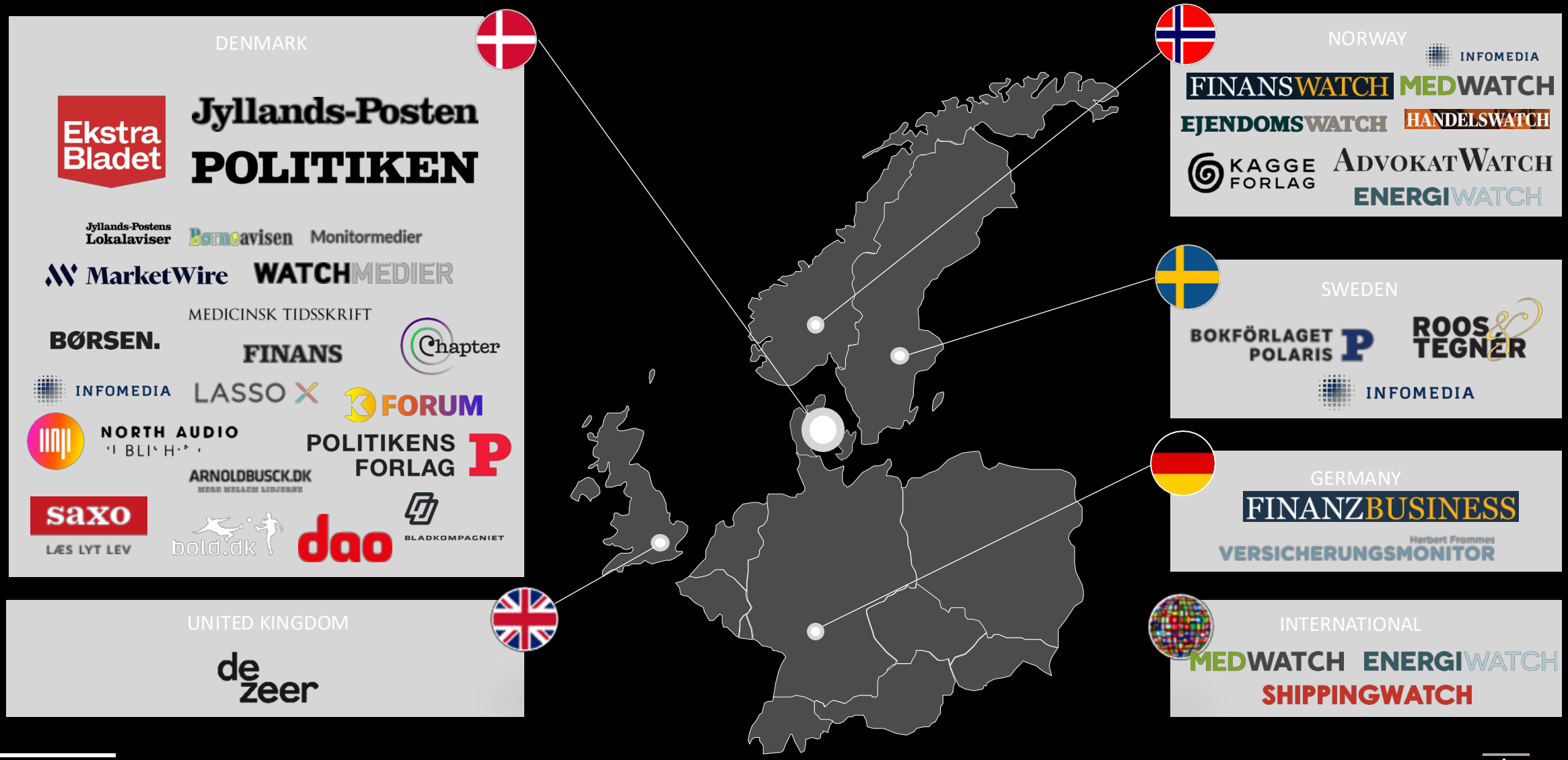
JP|Politiken Media Group

Paris Nakita Kejser
KubeCon Salt Lake City 2024

Transforming ML Engineering:
A Kubernetes Adoption Journey

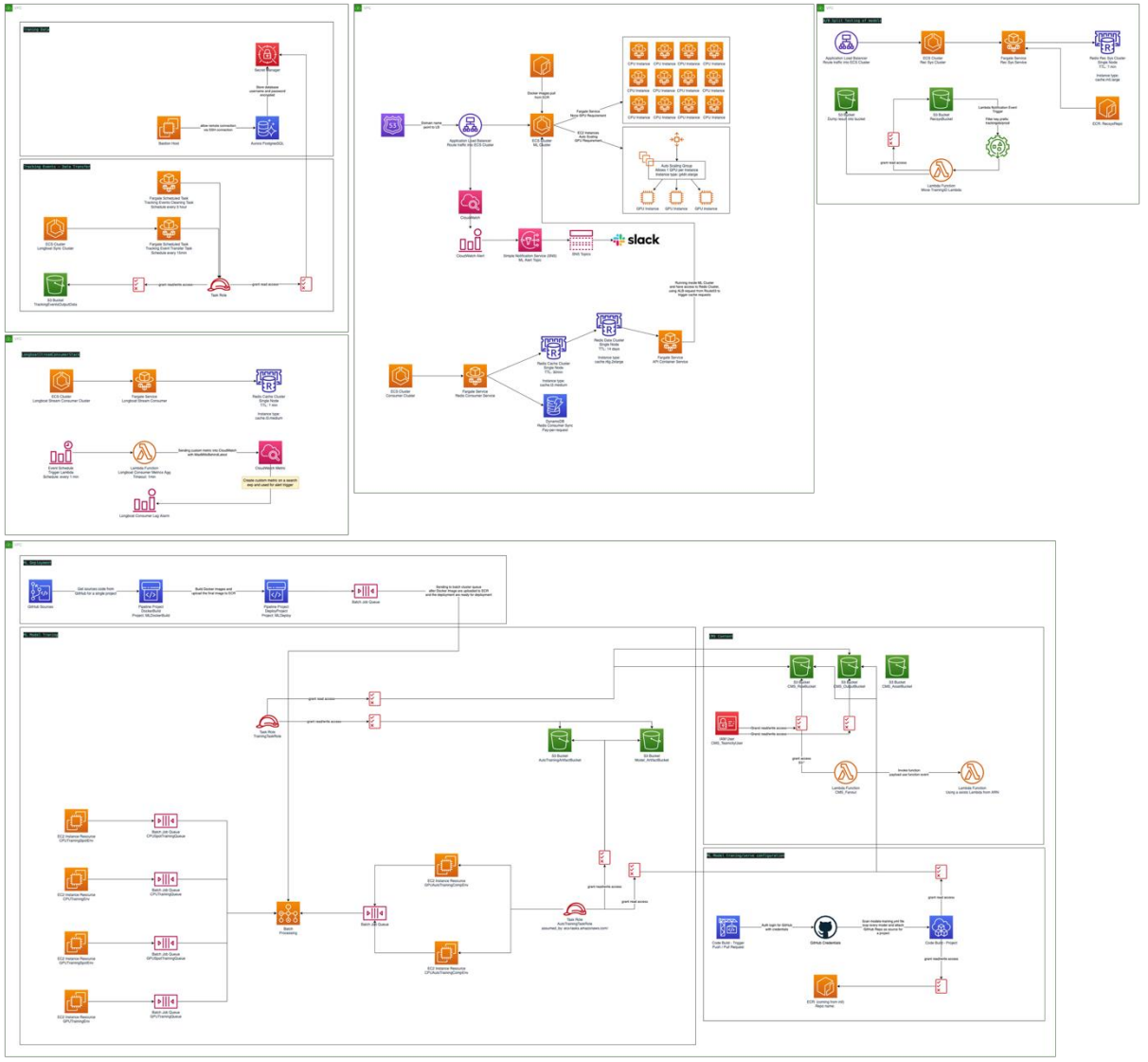
JP|Politiken Media Group



Building an ML platform for research

- Simplifying the training and serving of models
- Flexibility to use different libraries
- Self-service for model training and serving
- Error reporting and monitoring of the models
- Full control over the platform and resources

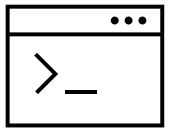
The chaos



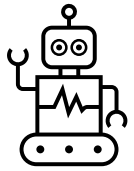
Understanding challenges and limitations of our ML platform

- Strict limits on ML model training and deployment
- Issues with testing and retraining models
- Knowledge requirements for ML platform changes
- Resource allocation challenges in cloud computing

