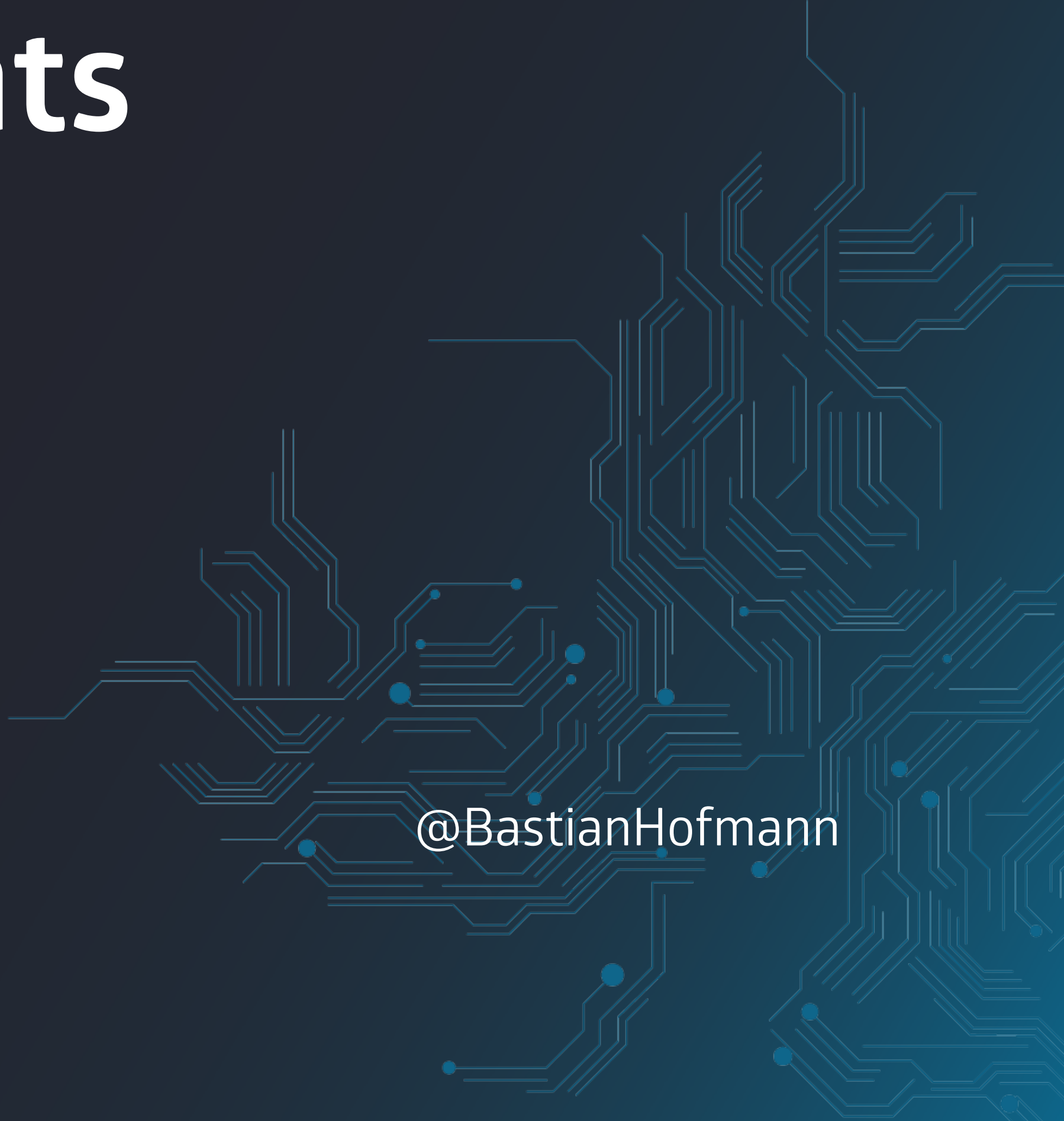


Highly available cross-region deployments with Kubernetes

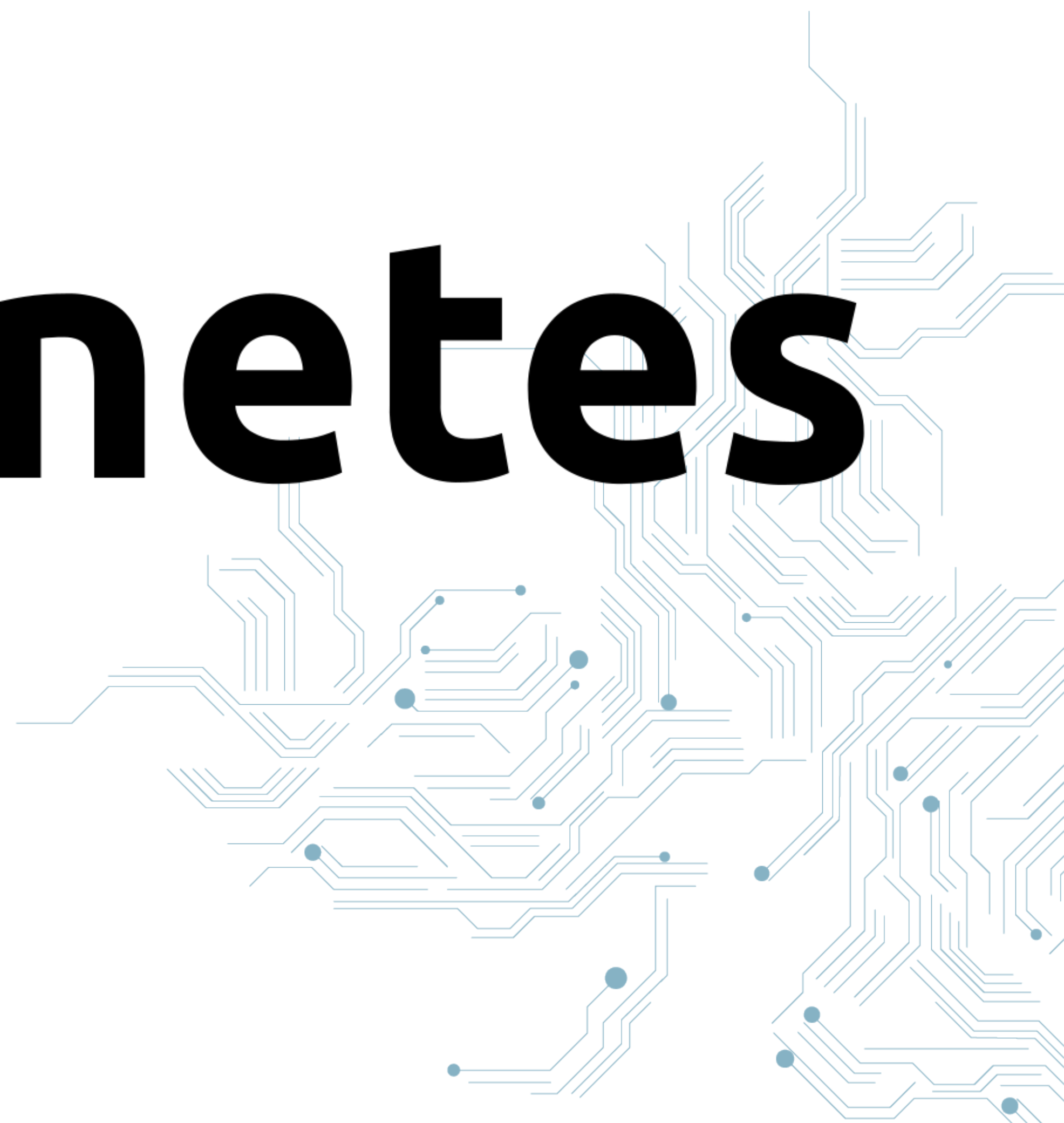
Bastian Hofmann

@BastianHofmann

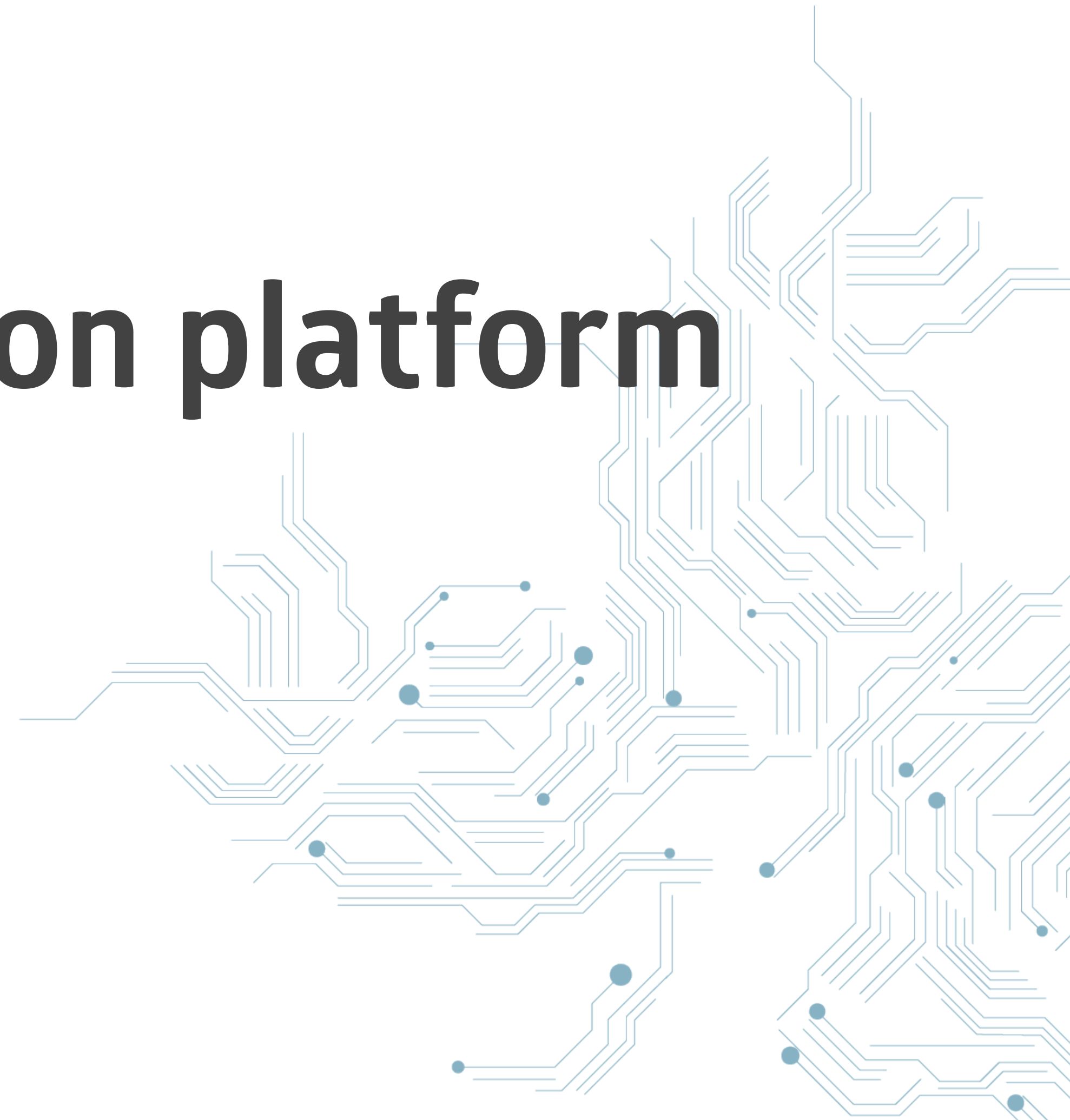
A decorative graphic in the bottom right corner consisting of a complex network of white lines and dots on a dark blue background, resembling a circuit board or a neural network.



kubernetes



Container orchestration platform



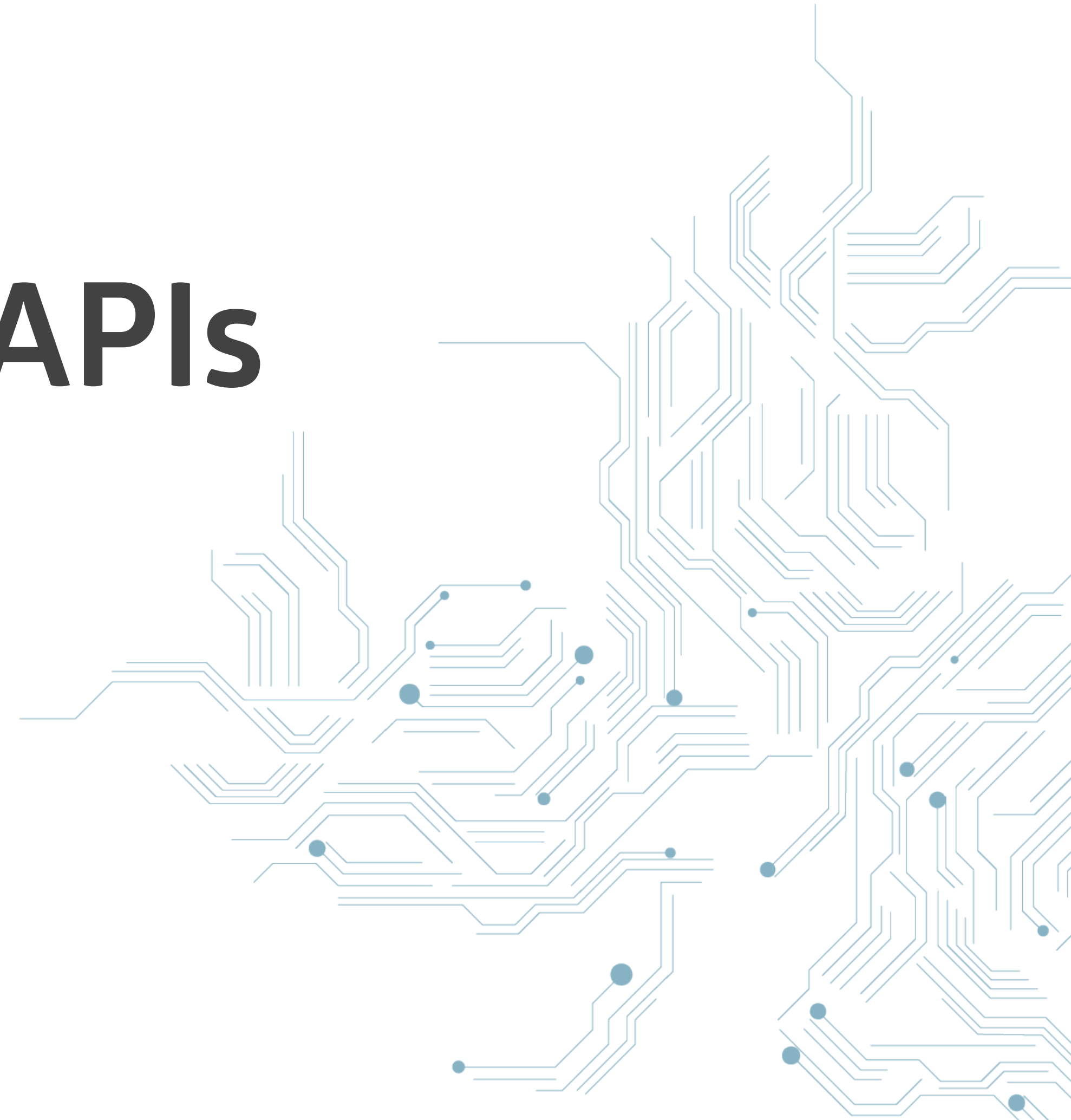
**Deploy, run and scale your services
in isolated containers**



No vendor lock in



Standardized APIs



Runs on



Your laptop



Bare metal



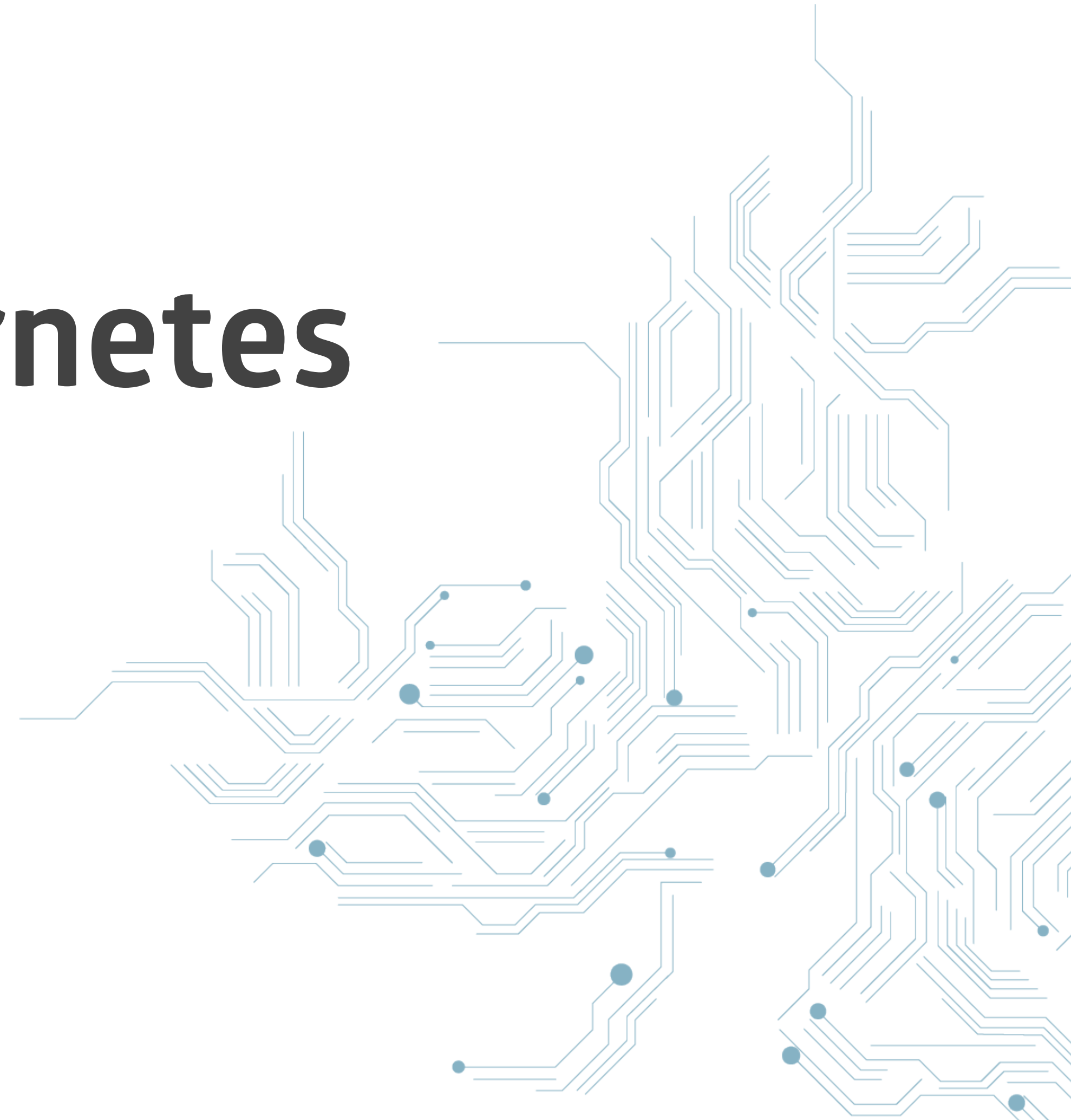
Cloud Providers



**And if you don't want to install and
maintain Kubernetes yourself**



Managed Kubernetes



CNCF Cloud Native Interactive Landscape



The Cloud Native Trail Map ([png](#), [pdf](#)) is CNCF's recommended path through the cloud native landscape. The cloud native landscape ([png](#), [pdf](#)) and serverless landscape ([png](#), [pdf](#)) are dynamically generated below. Please [open](#) a pull request to correct any issues. Greyed logos are not open source. Last Updated: 2019-02-12 06:44:45Z

You are viewing 33 cards with a total of 180 stars, market cap of \$4.01T and funding of \$1.19B.

Landscape **Card Mode** Serverless

Share 462

Platform - Certified Kubernetes - Hosted (33)

Alibaba Cloud Container Service for Kubernetes Alibaba Cloud MCap: \$434B	Amazon Elastic Container Service for Kubernetes (EKS) Amazon Web Services MCap: \$782B	Azure (ACS) Engine Microsoft ★ 174 MCap: \$808B	Azure Kubernetes Service (AKS) Microsoft MCap: \$808B	Azure Stack (ACS) Engine Microsoft ★ 4 MCap: \$808B	Baidu Cloud Container Engine Baidu MCap: \$58.6B	BoCloud BeyondcentContainer Bocloud Funding: \$15.3M	Catalyst Kubernetes Service Catalyst Cloud ★ 2	Cisco Container Platform Cisco MCap: \$214B
DigitalOcean Kubernetes DigitalOcean Funding: \$305M	EasyStack Kubernetes Service (EKS) EasyStack Funding: \$110M	eBaoCloud eBaoTech Corporation	eKing Cloud Container Platform Hainan eKing Technology	ELASTX Private Kubernetes ELASTX	Google Kubernetes Engine (GKE) Google MCap: \$763B	HarmonyCloud Container Platform Hangzhou Harmony Technology	Hasura Hasura Funding: \$1.6M	Huawei Cloud Container Engine (CCE) Huawei Technologies
IBM Cloud Kubernetes Service IBM MCap: \$122B	Intellect FABRIC Intellect Design Arena MCap: \$22.4B	Nirmata Managed Kubernetes Nirmata	Nutanix Karbon Nutanix MCap: \$9.48B	Oracle Container Engine Oracle MCap: \$184B	OVH Managed Kubernetes Service OVH.com Funding: \$737M	Rackspace Kubernetes-as-a-Service Rackspace Funding: \$17.8M	Samsung SDS Kubernetes Service Samsung SDS	SAP Certified Gardener SAP MCap: \$126B
SysEleven MetaKube SysEleven	Tencent Kubernetes Engine (TKE) Tencent Holdings MCap: \$421B	TenxCloud Container Engine (TCE) TenxCloud	VEXXHOST Kubernetes Container Service VEXXHOST	VMware Cloud PKS VMware MCap: \$65.2B	ZTE TECS ZTE			

Self-healing built in



**But what happens when a complete
datacenter is not available**



Multi-Region Kubernetes Setups

A decorative graphic on the right side of the slide, consisting of a complex, interconnected network of thin white lines and small blue dots, resembling a circuit board or a data network topology, set against a green gradient background.

Why



Availability



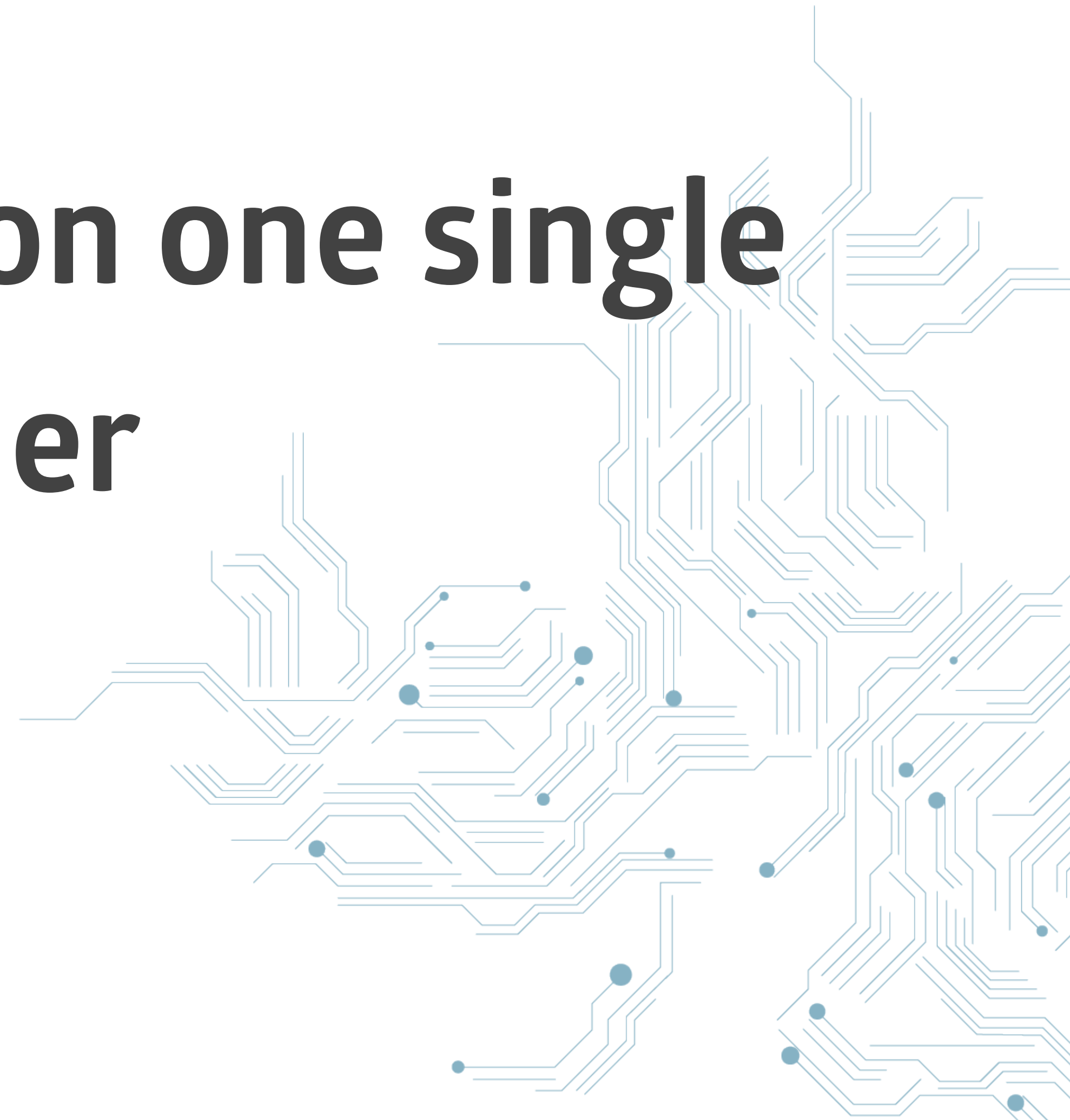
Scalability



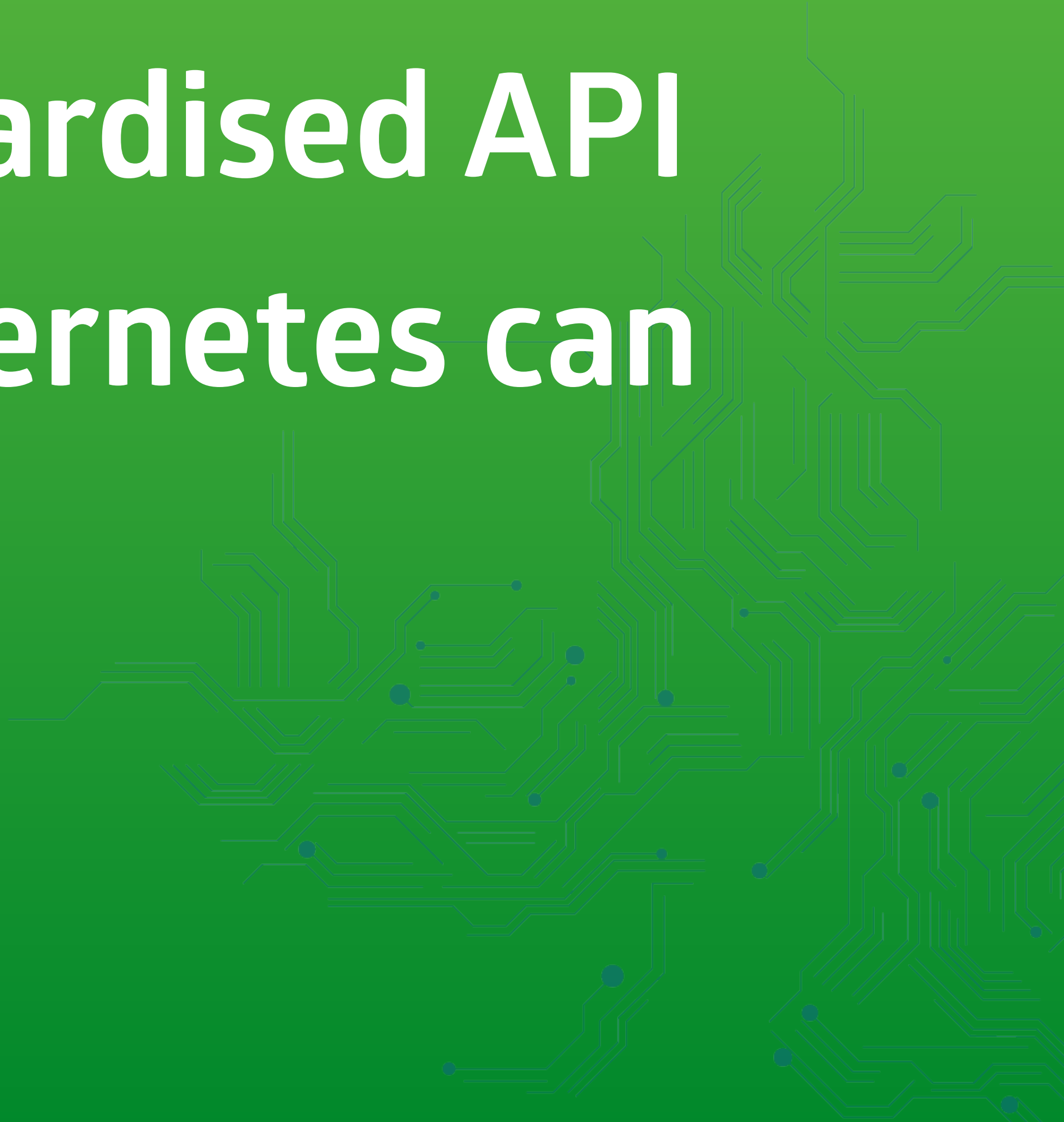
More Points-of-Presence



**Reduce dependencies on one single
cloud provider**



**Because of the standardised API
across providers Kubernetes can
help**

A decorative graphic on the right side of the slide, consisting of a complex, interconnected network of thin white lines and small blue dots, resembling a circuit board or a neural network, set against a dark green background.

