

Container Evolution



What is CoreOS?



What is CoreOS?

 etcd

 flannel

 rkt

Container is HOT



Compare Search terms ▼

linux container

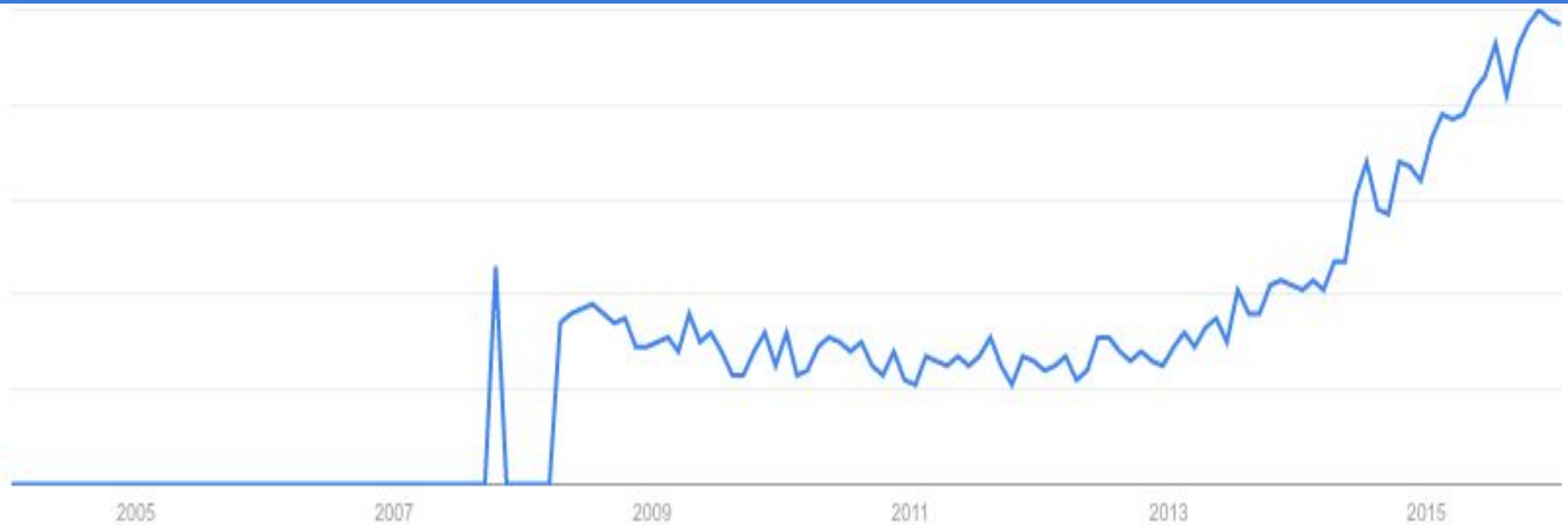
Search term

+ Add term

Interest over time ?

News headlines ? Forecast ?

Container is not a new technology



</>

FreeBSD ~ 2000



freeBSD®



Solaris Zones ~ 2005

Linux Container

Control group

- CPU
- Memory
- IO
- Devices
- ...

Namespaces

- Network
- IPC
- ProcessID
- ...



Linux

Borg at Google ~ 2005



MESOS

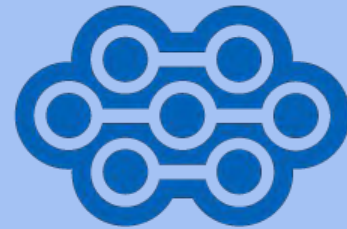
Mesos ~ 2011



heroku

Heroku ~ 2008

dotCloud ~ 2010



dotCloud

Microservices

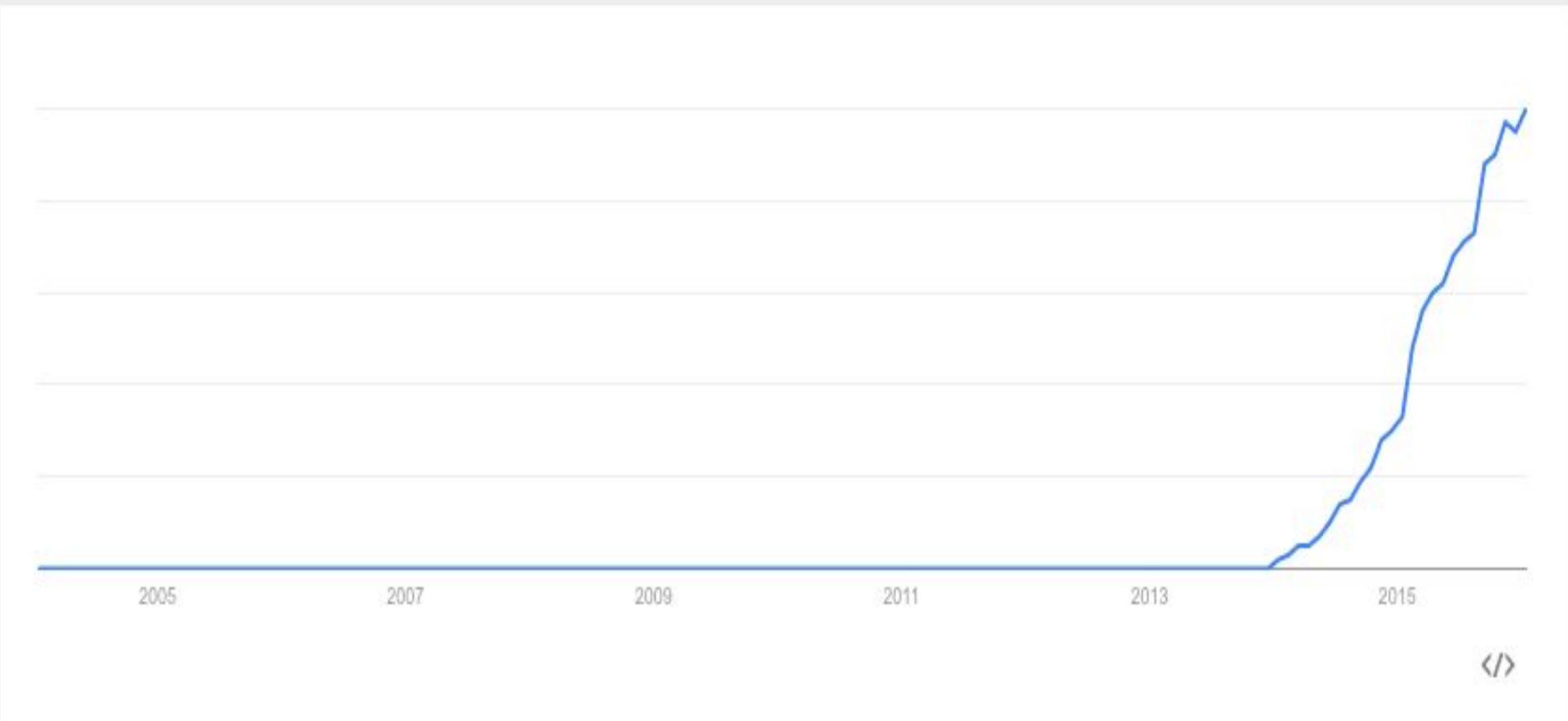
Search term

+ Add term

Microservices

Interest over time ?

News headlines ? Forecast ?

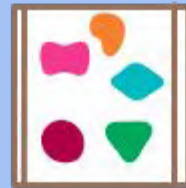


Microservices

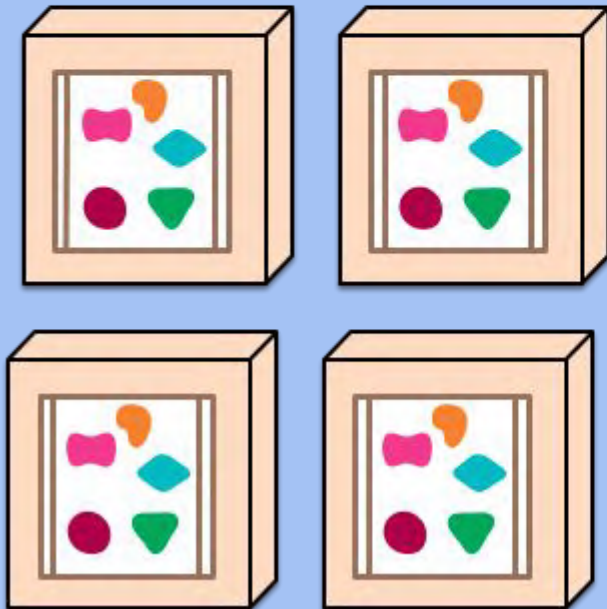
The background of the slide features a complex, abstract network diagram. It consists of numerous dark blue spheres of varying sizes, representing nodes, connected by a web of thin, dark blue lines. The lines crisscross the frame, creating a sense of interconnectedness and complexity. The overall aesthetic is modern and technological, typical of a presentation on cloud computing or software architecture.

Microservices

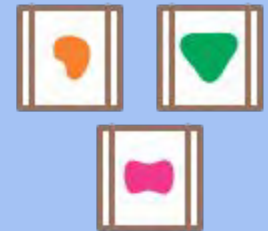
A monolithic application puts all its functionality into a single process...