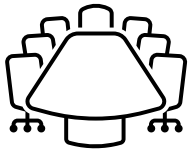
Lessons learned from building our ML platform



Streamlining infrastructure management



Gaining control over the ML platform



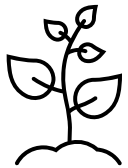
Enhancing collaboration and onboarding

Rethinking, rebuilding, and making informed decisions

Why we tried to optimize again

Platform stability

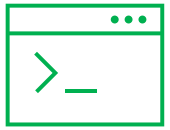
Balancing old and new



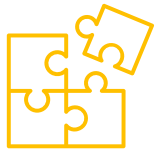
A growing team

Increased flexibility and deployment

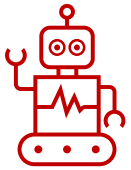
Was the new version a success?



Platform stability



Increased flexibility and deployment



Flexibility to run software as desired

Faster onboarding of new employees

The rethinking period: exploring options

It’s time to stop and rethink what our expected goal should be!

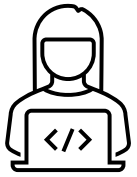
Consider another attempt to build our own ML Platform

Evaluate existing solutions

Leverage Kubernetes for flexibility

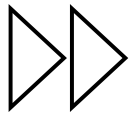
Why not existing solutions

Effort vs. benefit



Unreliable updates and limited support

Need for greater flexibility and transparency



Faster onboarding for new employees

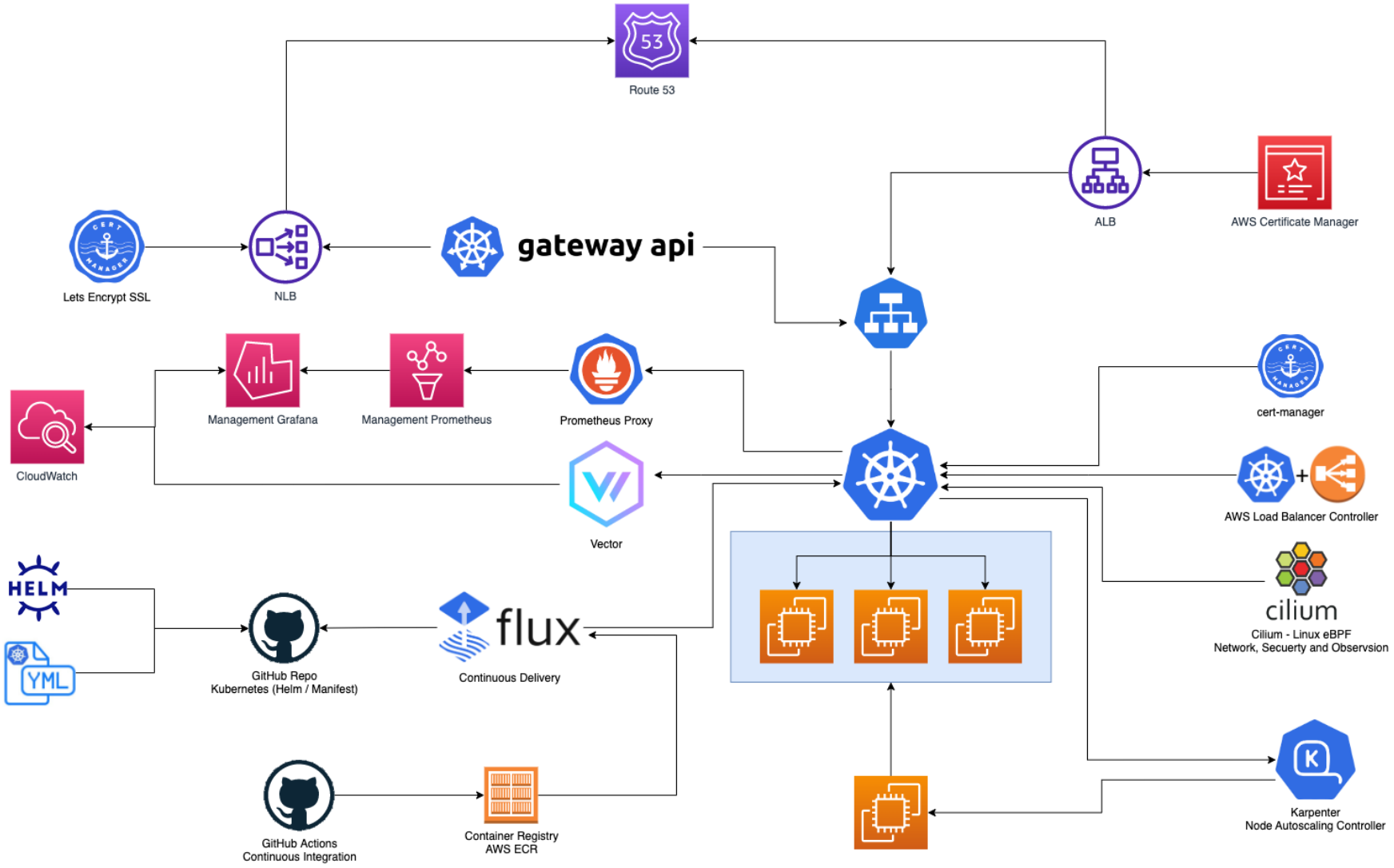
Why we landed on Kubernetes

- Simplified resource management
- Faster debugging, testing and development for ML specialists
- Freedom to prototype and test
- Faster onboarding
- Less infrastructure knowledge required

Benefits realized with Kubernetes

- Accelerated CI/CD with GitOps principles
- Enchanted debugging and development
- Seamless autoscaling
- Increased agility and efficiency
- Easier observability
- Easier security

Our new ML architecture is born!



