Running Large-Scale Scheduling Simulations with Virtual Kubelet

KubeCon NA 2023 – Nov. 8

David R. Morrison

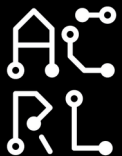Applied Computing Research Labs

https://appliedcomputing.io

Applied Computing does open-source research and development in distributed systems, with a particular focus on modeling, scheduling, and optimization.

David R. Morrison
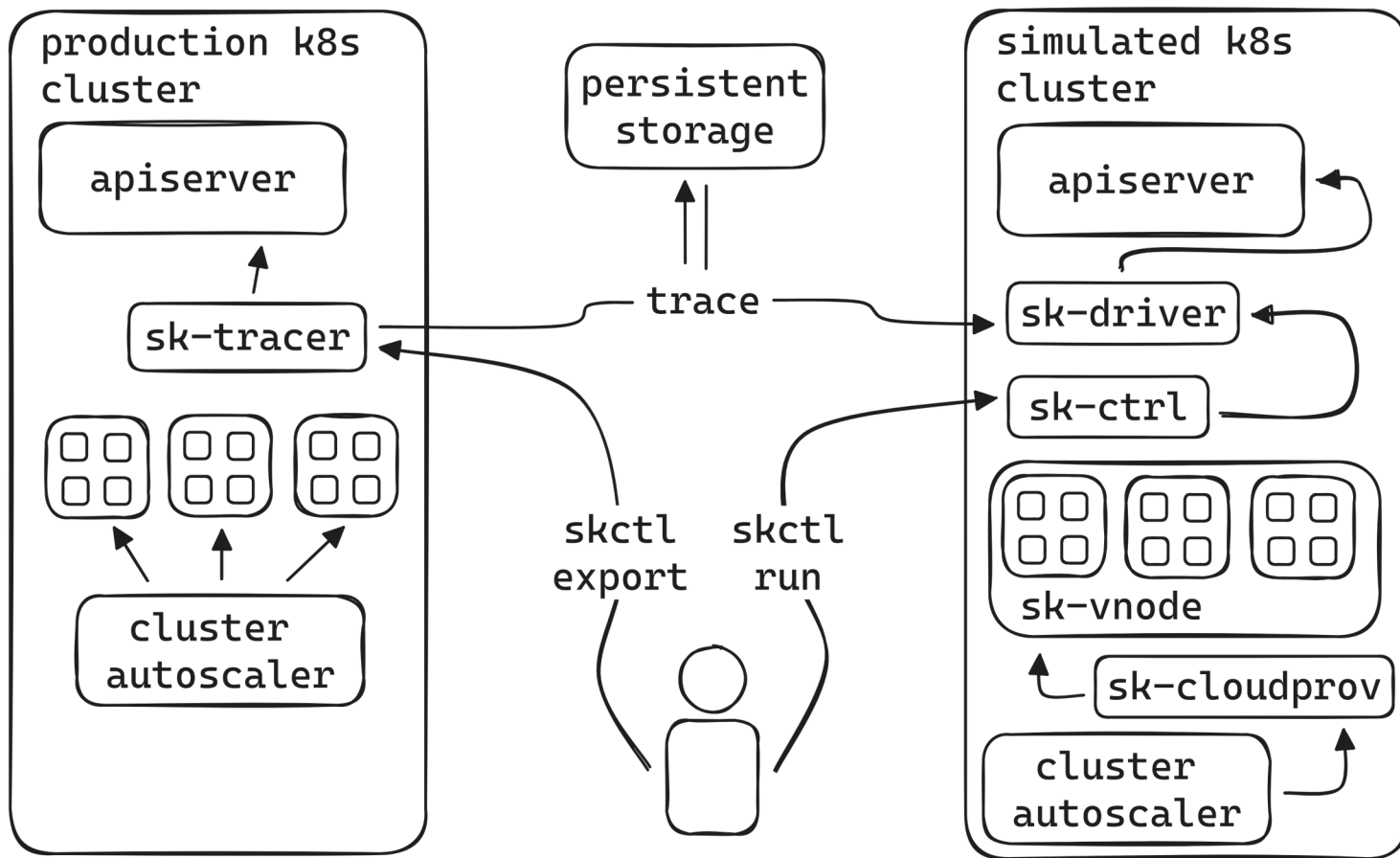
Applied Computing Research Labs

https://appliedcomputing.io

Why simulation?