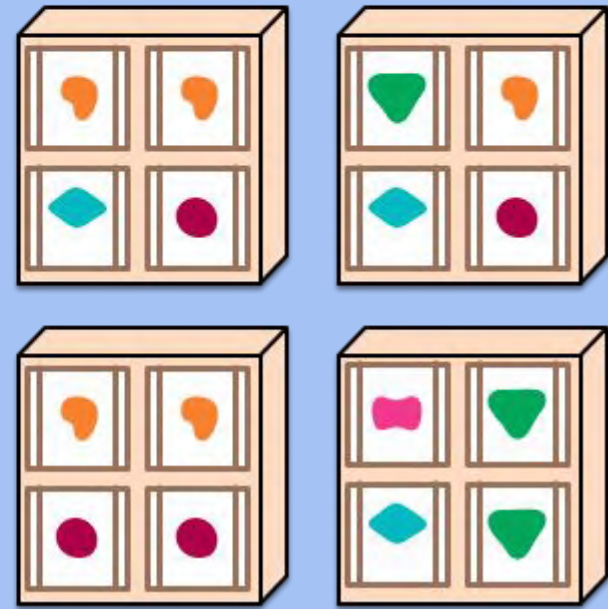
... and scales by replicating the monolith on multiple servers



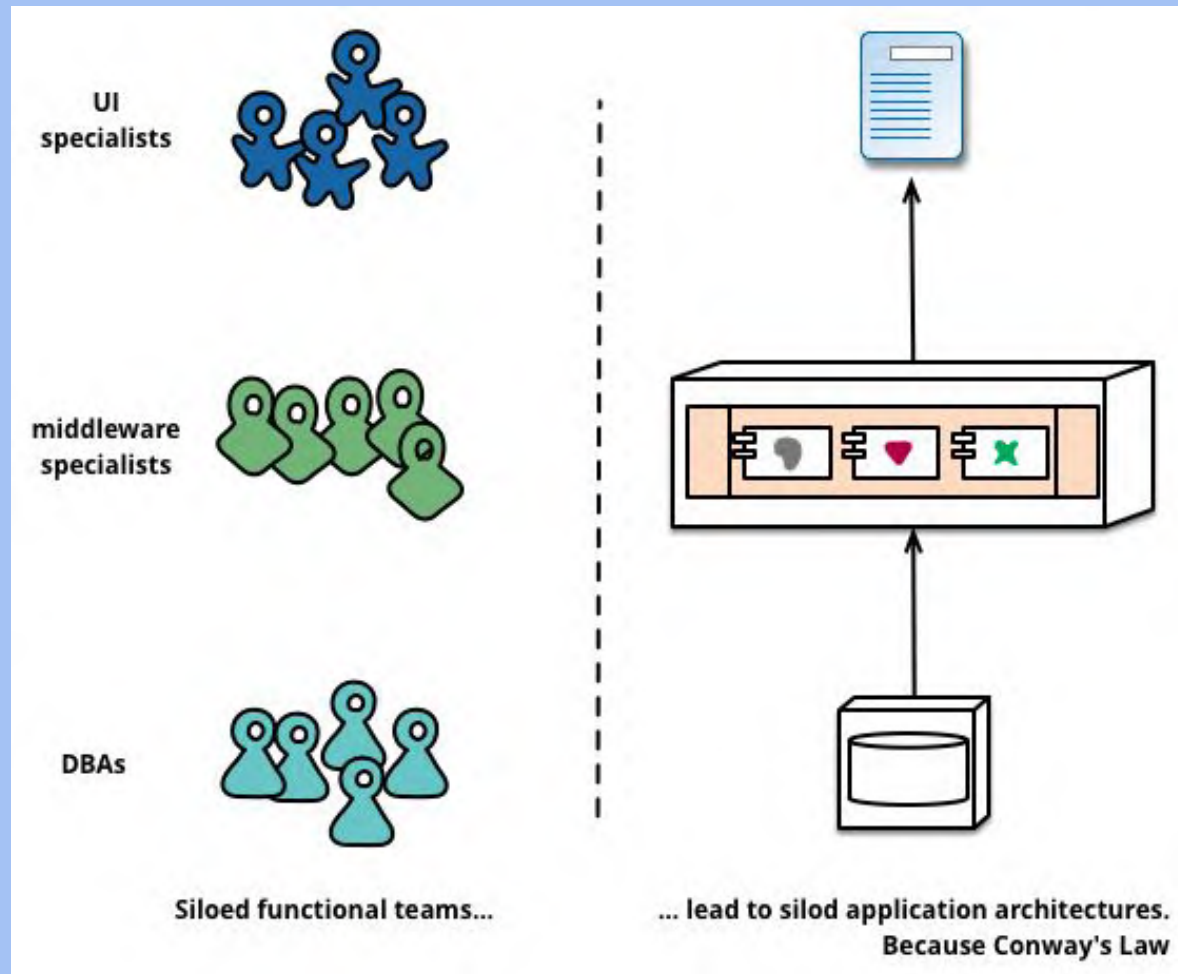
A microservices architecture puts each element of functionality into a separate service...



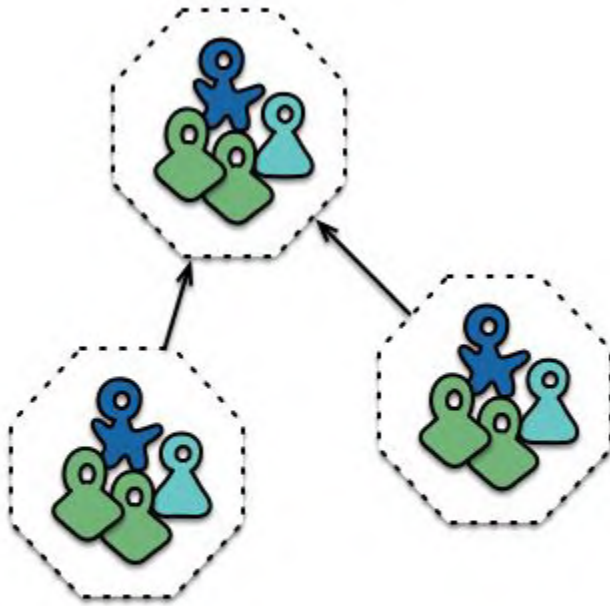
... and scales by distributing these services across servers, replicating as needed.



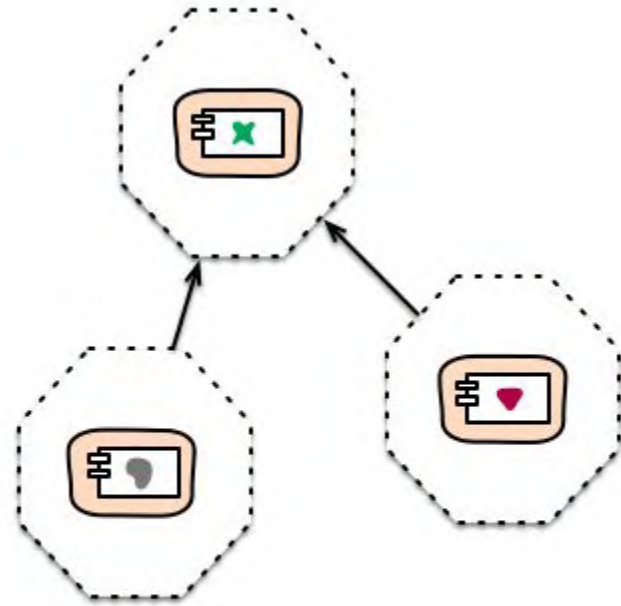
Microservices



Microservices

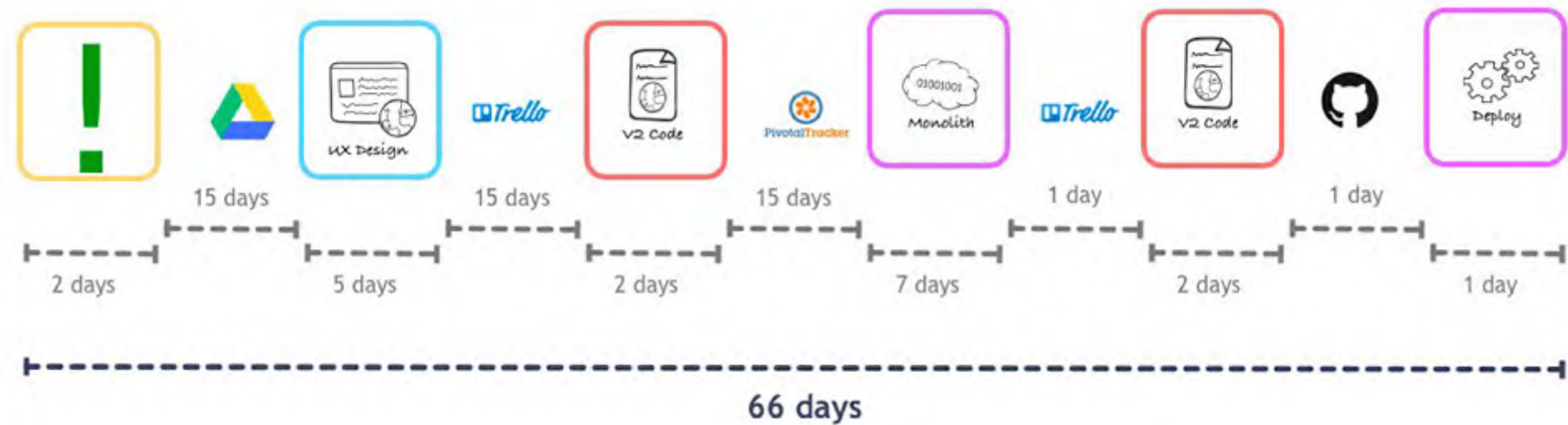


Cross-functional teams...