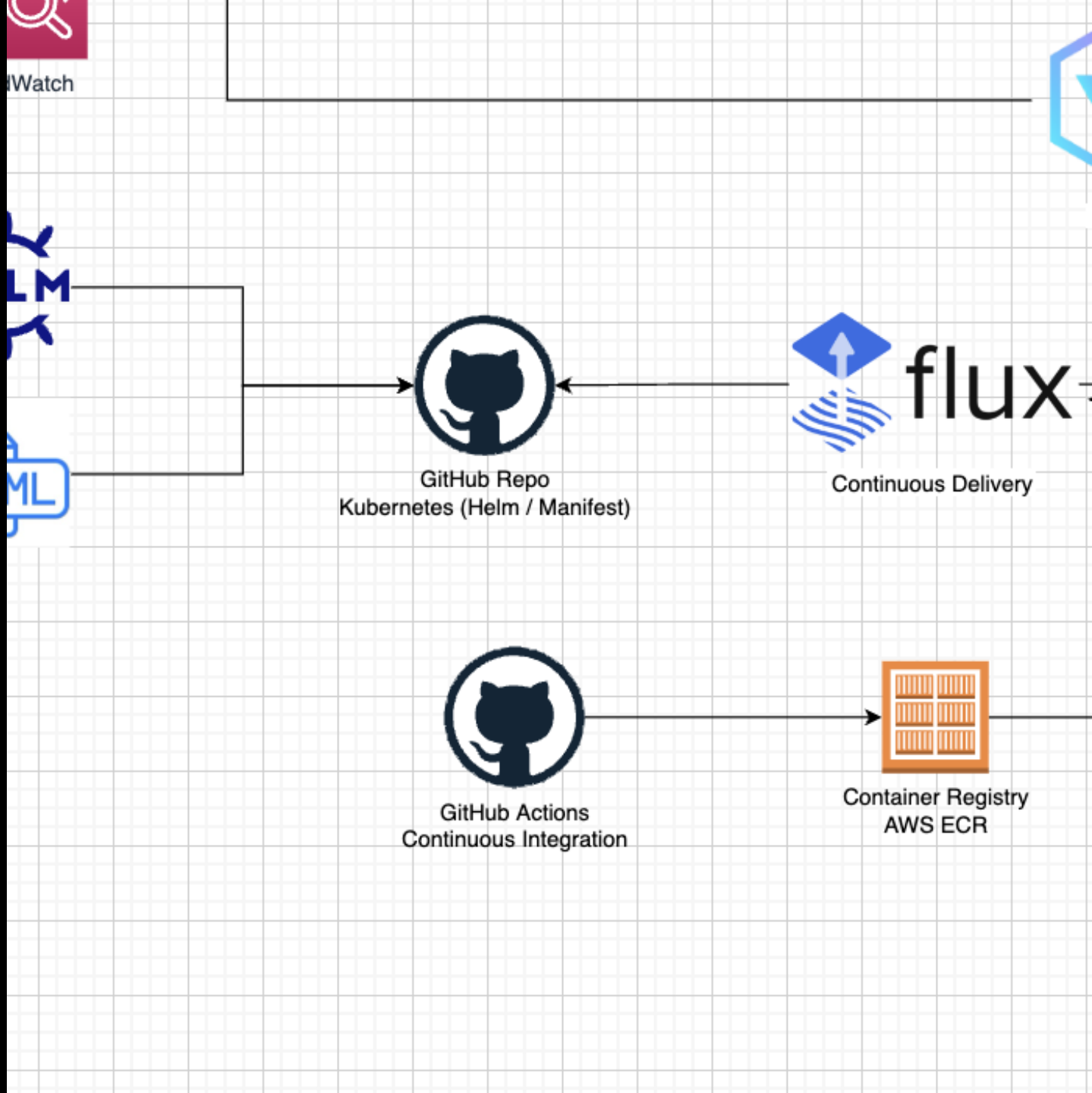
Applied software in Kubernetes

Automation (CI/CD)

