High Availability at Heroku

Mark McGranaghan

heroku

app

routing

runtime

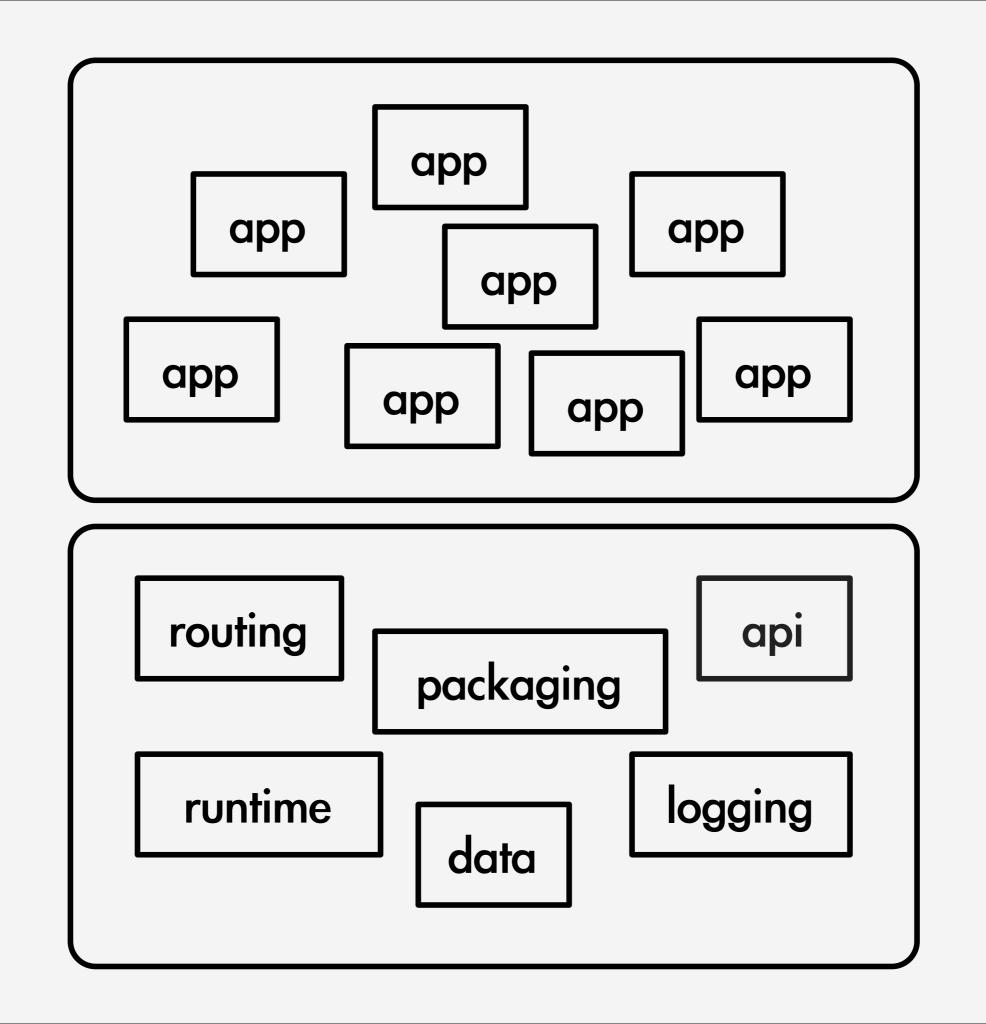
data

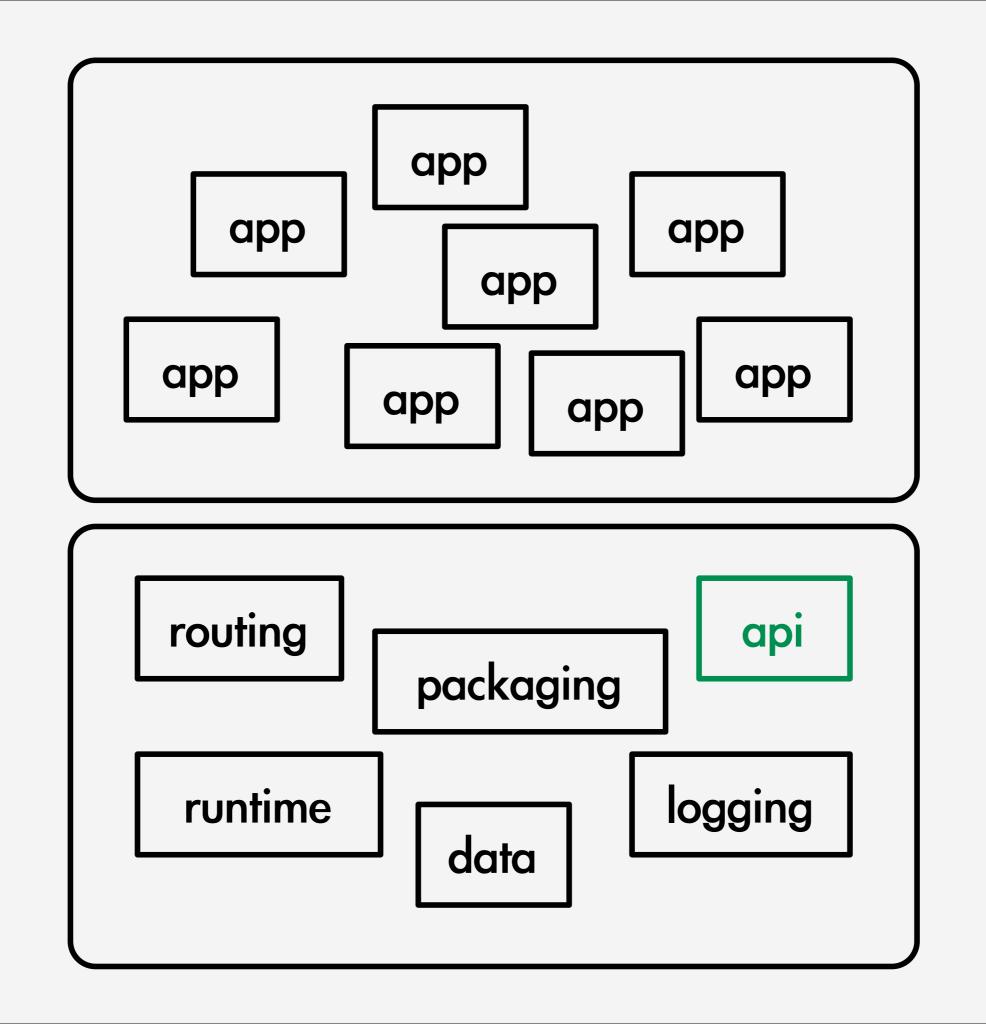
routing runtime data

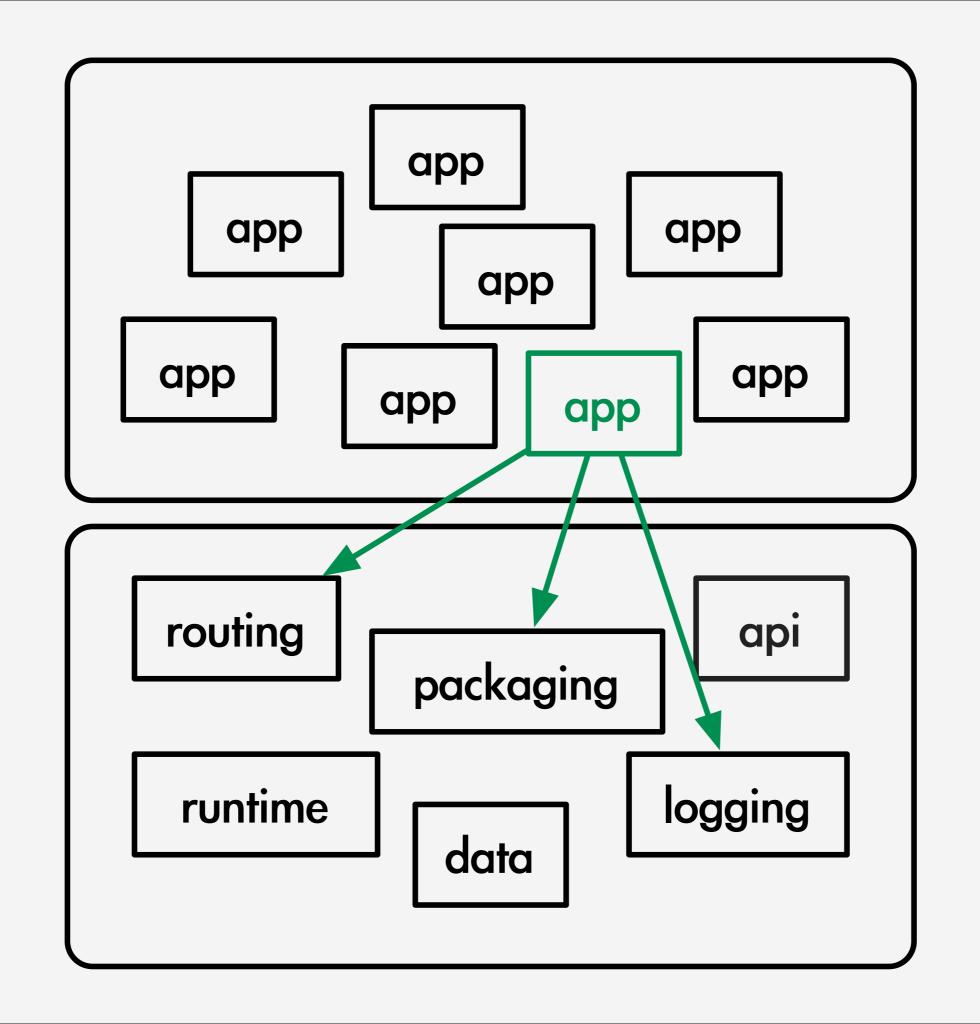
routing
runtime
data

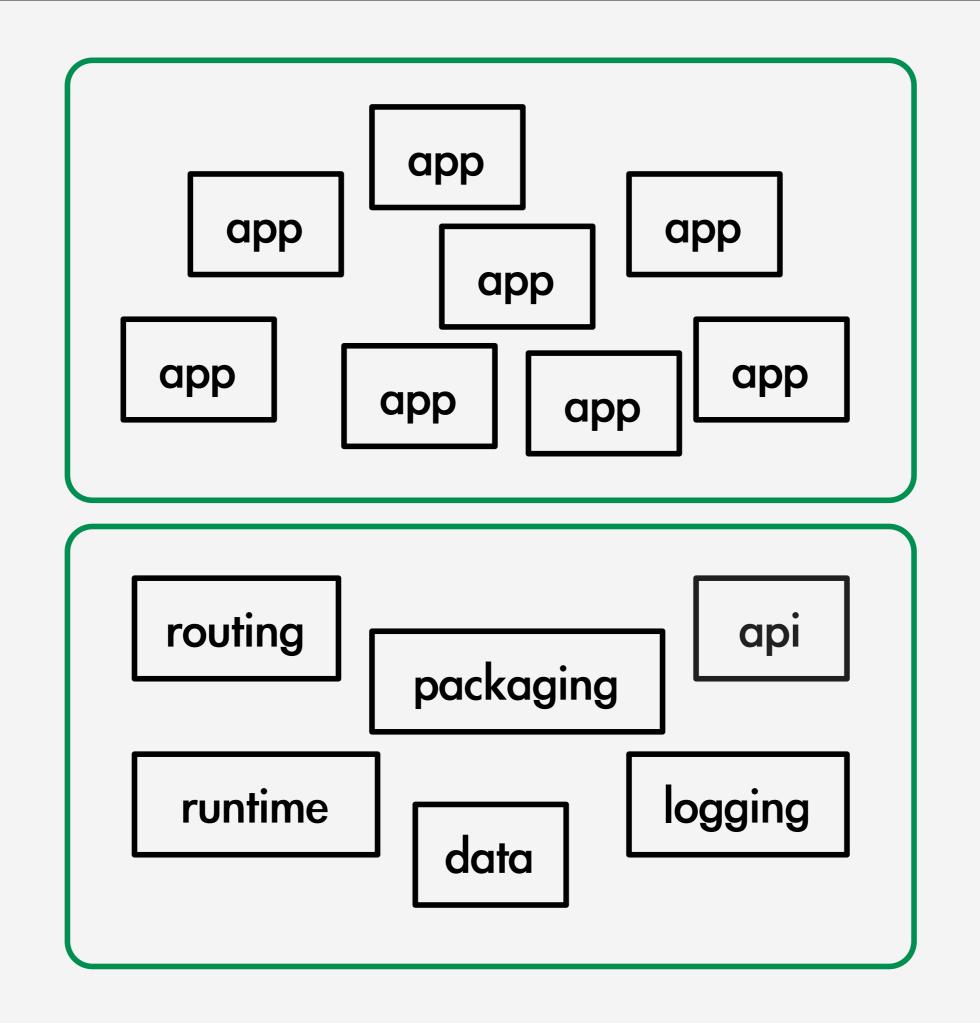
runtime

routing
runtime
data









Scale & Scope

O(1,000) instances

O(1,000,000) apps

Success & Failure

Architecture Execution