... organised around capabilities
Because Conway's Law

Microservices



Microservices



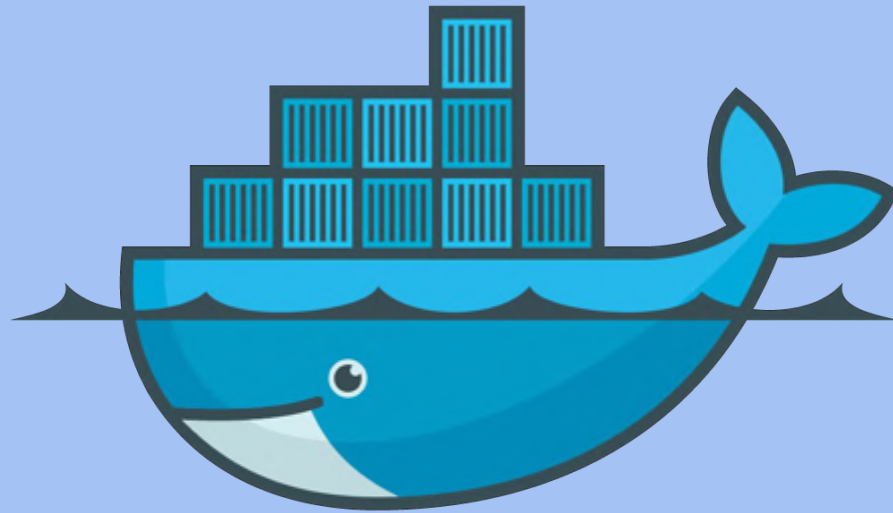
A composite image illustrating the 'Dev vs Ops' problem. The background shows a house engulfed in flames, with firefighters and a fire truck (numbered 38) at the scene. In the foreground, a young girl with brown hair is looking towards the camera with a slight smile. The text 'Ops problem' is overlaid in the top left, 'Dev and Ops' in the middle left, and 'Works fine in Dev' in the bottom left.

