Anmol Krishan Sachdeva

Sr. Hybrid Cloud Architect, Google

@greatdevaks

Cloud Native Gurugram

Infra In Cloud

June 07, 2025

CLOUD NATIVE COMMUNITY GROUPS GURUGRAM

THE FRONT DOOR TO YOUR APPS

EXPOSING SERVICES RELIABLY, SECURELY, AND AT SCALE

About Me

Anmol Krishan Sachdeva

Sr. Hybrid Cloud Architect, Google

@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

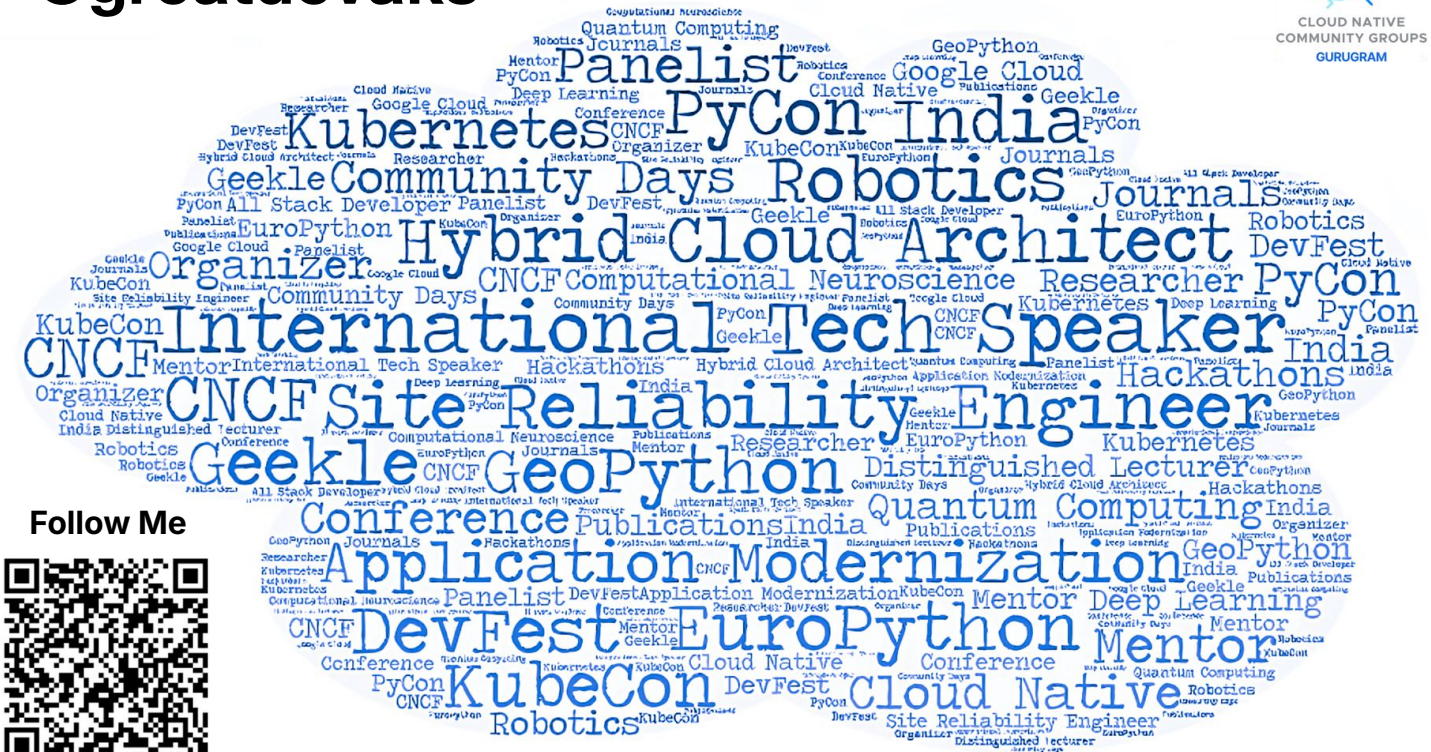
Cloud Native Gurugram

Infra In Cloud

June 07, 2025

@greatdevaks

CLOUD NATIVE COMMUNITY GROUPS GURUGRAM



Follow Me

Disclaimer

Anmol Krishan Sachdeva

Sr. Hybrid Cloud Architect, Google

@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram

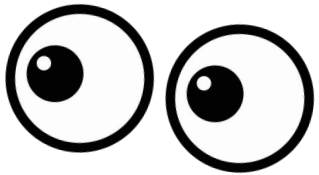
Infra In Cloud

June 07, 2025

CLOUD NATIVE COMMUNITY GROUPS GURUGRAM

THE CONTENT AND VIEWS PRESENTED DURING THE SESSION ARE AUTHOR'S OWN AND NOT OF ANY ORGANIZATIONS THEY ARE ASSOCIATED WITH.

