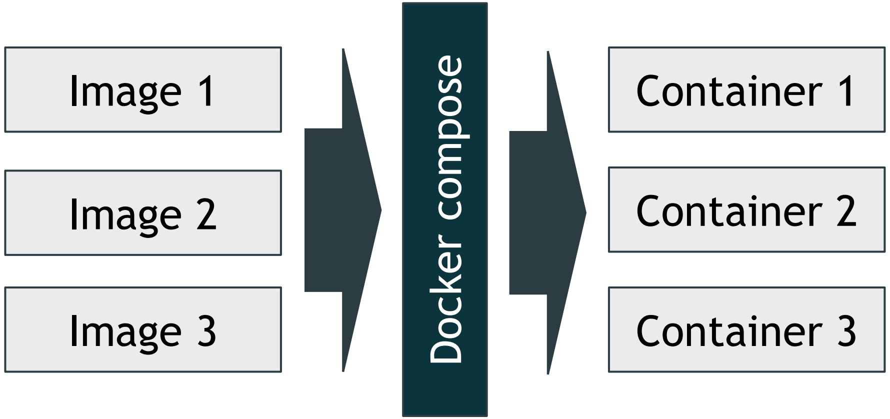
```
> docker pull <image_name:version>  # Pull docker image from docker registry
> docker build .                            # Build docker image from source code
> docker images                             # View docker image list
> docker rmi <image_name>                # Remove docker image
```

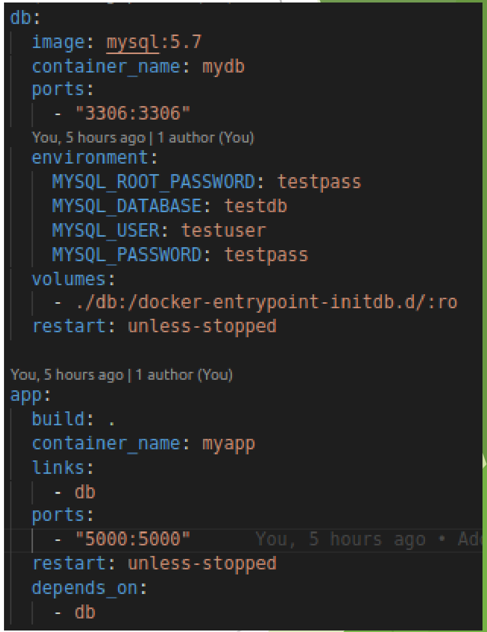
- **Docker Container**

```
> docker run -d -p <local_port>:<container_port> <image>   # Run docker container
> docker ps                                # View running container
```

https://docs.docker.com/engine/reference/commandline/docker/

How to run and manage Docker container simultaneously ?



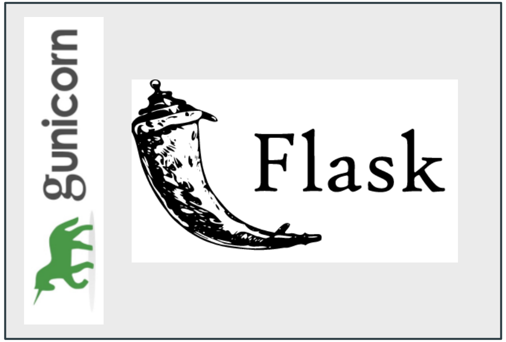


```
db:
  image: mysql:5.7
  container_name: mydb
  ports:
    - "3306:3306"
  You, 5 hours ago | 1 author (You)
  environment:
    MYSQL_ROOT_PASSWORD: testpass
    MYSQL_DATABASE: testdb
    MYSQL_USER: testuser
    MYSQL_PASSWORD: testpass
  volumes:
    - ./db:/docker-entrypoint-initdb.d/:ro
  restart: unless-stopped

You, 5 hours ago | 1 author (You)
app:
  build: .
  container_name: myapp
  links:
    - db
  ports:
    - "5000:5000"       You, 5 hours ago • Ad
  restart: unless-stopped
  depends_on:
    - db
```

docker-compose.yml

Back to our previous production architecture, let’s try to build docker container for all components that we have together

Docker Container 1 :
Webserver



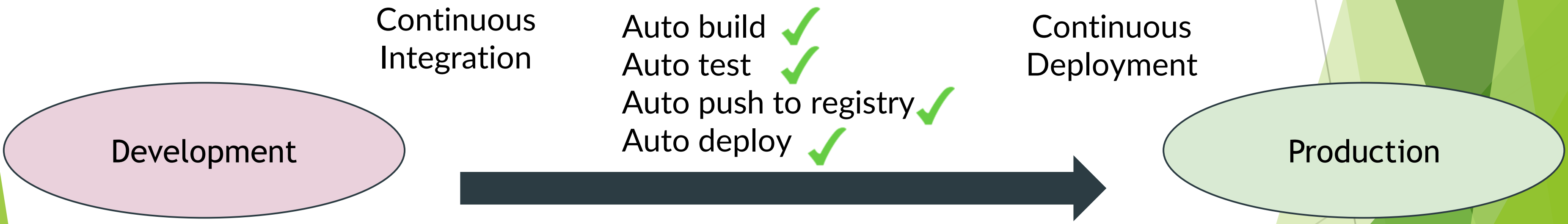
Docker Container 2 :
Our application

Docker Container 3 :
Database

Continuous Integration
Continuous Deployment (CICD)

Why we need CICD ?

AUTOMATION to make our life more **PRODUCTIVE**



Development

Continuous Integration

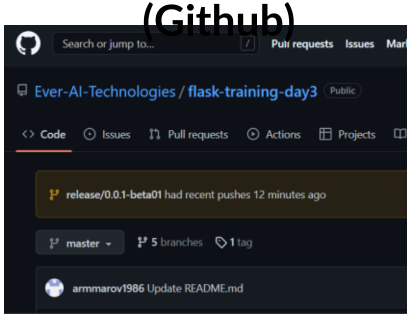
Auto build ✓
Auto test ✓
Auto push to registry✓
Auto deploy ✓

Continuous Deployment

Production

How we can automate CI part ?

Source Control (Github)



Ever-AI-Technologies / flask-training-day3 Public

release/0.0.1-beta01 had recent pushes 12 minutes ago

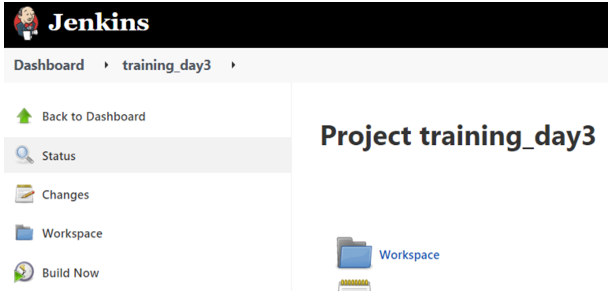
5 branches 1 tag

armmarov1986 Update README.md

Receive event (push)

Webhook

CICD Tool (Jenkins)



Project training_day3

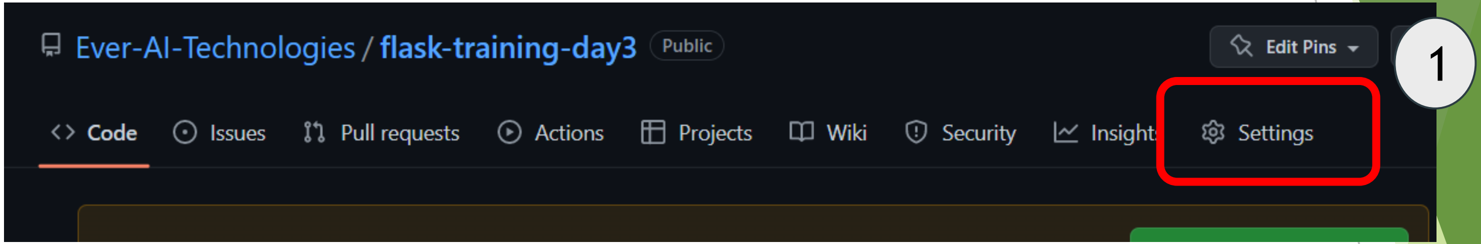
Clone → Build → Test → Push

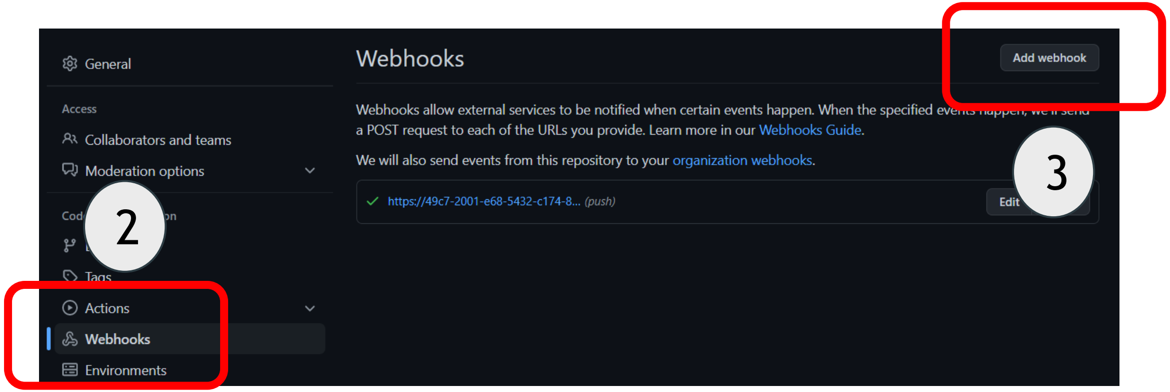
Push

Docker Registry (ECR)

Deploy to ECS/VPS

How to set Github for webhook ?



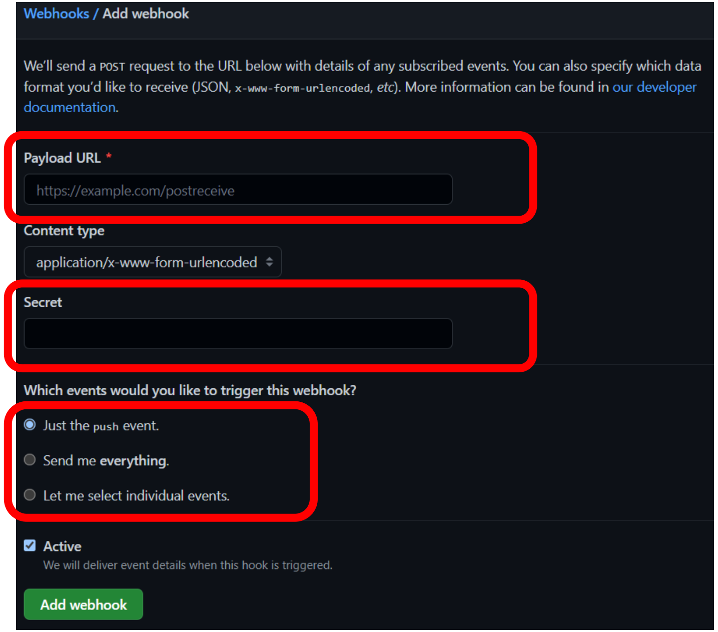


Webhooks allow external services to be notified when certain events happen. When the specified events happen, we'll send a POST request to each of the URLs you provide. Learn more in our Webhooks Guide.

We will also send events from this repository to your organization webhooks.

✓ https://49c7-2001-e68-5432-c174-8... (push)

How to set Github for webhook ?



Webhooks / Add webhook

We'll send a POST request to the URL below with details of any subscribed events. You can also specify which data format you'd like to receive (JSON, `x-www-form-urlencoded`, *etc*). More information can be found in our developer documentation.

Payload URL *

https://example.com/postreceive

Content type

application/x-www-form-urlencoded

Secret

Which events would you like to trigger this webhook?

● Just the `push` event.

○ Send me **everything**.

○ Let me select individual events.

☑ Active

We will deliver event details when this hook is triggered.

