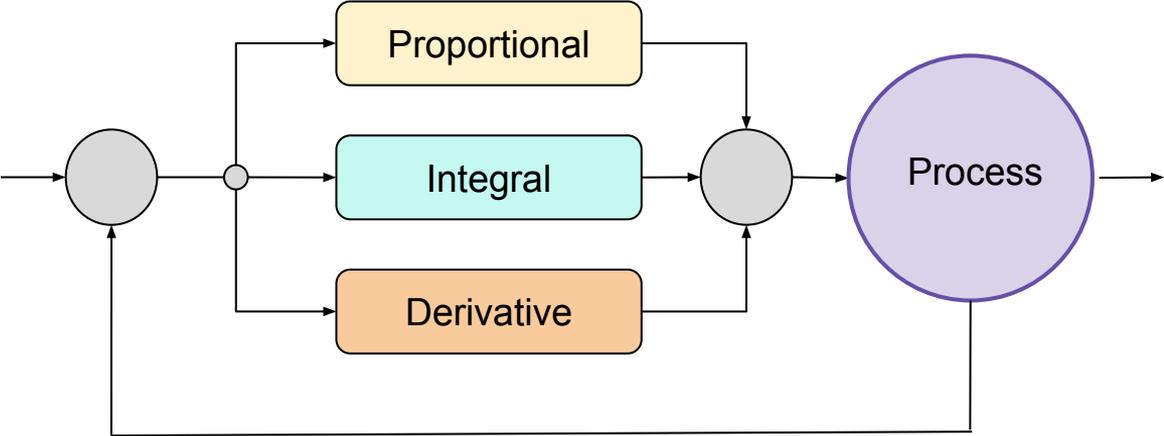
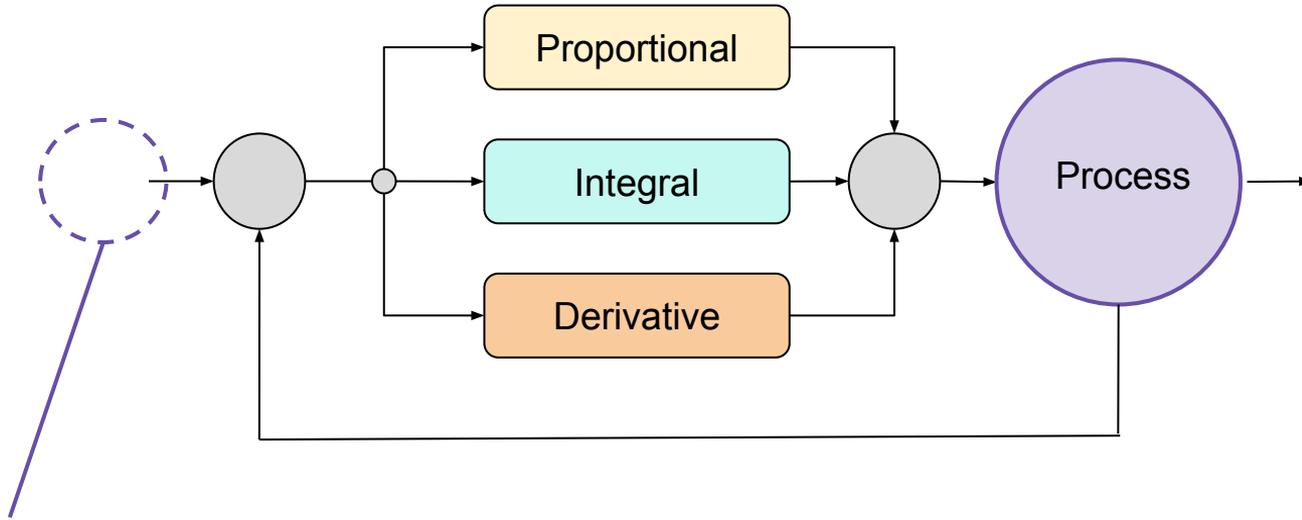

PID Auto-scaling a Linkerd Service Mesh

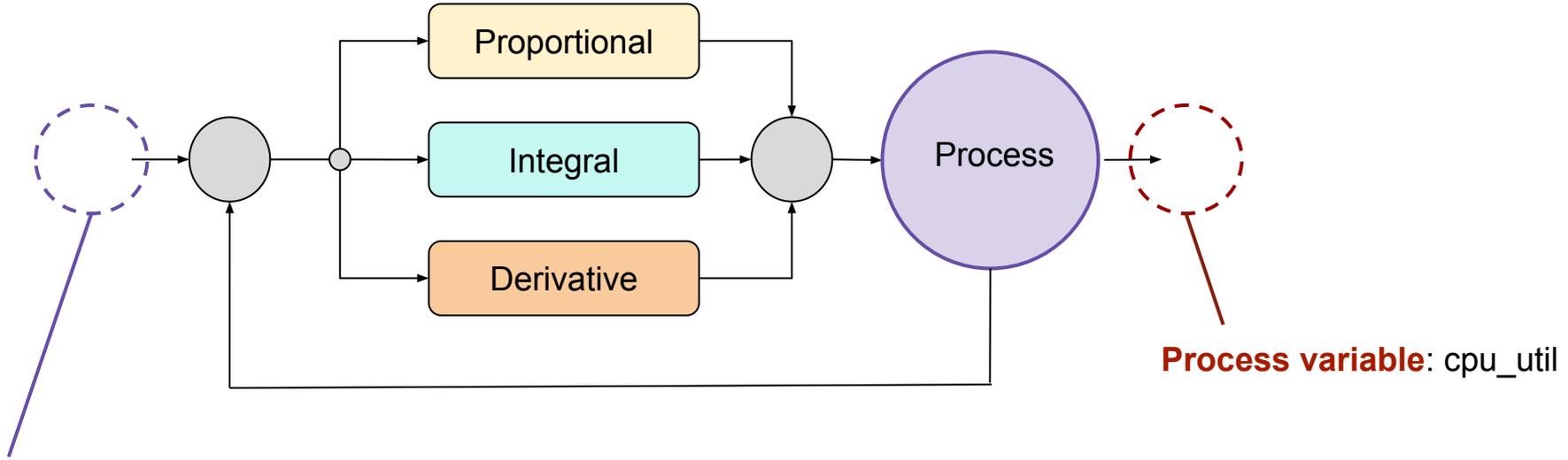
J Evans | Strava

What is a PID Controller?



