

MAELSTROM

MITIGATING DATACENTER-LEVEL DISASTERS BY
DRAINING INDEPENDENT TRAFFIC SAFELY AND EFFECTIVELY

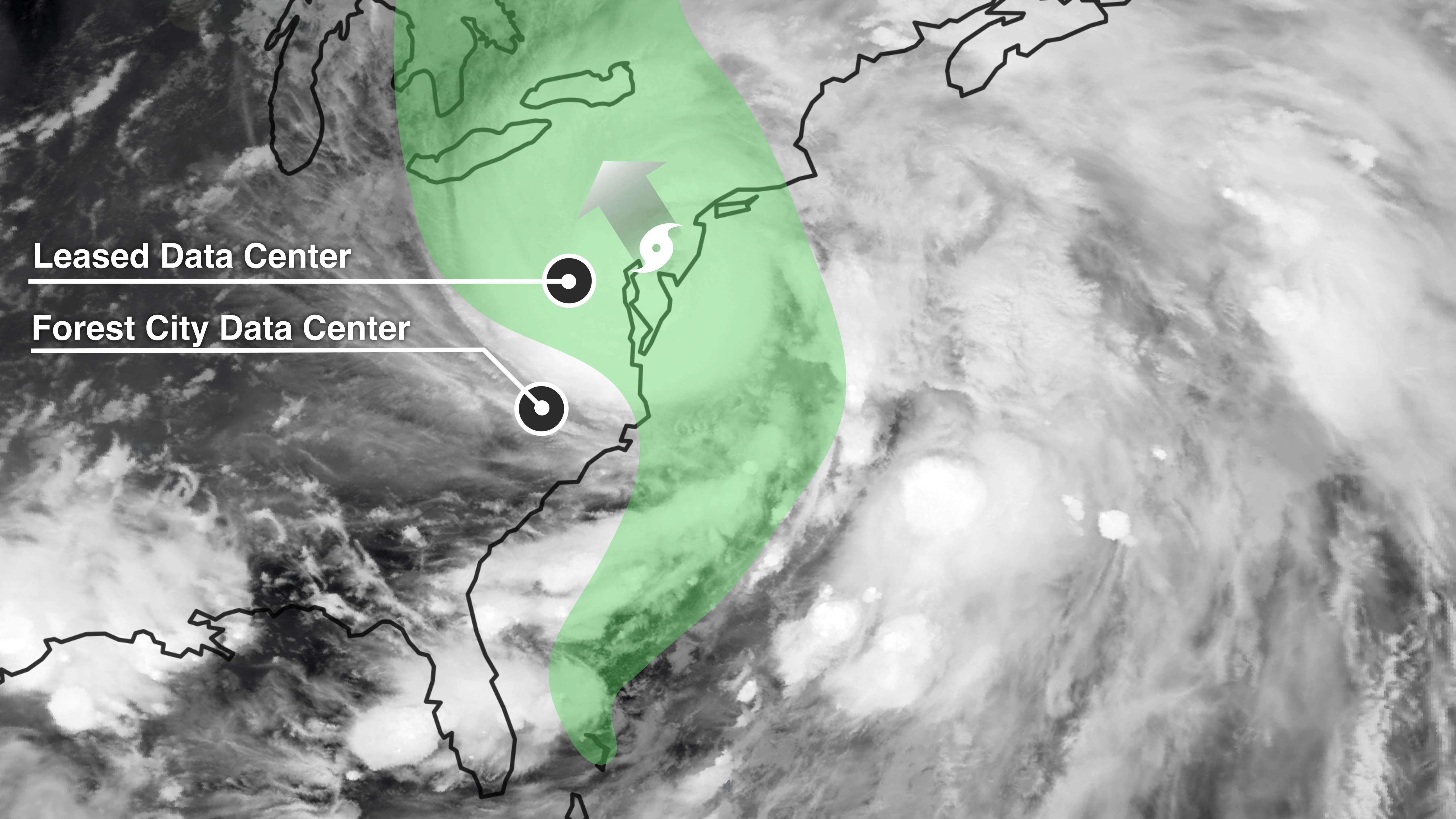
Kaushik Veeraraghavan, **Justin Meza**, Scott Michelson, Sankaralingam Paneerselvam, Alex Gyori, David Chou, Sonia Margulis, Daniel Obenshain, Shruti Padmanabha, Ashish Shah, Yee Jiun Song, Tianyin Xu*

Facebook, Inc. *UIUC

Forest City Data Center



But...



Leased Data Center

Forest City Data Center

Leased Data Center

Forest City Data Center

**Can we run Facebook
without one of our
data centers?**

An aerial photograph of a river with rapids and a waterfall. The water is dark and turbulent, with white foam from the rapids and waterfall. The background is a dark, textured surface, possibly a forest or a rocky terrain.

Fault tolerance without compromising correctness

Must track dependencies between micro services

Testing to verify dependencies and configuration

⑥ check Cactus

⑦ check Quasar

① Status of Network

④ War Room Testing
set up VIP or letu/hosts

Draining

② Start Testing Services

⑤ Employee Testing
set up latest

Unicorn

⑥ User Traffic - 1%

Dragon

⑦ Regional Load Test (Monday?)

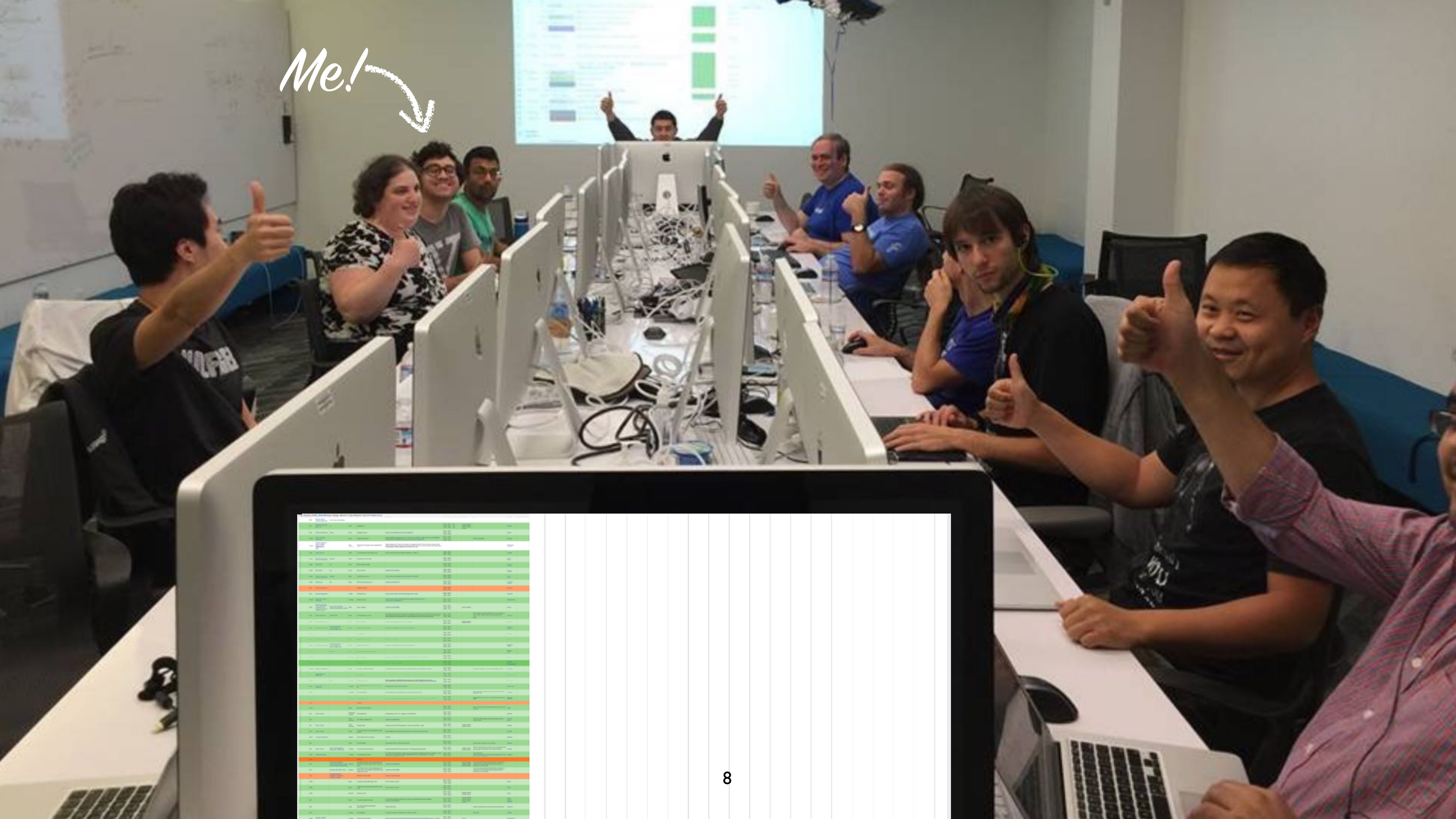
Search (Bandwidth Permitting)

Runbooks

⑧ All other services
all Nimble speak one

if significant bandwidth, coordinate in #encastorm

Me! →

A large data table is displayed on a computer monitor. The table has multiple columns and rows. The rows are color-coded: most are green, but several rows are highlighted in orange. The table appears to be a list of items or transactions with various numerical and text data points.