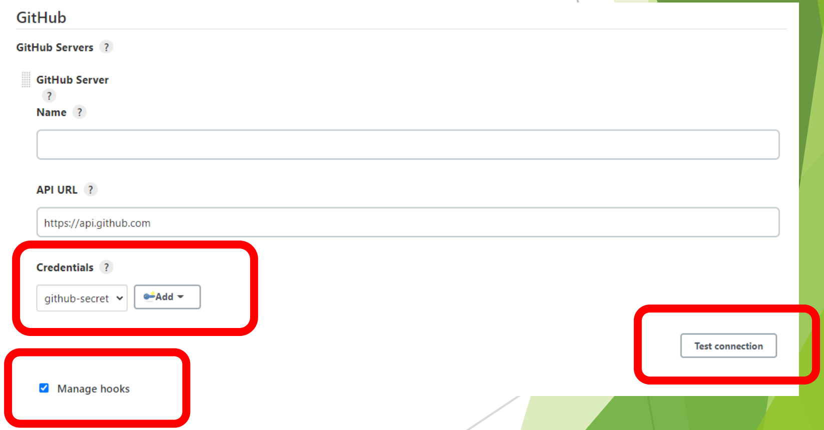
Add jenkins URL
https://jenkins_url:jenkins_port/github-webhook/

Add secret key for more secure

Select webhook trigger event

How to set Github Server in Jenkins for webhook ?

1. Open Manage Jenkins
2. Open Configure System
3. Add credential for Github Server
4. Click Test connection button
5. Tick Manage hooks



GitHub

GitHub Servers ?

GitHub Server ?

Name ?

API URL ?

https://api.github.com

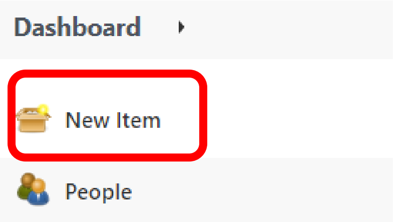
Credentials ?

github-secret

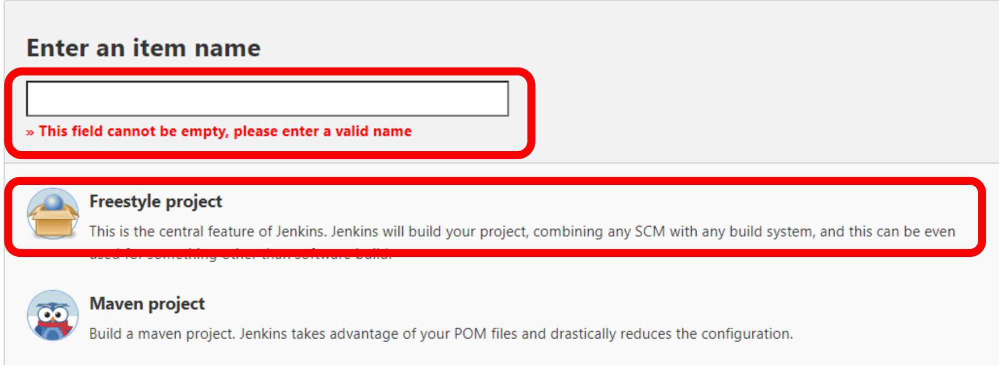
☑ Manage hooks

How to set up Jenkins project ?

At the dashboard, click New Item



Enter the project name, select “Freestyle project”



Enter an item name

» This field cannot be empty, please enter a valid name

Freestyle project

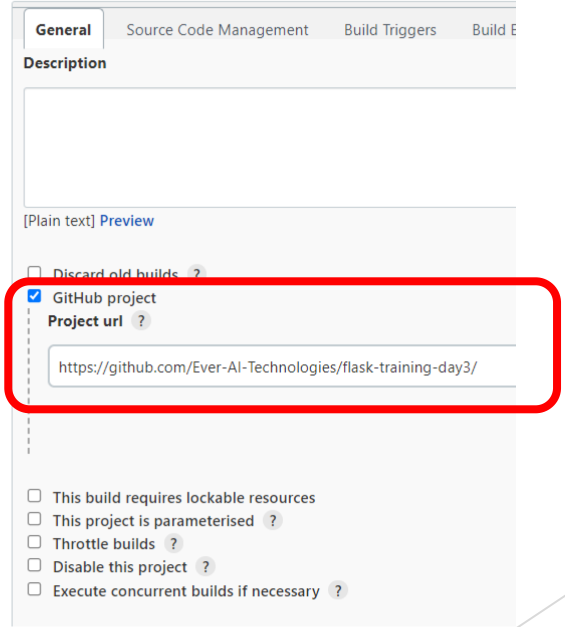
This is the central feature of Jenkins. Jenkins will build your project, combining any SCM with any build system, and this can be even used for something other than software build.

Maven project

Build a maven project. Jenkins takes advantage of your POM files and drastically reduces the configuration.

How to set up Jenkins project ?

At the General Tab, tick the Github project, fill in the github url



Description

[Plain text] Preview

□ Discard old builds ?

☑ GitHub project

Project url ?

https://github.com/Ever-AI-Technologies/flask-training-day3/

□ This build requires lockable resources

□ This project is parameterised ?

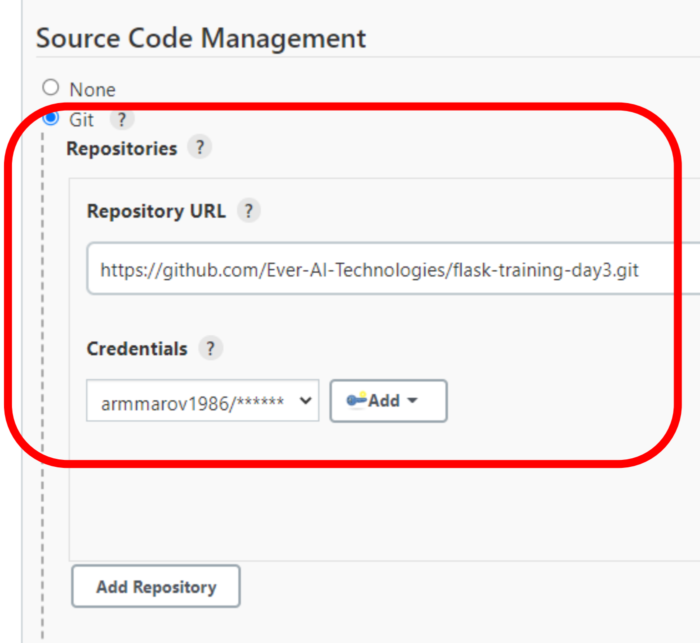
□ Throttle builds ?

□ Disable this project ?

□ Execute concurrent builds if necessary ?

How to set up Jenkins project ?

Scroll down to Source Code Management, tick Git and fill in Repository URL and credential



Source Code Management

○ None

● Git ?

Repositories ?

Repository URL ?

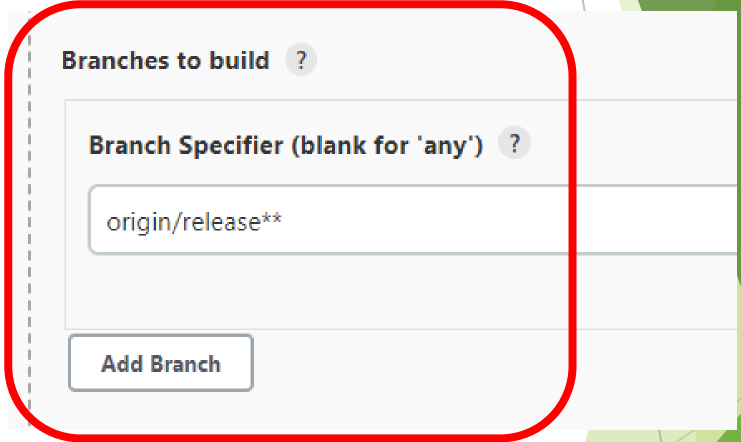
https://github.com/Ever-AI-Technologies/flask-training-day3.git

Credentials ?

armmarov1986/******

How to set up Jenkins project ?

Set branch to only origin/release** to allow only release branch, otherwise leave it blank to allow all branches



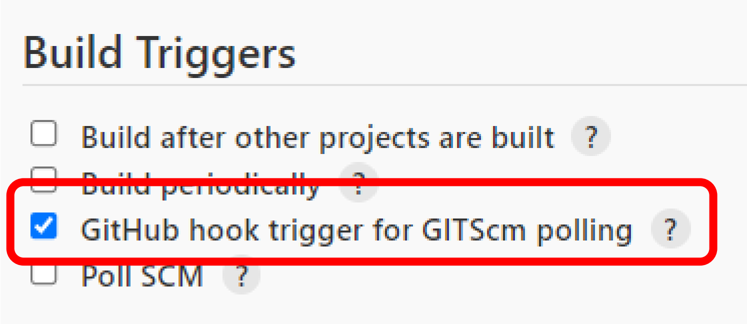
Branches to build

Branch Specifier (blank for 'any'): origin/release**

How to set up Jenkins project ?

Scroll to Build Triggers, tick the selection for Github hook trigger for GITScm polling

This will activate the webhook from Github



Build Triggers

□ Build after other projects are built ?

□ Build periodically ?

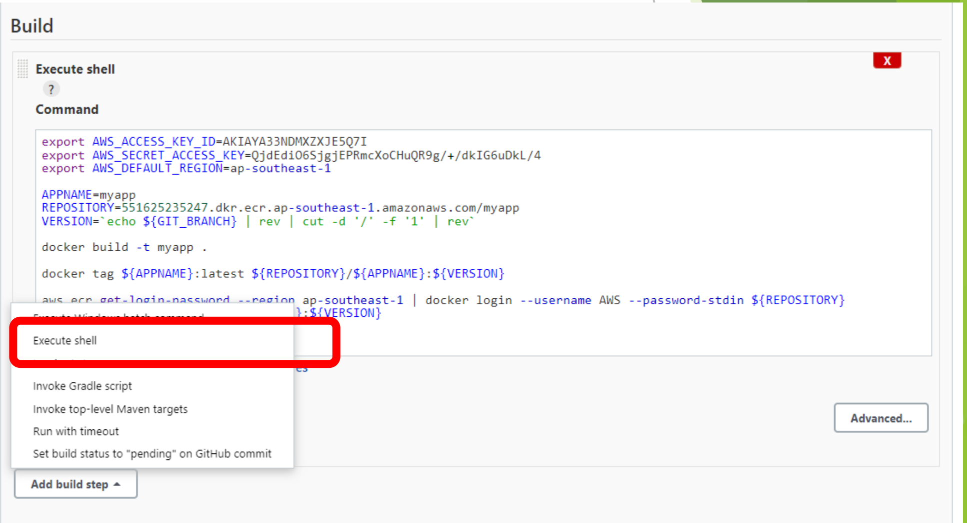
☑ GitHub hook trigger for GITScm polling ?

□ Poll SCM ?

How to set up Jenkins project ?

Scroll down to the Build section, click Add build step, select Execute shell.

Command window will appear



```
export AWS_ACCESS_KEY_ID=AKIAYA33NDMXZXJE5Q7I
export AWS_SECRET_ACCESS_KEY=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
export AWS_DEFAULT_REGION=ap-southeast-1

APPNAME=myapp
REPOSITORY=551625235247.dkr.ecr.ap-southeast-1.amazonaws.com/myapp
VERSION=`echo ${GIT_BRANCH} | rev | cut -d '/' -f '1' | rev`

docker build -t myapp .

docker tag ${APPNAME}:latest ${REPOSITORY}/${APPNAME}:${VERSION}

aws ecr get-login-password --region ap-southeast-1 | docker login --username AWS --password-stdin ${REPOSITORY}
                                      }:${VERSION}
```

Execute Windows batch command
Execute shell
Invoke Gradle script
Invoke top-level Maven targets
Run with timeout
Set build status to "pending" on GitHub commit
Add build step
Advanced...

How to set up Jenkins project ?

Copy paste this command into the Command window and Save

```
export AWS_ACCESS_KEY_ID=<YOUR_ACCESS_KEY>
export AWS_SECRET_ACCESS_KEY=<YOUR_SECRET_KEY>
export AWS_DEFAULT_REGION=ap-southeast-1

APPNAME=myapp
REPOSITORY=<YOUR_DOCKER_REPOSITORY>
VERSION=`echo ${GIT_BRANCH} | rev | cut -d '/' -f '1' | rev`

docker build -t ${APPNAME} .

docker tag ${APPNAME}:latest ${REPOSITORY}/${APPNAME}:${VERSION}

aws ecr get-login-password --region ap-southeast-1 | docker login --username AWS --password-stdin ${REPOSITORY}
docker push ${REPOSITORY}/${APPNAME}:${VERSION}
```

So now how can we execute all this ?