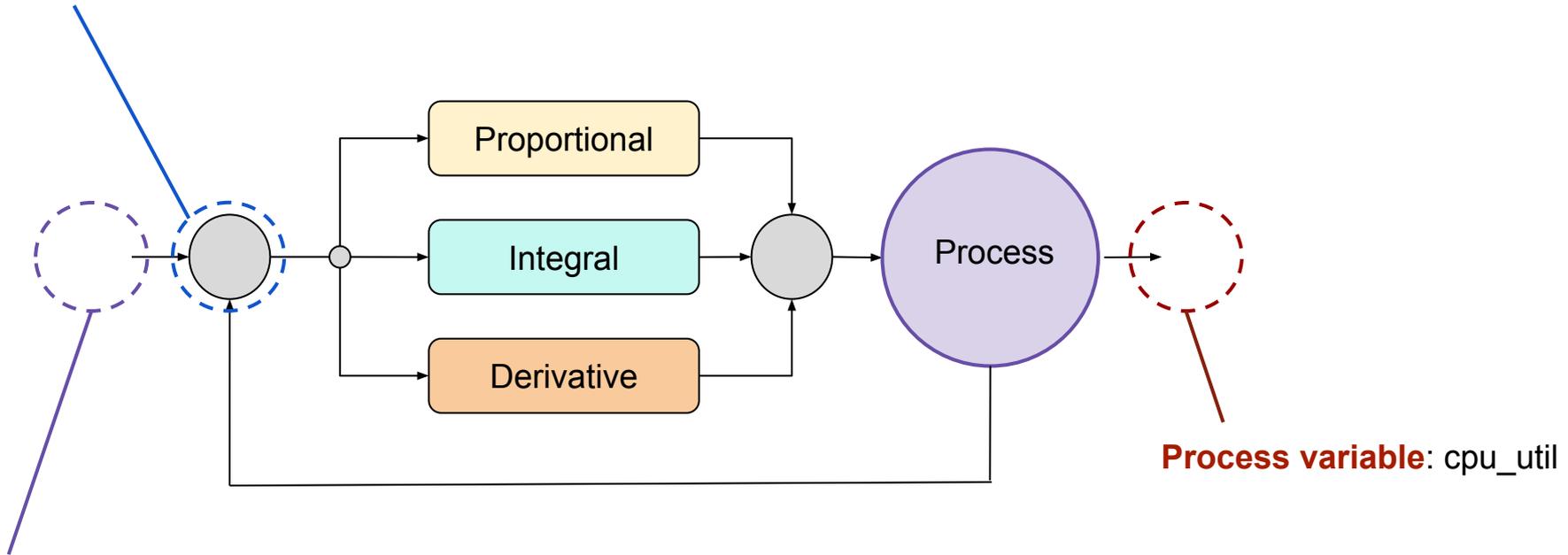


Setpoint: maintain 80% cpu_util



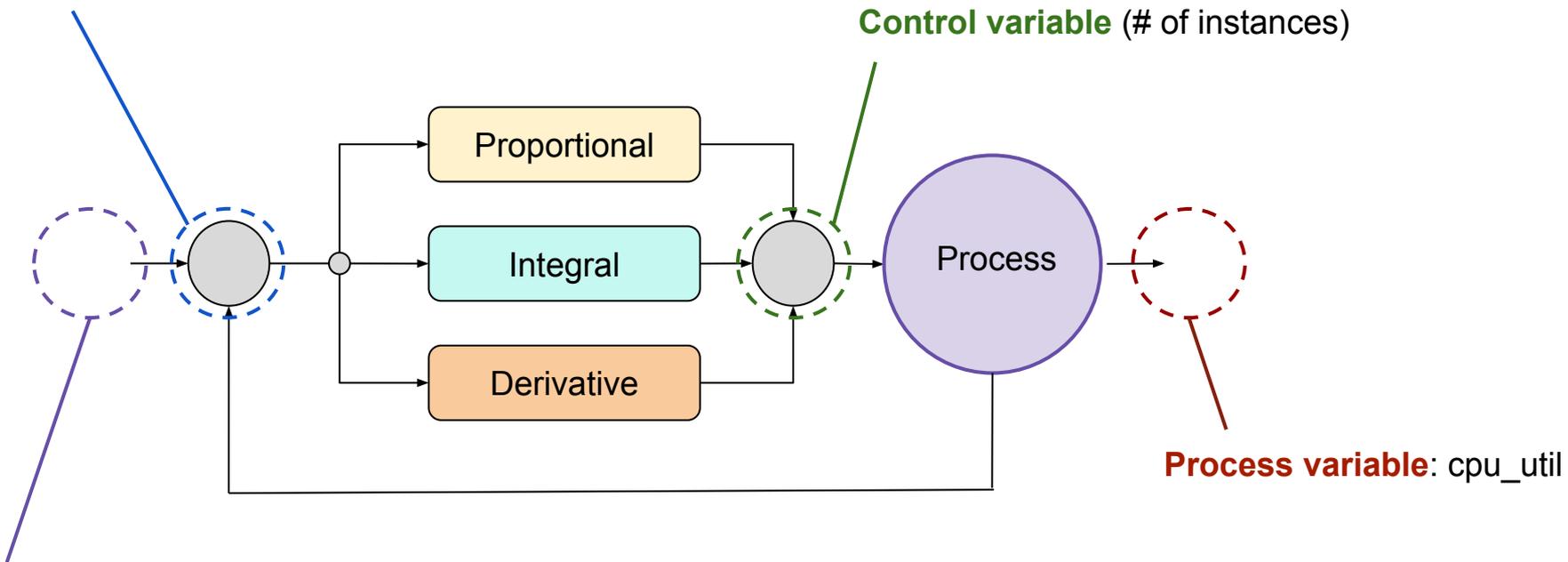
Setpoint: maintain 80% cpu_util

Error: Difference between what we want and what we observe



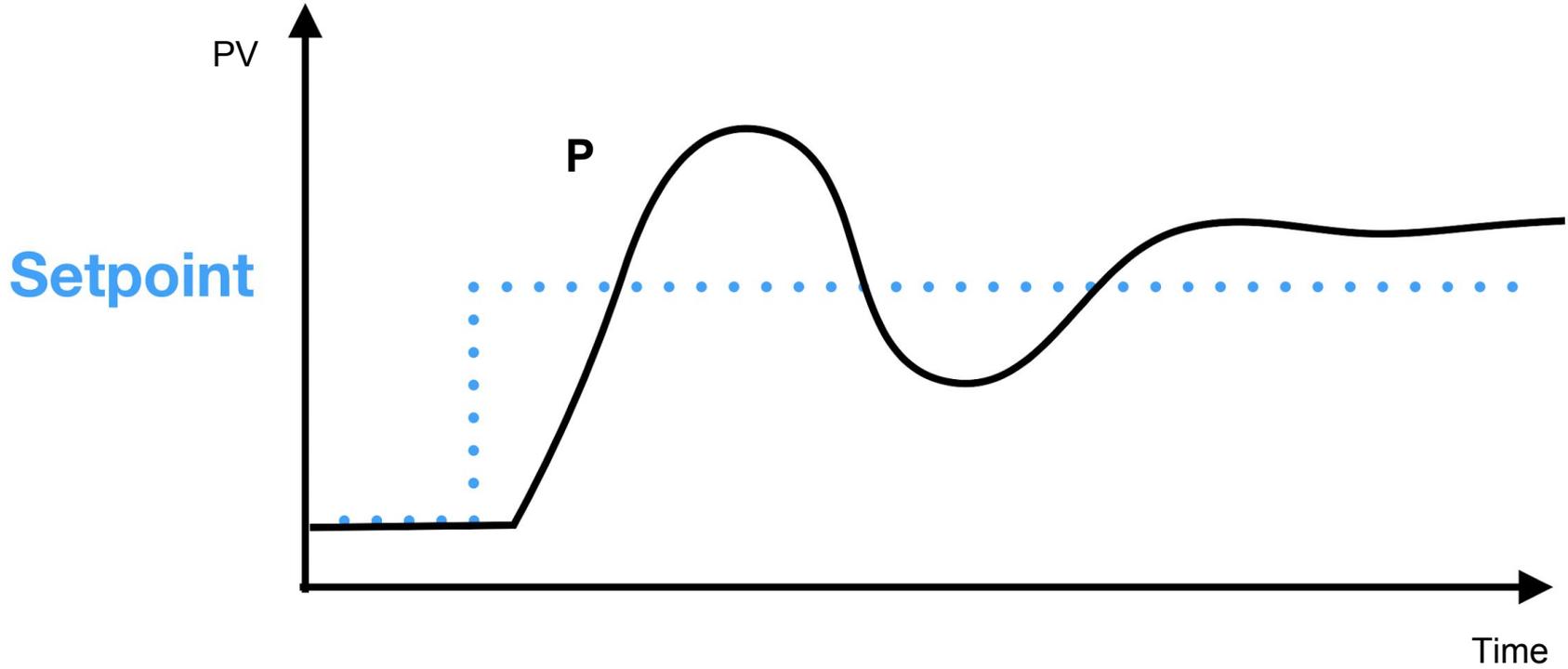
Setpoint: maintain 80% cpu_util

Error: Difference between what we want and what we observe

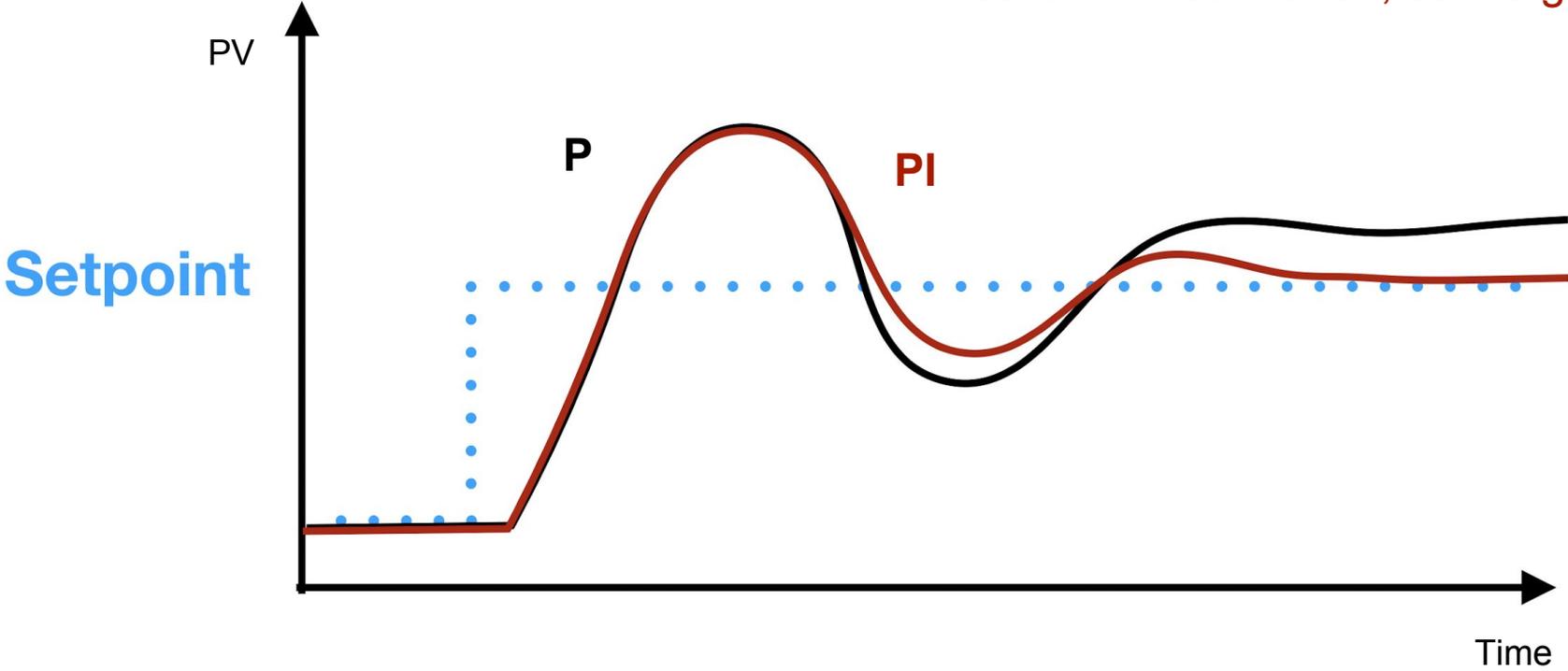


Setpoint: maintain 80% cpu_util

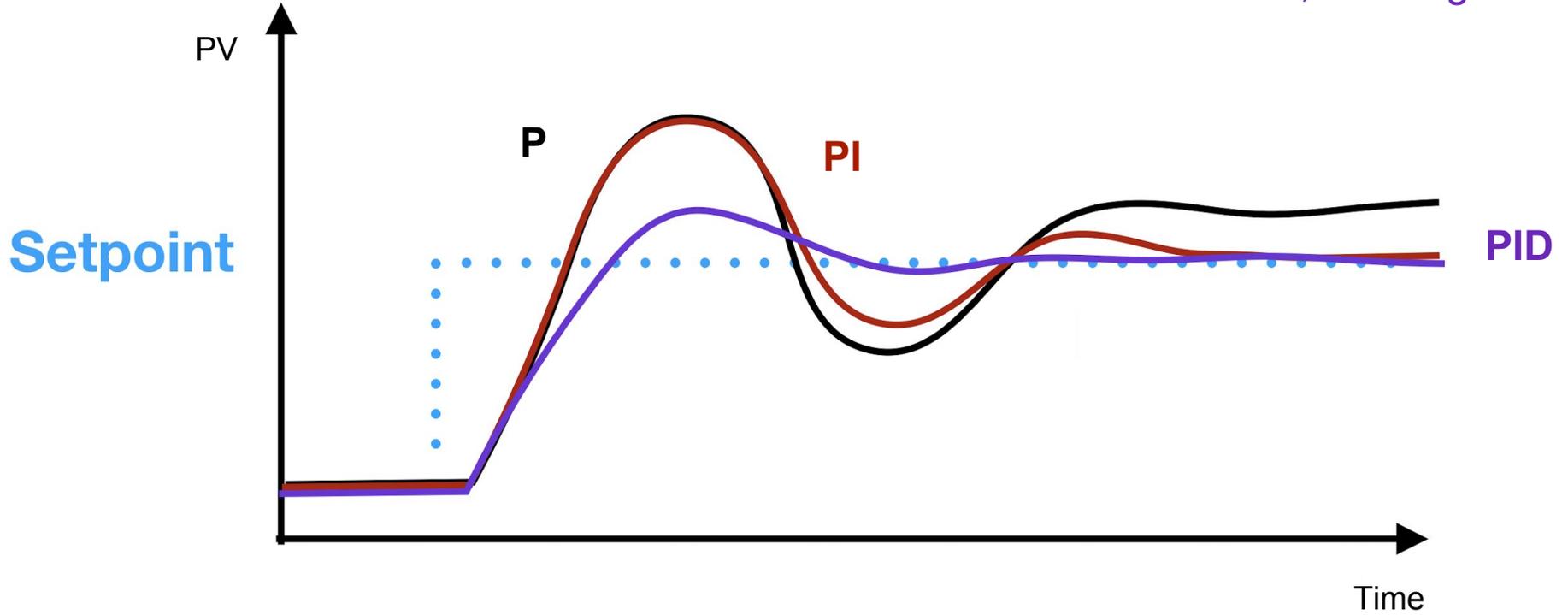
P controller: oscillation, no convergence



PI controller: oscillation, convergence



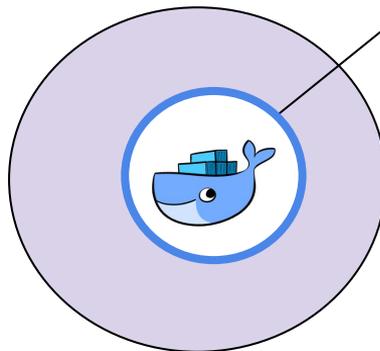
PID controller: minimal oscillation, convergence