SOME IMAGES IN THIS PRESENTATION WERE GENERATED WITH THE ASSISTANCE OF ARTIFICIAL INTELLIGENCE. SUCH ILLUSTRATIVE REPRESENTATIONS MAY NOT CONVEY ACCURATE OR FACTUALLY CORRECT INFORMATION.



Disclaimer : The content and views presented during the session are the author's personal views and are not representative of any organizations they are associated with or employed at.

Agenda

Anmol Krishan Sachdeva
Sr. Hybrid Cloud Architect, Google
@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram
Infra In Cloud
June 07, 2025

FLOW OF THE TALK

CLOUD NATIVE COMMUNITY GROUPS GURUGRAM

1. Building The Foundation: CNI & The Pod Network
2. The Core Problem: Ephemeral Pods
3. Kubernetes Services & kube-proxy Modes
4. Ingress and The Evolution
5. External Exposure For Bare-Metal
6. Modern Networking Patterns

Disclaimer : The content and views presented during the session are the author's personal views and are not representative of any organizations they are associated with or employed at.

A Primer on Kubernetes

Anmol Krishan Sachdeva

Sr. Hybrid Cloud Architect, Google

@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram

Infra In Cloud

June 07, 2025

CLOUD NATIVE COMMUNITY GROUPS GURUGRAM

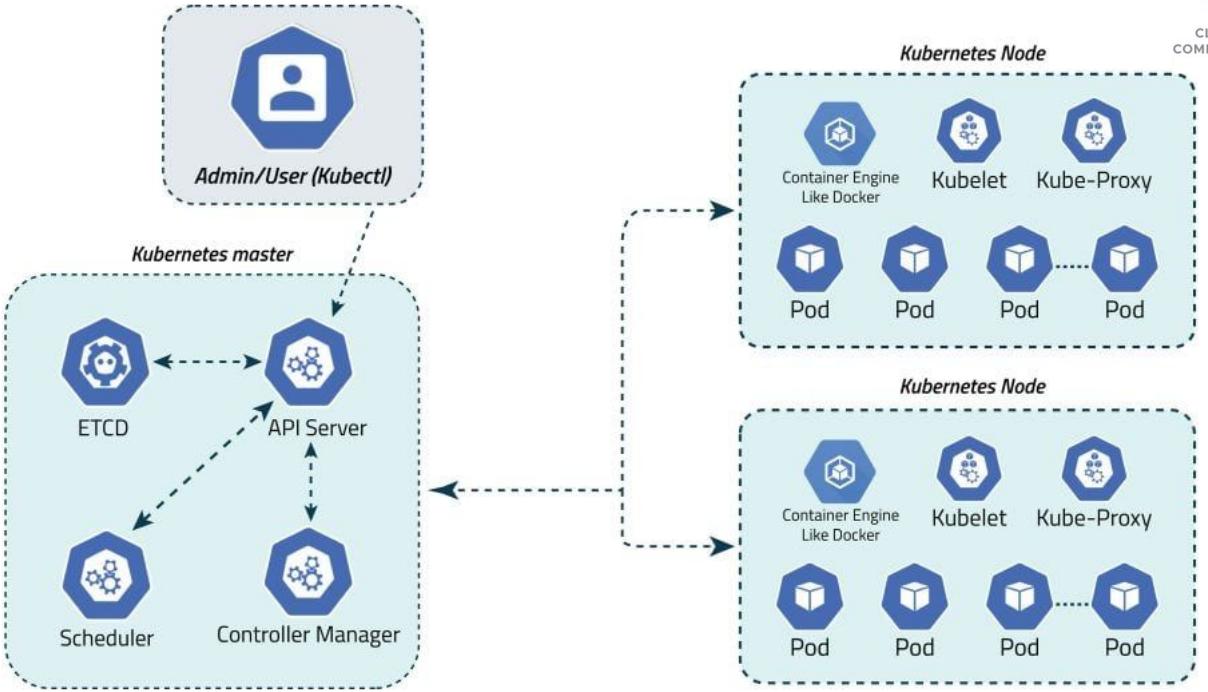


Diagram Source: CalCom

Disclaimer : The content and views presented during the session are the author's personal views and are not representative of any organizations they are associated with or employed at.