Ops problem

Dev and Ops

Works fine in Dev

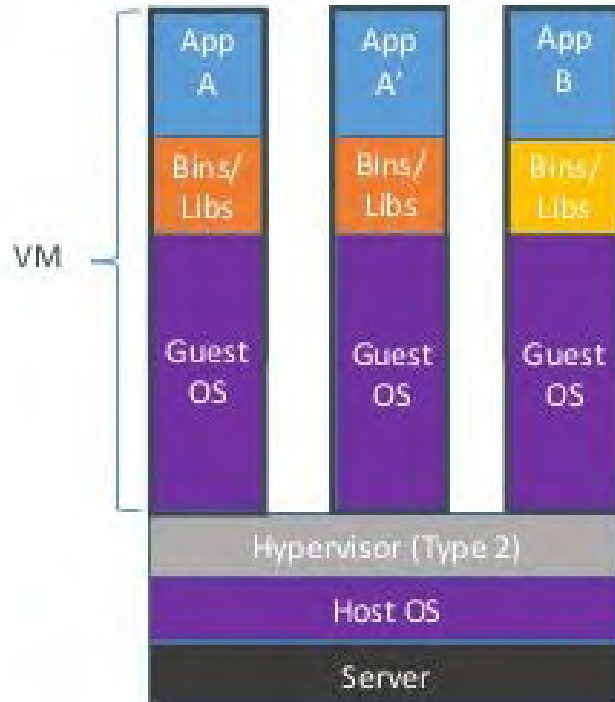
Docker



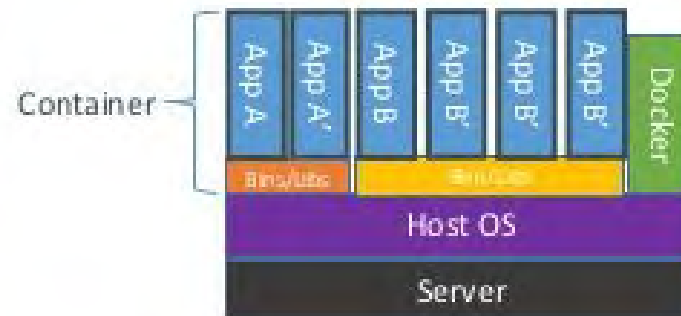
docker

Docker

Containers vs. VMs



Containers are isolated, but share OS and, where appropriate, bins/libraries



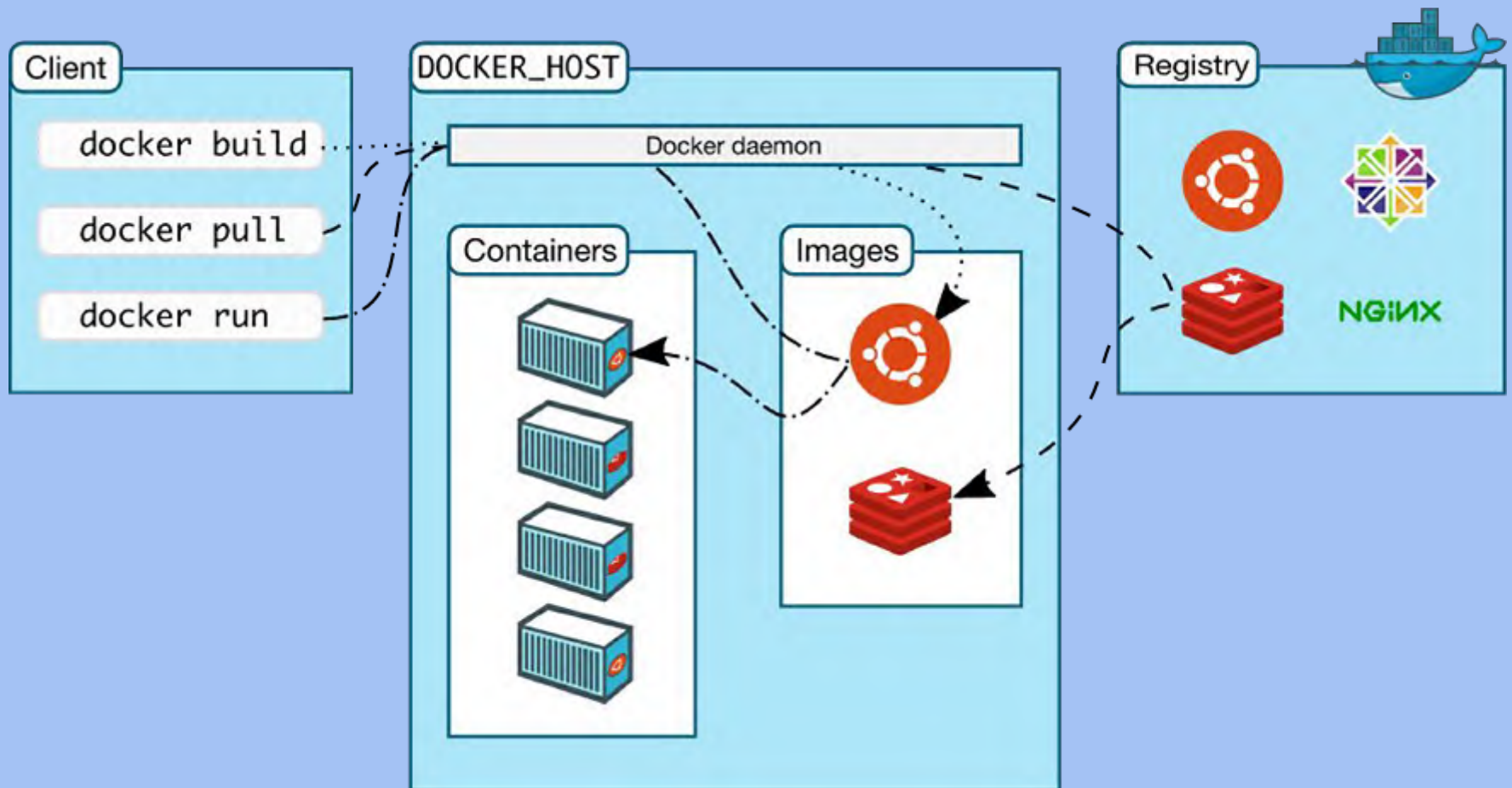


Container ~ 1950

Container ~ 2010



Container Spec



Container Spec

Open Container Specifications

- A container spec
 - config.json
 - runtime.json
 - rootfs

Container Spec

App Container Specifications

- An image spec
 - compressed
 - encrypted
 - signed

Container Spec

OCI and appc Intersections

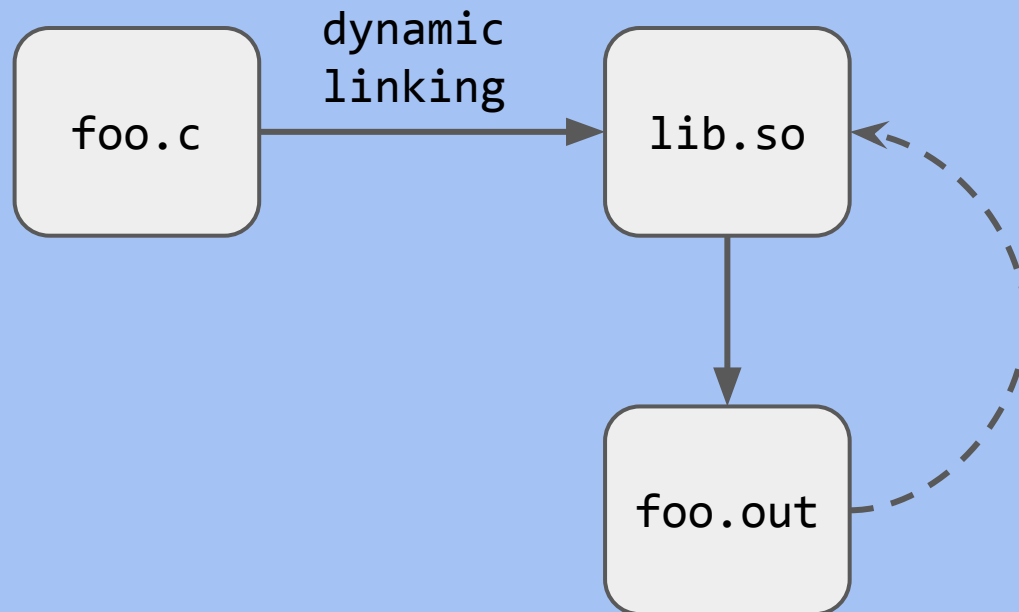
	Standards		Container Runtimes		Portability
	OCI	APPC	DOCKER	RKT	EXAMPLES
Container Image	✗	✓	Docker Format v2	appc Image Spec Docker Format v2	User builds container once, can run in docker or rkt
Image Distribution	✗	✓	Docker Registry Protocol	appc Discovery Spec Docker Registry Protocol	Docker and rkt can share container registry mechanism
Runtime	✓	✓	libcontainer	appc Runtime Spec	Docker and rkt can share exec drivers (impl. using LXC, runc, systemd-nspawn, etc)
On-disk Image Format	✓	✗	Previously unspecified	Previously unspecified	Exec drivers share a metadata format and filesystem layout for how to execute a container

Container Management

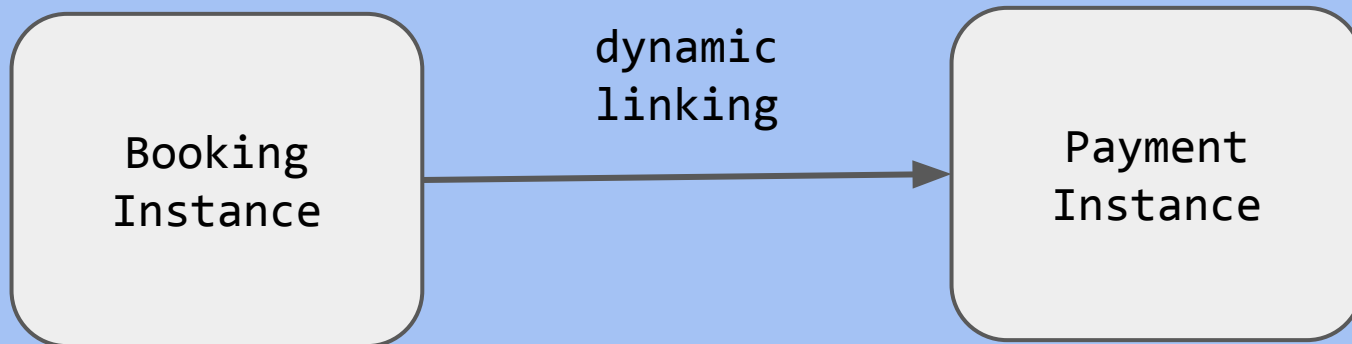


Discovery System

Dynamic Linking



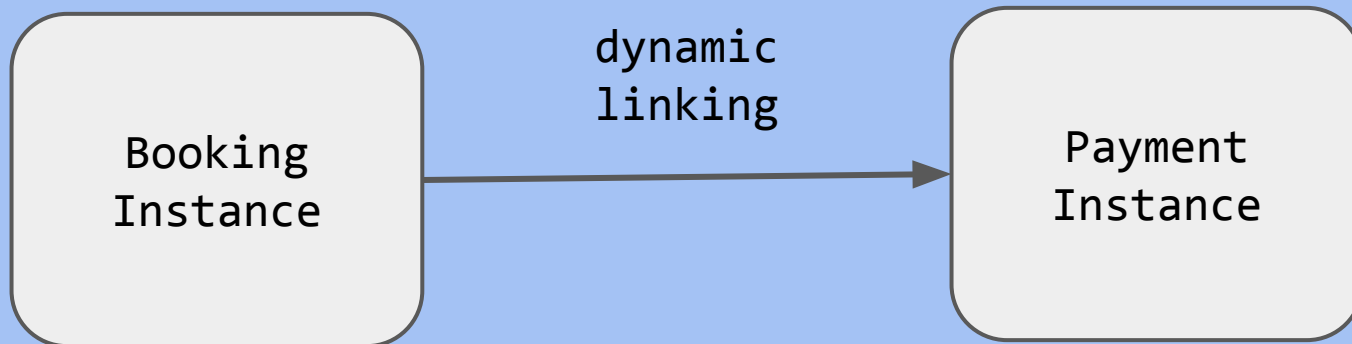
Discovery System



```
find_payment_service()  
{  
    return 10.0.0.1:1234  
}
```

10.0.0.1:1234

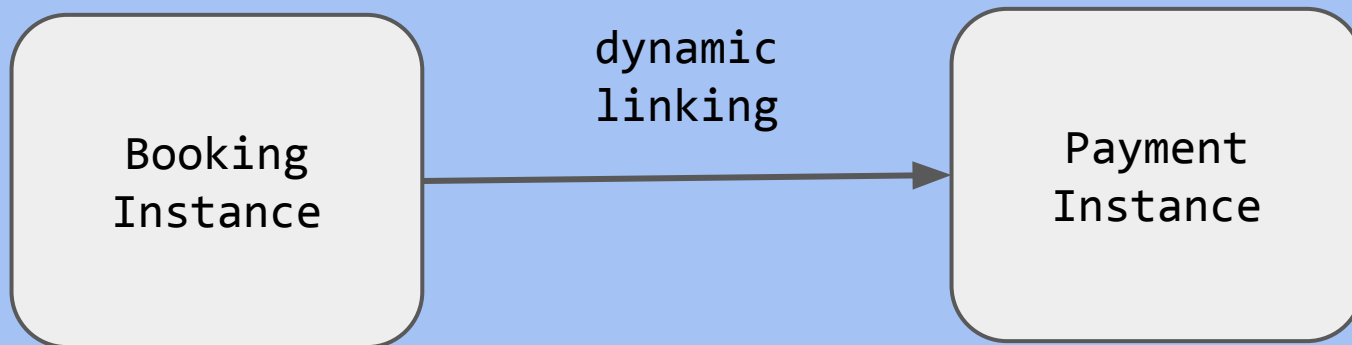
Discovery System



```
find_payment_service()  
{  
    return 10.0.0.100:1234  
}
```