Hints

1. Branch must be a release branch

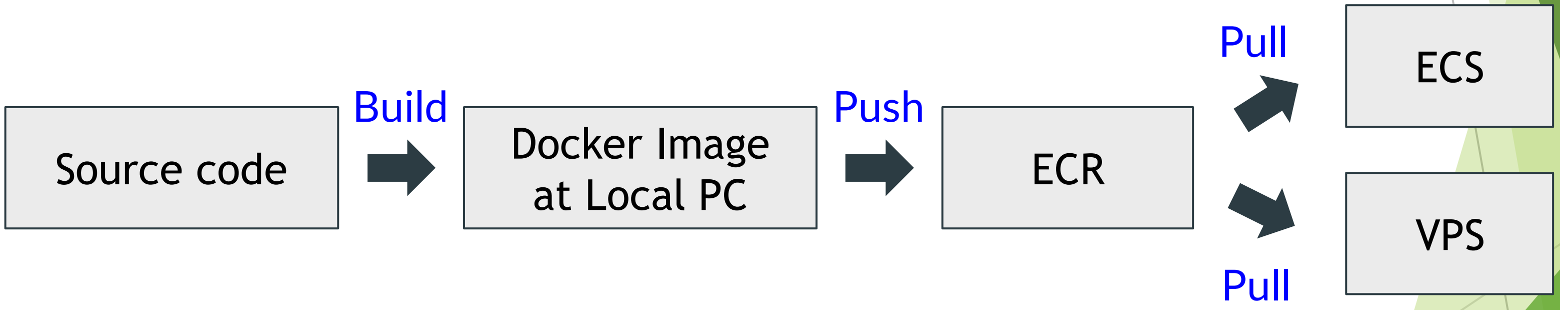
1. Trigger event is set to “push”

Is it working well now ? Or are we still missing something ?

Docker Registry

What is ECR ?

- AWS managed container image registry service that is secure, scalable and reliable.
- Support private / public repositories



Source code

Build

Docker Image at Local PC

Push

ECR

Pull

ECS

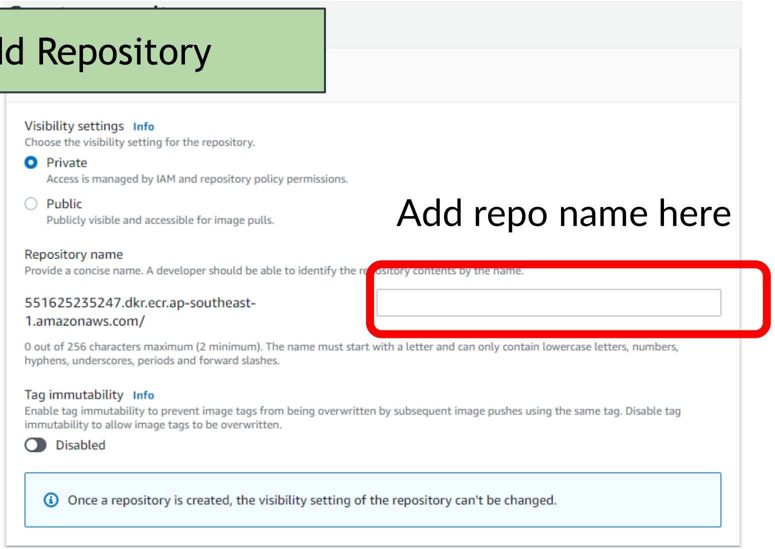
Pull

VPS

How to create registry repository ?

Navigate to https://ap-southeast-1.console.aws.amazon.com/ecr/get-started?region=ap-southeast-1

Add Repository



Visibility settings Info

Choose the visibility setting for the repository.

● Private
Access is managed by IAM and repository policy permissions.

○ Public
Publicly visible and accessible for image pulls.

Repository name

Provide a concise name. A developer should be able to identify the repository contents by the name.

551625235247.dkr.ecr.ap-southeast-1.amazonaws.com/

0 out of 256 characters maximum (2 minimum). The name must start with a letter and can only contain lowercase letters, numbers, hyphens, underscores, periods and forward slashes.

Tag immutability Info

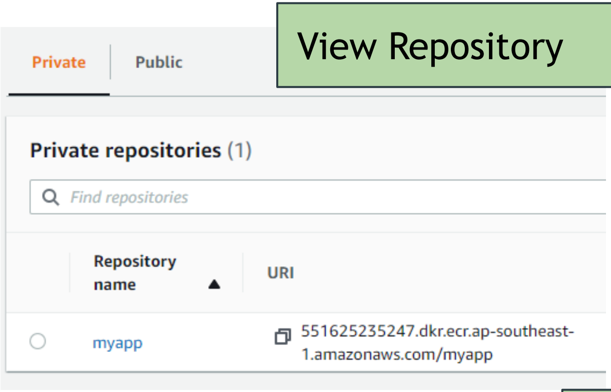
Enable tag immutability to prevent image tags from being overwritten by subsequent image pushes using the same tag. Disable tag immutability to allow image tags to be overwritten.

Disabled

Once a repository is created, the visibility setting of the repository can't be changed.

Add repo name here

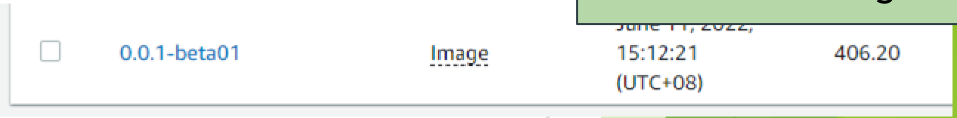
View Repository



Private repositories (1)

	Repository name ▲	URI
○	myapp	551625235247.dkr.ecr.ap-southeast-1.amazonaws.com/myapp

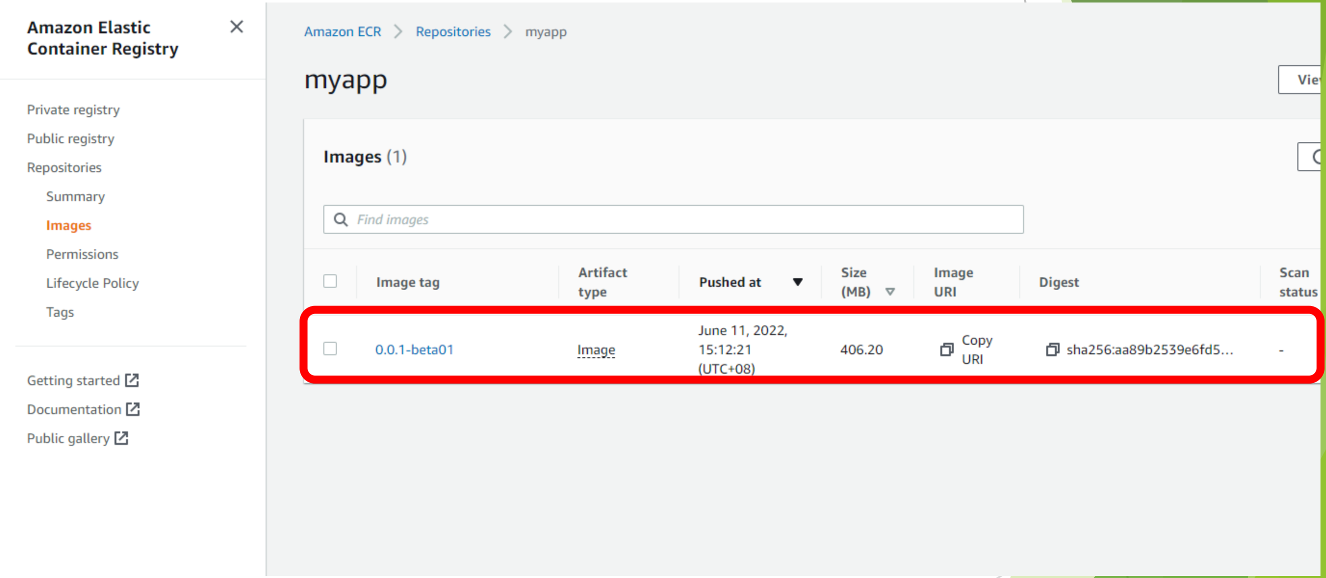
View Docker Image



□	0.0.1-beta01	Image	June 11, 2022, 15:12:21 (UTC+08)	406.20

Now, let’s try again and see if it works !!

Our target is to see this image in our list



myapp

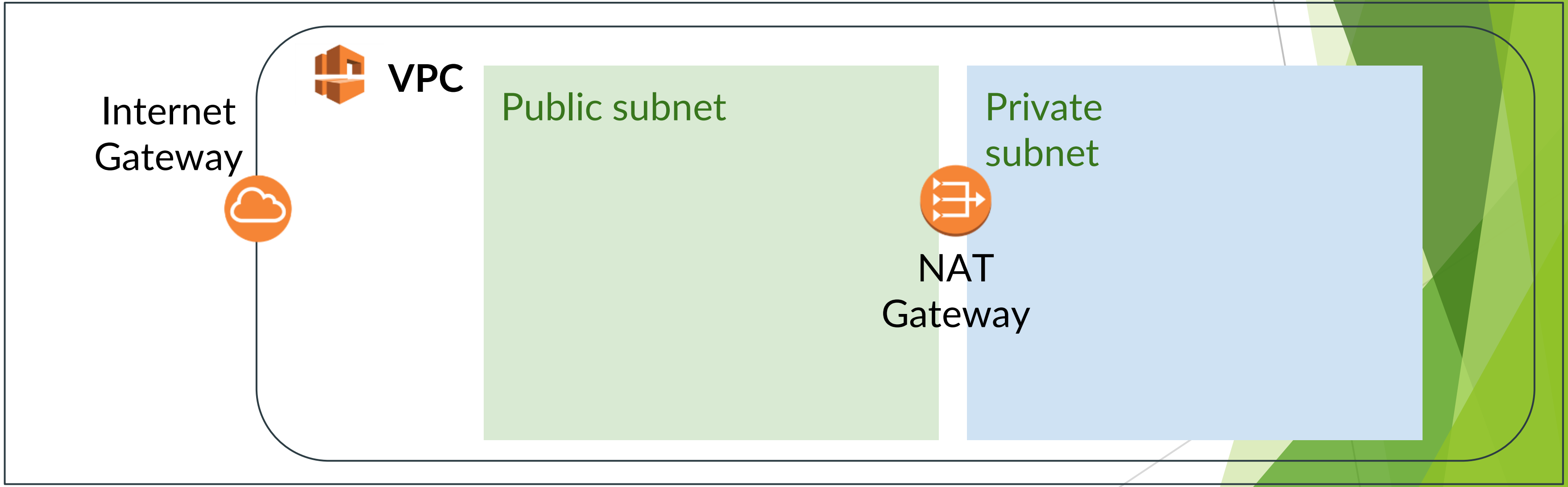
Images (1)

Find images

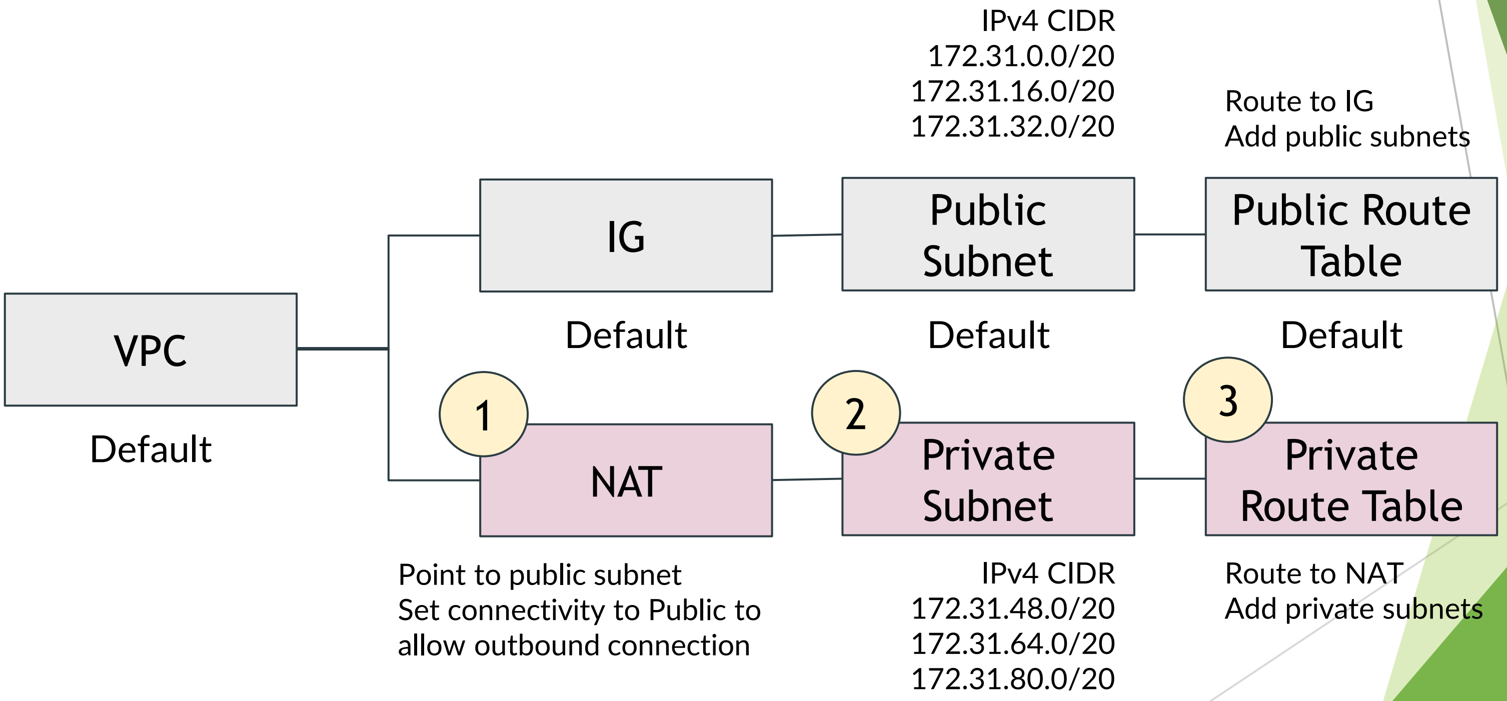
□	Image tag	Artifact type	Pushed at	Size (MB)	Image URI	Digest	Scan status
□	0.0.1-beta01	Image	June 11, 2022, 15:12:21 (UTC+08)	406.20	Copy URI	sha256:aa89b2539e6fd5...	-