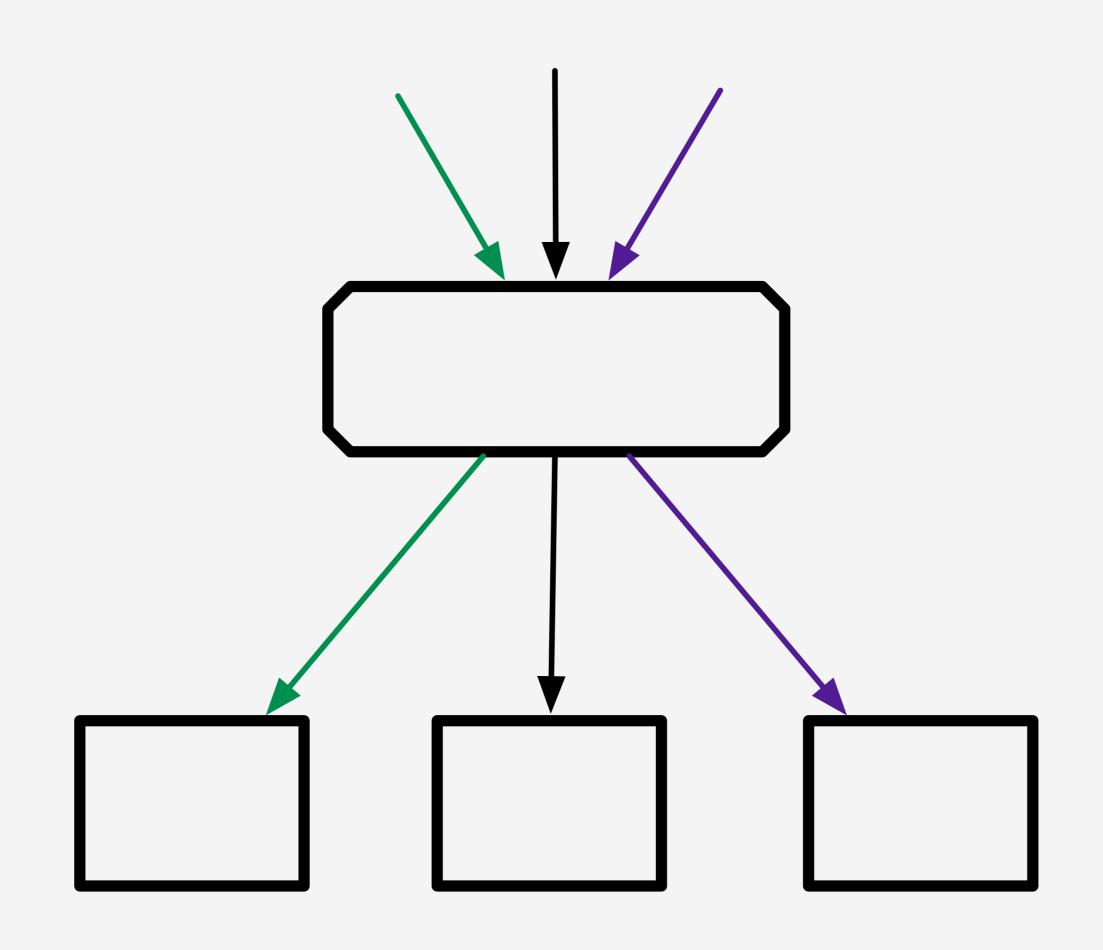
Architecture Execution

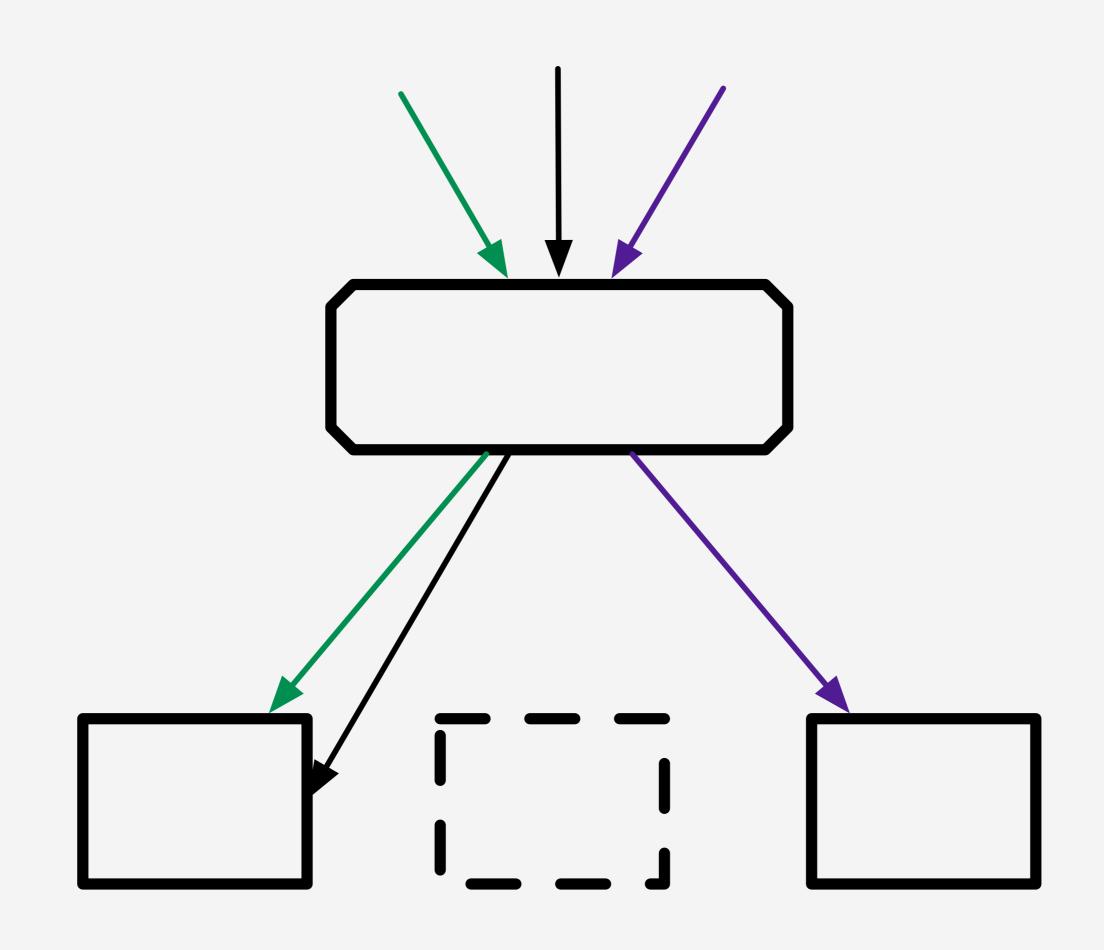
routing

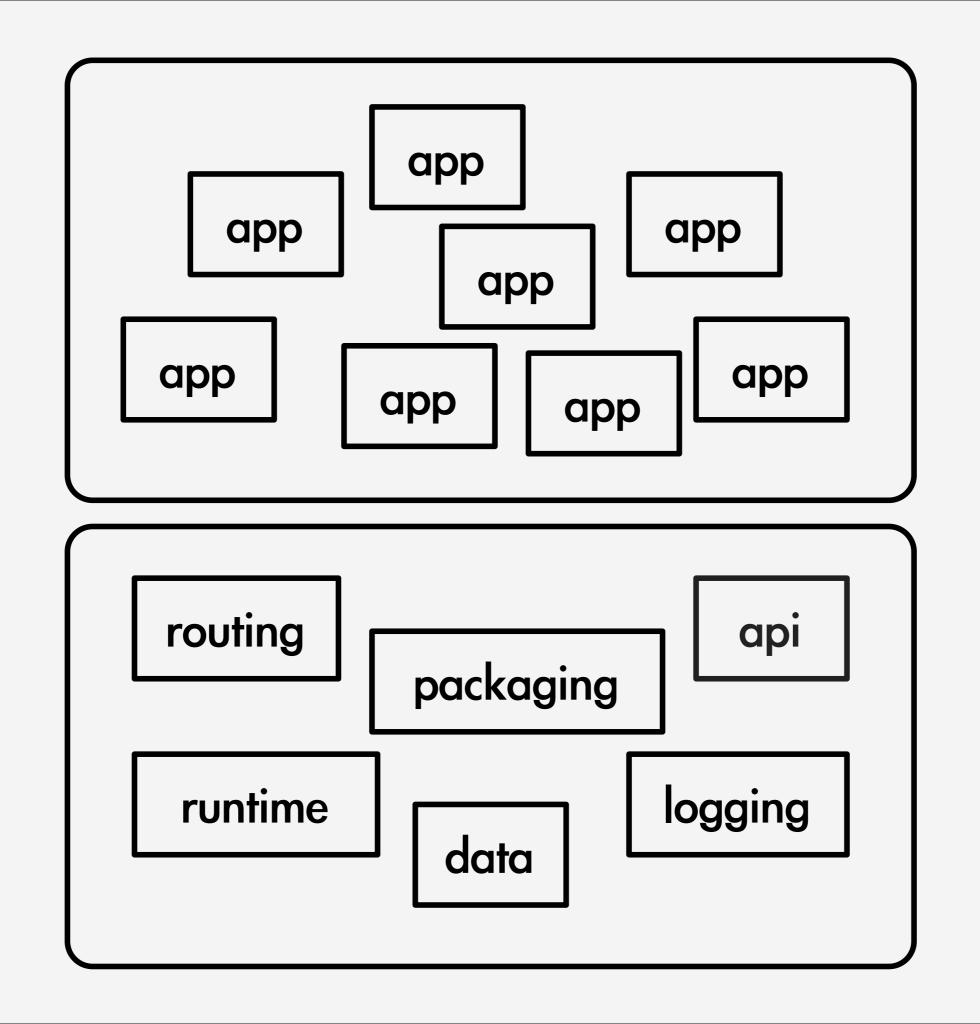
runtime

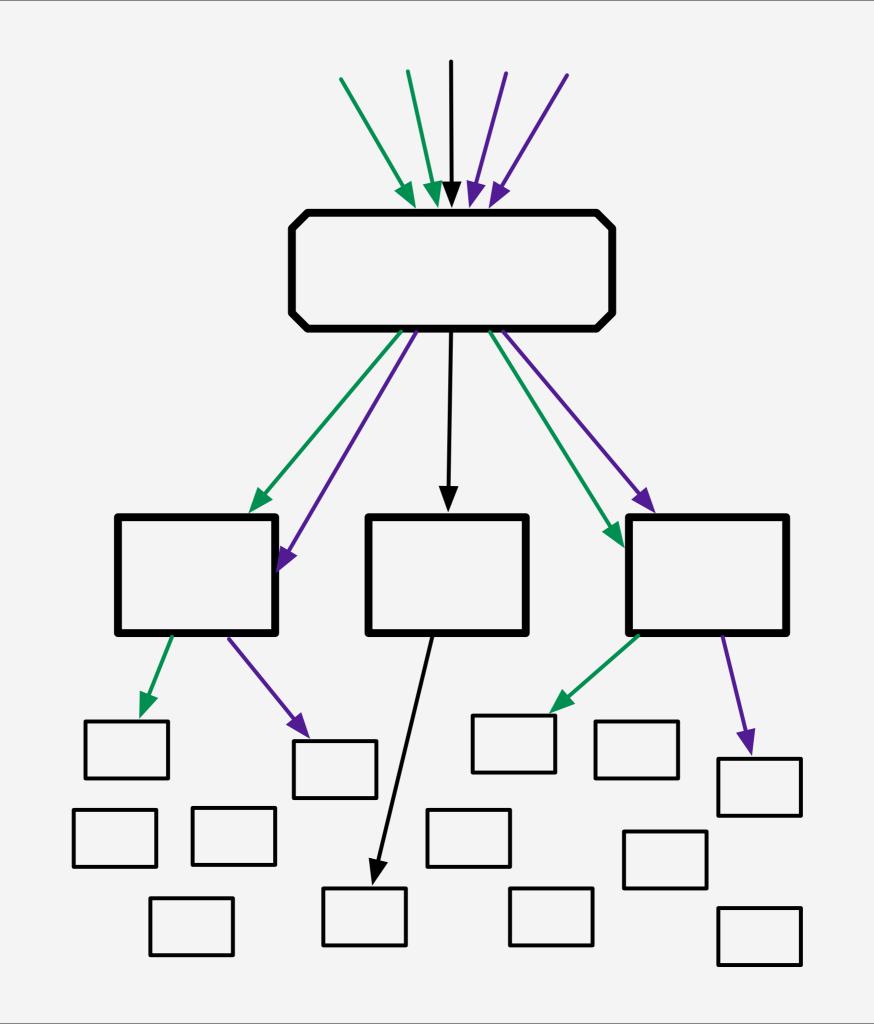
data

Platform-Enabled HA Routing

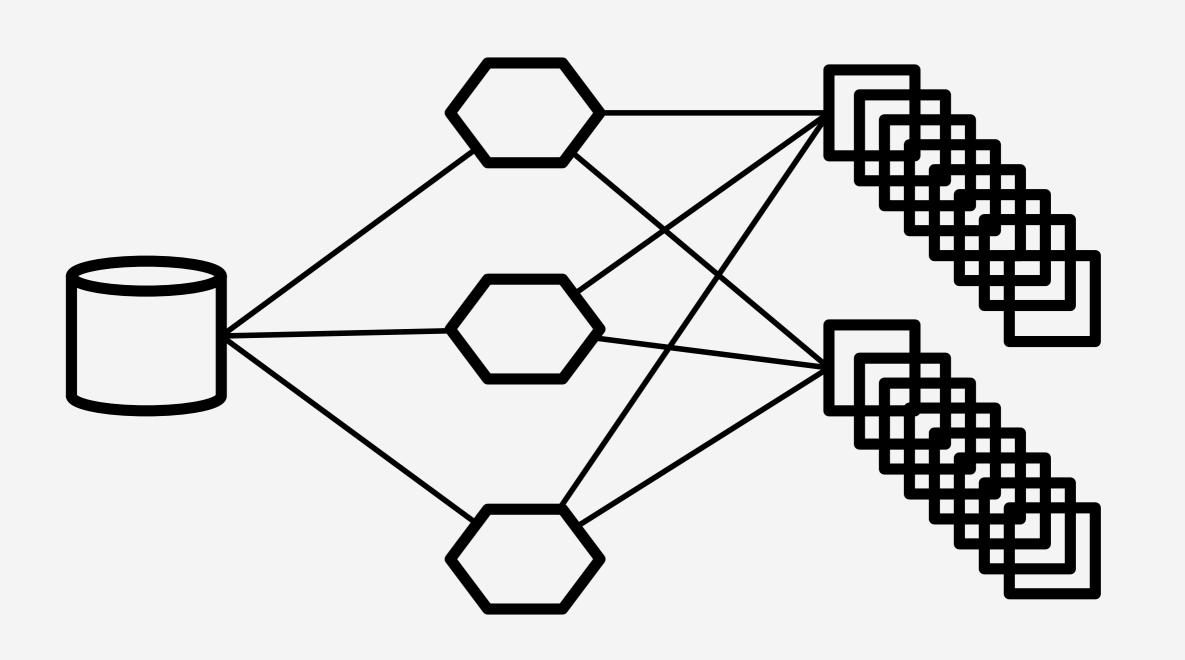


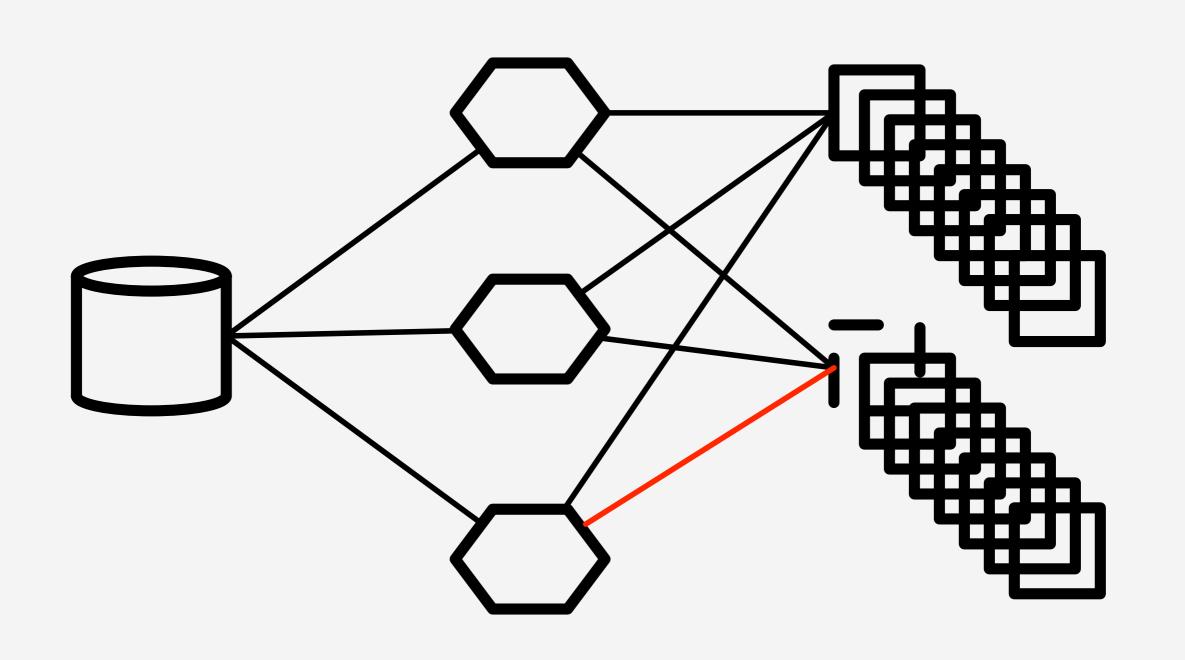


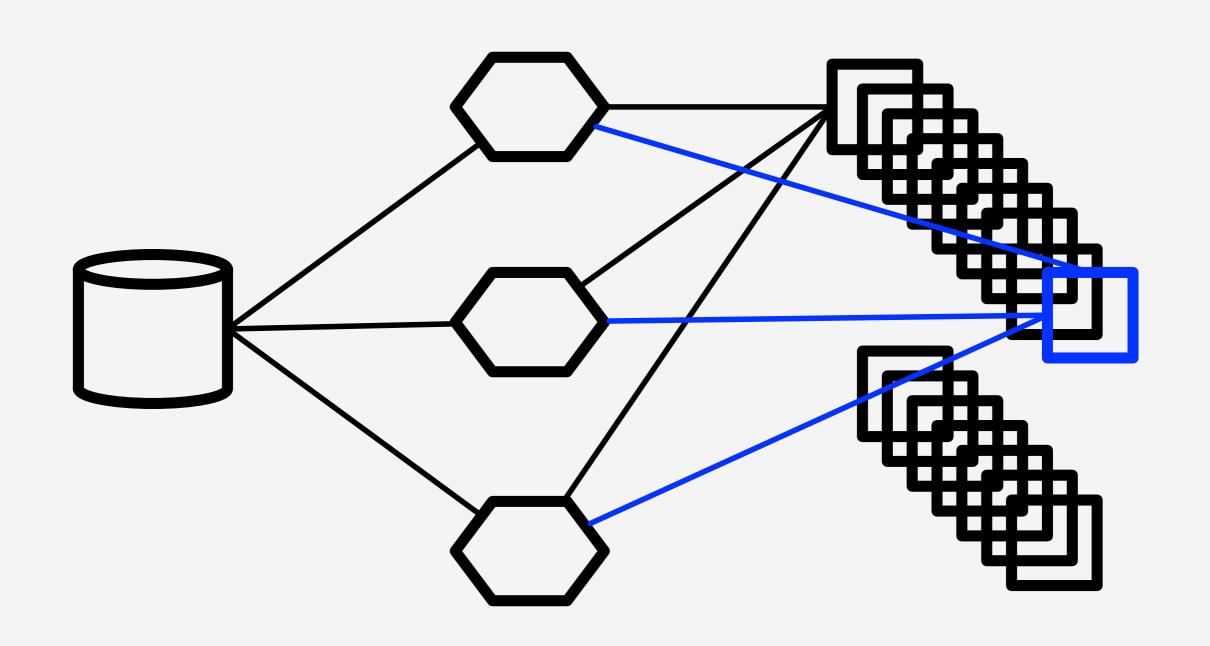




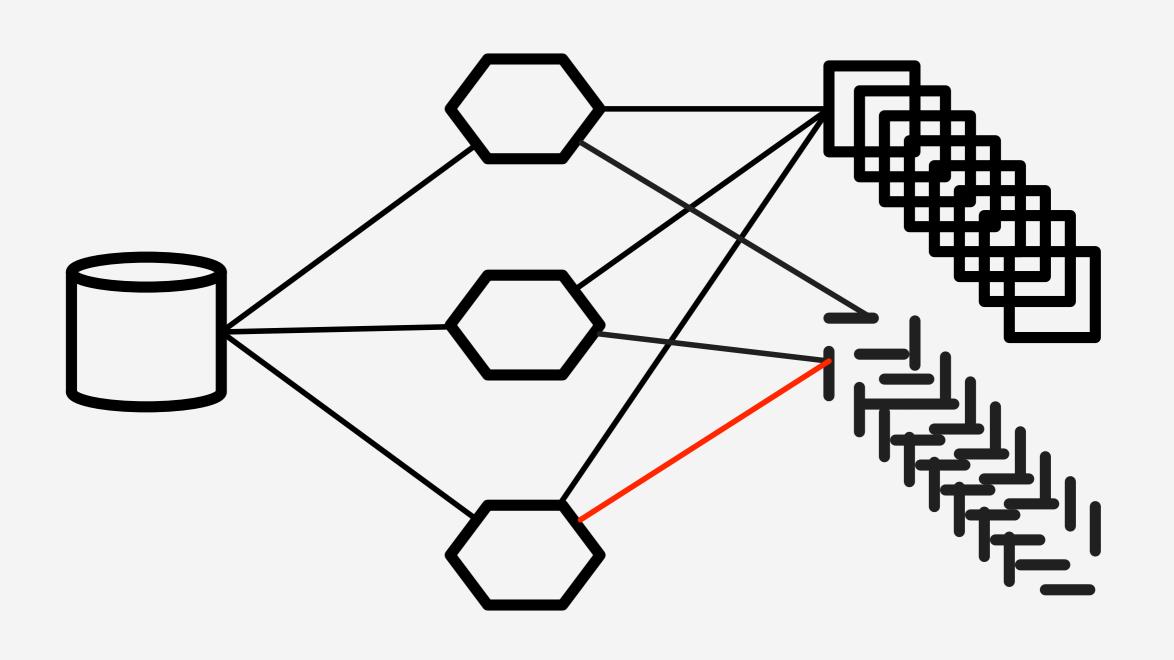
