

# DevOps, Software Evolution & Software Maintenance

Group P, 2024

## Authors

Name	Email
Andreas Andrä-Fredsted	aandr@itu.dk
Bence Luzsinszky	bluz@itu.dk
Christian Emil Nielsen	cemn@itu.dk
Michel Moritz Thies	mithi@itu.dk
Róbert Sluka	rslu@itu.dk

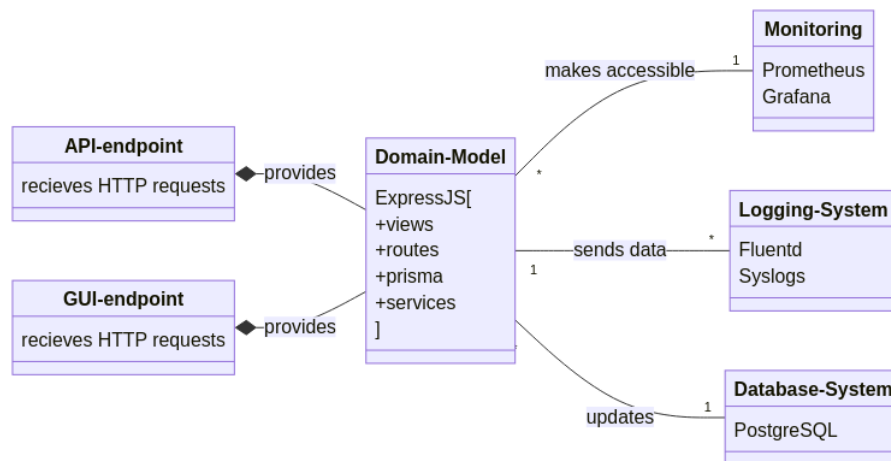
## System Perspective

### Architecture

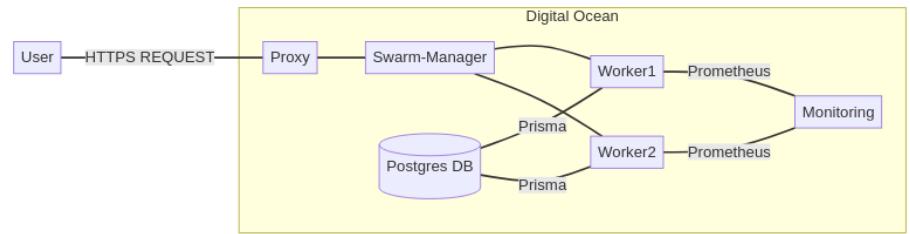
### Dependencies

### Viewpoints

**Module Viewpoint** To effectively capture this, the following class diagram presents the components of the web-app mapped to their respective dependencies.



### Components Viewpoint



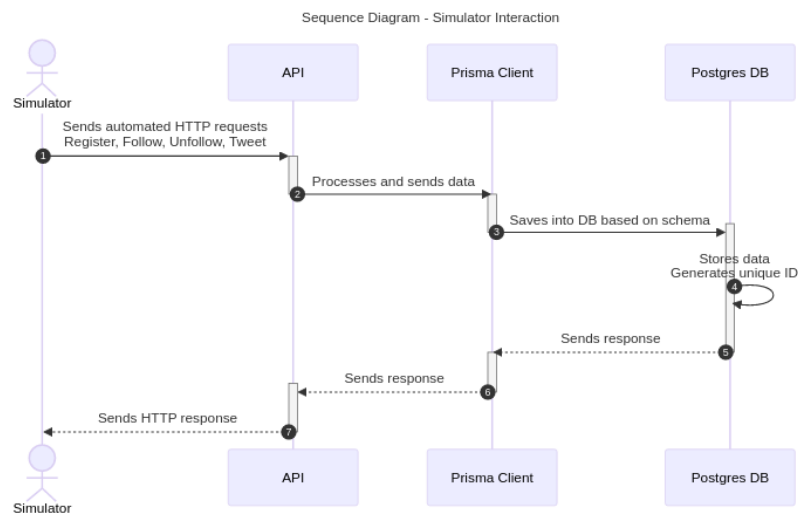
## Deployment Viewpoint

### Important interactions

The system can be interacted with in two ways:

- User Interface
- API for the simulator

A user (or the simulator) can register, follow/unfollow other users and send tweets.