Features



Scheduling by Node Labels



**failure-
domain.beta.kubernetes.io/
region=db1**

**failure-
domain.beta.kubernetes.io/
zone=db11**



Regions vs availability zone



**failure-
domain.beta.kubernetes.io/
region=db1**

**failure-
domain.beta.kubernetes.io/
zone=db11**



```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx
spec:
  template:
    spec:
      containers:
      - image: nginx
        name: nginx
      nodeSelector:
        failure-domain.beta.kubernetes.io/region: db1
```

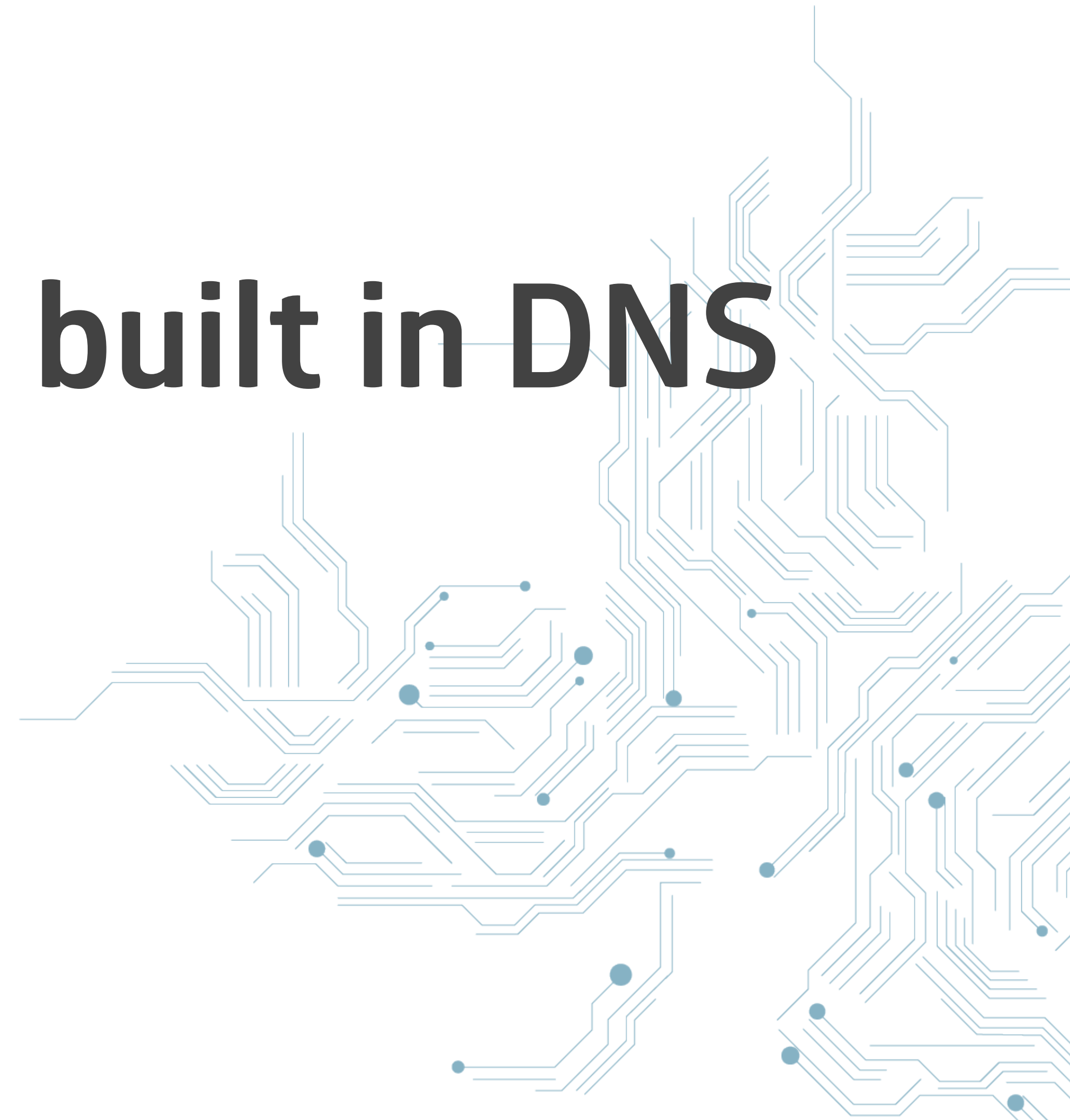
Affinities



```
spec:
  affinity:
    podAntiAffinity:
      requiredDuringSchedulingIgnoredDuringExecution:
        - topologyKey: "failure-domain.beta.kubernetes.io/
region"
      labelSelector:
        matchLabels:
          app: nginx
```

A decorative graphic on the right side of the slide, consisting of a complex network of light blue lines and dots, resembling a circuit board or a neural network diagram.

Service discovery with built in DNS



```
apiVersion: v1
kind: Service
metadata:
  name: nginx
spec:
  type: ClusterIP
  ports:
    - port: 80
      targetPort: 80
  selector:
    app: nginx
```

External LoadBalancers



```
apiVersion: v1
kind: Service
metadata:
  name: nginx
spec:
  type: LoadBalancer
  ports:
    - port: 80
      targetPort: 80
  selector:
    app: nginx
```

external-dns



StorageClasses



**Some storage providers support
dynamic volume provisioning**



```
apiVersion: storage.k8s.io/v1
kind: StorageClass
metadata:
  name: standard
provisioner: kubernetes.io/gce-pd
parameters:
  type: pd-standard
volumeBindingMode: WaitForFirstConsumer
allowedTopologies:
- matchLabelExpressions:
  - key: failure-domain.beta.kubernetes.io/zone
    values:
    - us-central1-a
    - us-central1-b
```

Challenges



Especially over different regions



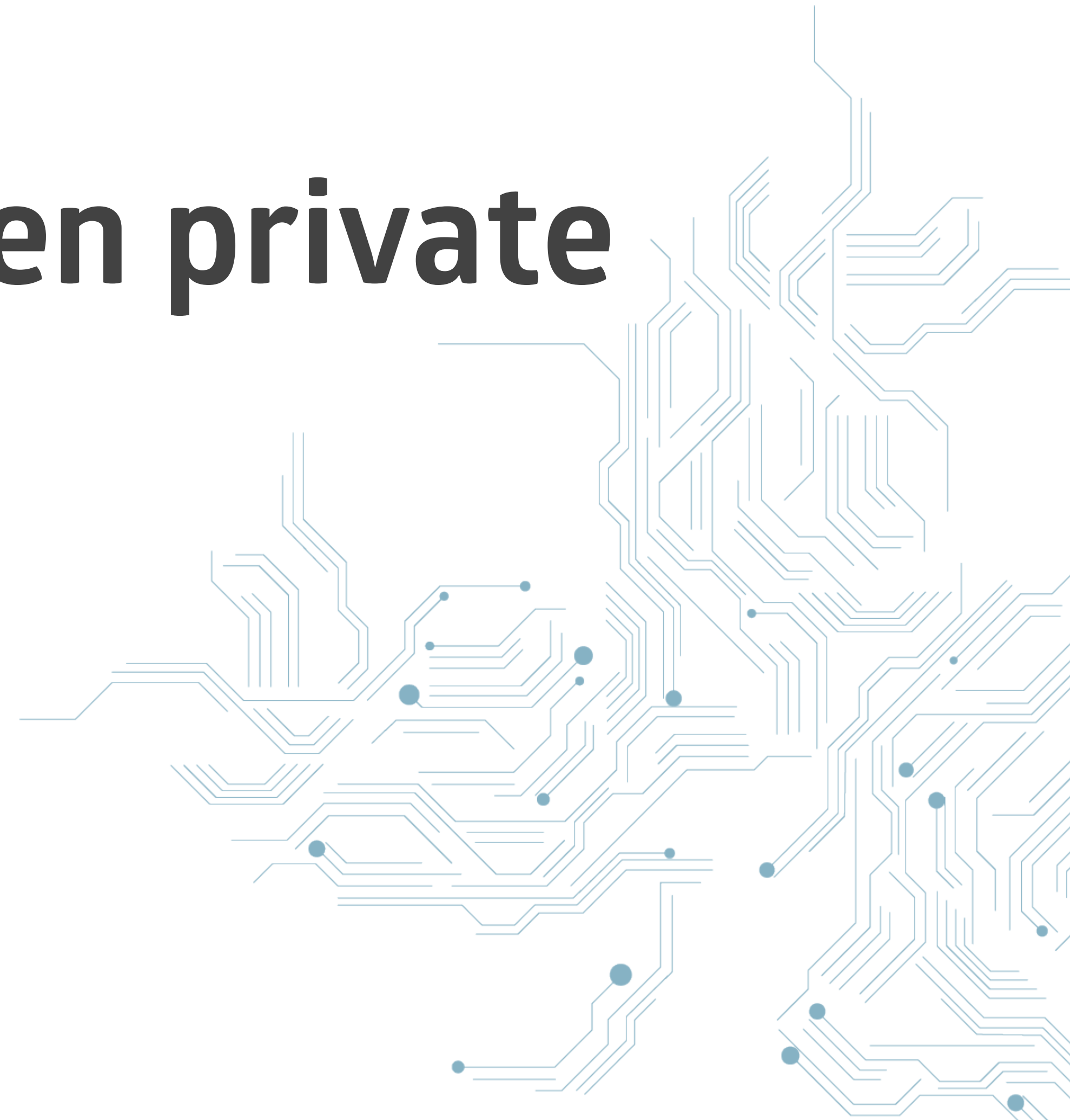
External load balancing and traffic management



Internal load balancing



Connectivity between private networks

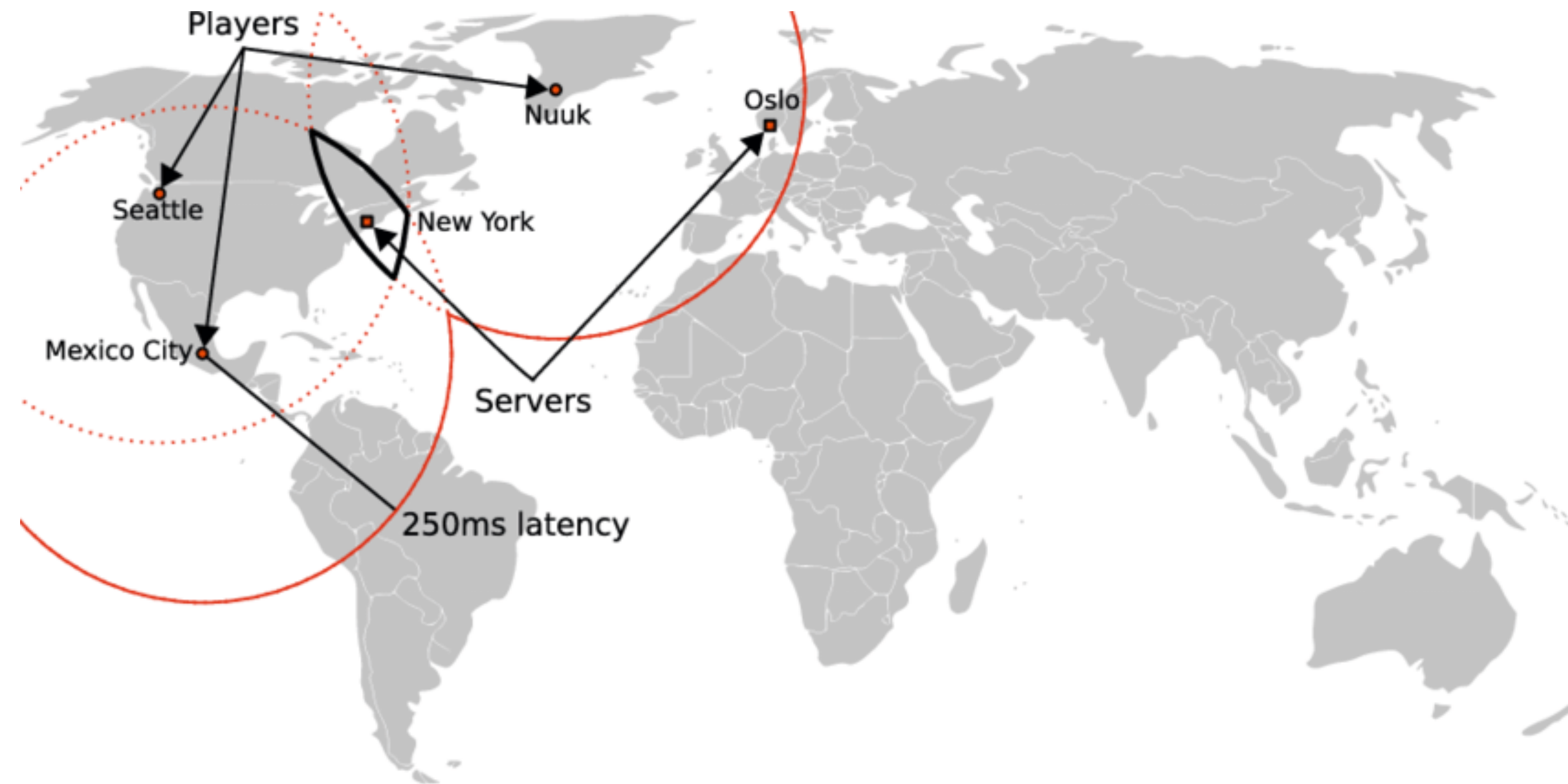


Kubernetes DNS across multiple clusters



Latencies





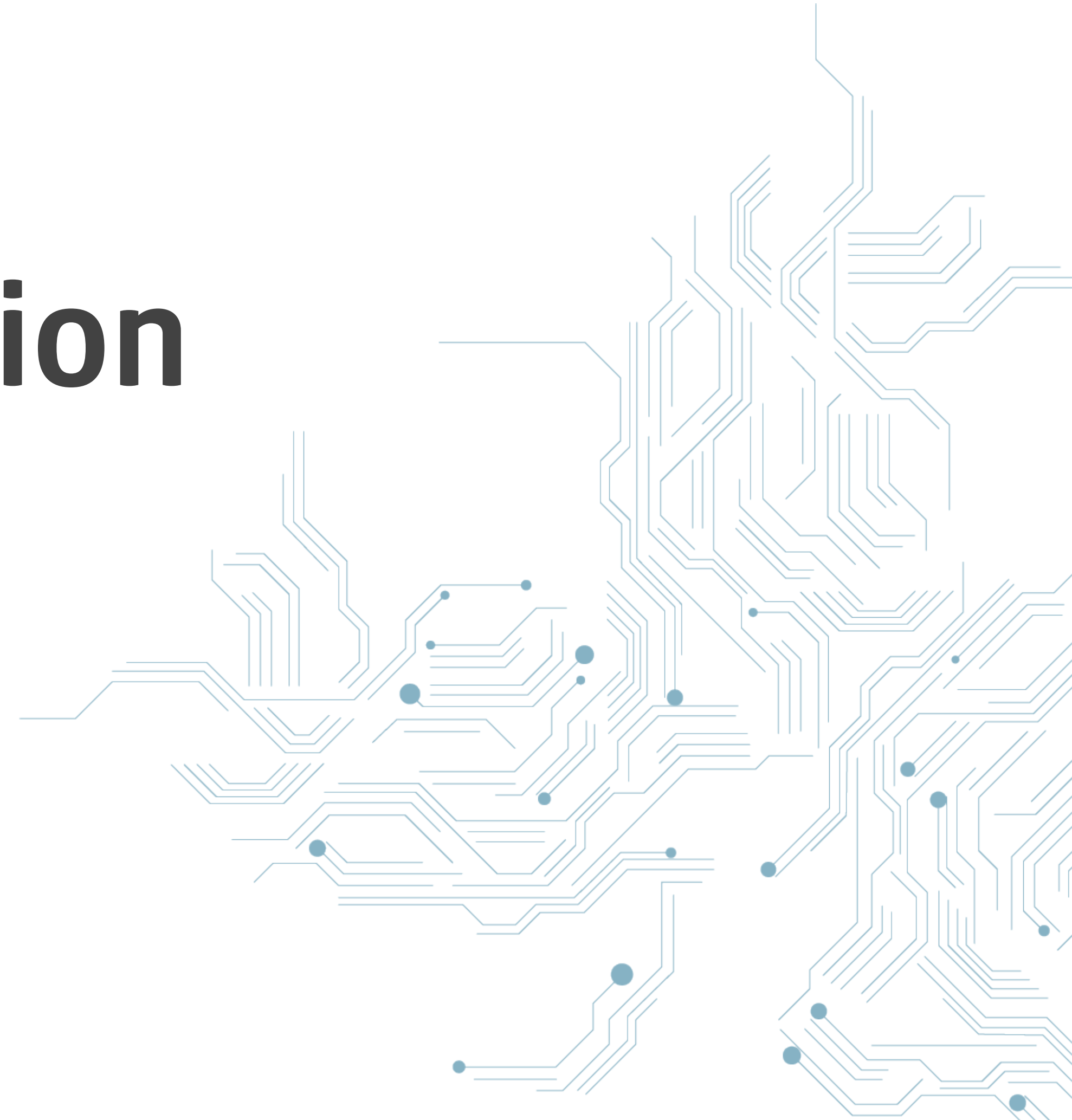
Synchronising Deployments across multiple clusters



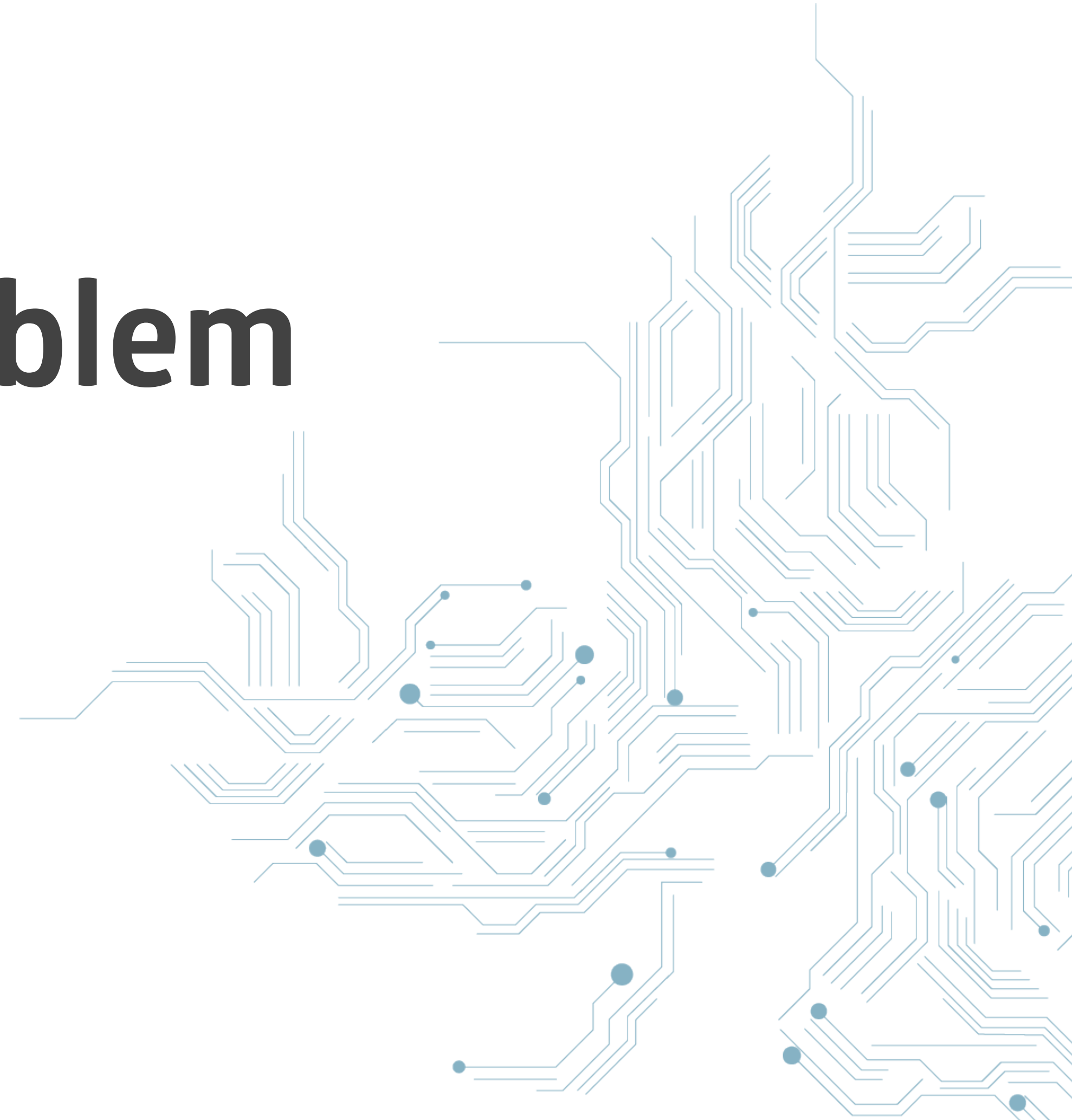
Storage



Data replication



Split-brain problem



Possible setups



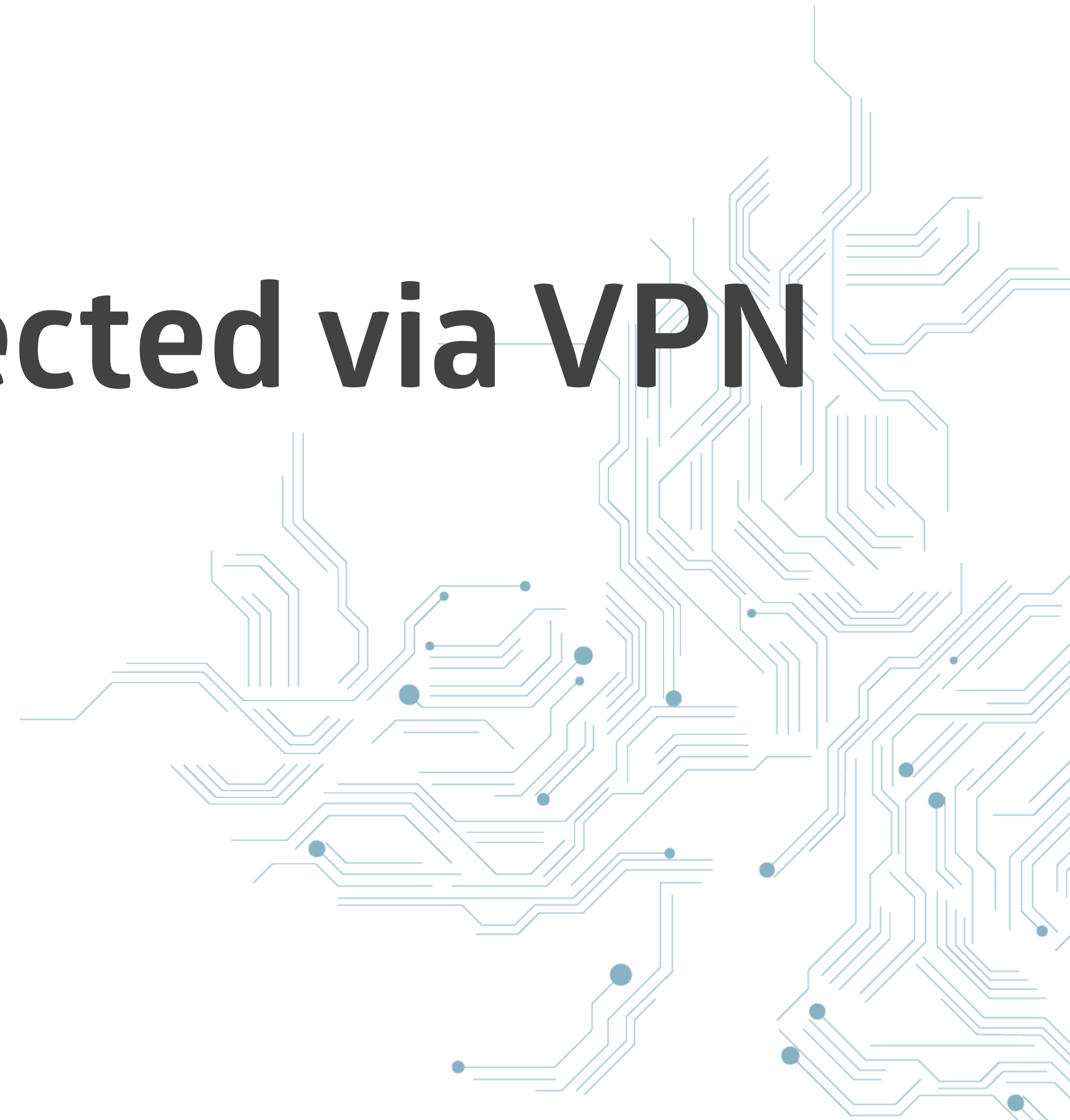
One cluster across *AZs* in one region



One cluster across regions



Multiple clusters connected via VPN



One cluster setup



**One Kubernetes cluster across
multiple availability zones**

A decorative graphic on the right side of the slide, consisting of a complex, stylized circuit board pattern. It features numerous thin, light-blue lines that form a dense, interconnected network, with several small, solid dark-blue dots scattered throughout, resembling components or nodes on a circuit.