Crashes Supervision

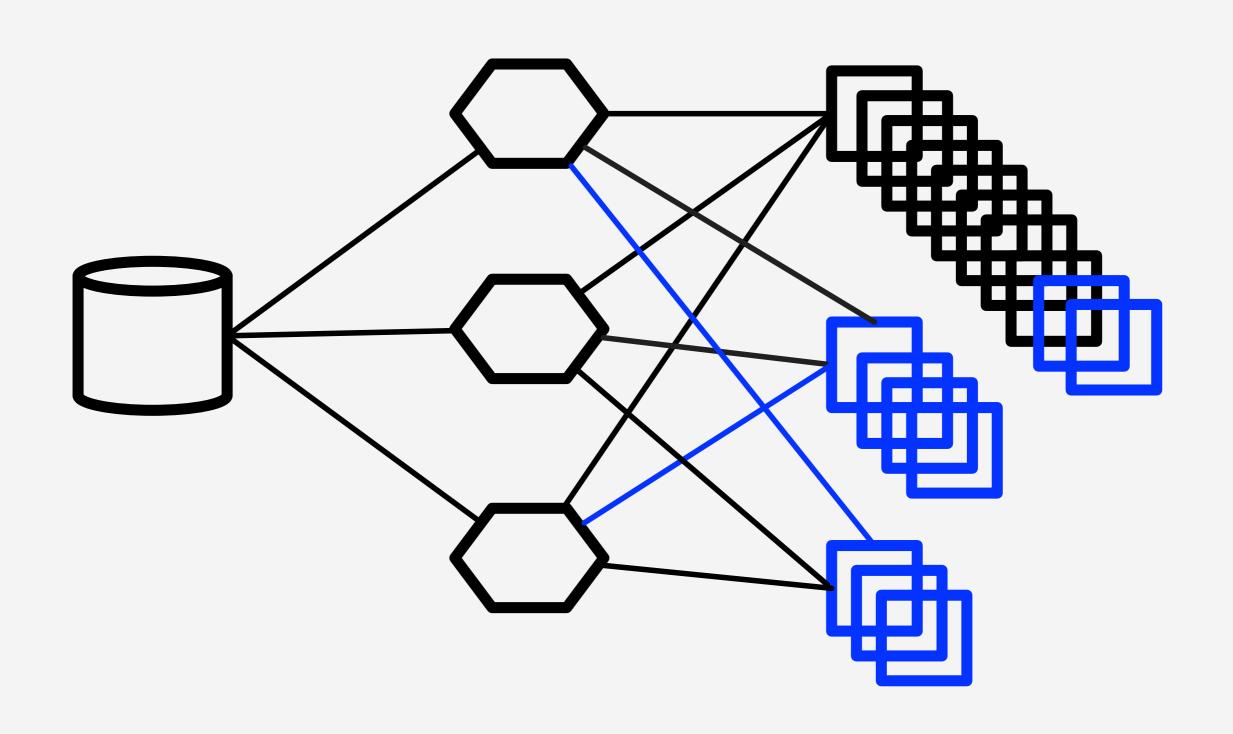




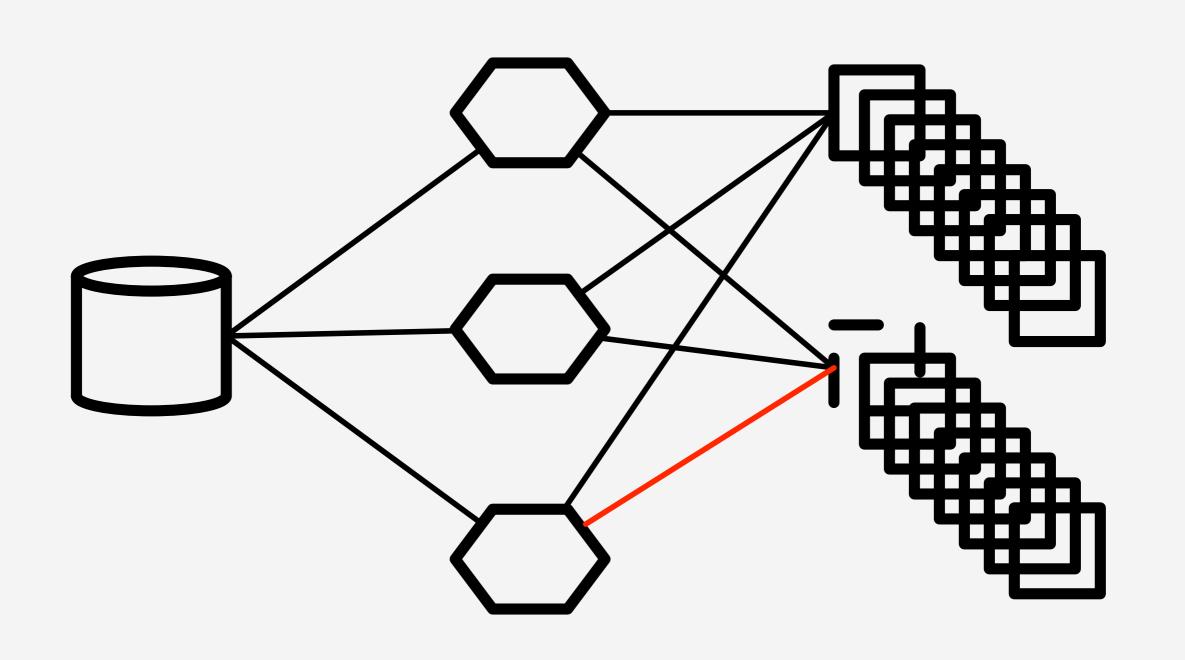


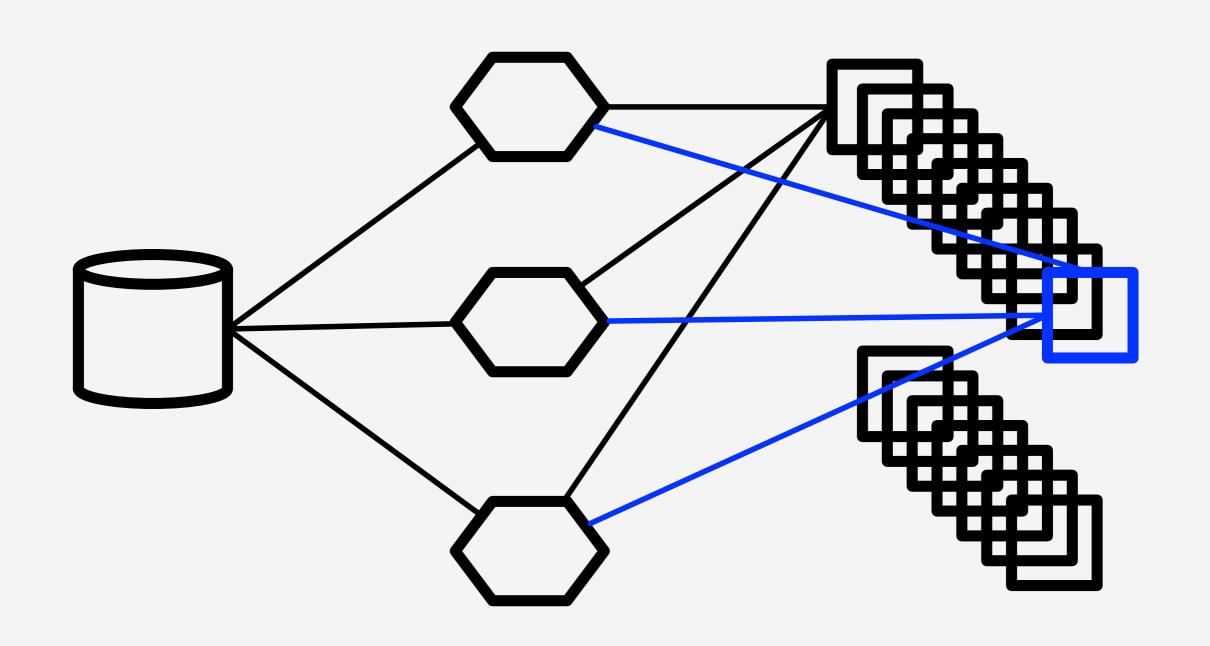
Crashes as the only code path



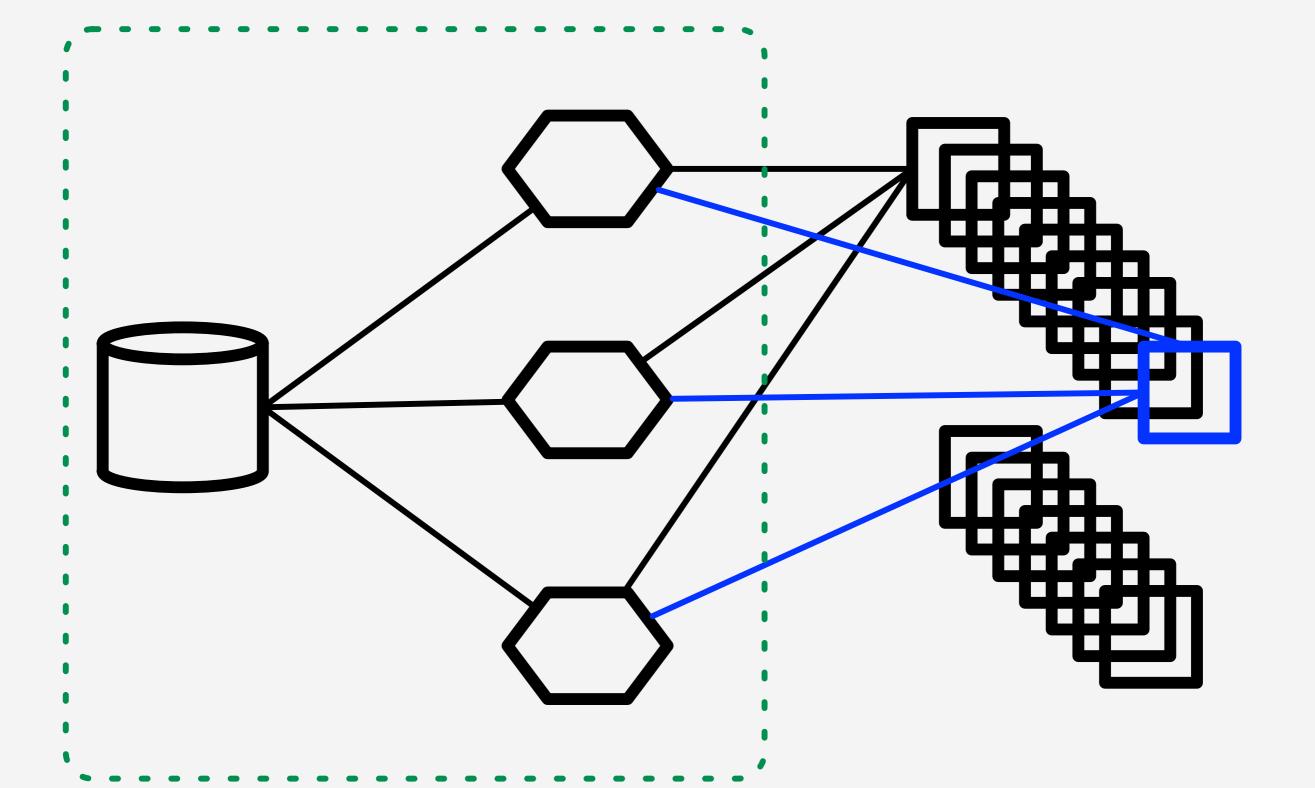


Crashes as a hot code path

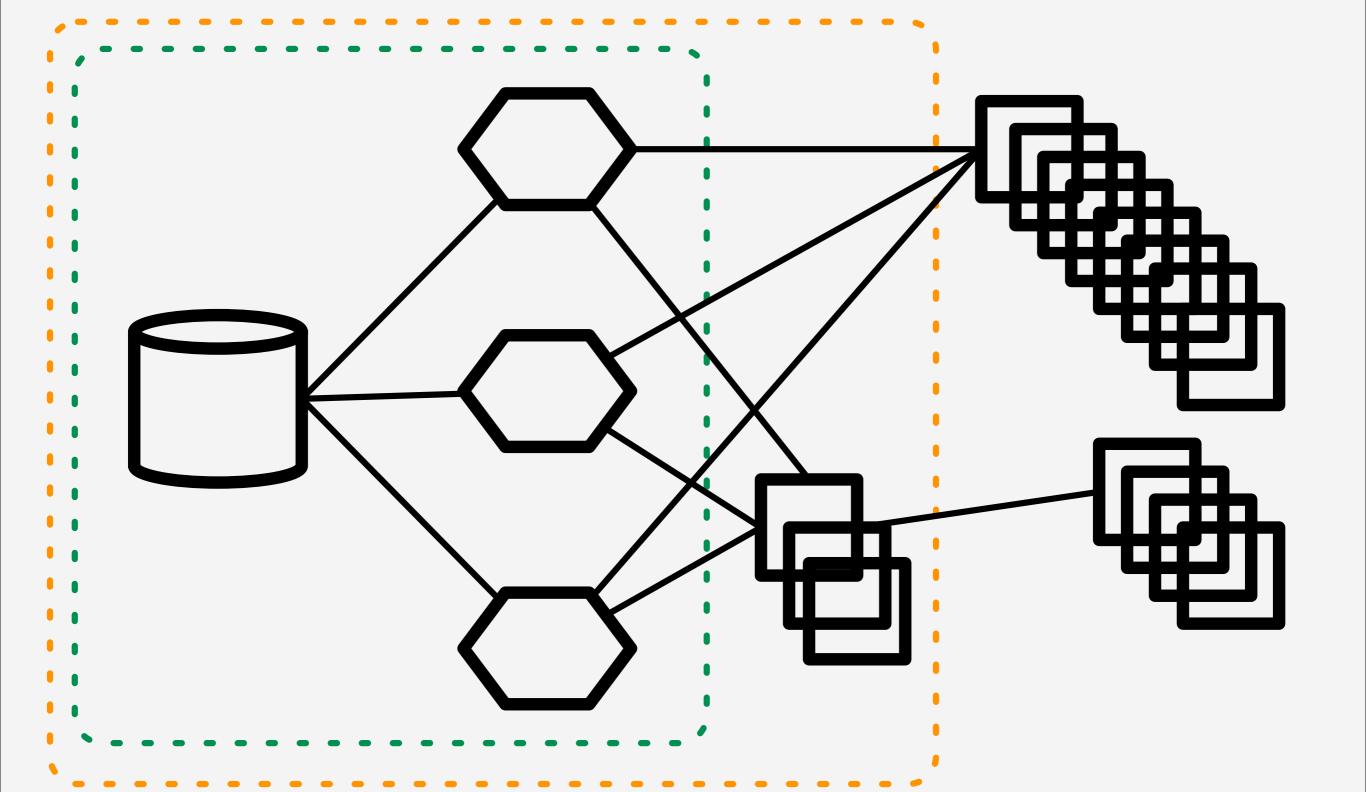




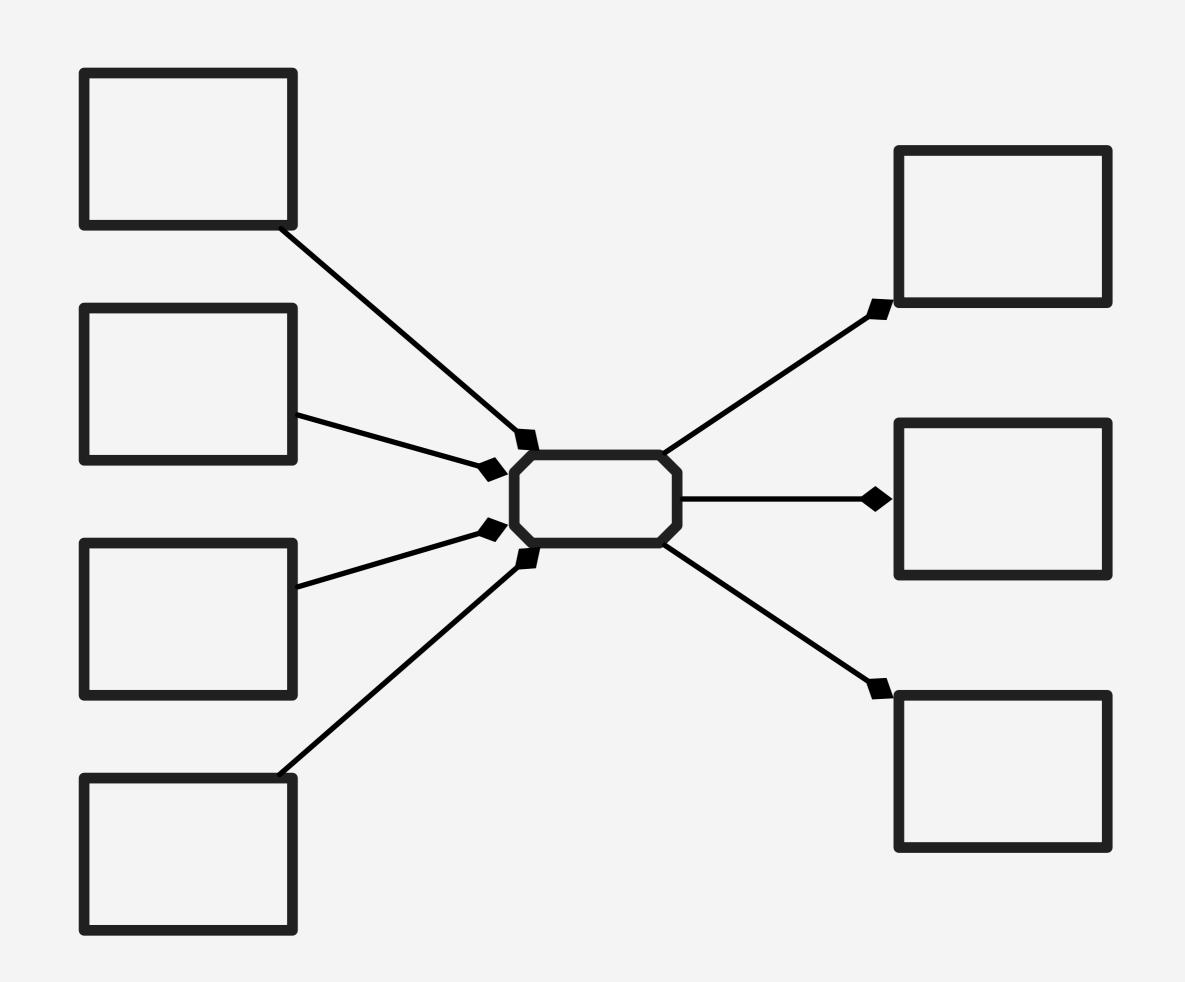
Error Kernel



Layered design

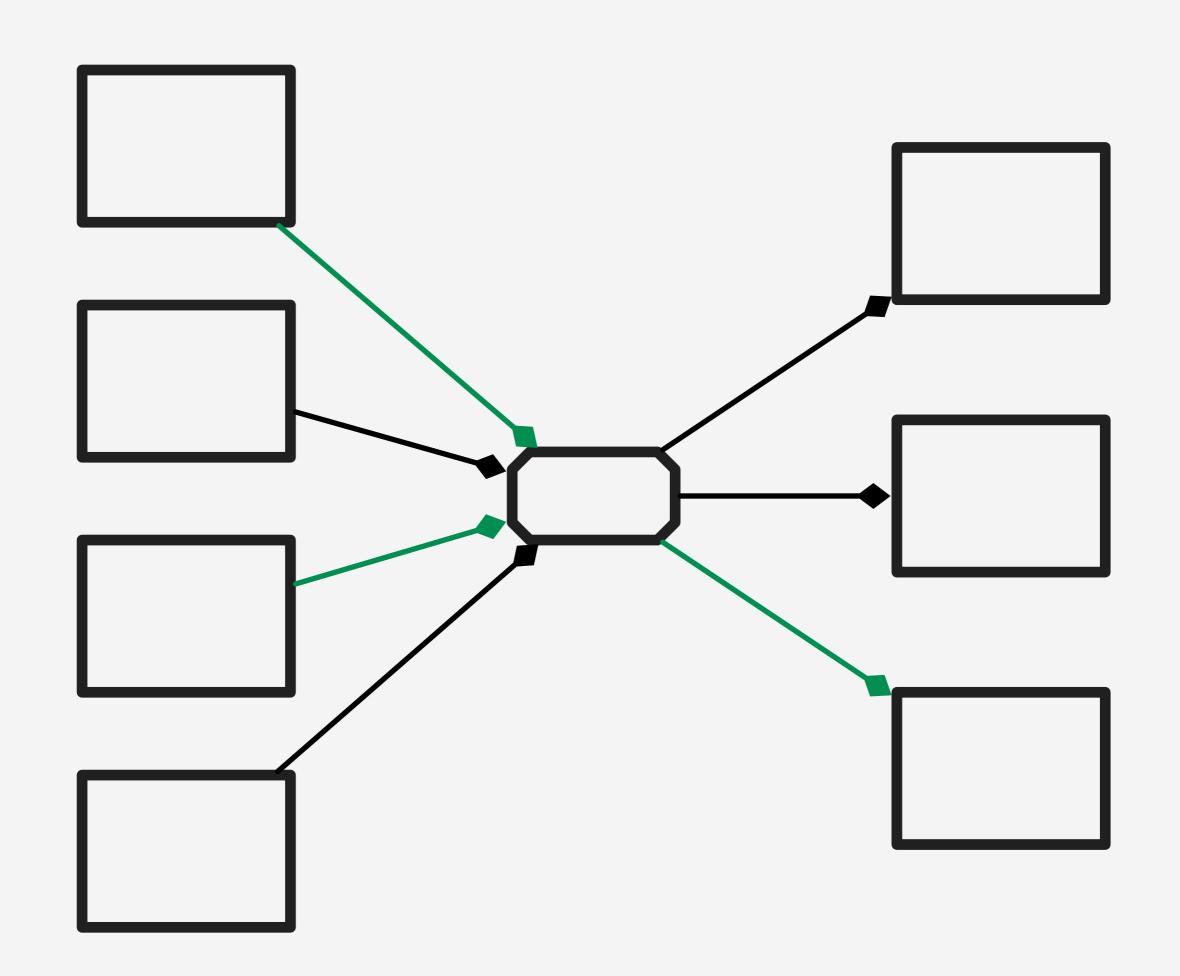


Message passing...