David R. Morrison

Applied Computing Research Labs

https://appliedcomputing.io

production k8s cluster

apiserver

sk-tracer

cluster autoscaler

persistent storage

trace

skctl export

skctl run

simulated k8s cluster

apiserver

sk-driver

sk-ctrl

sk-vnode

sk-cloudprov

cluster autoscaler

David R. Morrison

Applied Computing Research Labs

https://appliedcomputing.io

David R. Morrison
Applied Computing Research Labs
https://appliedcomputing.io

David R. Morrison
Applied Computing Research Labs
https://appliedcomputing.io

David R. Morrison

Applied Computing Research Labs

https://appliedcomputing.io

production k8s cluster

apiserver

sk-tracer

cluster autoscaler

persistent storage

trace

skctl export

skctl run

simulated k8s cluster

apiserver

sk-driver

sk-ctrl

sk-vnode

sk-cloudprov

cluster autoscaler
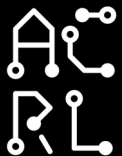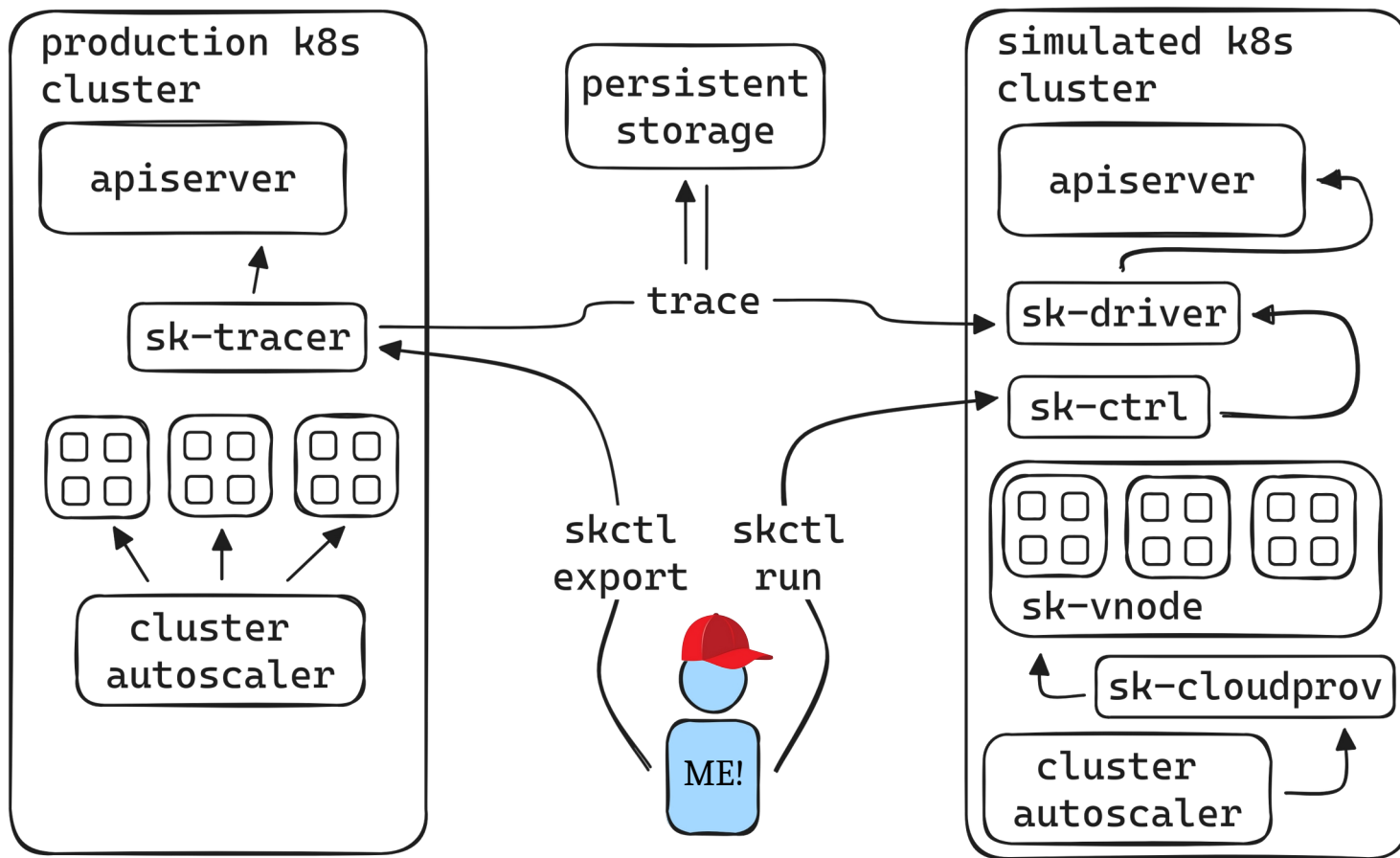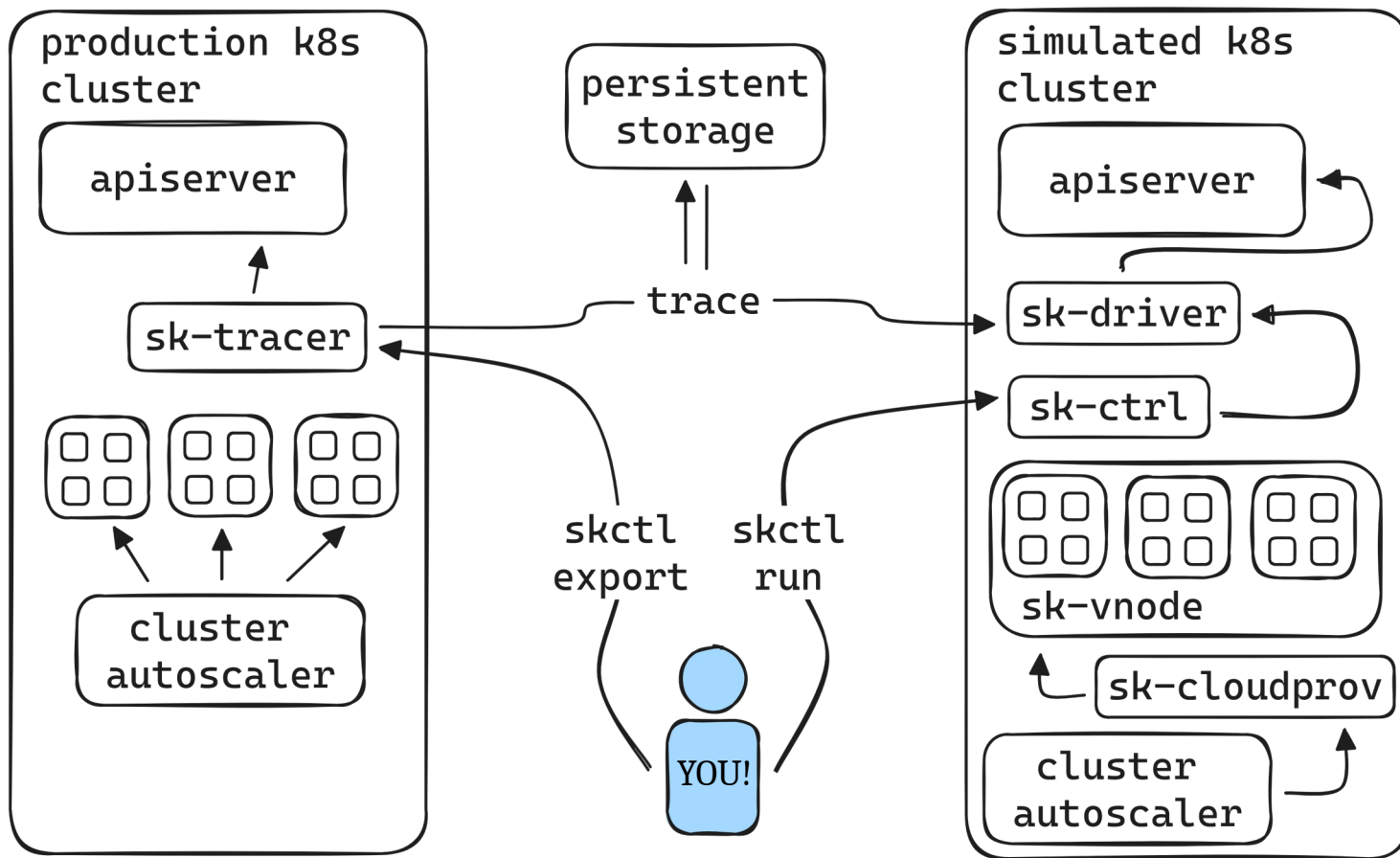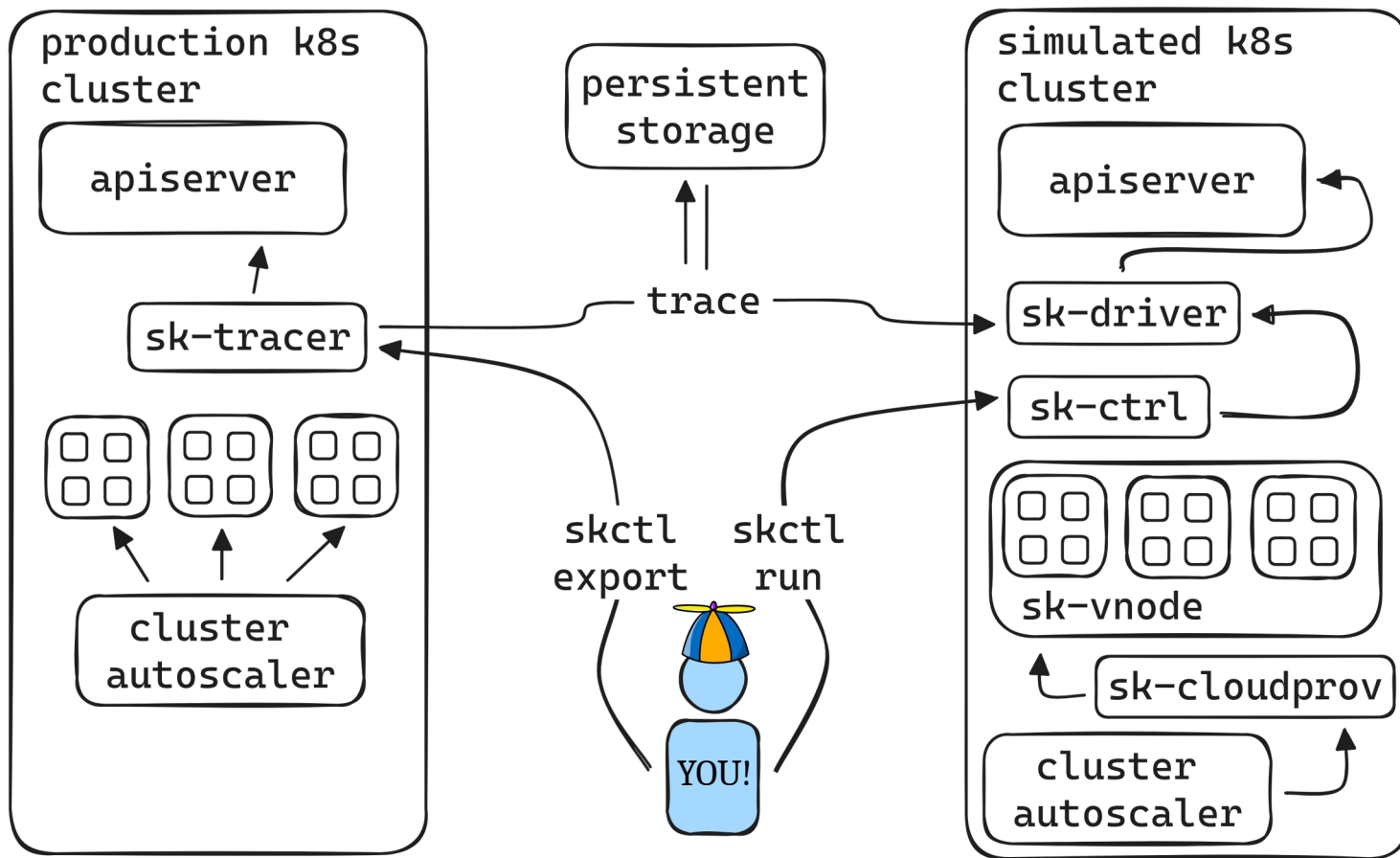
David R. Morrison
Applied Computing Research Labs
https://appliedcomputing.io
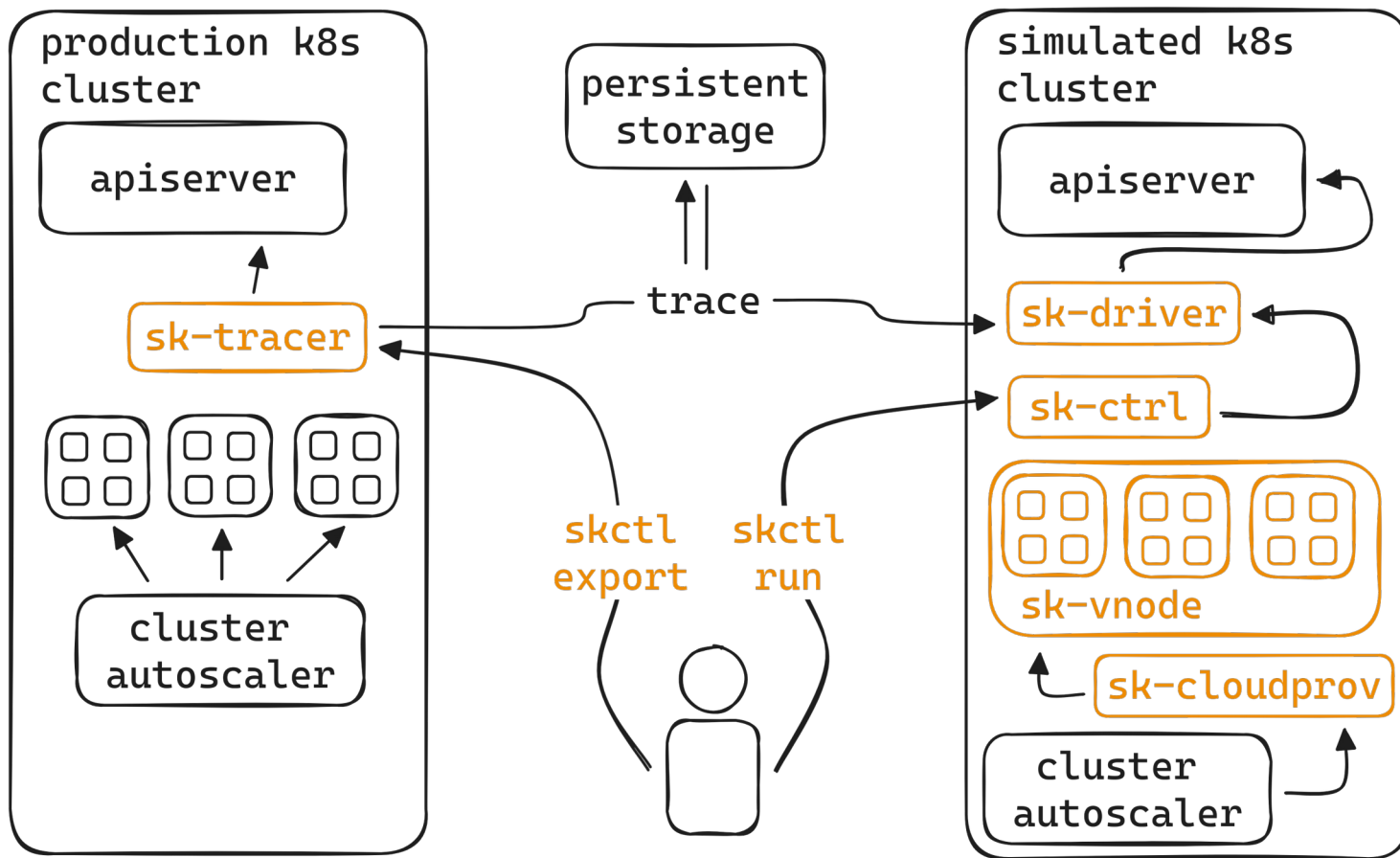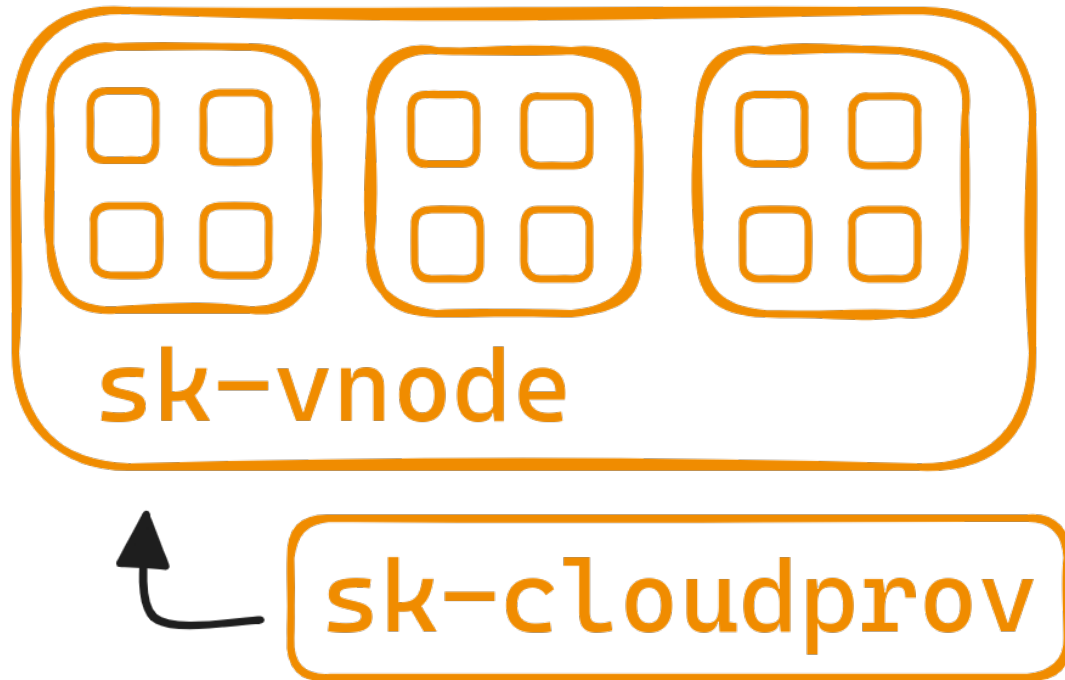
- sk-vnode:
  - Virtual Kubelet-based
  - Configurable node properties
  - Watches annotations to control pod lifecycle
- sk-cloudprov
  - gRPC cloud provider for Cluster Autoscaler
  - Scales the sk-vnode deployment up and down
  - Uses pod deletion cost feature to select nodes to terminate