Flux

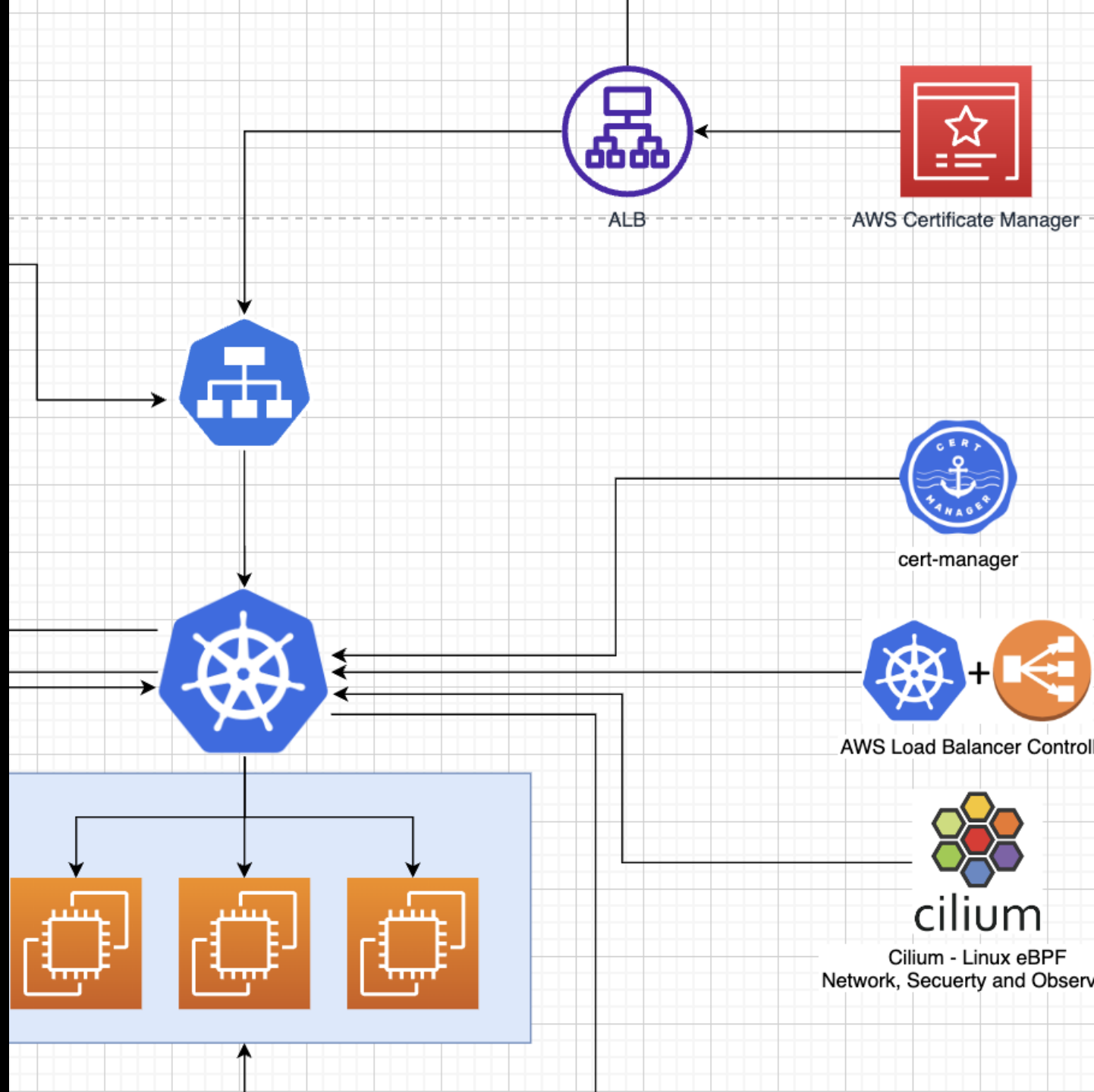
Helm Charts

GitHub



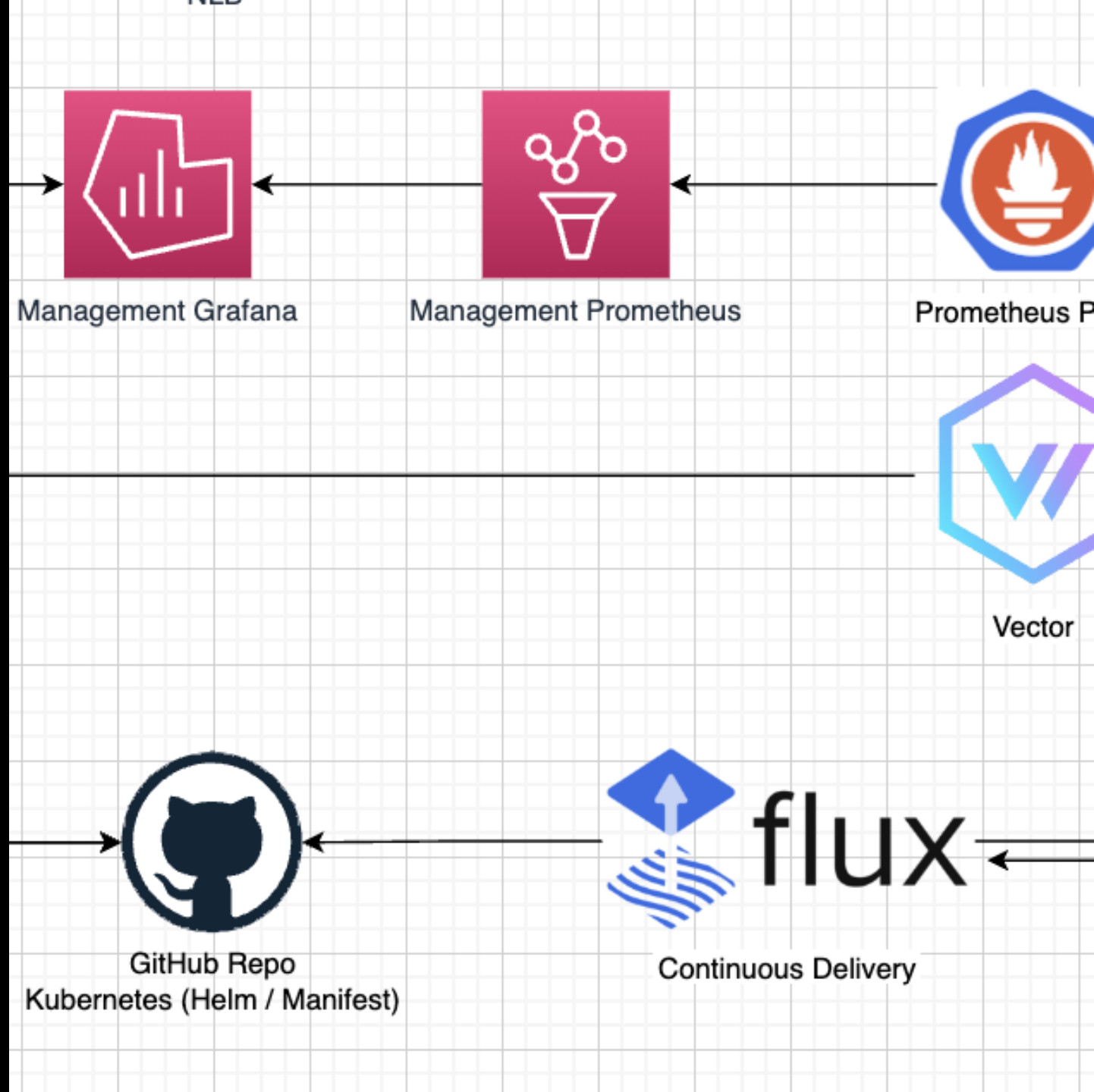
Security and Networking