CNI & The Pod Network: SDN

Anmol Krishan Sachdeva

Sr. Hybrid Cloud Architect, Google

@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

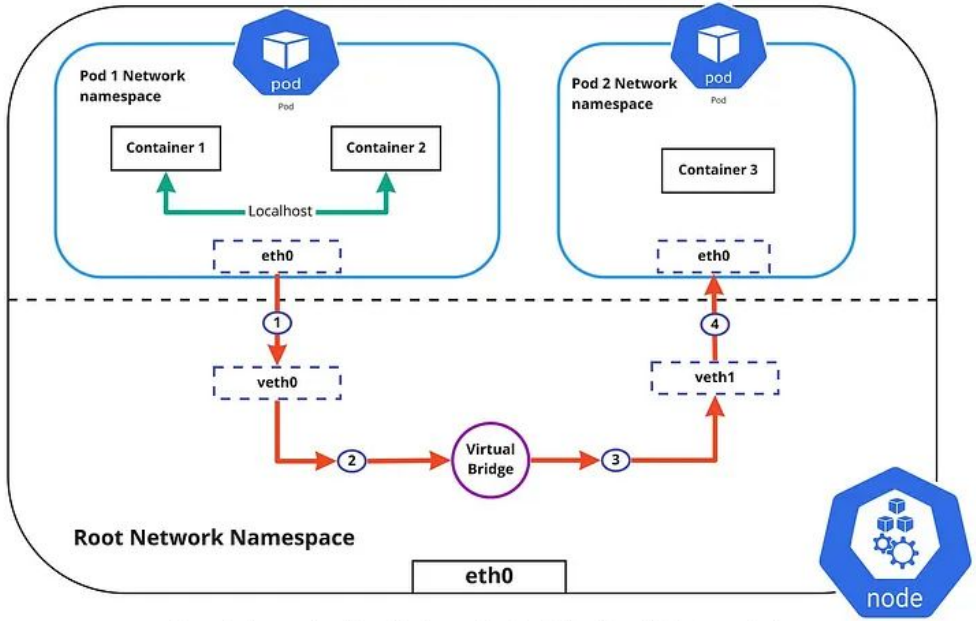
Cloud Native Gurugram

Infra In Cloud

June 07, 2025

Container Network Interface

- A framework for dynamically configuring container networking resources



Container-to-Container & Pod-to-Pod Networking

CLOUD NATIVE COMMUNITY GROUPS GURUGRAM

Diagram Source: Medium

Disclaimer : The content and views presented during the session are the author's personal views and are not representative of any organizations they are associated with or employed at.

The Hidden Gem

~~Well-Known~~ FUNDAMENTAL

Pause Container

CLOUD NATIVE COMMUNITY GROUPS GURUGRAM

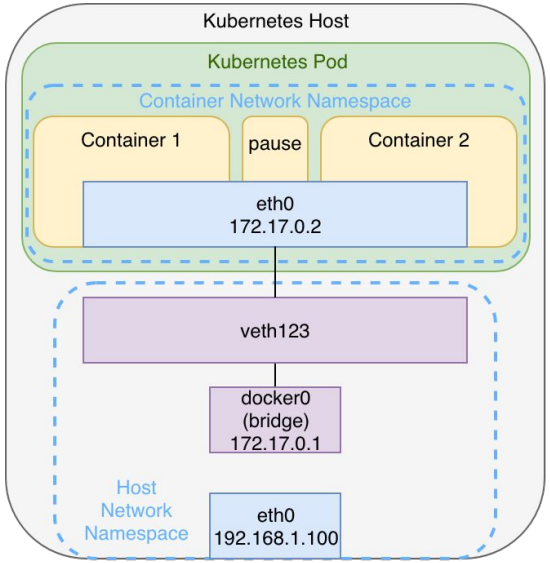


Diagram Source: Inovex

Anmol Krishan Sachdeva
Sr. Hybrid Cloud Architect, Google
@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram
Infra In Cloud
June 07, 2025

Disclaimer : The content and views presented during the session are the author's personal views and are not representative of any organizations they are associated with or employed at.

CLOUD NATIVE COMMUNITY GROUPS GURUGRAM

Why Can't We Just Connect To A Pod's IP?

- Ephemeral
- Dynamic IPs
- Autoscaled
- Unbalanced Traffic Routing

Disclaimer : The content and views presented during the session are the author's personal views and are not representative of any organizations they are associated with or employed at.