Zone 1

Zone 2

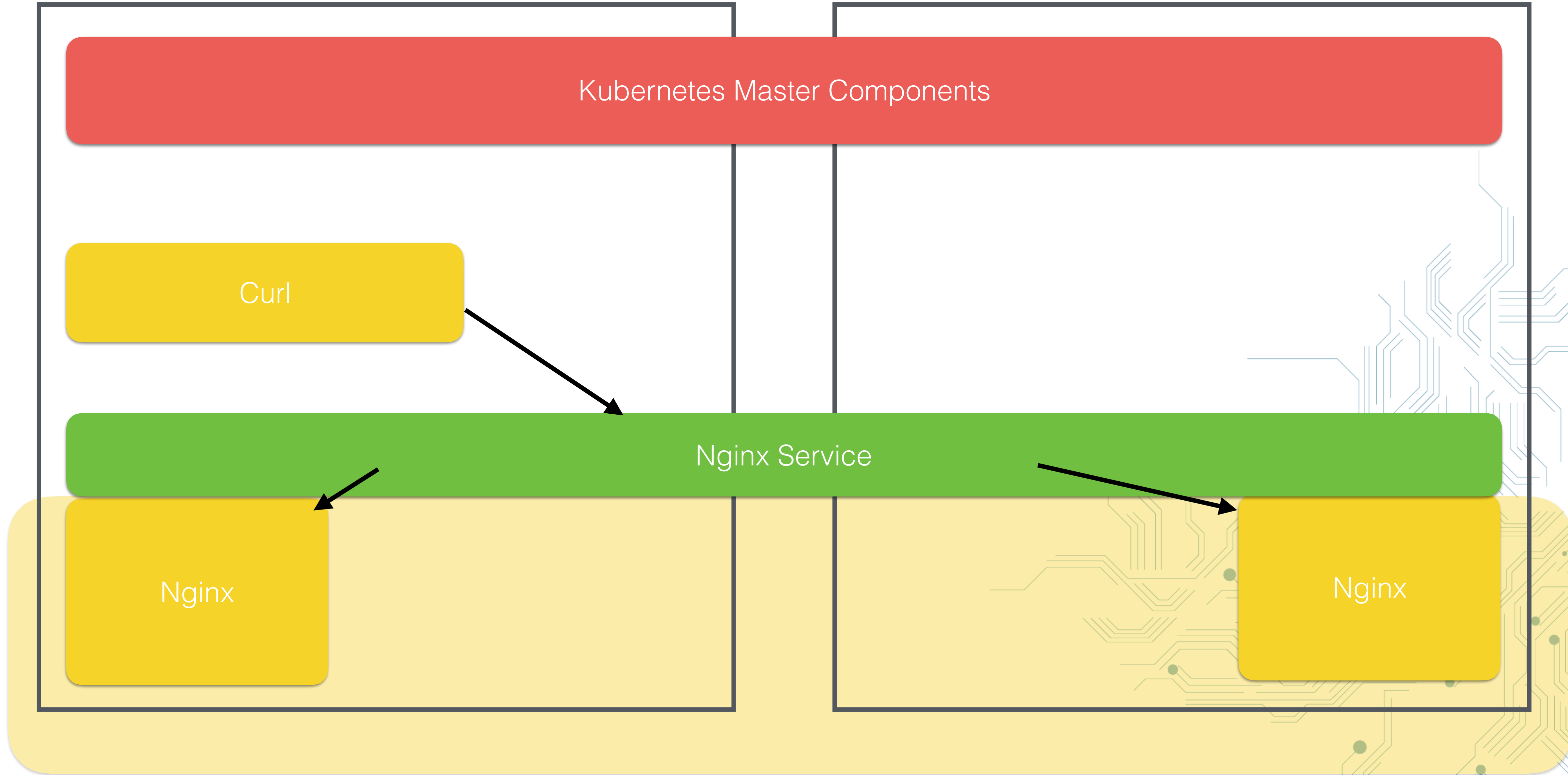
Kubernetes Master Components

Curl

Nginx Service

Nginx

Nginx



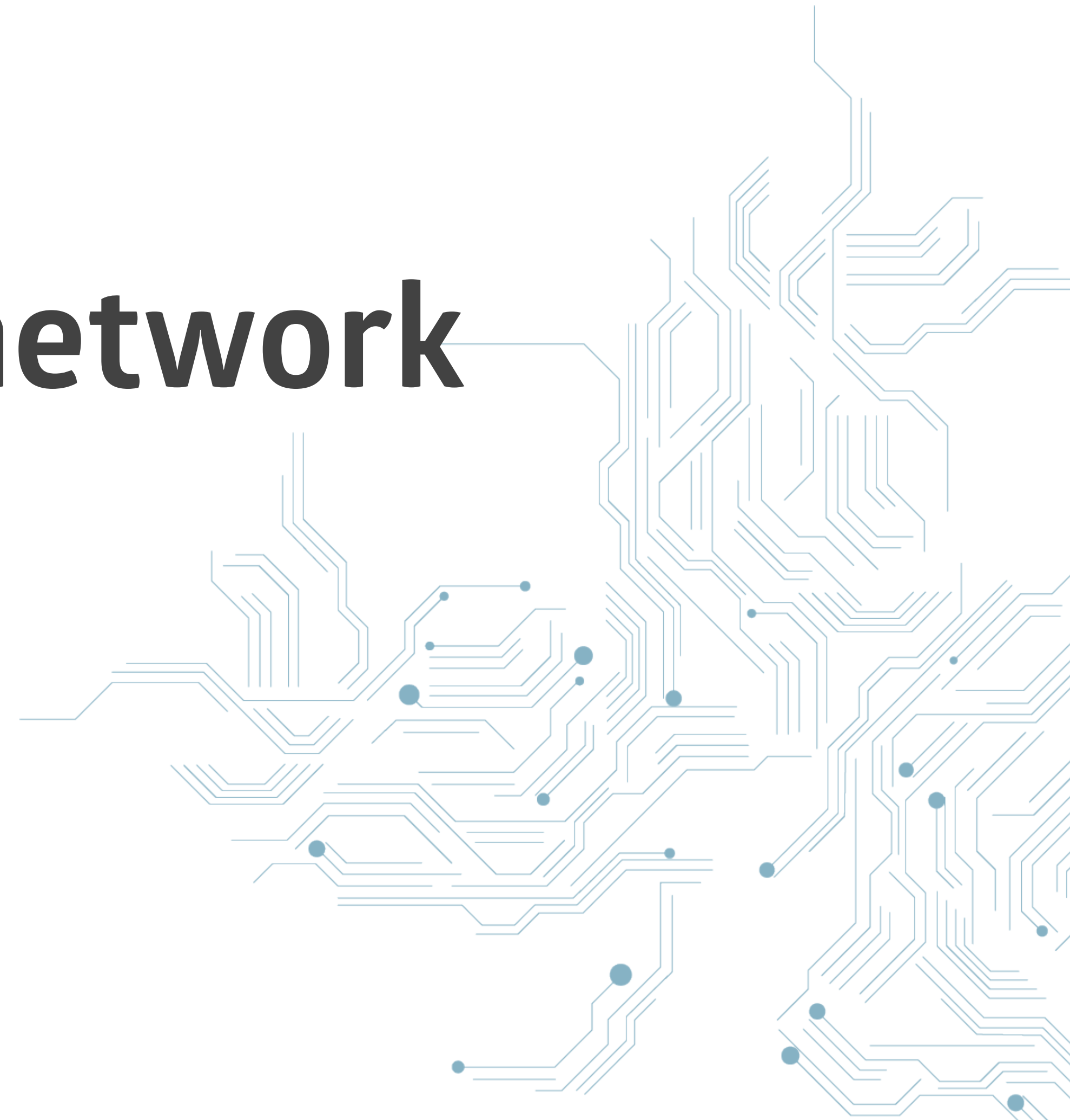
Easiest option



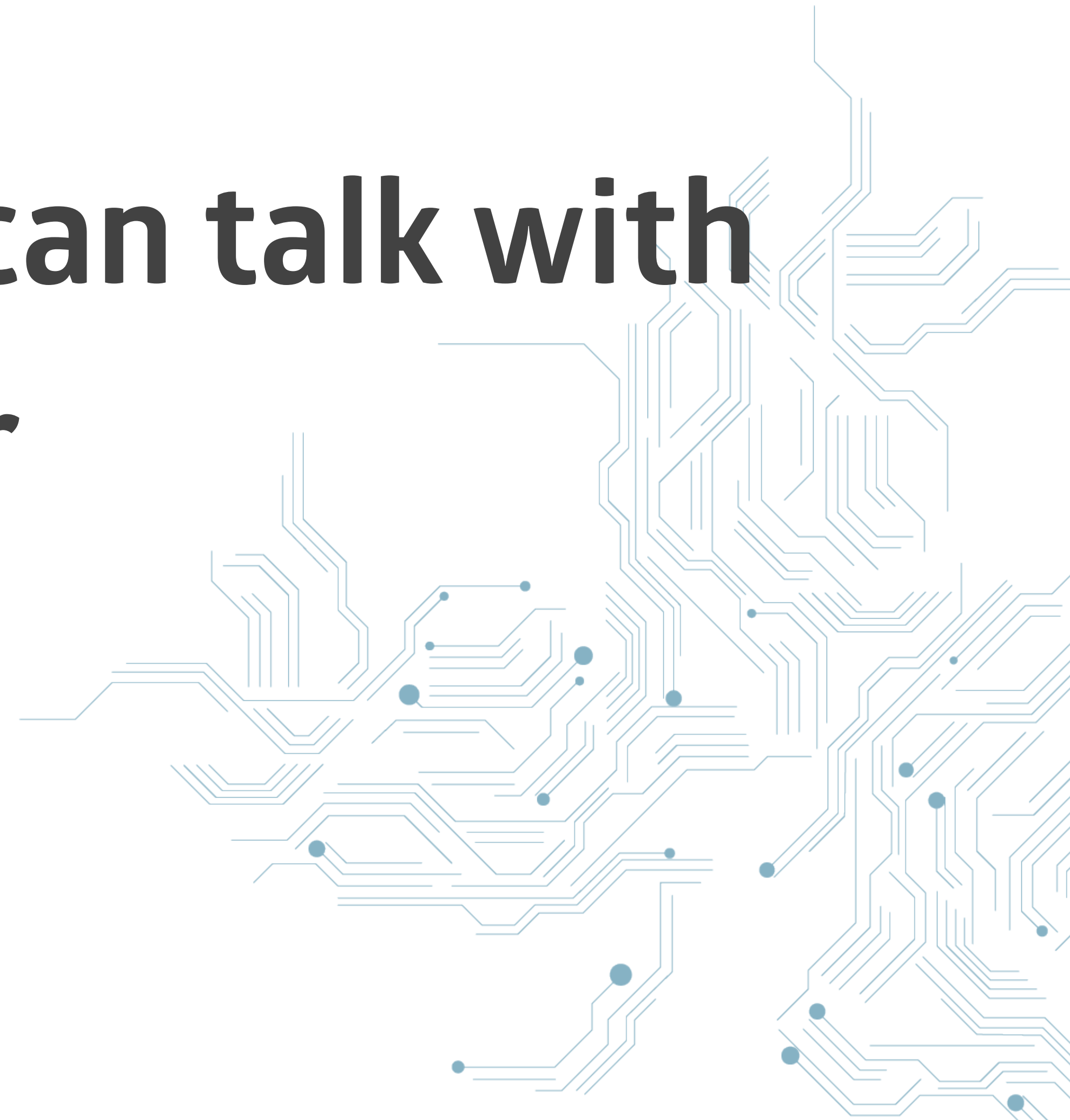
Low latencies



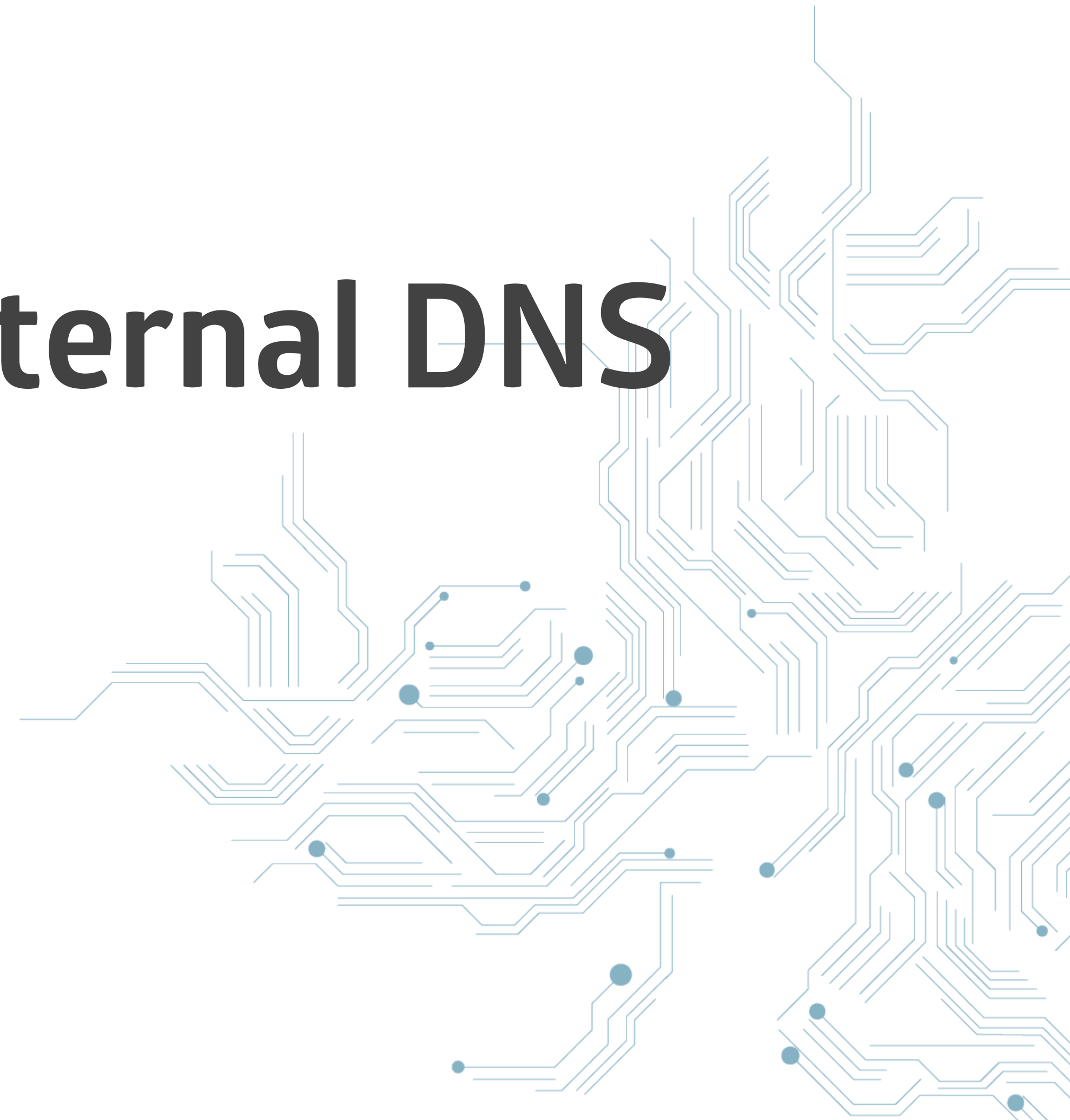
You can use one network



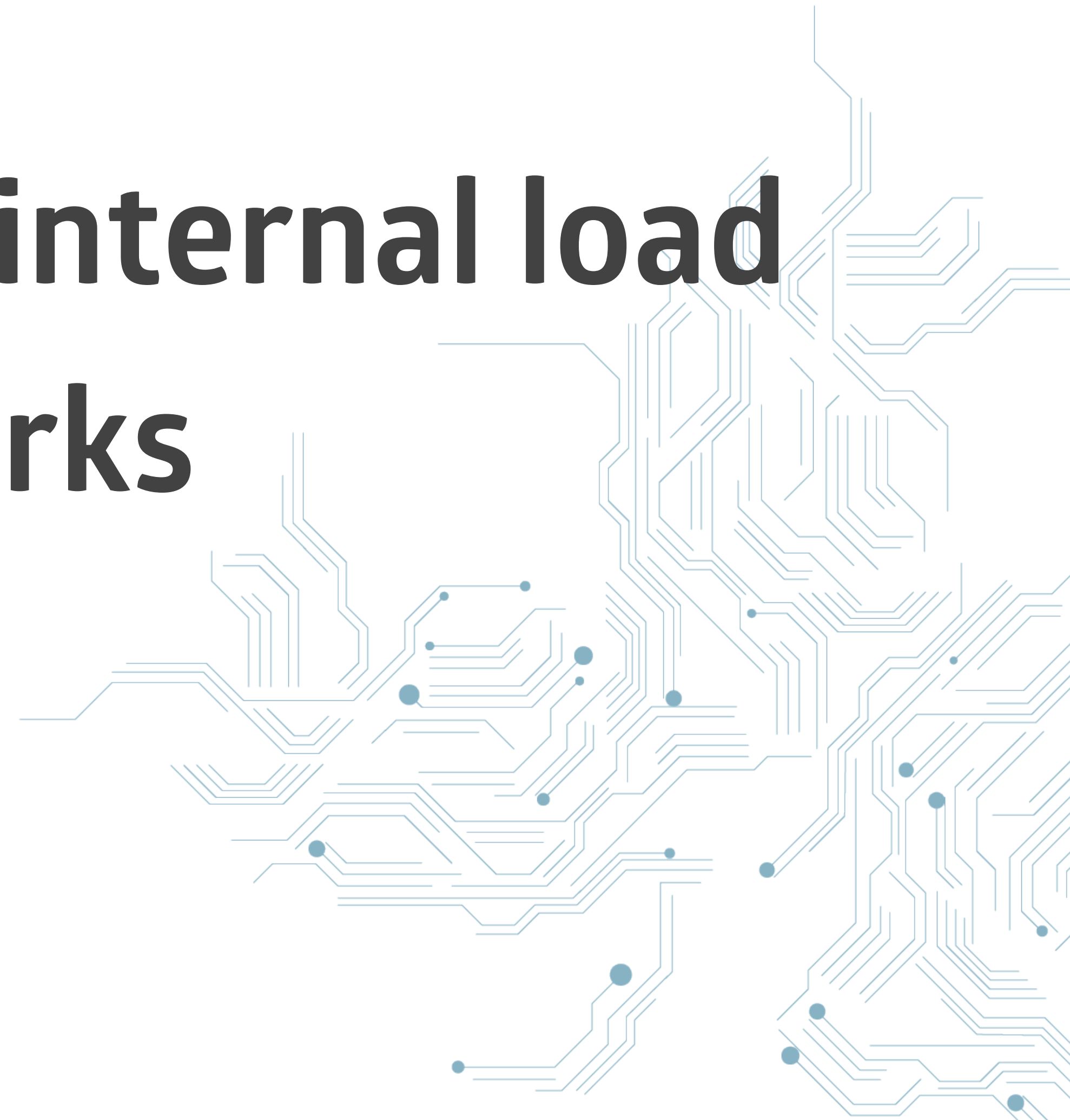
**All pods and services can talk with
each other**



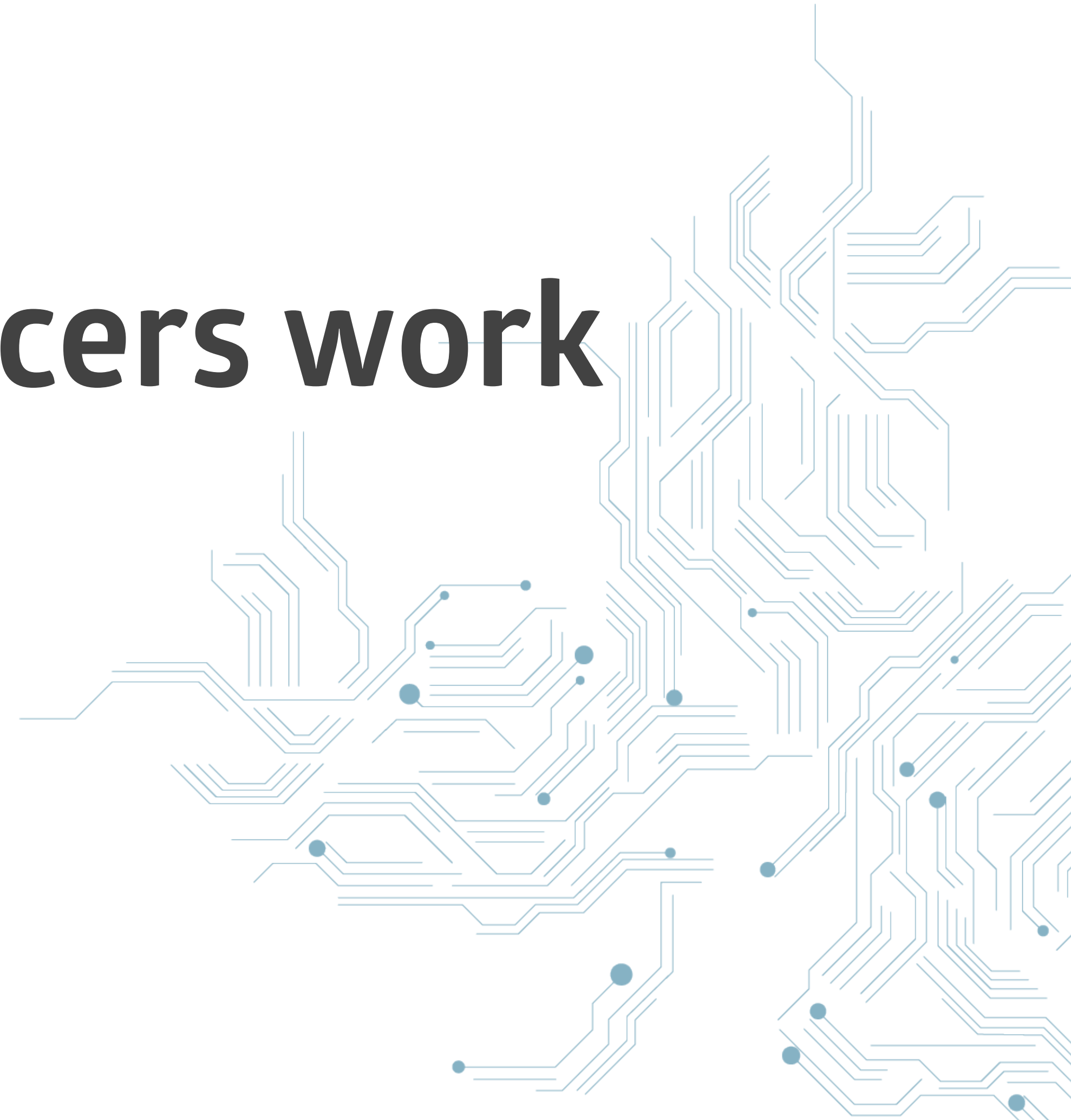
One cluster => one internal DNS



Service Discovery and internal load balancing works



External LoadBalancers work



Zone 1

LB

Zone 2

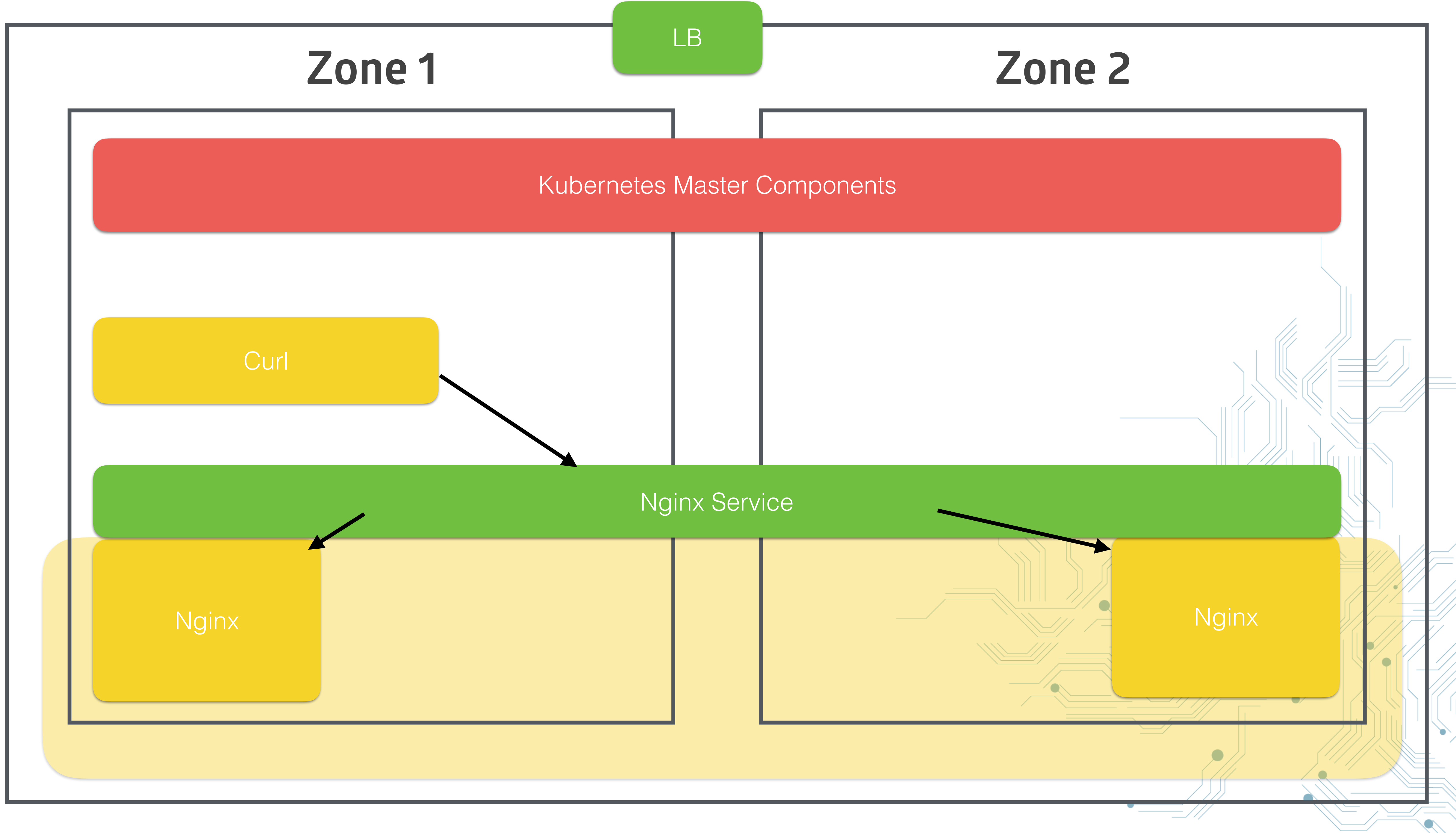
Kubernetes Master Components

Curl

Nginx Service

Nginx

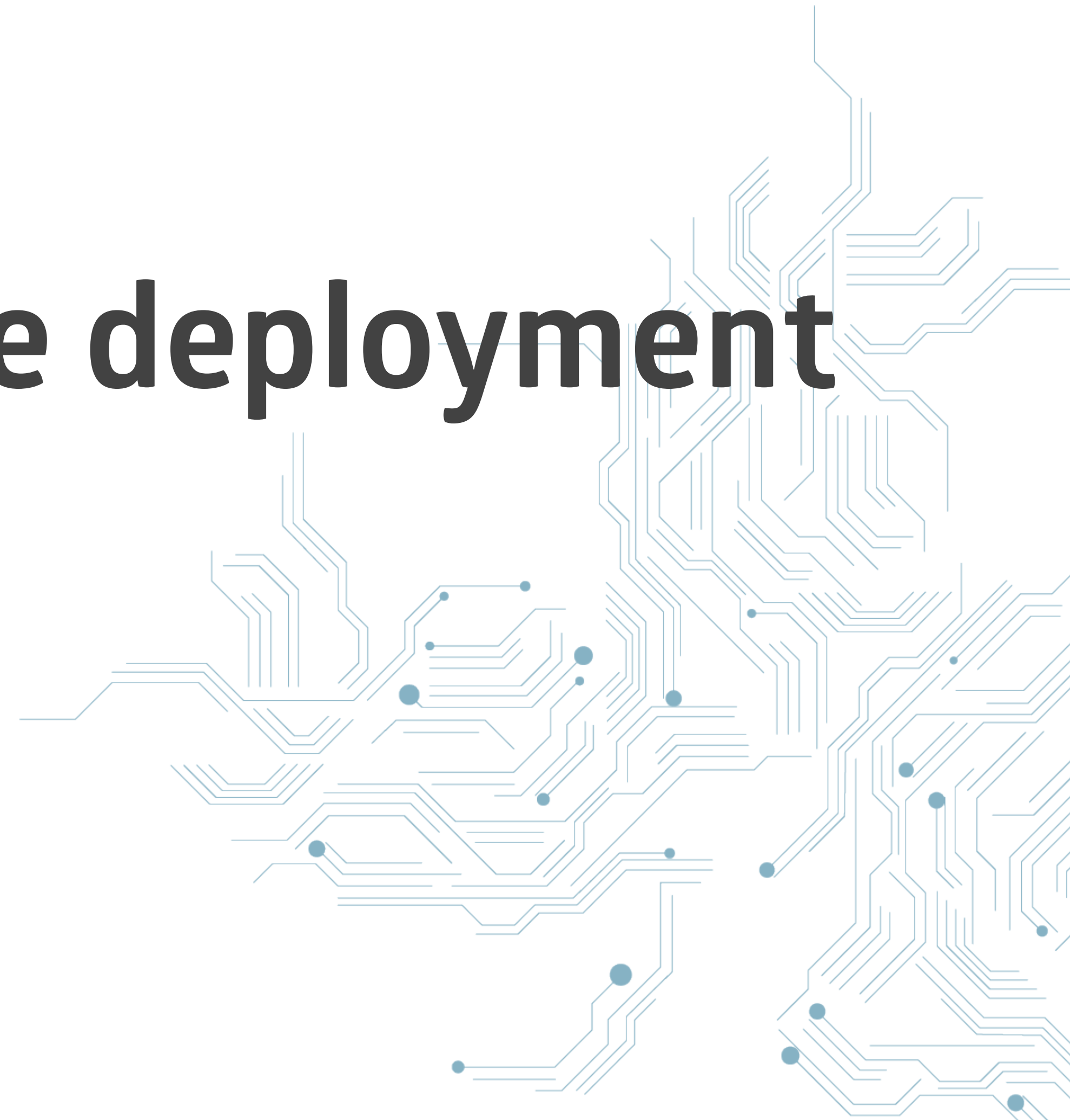
Nginx



You can mount storage across AZ



One cluster state => One deployment



**But what if the whole region is
down?**



**One Kubernetes cluster across
multiple regions and VPN connection
between networks**

A decorative graphic of a circuit board pattern, consisting of white lines and dots, is located in the bottom right corner of the slide.