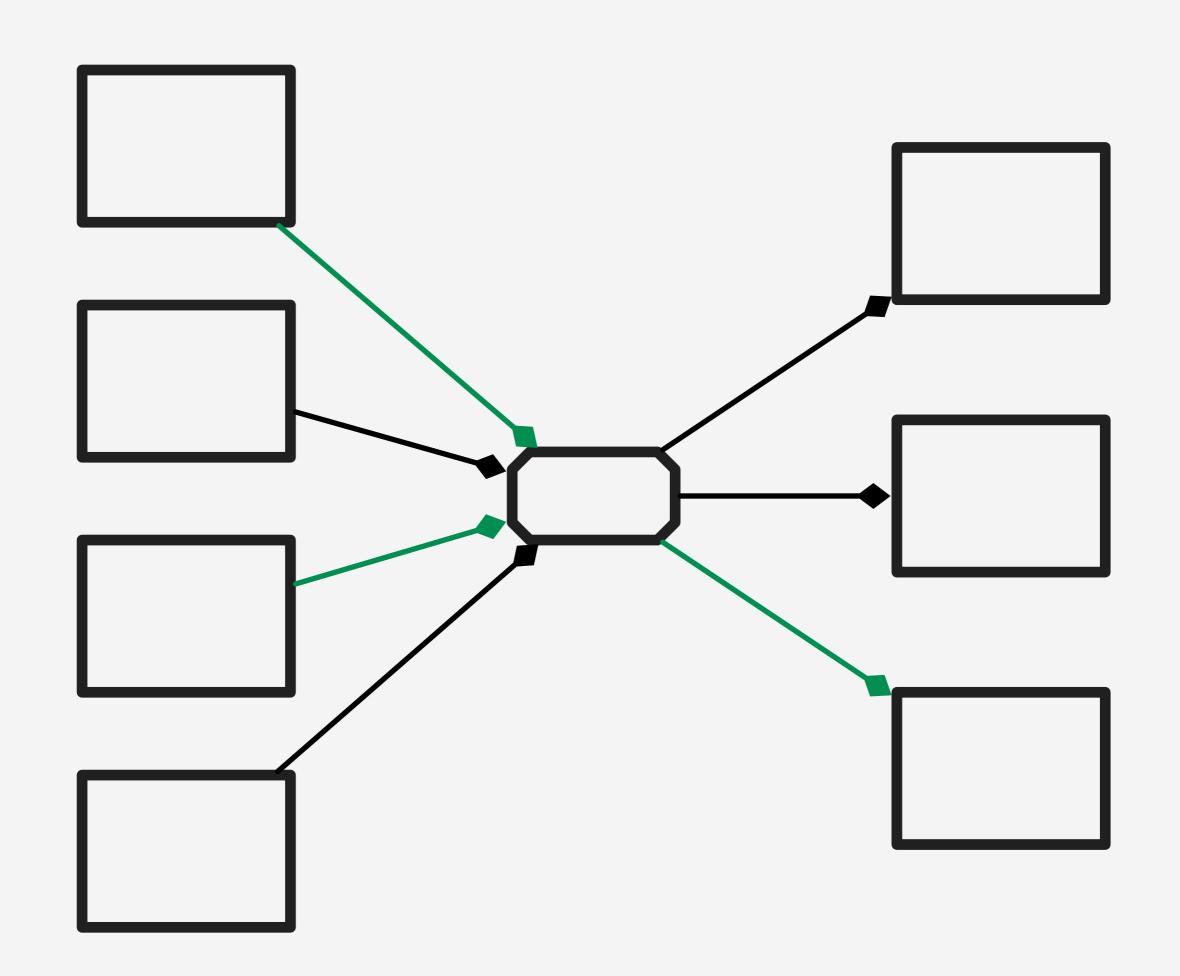


...of narrow, versioned values

```
{slug: "https://s3...",
cmd: "java -cp ...",
env: {"JAVA_OPTS": ...,
        "DATABASE_URL": ...,
        "SESSION_SECRET": ...},
flag: "extra_cpu"}
```



No Stopping the World



Load balancing Supervision Crash-only Error kernels Layered design Message-passing