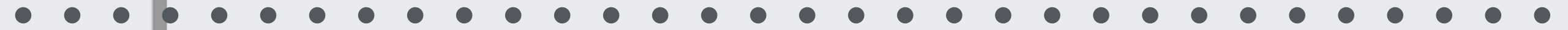
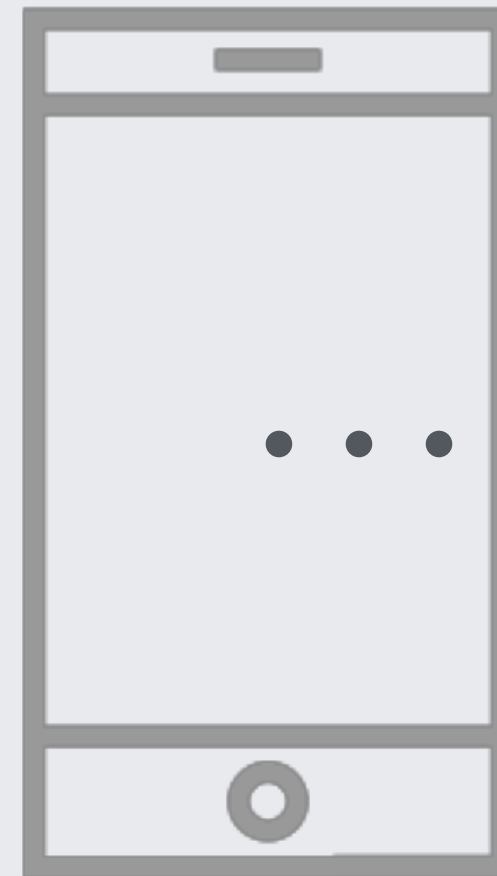