Let’s properly set up our AWS for Production

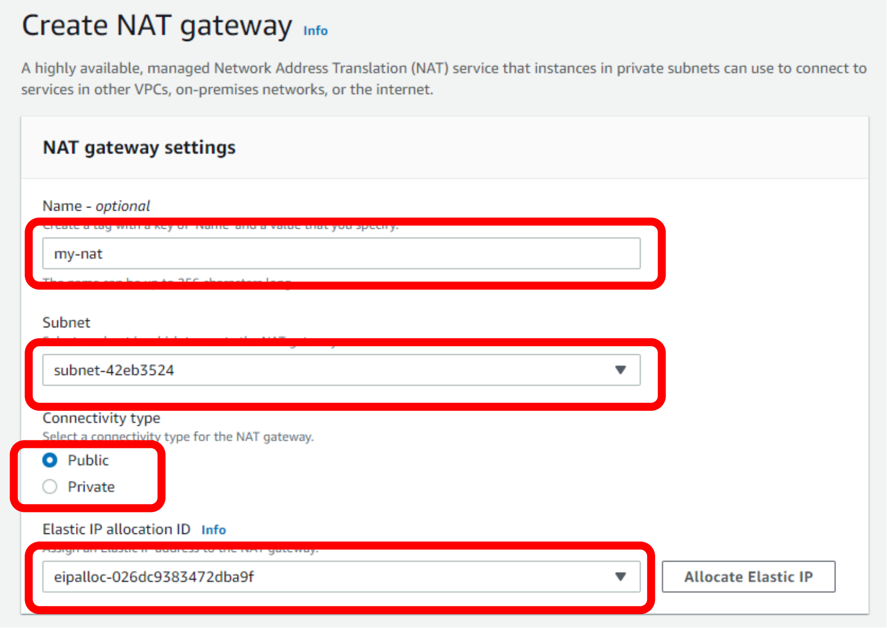
Before we jump in, let’s set up the network first



Let’s see step by step



1. Create NAT Gateway



Create NAT gateway Info

A highly available, managed Network Address Translation (NAT) service that instances in private subnets can use to connect to services in other VPCs, on-premises networks, or the internet.

NAT gateway settings

Name - *optional*
Create a tag with a key of 'Name' and a value that you specify.
my-nat
The name can be up to 256 characters long.

Subnet
Select a subnet in which to create the NAT gateway.
subnet-42eb3524 ▼

Connectivity type
Select a connectivity type for the NAT gateway.
● Public
○ Private

Elastic IP allocation ID Info
Assign an Elastic IP address to the NAT gateway.
eipalloc-026dc9383472dba9f ▼

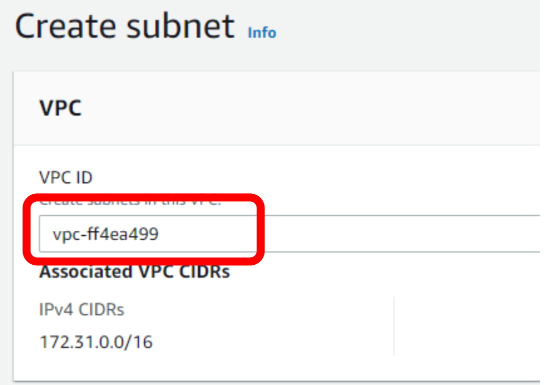
Fill in the name for reference

Attach to any public

Set to public to allow outbound

Allocate elastic IP

2. Create Private Subnet



Create subnet Info

VPC

VPC ID

Create subnets in this VPC.

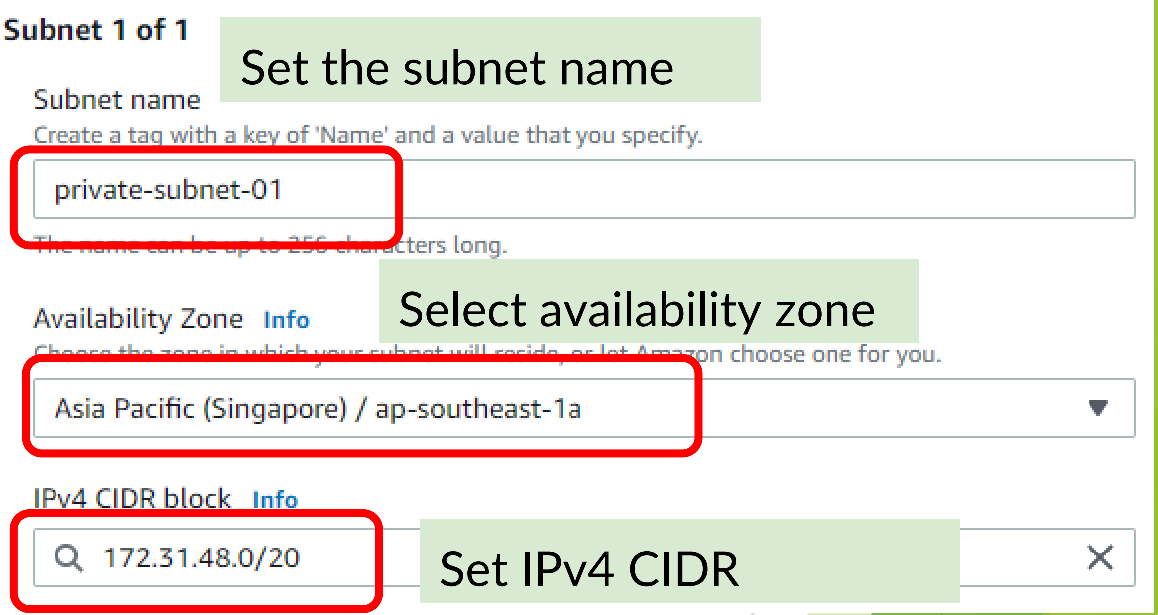
vpc-ff4ea499

Associated VPC CIDRs

IPv4 CIDRs

172.31.0.0/16

Select current VPC



Subnet 1 of 1

Set the subnet name

Subnet name

Create a tag with a key of 'Name' and a value that you specify.

private-subnet-01

The name can be up to 256 characters long.

Availability Zone Info

Select availability zone

Choose the zone in which your subnet will reside, or let Amazon choose one for you.

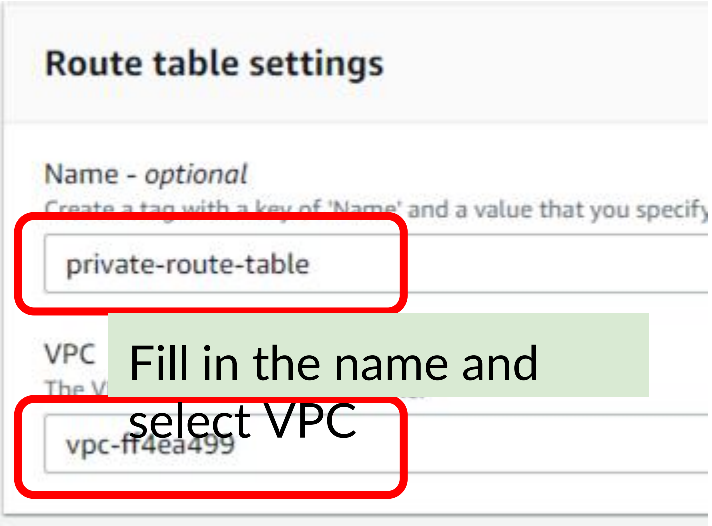
Asia Pacific (Singapore) / ap-southeast-1a

IPv4 CIDR block Info

172.31.48.0/20

Set IPv4 CIDR

3. Create Private Route Table



Route table settings

Name - *optional*

Create a tag with a key of 'Name' and a value that you specify

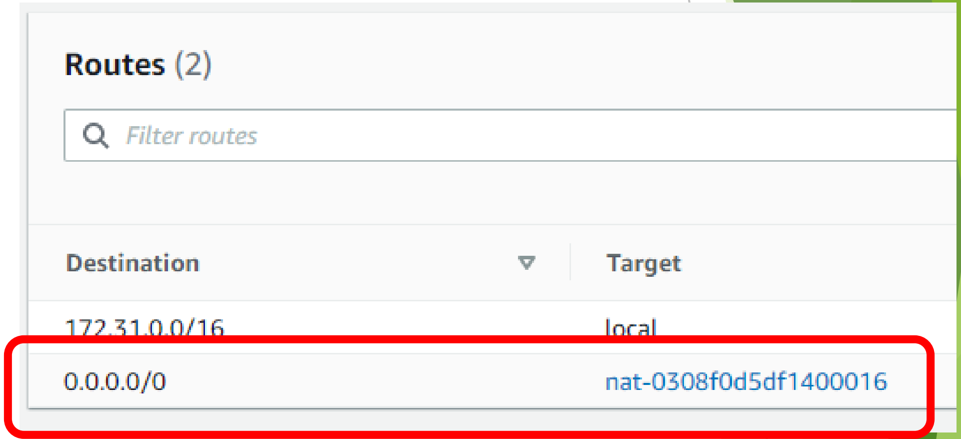
private-route-table

VPC

The V

vpc-ff4ea499

Fill in the name and select VPC



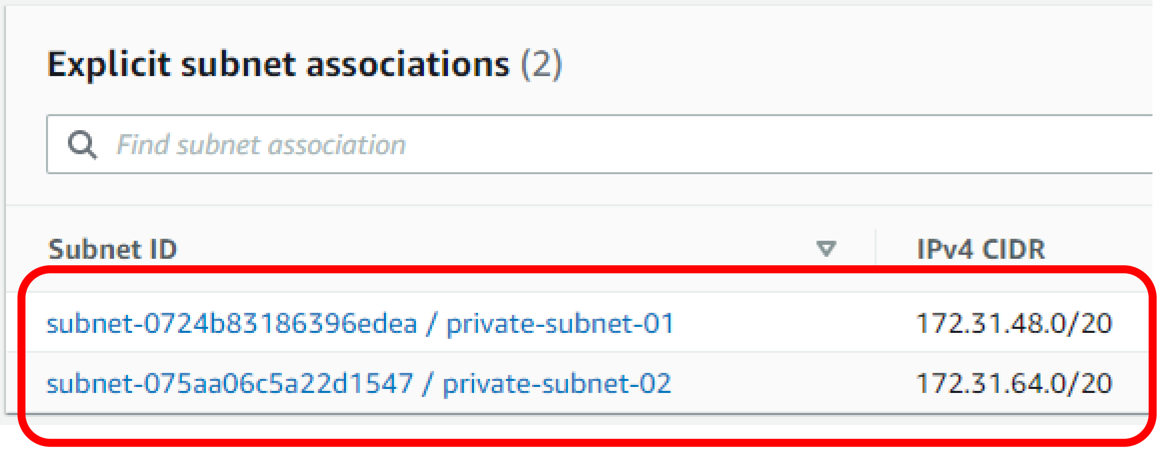
Routes (2)