10.0.0.100:1234

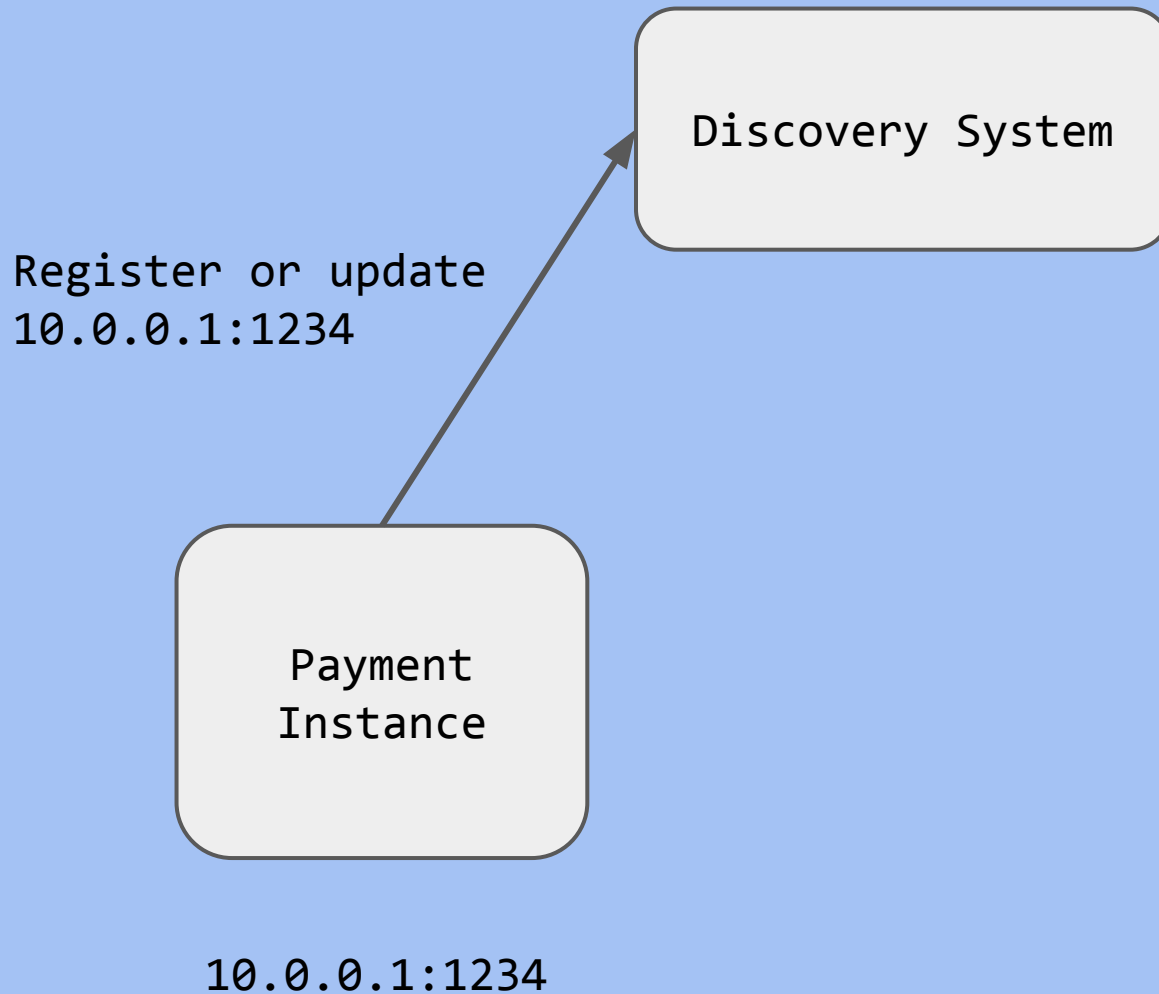
Discovery System



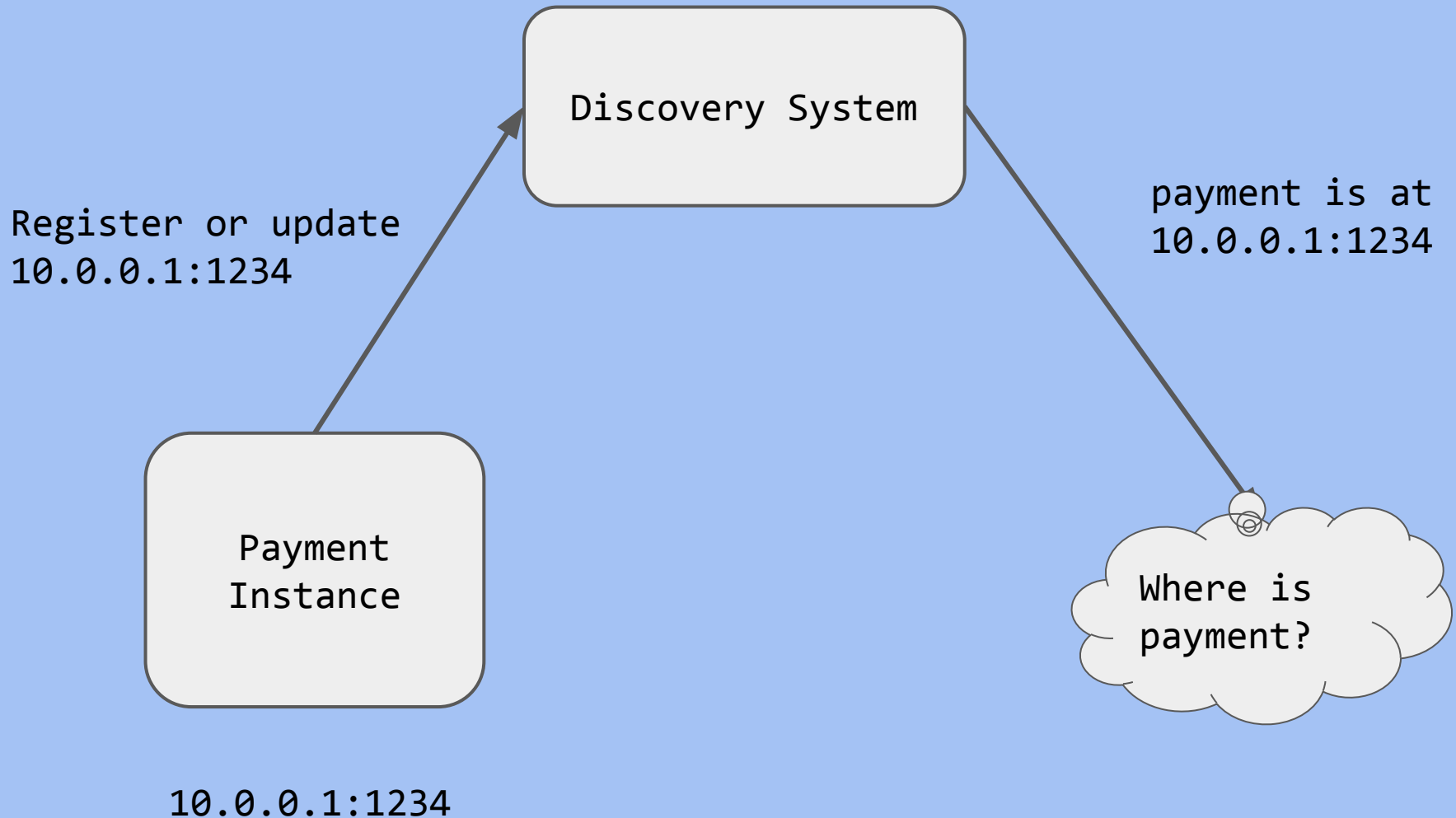
```
find_payment_service()  
{  
    return payment.example.com  
}
```

10.0.0.1:1234

Discovery System



Discovery System



Discovery System

Real problems

- Load balancing across multiple instances
- Auto removal dead instances
- Rolling upgrades

Cluster Management



Scheduler

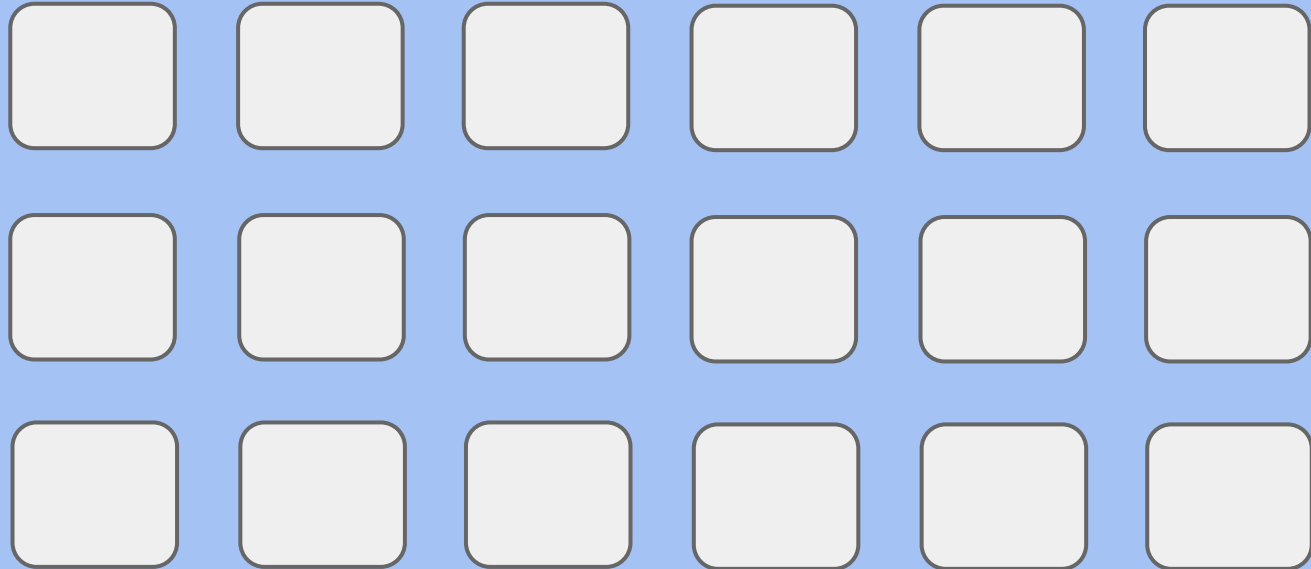
- Schedule pods to run on nodes
 - Global state
 - Multiple schedulers
 - Pluggable
 -

