CSML

A Collaborative and Secure Machine Learning Framework

Undecidable:

Wenjie Qiu, Dyon Buitenkamp, Yongcheng Song, Yu Hong

Contents

- Collaborative:

  Docker and swarm.

- Secure:

  TEE, SCONE, randomness and anonymous.

- Machine Learning:

  opencv, dynet, caffee and more.

- Summary

Collaborative?

Make it accessible to everyone & user-friendly.

- Applications in Docker

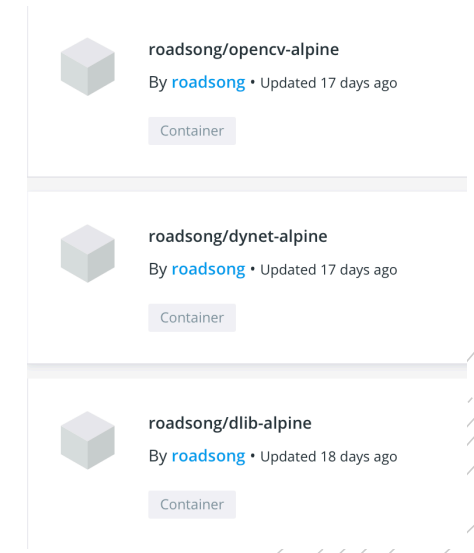
```
REPOSITORY                TAG                 IMAGE ID            CREATED             SIZE
roadsong/opencv-alpine    latest              b82178f1b66d        2 weeks ago         2GB
roadsong/dynet-alpine     latest              8d62ca960594        2 weeks ago         1.15GB
roadsong/dlib-alpine      latest              0ce6aac0631b        2 weeks ago         417MB
```

```
docker pull roadsong/opencv-alpine
```

```
docker pull roadsong/dynet-alpine
```

```
docker pull roadsong/dlib-alpine
```

Each docker for one specific application!



https://hub.docker.com/u/roadsong

Collaborative?

- Docker network
- Docker swarm

```
NAME                DRIVER                                  SCOPE
bridge              bridge                                  local
csml-cluster        weaveworks/net-plugin:latest_release    swarm
docker_gwbridge     bridge                                  local
host                host                                    local
ingress             overlay                                 swarm
```

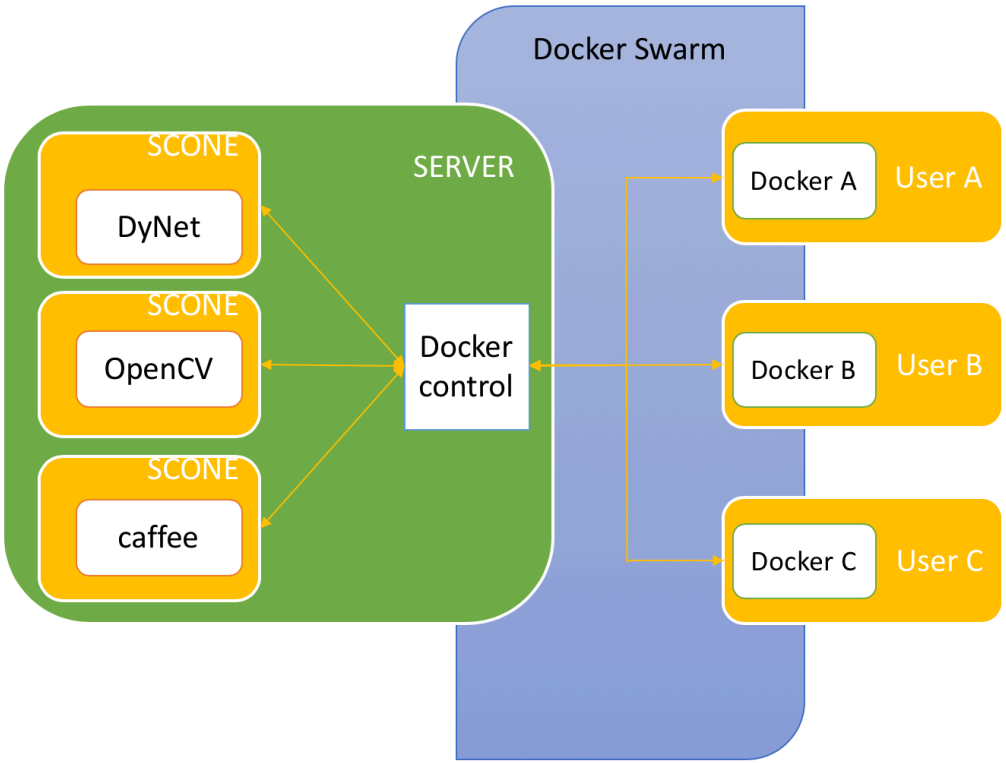
So that every application docker can work together and users can join our network!