Filter routes

Destination	Target
172.31.0.0/16	local
0.0.0.0/0	nat-0308f0d5df1400016

Add NAT gateway to the table

3. Create Private Route Table



Explicit subnet associations (2)

Find subnet association

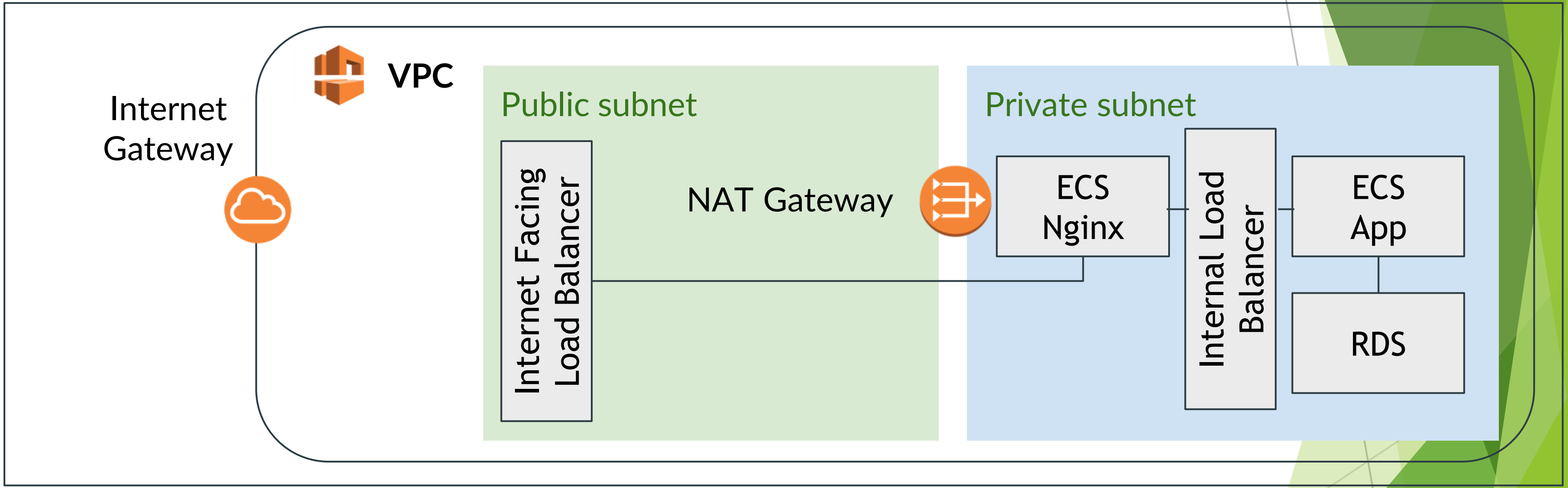
Subnet ID	IPv4 CIDR
subnet-0724b83186396edea / private-subnet-01	172.31.48.0/20
subnet-075aa06c5a22d1547 / private-subnet-02	172.31.64.0/20

Associate all private subnets to the route table

Now, let's deploy to production !!

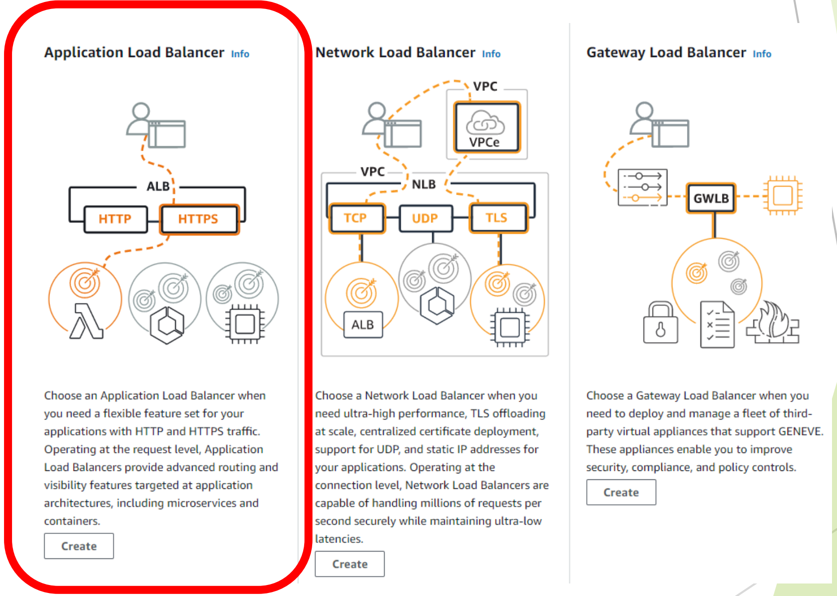
Auto Scaling with Higher Cost

Let’s design high level architecture first



Prior to set up the ECS, we need to bring both internet facing and internal Load Balancer up first. Let’s do this

For both cases, since we are dealing with HTTP/s protocol, let use application load balancer



Application Load Balancer Info

Choose an Application Load Balancer when you need a flexible feature set for your applications with HTTP and HTTPS traffic. Operating at the request level, Application Load Balancers provide advanced routing and visibility features targeted at application architectures, including microservices and containers.

Create

Network Load Balancer Info

Choose a Network Load Balancer when you need ultra-high performance, TLS offloading at scale, centralized certificate deployment, support for UDP, and static IP addresses for your applications. Operating at the connection level, Network Load Balancers are capable of handling millions of requests per second securely while maintaining ultra-low latencies.

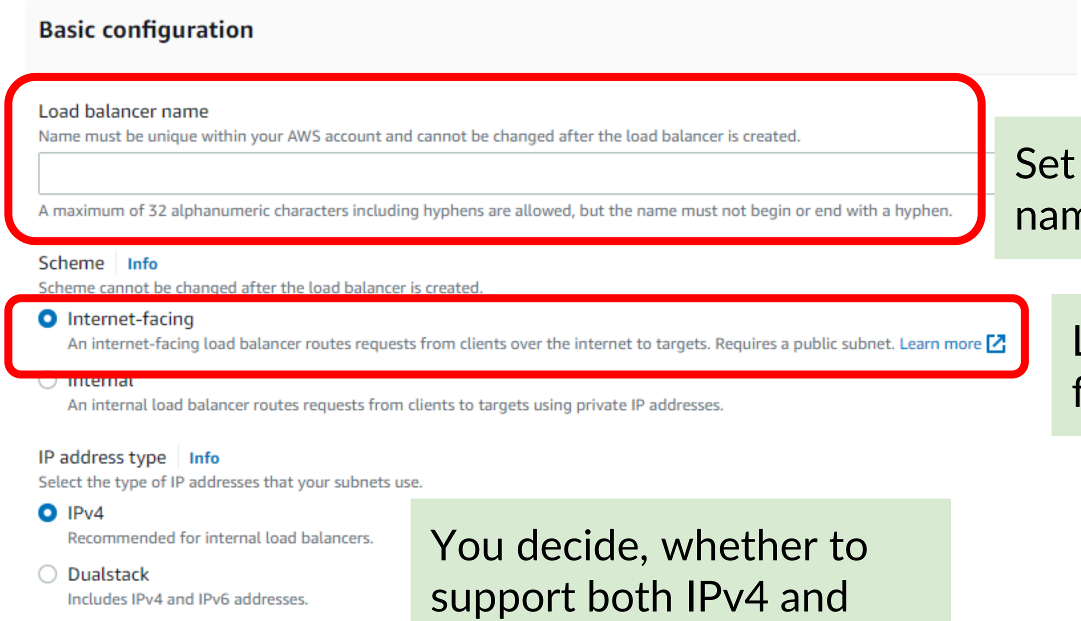
Create

Gateway Load Balancer Info

Choose a Gateway Load Balancer when you need to deploy and manage a fleet of third-party virtual appliances that support GENEVE. These appliances enable you to improve security, compliance, and policy controls.

Create

Configuration for the internet-facing load balancer



Basic configuration

Load balancer name

Name must be unique within your AWS account and cannot be changed after the load balancer is created.

A maximum of 32 alphanumeric characters including hyphens are allowed, but the name must not begin or end with a hyphen.

Scheme | Info

Scheme cannot be changed after the load balancer is created.

● Internet-facing

An internet-facing load balancer routes requests from clients over the internet to targets. Requires a public subnet. Learn more

○ Internal

An internal load balancer routes requests from clients to targets using private IP addresses.

IP address type | Info

Select the type of IP addresses that your subnets use.

● IPv4

Recommended for internal load balancers.

○ Dualstack

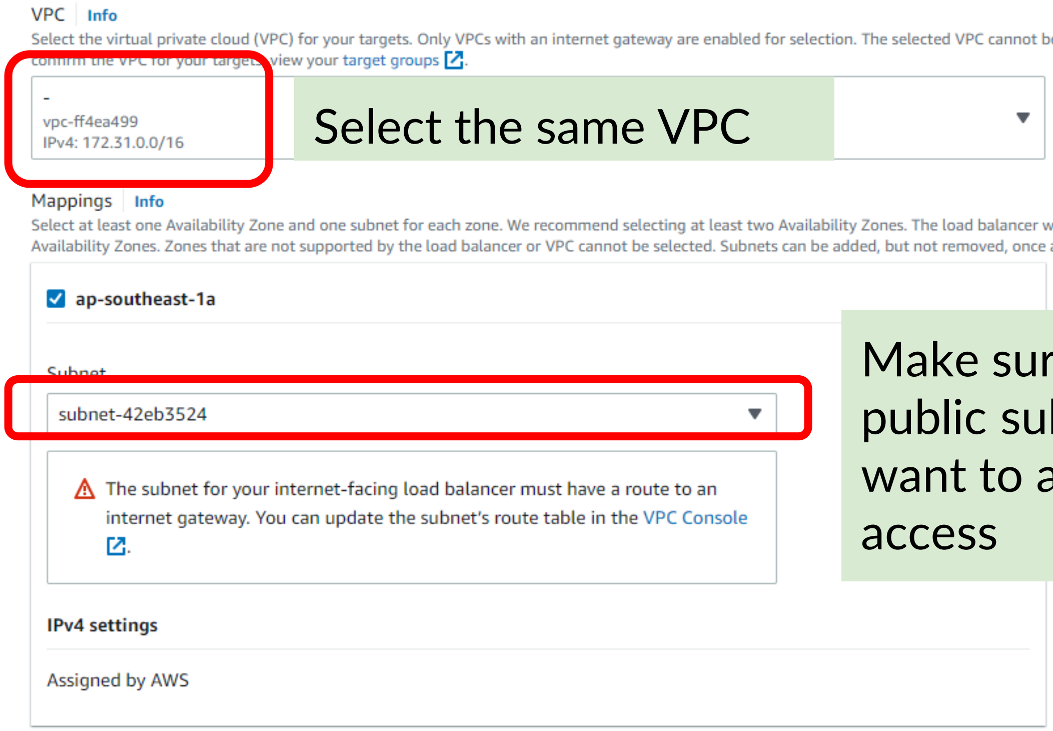
Includes IPv4 and IPv6 addresses.

Set the load balancer name

Let’s choose internet-facing for this one

You decide, whether to support both IPv4 and IPv6

Configuration for the internet-facing load balancer



VPC | Info

Select the virtual private cloud (VPC) for your targets. Only VPCs with an internet gateway are enabled for selection. The selected VPC cannot be confirm the VPC for your targets, view your target groups.

-
vpc-ff4ea499
IPv4: 172.31.0.0/16

Select the same VPC

Mappings | Info

Select at least one Availability Zone and one subnet for each zone. We recommend selecting at least two Availability Zones. The load balancer wil Availability Zones. Zones that are not supported by the load balancer or VPC cannot be selected. Subnets can be added, but not removed, once a

☑ ap-southeast-1a

Subnet

subnet-42eb3524

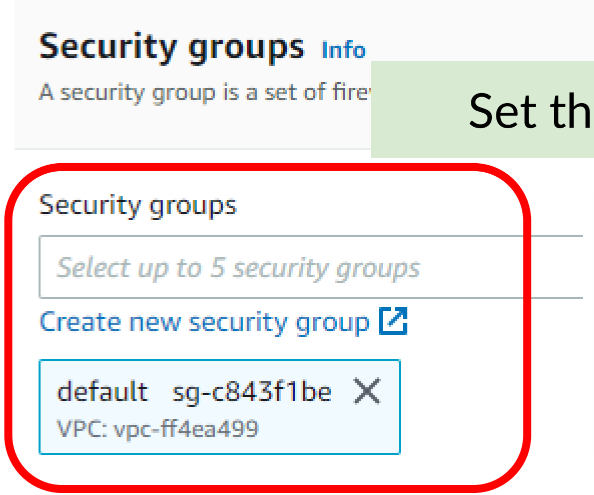
⚠ The subnet for your internet-facing load balancer must have a route to an internet gateway. You can update the subnet's route table in the VPC Console.