Scheduler

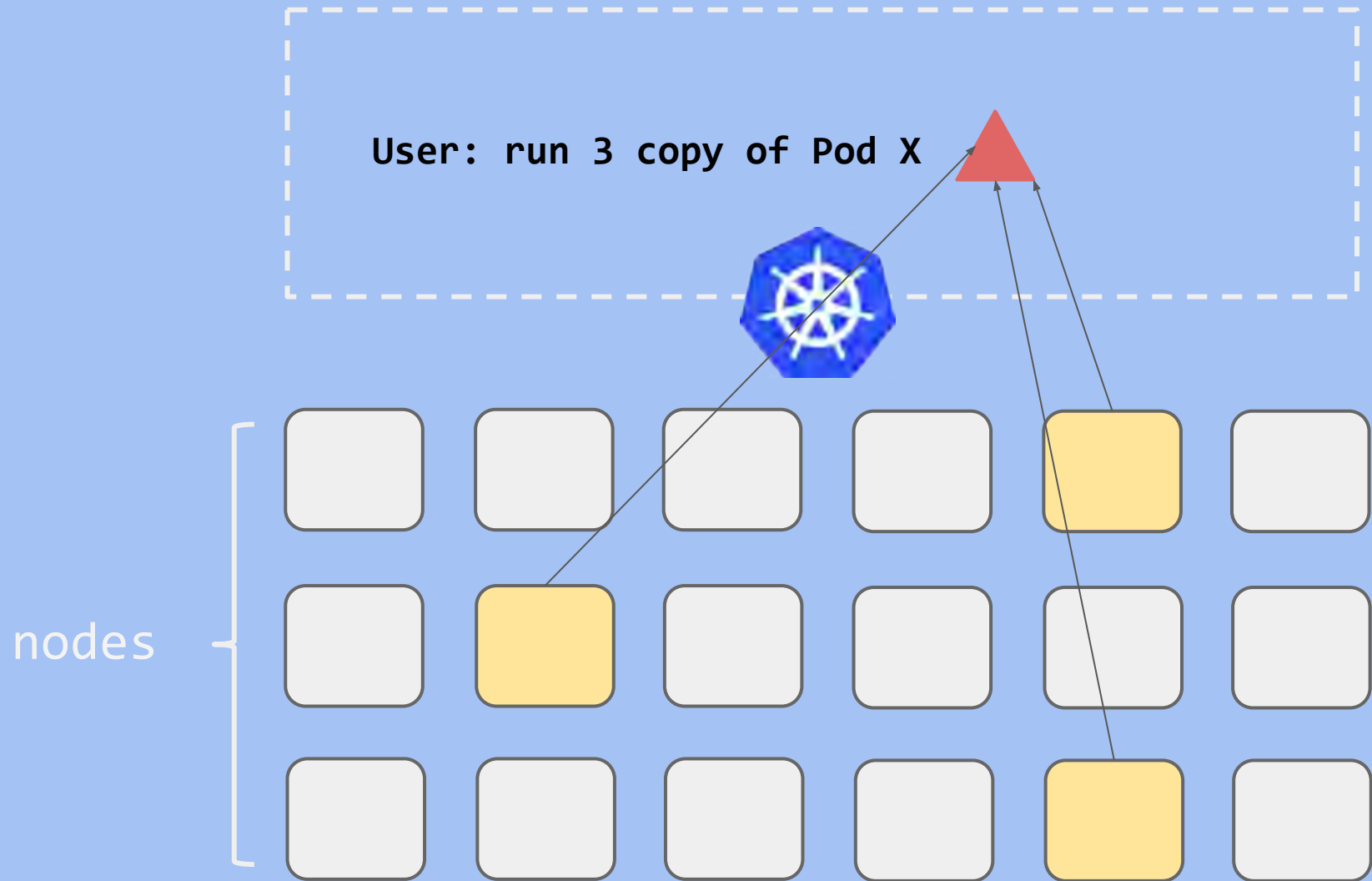
User: run 3 copy of Pod X



nodes



Scheduler

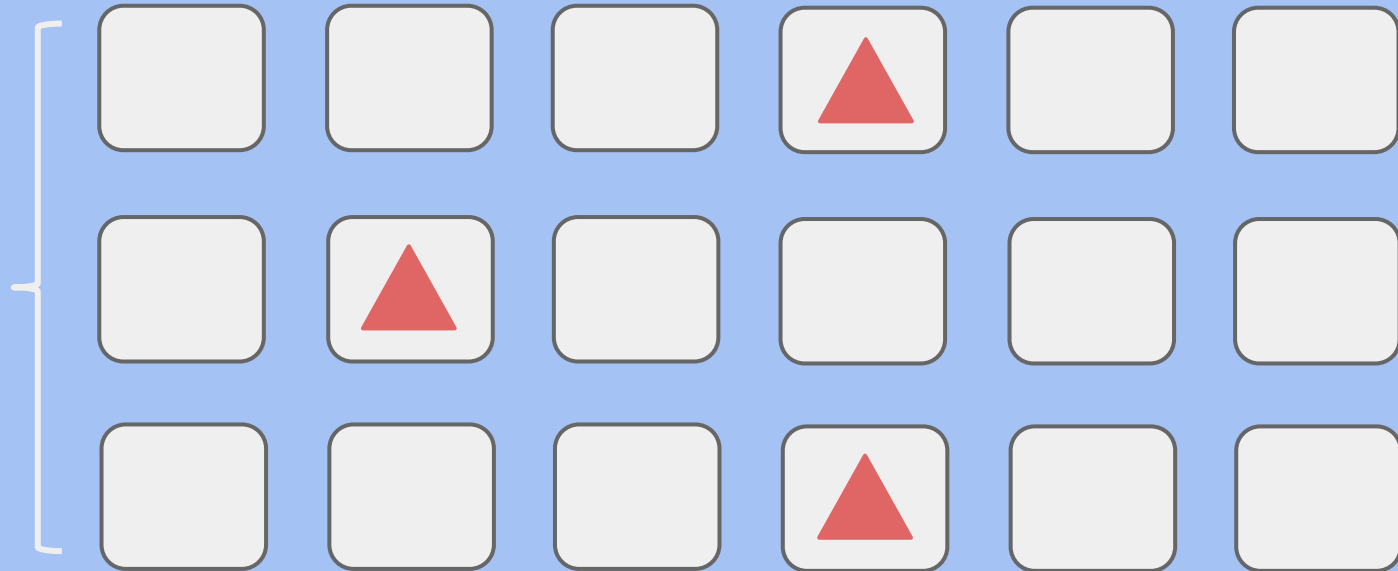


Scheduler

User: what is the status of X Pod?