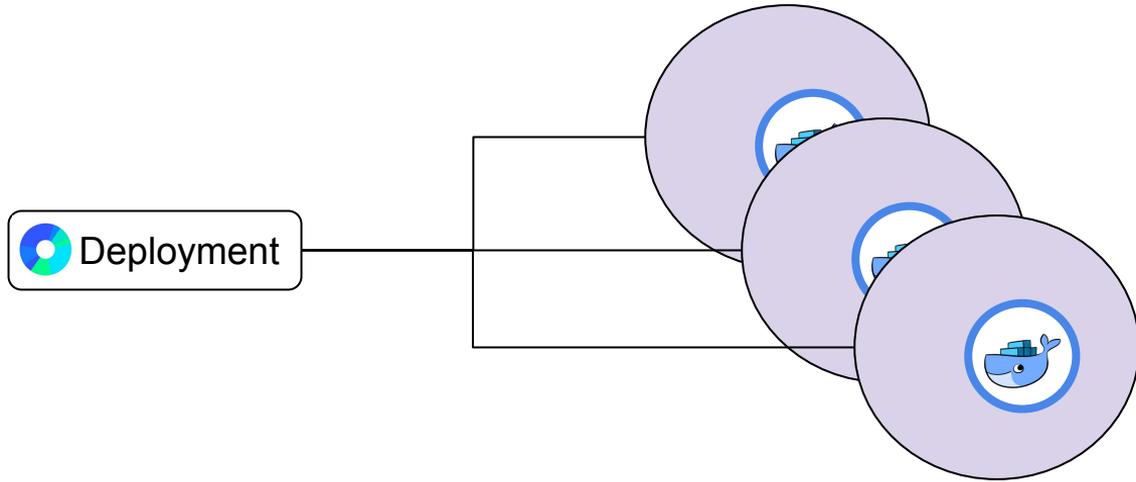
How Strava uses PID for autoscaling



Mesos task



Application
container

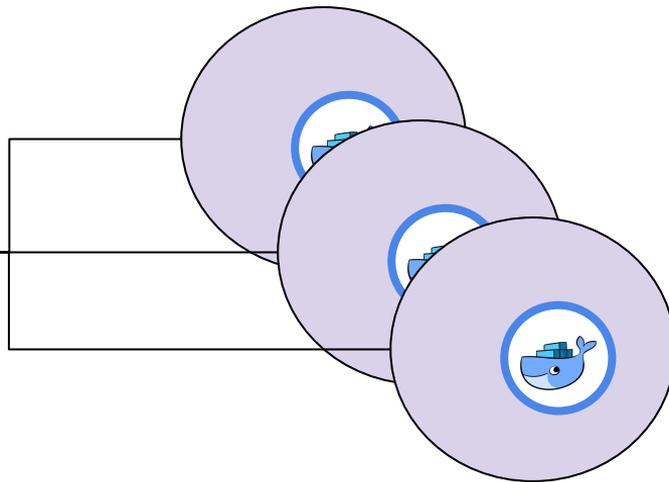




PID
Controller
service



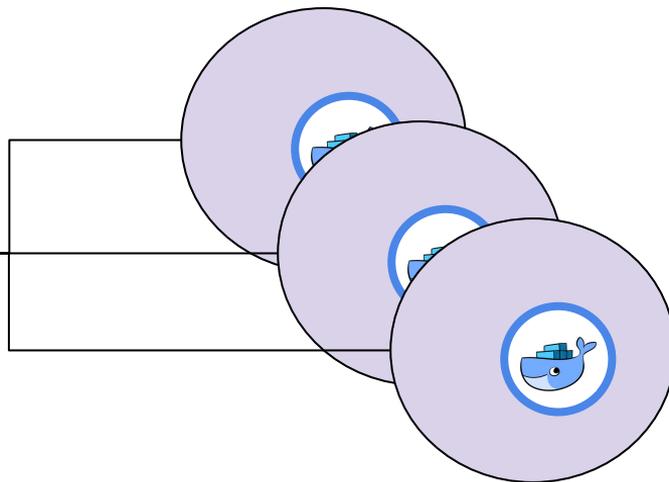
Deployment