- Cilium
- Cert-manager
- Gateway API
- AWS LB-Controller



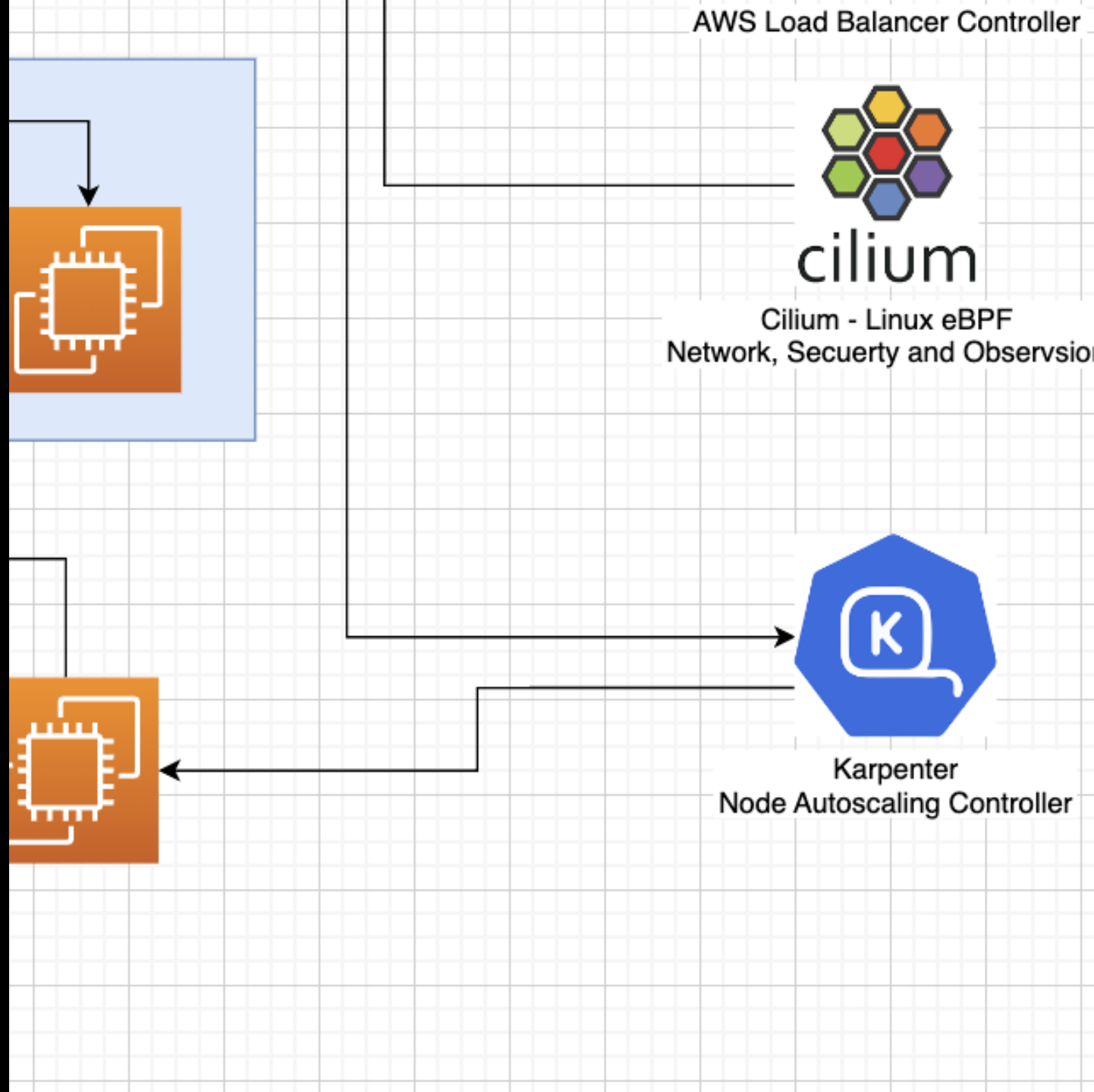
Observability

- Prometheus
- Grafana
- Vector
- Hubble



Scalability

- Karpenter
- K8s metrics API
- HPA



Thanks

parisnakitakejser
paris.kejser@jppol.dk