IPv4 settings

Assigned by AWS

Make sure to select the public subnet since we want to allow the external access

Configuration for the internet-facing load balancer



Security groups Info

A security group is a set of fire

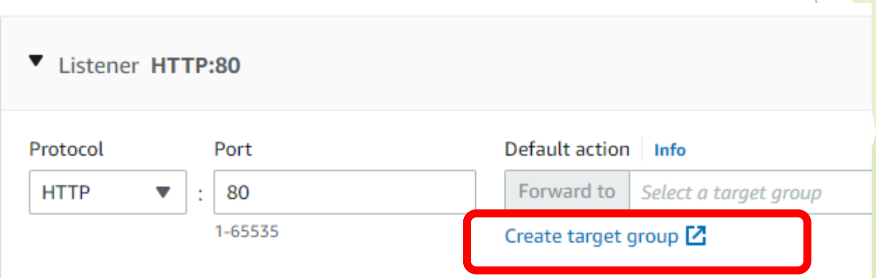
Security groups

Select up to 5 security groups

Create new security group

default sg-c843f1be ✕
VPC: vpc-ff4ea499

Set the right security groups

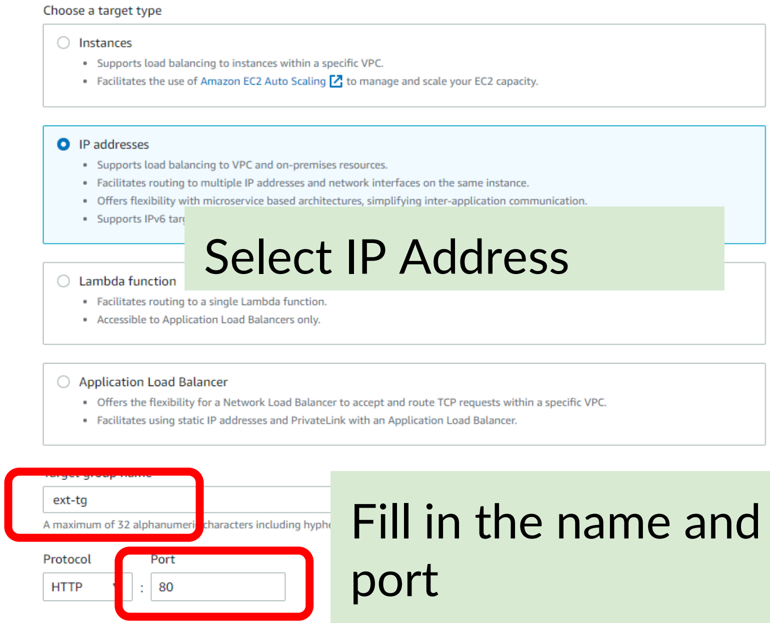


▼ Listener **HTTP:80**

Protocol	Port	Default action Info
HTTP ▼	80	Forward to Select a target group
	1-65535	Create target group

Click create target group

Create target group



Choose a target type

○ Instances
- Supports load balancing to instances within a specific VPC.
- Facilitates the use of Amazon EC2 Auto Scaling to manage and scale your EC2 capacity.

● IP addresses
- Supports load balancing to VPC and on-premises resources.
- Facilitates routing to multiple IP addresses and network interfaces on the same instance.
- Offers flexibility with microservice based architectures, simplifying inter-application communication.
- Supports IPv6 targ

○ Lambda function
- Facilitates routing to a single Lambda function.
- Accessible to Application Load Balancers only.

○ Application Load Balancer
- Offers the flexibility for a Network Load Balancer to accept and route TCP requests within a specific VPC.
- Facilitates using static IP addresses and PrivateLink with an Application Load Balancer.

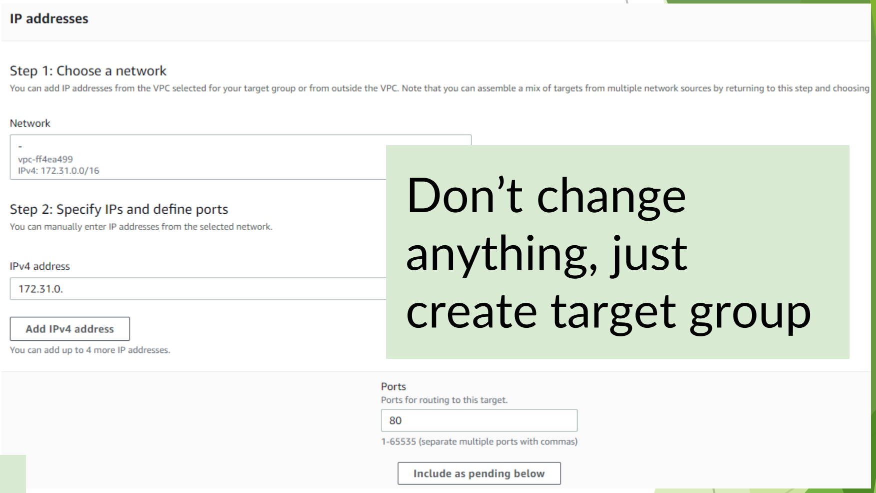
Target group name: ext-tg

A maximum of 32 alphanumeri characters including hyphe

Protocol: HTTP : Port: 80

Select IP Address

Fill in the name and port



IP addresses

Step 1: Choose a network

You can add IP addresses from the VPC selected for your target group or from outside the VPC. Note that you can assemble a mix of targets from multiple network sources by returning to this step and choosing

Network

-
vpc-ff4ea499
IPv4: 172.31.0.0/16

Step 2: Specify IPs and define ports

You can manually enter IP addresses from the selected network.

IPv4 address: 172.31.0.

You can add up to 4 more IP addresses.

Ports

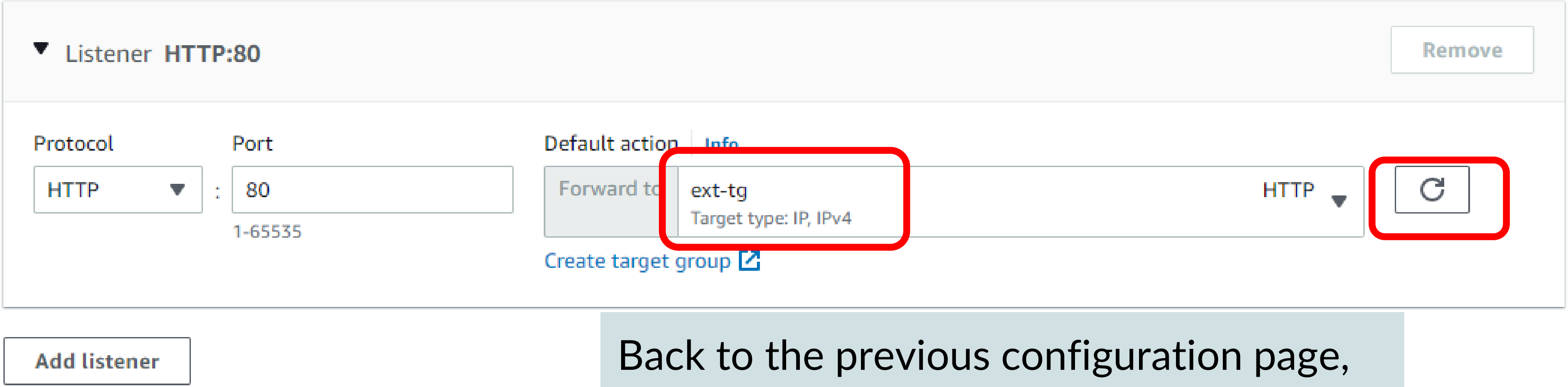
Ports for routing to this target.

80

1-65535 (separate multiple ports with commas)

Don’t change anything, just create target group

Configuration for the internet-facing load balancer



▼ Listener **HTTP:80**

Remove

Protocol: HTTP ▼

Port: 80

1-65535

Default action | Info

Forward to: ext-tg
Target type: IP, IPv4
HTTP ▼

Create target group

Add listener

Back to the previous configuration page, click refresh, select the right target group

Configuration for the internal load balancer

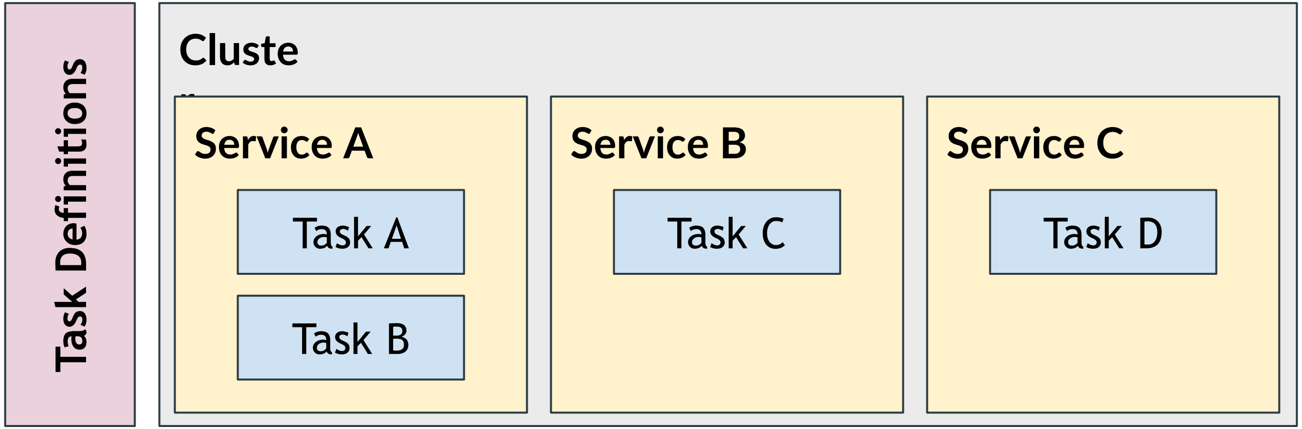
All are the same except:

1. Select “Internal” instead of “Internet-facing”
2. Select “Private Subnet” instead of “Public Subnet”

Can you try ?

What is Elastic Container Service (ECS) ?

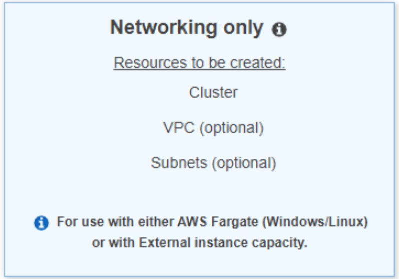
ECS is similar to Kubernetes whereby it intended to make an easy deployment, managing and scaling for the container such as Docker



Task

- Fargate
- EC2

Let’s create our first cluster



Networking only

Resources to be created:

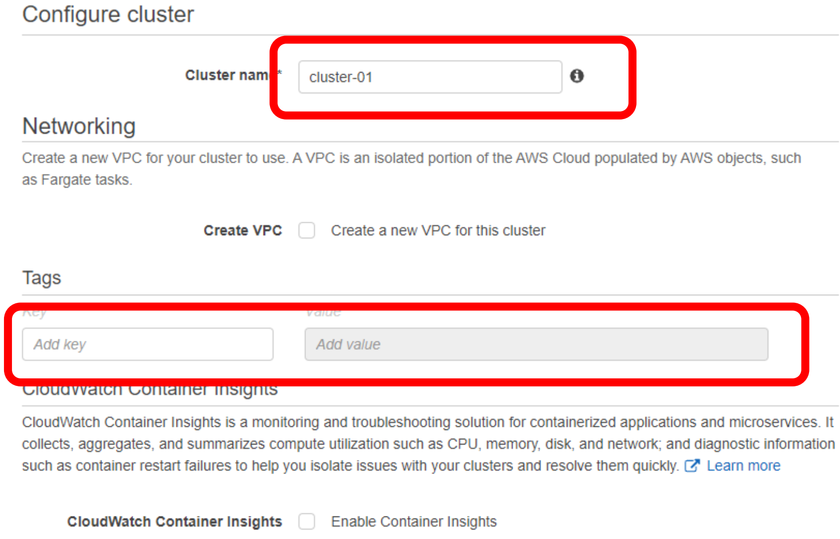
Cluster

VPC (optional)

Subnets (optional)

For use with either AWS Fargate (Windows/Linux) or with External instance capacity.