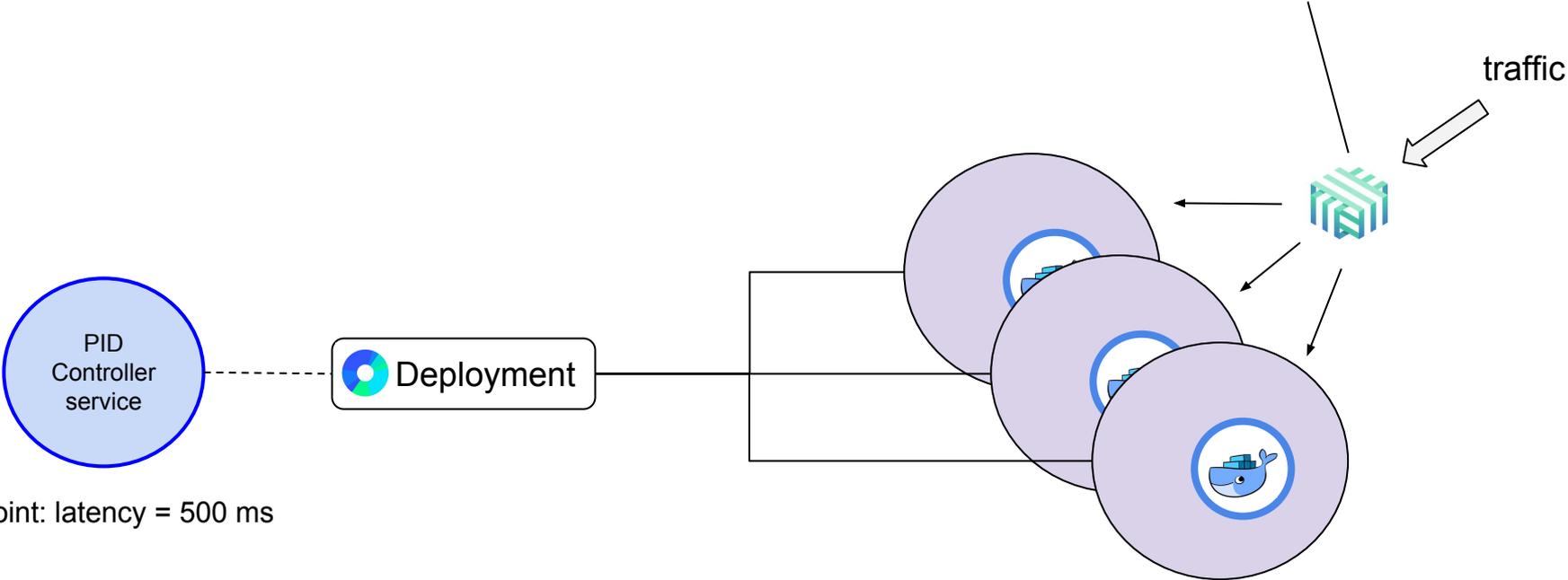
Setpoint: latency = 500 ms



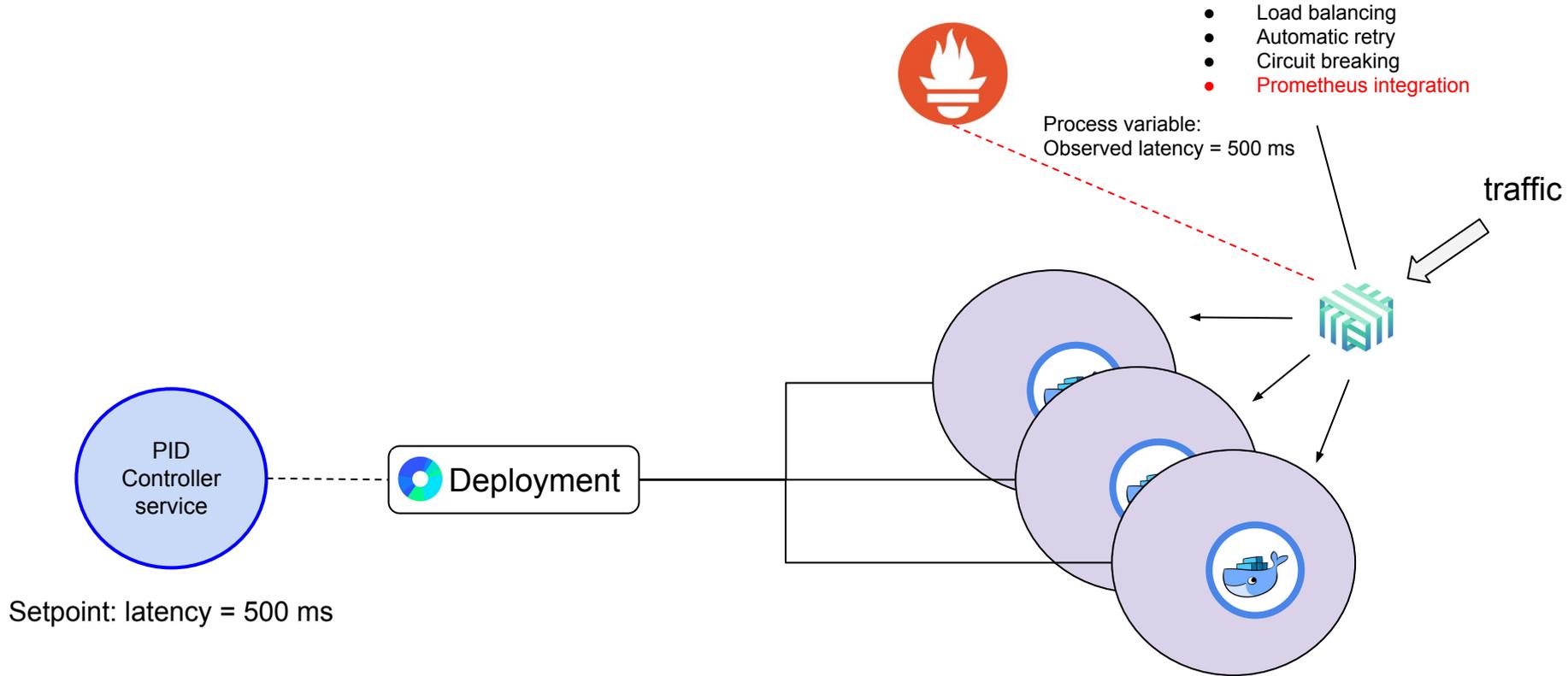
Setpoint: latency = 500 ms

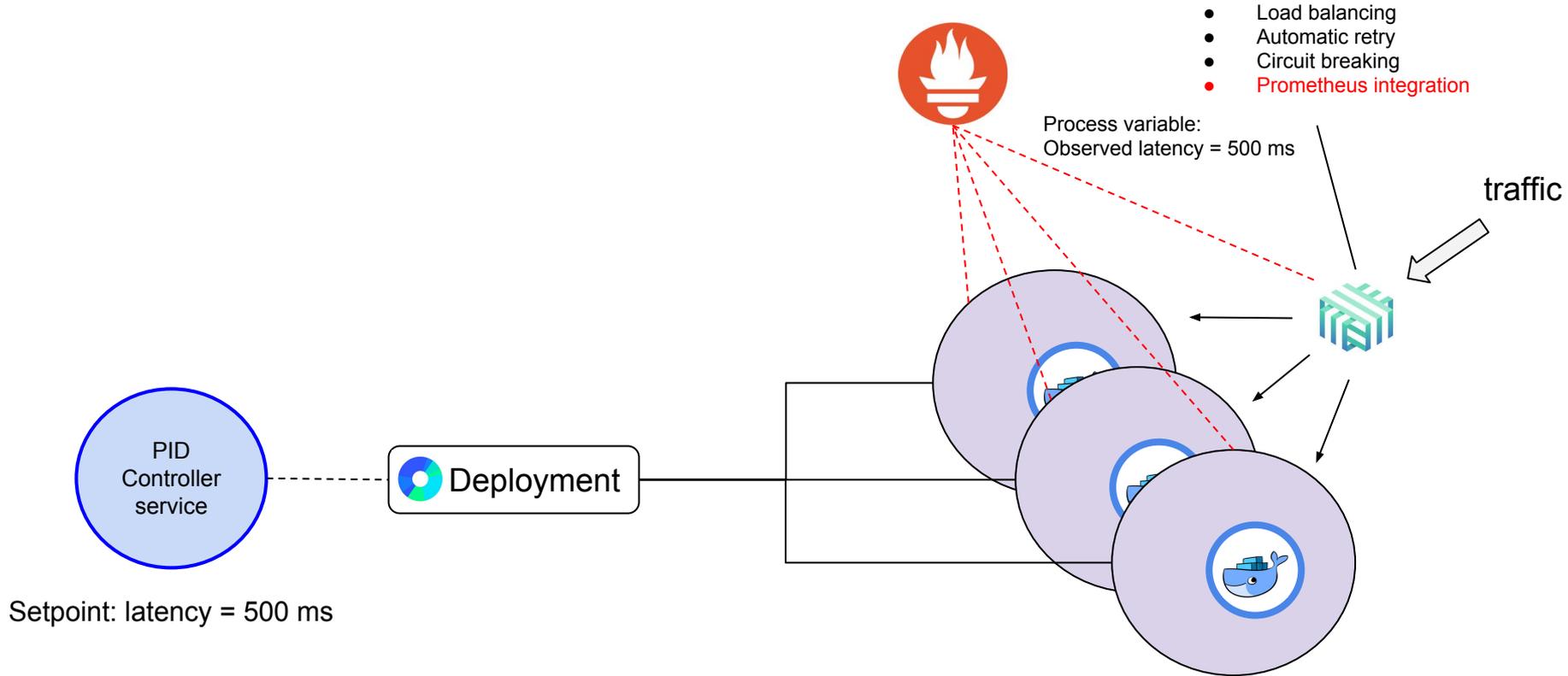


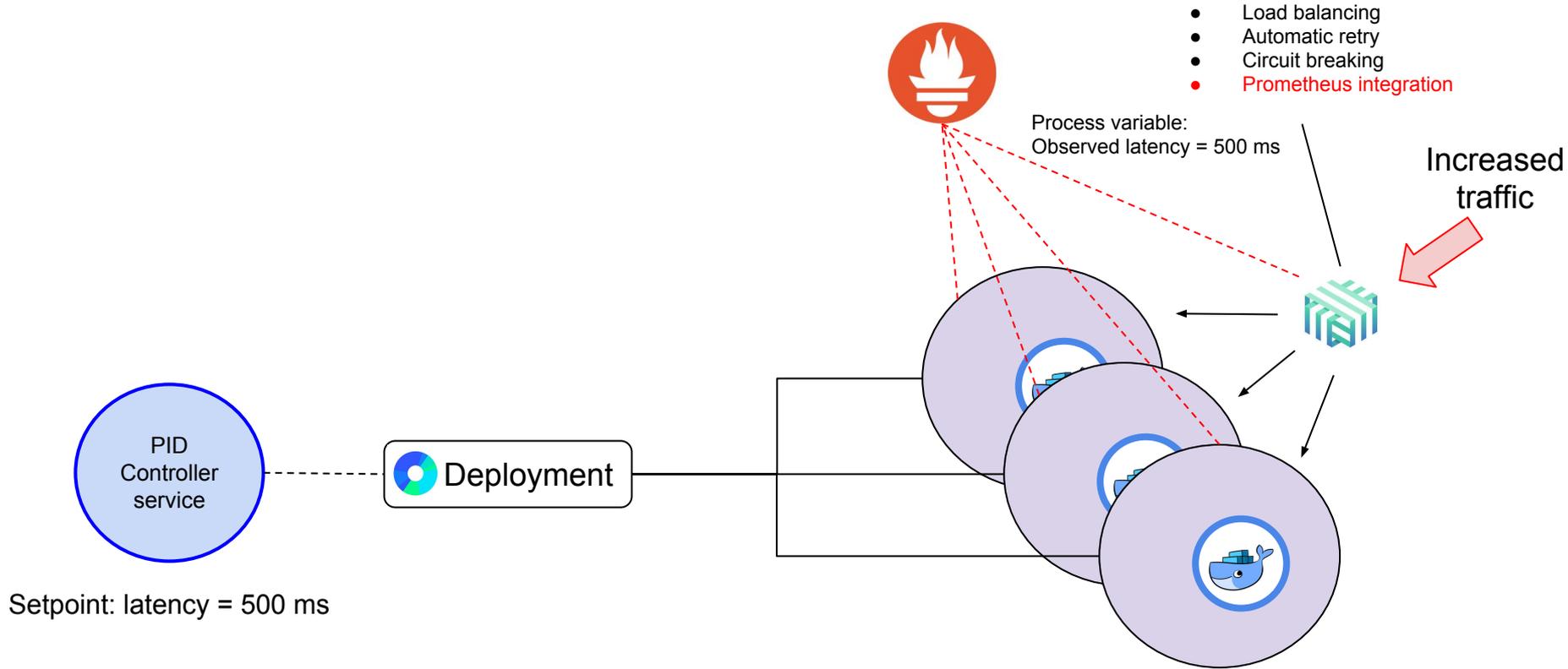
- Load balancing
- Automatic retry
- Circuit breaking