## Current State

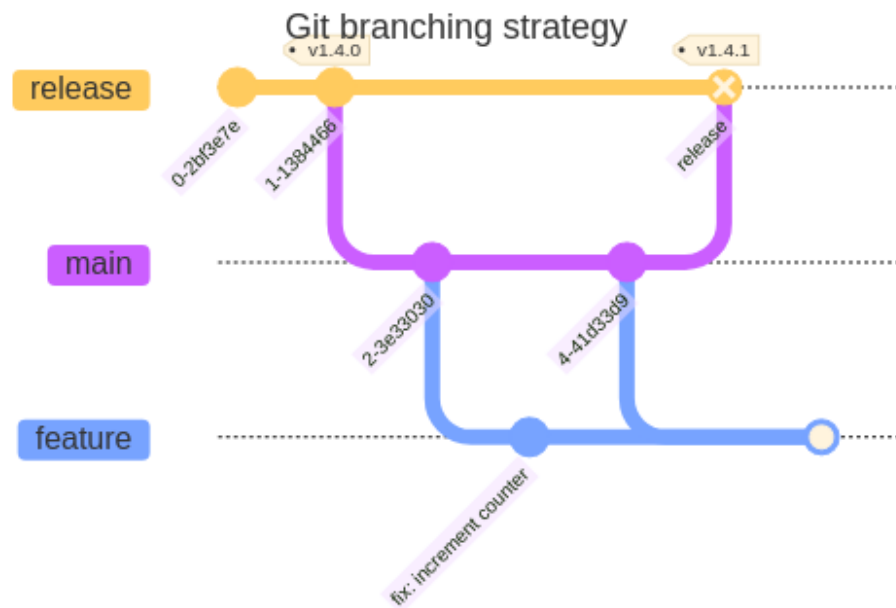
Security	Reliability	Maintainability
<b>4</b> Open issues <span style="color: red;">E</span> <div> <span style="background-color: #f8d7da; padding: 2px;">4 H</span> <span style="background-color: #d1ecf1; padding: 2px;">0 M</span> <span style="background-color: #d1ecf1; padding: 2px;">0 L</span> </div>	<b>0</b> Open issues <span style="color: green;">A</span> <div> <span style="background-color: #d1ecf1; padding: 2px;">0 H</span> <span style="background-color: #d1ecf1; padding: 2px;">0 M</span> <span style="background-color: #d1ecf1; padding: 2px;">0 L</span> </div>	<b>9</b> Open issues <span style="color: green;">A</span> <div> <span style="background-color: #d1ecf1; padding: 2px;">0 H</span> <span style="background-color: #fff3cd; padding: 2px;">1 M</span> <span style="background-color: #d1ecf1; padding: 2px;">8 L</span> </div>

The application is practically fully functional, apart from a single outstanding bug. While the application has minimal technical debt, it relies on legacy code and dependencies to test the application (test suite and simulator).

## Process Perspective

Why: ExpressJS, Prisma, Postgres

## Branching strategy



The chosen branching strategy loosely follows the Gitflow workflow. We chose to omit hotfix branches and merge the concept of a main/develop branch for simplicity. CI pipeline is triggered on open pull-requests from feature to main. CD pipeline is triggered on new commits on release. Release tag is bumped according to the contents of the release, using the semantic versioning protocol.

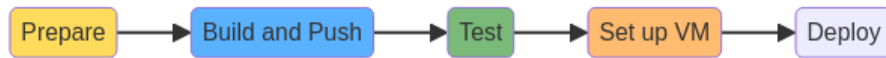
## Commit hooks

A pre-commit hook was added in d40fcba to lint and enforce commit messages and to follow the semantic versioning protocol. A CLI-tool was also added to aid developers write commit messages that follows the chosen protocol. Effectively standardizing a common development process, improving our process quality and readability of the git log.

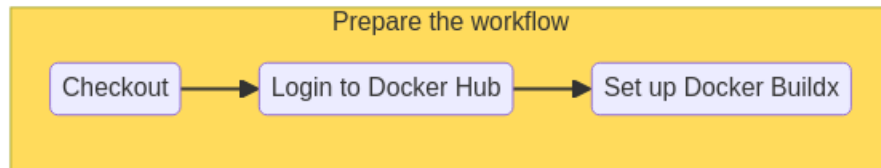
## CI/CD pipeline

Our CI/CD pipeline is based on **Github Actions**. We have a `deploy.yml` file that is automatically triggered when new data is pushed to the **release branch**.