CLOUD NATIVE COMMUNITY GROUPS GURUGRAM

Enter 🚪 Kubernetes Services

ClusterIP

NodePort

LoadBalancer

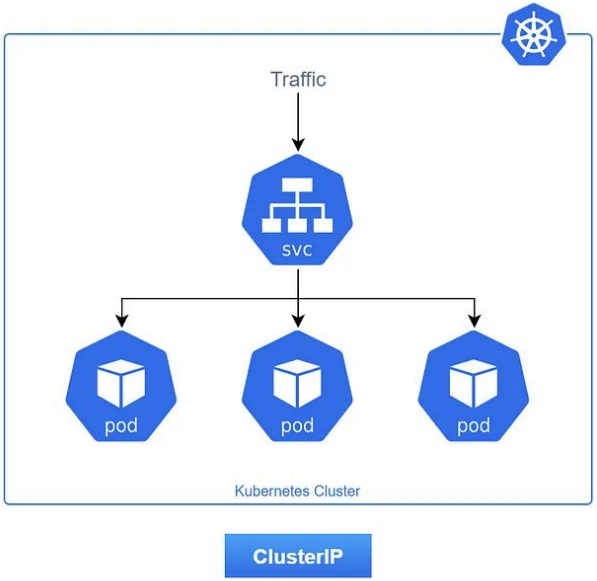
ExternalName

Headless

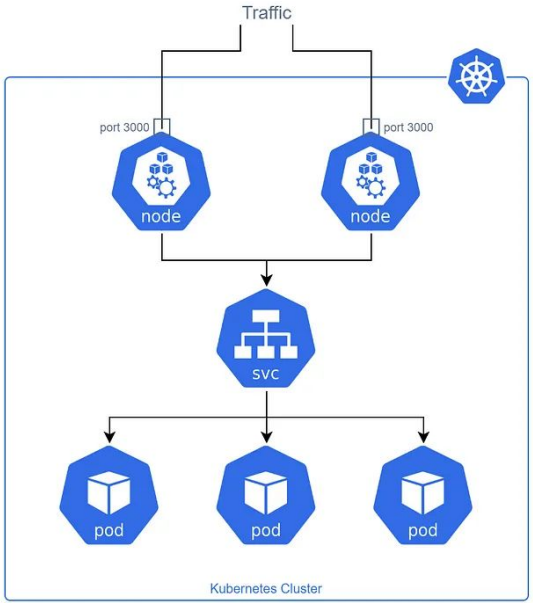
Disclaimer : The content and views presented during the session are the author's personal views and are not representative of any organizations they are associated with or employed at.

Kubernetes Services

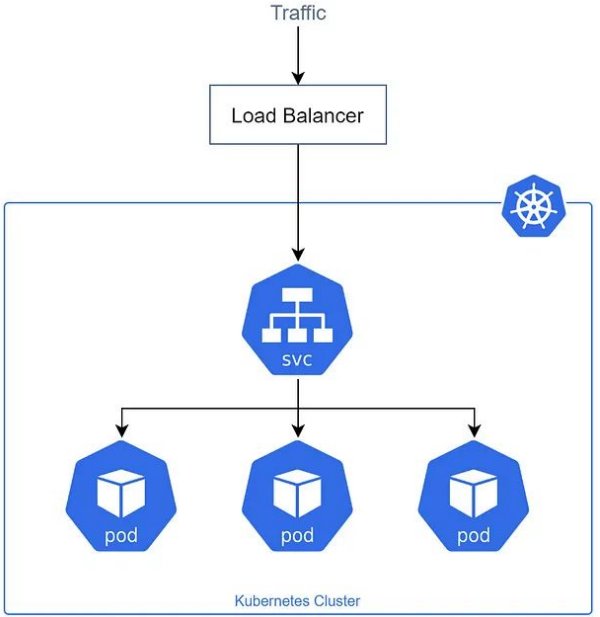
CLOUD NATIVE COMMUNITY GROUPS GURUGRAM



ClusterIP



NodePort



LoadBalancer

Diagram Source: Medium

Disclaimer : The content and views presented during the session are the author's personal views and are not representative of any organizations they are associated with or employed at.

ClusterIP Service

Noteworthy Points

CLOUD NATIVE COMMUNITY GROUPS
GURUGRAM

```yaml
apiVersion: v1
kind: Service
metadata:
  name: my-service
spec:
  type: ClusterIP
  selector:
    app.kubernetes.io/name: MyApp
  ports:
    - name: web-app
      protocol: TCP
      port: 80
      targetPort: 9376
```

Internal Networking

Default Service Type

Even Leveraged By NodePort and LoadBalancer BTS

Kubernetes DNS (CoreDNS) Creates DNS Records for ClusterIP Services

<service>.<ns>.svc.cluster.local

Every Pod Gets Domain Suffixes In /etc/resolv.conf

Know How to Deal With The Performance Tax: ndots

Anmol Krishan Sachdeva
Sr. Hybrid Cloud Architect, Google
@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram
Infra In Cloud
June 07, 2025

NodePort Service

Noteworthy Points

CLOUD NATIVE COMMUNITY GROUPS GURUGRAM

```yaml
apiVersion: v1
kind: Service
metadata:
  name: my-service
spec:
  type: NodePort
  selector:
    app.kubernetes.io/name: MyApp
  ports:
    - port: 80
      targetPort: 80
      nodePort: 30007
```

NodePort is open on every Node

Doesn't preserve True Client IP by default

externalTrafficPolicy: cluster

Port Range: 30000-32767

No two services can share the same NodePort

externalTrafficPolicy: Local

kube-proxy & healthCheckNodePort

Anmol Krishan Sachdeva
Sr. Hybrid Cloud Architect, Google
@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram
Infra In Cloud
June 07, 2025

LoadBalancer Service

Noteworthy Points

CLOUD NATIVE COMMUNITY GROUPS
GURUGRAM

```yaml
apiVersion: v1
kind: Service
metadata:
  name: my-service
spec:
  type: LoadBalancer
  selector:
    app.kubernetes.io/name: MyApp
  ports:
    - protocol: TCP
      port: 80
      targetPort: 9376
```