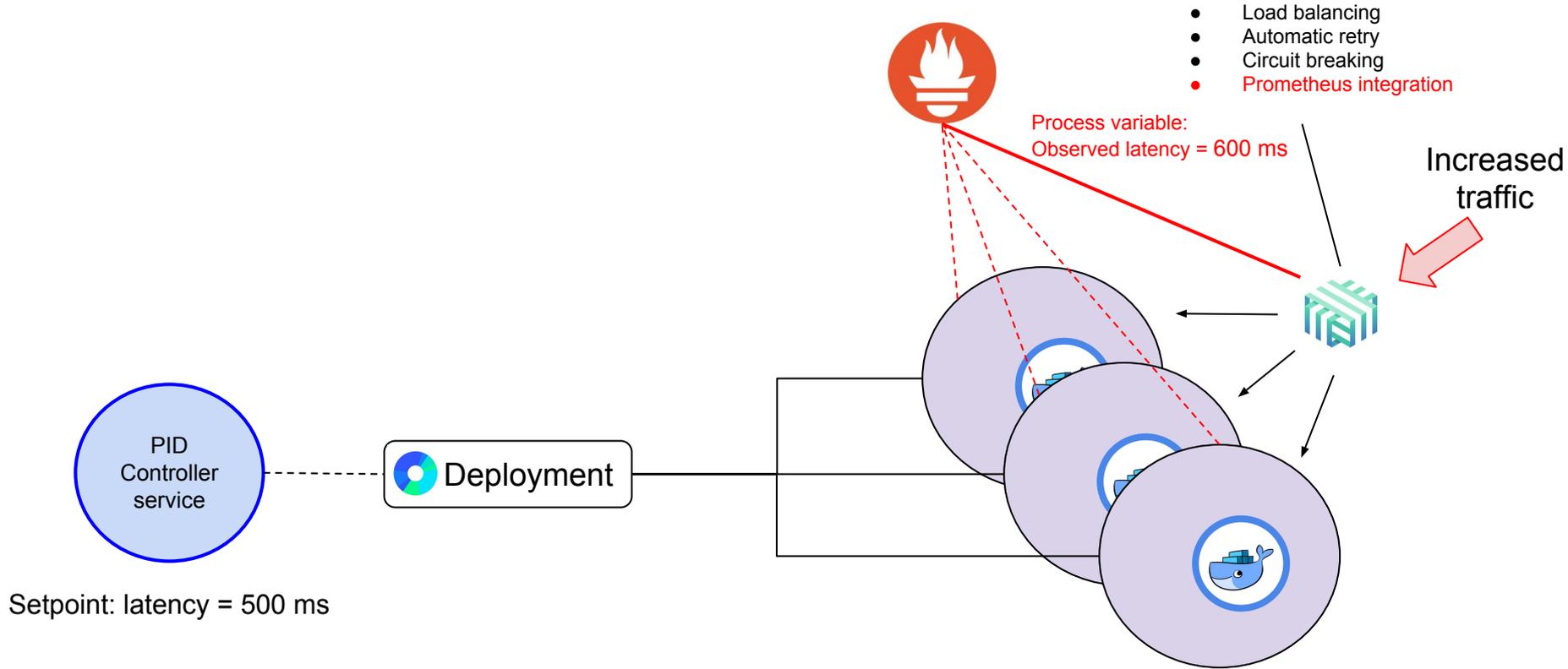


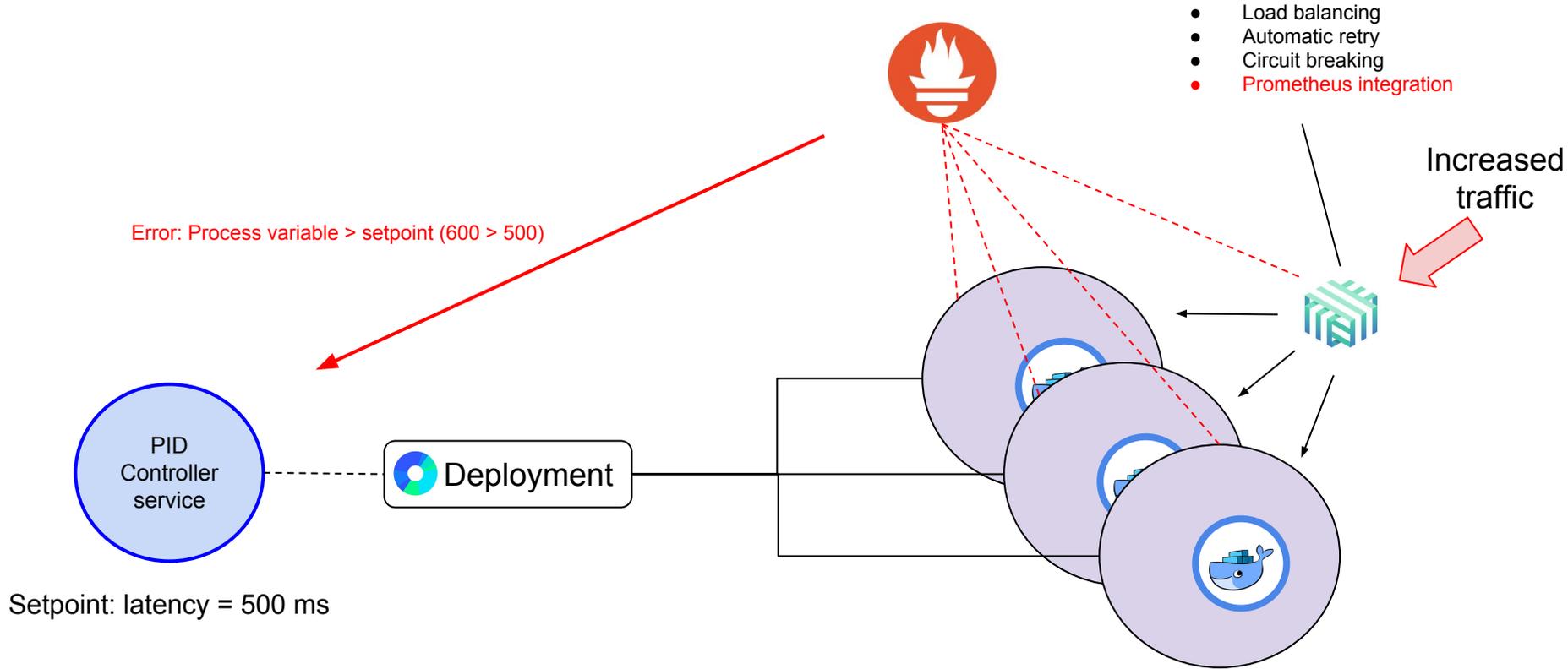
Setpoint: latency = 500 ms

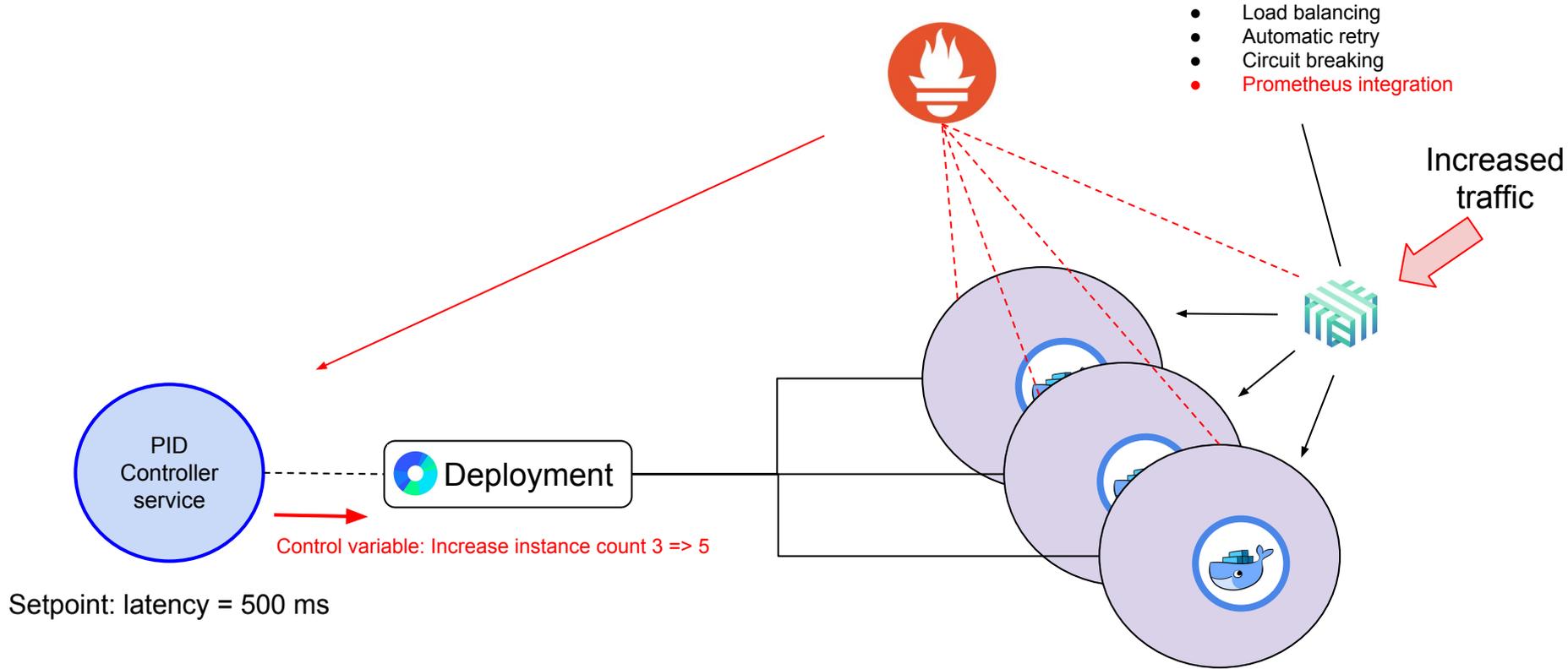


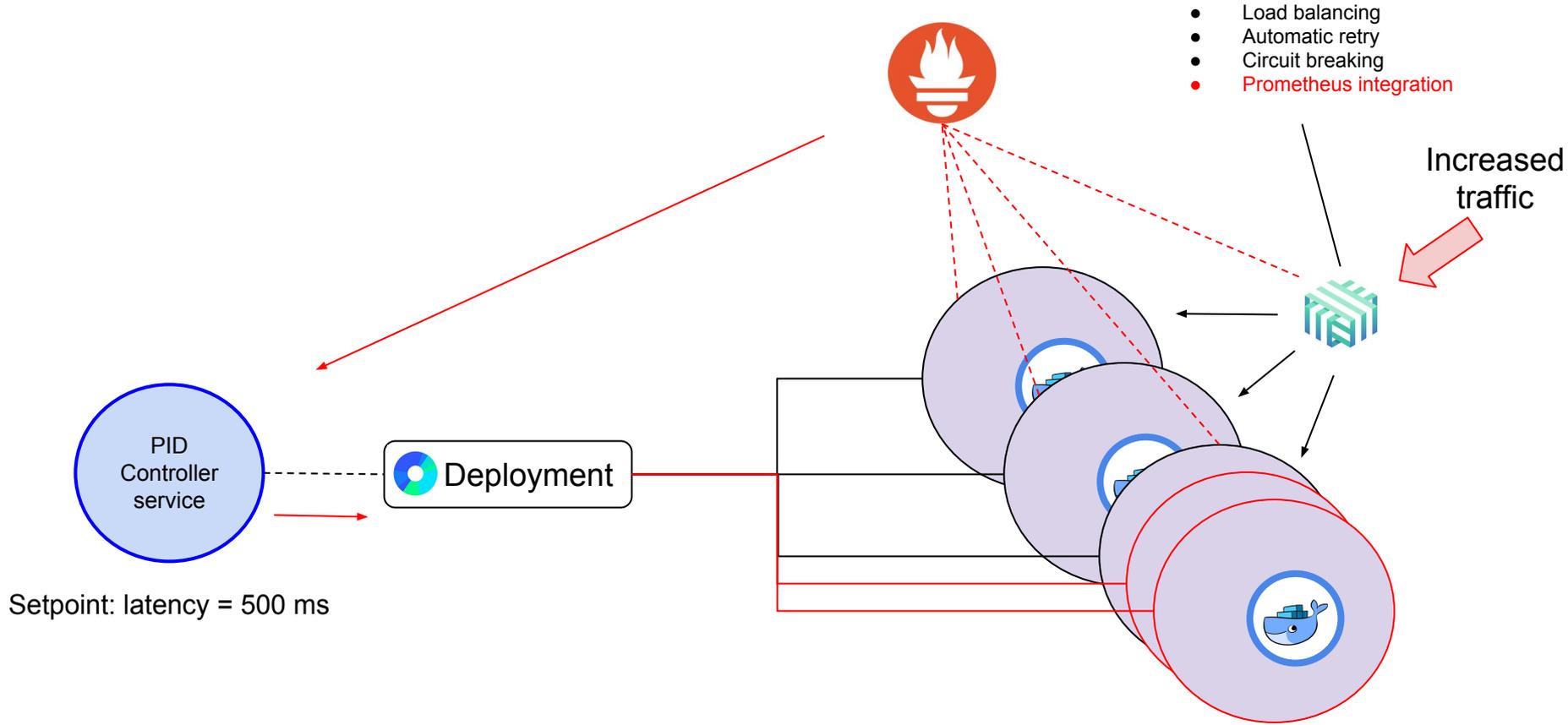


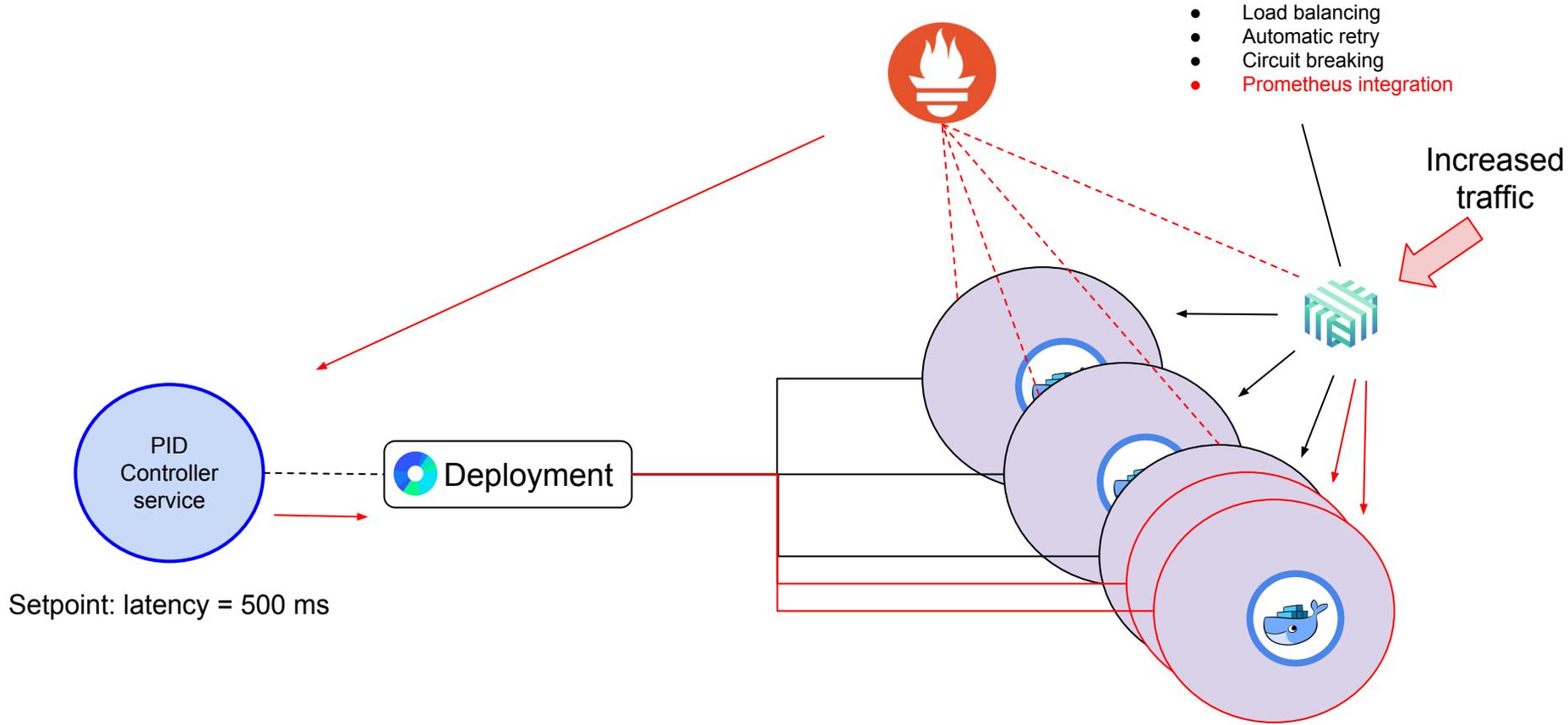