nodes

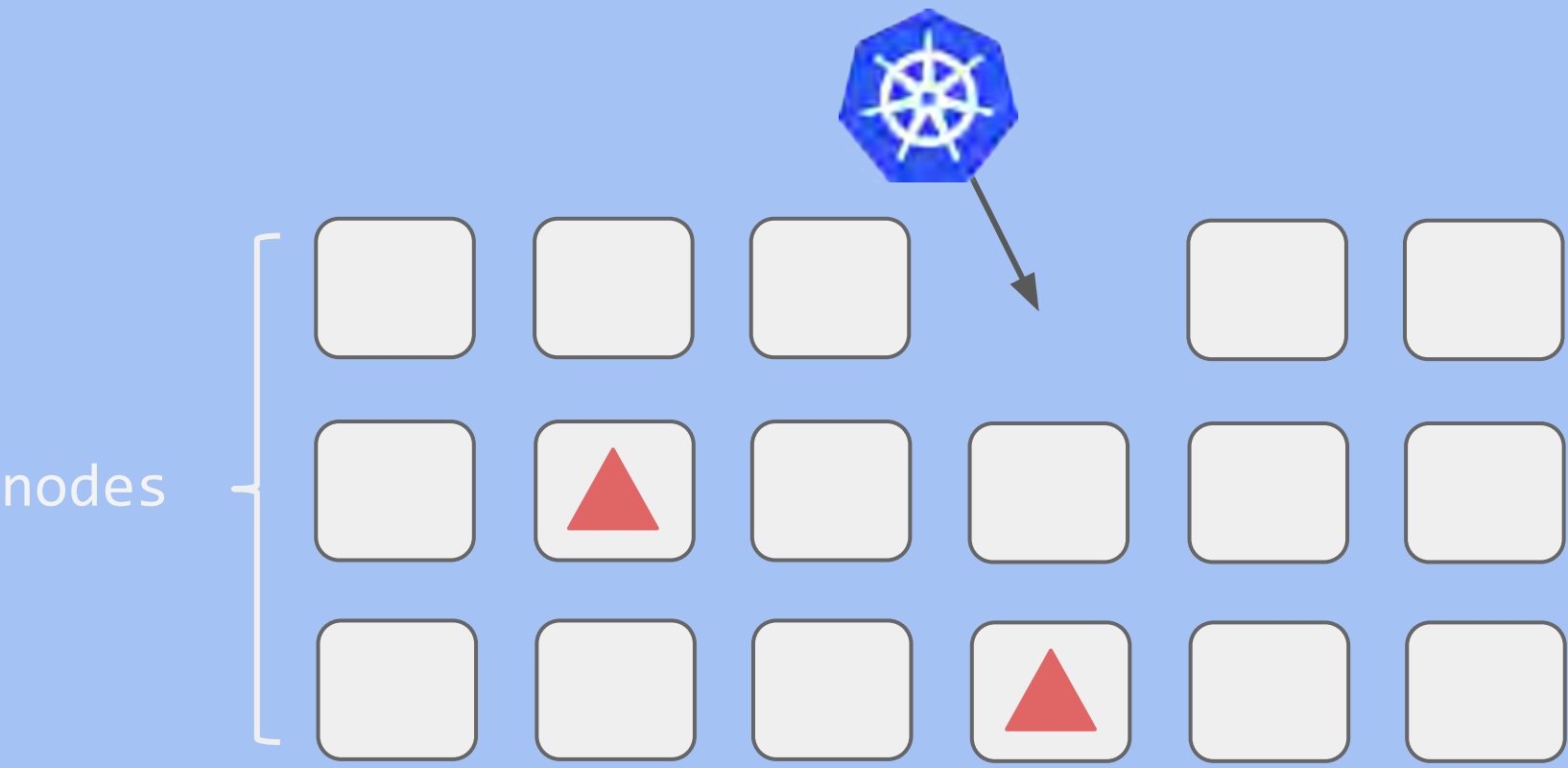


Replication Controller

- Manage a replicated set of pods
 - ensure the desired number
 - resizing
-

Replication Controller

Node failure

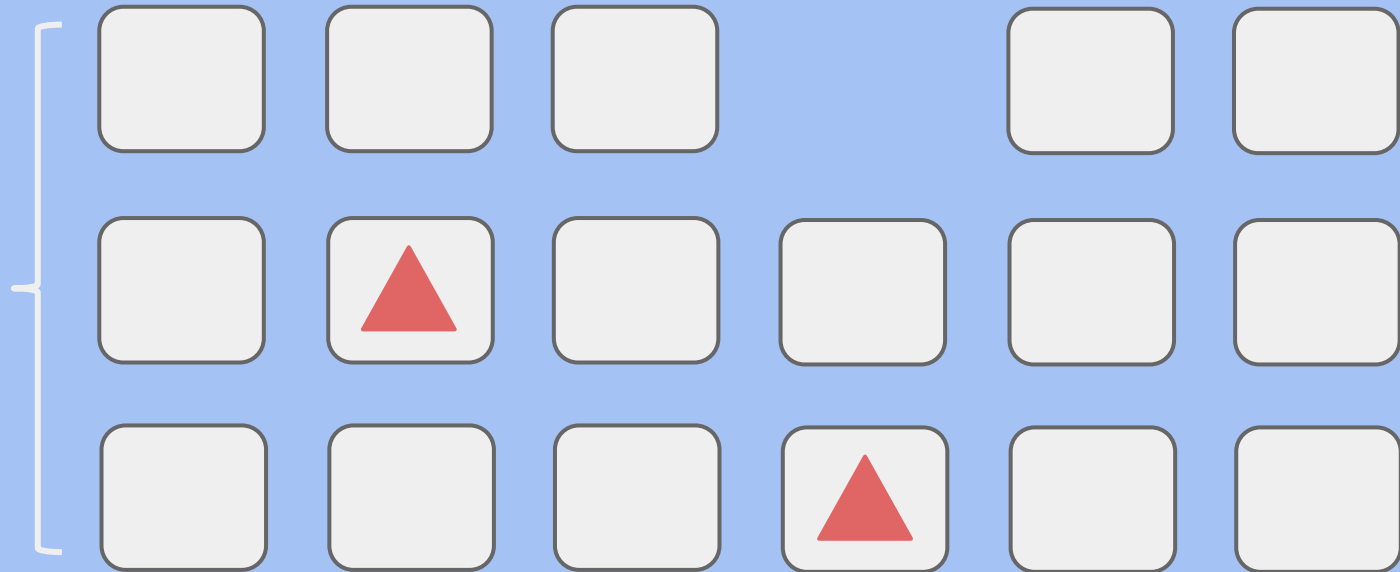


Replication Controller

Need to replicate one more X



nodes

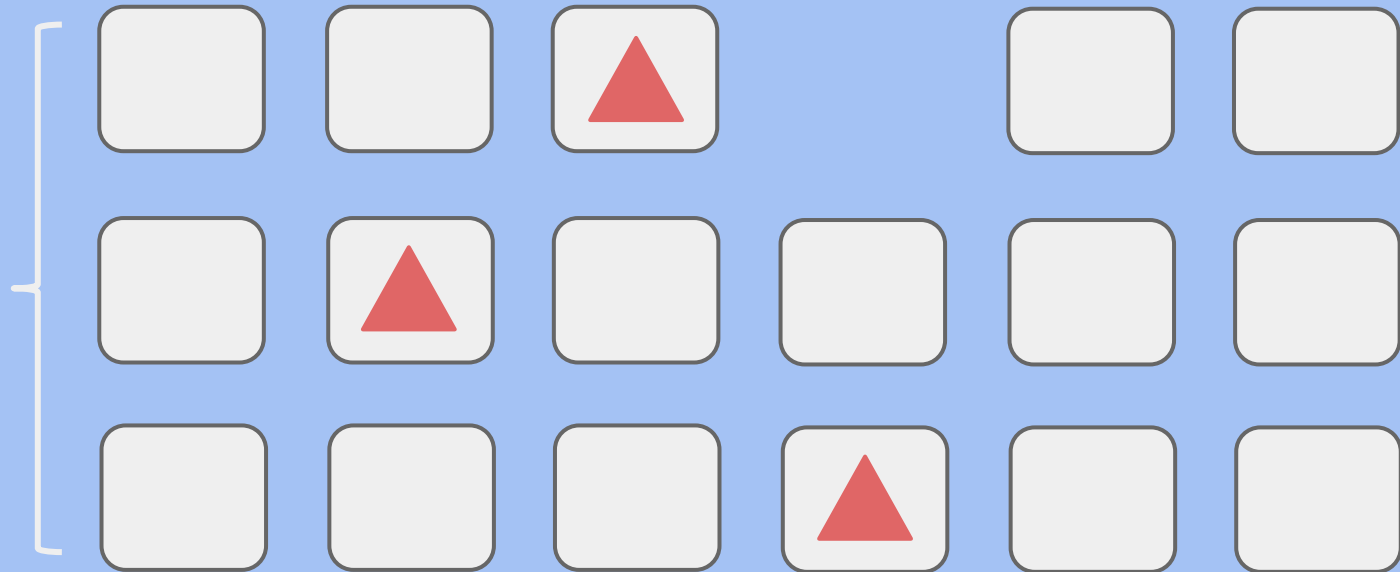


Replication Controller

Need to replicate one more X

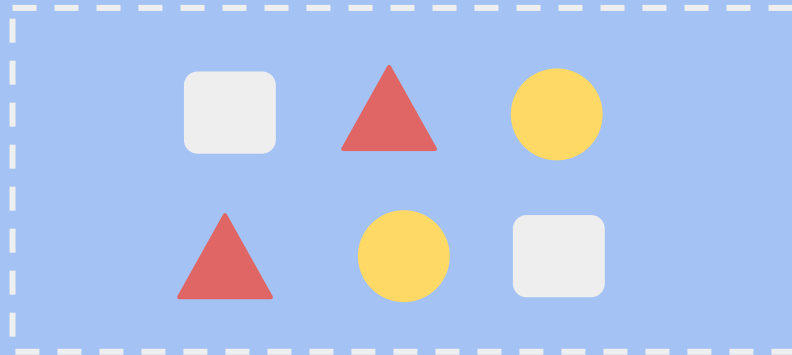


nodes



Service

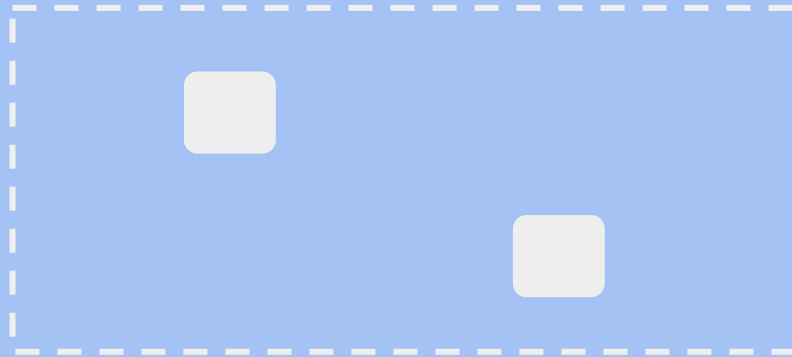
- A dynamic collection of pods
 - select by label query



Service

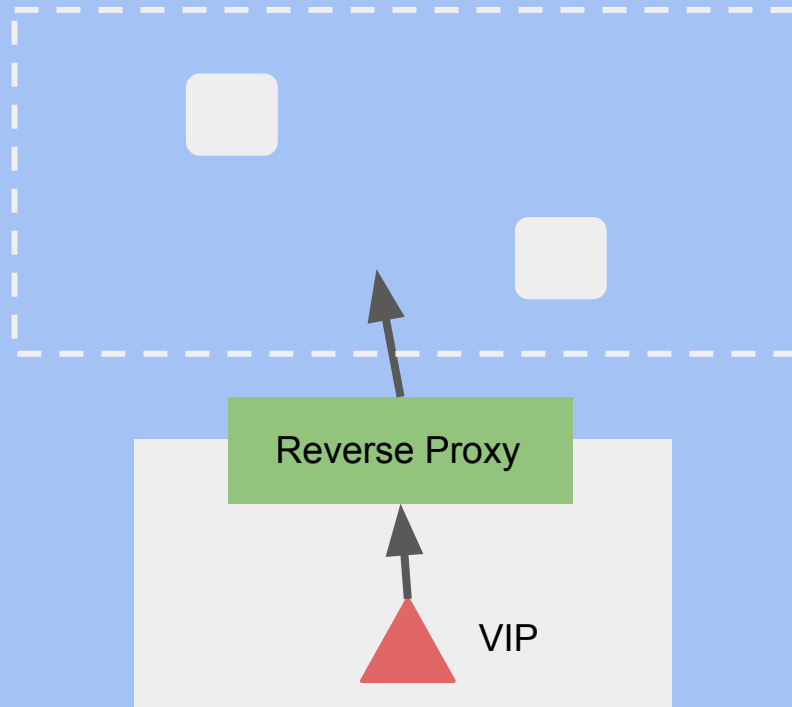
- A dynamic collection of pods
 - select by label query

White && Square



Service

- Native support for service discovery
 - Local proxy on each node
 - Virtual IP per service