Region 1

Region 2

Kubernetes Master Components

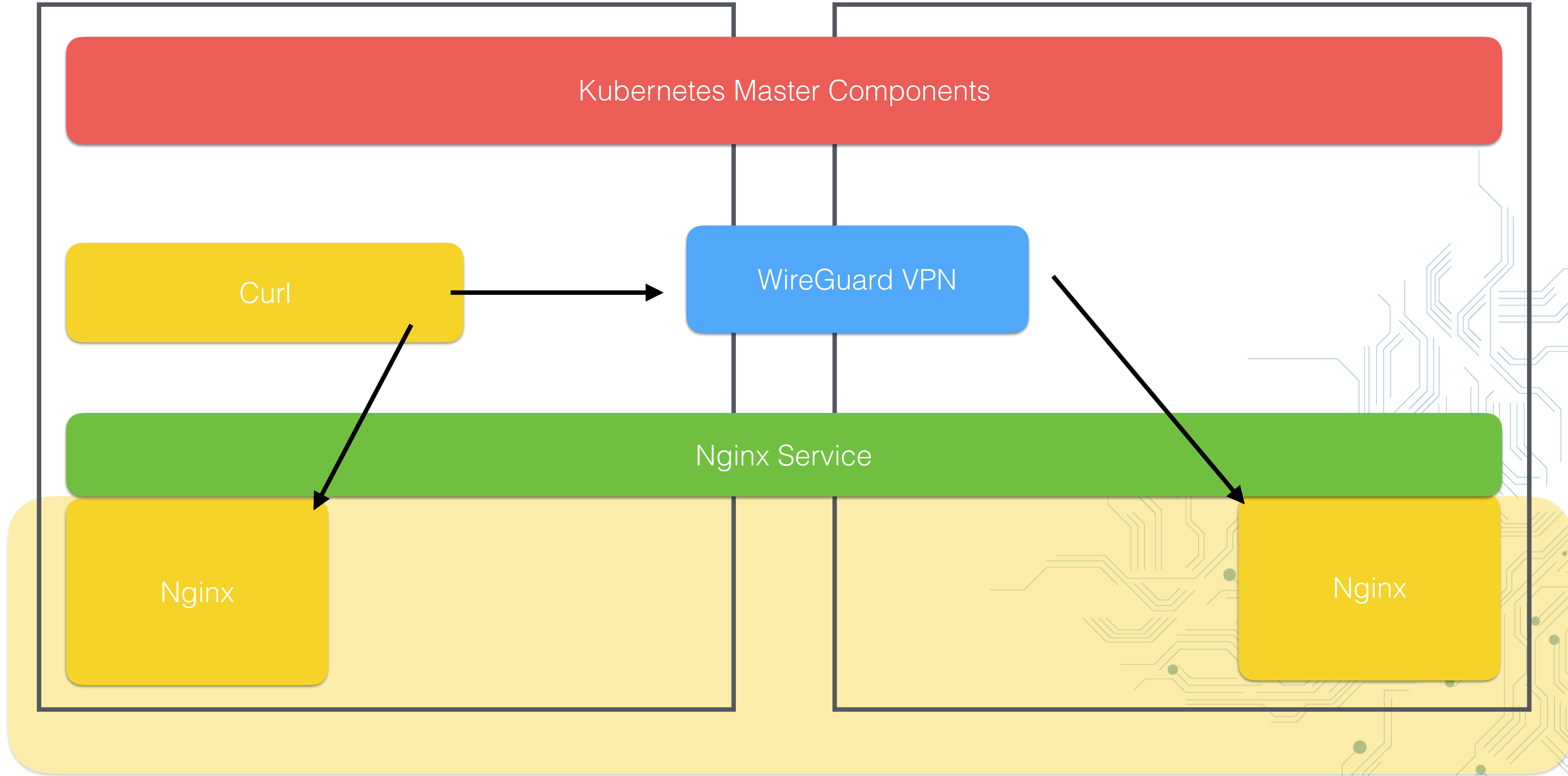
Curl

WireGuard VPN

Nginx Service

Nginx

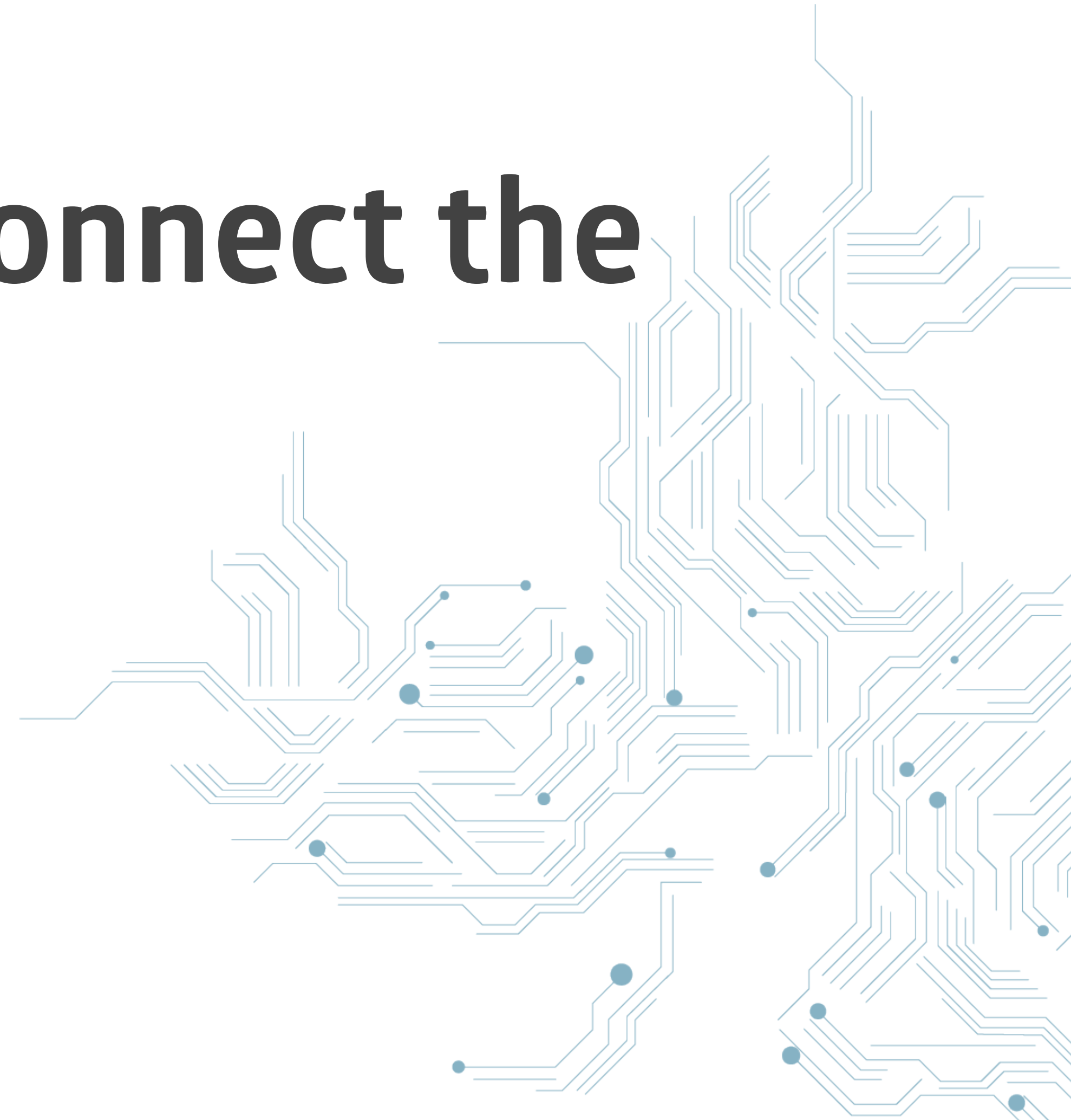
Nginx



Latencies get higher by distance



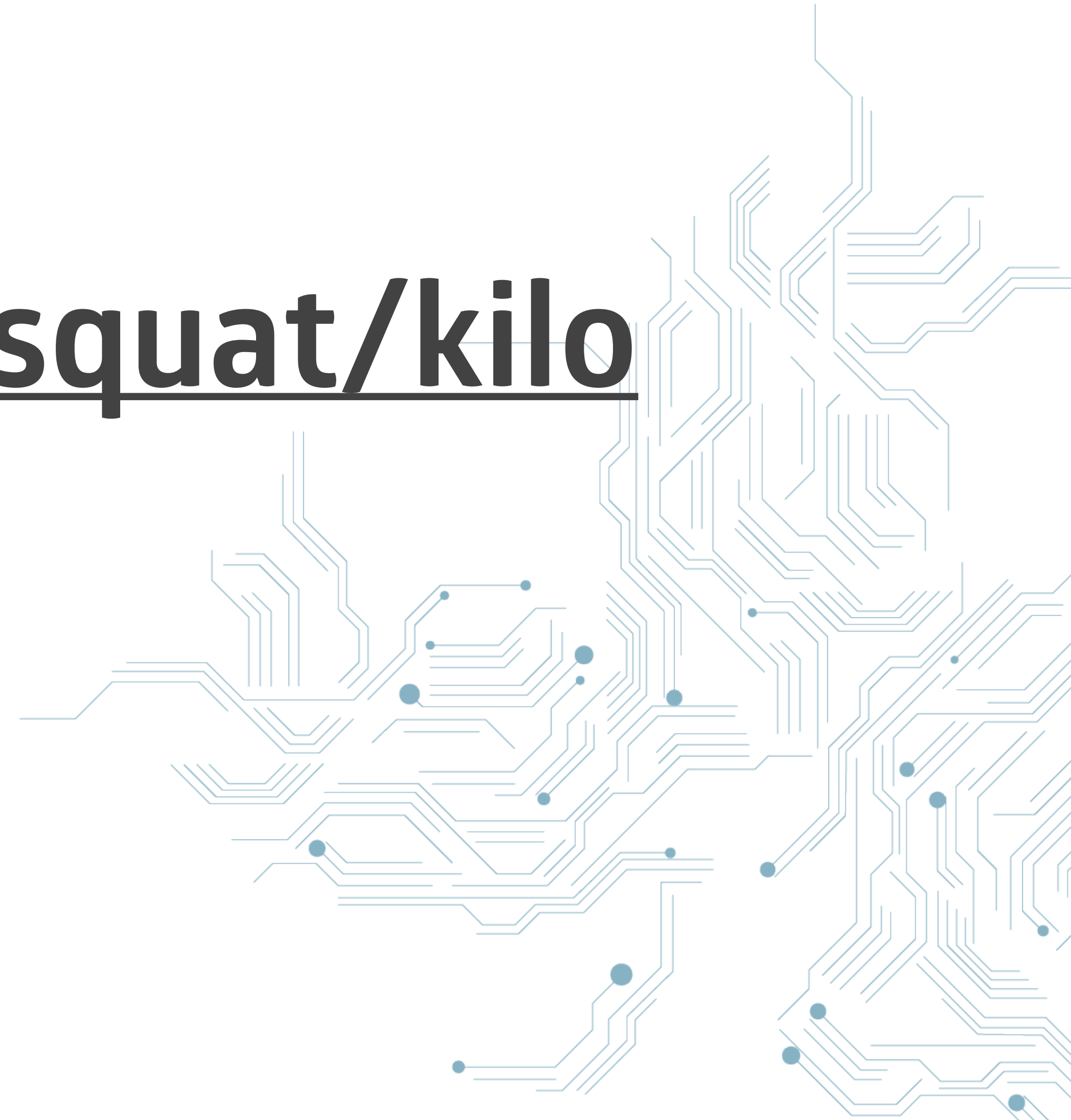
**You need a VPN to connect the
networks**



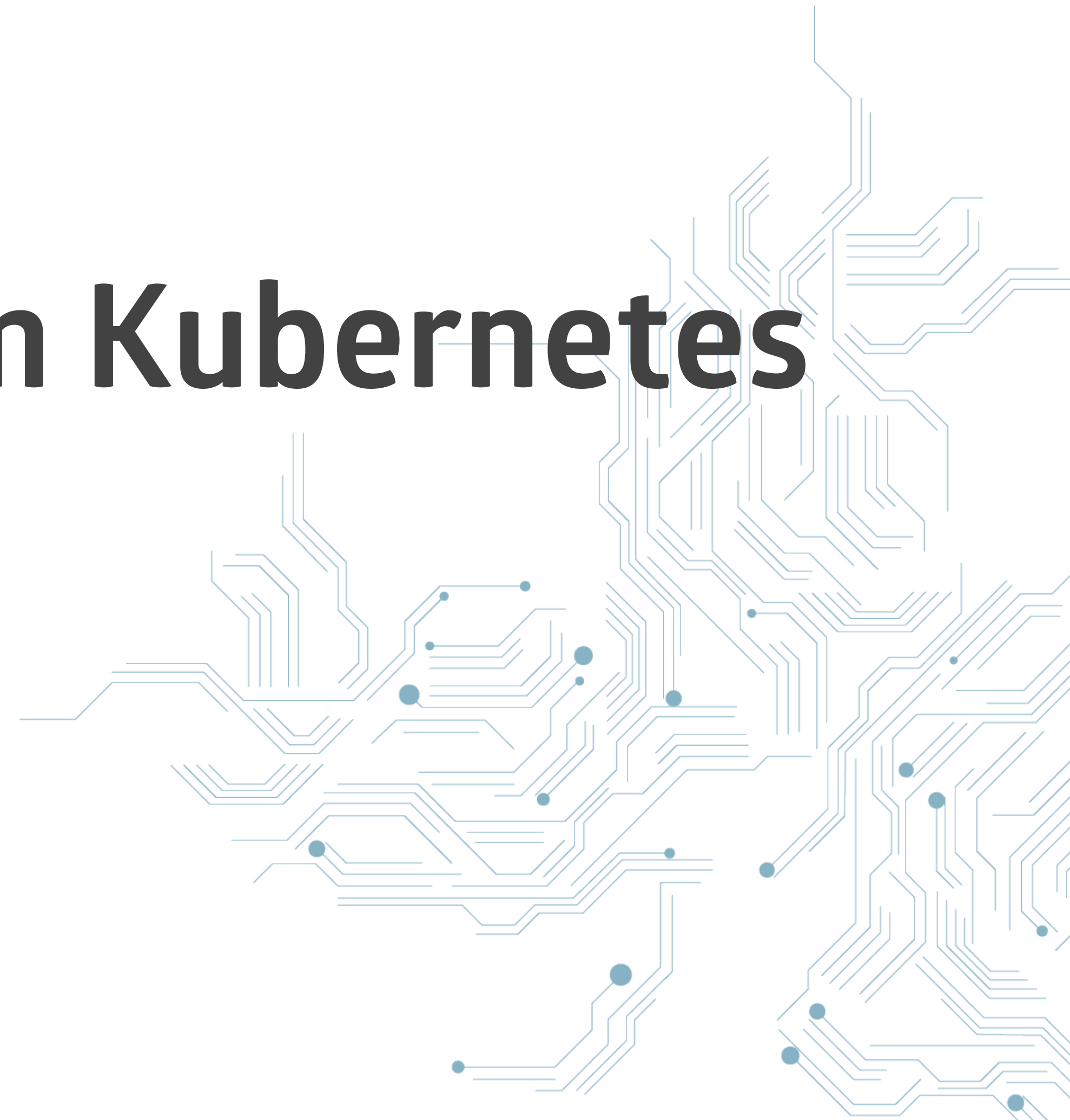
WireGuard as a VPN



<https://github.com/squat/kilo>



Kilo gets all Nodes from Kubernetes



**Discovers regions and zones by
Kubernetes Node labels**



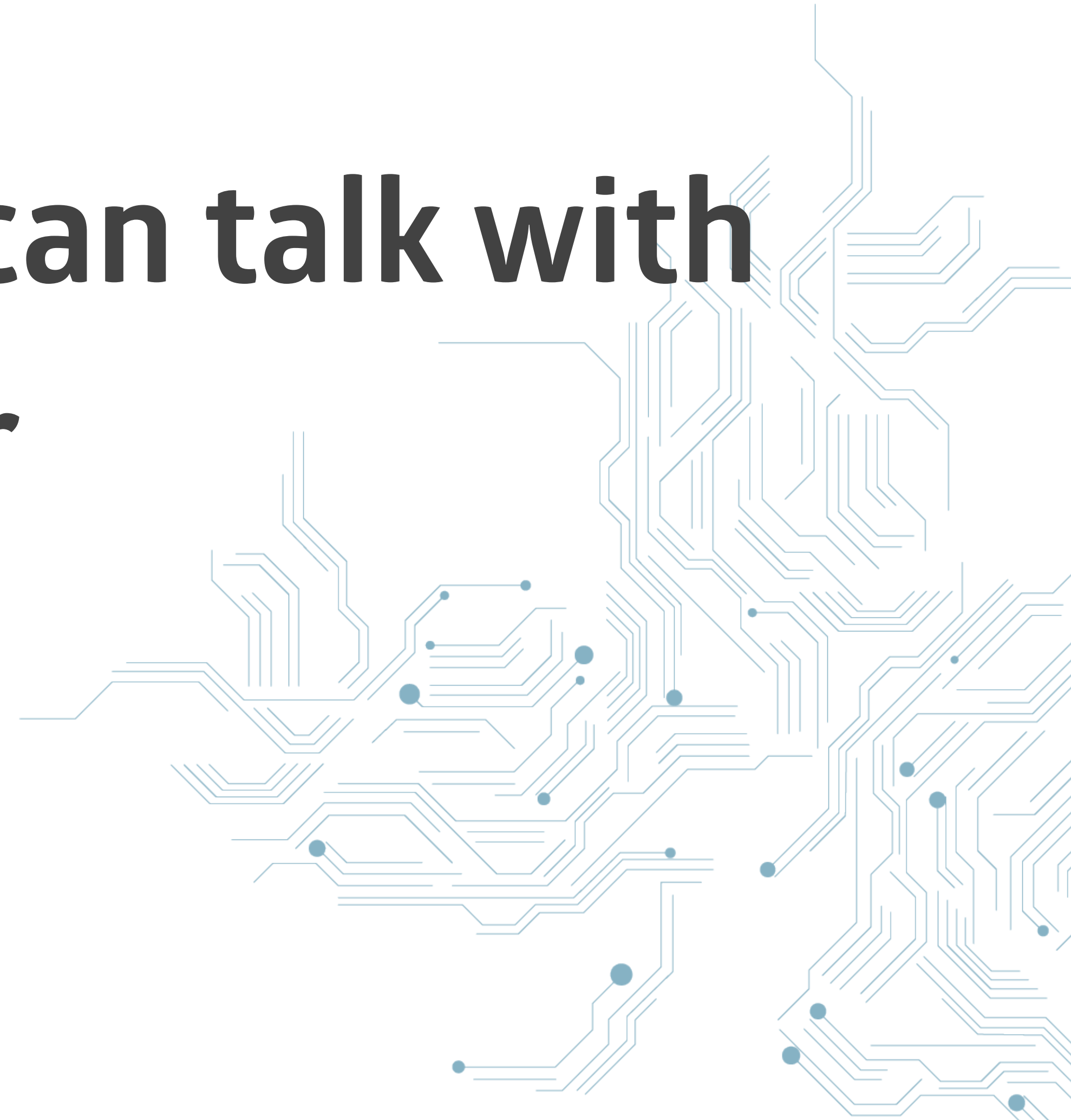
Configures WireGuard



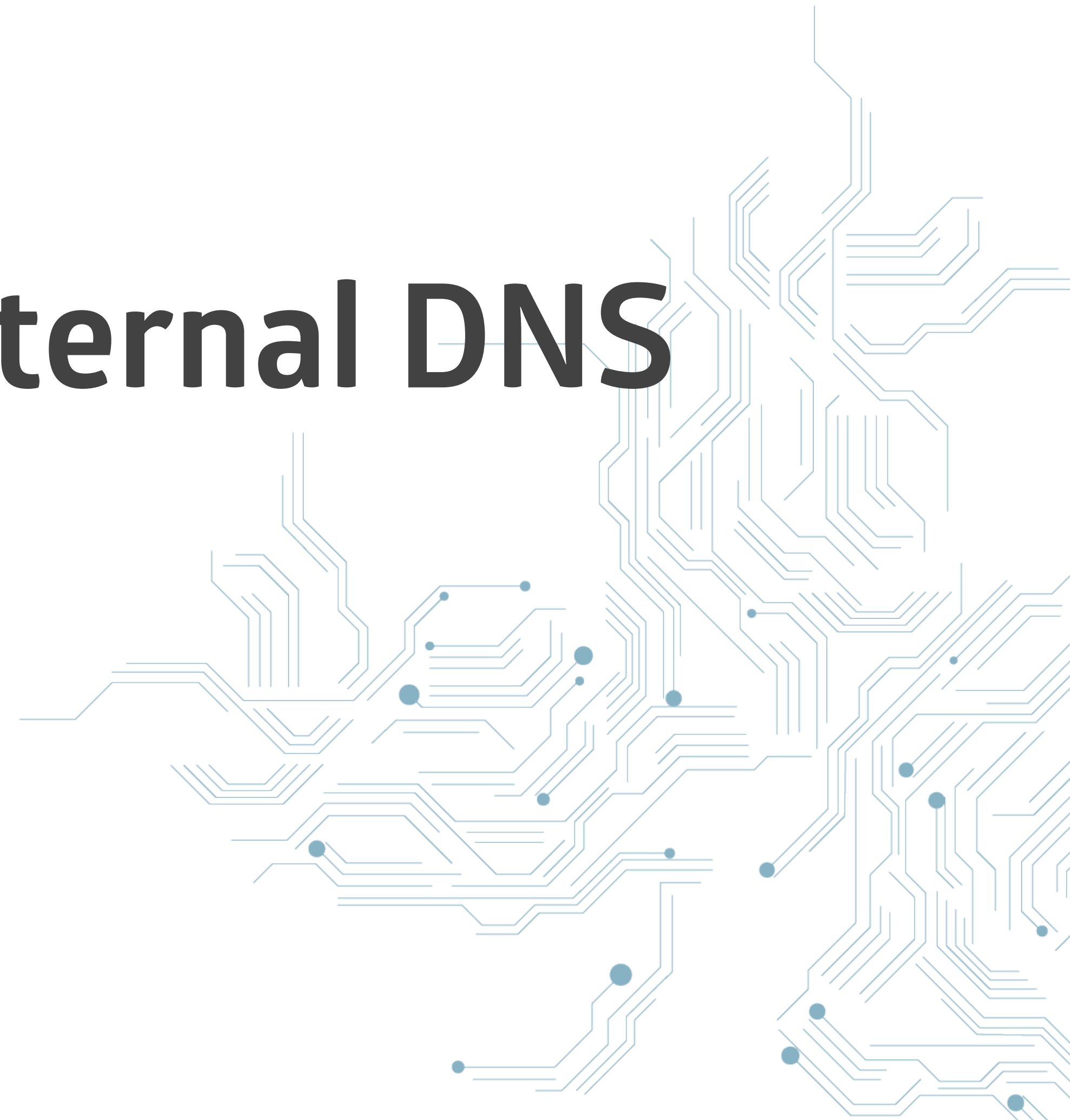
Routing tables



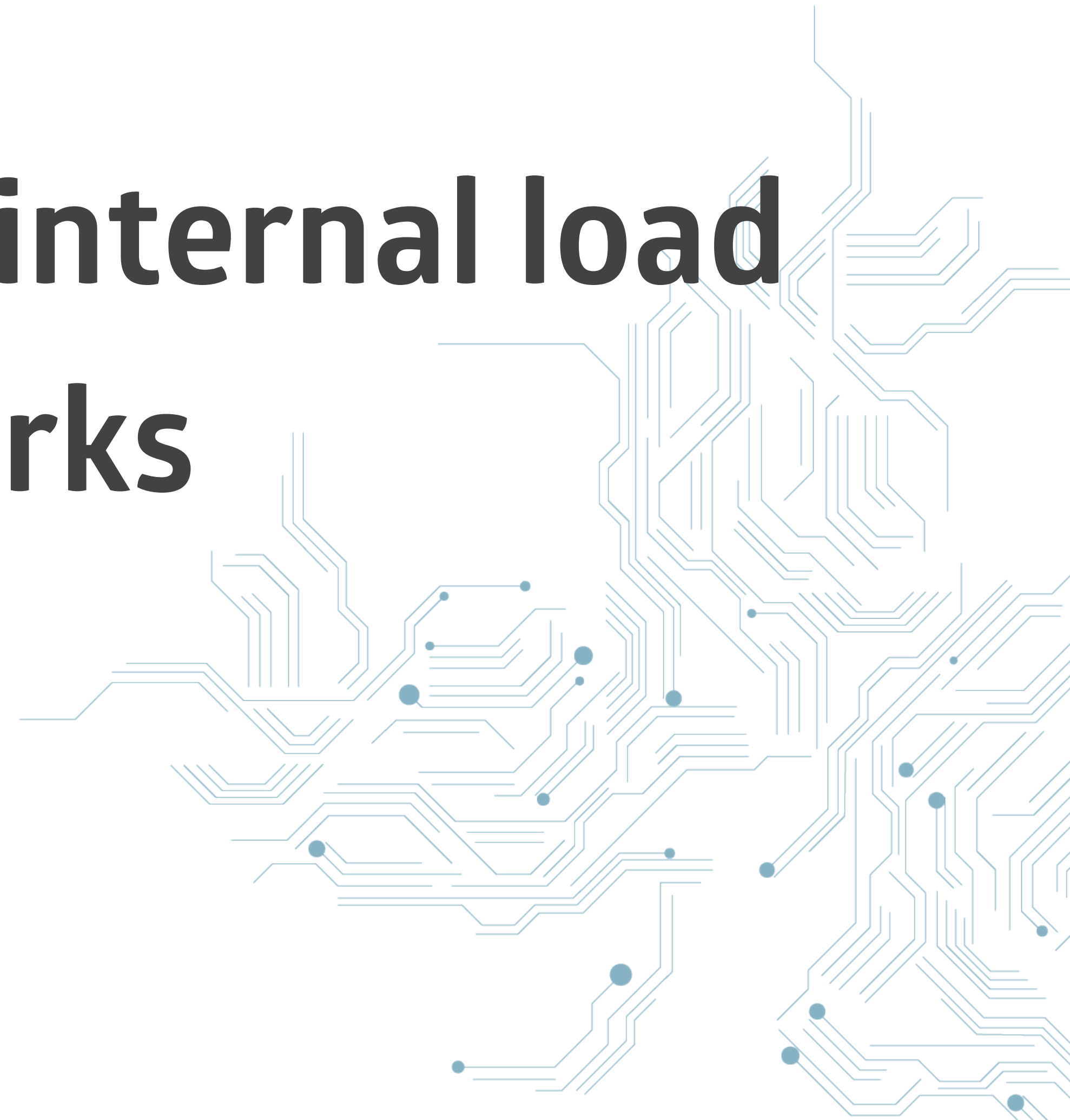
**All pods and services can talk with
each other**