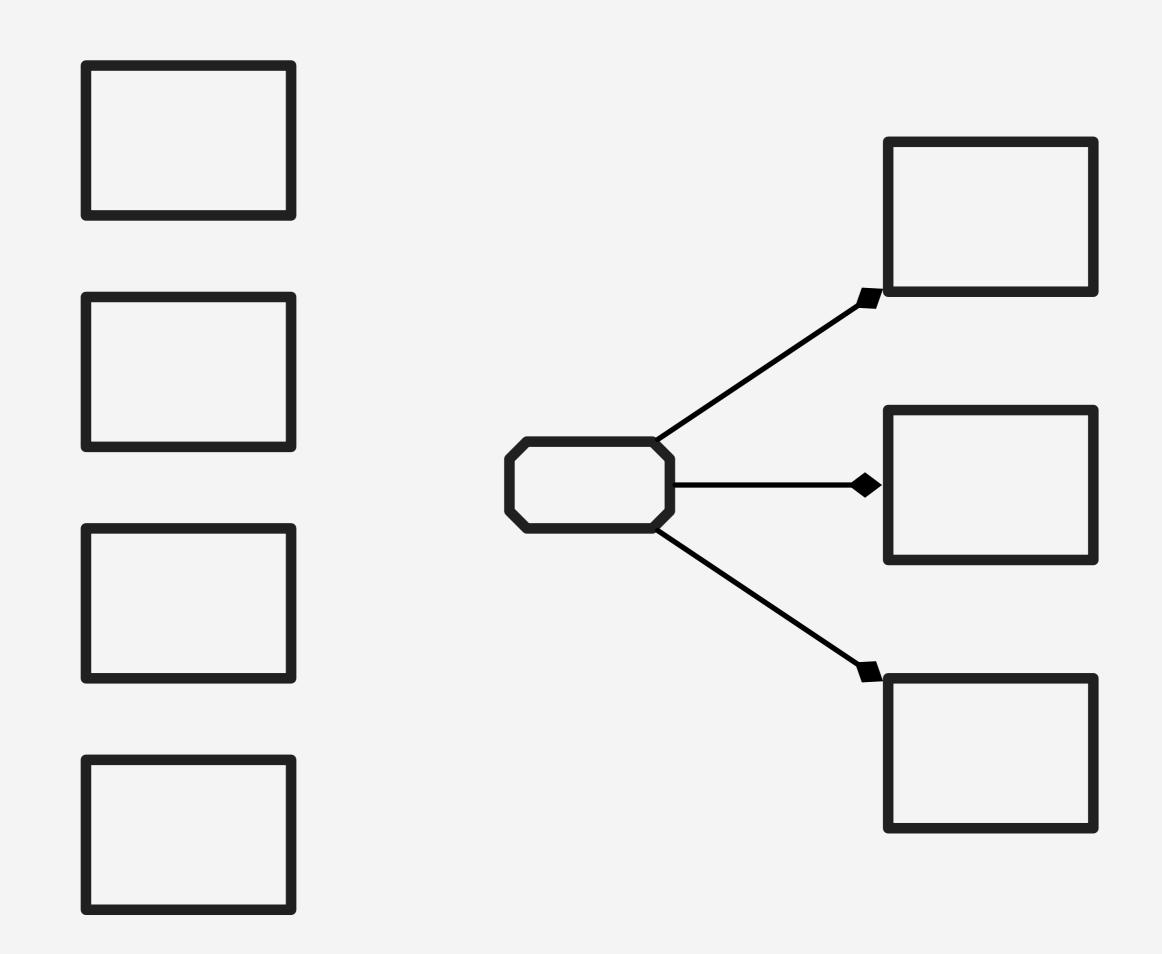
Erlang

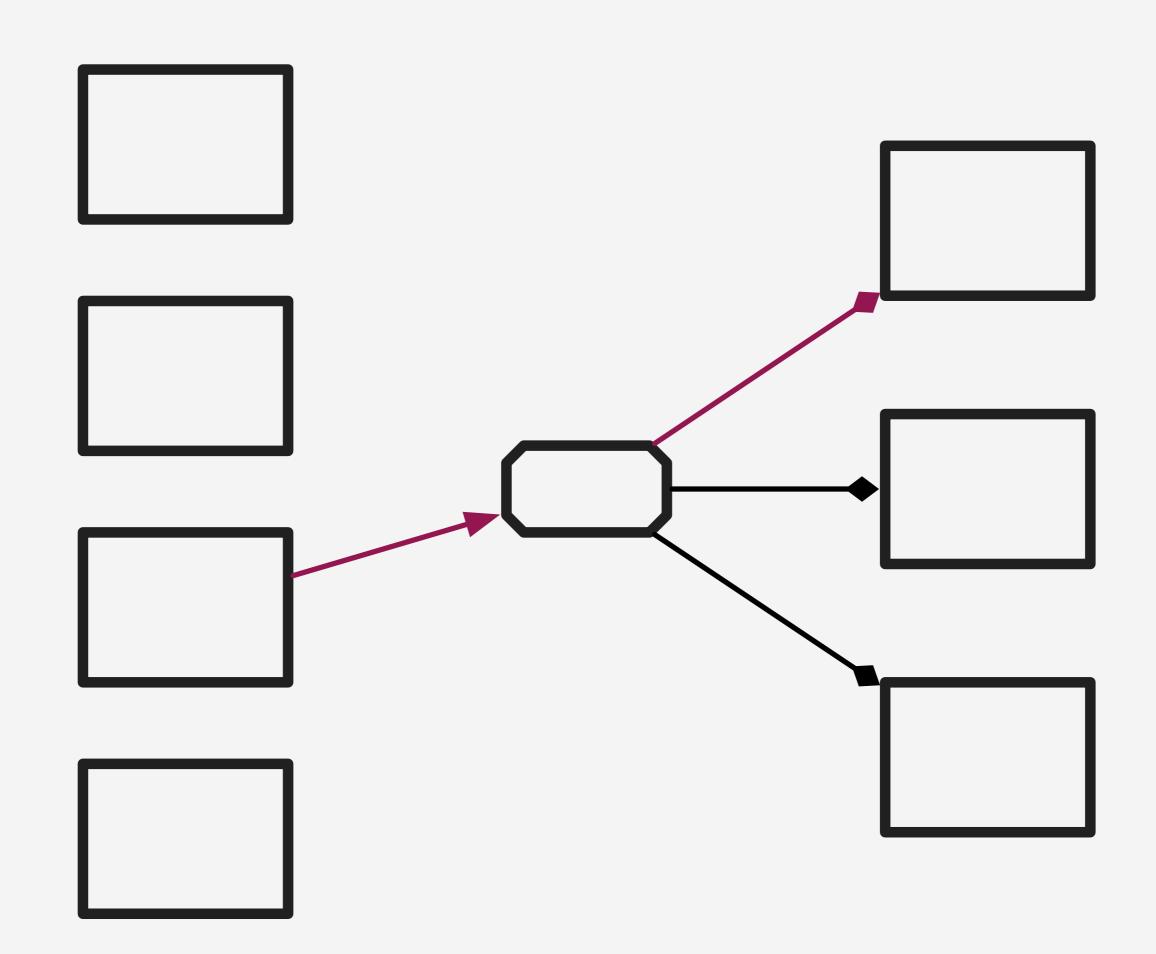
Designed for granular failure

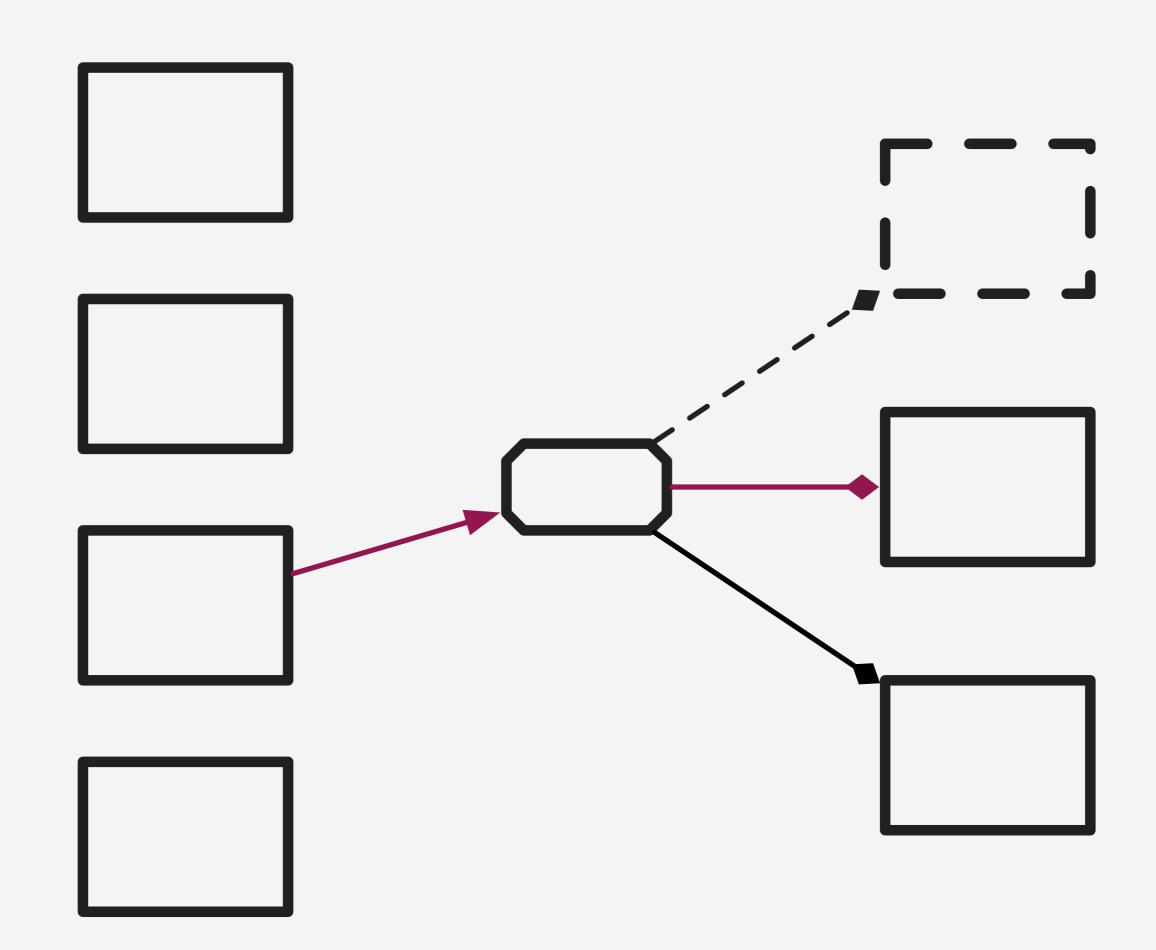
Distributed Systems

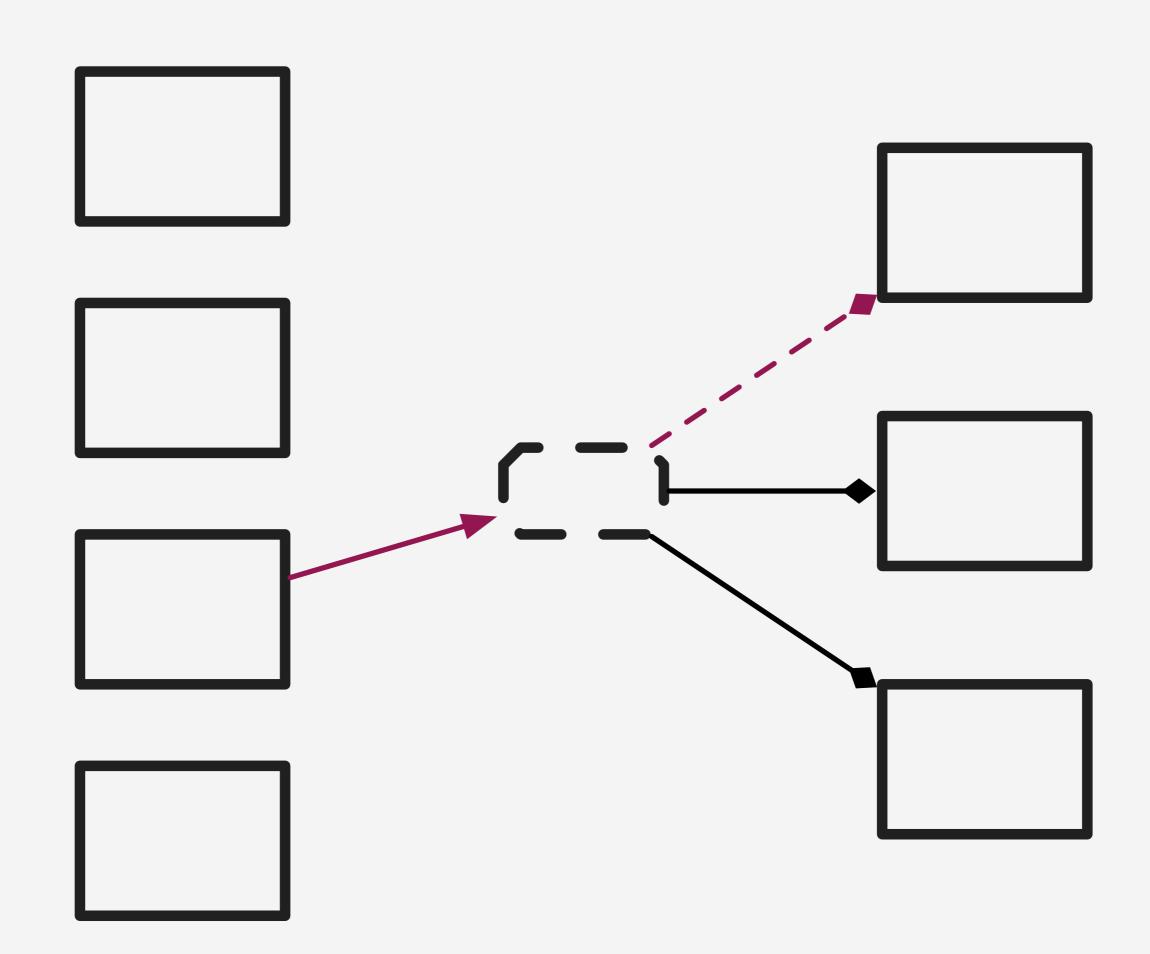
Defined as granular failure

Brokered Queueing

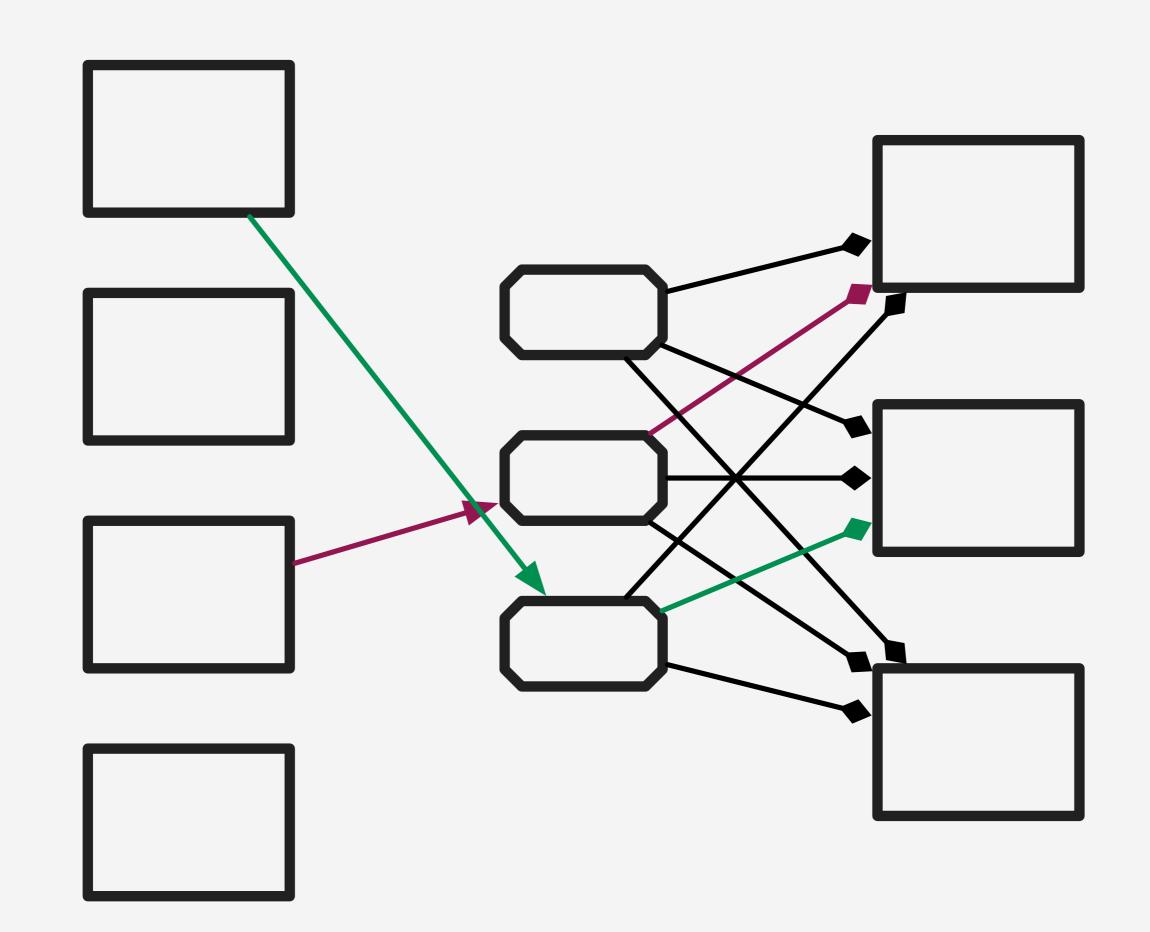


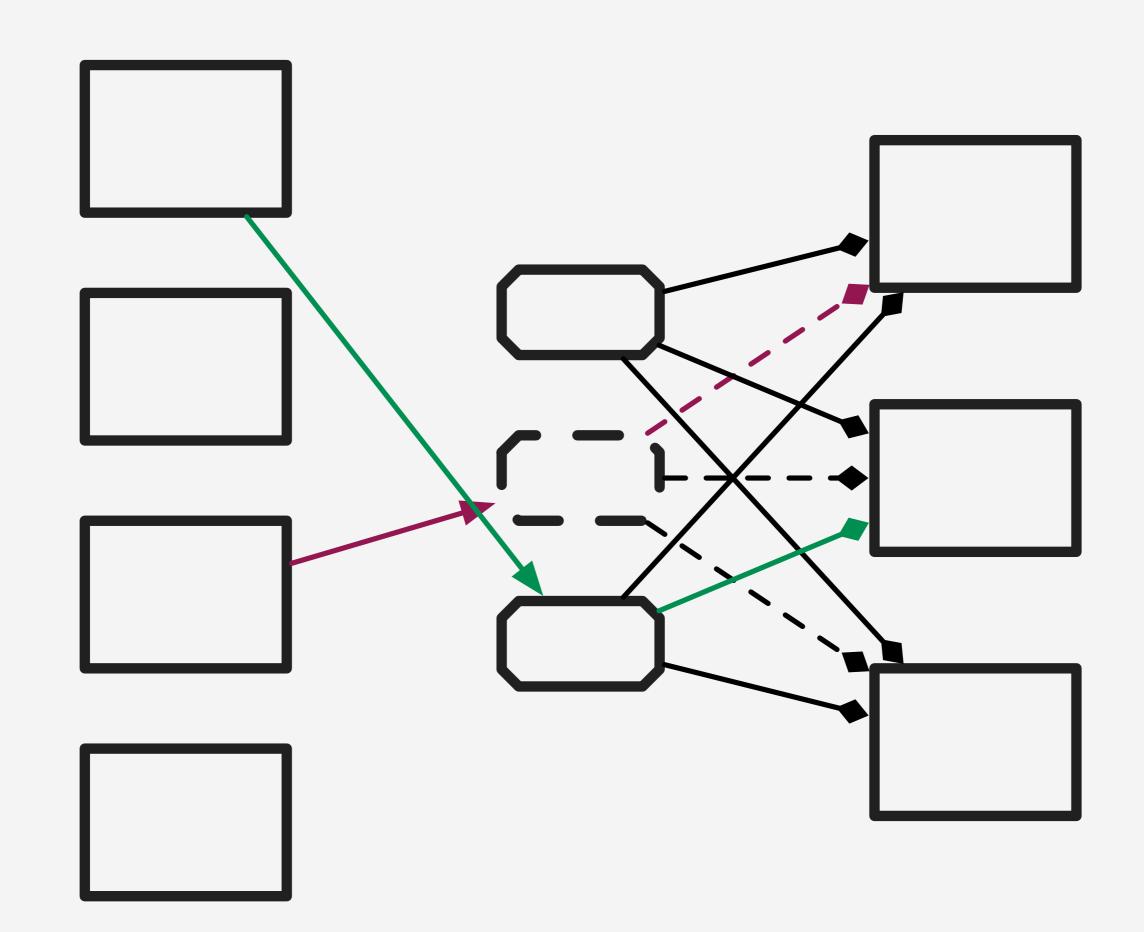


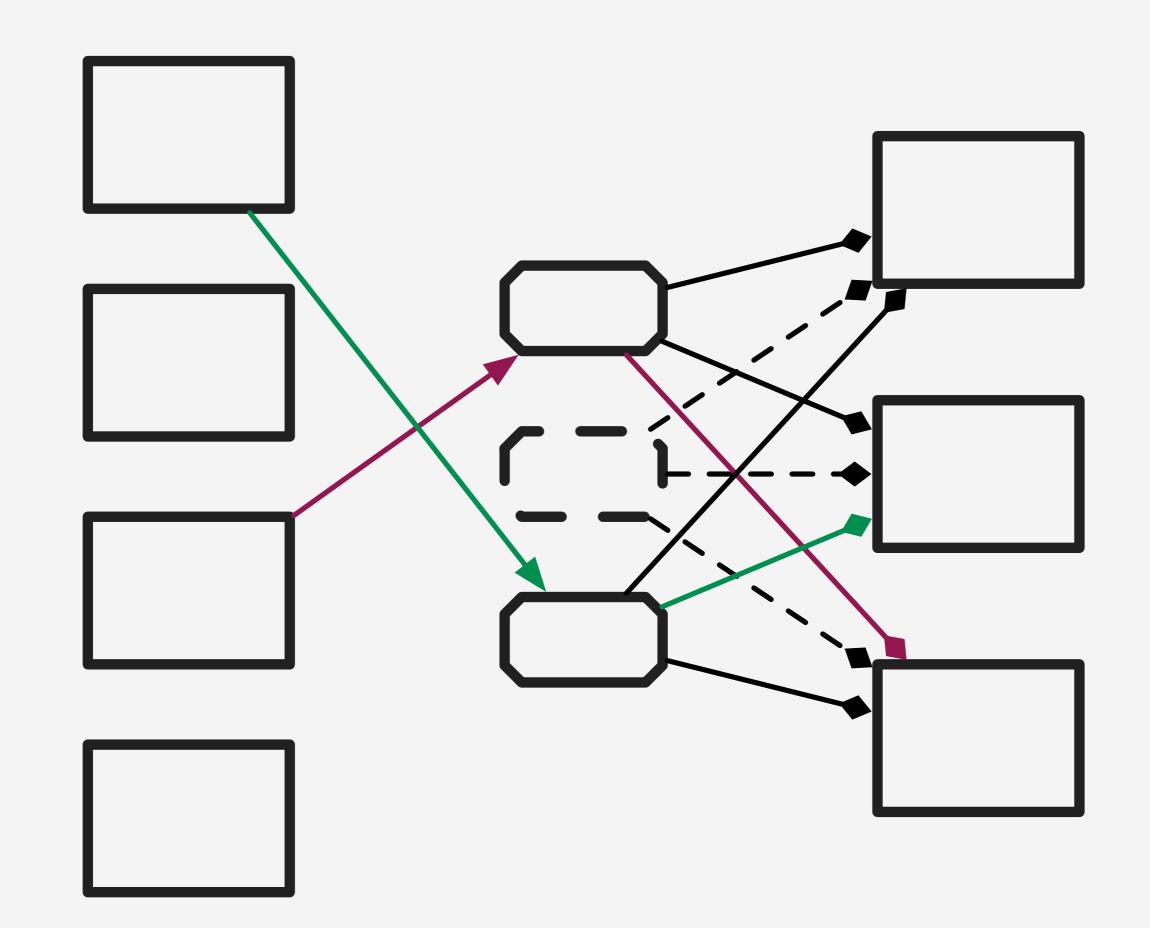




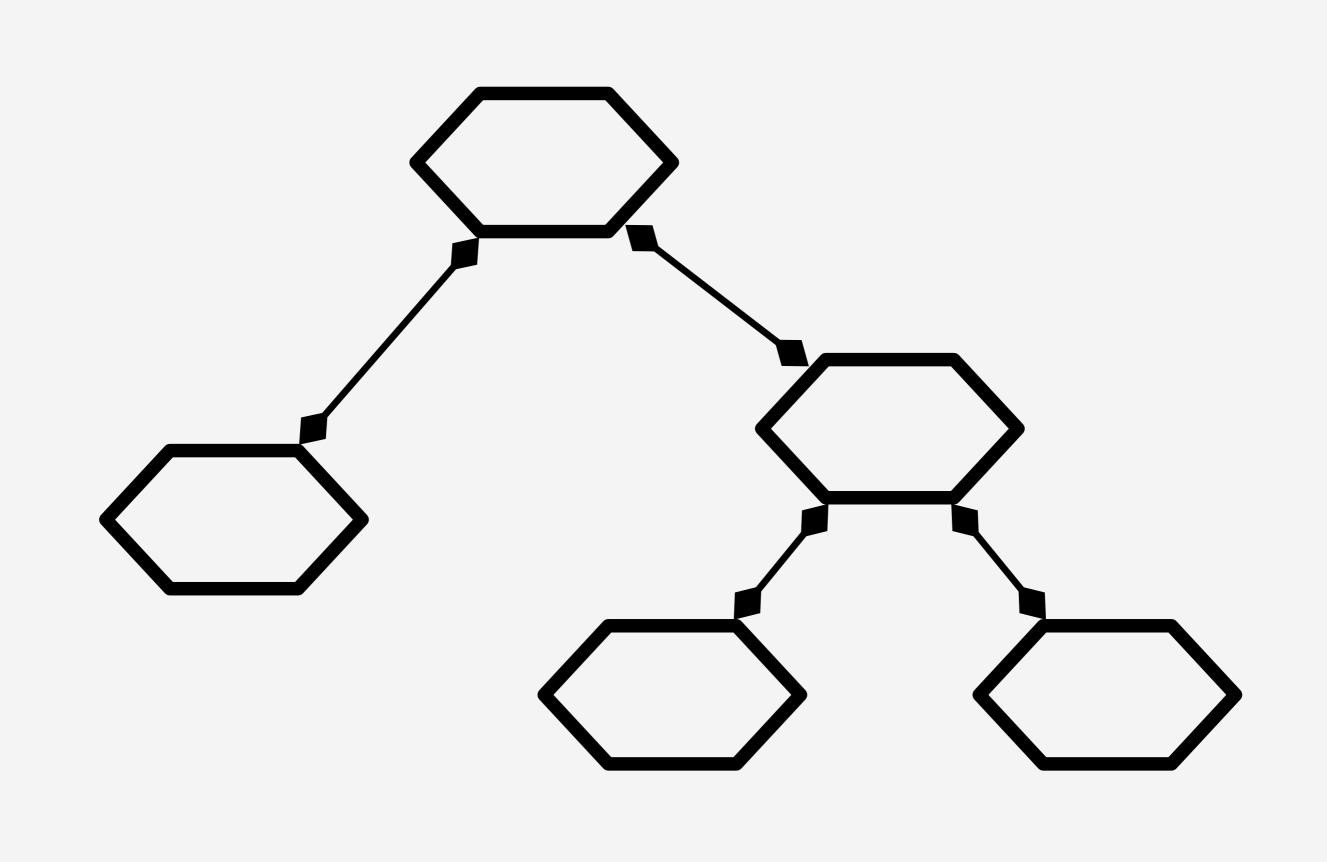
Publish one / Subscribe many

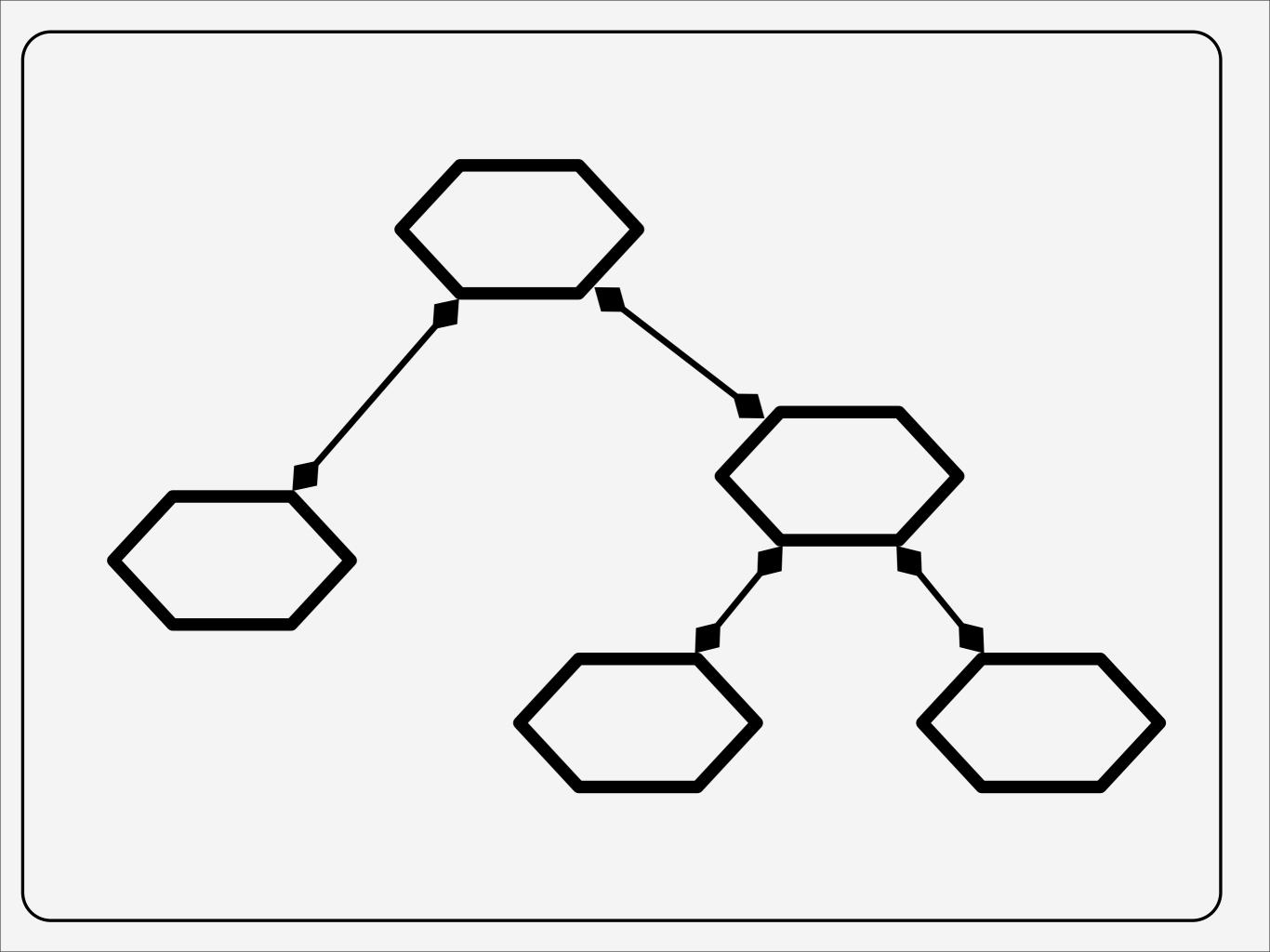


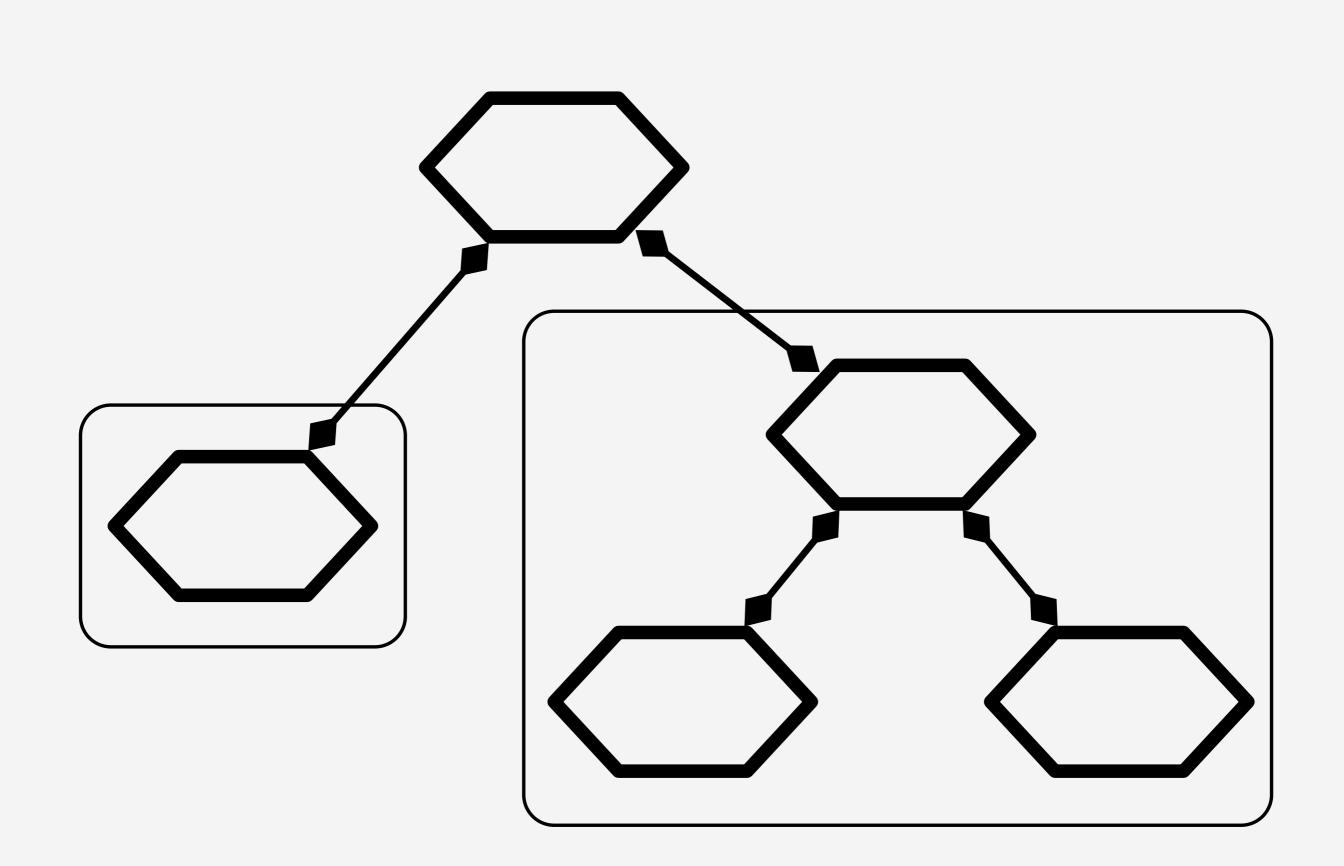


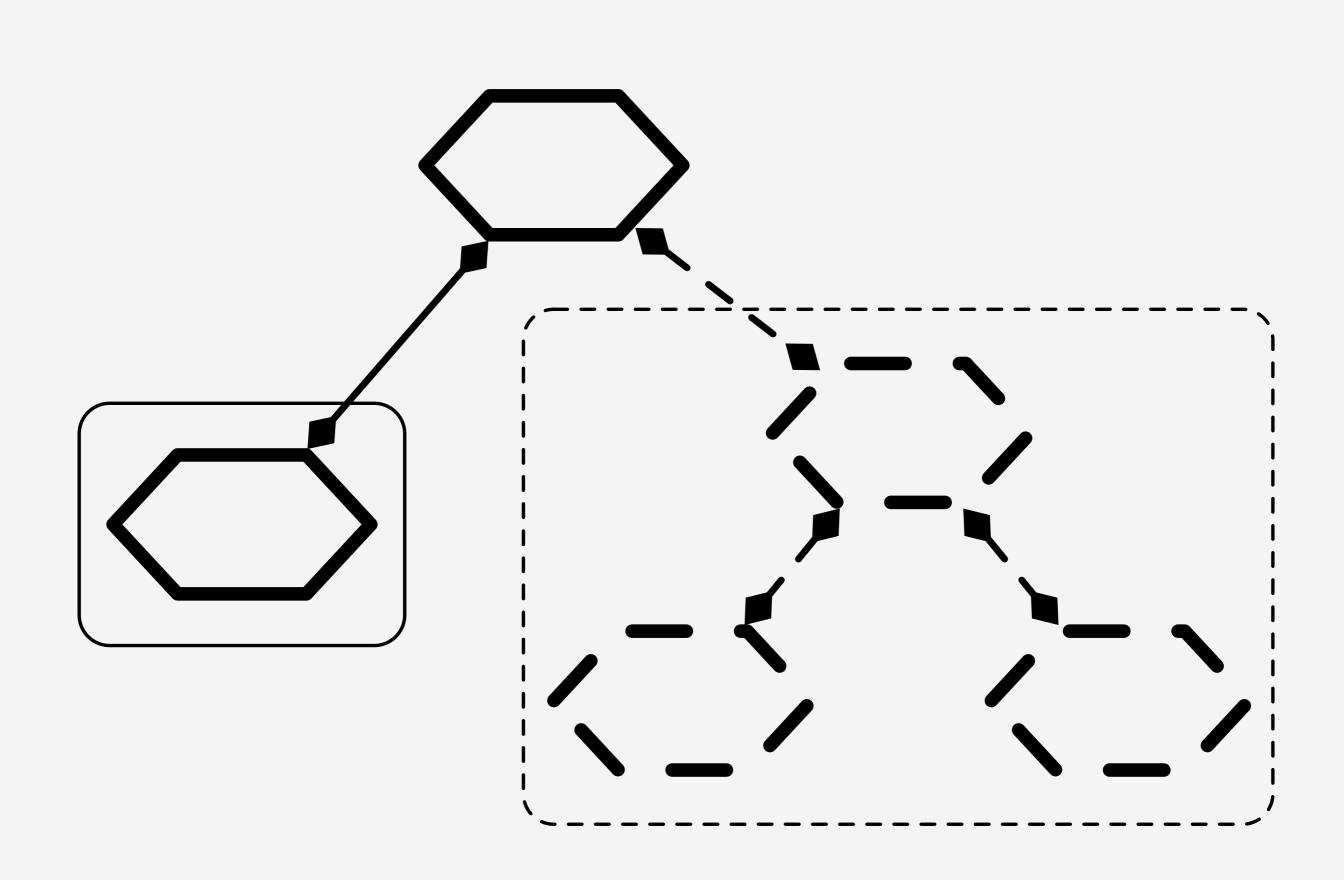