MAELSTROM

Drains traffic from data centers

Minimal operator interaction

Used in tests and production events



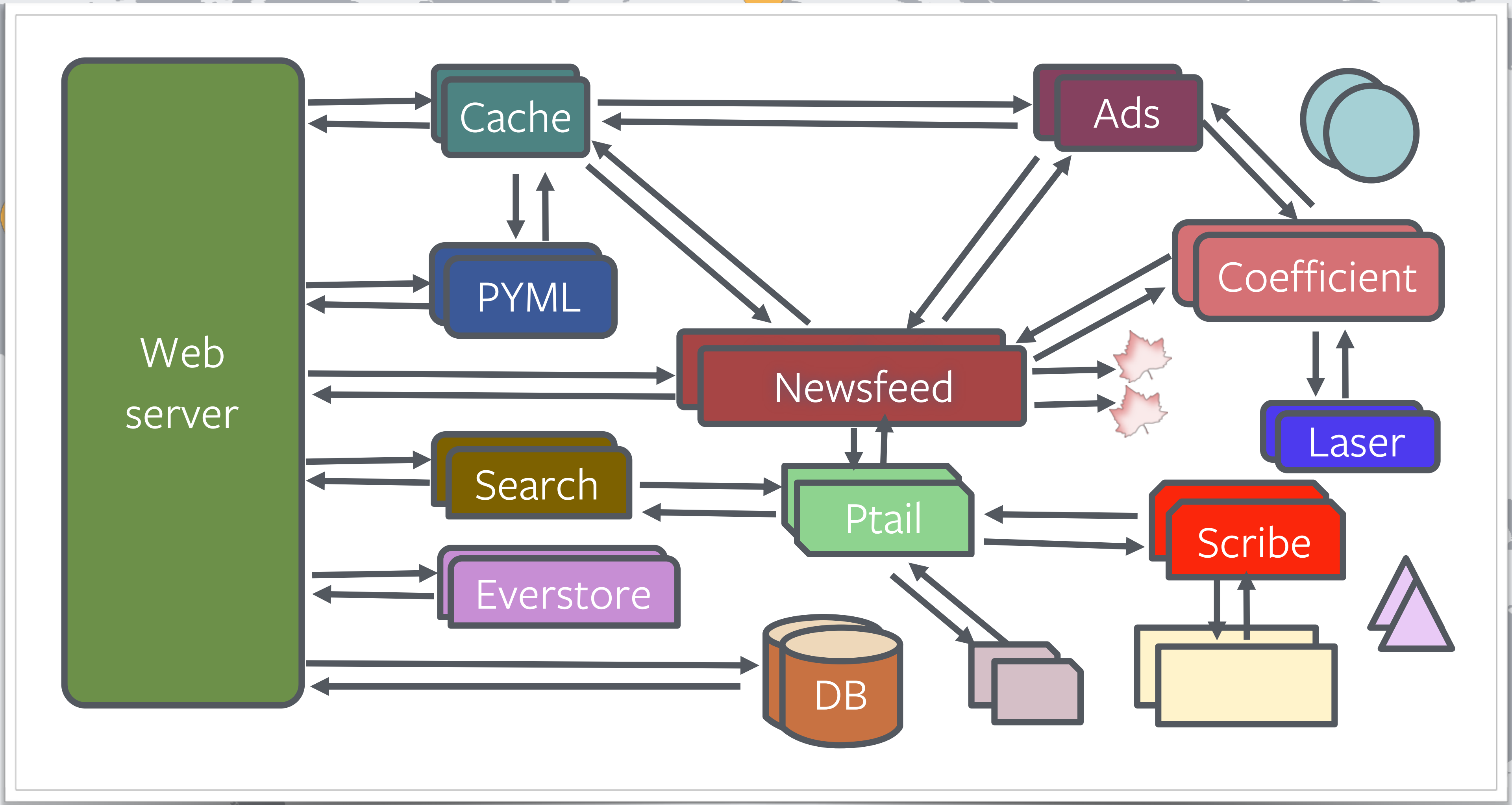




~~Monolithic
system~~



Microservices



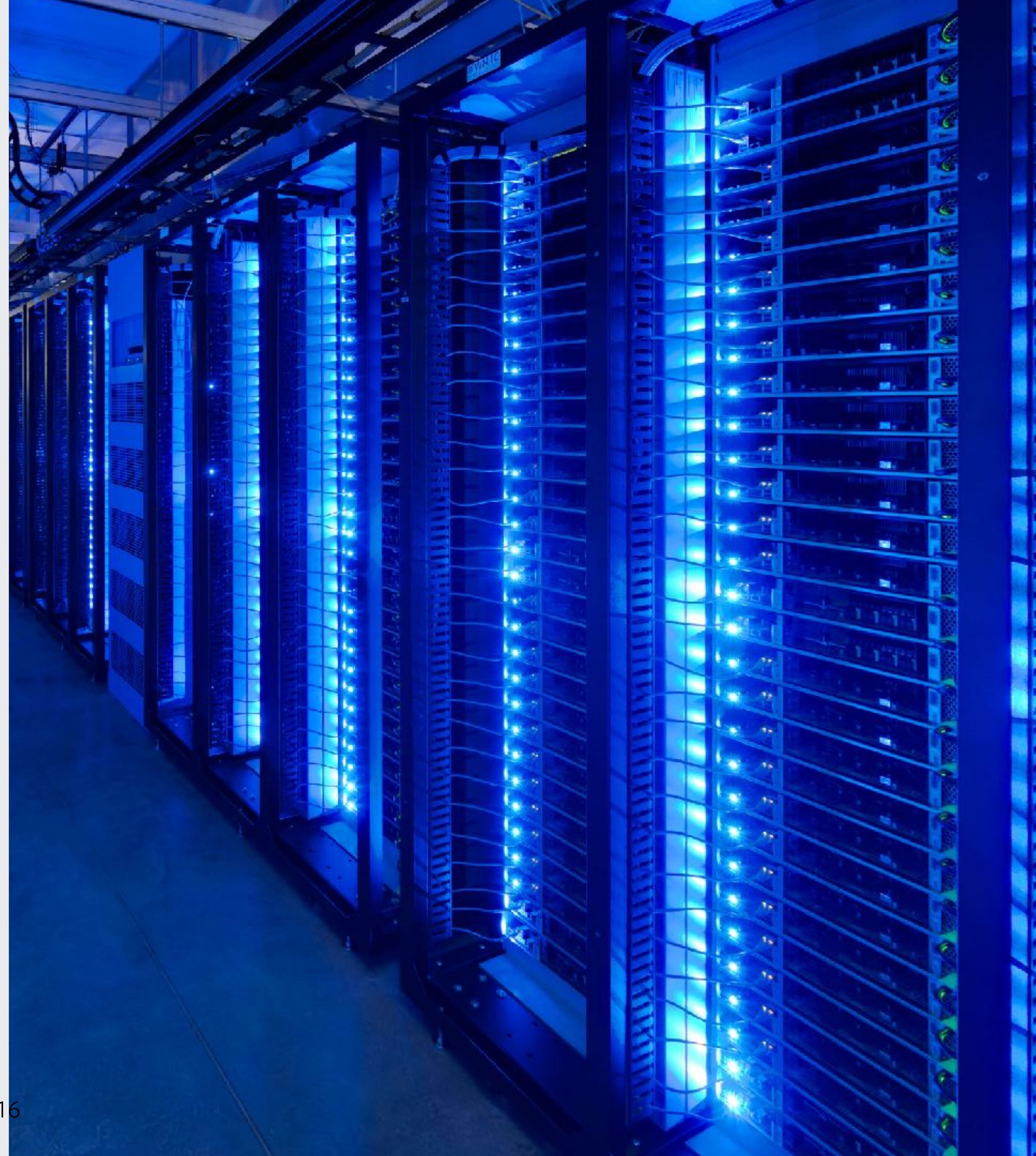
Traffic types

Stateless