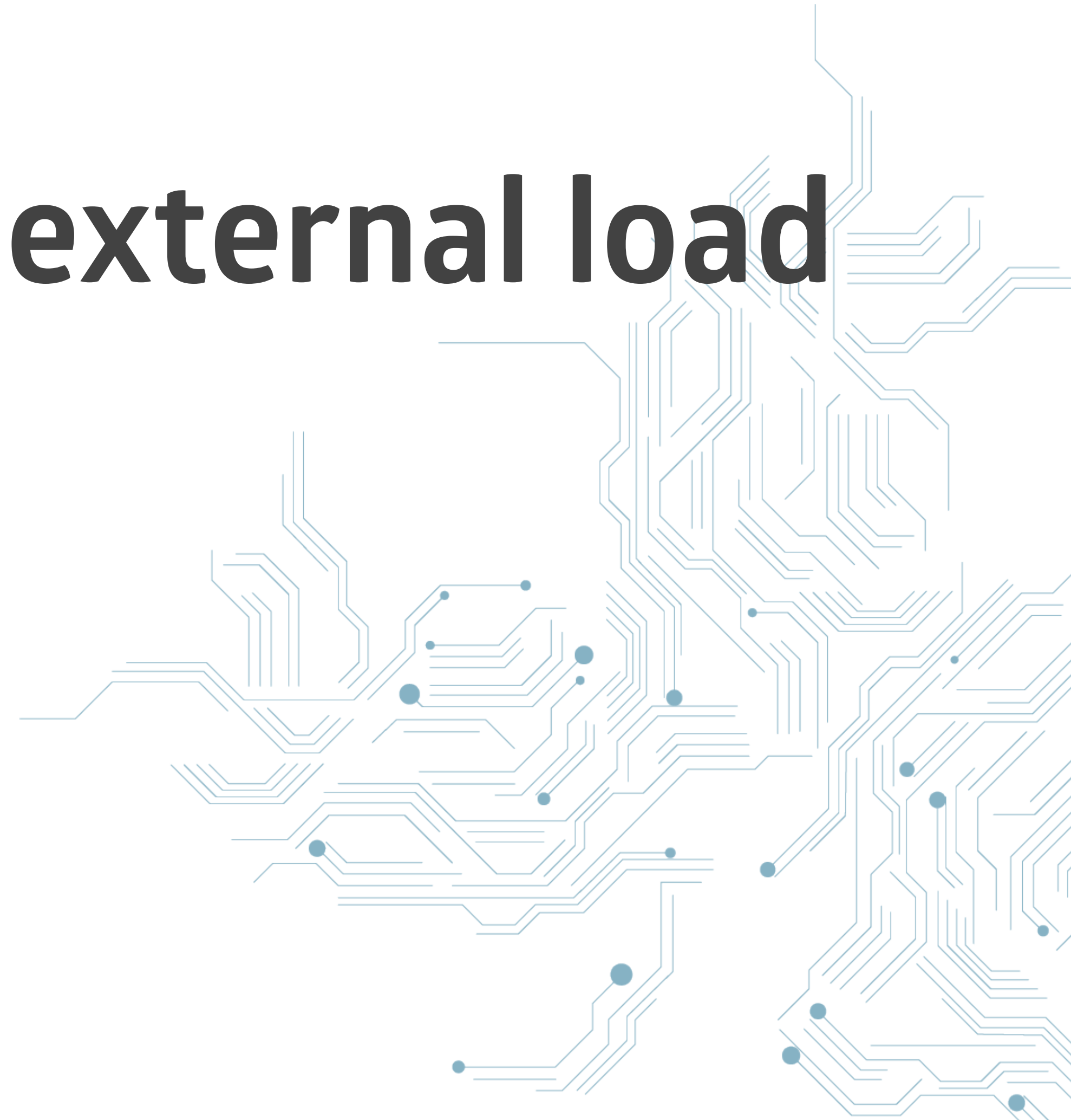
One cluster => one internal DNS



Service Discovery and internal load balancing works



**You need a solution for external load
balancing**



Region 1

Region 2

LB

LB

Kubernetes Master Components

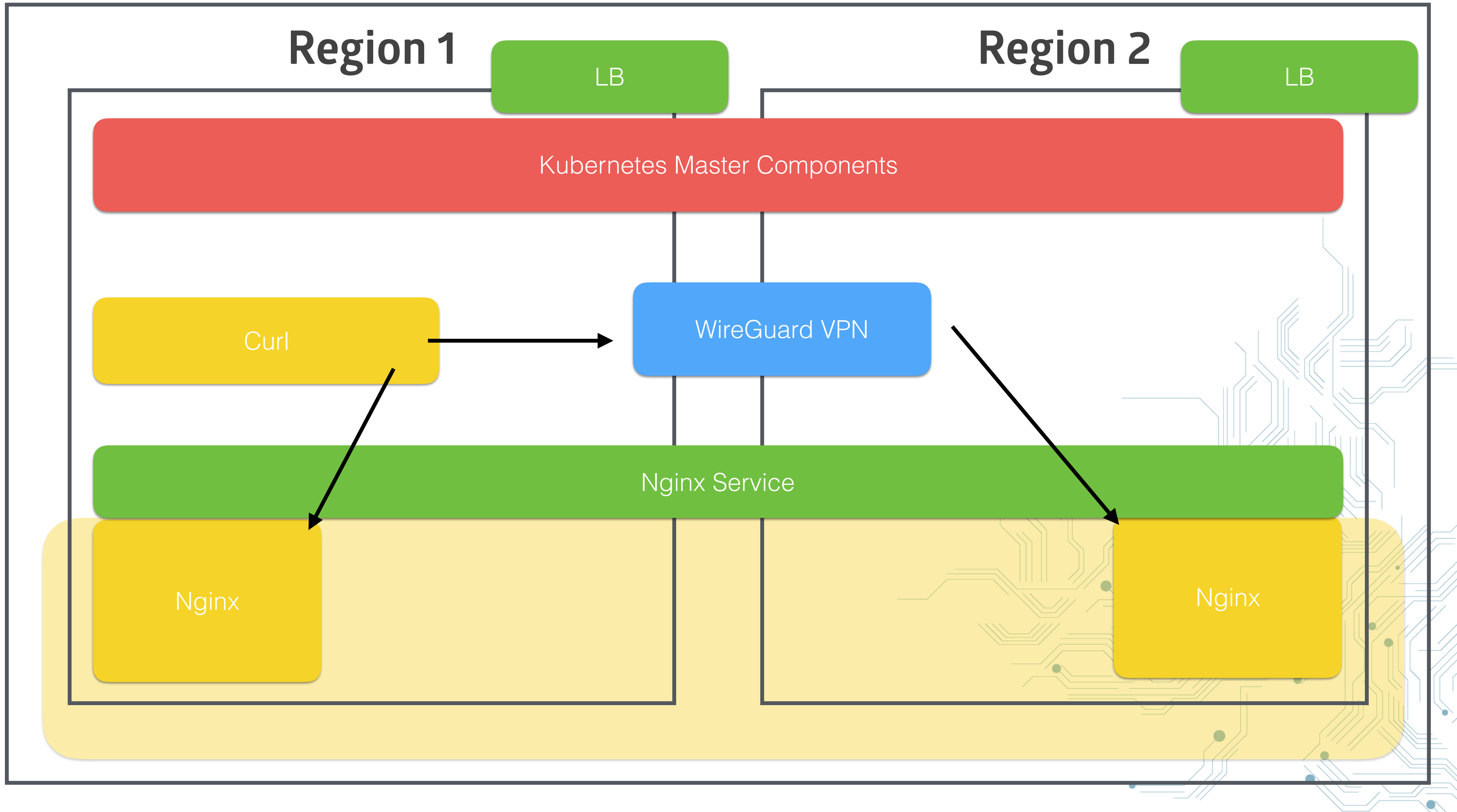
Curl

WireGuard VPN

Nginx Service

Nginx

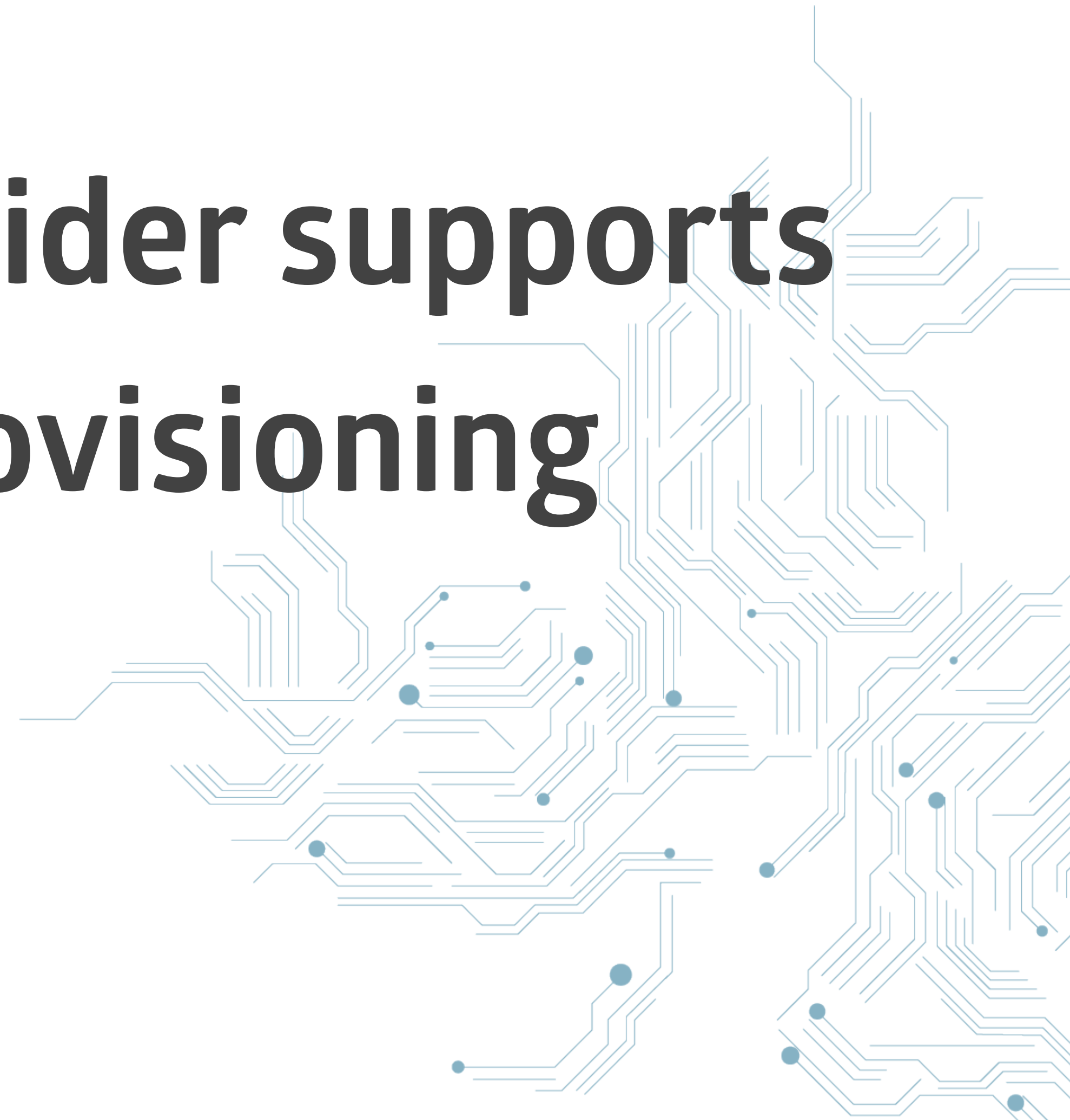
Nginx



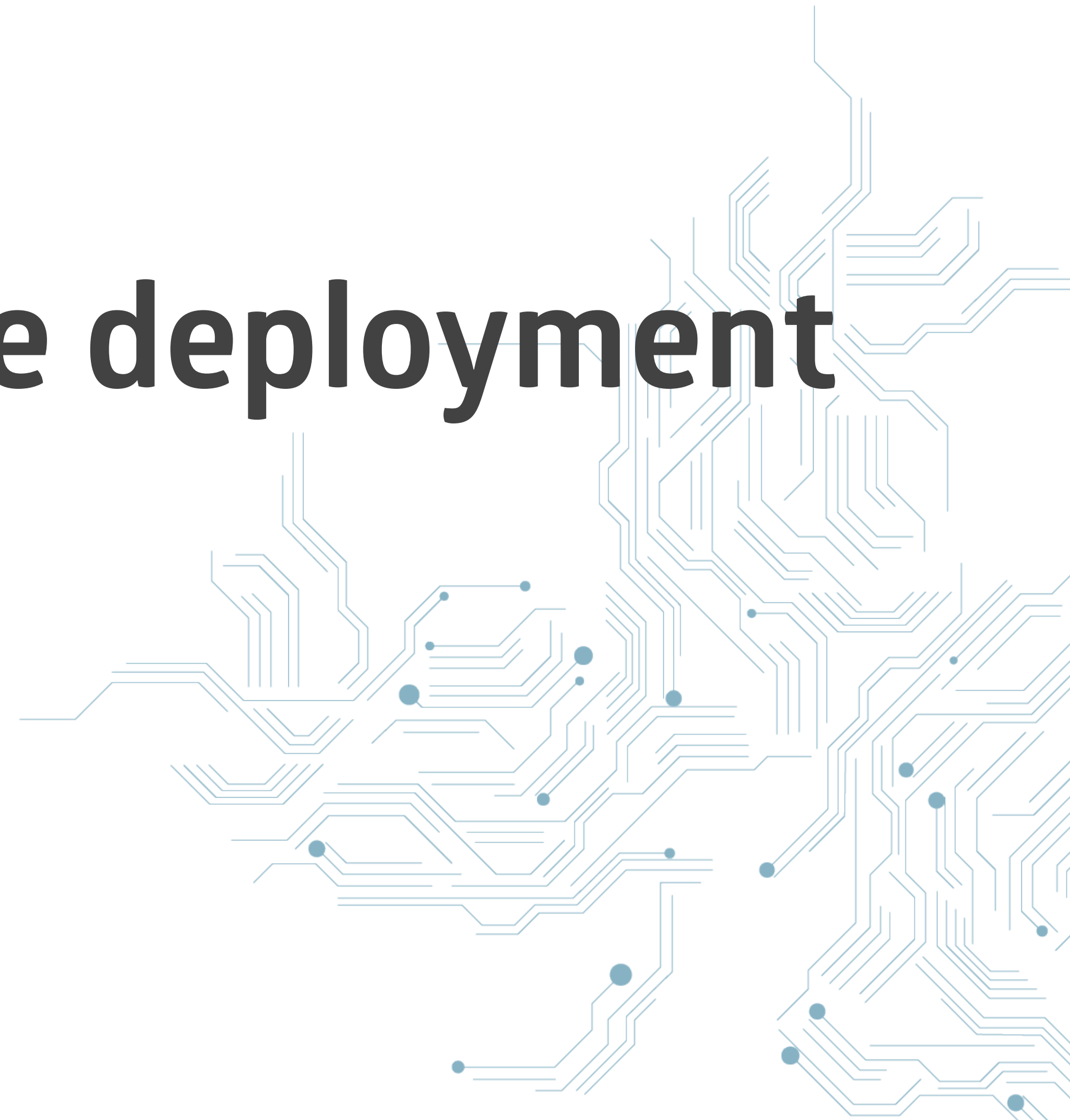
**You have to replicate storage
yourself**



**Not every Storage provider supports
dynamic volume provisioning**



One cluster state => One deployment



Demo



Multi cluster setup

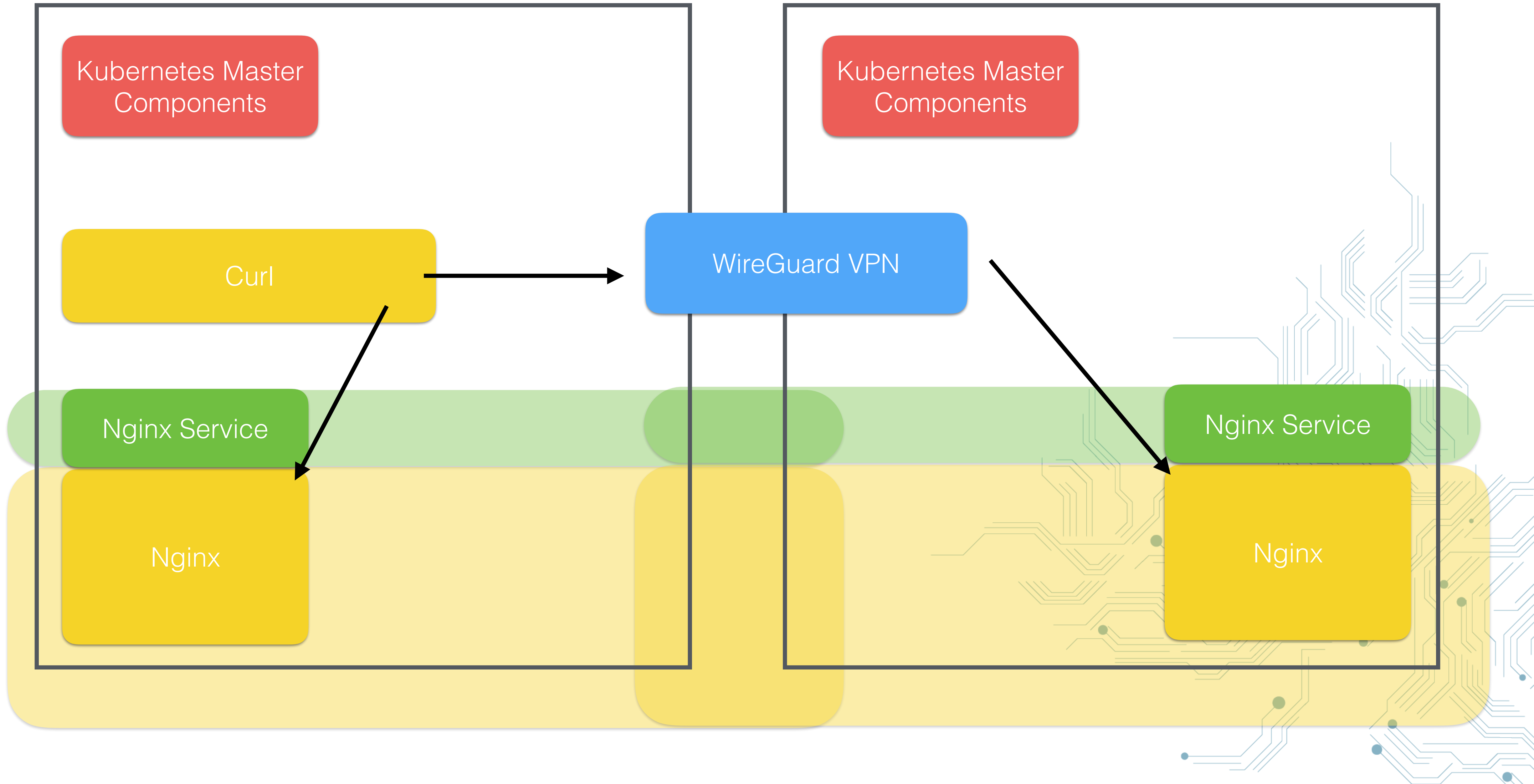


Connecting multiple clusters with a VPN



Cluster 1

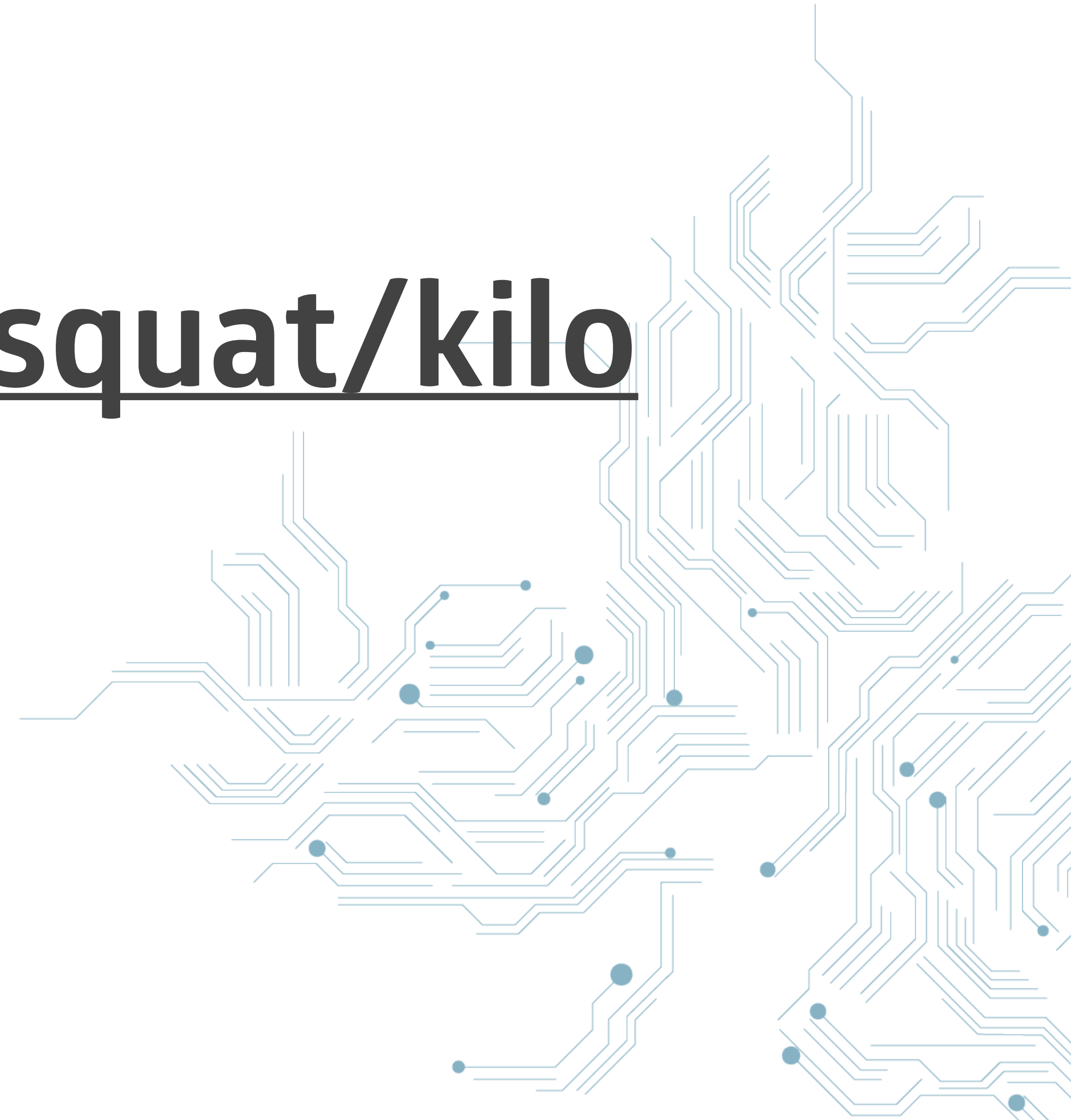
Cluster 2



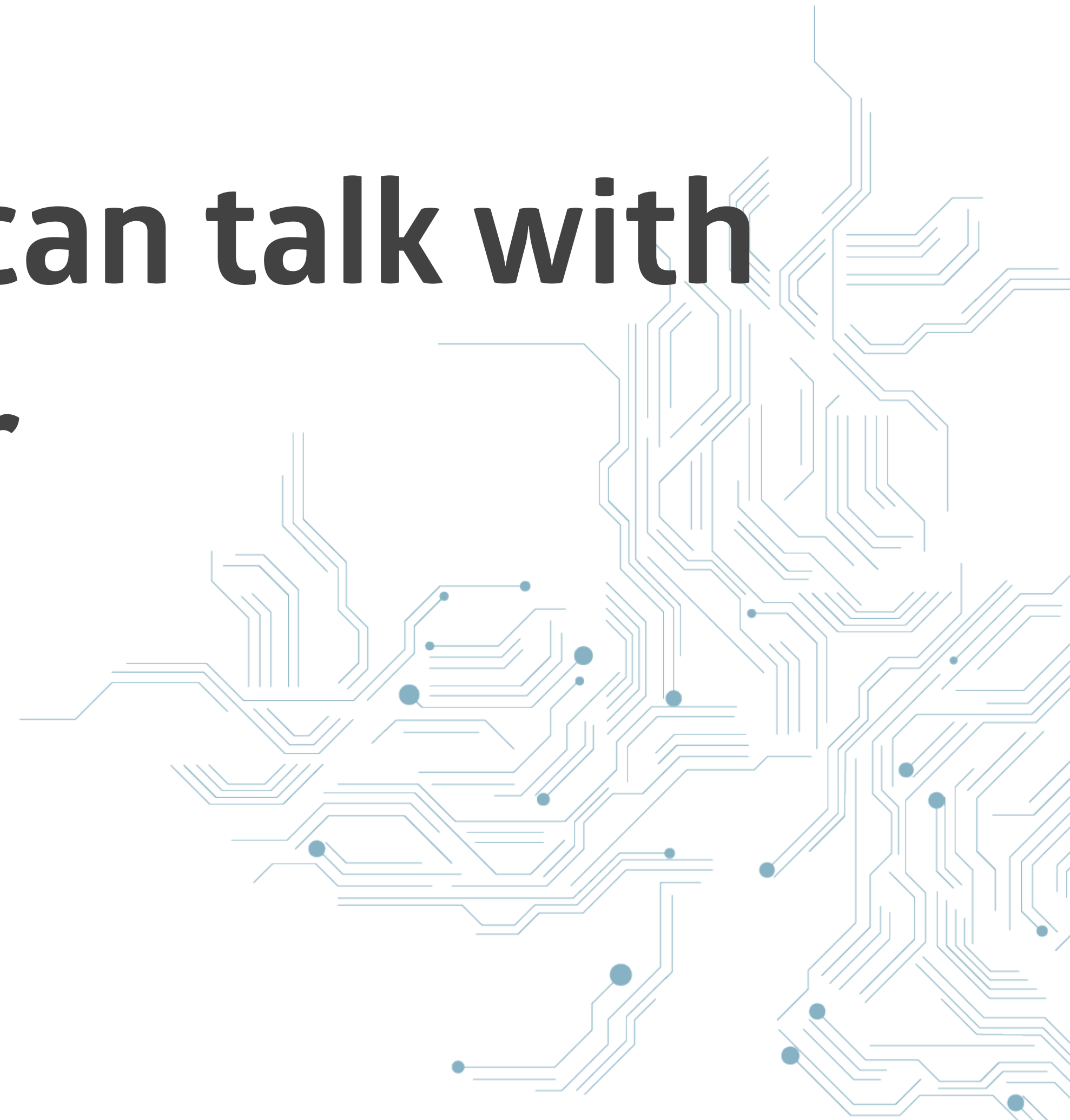
WireGuard as a VPN



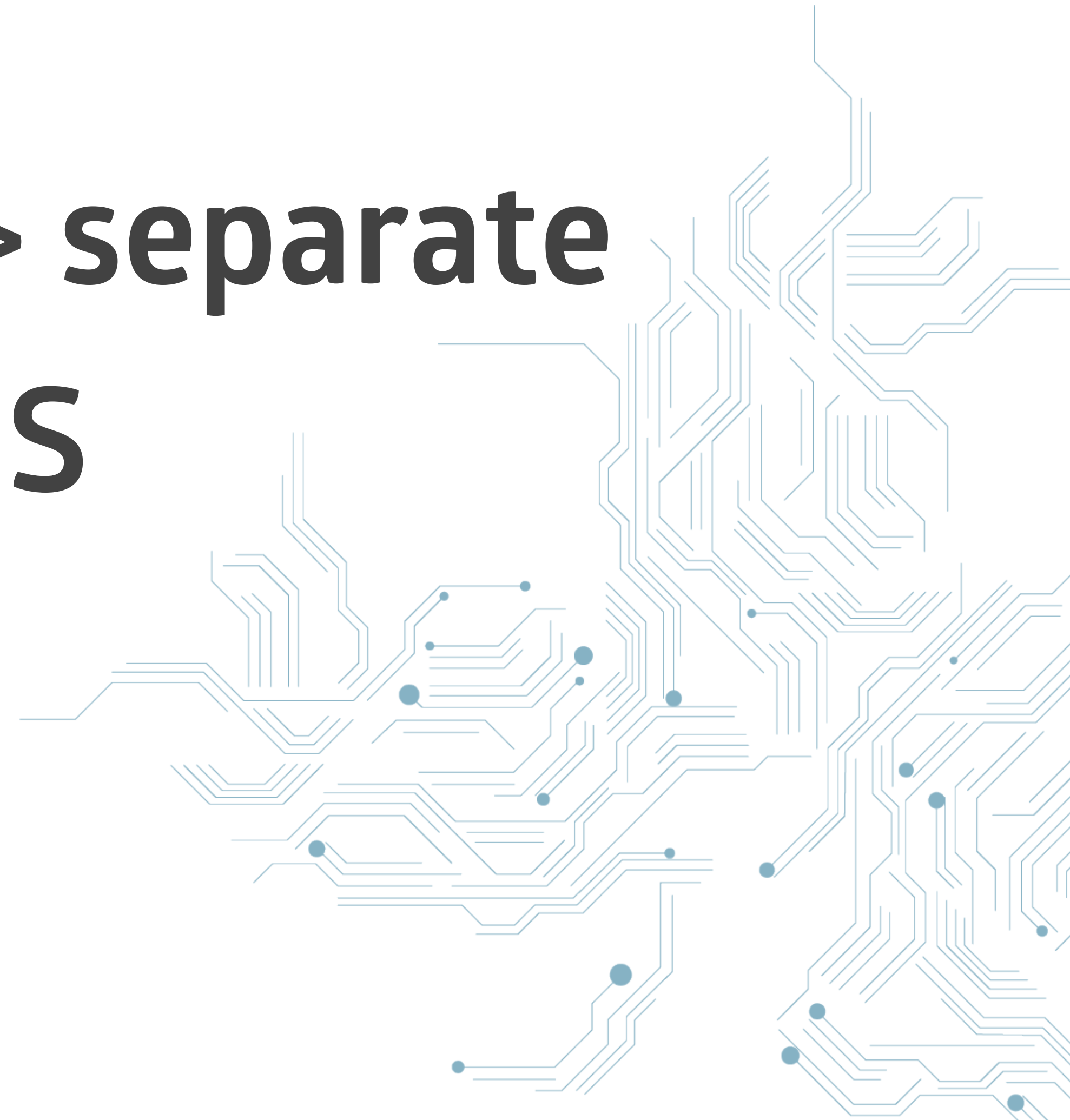
<https://github.com/squat/kilo>



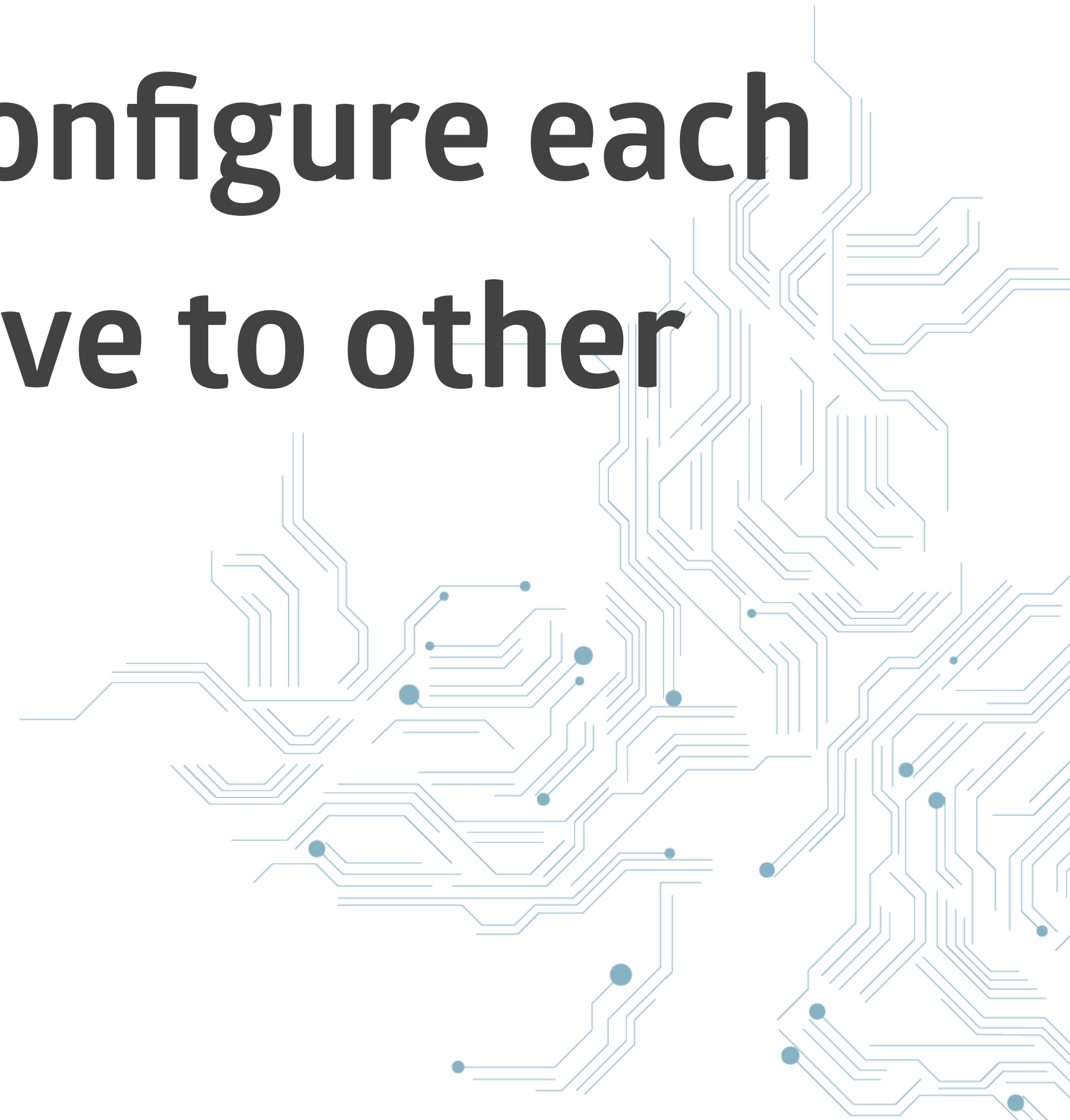
**All pods and services can talk with
each other**