Distributed call graphs

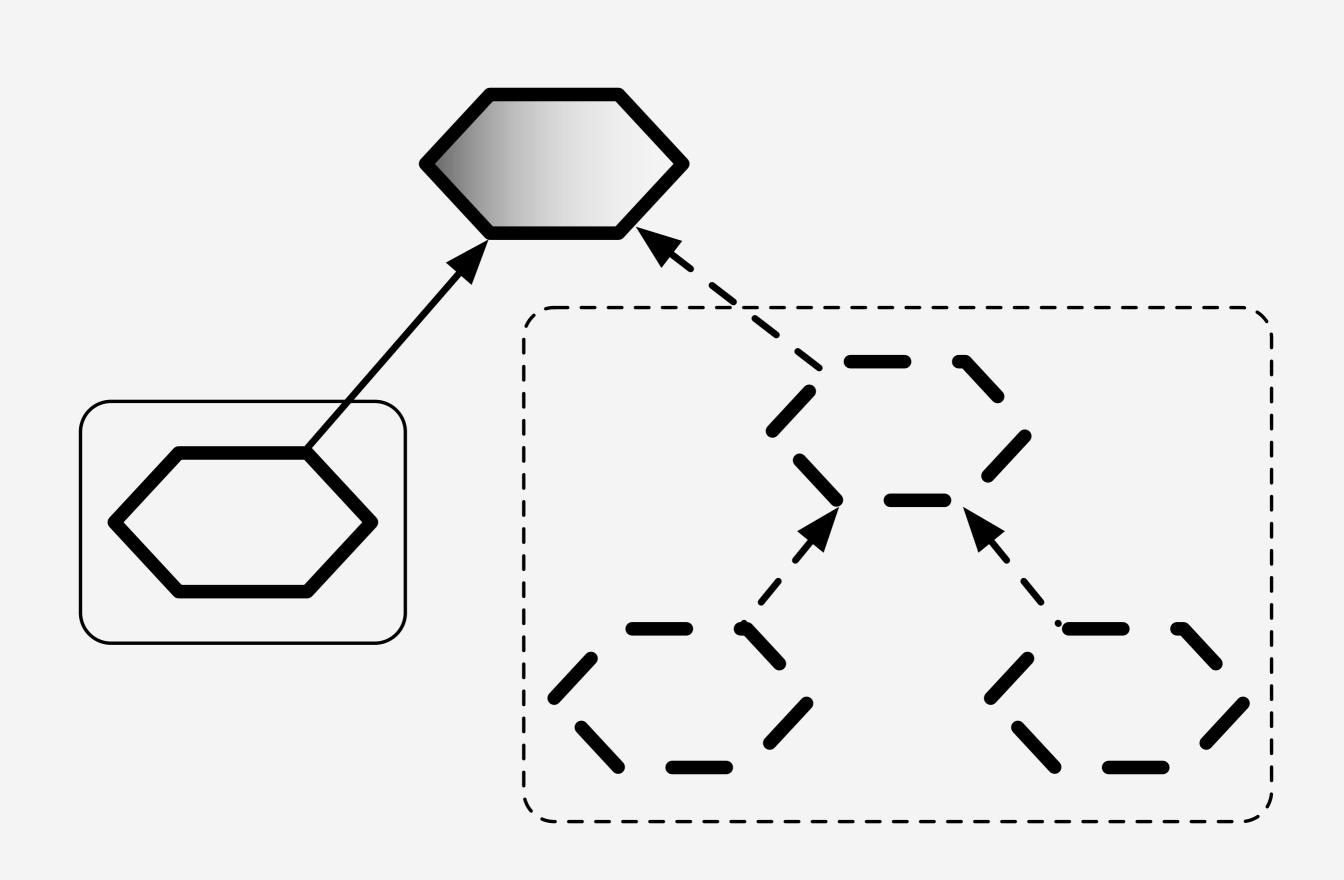




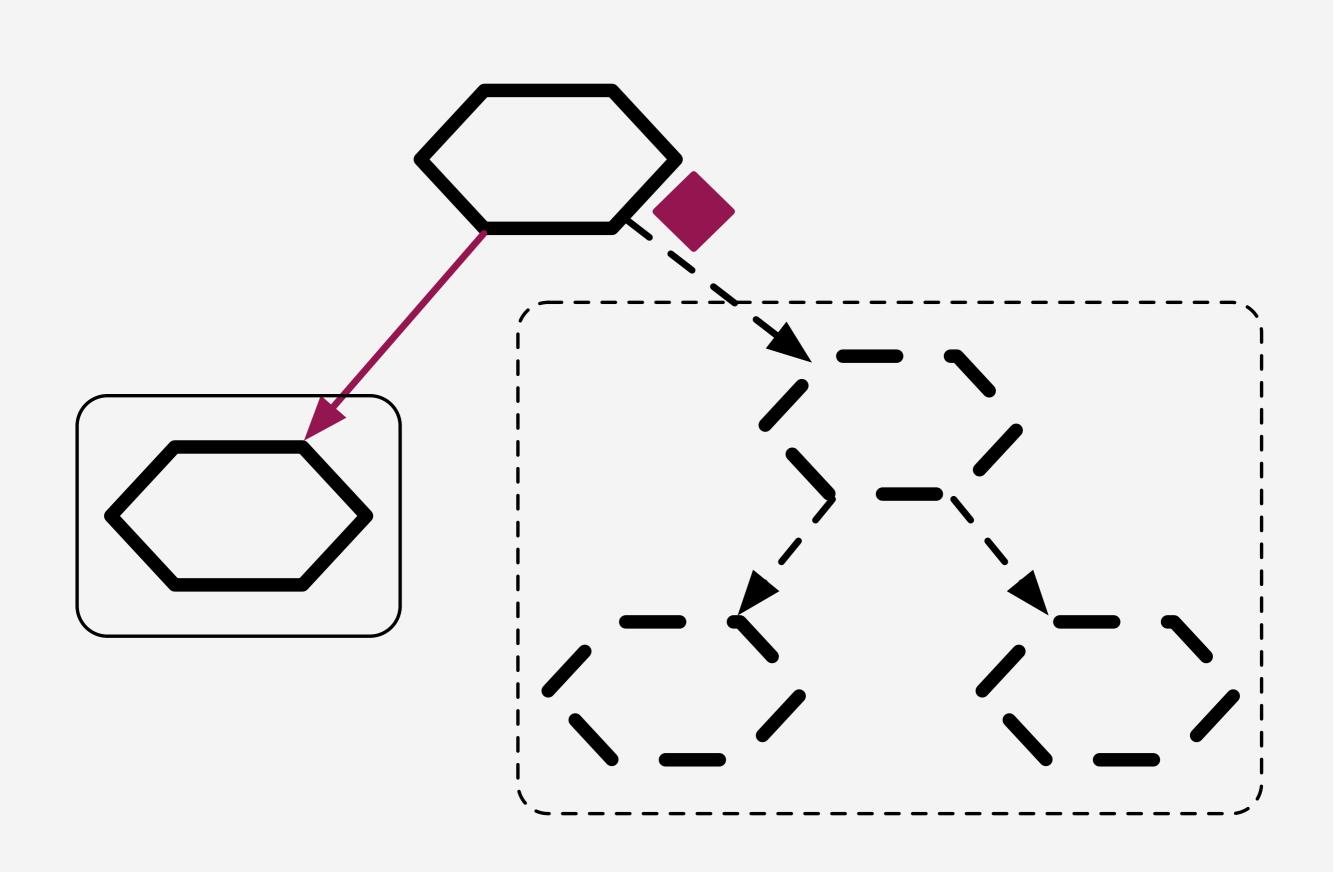


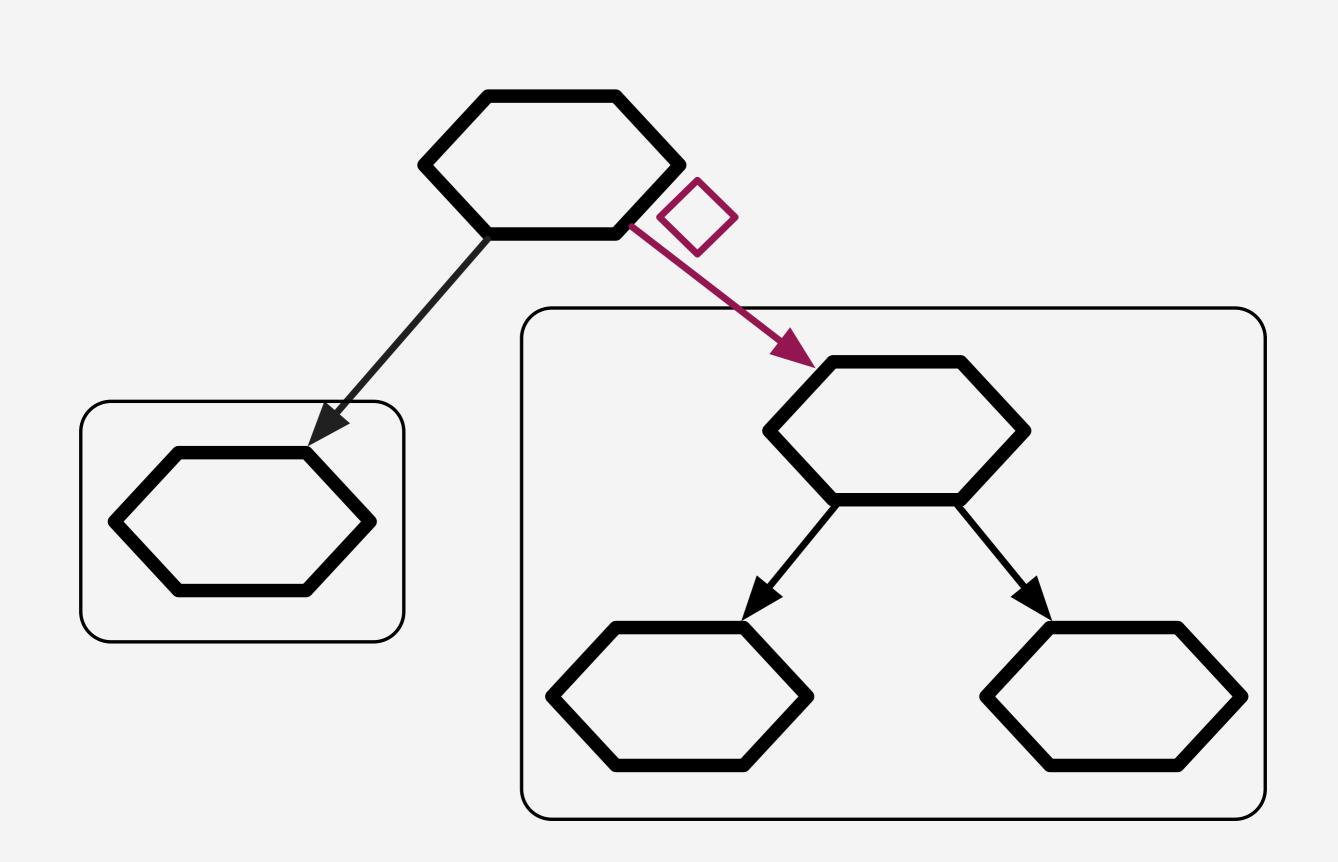


Read call graph Partial failure



Write call graph de-synchronizing





Architecture Execution

Architecture Execution

"...we deployed a code change...

...introduced a...problem...

...visible under unusual...conditions...

...engineers noticed a deviation...

...began to escalate...

...system...entered into a feedback loop...

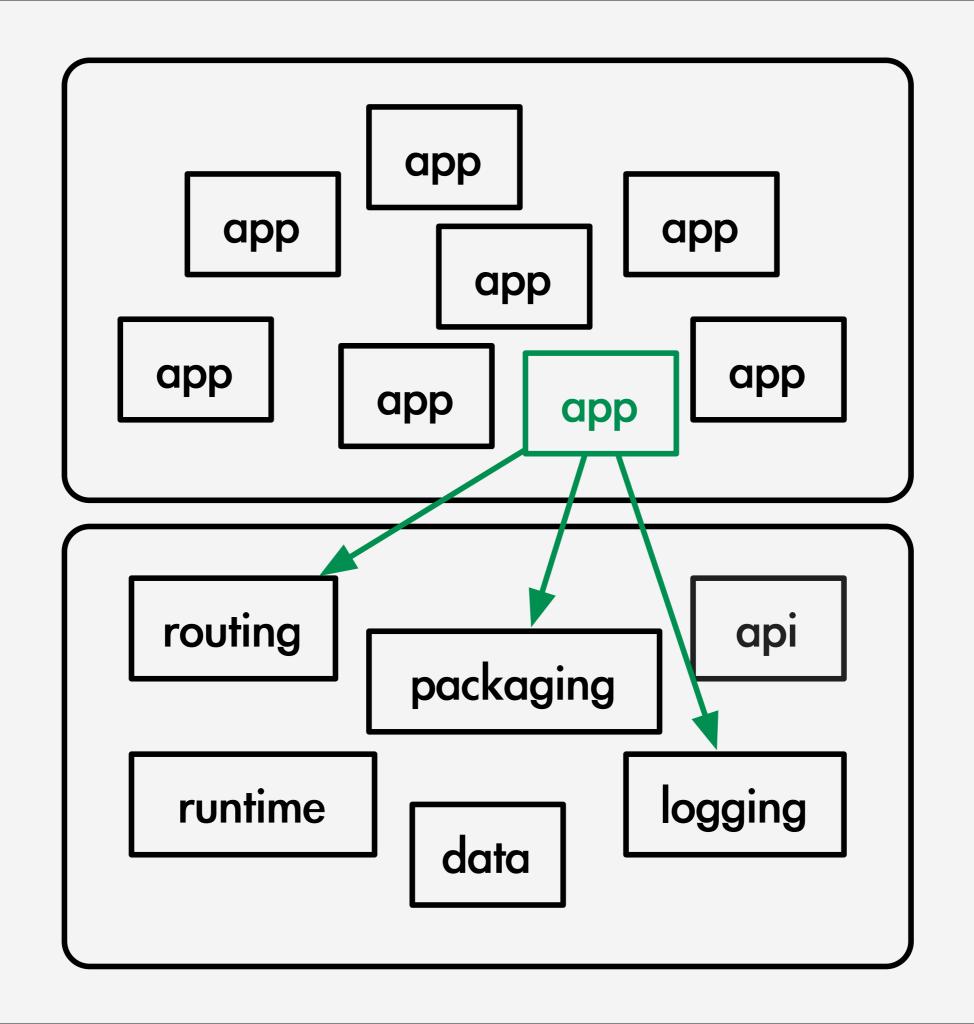
...engineers...deactivated the feedback...