Select Networking only



Configure cluster

Cluster nam* cluster-01

Networking

Create a new VPC for your cluster to use. A VPC is an isolated portion of the AWS Cloud populated by AWS objects, such as Fargate tasks.

Create VPC □ Create a new VPC for this cluster

Tags

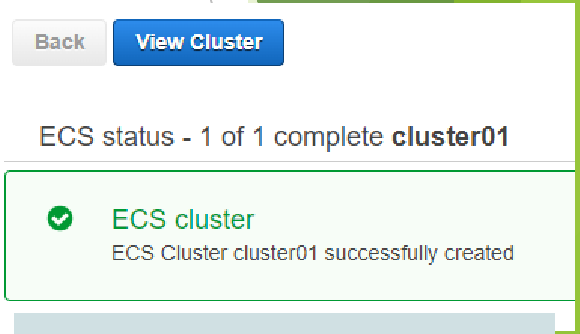
Key	Value
Add key	Add value

CloudWatch Container Insights

CloudWatch Container Insights is a monitoring and troubleshooting solution for containerized applications and microservices. It collects, aggregates, and summarizes compute utilization such as CPU, memory, disk, and network; and diagnostic information such as container restart failures to help you isolate issues with your clusters and resolve them quickly. Learn more

CloudWatch Container Insights □ Enable Container Insights

Fill in cluster name and Tags



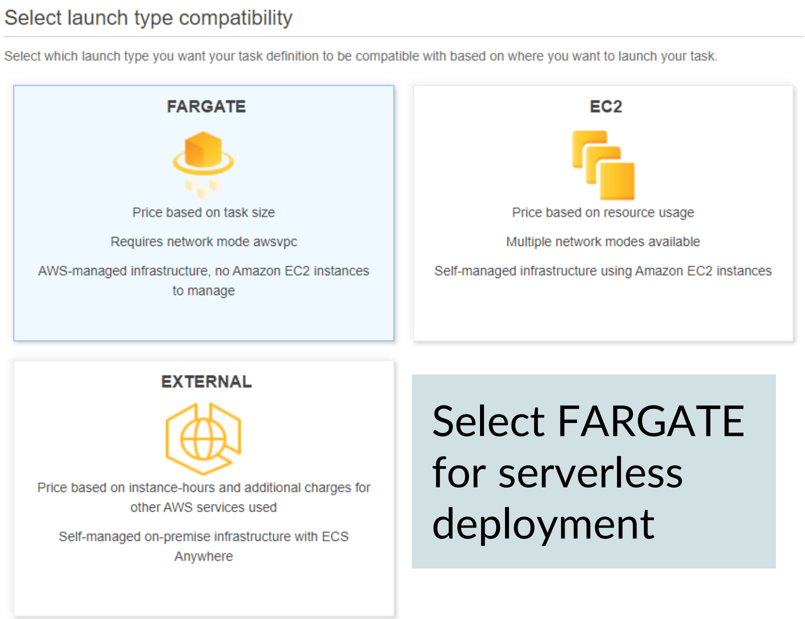
ECS status - 1 of 1 complete **cluster01**

ECS cluster

ECS Cluster cluster01 successfully created

Create the cluster

Before we can create any service, we need to have our Task Definition



Select launch type compatibility

Select which launch type you want your task definition to be compatible with based on where you want to launch your task.

FARGATE

Price based on task size

Requires network mode awsvpc

AWS-managed infrastructure, no Amazon EC2 instances to manage

EC2

Price based on resource usage

Multiple network modes available

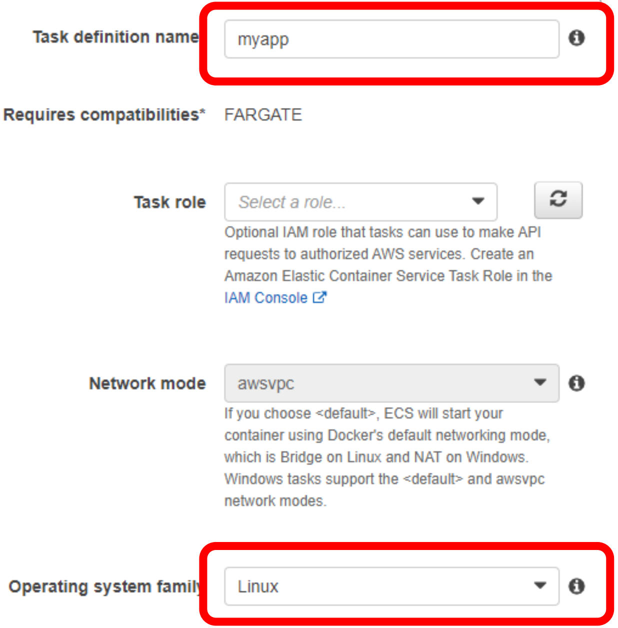
Self-managed infrastructure using Amazon EC2 instances

EXTERNAL

Price based on instance-hours and additional charges for other AWS services used

Self-managed on-premise infrastructure with ECS Anywhere

Select FARGATE for serverless deployment



Task definition name myapp

Requires compatibilities* FARGATE

Task role Select a role...

Optional IAM role that tasks can use to make API requests to authorized AWS services. Create an Amazon Elastic Container Service Task Role in the IAM Console

Network mode awsvpc

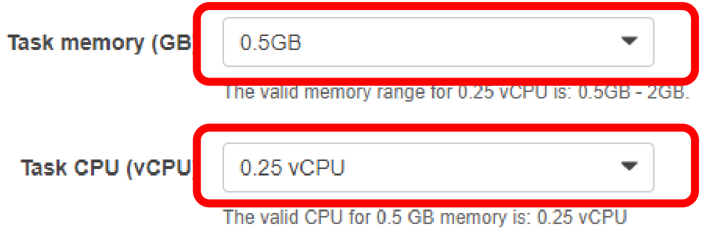
If you choose <default>, ECS will start your container using Docker's default networking mode, which is Bridge on Linux and NAT on Windows. Windows tasks support the <default> and awsvpc network modes.

Operating system family Linux

Enter name

Select Linux

Before we can create any service, we need to have our Task Definition



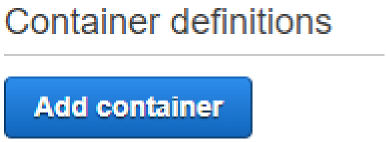
Task memory (GB 0.5GB

The valid memory range for 0.25 vCPU is: 0.5GB - 2GB.

Task CPU (vCPU 0.25 vCPU

The valid CPU for 0.5 GB memory is: 0.25 vCPU

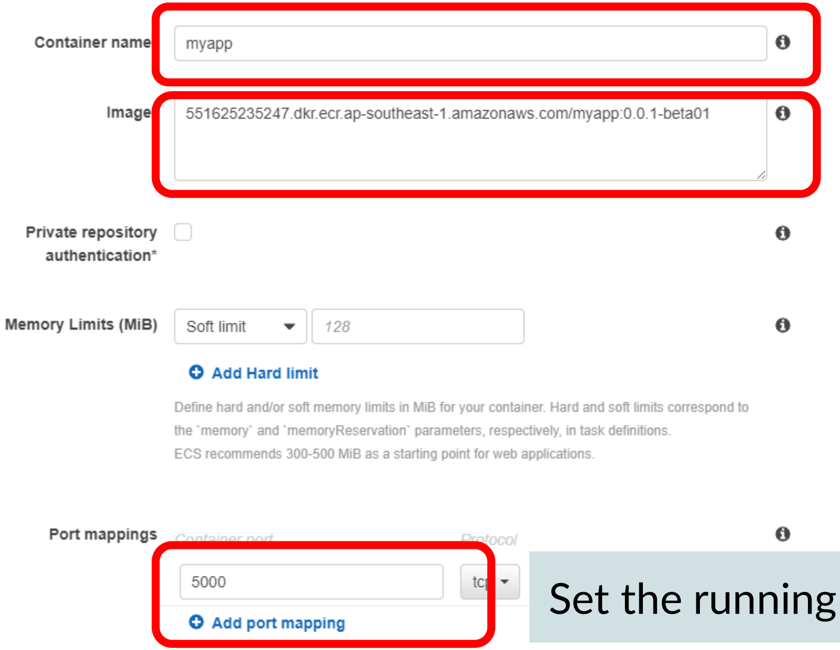
Set minimum resources first



Container definitions

Click add container to add docker container

Add docker container at the Task Definition



Container name: myapp

Image: 551625235247.dkr.ecr.ap-southeast-1.amazonaws.com/myapp:0.0.1-beta01

Private repository authentication* □

Memory Limits (MiB): Soft limit ▾ 128

⊕ Add Hard limit

Define hard and/or soft memory limits in MiB for your container. Hard and soft limits correspond to the `memory` and `memoryReservation` parameters, respectively, in task definitions.
ECS recommends 300-500 MiB as a starting point for web applications.

Port mappings: Container port 5000 Protocol tcp ▾

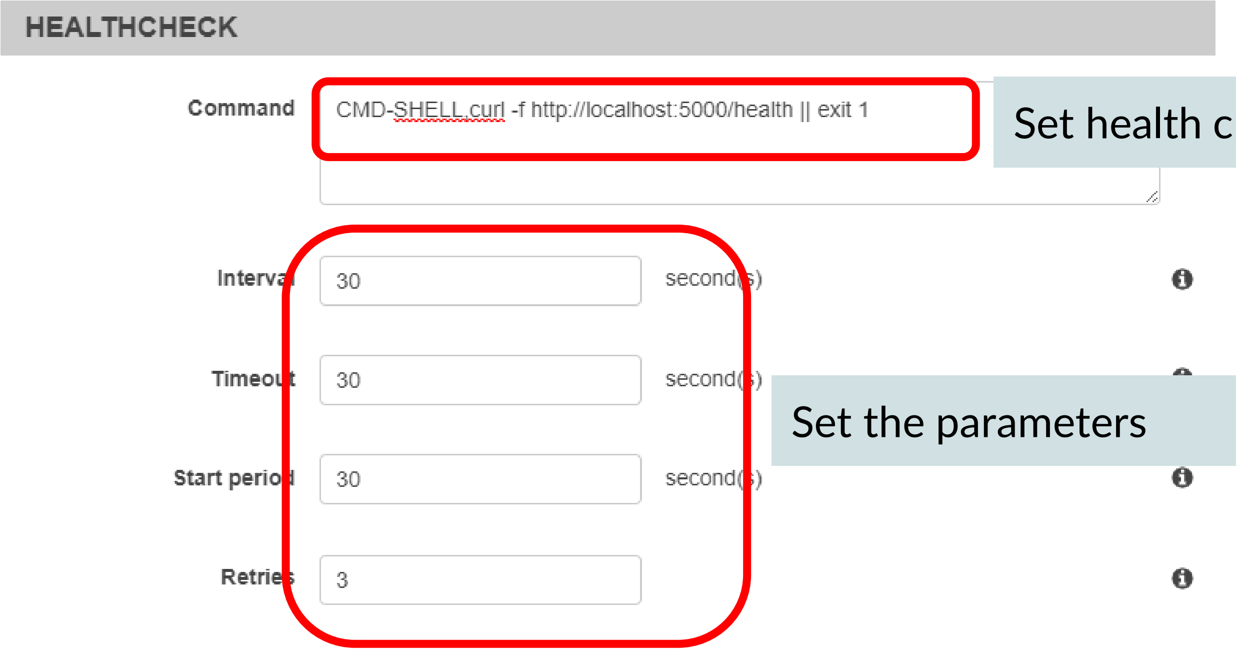
⊕ Add port mapping

Enter the container name

Copy the image url from ECR

Set the running Port

Add docker container at the Task Definition



HEALTHCHECK

Command: CMD-SHELL,curl -f http://localhost:5000/health || exit 1

Interval: 30 second(s)

Timeout: 30 second(s)