- Kubernetes

Overkill in this small project, practical choice for the real world.

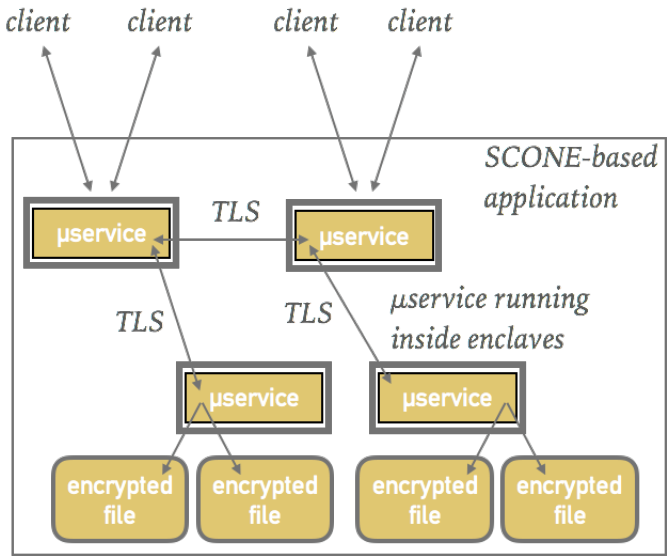
Users in user docker can join the platform remotely, submitting their jobs, waiting for server training, inferencing, and getting the results.

CSML
System Design



Secure

- Trusted Execution Environment (TEE)
- But TEE has usability problem…
- SCONE



Secure

- Randomness

```
func generateID(crypto bool) string {
        b := make([]byte, 32)
        var r io.Reader = random.Reader
        if crypto {
                r = rand.Reader
        }
        for {
                if _, err := io.ReadFull(r, b); err != nil {
                        panic(err) // This shouldn't happen
                }
                id := hex.EncodeToString(b)
                // if we try to parse the truncated for as an int and we don't have
                // an error then the value is all numberic and causes issues when
                // used as a hostname. ref #3869
                if _, err := strconv.ParseInt(TruncateID(id), 10, 64); err == nil {
                        continue
                }
                return id
        }
}
```

Start up a new docker container, its ID is RANDOM!

Guaranteed by GO programming language.

- Anonymous

We don't know you, we we do know your ID!

Solve privacy issue!

```
CONTAINER ID
b7dbb2e960fb
60955c62b830
ea3499c80b2e
f3bb1eabb08d
```

Machine Learning

- Although SCONE is secure and easy to use (NO…)
- We must build ML application (C++) into SCONE!

  See their Dockerfile(s) for detail

- Sending email to SCONE group to get access…
- I cannot redistribute them, so dockers I created are only for evaluation use only, they are not truly secure.

- What do we have?
- opencv - Open souce computer vision library

  https://github.com/opencv/opencv
- dynet – The dynamic

  https://github.com/clab/dynet
- dlib – C++ ML algorithms and tools

  https://github.com/davisking/dlib

dy/net

Summary

- Security

  TEE -> SGX -> SCONE

- Privacy

  We don't know your identity; we only know your (container) ID.

- Usability

  Sending requests, waiting, getting results…

- Scalability & Extensibility

  Lots of users?

  No worry, this is docker! Industry level kubernetes!

  Add some other applications?

  Build the docker, write code and join us!

- See our github repo:

  https://github.com/Roadsong/CSML

Reference

- https://github.com/opencv/opencv
- https://github.com/clab/dynet
- https://github.com/davisking/dlib
- https://sconedocs.github.io/
- https://hub.docker.com/u/roadsong
- https://github.com/Roadsong/CSML