Requires Cloud Controller Manager or Bare-Metal Service Controller

Almost Always Layer 4 Load Balancer

Supports External and Internal Exposure

Built on top of NodePort

Same limitation BTS… Port Range: 30000-32767

externalTrafficPolicy configuration holds true

Anmol Krishan Sachdeva
Sr. Hybrid Cloud Architect, Google
@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram
Infra In Cloud
June 07, 2025

kube-proxy

Anmol Krishan Sachdeva
Sr. Hybrid Cloud Architect, Google
@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram
Infra In Cloud
June 07, 2025

kube-proxy

CLOUD NATIVE COMMUNITY GROUPS GURUGRAM

- Runs as a DaemonSet - provisioning can be skipped in latest implementations leveraging eBPF
- Helps implement a Virtual IP mechanism for Services; exception is ExternalName type
- Helps in redirecting (DNAT) traffic from the ClusterIP:Port to PodEndpoint:Port
- Might influence kernel level rules and netfilter configurations
- Can be configured in the below-shown modes:

userspace	iptables	ipvs	nftables	kernelspace
Older method; kube-proxy used to load balance	Netfilter rule chains and evaluations Default but slow	HashMap-based implementation for O(1) performance Actual LB	Maps and Sets leveraged for O(1) performance	Packet forwarding rules for Windows kernel

- Conntrack helps in connection and connection state tracking; session affinity is possible

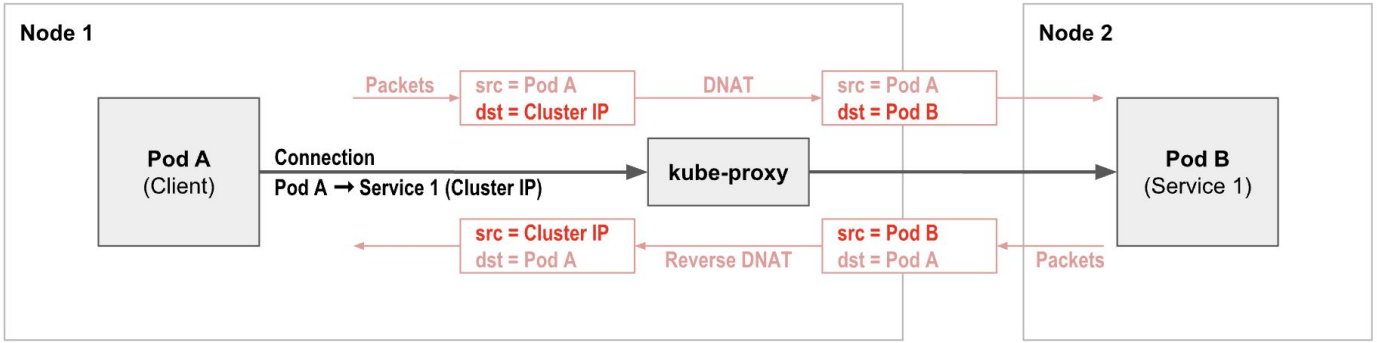
Disclaimer : The content and views presented during the session are the author's personal views and are not representative of any organizations they are associated with or employed at.

kube-proxy

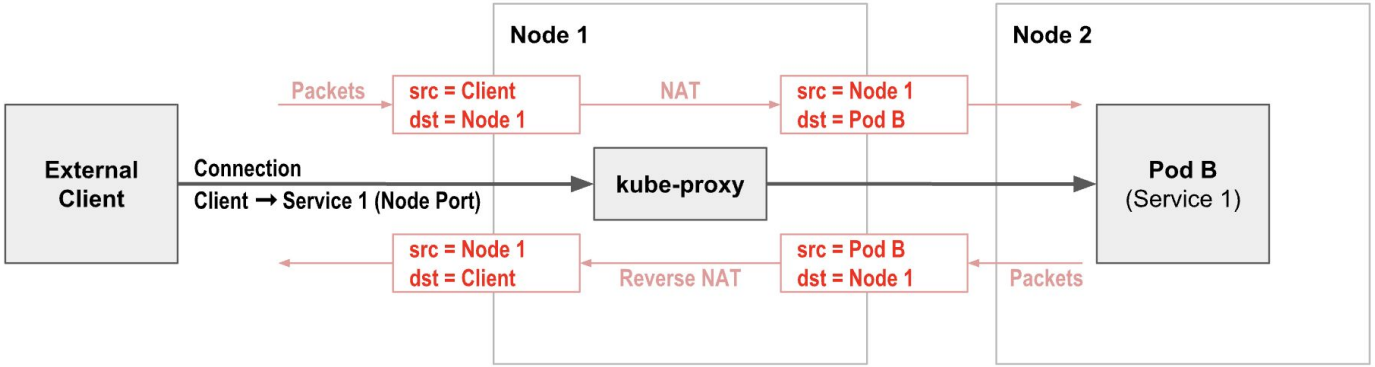
Anmol Krishan Sachdeva
Sr. Hybrid Cloud Architect, Google
@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram
Infra In Cloud
June 07, 2025