Stateful

Sticky

Replication

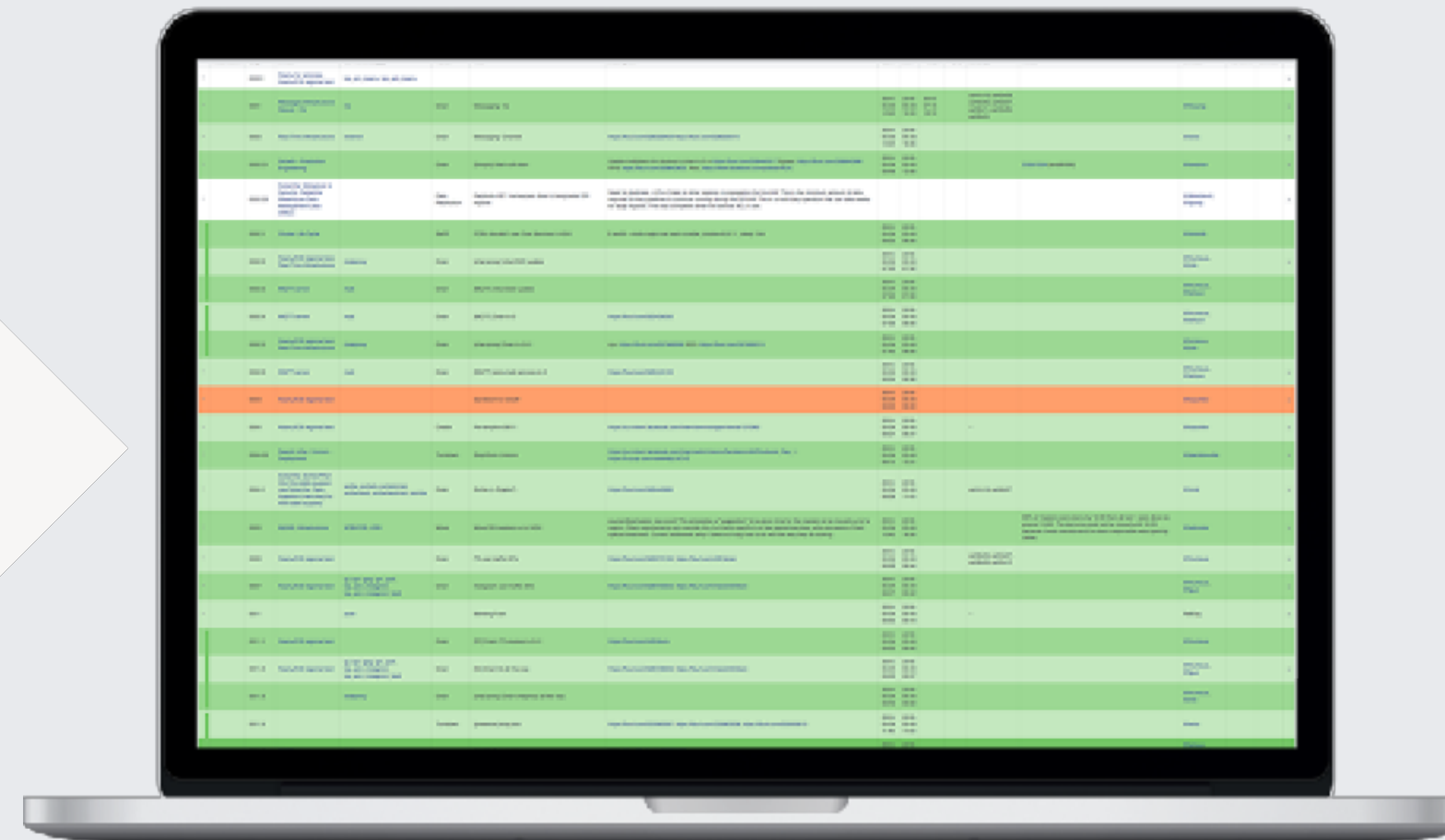


TEMPLATE

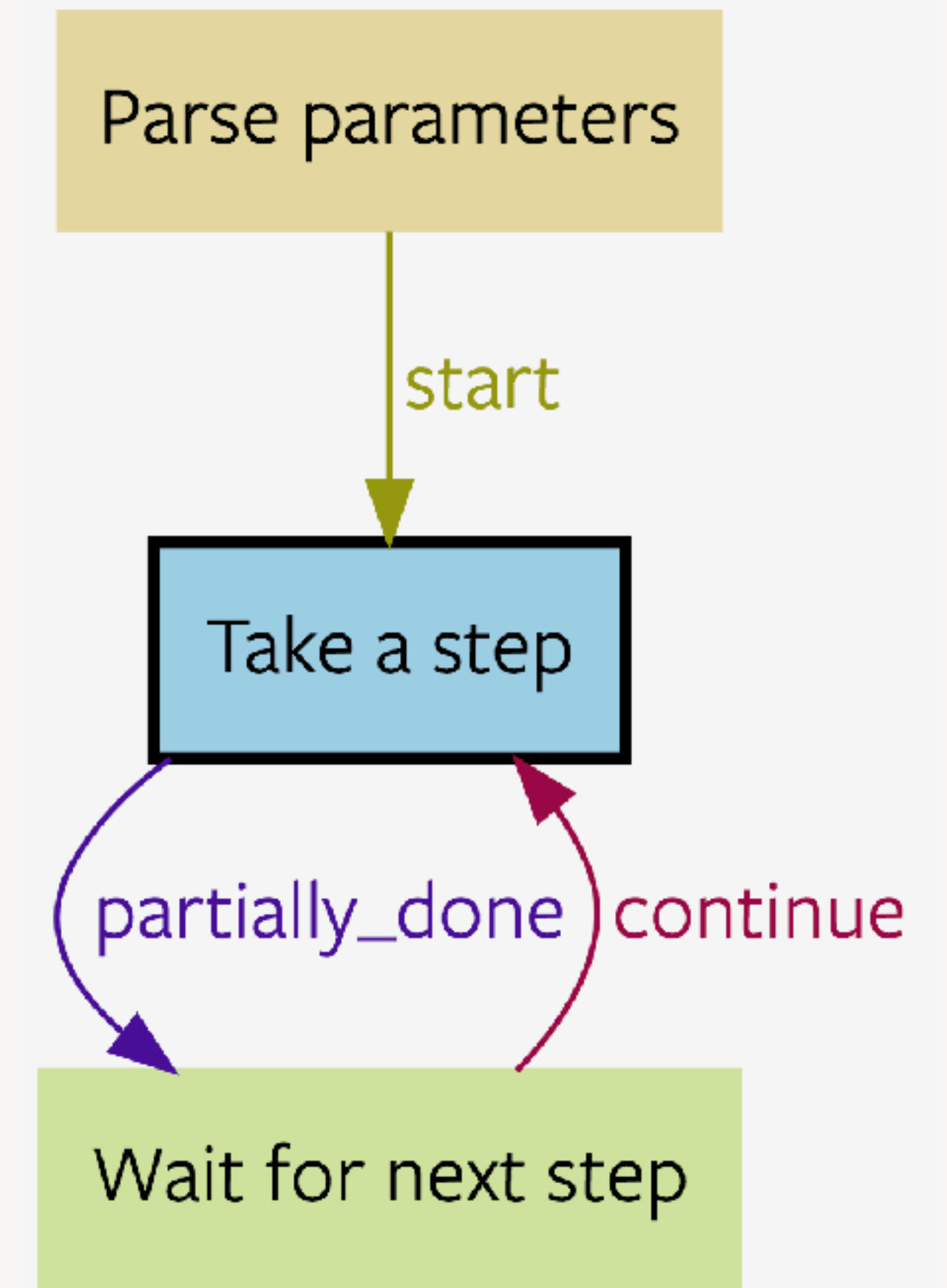
Task TrafficShift

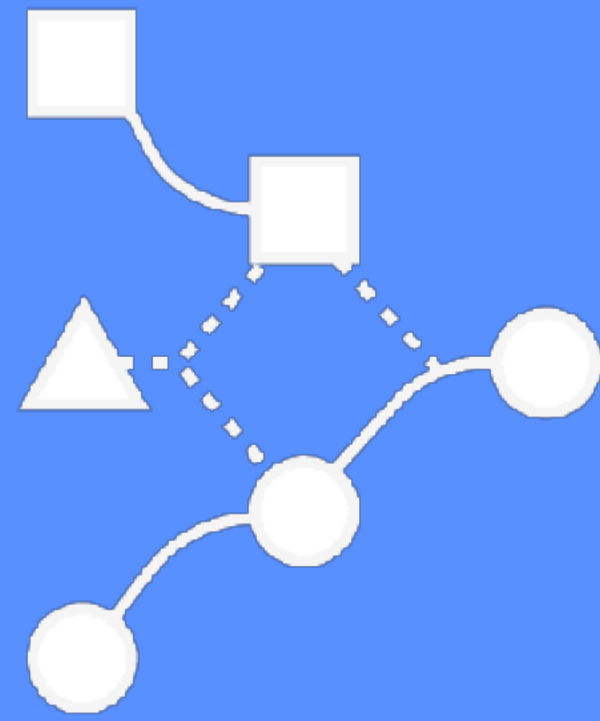
```
{  
  vip_type: <virtualIP>,  
  target: <datacenter>,  
  ratio: <amount>  
}
```