David R. Morrison

Applied Computing Research Labs

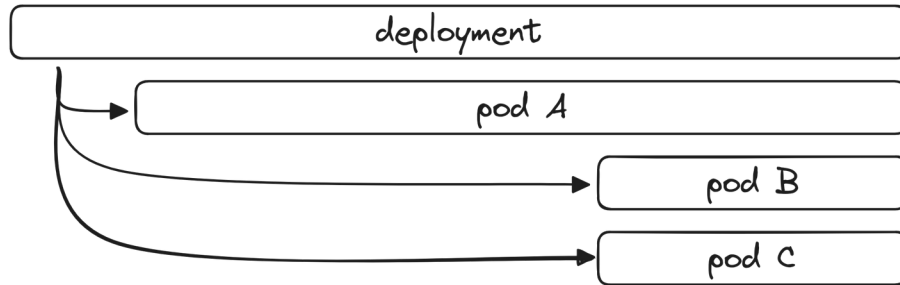https://appliedcomputing.io

- sk-tracer:
  - Watches apiserver for resources and pods
  - Records a timeline of "important" events
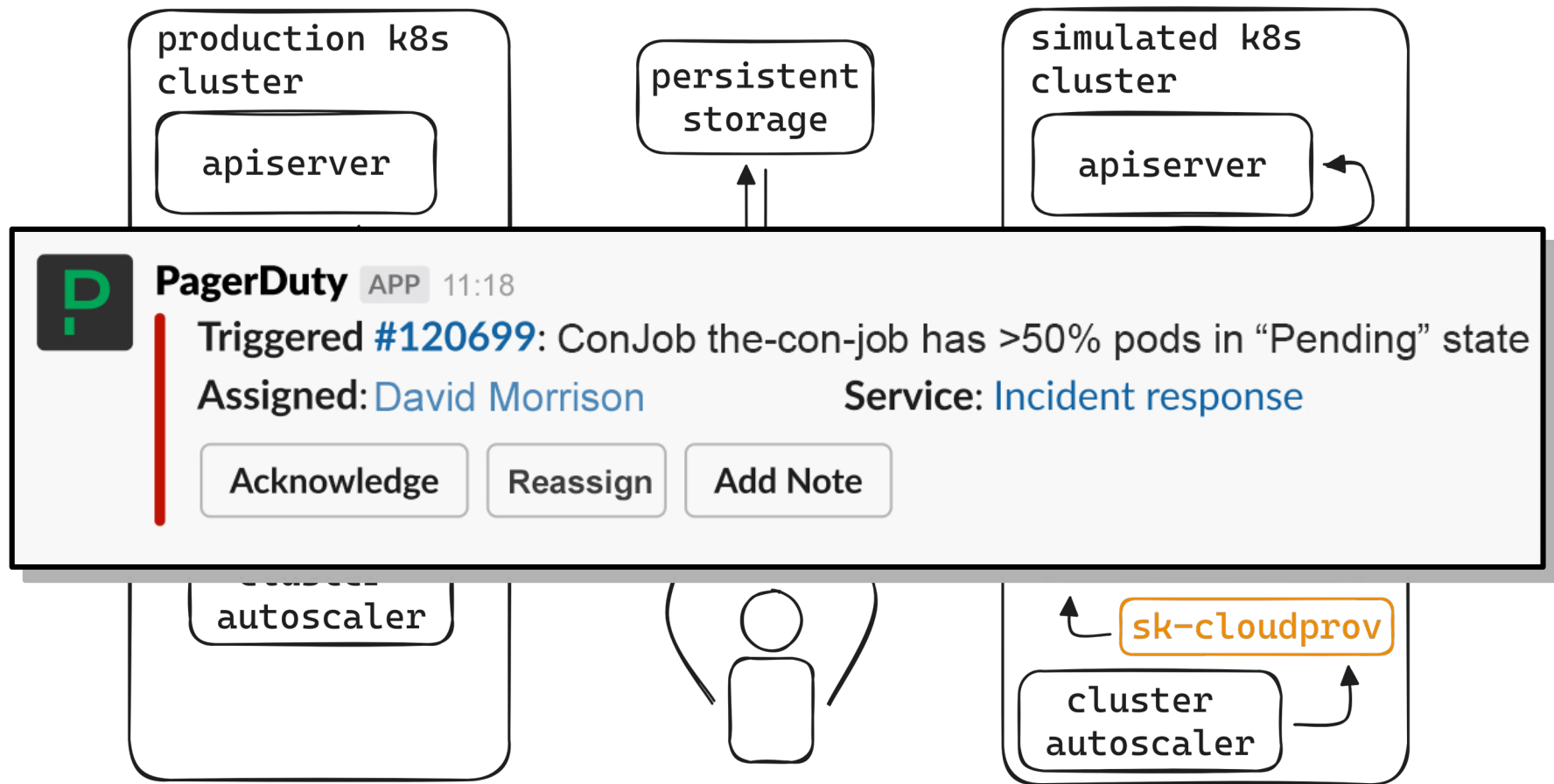 - Exports data to persistent storage on-demand