ClientIP and kube-proxy

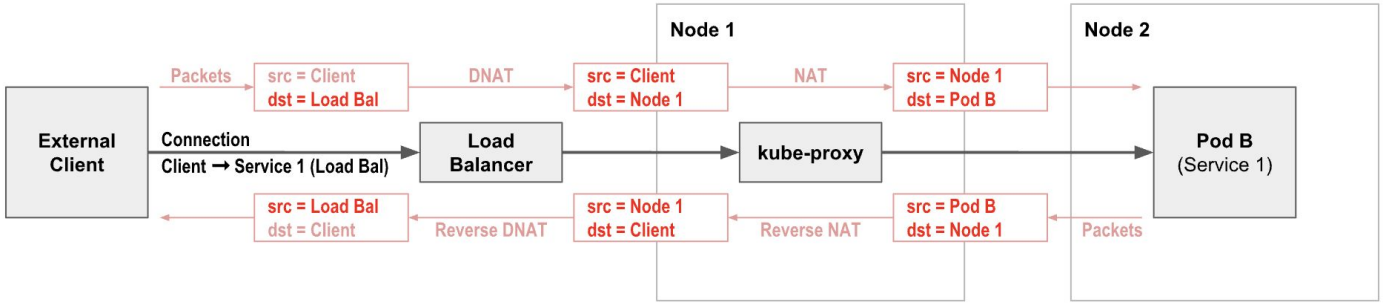


NodePort and kube-proxy

Diagram Source: Tigera

Disclaimer : The content and views presented during the session are the author's personal views and are not representative of any organizations they are associated with or employed at.

kube-proxy



LoadBalancer and kube-proxy

Anmol Krishan Sachdeva

Sr. Hybrid Cloud Architect, Google

@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram

Infra In Cloud

June 07, 2025

Diagram Source: Tigera

CLOUD NATIVE COMMUNITY GROUPS GURUGRAM

Can Layer 7 Load Balancing Be Also Done?

- Ingress Controller
- Multiplexing

  One IP, Many Services
- TLS/mTLS Termination
- Namespaced Resource
- Catch-All Default Backend

Ingress API and Layer 7 Load Balancing

Anmol Krishan Sachdeva
Sr. Hybrid Cloud Architect, Google
@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram
Infra In Cloud
June 07, 2025

```yaml
apiVersion: networking.k8s.io/v1
kind: Ingress
metadata:
  name: my-multi-service-ingress
  annotations:
    nginx.ingress.kubernetes.io/force-ssl-redirect: "true"
spec:
  ingressClassName: nginx
  tls:
  - hosts:
    - app1.example.com
    - app2.example.com
    secretName: my-tls-secret
  rules:
  - host: "app1.example.com"
    http:
      paths:
      - path: /
        pathType: Prefix
        backend:
          service:
            name: app1-service
            port:
              number: 80
  - host: "app2.example.com"
    http:
      paths:
      - path: /
        pathType: Prefix
        backend:
          service:
            name: app2-service
            port:
              number: 8080
```

Noteworthy Points

Backend Services and TLS Secret should be in the same Namespace as Ingress

default Namespace in this example

Annotation-based Advanced Features

Annotations are Vendor specific

HTTP/HTTPS only supported

Other protocols like gRPC not supported; L4 ones could work with customizations

Advanced Traffic Splitting, Canary / Blue/Green, and Mirroring Features not natively and commonly available

Disclaimer : The content and views presented during the session are the author's personal views and are not representative of any organizations they are associated with or employed at.

Gateway API and Layer 7 Load Balancing

Anmol Krishan Sachdeva

Sr. Hybrid Cloud Architect, Google

@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram

Infra In Cloud

June 07, 2025

```yaml
apiVersion: gateway.networking.k8s.io/v1
kind: Gateway
metadata:
  name: my-multi-service-gateway
spec:
  gatewayClassName: nginx
  listeners:
    - name: https
      protocol: HTTPS
      port: 443
      hostname: "*.example.com"
      tls:
        mode: Terminate
        certificateRefs:
          - kind: Secret
            name: my-tls-secret
---
apiVersion: gateway.networking.k8s.io/v1
kind: HTTPRoute
metadata:
  name: app1-route
spec:
  parentRefs:
    - name: my-multi-service-gateway
      sectionName: https
  hostnames:
    - "app1.example.com"
  rules:
    - matches:
        - path:
            type: PathPrefix
            value: /
      backendRefs:
        - name: app1-service
          port: 80

---
apiVersion: gateway.networking.k8s.io/v1
kind: HTTPRoute
metadata:
  name: app2-route
spec:
  parentRefs:
    - name: my-multi-service-gateway
      sectionName: https
  hostnames:
    - "app2.example.com"
  rules:
    - matches:
        - path:
            type: PathPrefix
            value: /
      backendRefs:
        - name: app2-service
          port: 8080
```

Noteworthy Points

No Annotations Hassle

Cross-Namespace Referencing Possible

Role-Oriented and Composable

Route Attachment can be controlled per Namespace

Advanced Traffic Splitting, Canary / Blue/Green, and Mirroring Features natively and commonly available

Disclaimer : The content and views presented during the session are the author's personal views and are not representative of any organizations they are associated with or employed at.