**Separate clusters => separate
internal DNS**



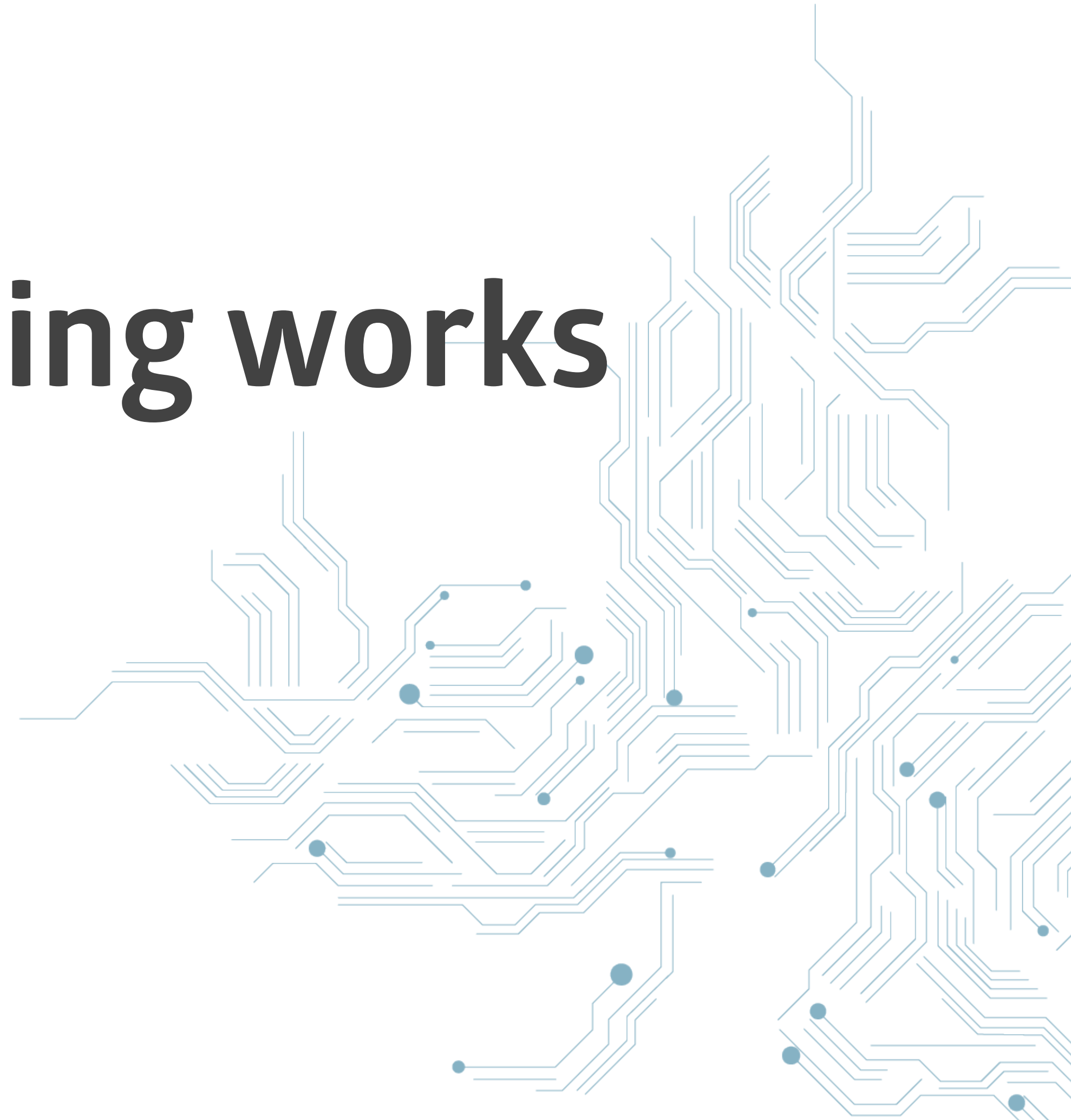
**For service discovery configure each
internal DNS to resolve to other
clusters**



```
cluster.region2:53 {  
    forward . 10.10.11.10  
}  
.:53 {  
    kubernetes cluster.local cluster.region1 in-addr.arpa  
ip6.arpa {  
    pods insecure  
    upstream  
    fallthrough in-addr.arpa ip6.arpa  
}  
forward . /etc/resolv.conf  
loop  
loadbalance  
}
```

```
cluster.region1:53 {  
    forward . 10.10.10.10  
}  
.:53 {  
    kubernetes cluster.local cluster.region2 in-addr.arpa  
ip6.arpa {  
    pods insecure  
    upstream  
    fallthrough in-addr.arpa ip6.arpa  
}  
forward . /etc/resolv.conf  
loop  
loadbalance  
}
```

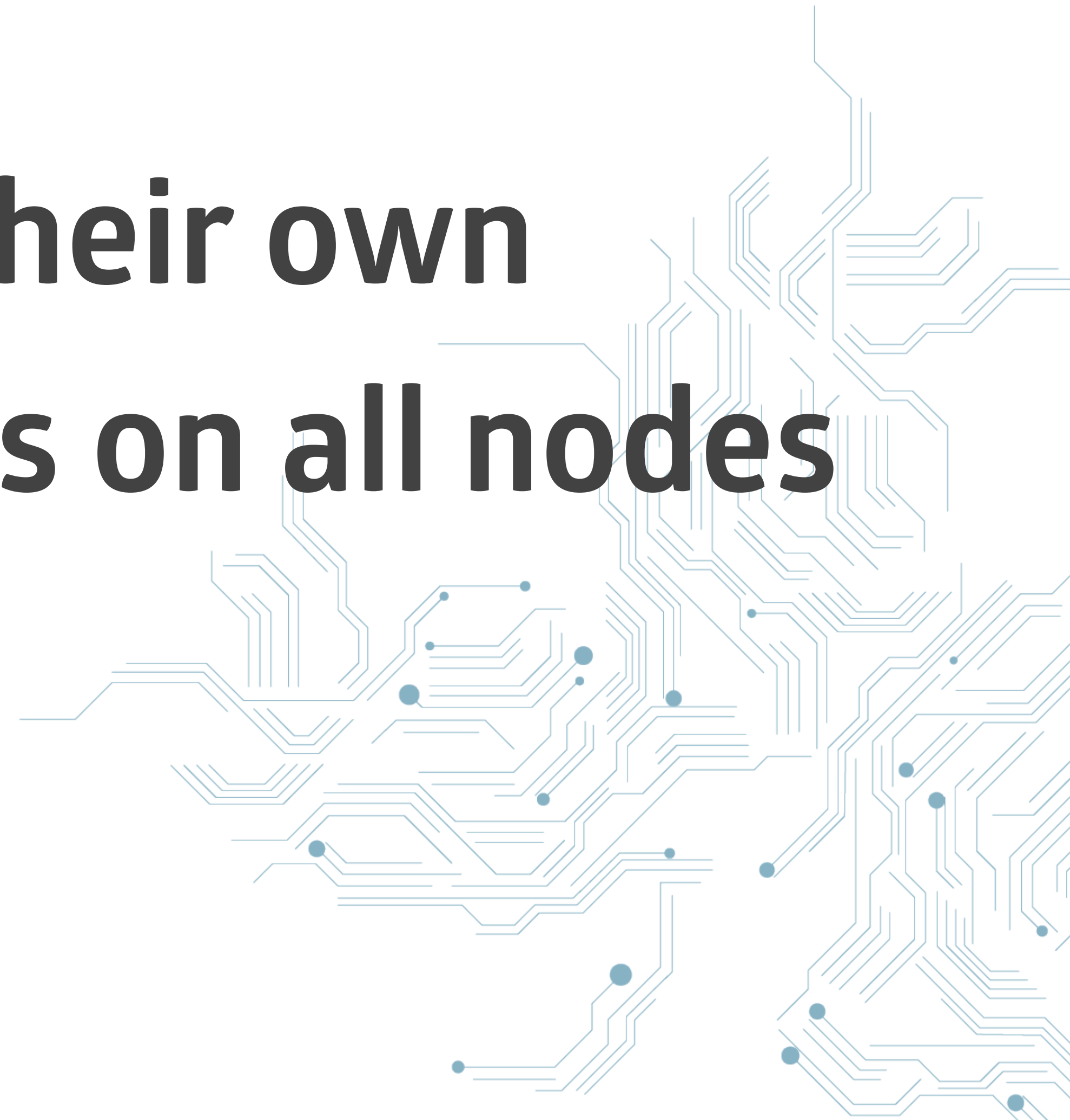
Internal load balancing works



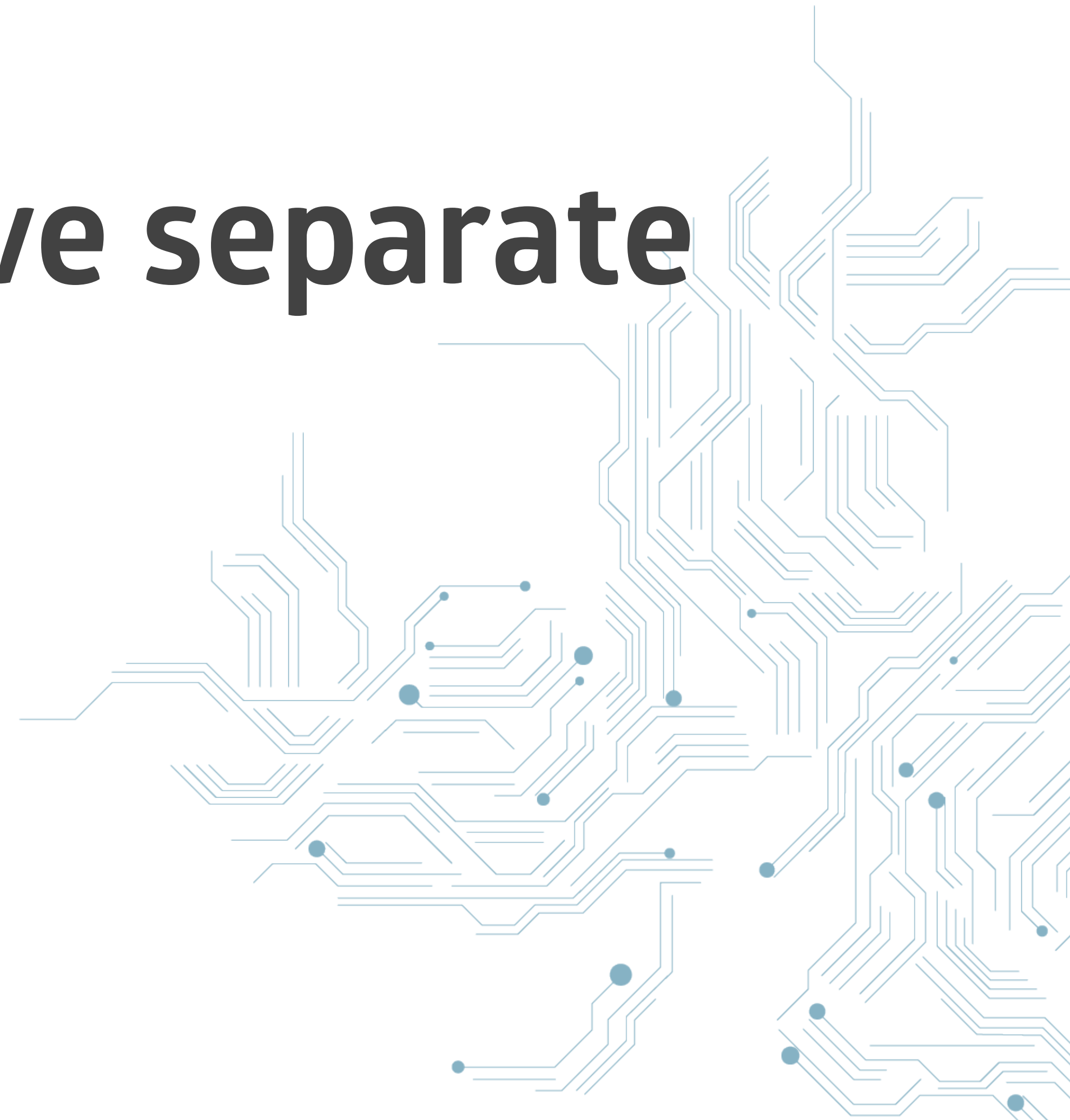
**You have to replicate storage
yourself**



**Every cluster has their own
StorageClass that works on all nodes**



**Separate clusters have separate
state**



Management of deployments across clusters

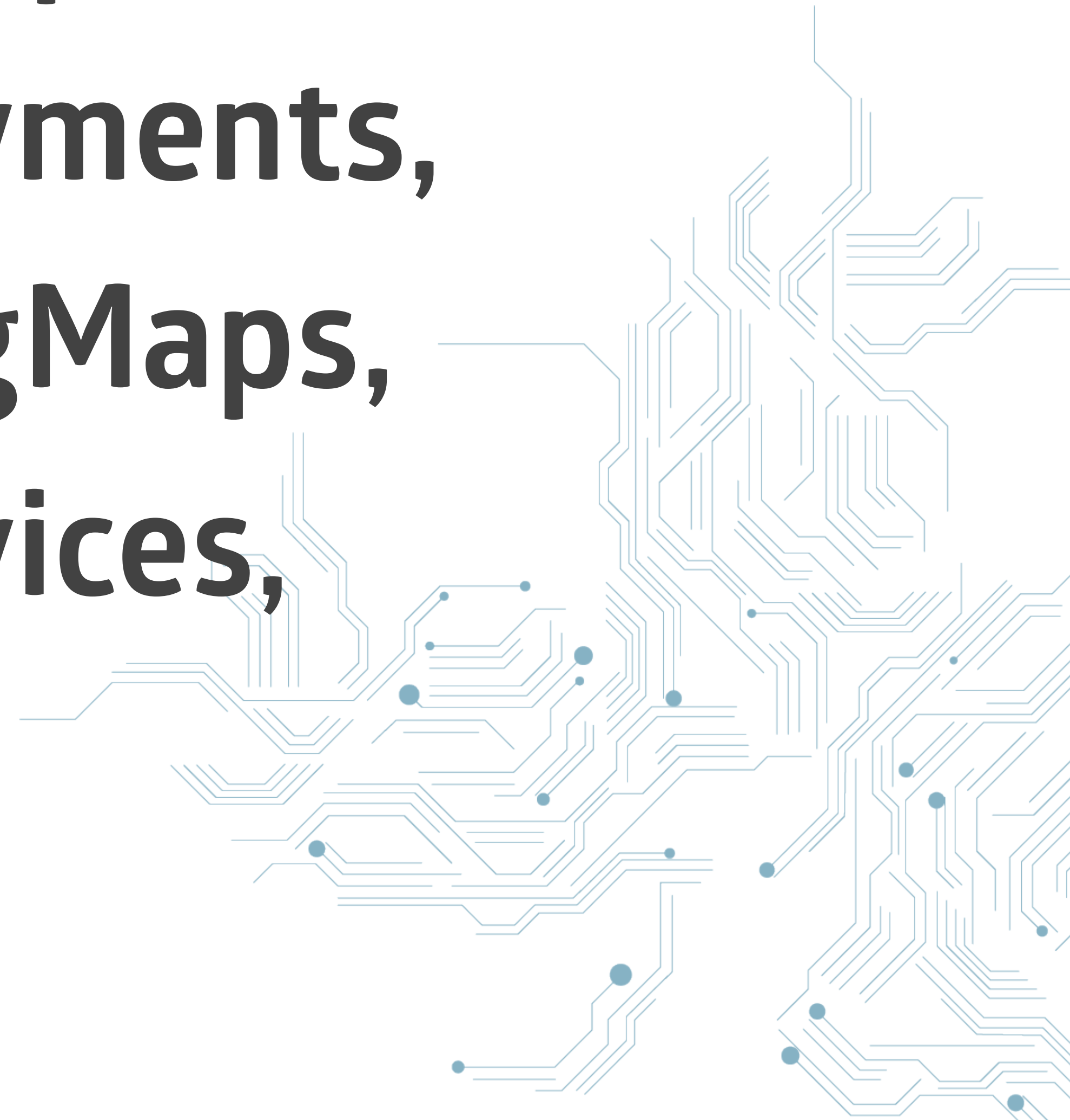


Kubefed

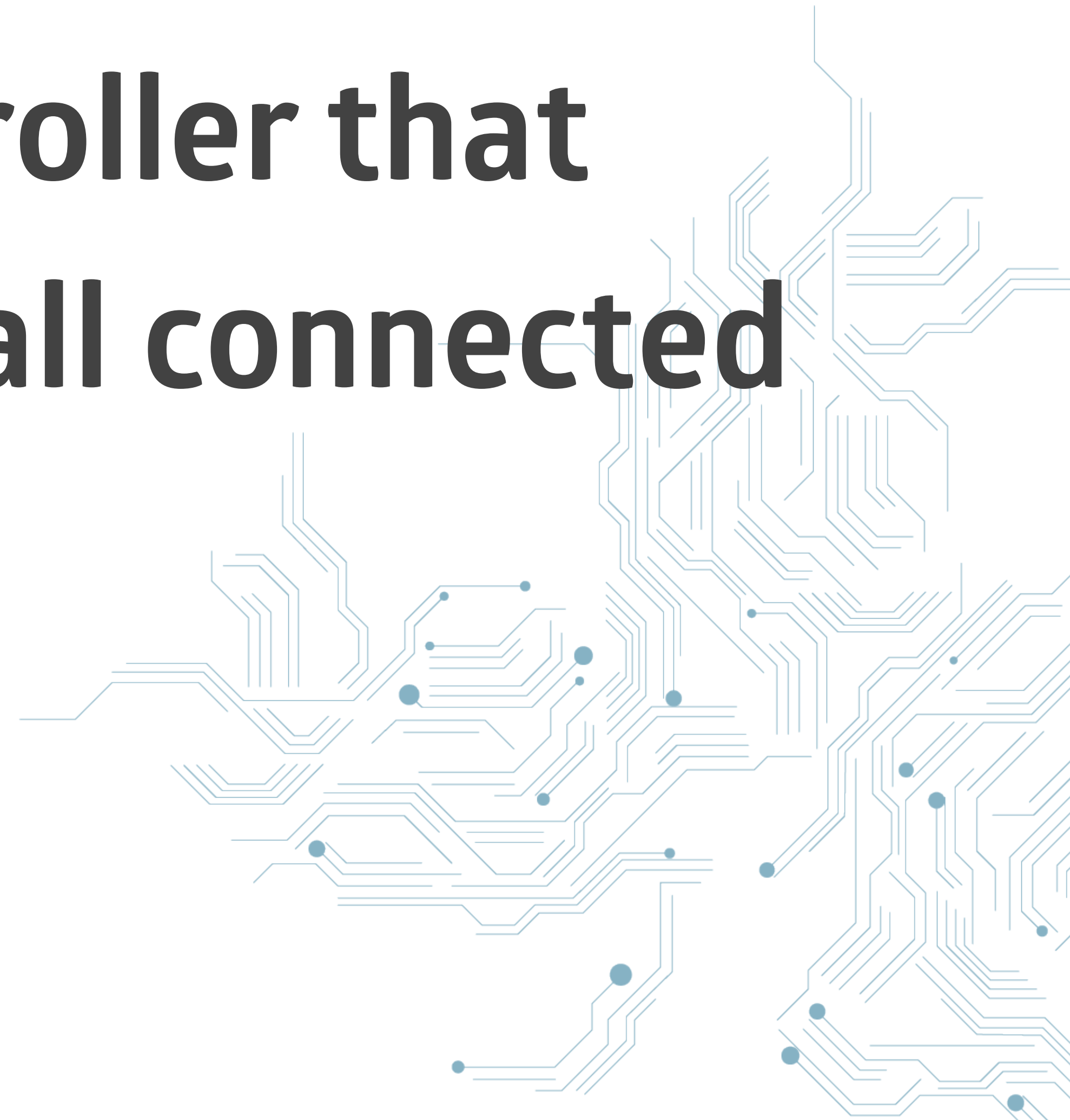
<https://github.com/kubernetes-sigs/kubefed>

**FederatedNamespaces,
FederatedDeployments,
FederatedConfigMaps,
FederatedServices,**

...