David R. Morrison

Applied Computing Research Labs

https://appliedcomputing.io

production k8s cluster

apiserver

persistent storage

simulated k8s cluster

apiserver

P **PagerDuty** APP 11:18

Triggered #120699: ConJob the-con-job has >50% pods in "Pending" state

Assigned: David Morrison   Service: Incident response

Acknowledge   Reassign   Add Note

cluster autoscaler

sk-cloudprov

cluster autoscaler

David R. Morrison

Applied Computing Research Labs

https://appliedcomputing.io

production k8s cluster

apiserver

persistent storage

simulated k8s cluster

apiserver

PagerDuty APP 11:18

PagerDuty APP 11:18

Triggered #120699: ConJob the-con-job has >50% pods in "Pending" state

Assigned: David Morrison   Service: Incident response

Acknowledge   Reassign   Add Note

cluster autoscaler

David R. Morrison

Applied Computing Research Labs

https://appliedcomputing.io

Future directions for SimKube

- Comparing and visualizing results
- Trace generation and modification
- Integrate with KWOK or similar
  - (two KWOK talks at KubeCon! 1, 2)
- Golang rewrite?

David R. Morrison

Applied Computing Research Labs

https://appliedcomputing.io

Are you interested in contributing to, using, or sponsoring SimKube? Come talk to me!

David R. Morrison

Applied Computing Research Labs

https://appliedcomputing.io
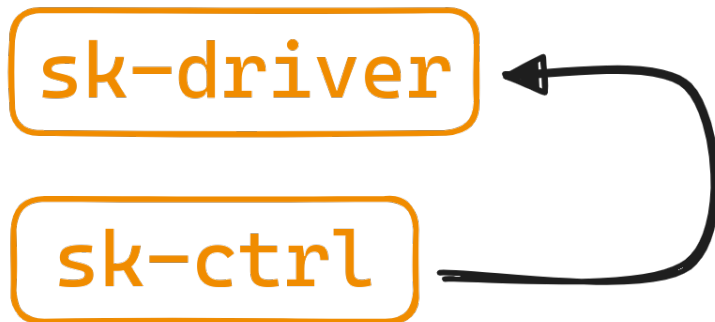
Thank you! Any questions?

https://github.com/acrlabs/simkube
https://github.com/acrlabs/kubecon-na2023
https://blog.appliedcomputing.io
https://hachyderm.io/@drmorr

David R. Morrison
Applied Computing Research Labs
https://appliedcomputing.io

- sk-ctrl: watches for simulation CR, creates driver
- sk-driver
  - Runner: replays recorded trace
  - Mutator: watches incoming pod objects, applies appropriate labels, annotations, and node-selectors

David R. Morrison

Applied Computing Research Labs

https://appliedcomputing.io