Ingress API vs. Gateway API

CLOUD NATIVE COMMUNITY GROUPS

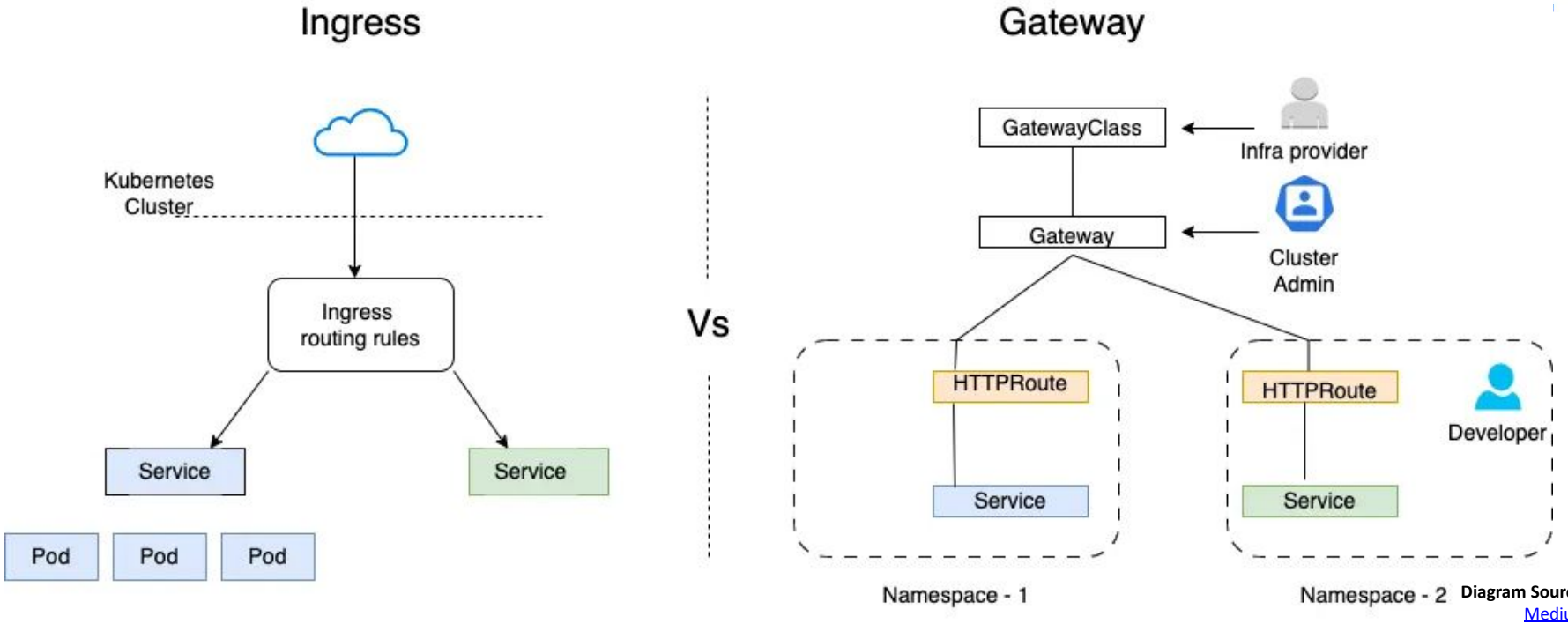


Diagram Source: Medium

Disclaimer : The content and views presented during the session are the author's personal views and are not representative of any organizations they are associated with or employed at.

Battlecard: Ingress API vs. Gateway API

Feature / Aspect	Ingress API	Gateway API
Operational Mode	Application Centric and Mostly Monolithic	Role-Oriented and Composable
Portability	Low and Vendor-Specific Annotations and Extensions	Leverages Standard Portable Specifications
Protocol Support	HTTP/HTTPS	HTTP/S, TCP, UDP, TLS, gRPC are native
Routing Rules and Traffic Splitting	Host/Path-based	Header-based additionally supported along with Advanced Traffic Splitting and Mirroring
Infrastructure Sharing and Multi-Tenancy	No clean and safe way for sharing a Load Balancer across Namespaces	Gateway operator can control the Namespaces which could participate in attaching routes
Extension	Non-Standard Annotation-based	Allows Custom Filters and Policy Attachments

Anmol Krishan Sachdeva
Sr. Hybrid Cloud Architect, Google
@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram
Infra In Cloud
June 07, 2025