RUNBOOK

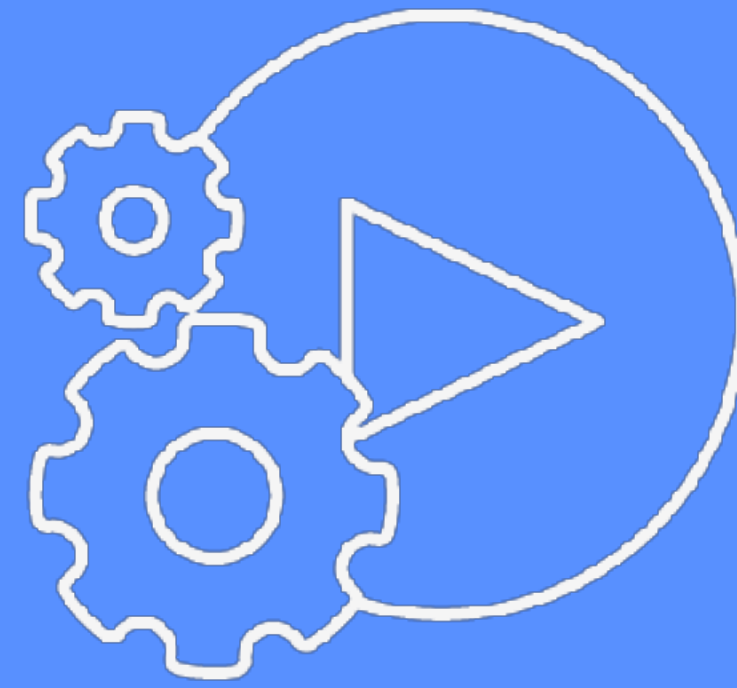


EXECUTION





DEPENDENCE



AUTOMATION



VALIDATION



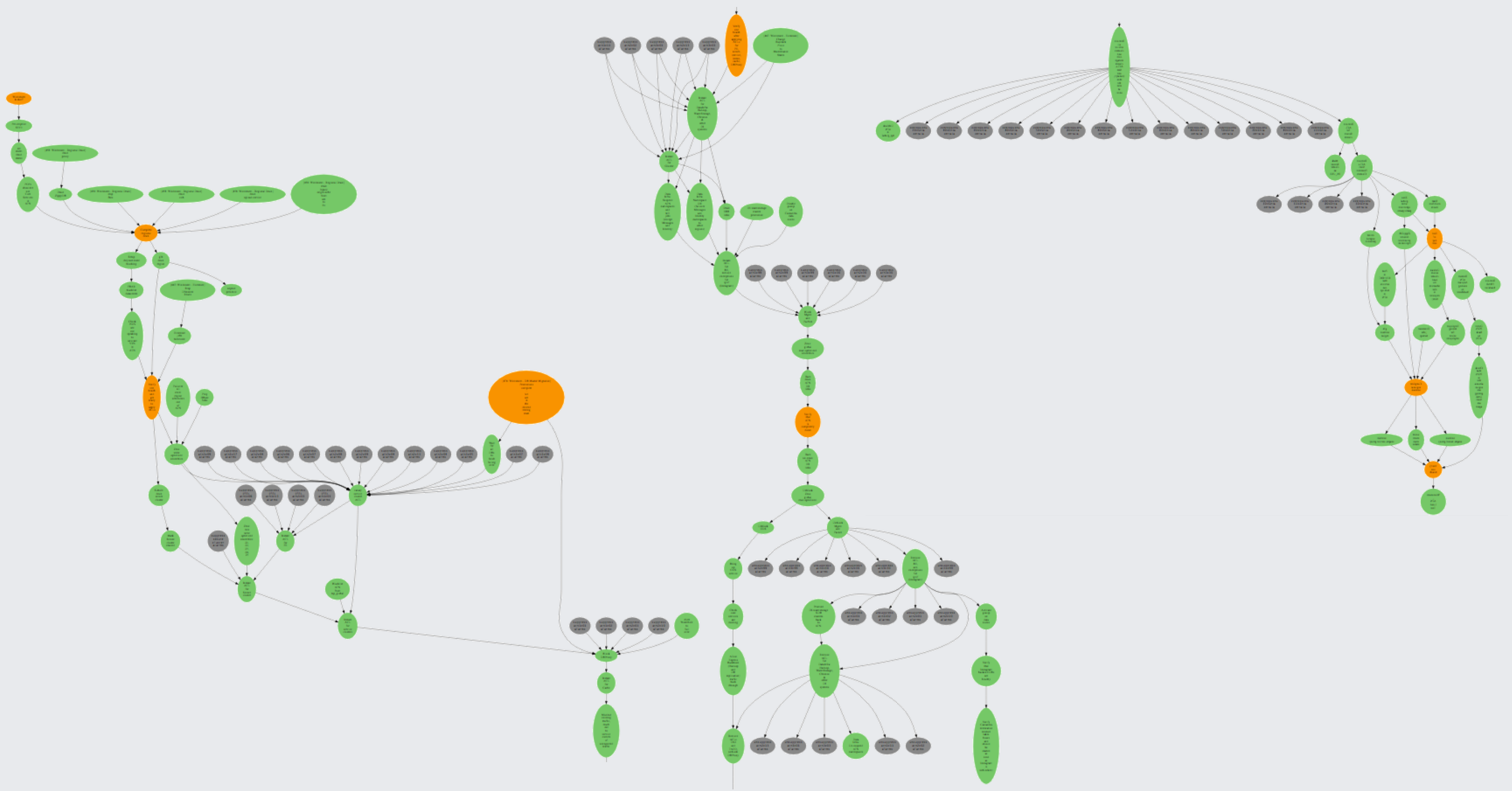
DEPENDENCE

Captures service dependence

Directed acyclic graph

Manually curated in runbooks







AUTOMATION

Traffic shifting

Job manipulation

Configuration changes

Shard movement





VALIDATION

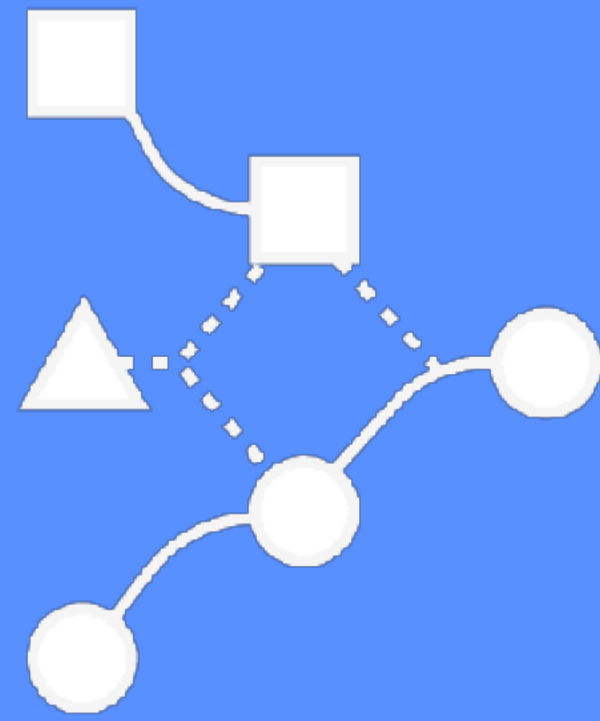
Metrics and thresholds

When to launch a step