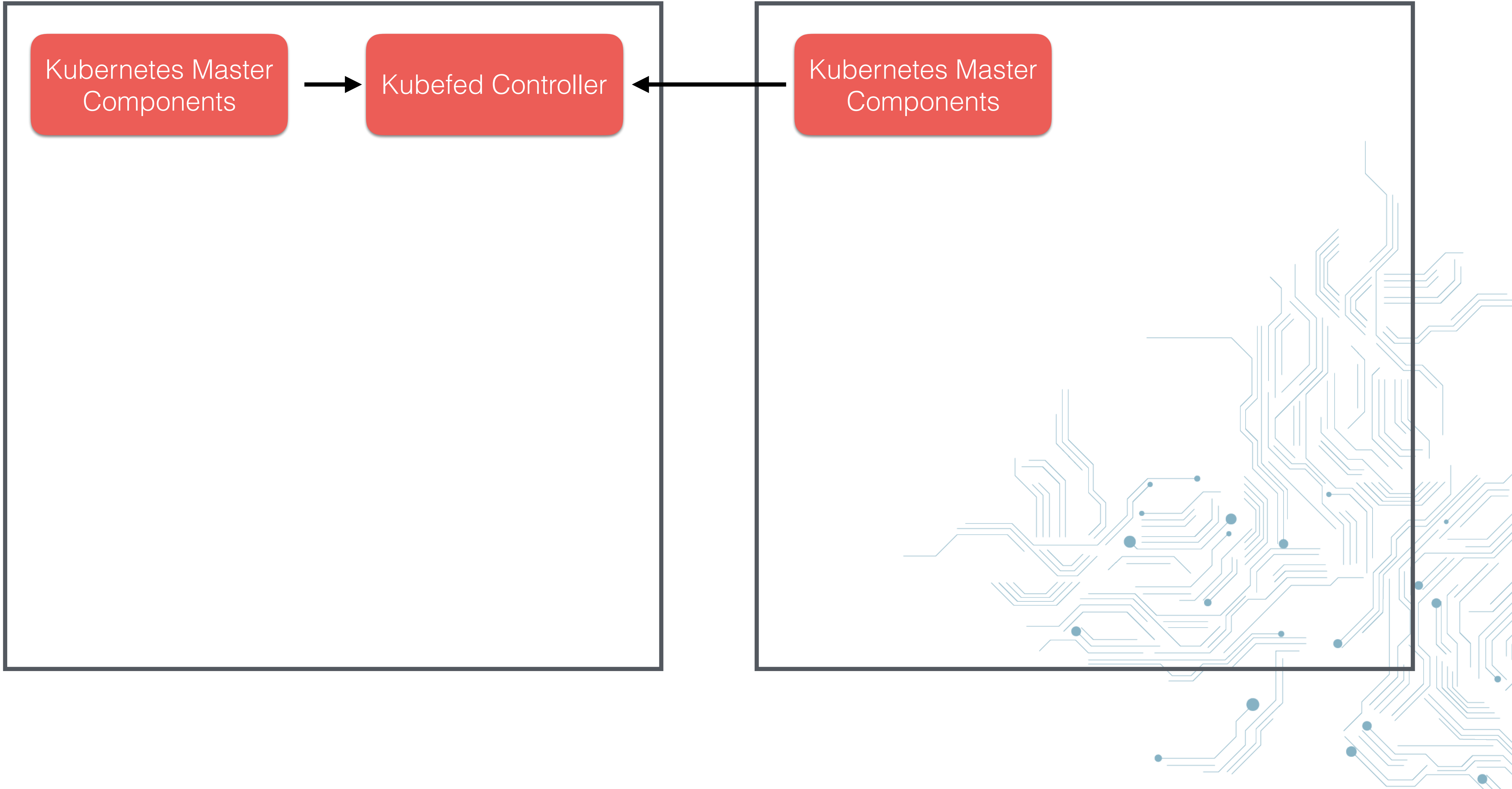


**Cluster aware controller that
manages resources in all connected
clusters**



Cluster 1

Cluster 2



Cluster 1

Kubernetes Master
Components

Kubefed Controller

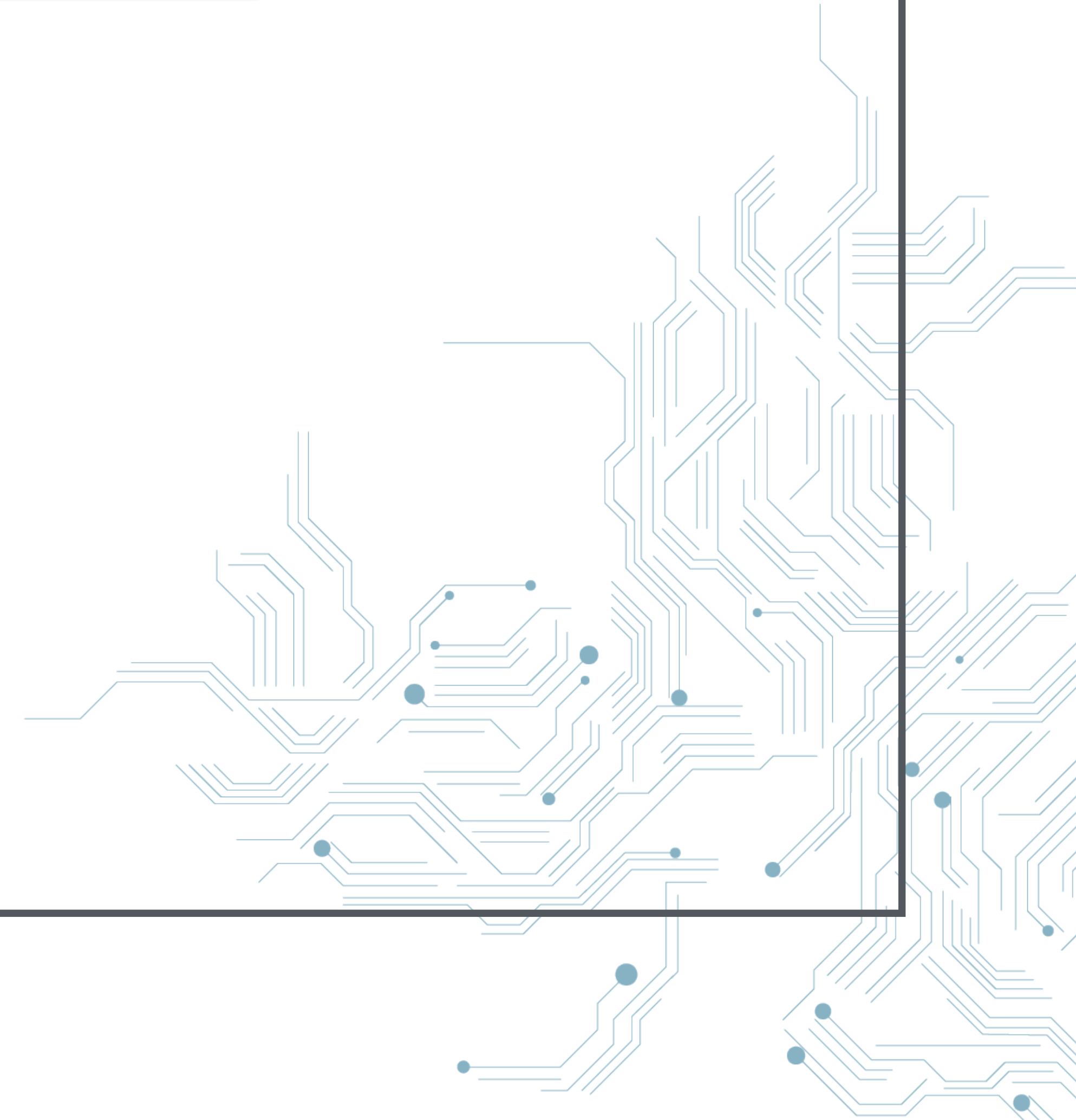


FederatedService

FederatedDeployment

Cluster 2

Kubernetes Master
Components



Cluster 1

Cluster 2

Kubernetes Master
Components

Kubefed Controller

Kubernetes Master
Components

Kubefed Controller

Nginx Service

Nginx

Nginx Service

Nginx



Demo



Solution for external traffic



Cluster 1

Cluster 2

LB

LB

Kubernetes Master Components

Kubernetes Master Components

Curl

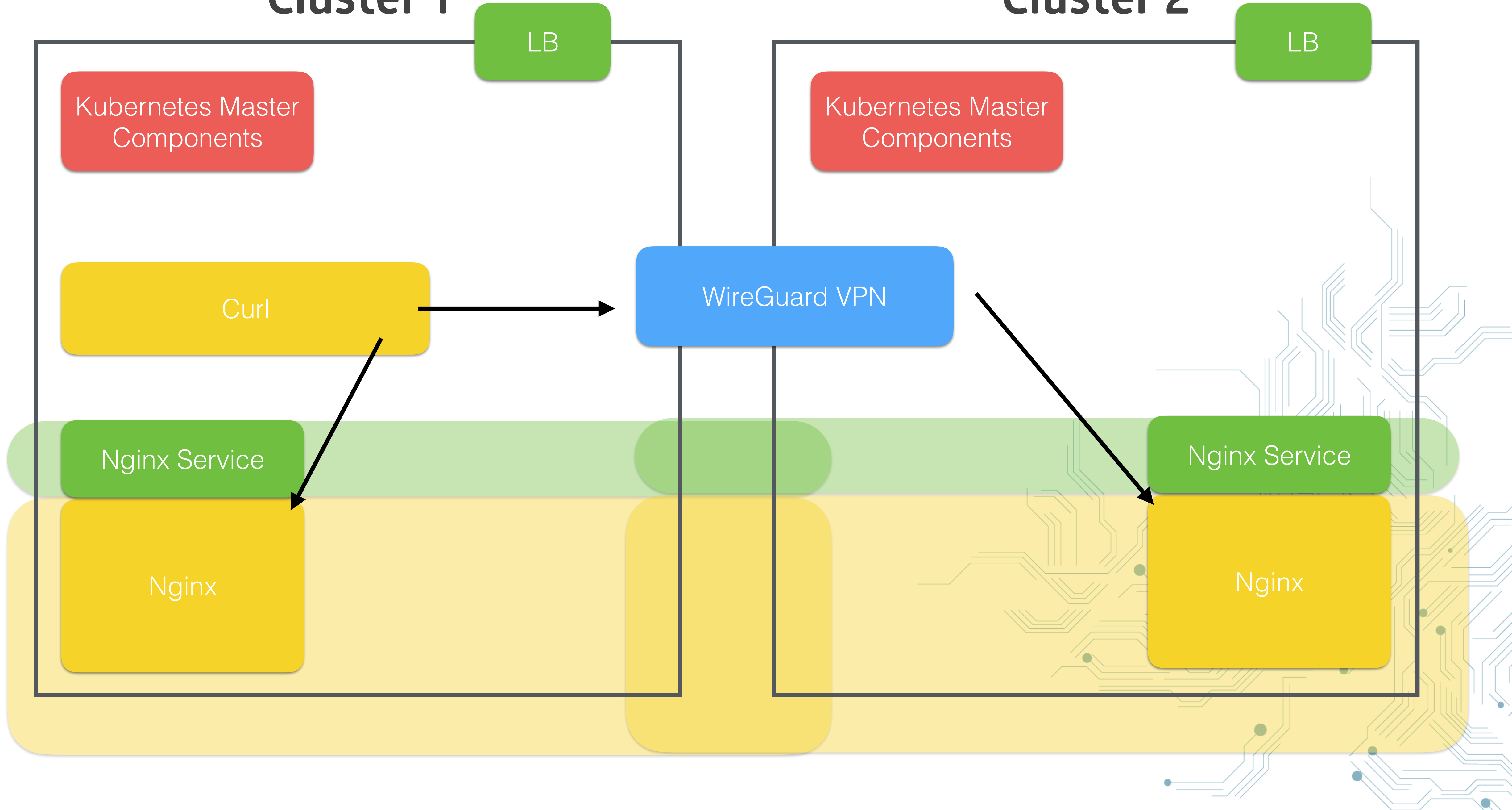
WireGuard VPN

Nginx Service

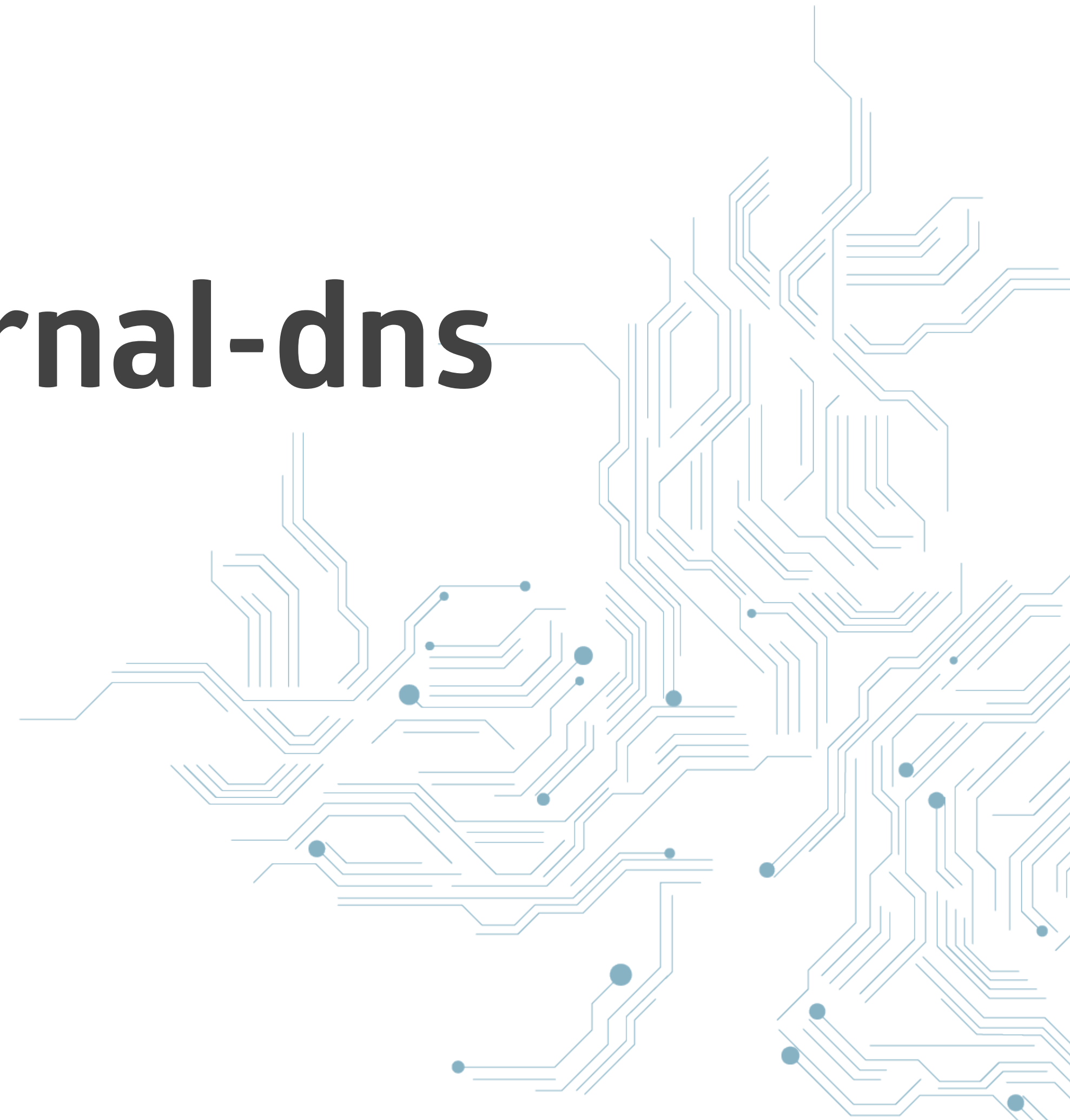
Nginx

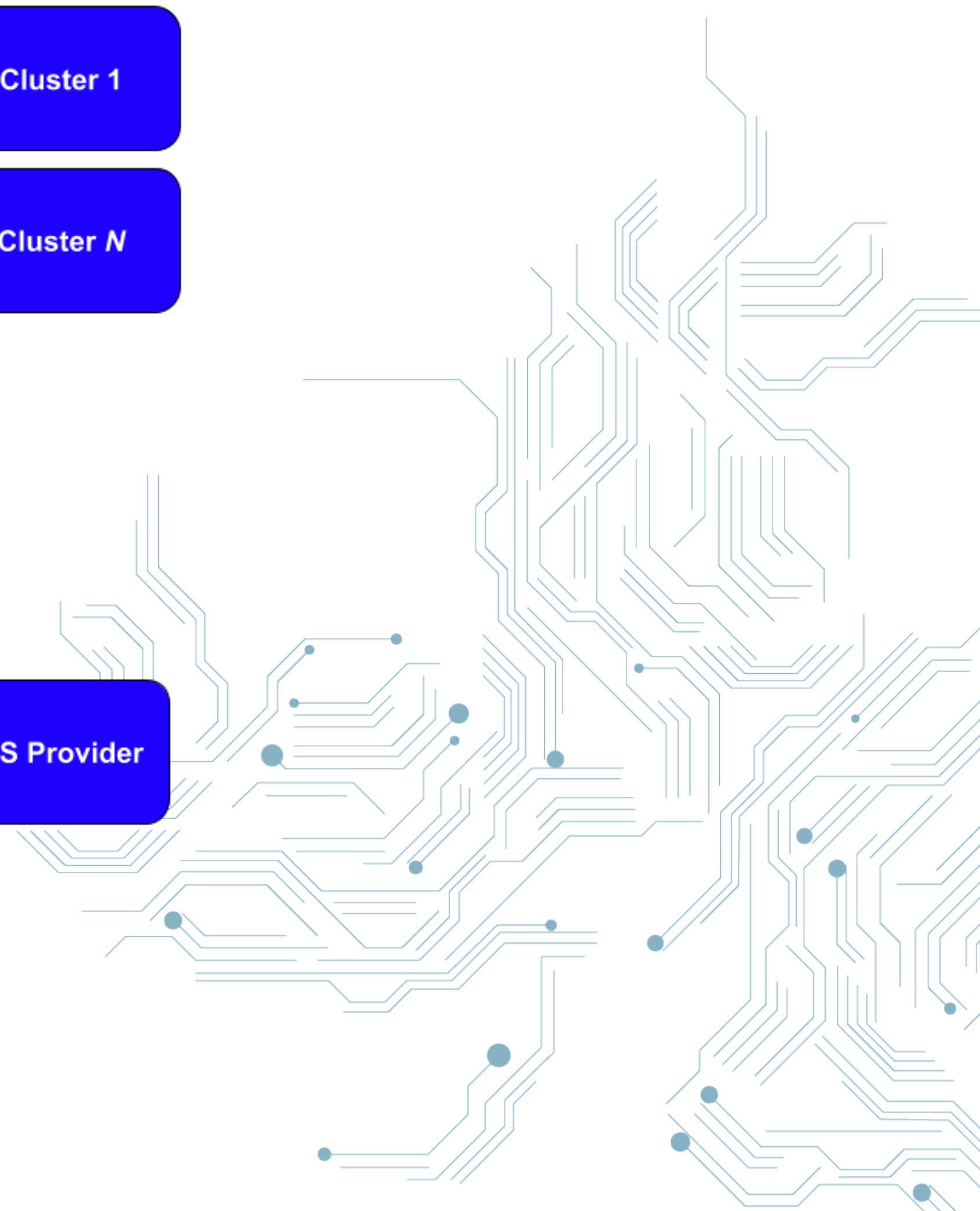
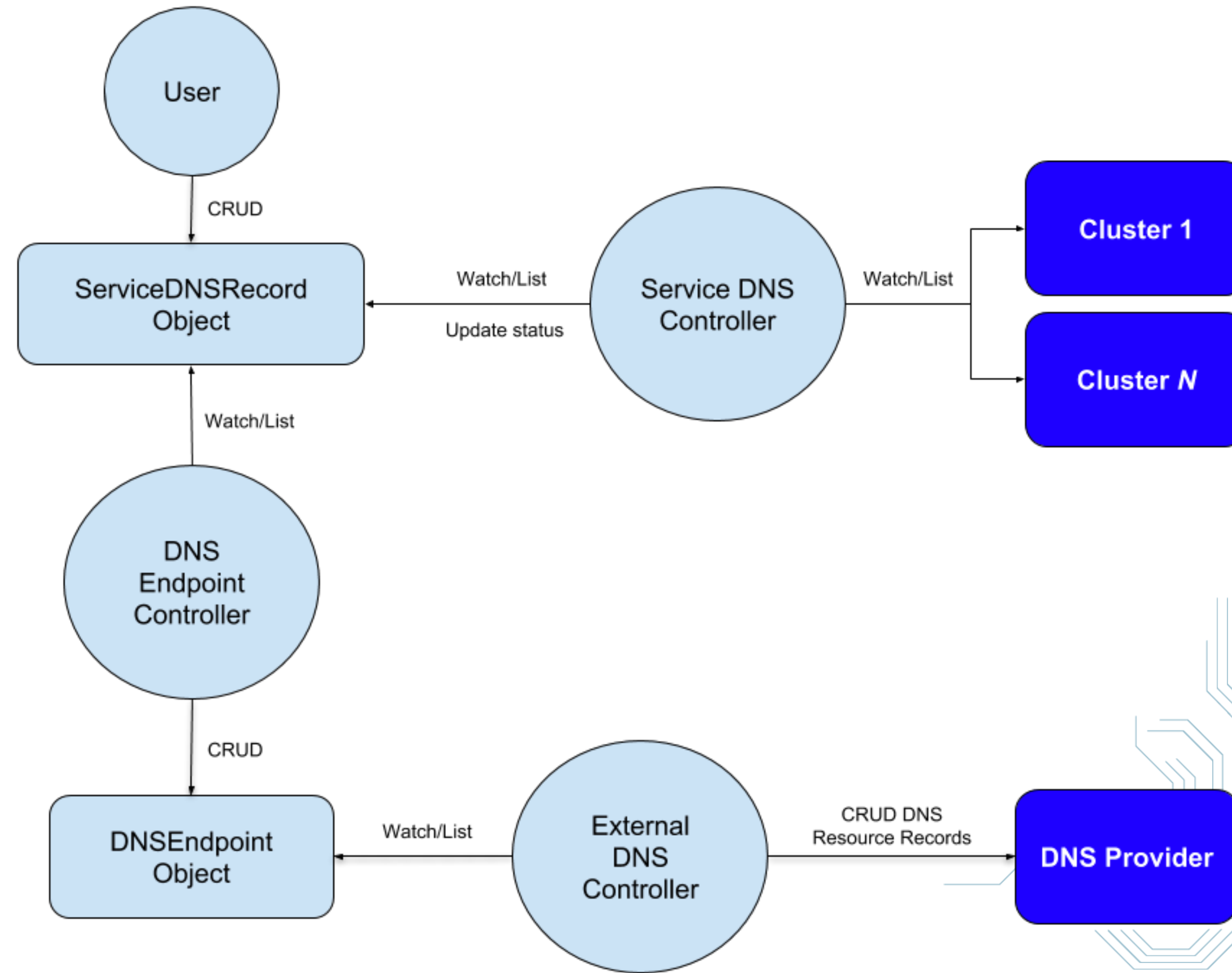
Nginx Service

Nginx



Multi cluster external-dns





service.namespace.domain.svc.example.com
service.namespace.domain.svc.region1.example.com
service.namespace.domain.svc.region2.example.com

A decorative graphic in the bottom right corner consisting of a complex, light blue circuit board pattern with various lines, dots, and geometric shapes.

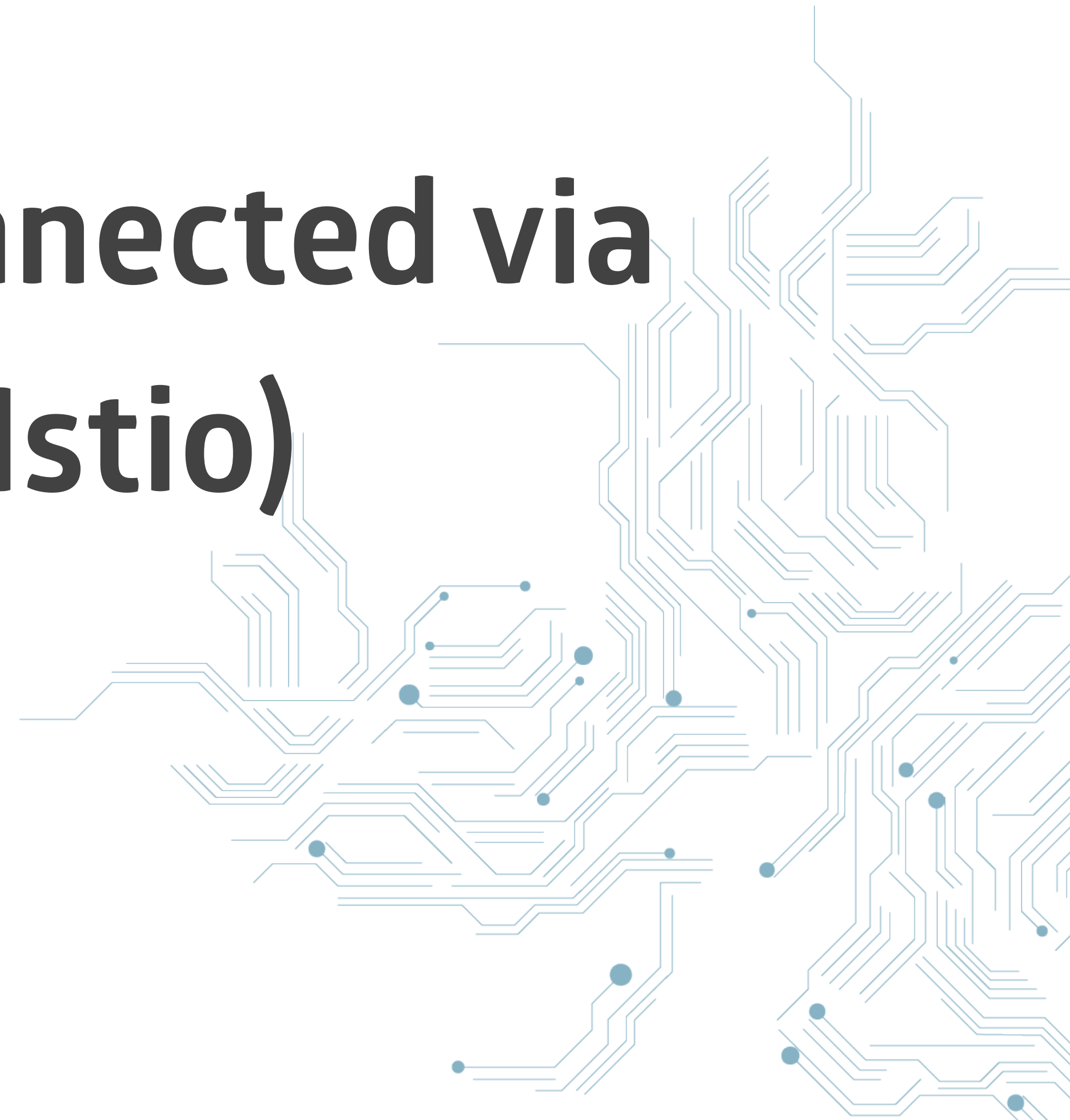
Demo



More options



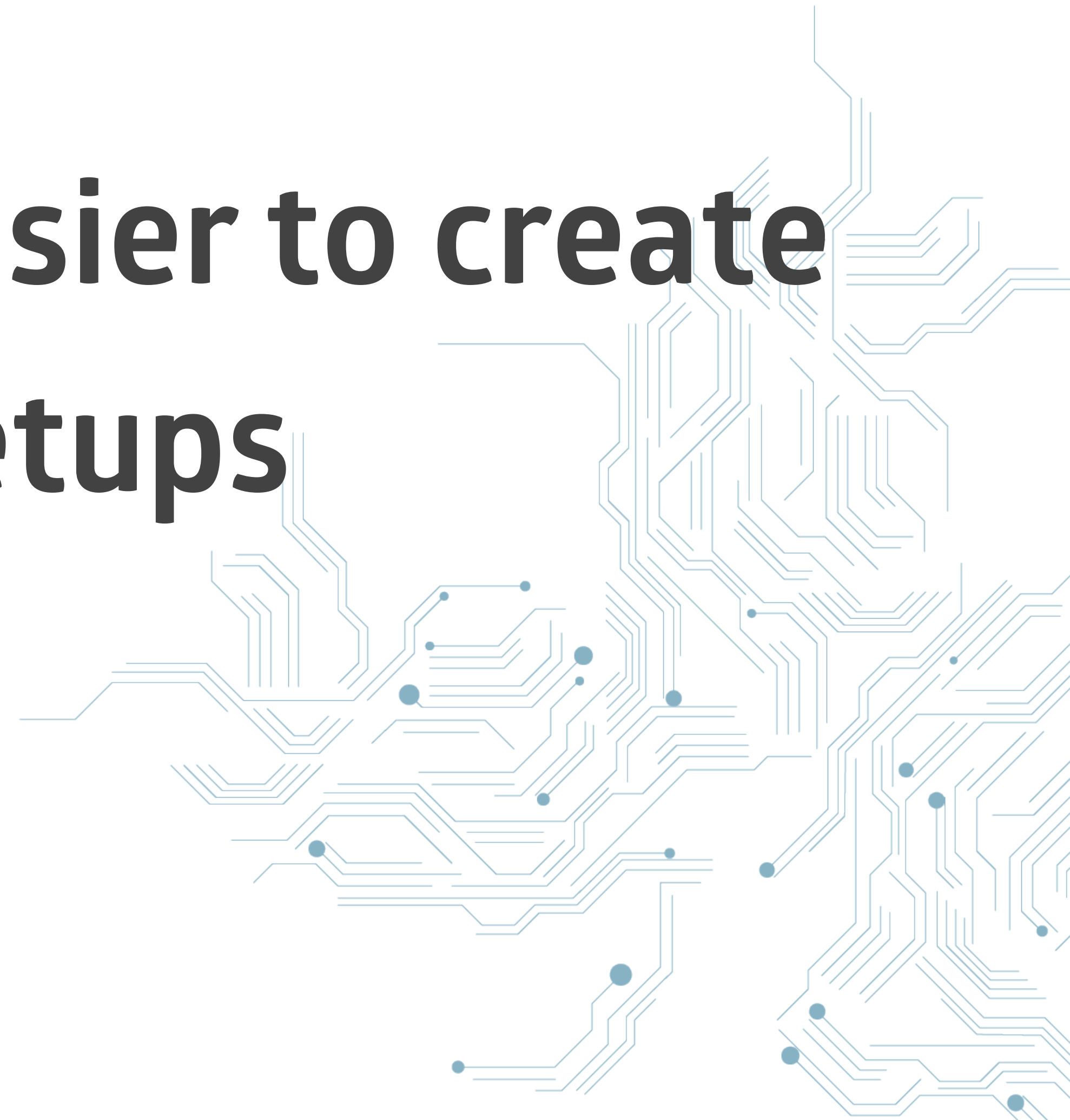
Multiple clusters connected via Service Mesh (Istio)



Conclusion



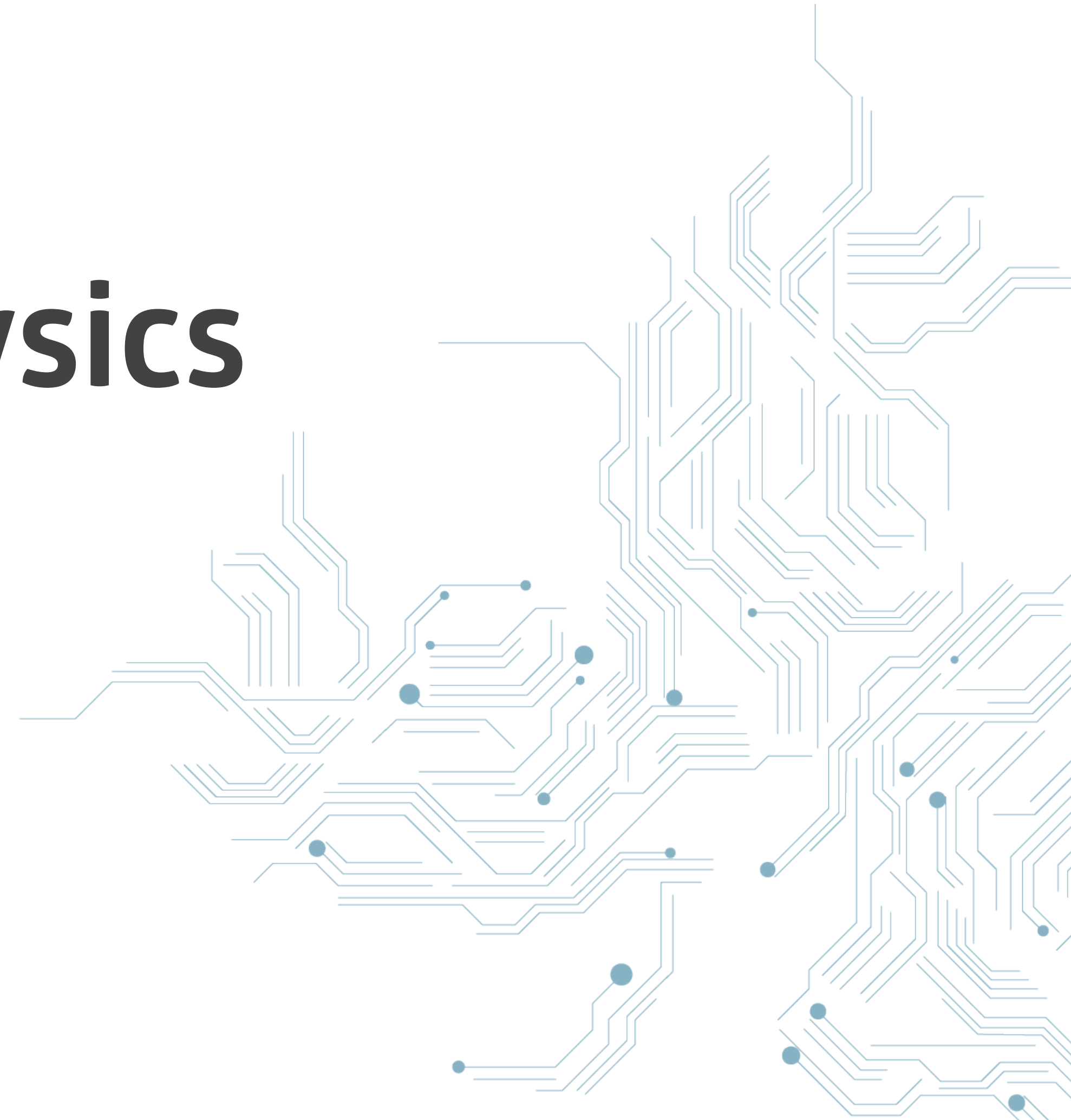
**Kubernetes makes it easier to create
multi region setups**



**There are still challenges you have to
overcome**



Often just physics



**Federation Tooling is just getting
started**





**Test it
30 days
For free**

**Visit us at our
booth**

Rate today's session

Cyberconflict: A new era of war, sabotage, and fear

See passes & pricing

David Sanger (The New York Times)
9:55am-10:10am Wednesday, March 27, 2019
Location: Ballroom
Secondary topics: Security and Privacy

Rate This Session

We're living in a new era of constant sabotage, misinformation, and fear, in which everyone is a target, and you're often the collateral damage in a growing conflict among states. From crippling infrastructure to sowing discord and doubt, cyber is now the weapon of choice for democracies, dictators, and terrorists.