Evolving Socio-Technical Systems

routing

runtime

data



Availability >> Architecture

Failed deploys Bad visibility Cascading feedback

Evolving Socio-Technical Systems

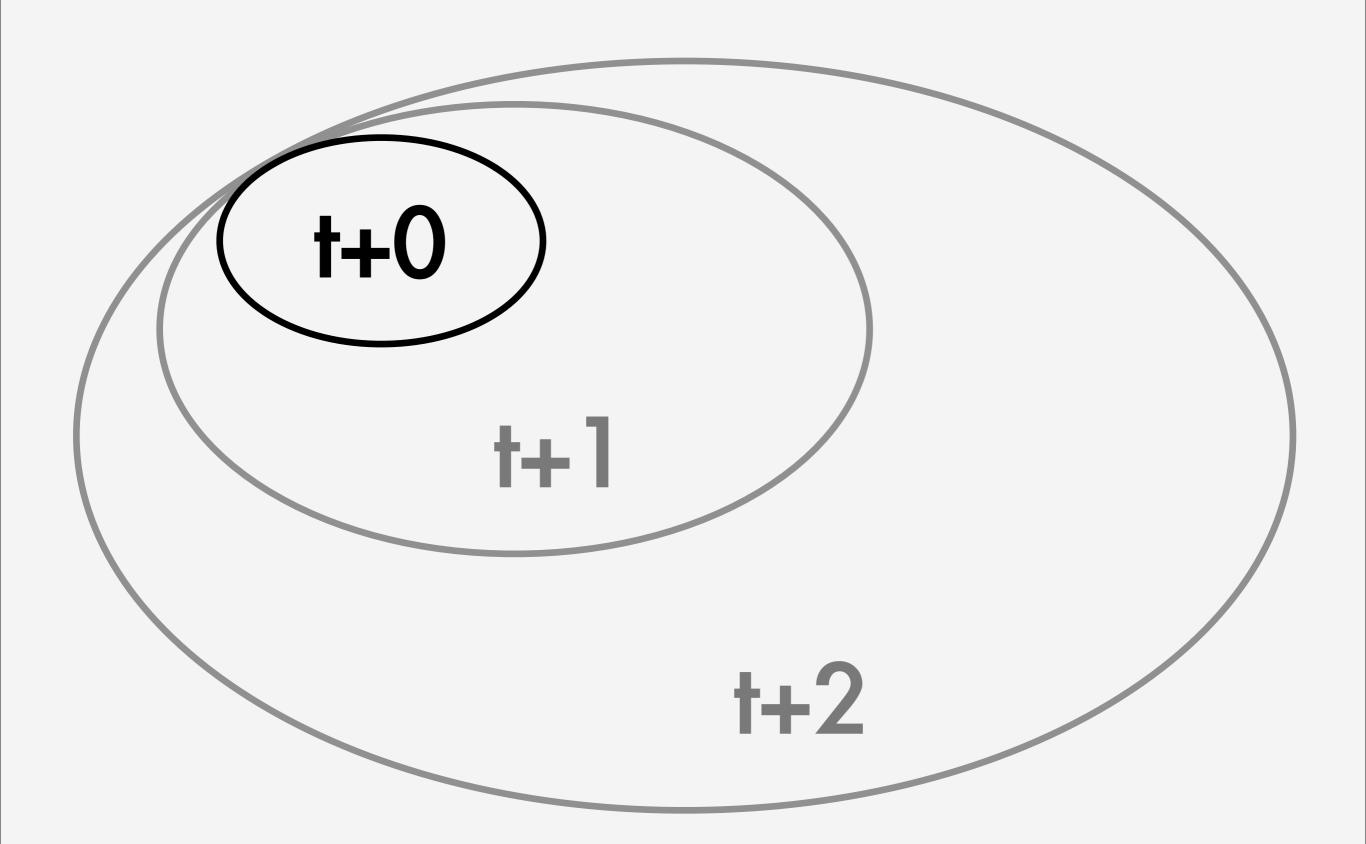
Failed deploys Bad visibility Cascading feedback

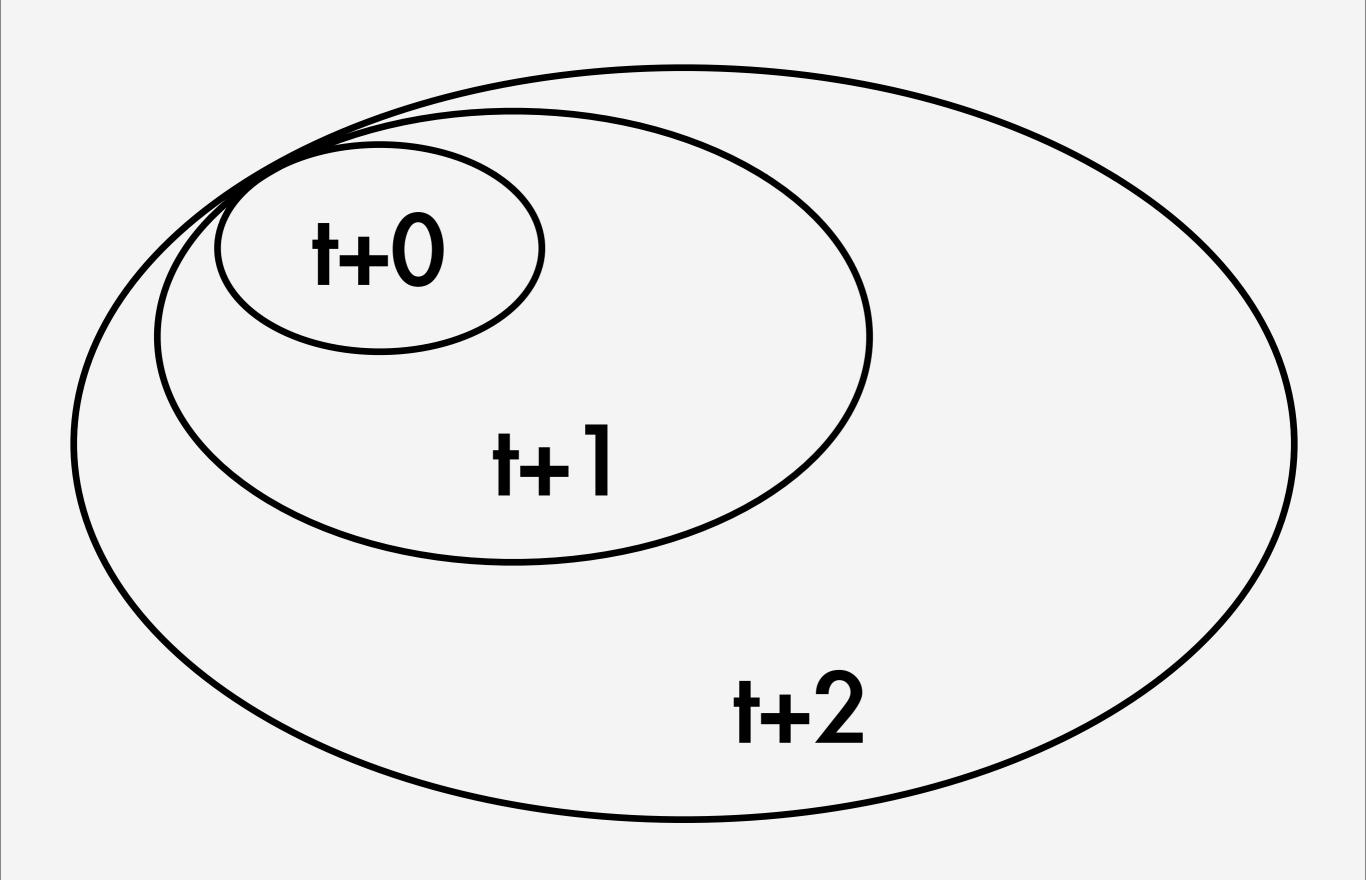
Deploy tooling Visibility services Feedback controls

bin/ship

bin/ship \ --component api \ --version v408

Incremental deploys





Incremental rollouts

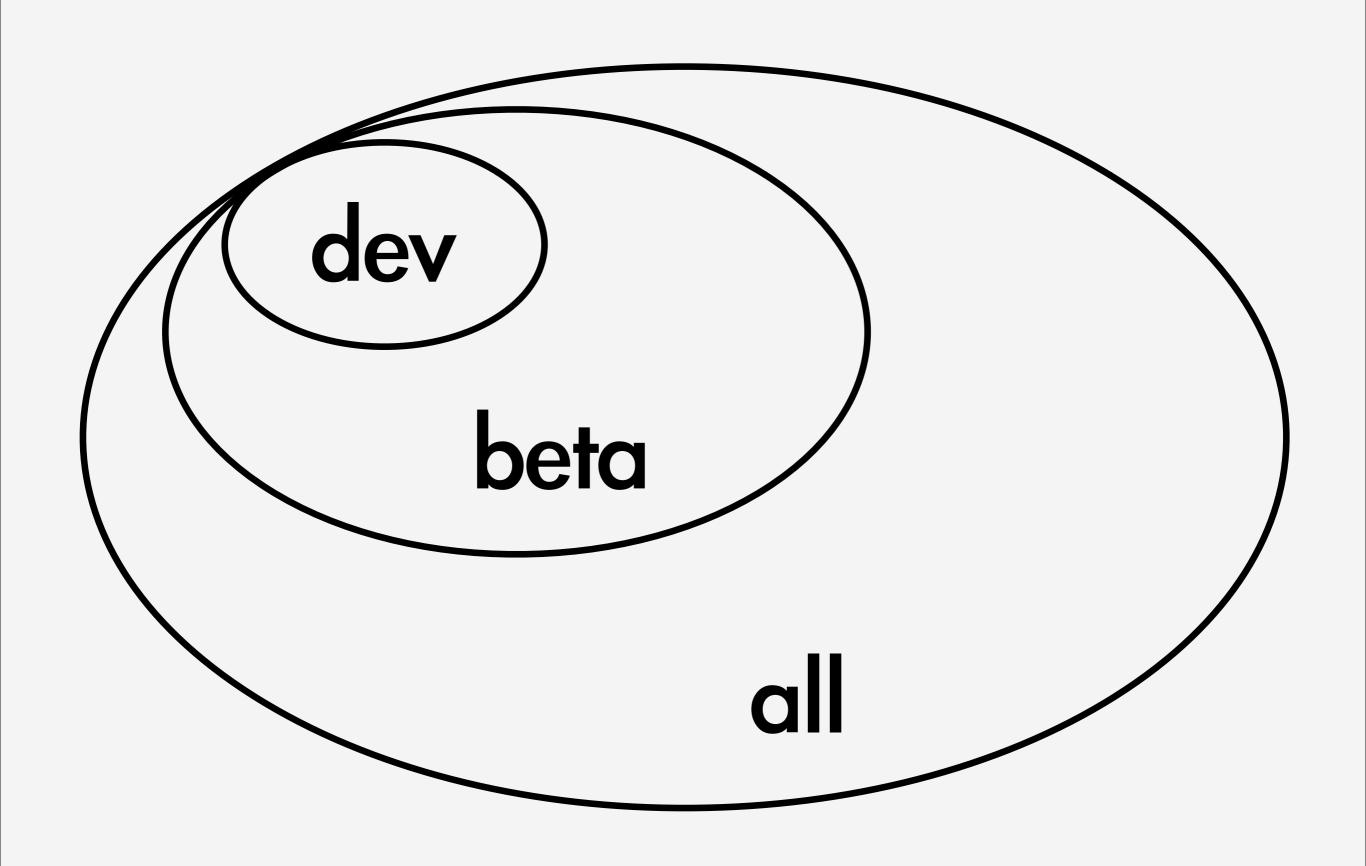
prep_launch

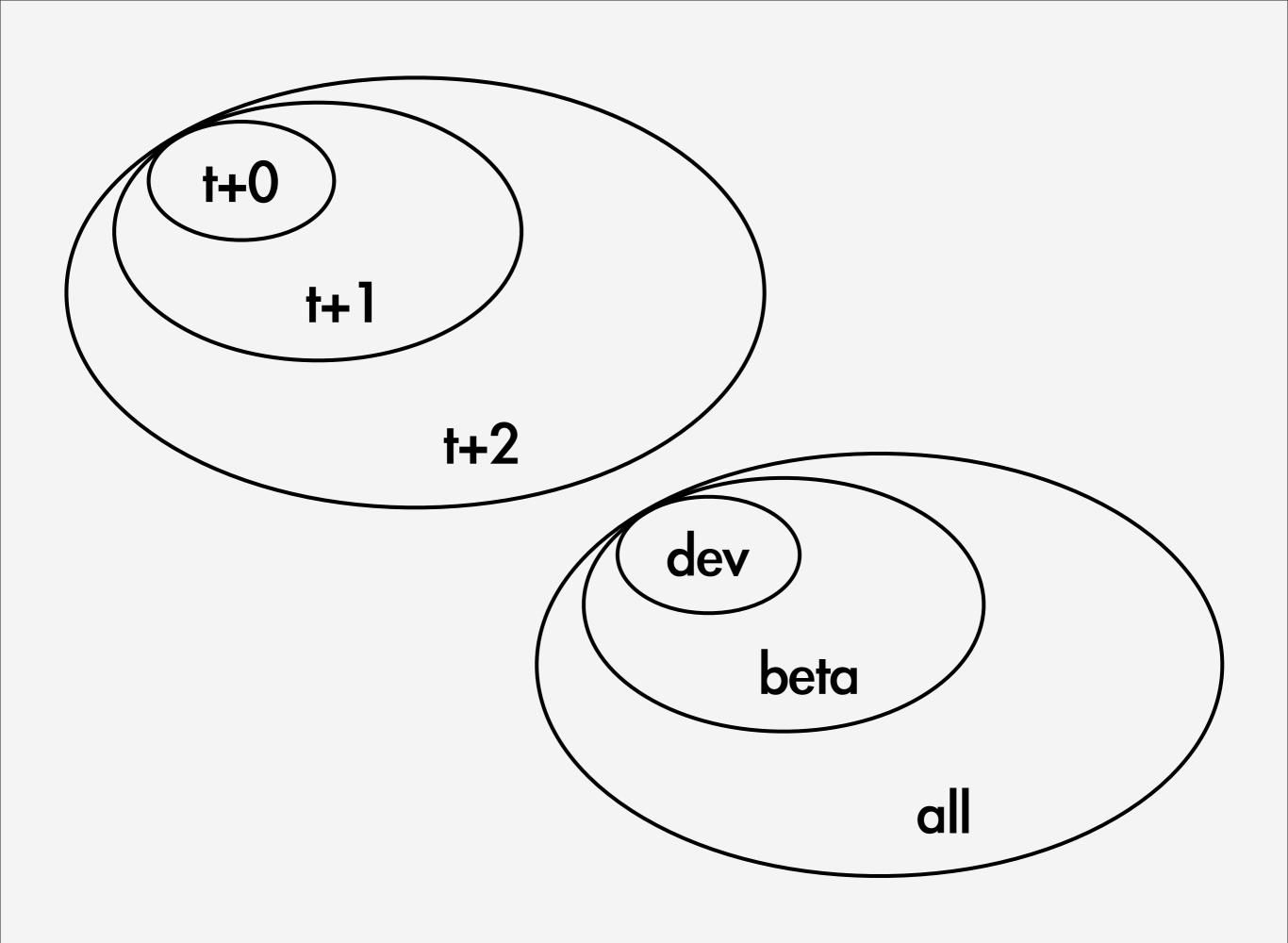
- launch_without_lxc
- + launch-with_lxc monitor_launch

```
if flag_on?("lxc")
  launch_with_lxc
else
  launch_without_lxc
```

app.flag_on("lxc")

app.flag_off("lxc")





Real-time visibility

nginx req/sec:	nginx 500 dom/min:	nginx 502 dom/min:	nginx 503 dom/min:	nginx 504 dom/min:	nginx err/min:	event/sec:
varnish req/sec:	varnish 500/min:	vamish 502/min:	vamish 503/min:	vamish 504/min:	varnish purge/min:	rdv join/min:
		_				
hermes req/sec:	hermes h10 app/min:	hermes h11 app/min:	hermes h12 app/min:	hermes h13 app/min:	hermes h14 app/min:	hermes h99 app/min:
psmgr ps up total:	psmgr ps up web:	psmgr ps up worker:	psmgr ps up other:	psmgr ps starting:	psmgr ps crashed:	psmgr ps created:
mileus es tetale	ilava aa waba					
railgun ps total:	railgun ps web:	railgun ps worker:	railgun ps clock:	psmgr run req/min:	psmgr kill req/min:	psmgr ps lost:
psmgr idle/min:	psmgr unidle/min:	psmgr conv/sec:	railgun save time:	railgun run/min:	railgun kill/min:	railgun launch time:
runtime bus depth:	psmgr bus rec/min:	psmgr bus tout/min:	psmgr bus pub/min:	amqp pub/sec:	amqp rec/sec:	amqp tout/min:
		_				
api release/min:	slugc compile/min:	slugc fail/min:	slugc err/min:			
railgun excp/min:	psmgr excp/min:	api excp/min:	codex err/min:	gitproxy err/min:	shen err/min:	hermes error/min:

Service-level assertions

nginx req/sec:	nginx 500 dom/min:	nginx 502 dom/min:	nginx 503 dom/min:	nginx 504 dom/min:	nginx err/min:	event/sec:
varnish req/sec:	varnish 500/min:	vamish 502/min:	vamish 503/min:	vamish 504/min:	varnish purge/min:	rdv join/min:
		_				
hermes req/sec:	hermes h10 app/min:	hermes h11 app/min:	hermes h12 app/min:	hermes h13 app/min:	hermes h14 app/min:	hermes h99 app/min:
psmgr ps up total:	psmgr ps up web:	psmgr ps up worker:	psmgr ps up other:	psmgr ps starting:	psmgr ps crashed:	psmgr ps created:
mileus es tetale	ilava aa waba					
railgun ps total:	railgun ps web:	railgun ps worker:	railgun ps clock:	psmgr run req/min:	psmgr kill req/min:	psmgr ps lost:
psmgr idle/min:	psmgr unidle/min:	psmgr conv/sec:	railgun save time:	railgun run/min:	railgun kill/min:	railgun launch time:
runtime bus depth:	psmgr bus rec/min:	psmgr bus tout/min:	psmgr bus pub/min:	amqp pub/sec:	amqp rec/sec:	amqp tout/min:
		_				
api release/min:	slugc compile/min:	slugc fail/min:	slugc err/min:			
railgun excp/min:	psmgr excp/min:	api excp/min:	codex err/min:	gitproxy err/min:	shen err/min:	hermes error/min:

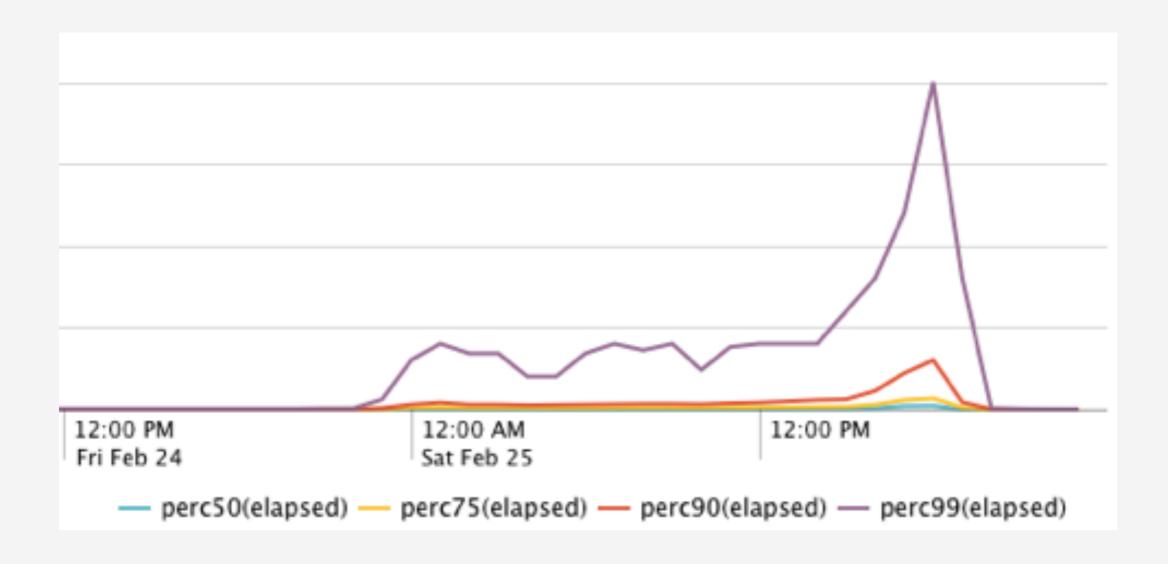
assert(index > 0)

assert(index > 0)

objects[index]

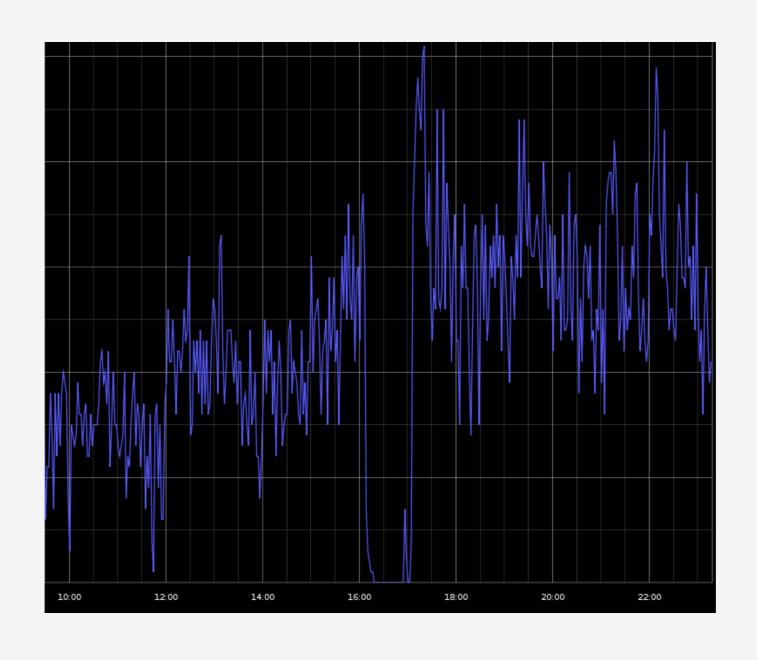
```
assert(p99_latency < 50)
```

assert(p99_latency < 50)

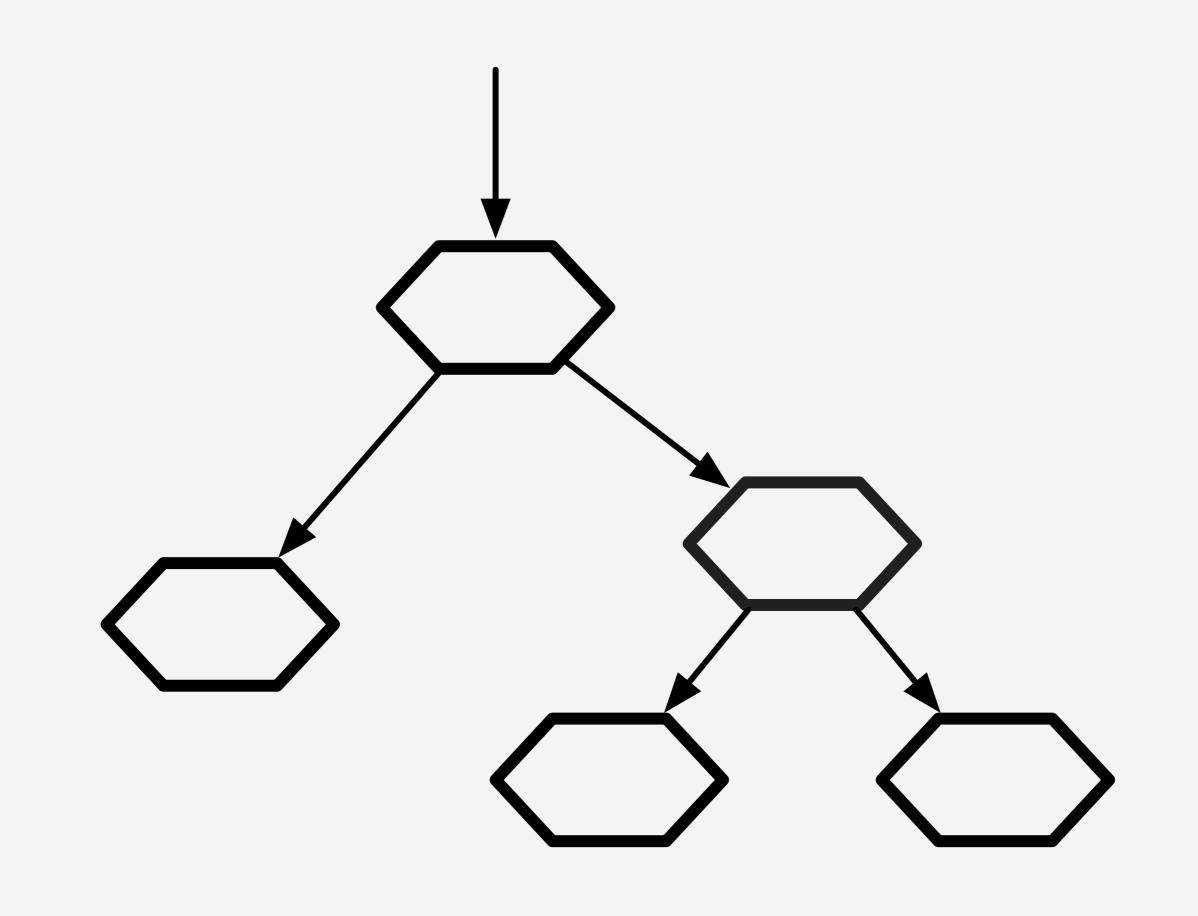


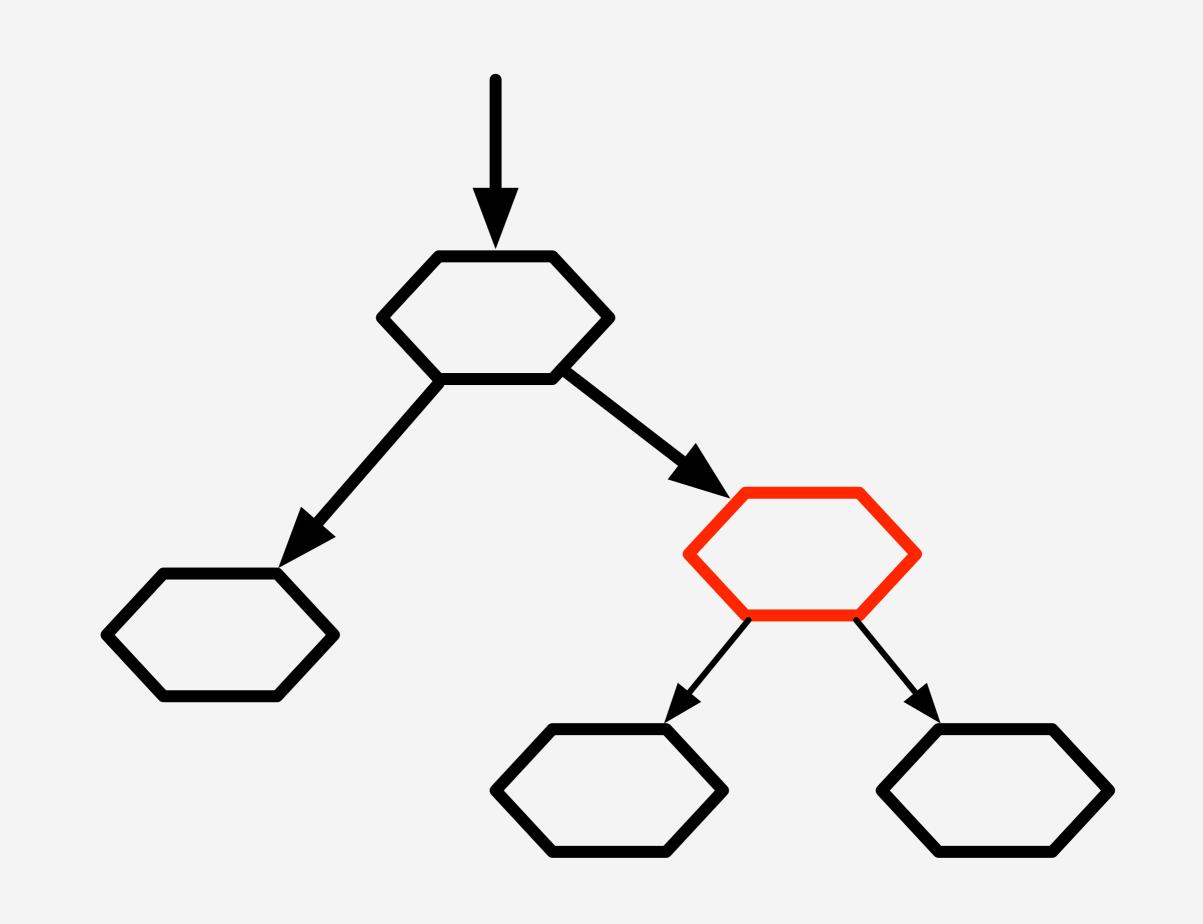
```
assert(active_cons > 10)
```

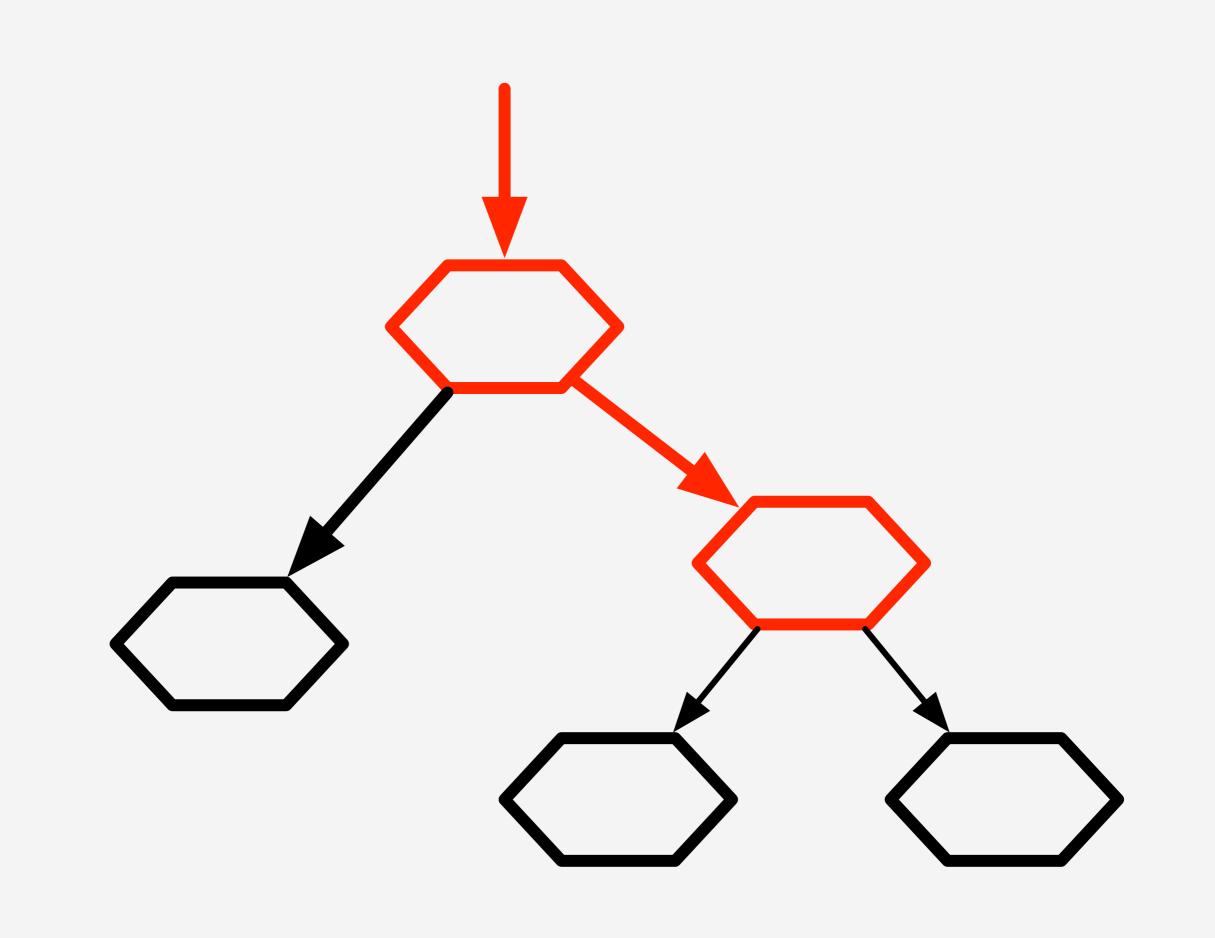
assert(active_cons > 10)

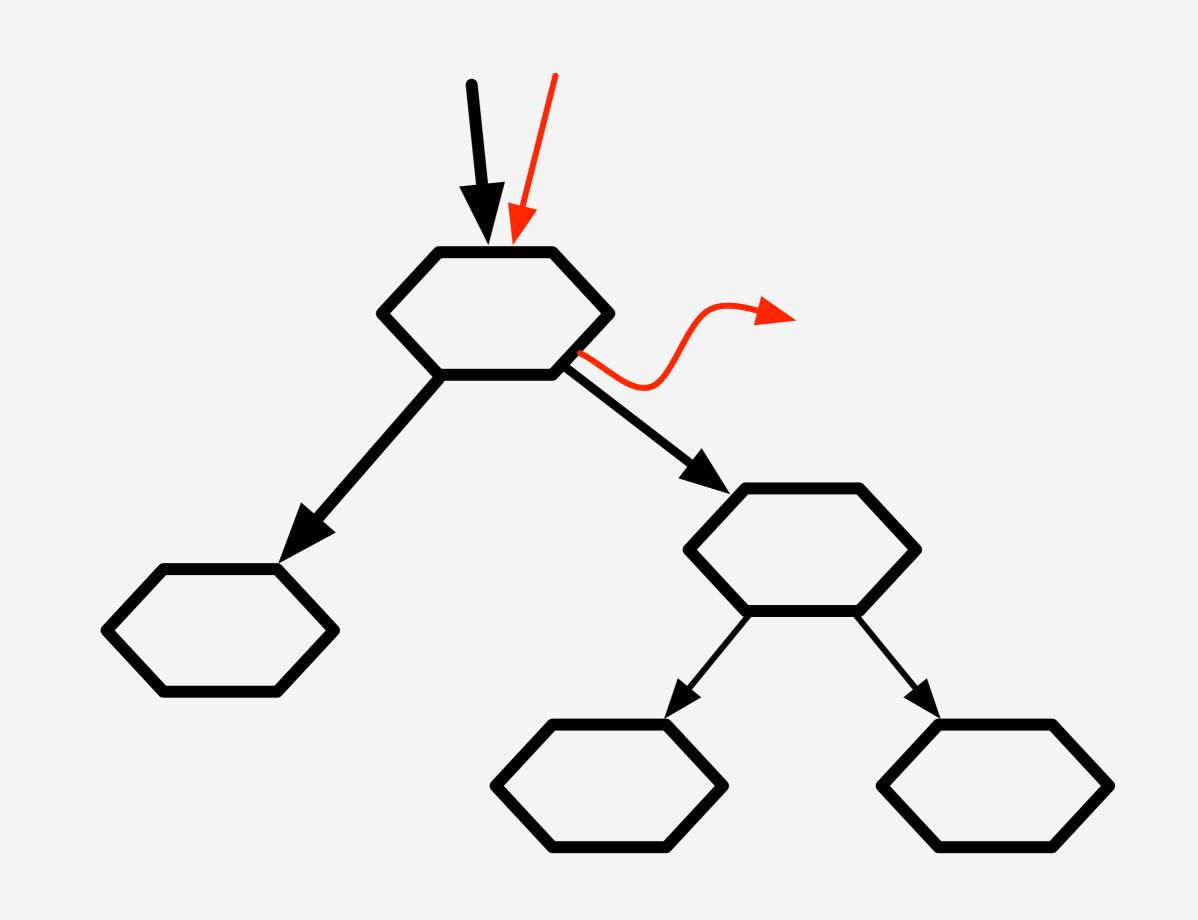


Flow control & Backpressure









```
limit("publish", 10) do
Bus.publish(msg)
end
```

```
limit("publish", prate) do
  publish(msg)
end
```

echo 0 >/etc/rates/publish

Evolving Socio-Technical Systems

Architecture Execution

heroku

Thanks

Questions?