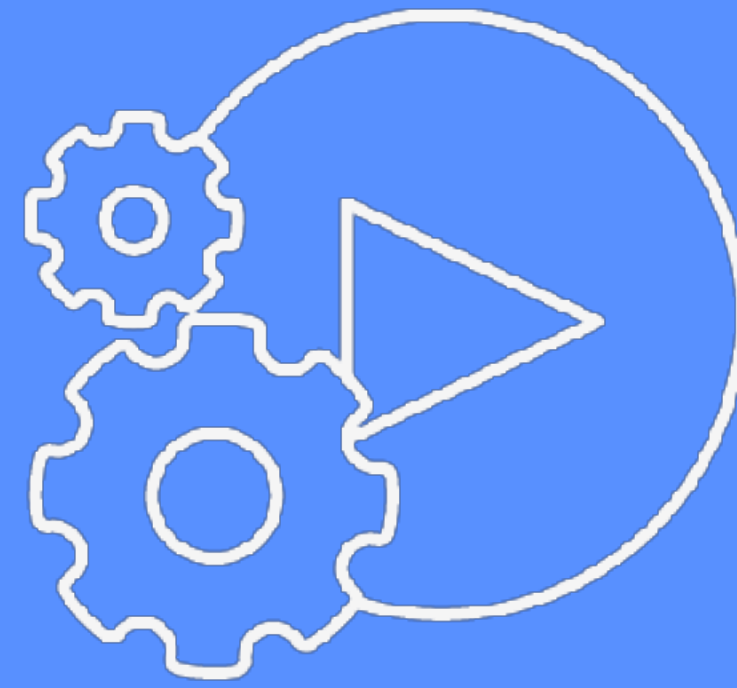
When a step has completed

Service health checks





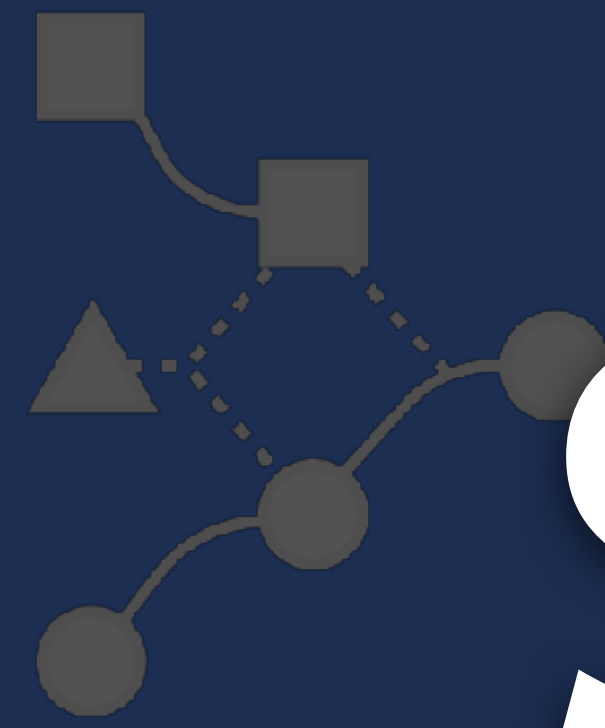
DEPENDENCE



AUTOMATION



VALIDATION



Scheduler

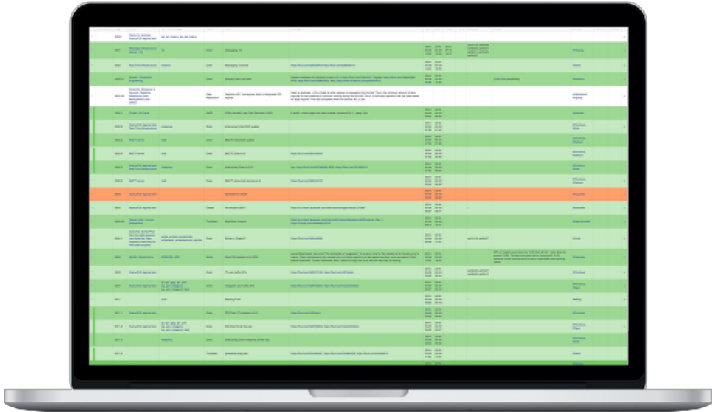


DEPENDENCE

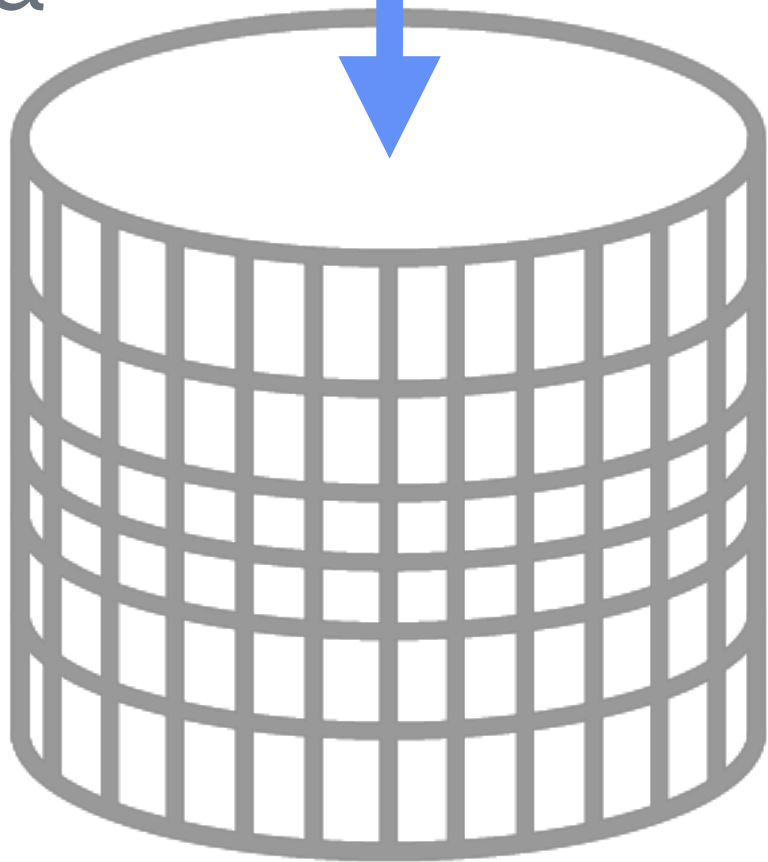
AUTOMATION

VALIDATION

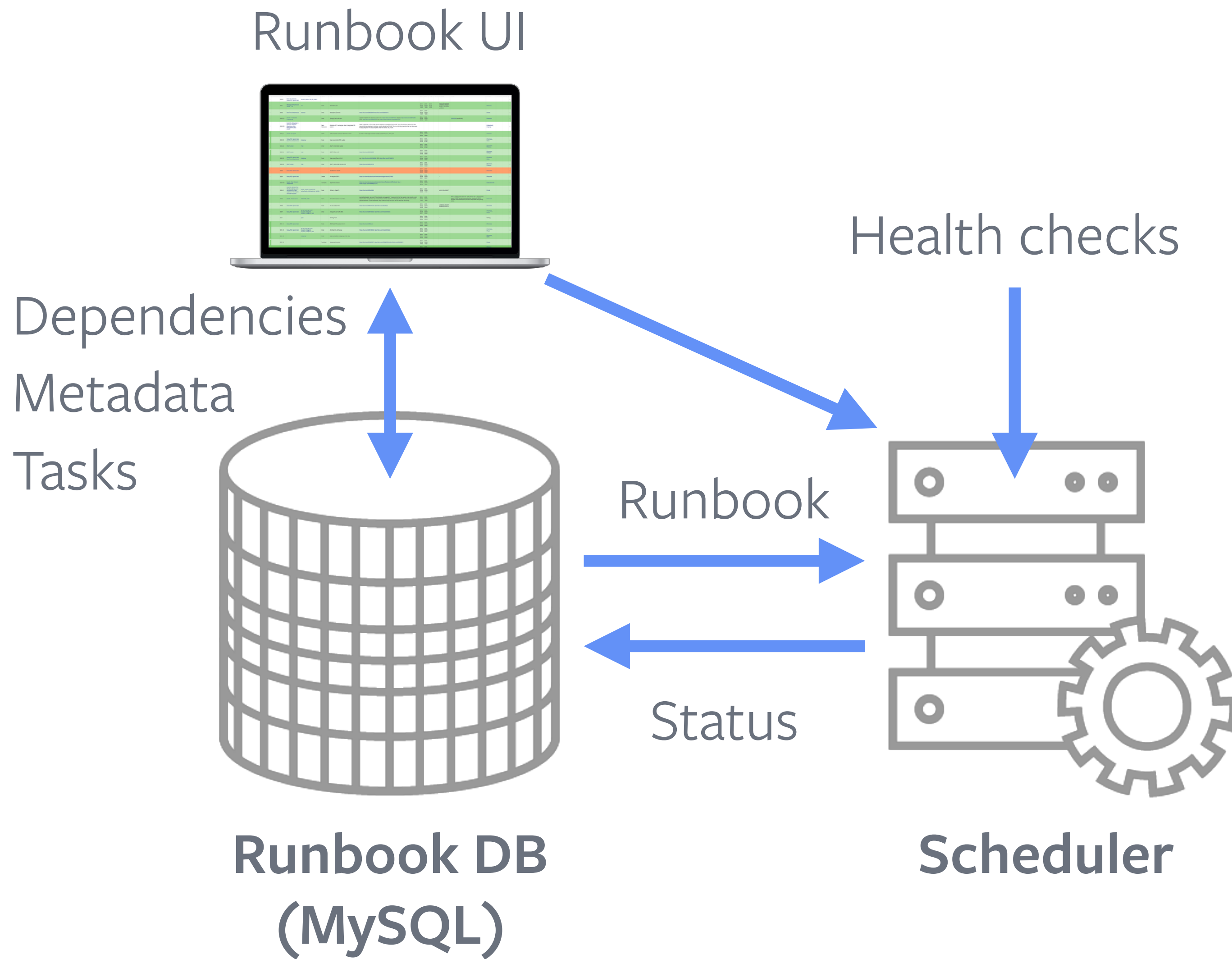
Runbook UI

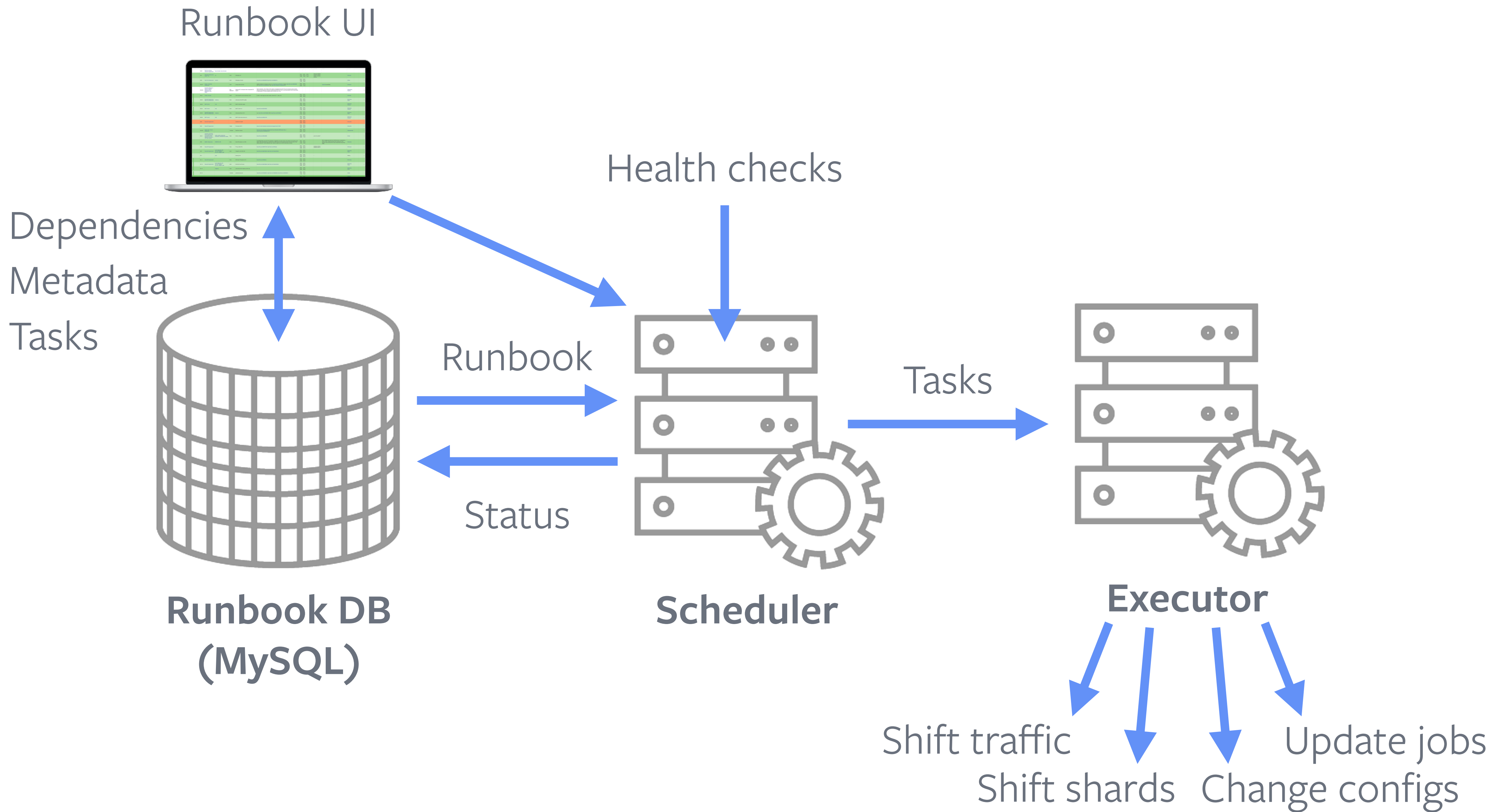


Dependencies
Metadata
Tasks



Runbook DB
(MySQL)