Bare-Metal Load Balancing

CLOUD NATIVE COMMUNITY GROUPS GURUGRAM

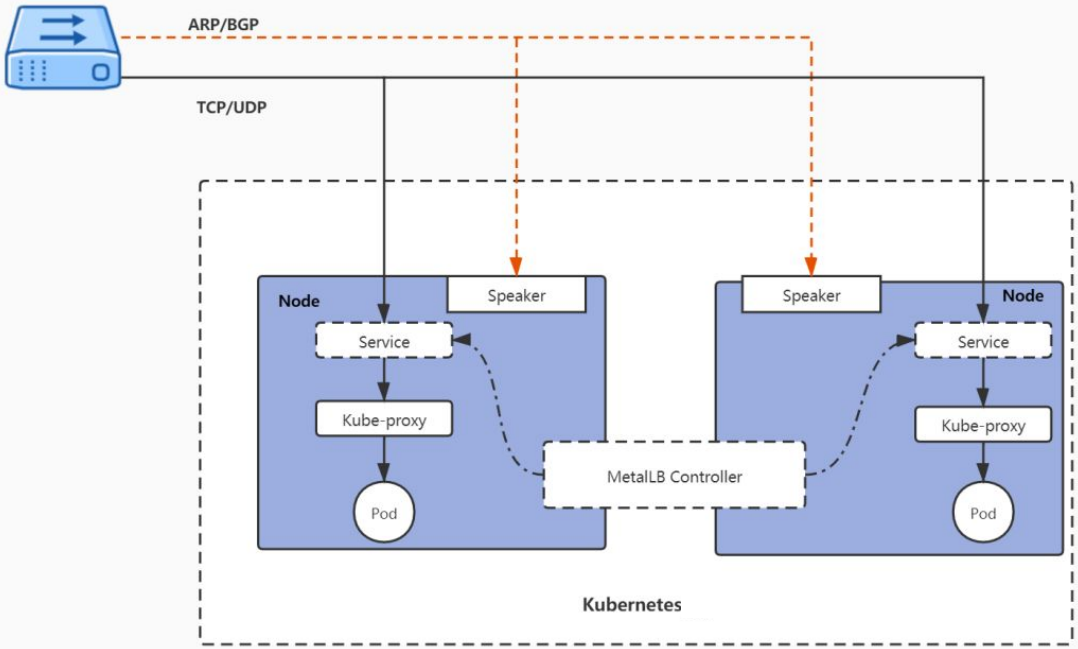


Diagram Source: ZenTao

Bare-Metal Load Balancing with MetalLB

Anmol Krishan Sachdeva
Sr. Hybrid Cloud Architect, Google
@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram
Infra In Cloud
June 07, 2025

Feature / Aspect	Layer 2 Mode (ARP)	Layer 3 Mode (BGP)
How it Works	A single "leader" node broadcasts an ARP message on the local network, claiming a Service's IP for its own MAC Address	Multiple nodes peer with a network router using BGP, each advertising themselves as a valid, equal-cost path to the Service IP
Traffic Path	All traffic funneled through a single leader node for a Service	Equal spread of traffic across multiple nodes
Fitment	Home lab or small scale setups	Production-grade setups needing HA
Drawback(s)	Single-Node Bottleneck and Memory/Page Size Limits	Needs BGP-capable Router and Network Configuration outside Kubernetes

And ofcourse, alternatively, routing from F5/Cisco to NodePort also works - L7 LBs work too

Conclusion

Key Takeaways

- eBPF is the New Data Plane; kube-proxy is getting bypassed
- BGP with eBPF could give immense scale to On-Premises/Bare-Metal setups
- Gateway API is the near-term future for North-South traffic management
- Gateway API is well-suited for advanced traffic routing and multi-cluster routing
- NetworkPolicies are evolving and CNI providers are adding many L7 capabilities also to the NetworkPolicies
- Use IPsec or Wireguard for Multi-Cluster Pod-to-Pod communication
- Ingress and Mesh capabilities are getting unified by GAMMA Gateway API project
- Egress Gateways are non-negotiable for security
- Tools like Hubble can help in network observability for Cilium powered implementations

CLOUD NATIVE COMMUNITY GROUPS GURUGRAM

Anmol Krishan Sachdeva
Sr. Hybrid Cloud Architect, Google
@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram
Infra In Cloud
June 07, 2025

Diagram Source:
Medium

Follow Me

Anmol Krishan Sachdeva

Sr. Hybrid Cloud Architect, Google

@greatdevaks

The Front Door To Your Apps: Exposing Services Reliably, Securely, And At Scale

Cloud Native Gurugram

Infra In Cloud

June 07, 2025

#OpenSource
#CNCF #Kubernetes
#Networking
@greatdevaks

Thanks Everyone !

CLOUD NATIVE COMMUNITY GROUPS GURUGRAM