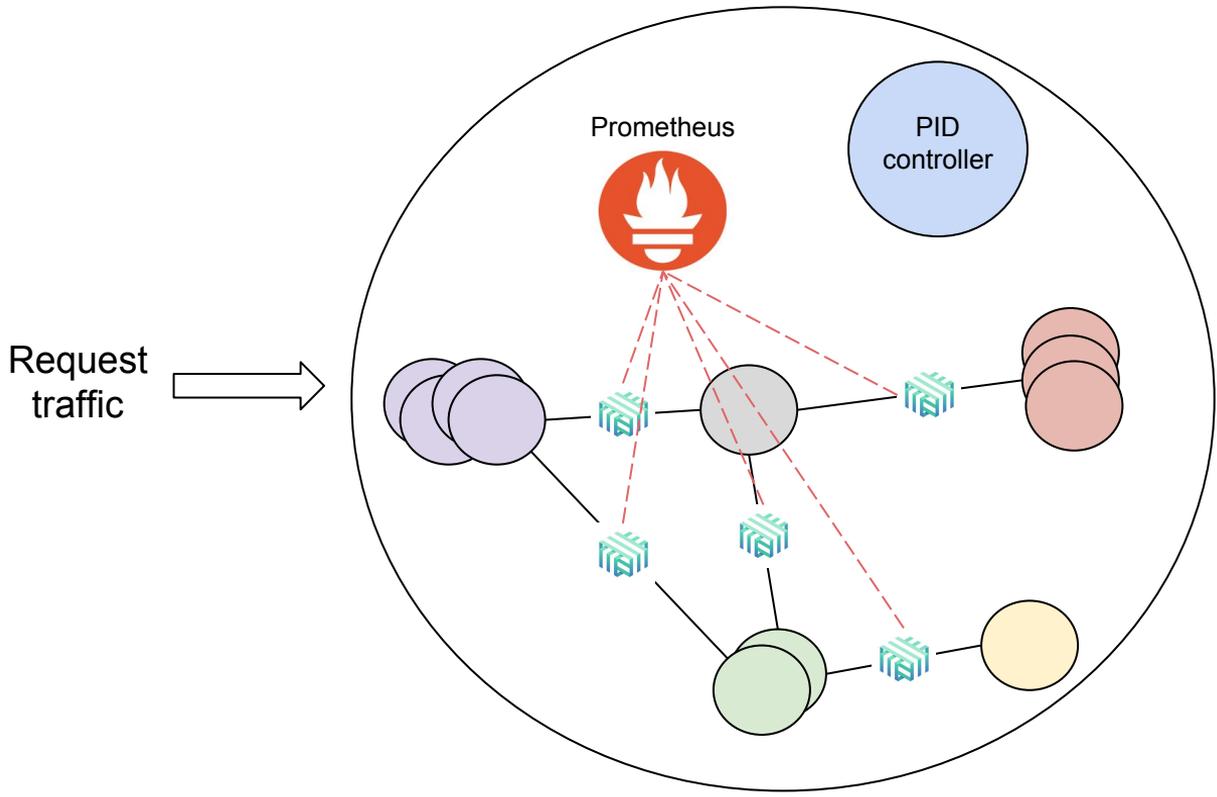






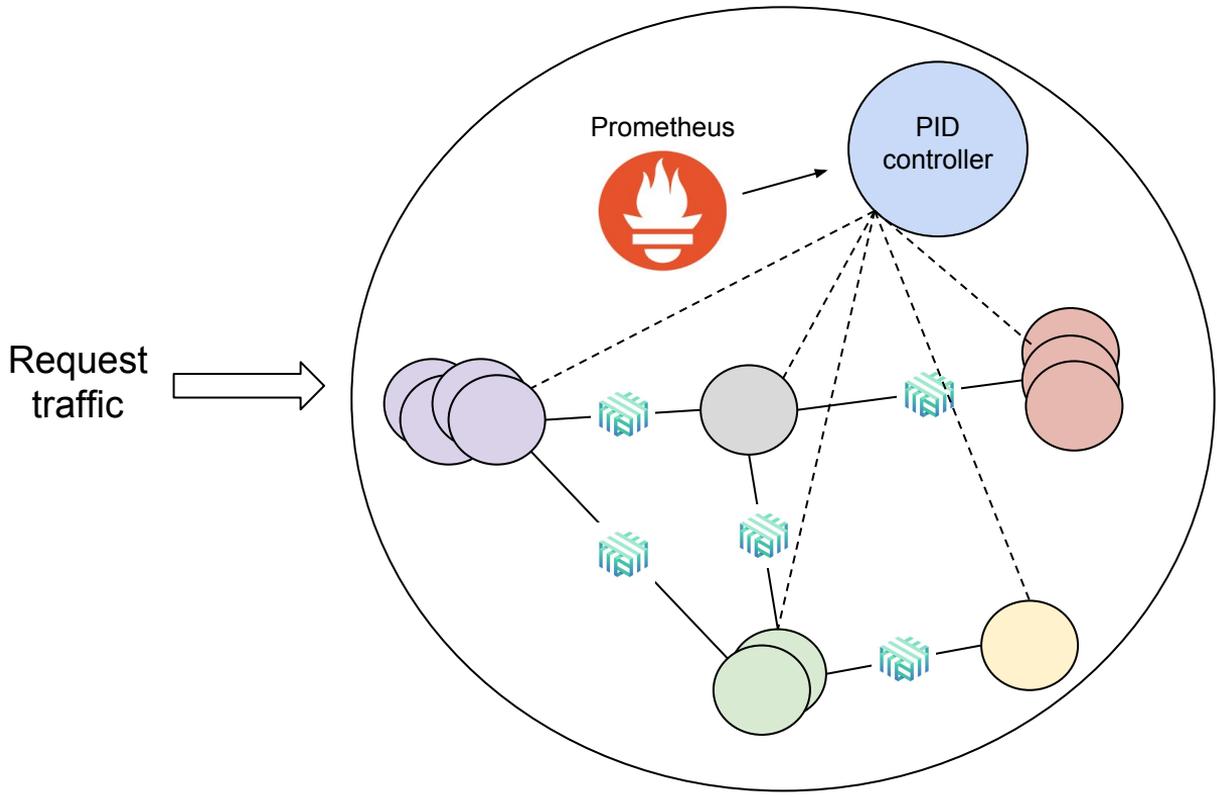


Mesos cluster

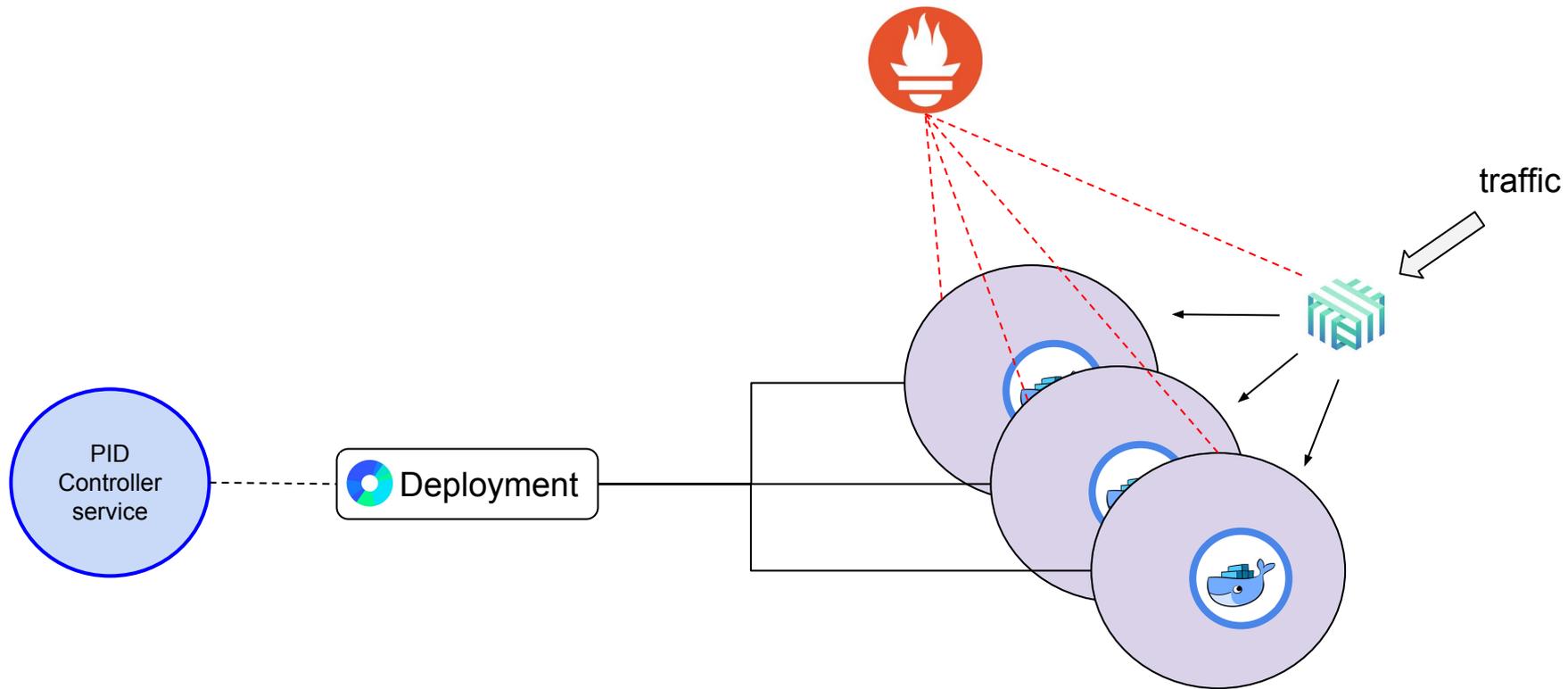


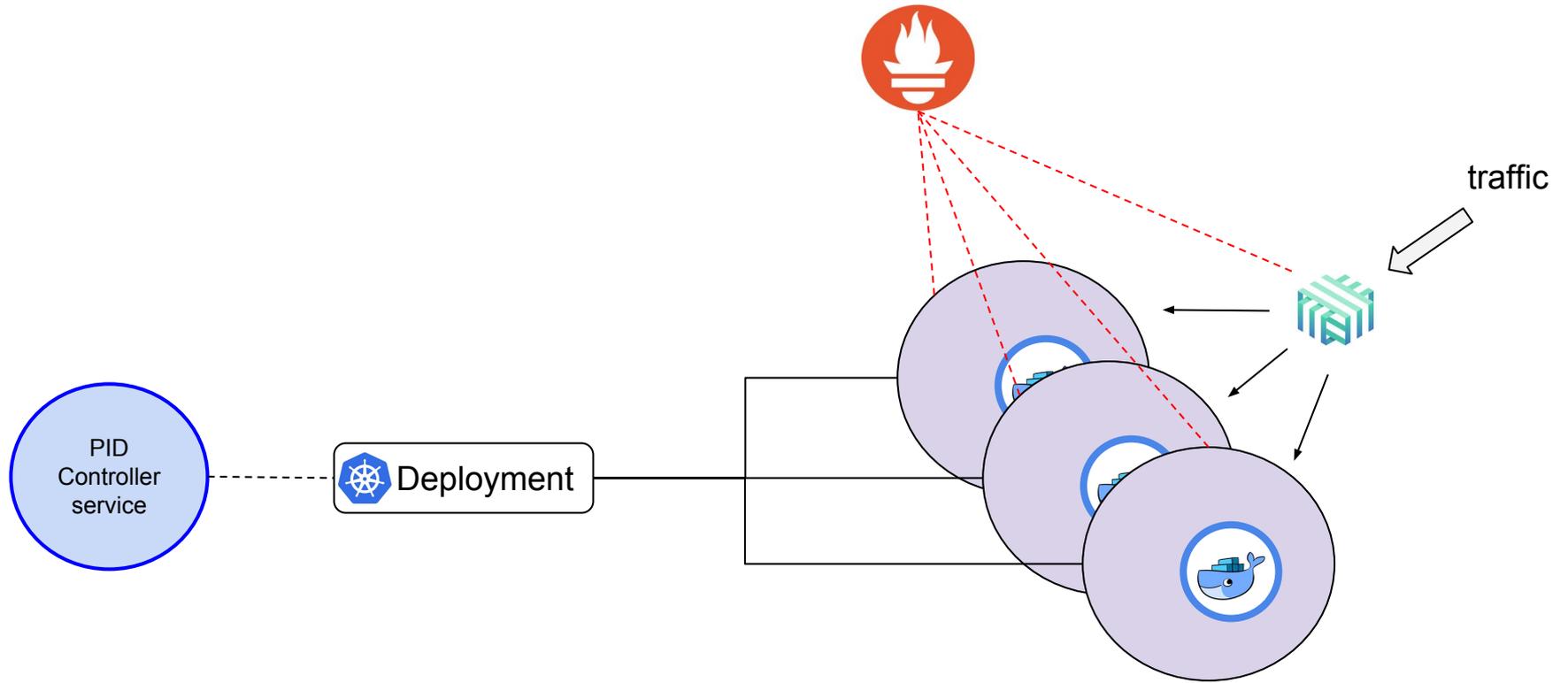


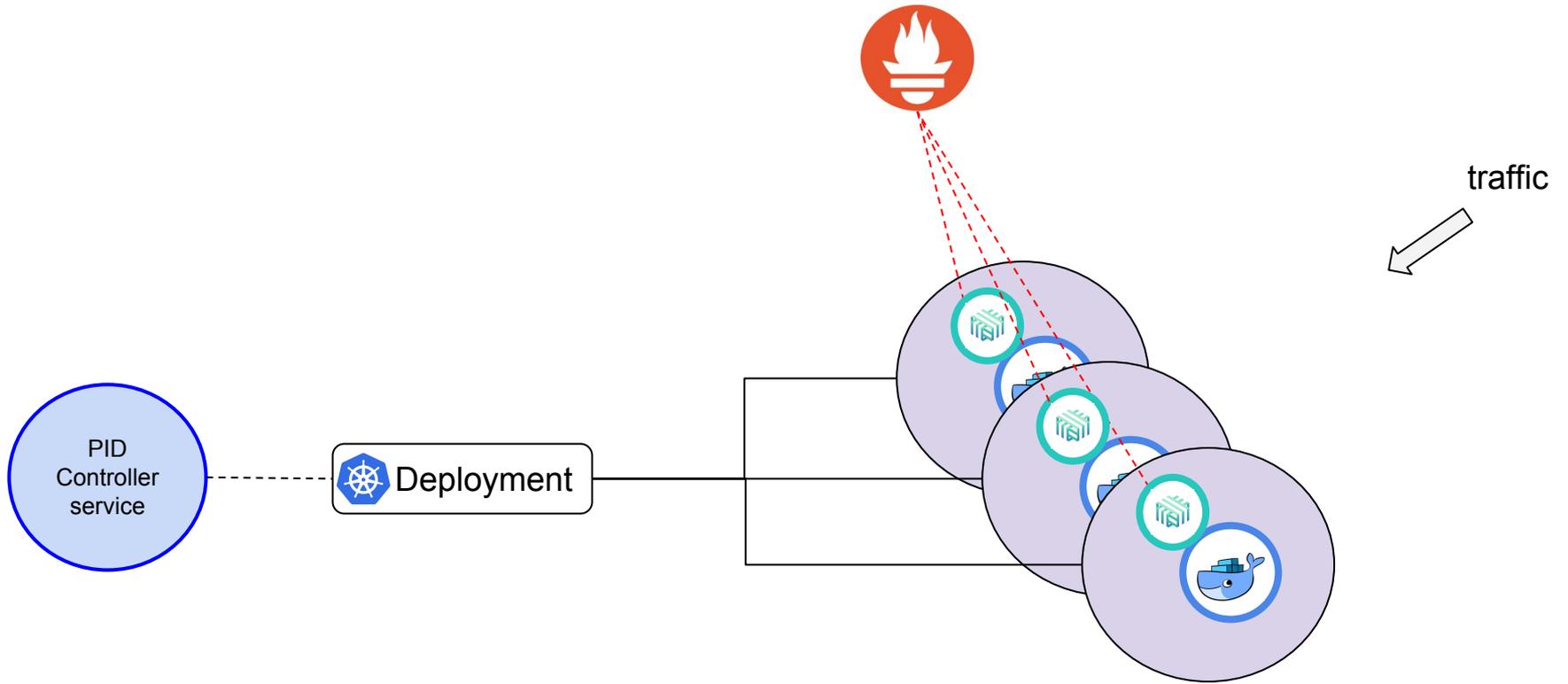