<input checked="" type="checkbox"/>	Cluster Traffic		...	Drain snapshots-formal-proxy-regional					-	x	link
<input checked="" type="checkbox"/>	Cluster Traffic		...	Drain snapshots-service-proxy-regional					-	x	link
<input checked="" type="checkbox"/>	Topology		...	Restart snapshots-utility-norm					-	x	link
<input checked="" type="checkbox"/>	Cluster Traffic		...	Drain thum-gateway-regional					-	x	link
<input checked="" type="checkbox"/>	Cluster Traffic		...	Drain thum-gateway-stage-regional					-	x	link
<input checked="" type="checkbox"/>	Cluster Traffic		...	Drain vltv-regional-proxy					-	x	link
<input checked="" type="checkbox"/>	Enable Config Change			Disable team's automated cluster test testing					-	x	link
<input type="checkbox"/>	Configure Route	...		Disable CI path					-	x	link
<input checked="" type="checkbox"/>	Cluster Traffic		...	Drain upload-whatsapp-proxy-regional					-	x	link
<input checked="" type="checkbox"/>	Cluster Traffic		...	Drain livestream-upload-proxy					-	x	link
<input checked="" type="checkbox"/>	Cluster Traffic		...	Drain livestream-api-upload-proxy					-	x	link
<input checked="" type="checkbox"/>	Cluster Traffic		...	Drain livestream-api-connector-upload-proxy					-	x	link
<input checked="" type="checkbox"/>	Cluster Traffic		...	Drain livestream-instagram-upload-proxy					-	x	link
<input checked="" type="checkbox"/>	Cluster Traffic		...	Drain livestream-360-upload-proxy					-	x	link
<input checked="" type="checkbox"/>	Cluster Traffic		...	Drain livestream-360-upload-proxy-test-utility					-	x	link
<input checked="" type="checkbox"/>	Cluster Traffic		...	Drain livestream-api-dev-upload-proxy					-	x	link