David Sanger explains how the rise of cyberweapons has transformed geopolitics like nothing since the invention of the atomic bomb. Moving from the White House Situation Room to the dens of Chinese, Russian, North Korean, and Iranian hackers to the boardrooms of Silicon Valley, David reveals a world coming face-to-face with the perils of technological revolution—a conflict that the United States helped start when it began using cyberweapons against Iranian nuclear plants and North Korean missile launches. But now we find ourselves in a conflict we're uncertain how to control, as our adversaries exploit vulnerabilities in our hyperconnected nation and we struggle to figure out how to deter these complex, short-of-war attacks.

David Sanger
The New York Times

David E. Sanger is the national security correspondent for the *New York Times* as well as a national security and political contributor for CNN and a frequent guest on *CBS This Morning*, *Face the Nation*, and many PBS shows.

[Add to Your Schedule](#)
[Add Comment or Question](#)

Session page on conference website

Attending Notes Remove

Cyberconflict: A new era of war, sabotage, and fear

9:55 AM - 10:10 AM, Wed, Mar 27, 2019

Speakers

**David Sanger**
National Security Correspondent
The New York Times

Ballroom

Keynotes

David Sanger explains how the rise of cyberweapons has transformed geopolitics like nothing since the invention of the atomic bomb. From crippling infrastructure to sowing discord and doubt, cyber is now the weapon of choice for democracies, dictators, and terrorists.

 **SESSION EVALUATION**

O'Reilly Events App

**[https://github.com/bashofmann/
kubernetes-multicloud-demos](https://github.com/bashofmann/kubernetes-multicloud-demos)**





mail@bastianhofmann.de

<https://twitter.com/BastianHofmann>



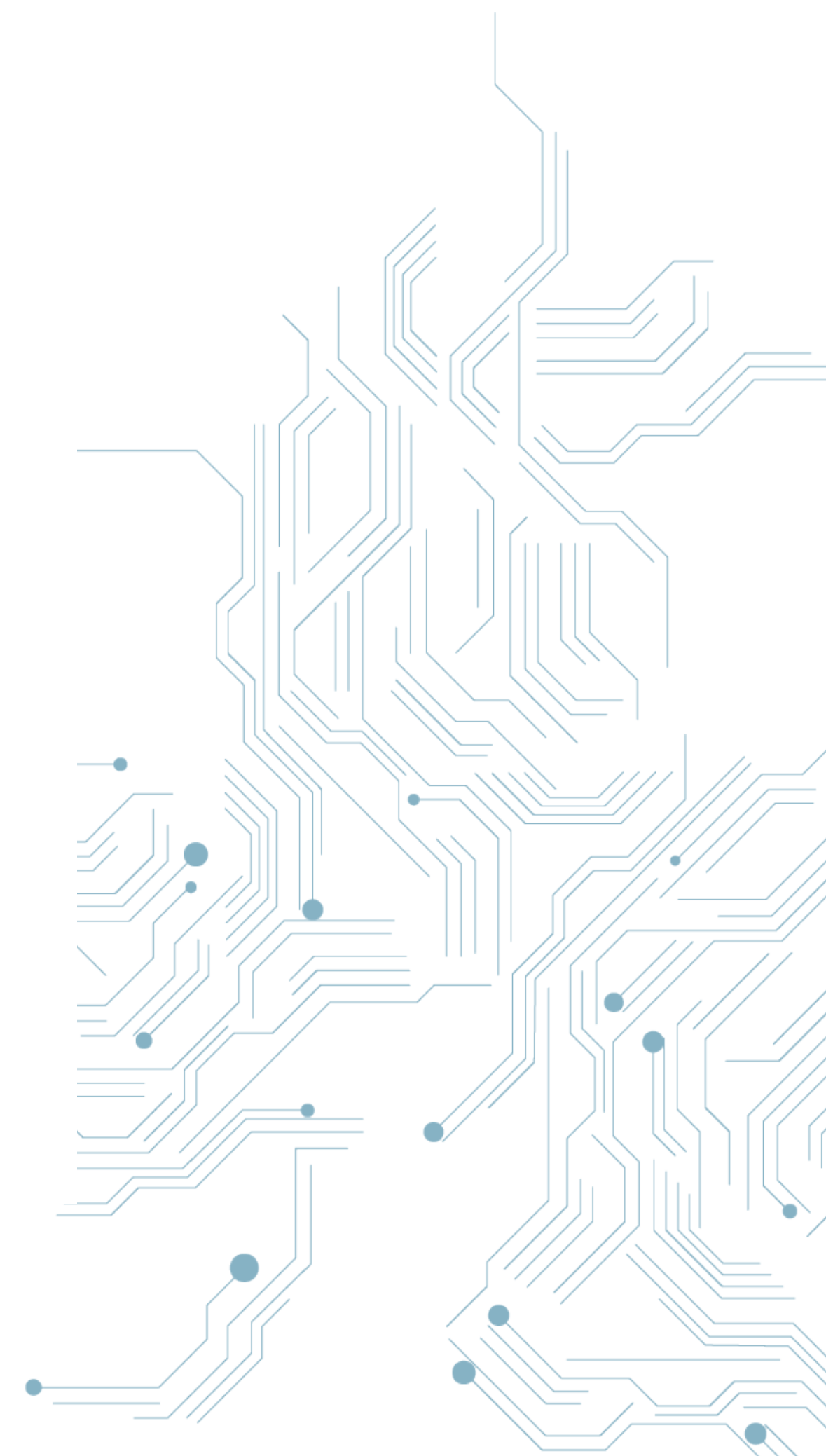
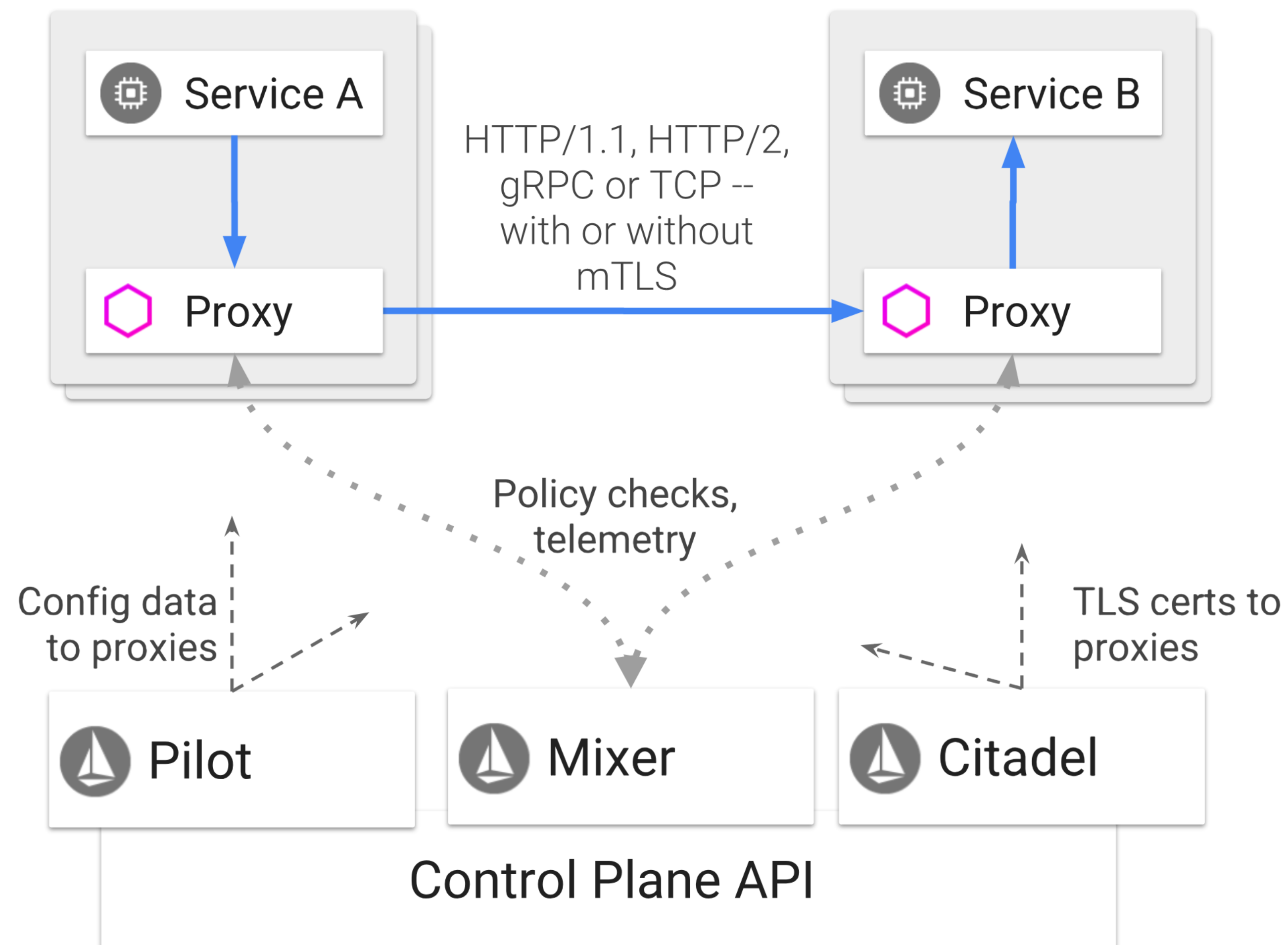
Connecting multiple clusters with a Service Mesh Gateway





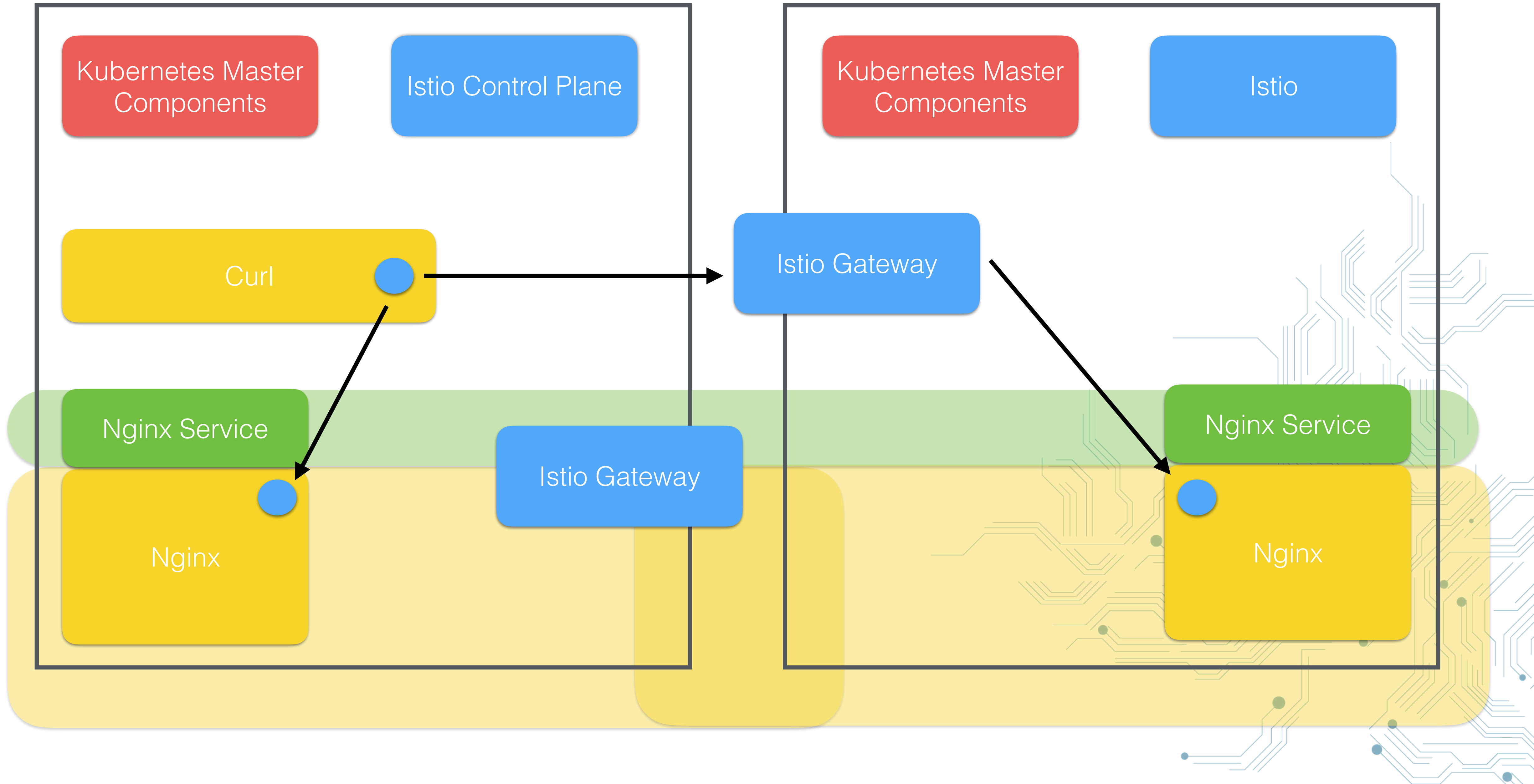
Istio





Cluster 1

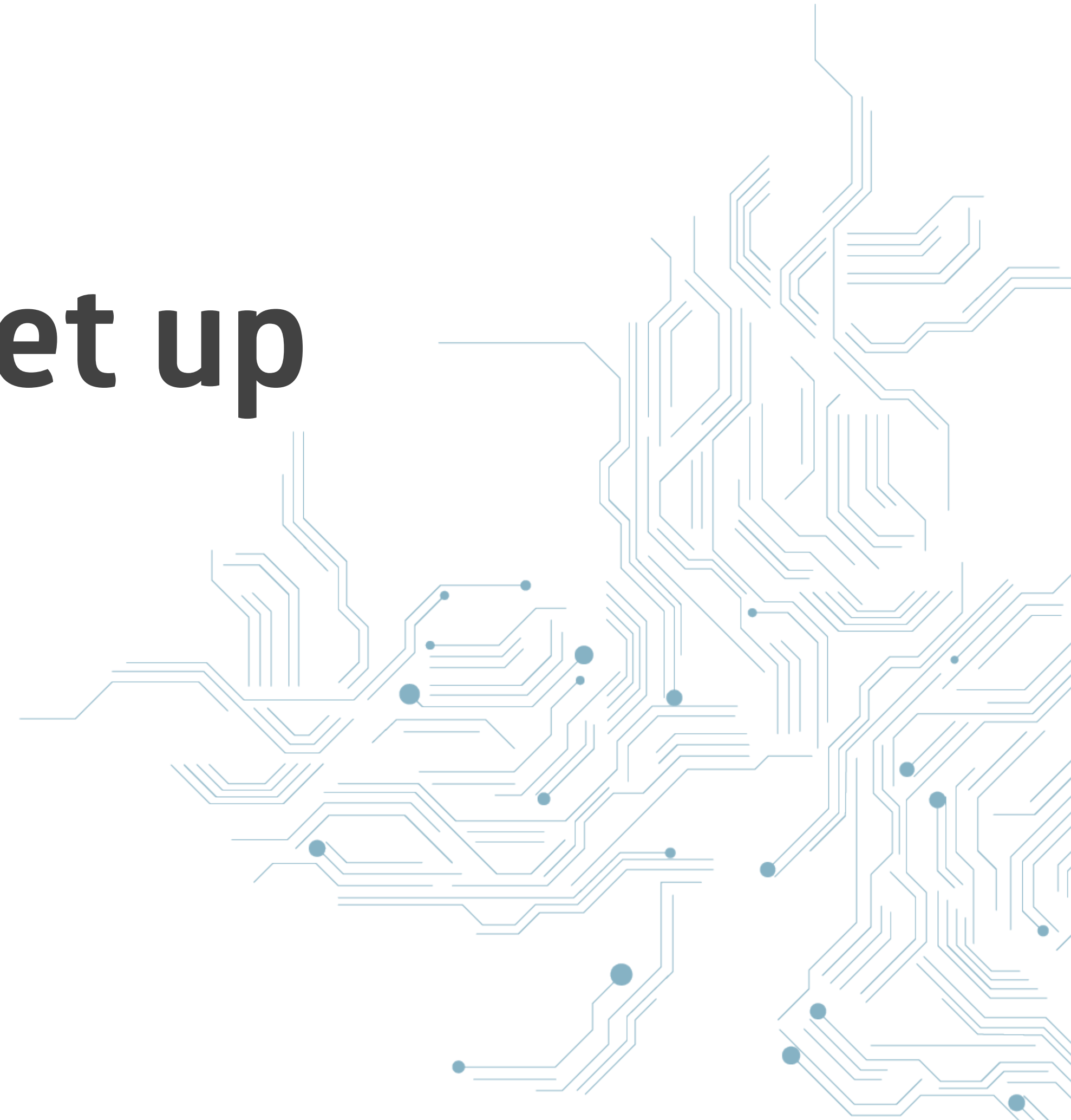
Cluster 2



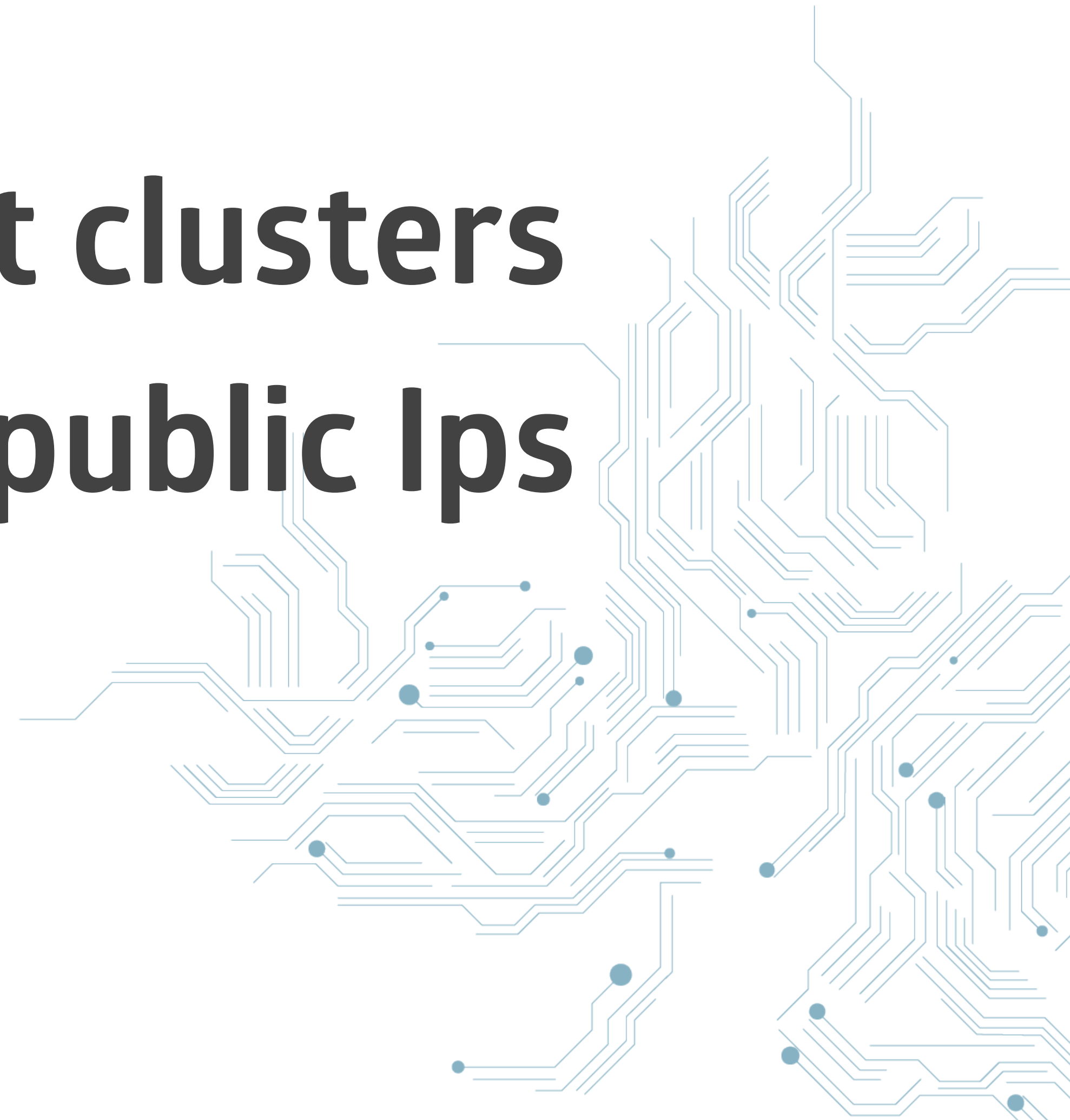
No VPN necessary



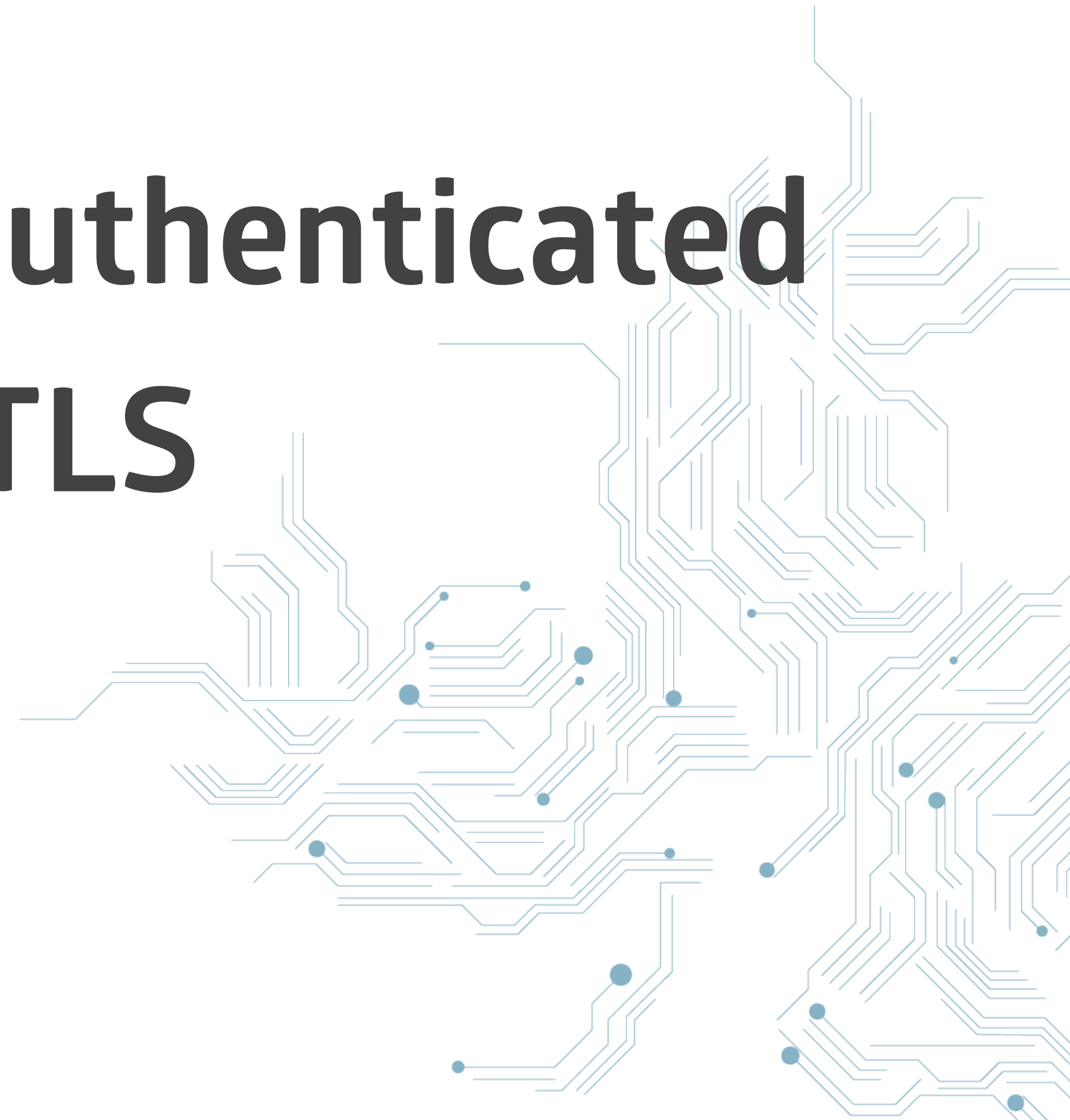
Fairly easy to set up



**Pods from different clusters
communicate over public Ips**



**Traffic encrypted and authenticated
with mutual TLS**



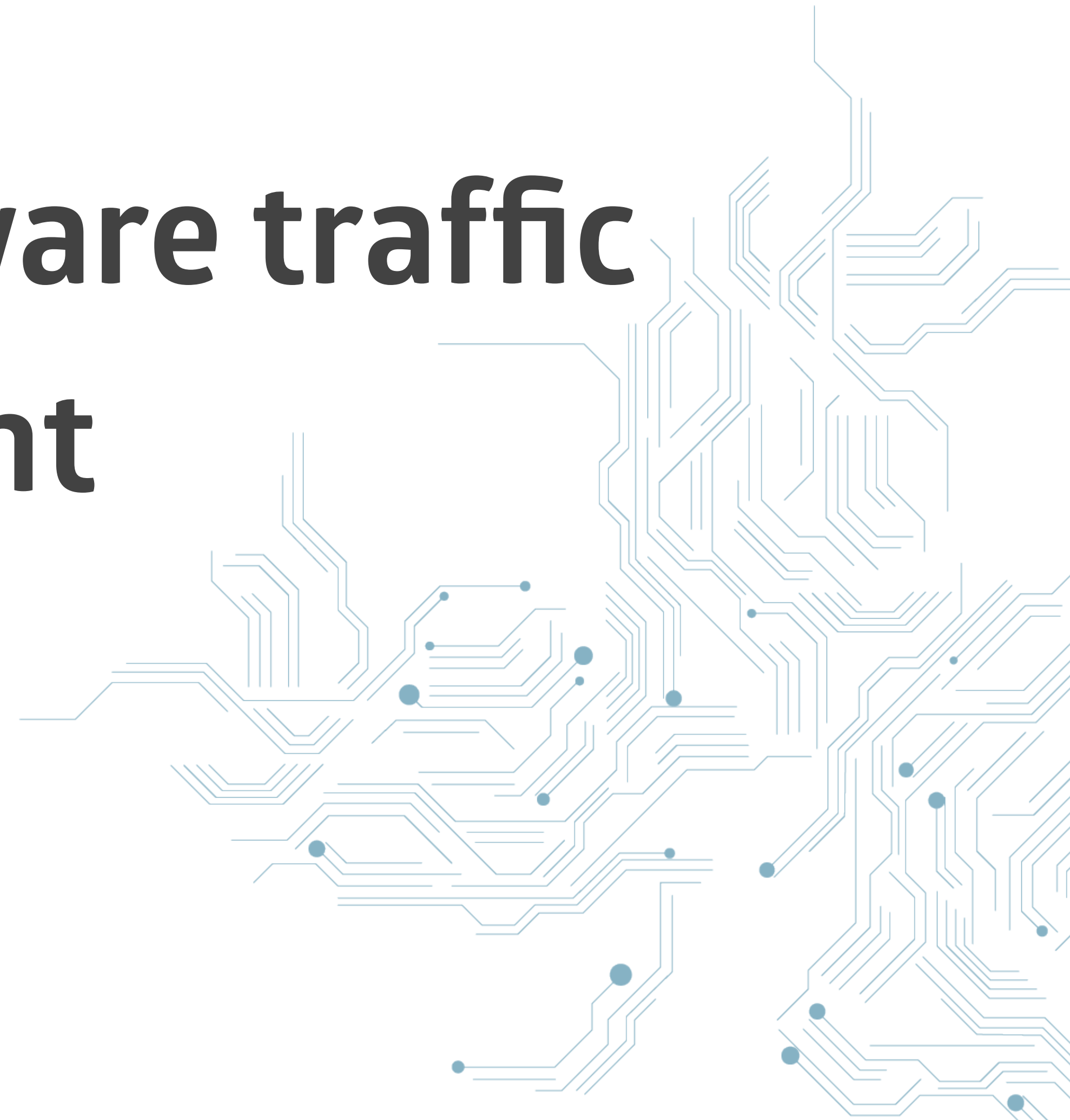
**Communication is only possible
through Istio proxies**



Easy service discovery



Flexible, location aware traffic management



Telemetry and tracing included



Demo