Mesos cluster

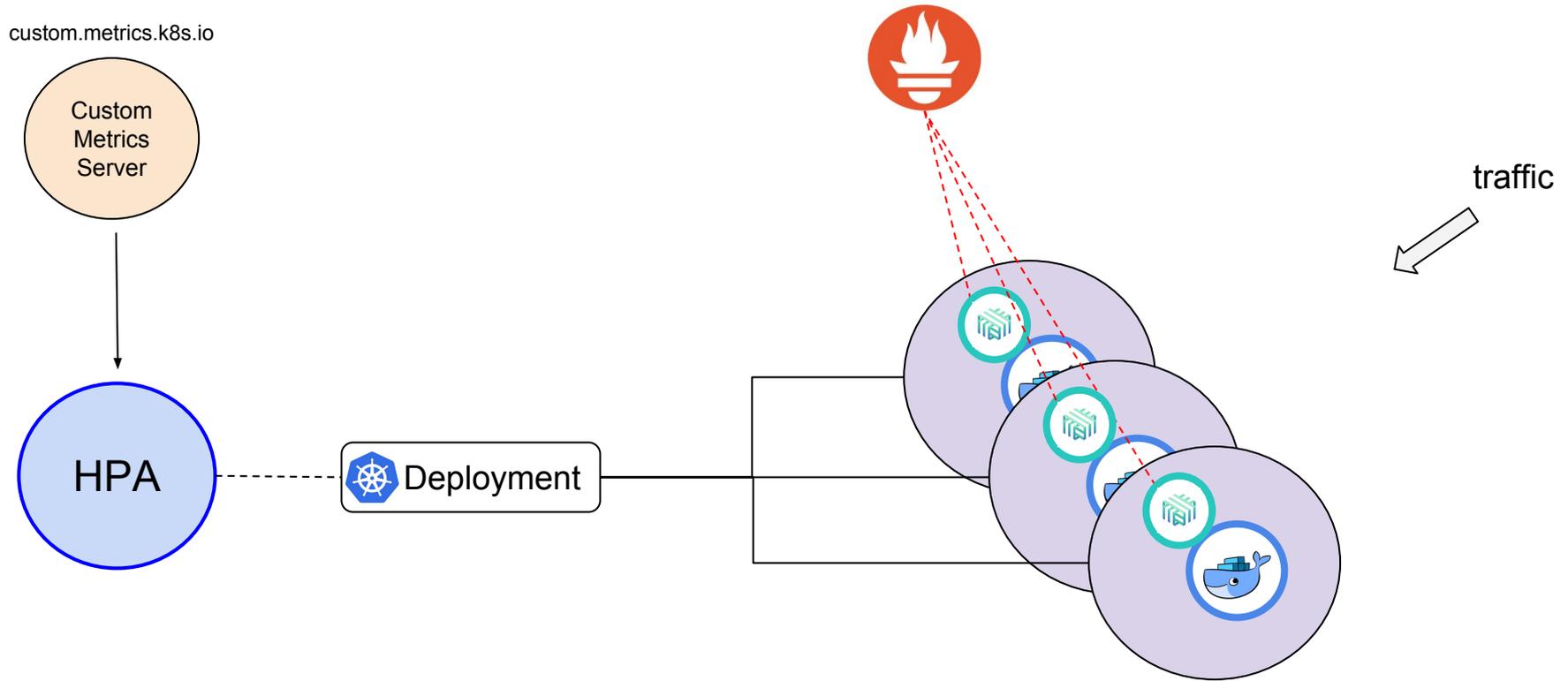


How you can autoscale in kubernetes



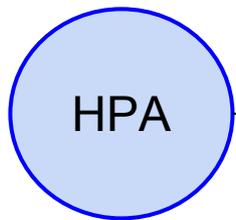
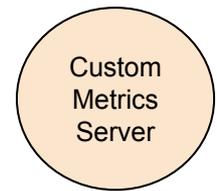






Implementation of custom.metrics.k8s.io API using Prometheus
<https://github.com/DirectXMan12/k8s-prometheus-adapter.git>

custom.metrics.k8s.io



Increased traffic

A large red arrow pointing downwards and to the left, indicating the direction of increased traffic.