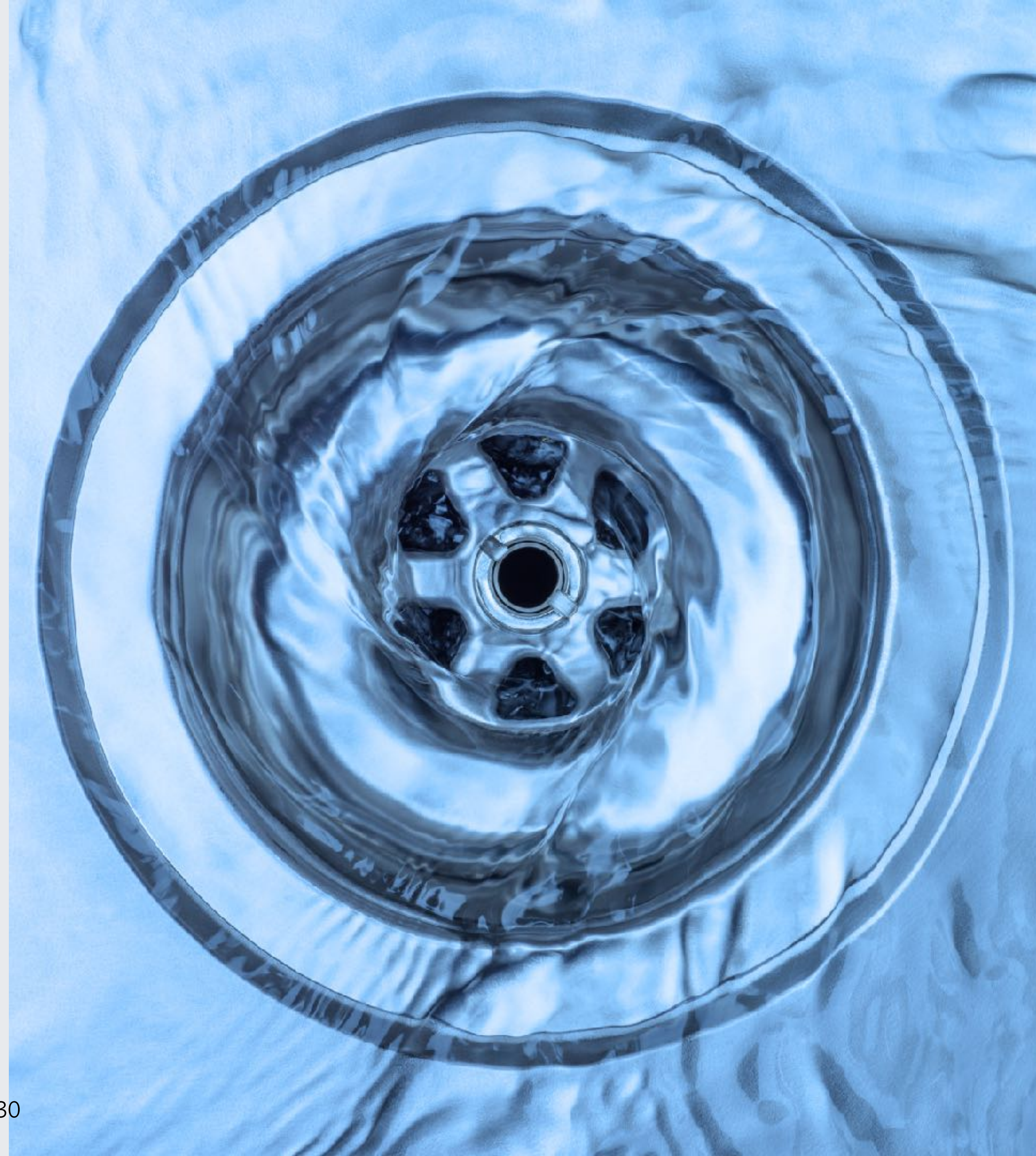
Ready
 Running
 Done

Drain tests

Performed weekly

Verify dependencies

Ensure tools work





Limitations

Must plan for shared resources

Reactive approach to mitigation

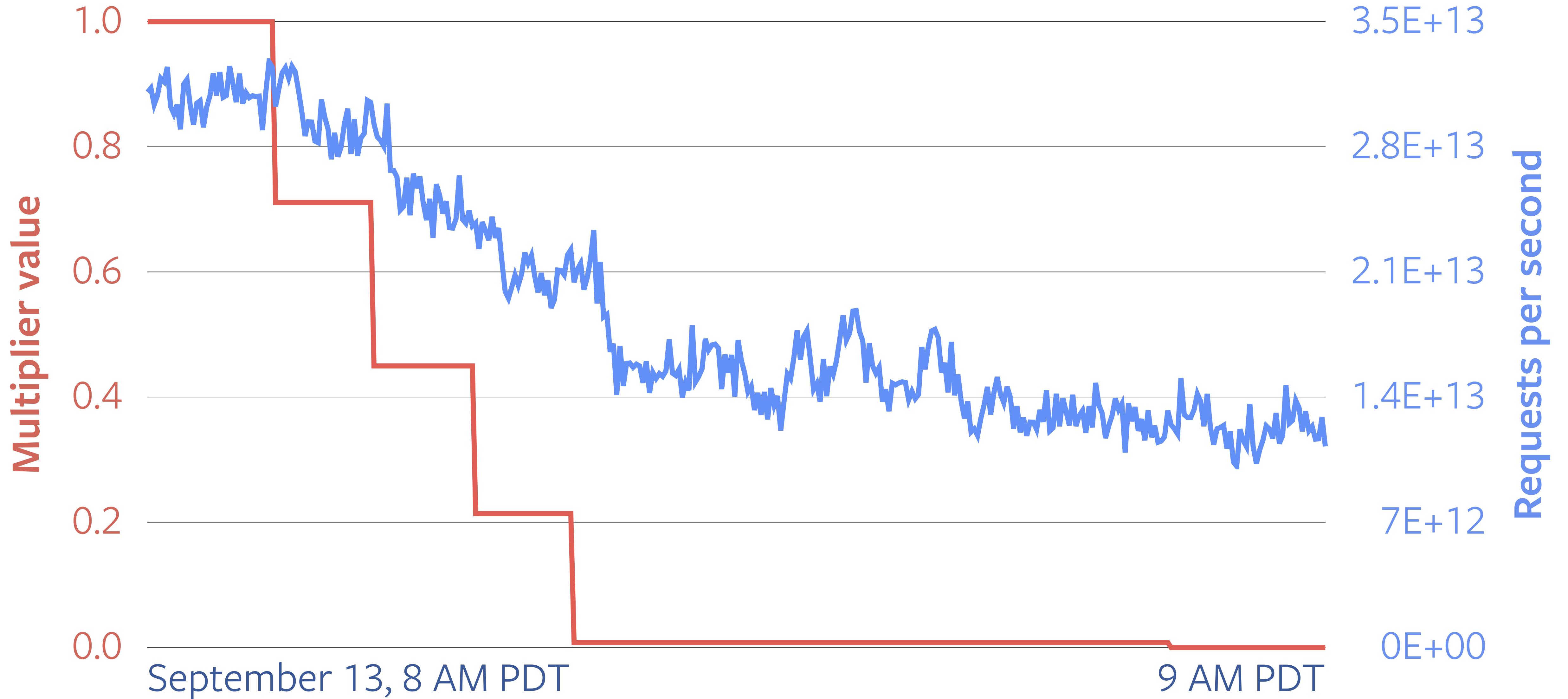
Need to avoid cascading failures

Not fastest for every disaster

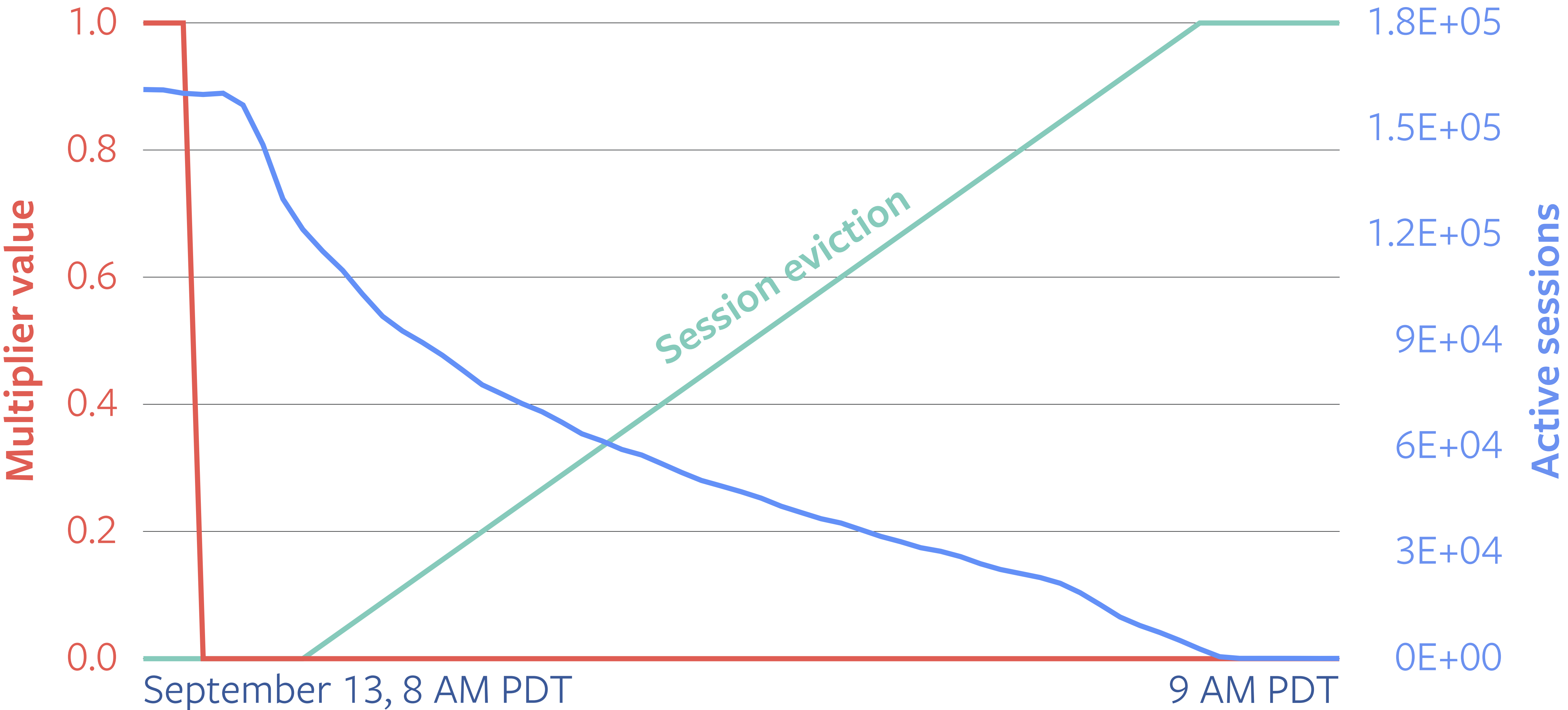
Forest City Data Center



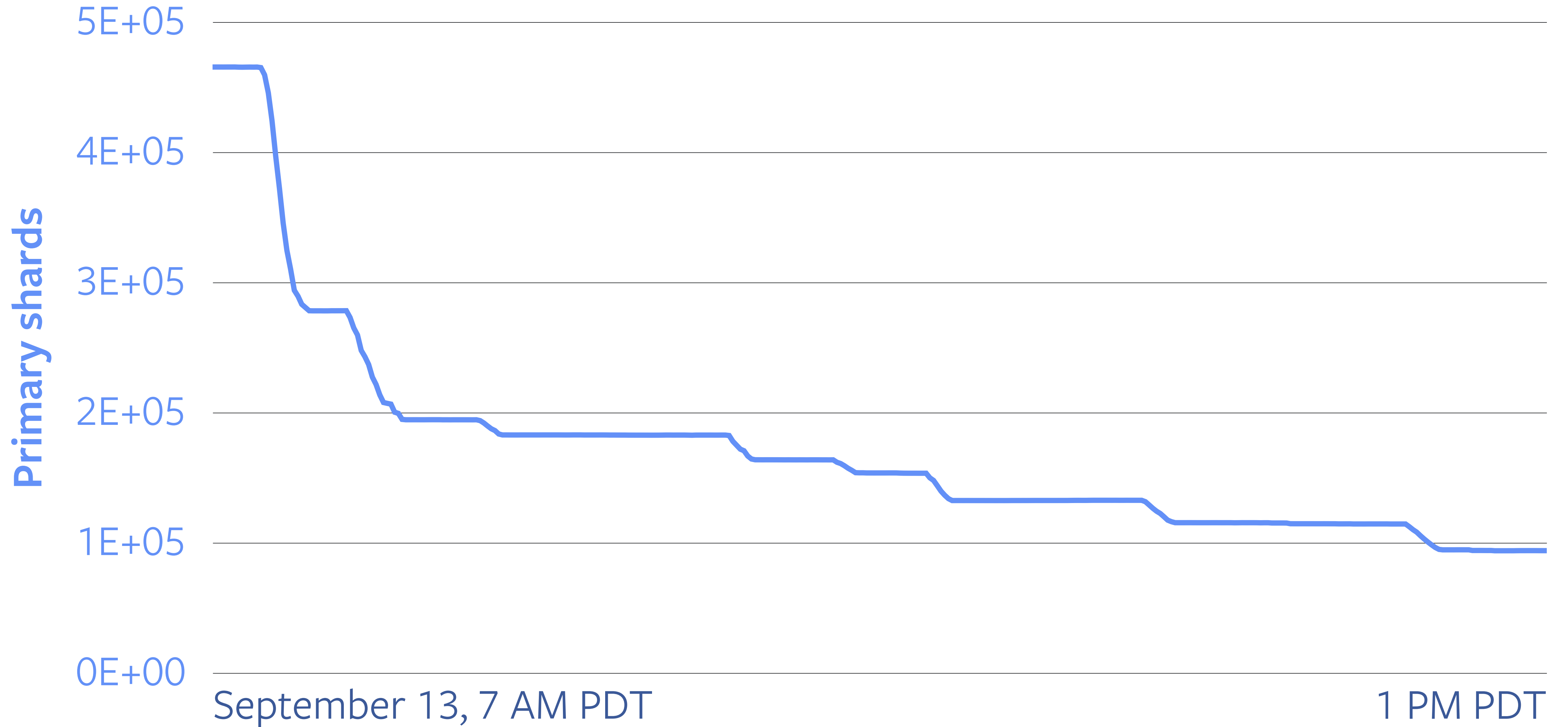
Stateless traffic



Sticky traffic



Stateful traffic





MAELSTROM

Fault tolerance without compromising correctness

Must track dependencies between micro services

Tests verify dependencies and configuration