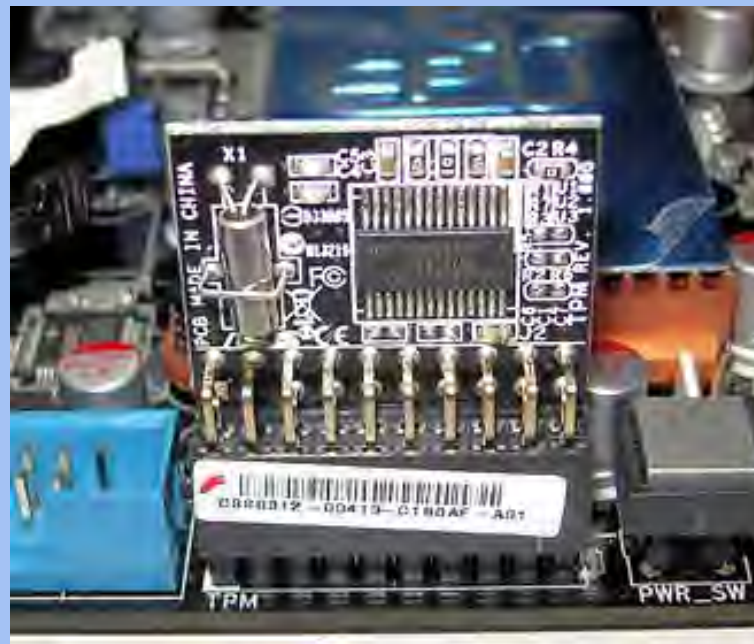
Distributed Trusted Computing



Distributed Trusted Computing

TPM

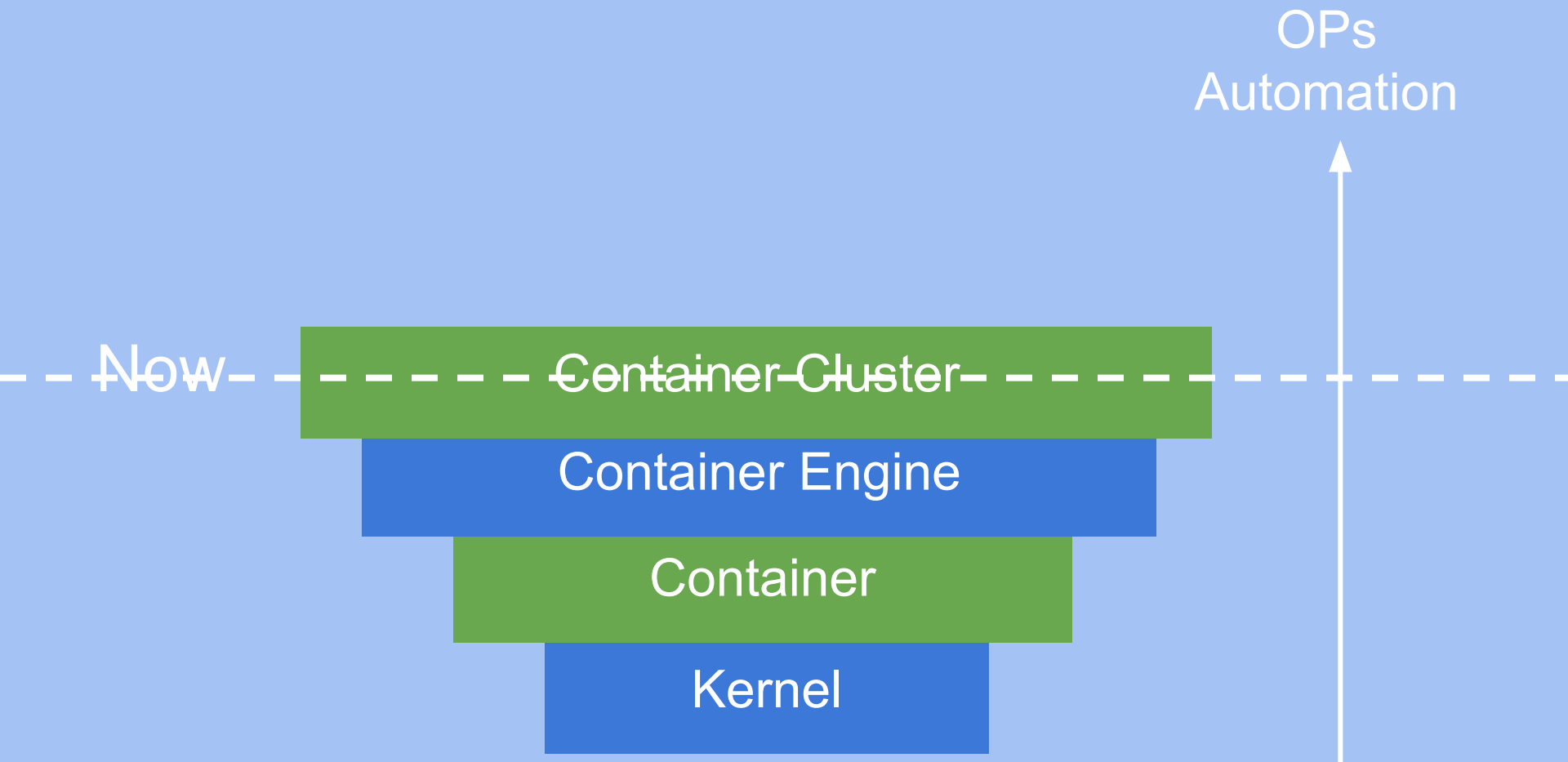
- Store and generate RSA keys
- Records measurements
- Provide Attestation



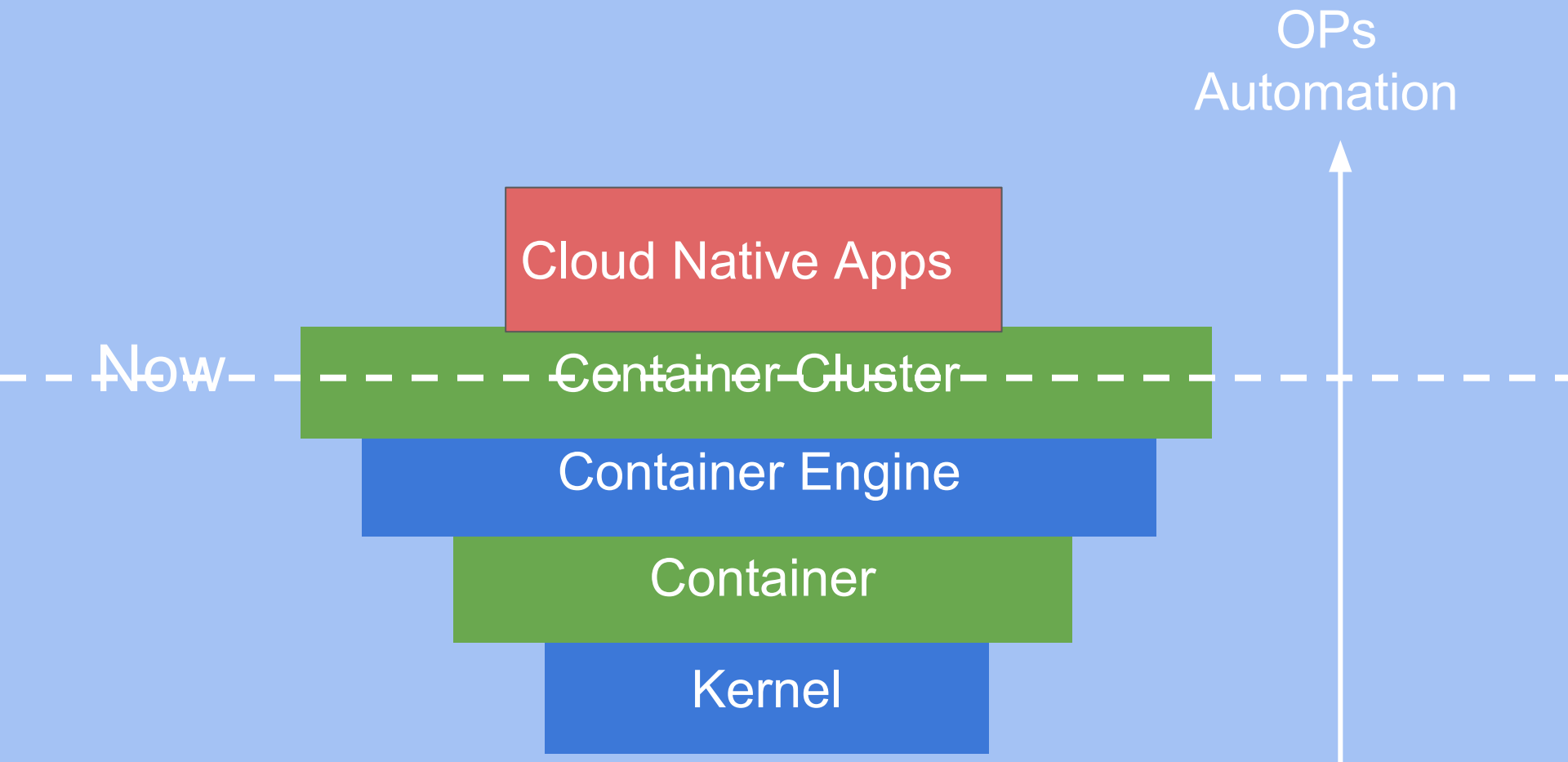
Distributed Trusted Computing



Evolution



Evolution



Thanks

We are hiring!

San francisco, New York, Berlin and Beijing

xiang.li@coreos.com