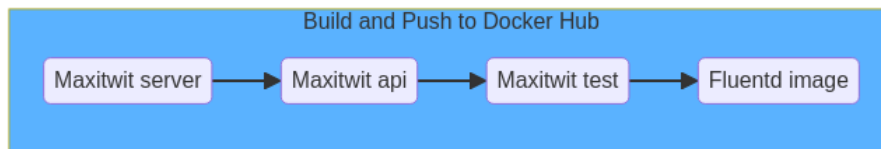
## CI/CD Pipeline



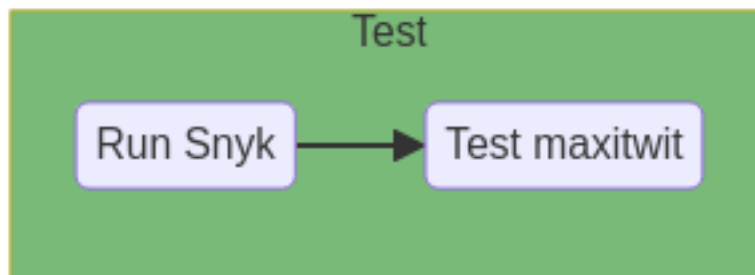
We prepare the workflow by checking out to our release branch, logging in to Docker Hub and setting up Docker Buildx so it can build the images.



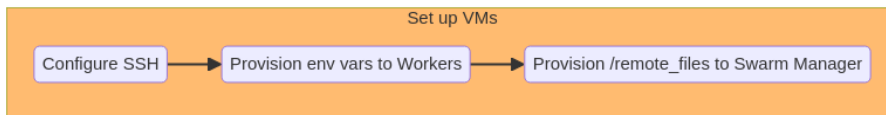
The workflow builds our images and pushes them to Docker Hub.



The workflow runs snyk to check for vulnerabilities, then builds our images and runs our tests suite against them.



The environment variables stored in GitHub Actions Secrets are given to the workers and the most recent `/remote_files` are SCPd to the Swarm Manager.



Finally we SSH onto the Swarm Manager and run the `deploy.sh` script to pull and build the new images.

## Monitoring

We use Prometheus and Grafana for monitoring. There are multiple metrics set up in our backend, that are sent to /metrics endpoint on our both our GUI and the API. Prometheus scrapes these endpoints and Grafana visualizes the data.

We set up a separate Droplet on DigitalOcean for monitoring, because we had issues with its resource consumption. The monitoring droplet runs Prometheus and Grafana, and scrapes the metrics from the Worker nodes of the Docker swarm.

## Security Assesment

- TODO sentence about our pipelines using root users which violates PloP

## Scaling strategy

We used Docker Swarm for horizontal scaling. The strategy is defined in compose.yml. One manager node is responsible for the load balancing and the health checks of two worker nodes. Worker nodes we have 6 replicas of the service running. We update our system with rolling upgrades. The replicas are updated 2 at a time, with 10s interval between the updates. The health of the service is monitored every 10s. If the service fails, it will be restarted with a maximum of 2 attempts.

## Lessons Learned

- evolution and refactoring
  - Implementation of Logging The evolution to the docker swarm architecture made the reconfiguration of subsystems a necessity. Specifically the reimplementaion of logs proved difficult, as we had to sync logs across workers. To achieve this, it was attempted to have fluentd containers in each worker gather logs and send them to a seperate digital ocean droplet with elasticsearch and kibana running. This however proved infeasible within the constraints of this class, as elasticsearch kept crashing due to the limited resources provided to it on the droplet. Thus, we defaulted to store logs in a /logs folder on the droplet also containing the load balancer. Overall, this proved a learning experience for how a change to the tech-stack can make other segments obsolete, thus increasing the amount of refactoring required for a change to be feasible. In this specific case, a migration to the EFK-stack before the implementation of the docker swarm would have allowed for a more seamless and robust evolution of the application.
- operation
- maintenance

### **Maintaining a performant DB**

We noticed the performance of the public timeline endpoint getting slower as the database grew. To remedy this, we wrote a shell script to query the performance table of our production database to identify which relations needed indices.