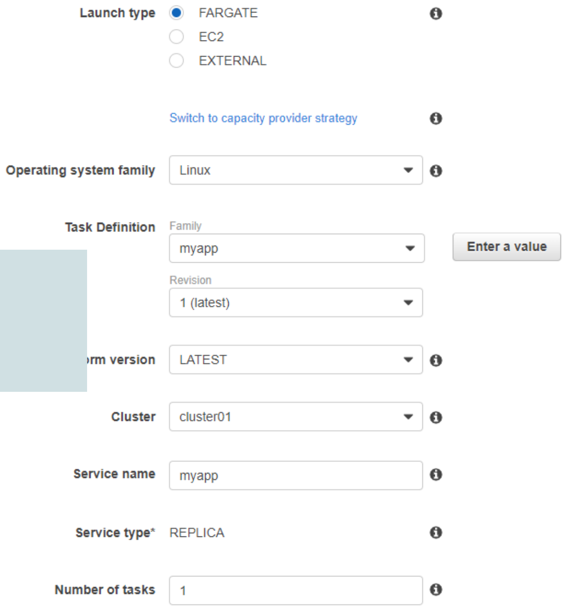
Start period: 30 second(s)

Retries: 3

Set health check command

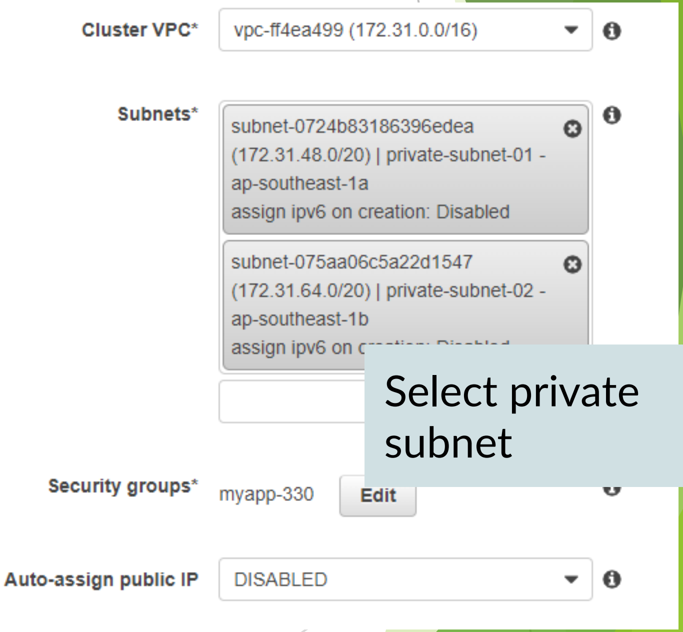
Set the parameters

Now let’s create our first service and execute the task



Launch type	● FARGATE ○ EC2 ○ EXTERNAL
	Switch to capacity provider strategy
Operating system family	Linux
Task Definition	Family: myapp Revision: 1 (latest)
rm version	LATEST
Cluster	cluster01
Service name	myapp
Service type*	REPLICA
Number of tasks	1

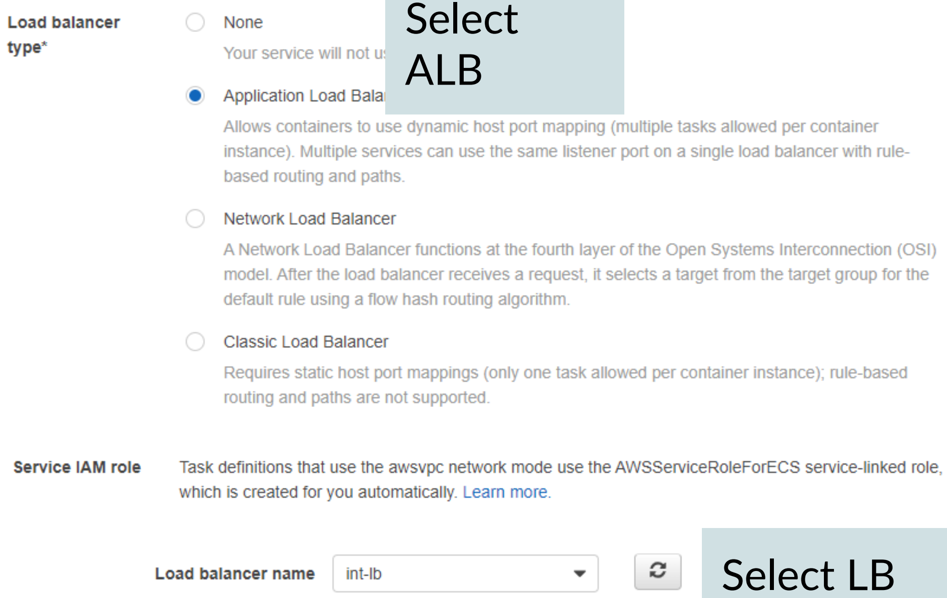
Fill in all required fields



Cluster VPC*	vpc-ff4ea499 (172.31.0.0/16)
Subnets*	subnet-0724b83186396edea (172.31.48.0/20) | private-subnet-01 - ap-southeast-1a assign ipv6 on creation: Disabled subnet-075aa06c5a22d1547 (172.31.64.0/20) | private-subnet-02 - ap-southeast-1b assign ipv6 on c
Security groups*	myapp-330
Auto-assign public IP	DISABLED

Select private subnet

Set the load balancer for the service



Load balancer type*

○ None

Your service will not us

● Application Load Bala

Allows containers to use dynamic host port mapping (multiple tasks allowed per container instance). Multiple services can use the same listener port on a single load balancer with rule-based routing and paths.

○ Network Load Balancer

A Network Load Balancer functions at the fourth layer of the Open Systems Interconnection (OSI) model. After the load balancer receives a request, it selects a target from the target group for the default rule using a flow hash routing algorithm.

○ Classic Load Balancer

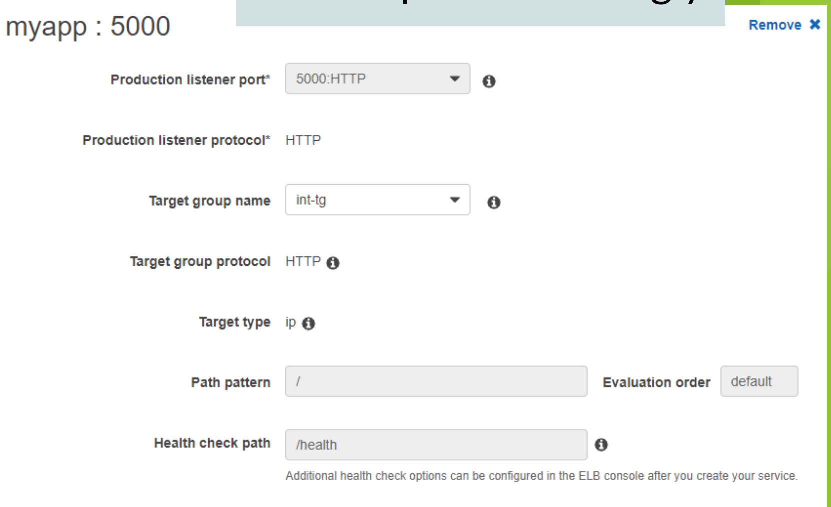
Requires static host port mappings (only one task allowed per container instance); rule-based routing and paths are not supported.

Service IAM role Task definitions that use the awsvpc network mode use the AWSServiceRoleForECS service-linked role, which is created for you automatically. Learn more.

Load balancer name int-lb

Select ALB

Select LB name



myapp : 5000

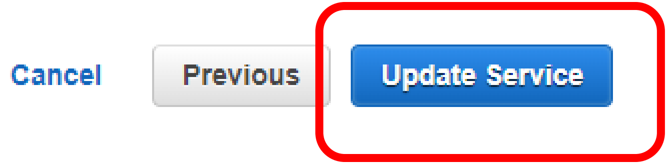
Remove ✖

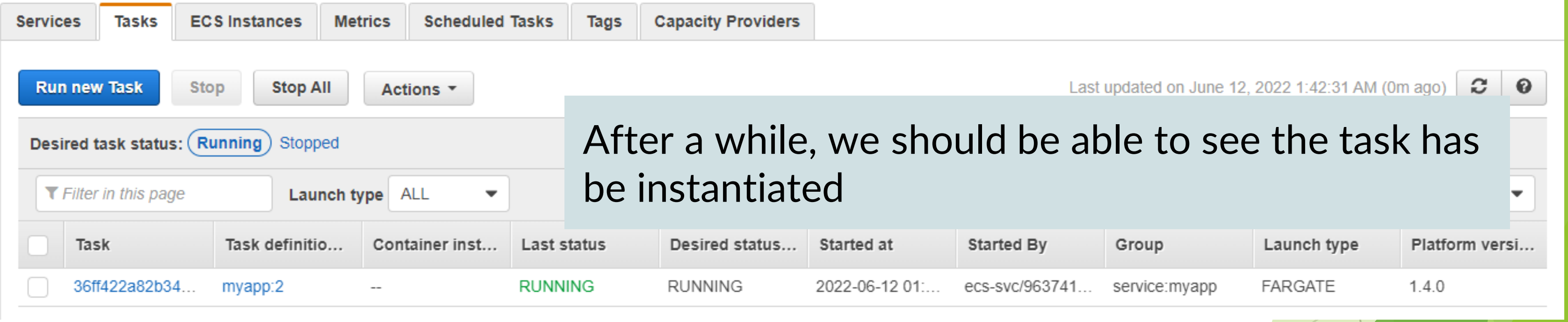
Production listener port*	5000:HTTP
Production listener protocol*	HTTP
Target group name	int-tg
Target group protocol	HTTP
Target type	ip
Path pattern	/
Evaluation order	default
Health check path	/health

Additional health check options can be configured in the ELB console after you create your service.

Set the port accordingly

Click Next until last page, then click Update Service





Last updated on June 12, 2022 1:42:31 AM (0m ago)

Desired task status: Running Stopped

Launch type ALL

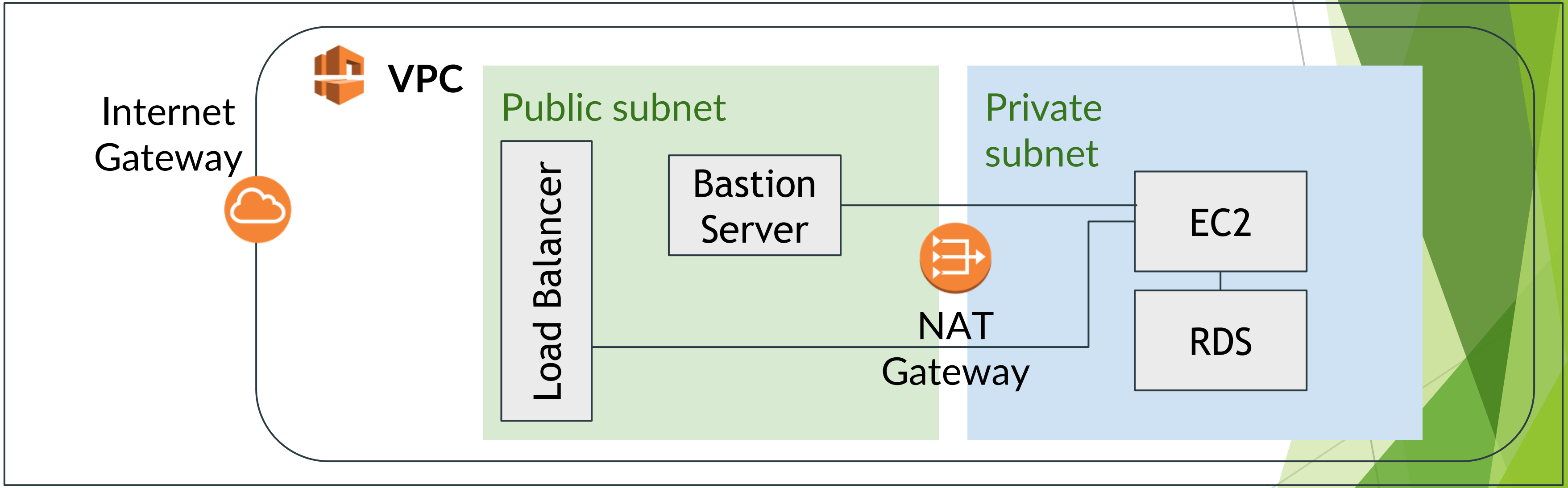
☐	Task	Task definitio...	Container inst...	Last status	Desired status...	Started at	Started By	Group	Launch type	Platform versi...
☐	36ff422a82b34...	myapp:2	--	RUNNING	RUNNING	2022-06-12 01:...	ecs-svc/963741...	service:myapp	FARGATE	1.4.0

After a while, we should be able to see the task has be instantiated

Let’s reduce deployment cost in Production

Manual Scaling with Lower Cost

Then, let’s design a simple architecture



Thank you.