



KubeCon



CloudNativeCon

Europe 2026

#KubeCon #CloudNativeCon

From NLB Sprawl to Mesh Efficiency

How Skyscanner Handles 60M Requests per Minute

Skyscanner, Solo.io





KubeCon



CloudNativeCon

Europe 2026



John Clark

Skyscanner



Steven Thwaites

Solo.io



Skyscanner Platform



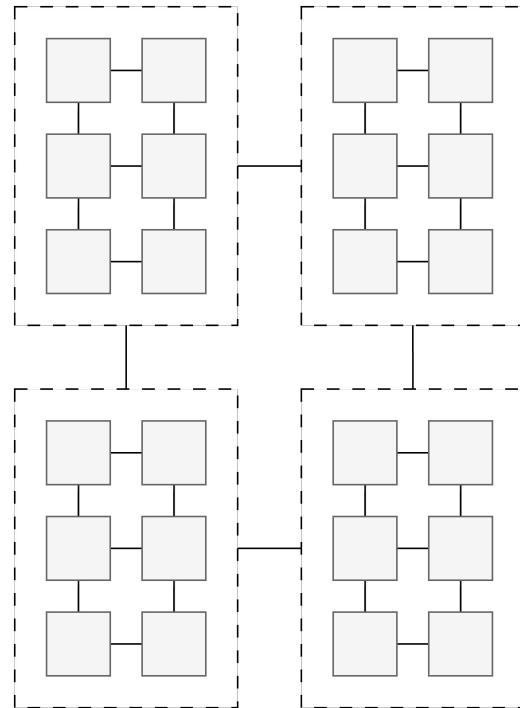
KubeCon



CloudNativeCon

Europe 2026

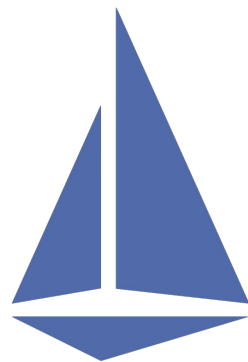
- 4 production regions
- 6 EKS clusters per region, 2 per AZ
- Cells architecture, flat networks
- EKS fully on Spot with Karpenter
- Connected worldwide with Istio
- 600+ services, 35,000 pods, 1,000 nodes
- 60+ million requests per minute



Project Globe

Moving to a new architecture (2019)

- Modernise - minimise single points of failure
- New resilience model - separate and replicated clusters per AZ
- How to:
 - Securely route internally and across clusters
 - Support automatic failover and locality routing
 - Abstract all away from service logic/code
- **Istio 1.4** - mTLS, telemetry, DNS abstraction, etc.



Service Mesh with Istio

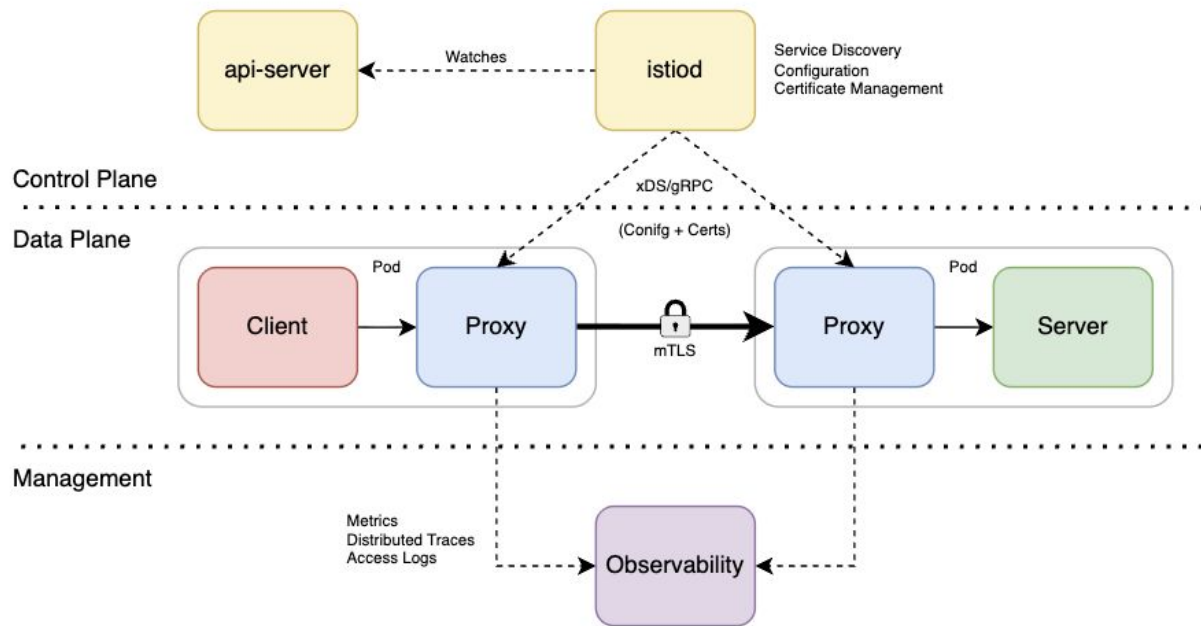


KubeCon



CloudNativeCon

Europe 2026



- Infra layer for **service-service communication**
- Transparent to the application, **no code changes**
- Connect, Secure, Control, Observe

The Core Capability

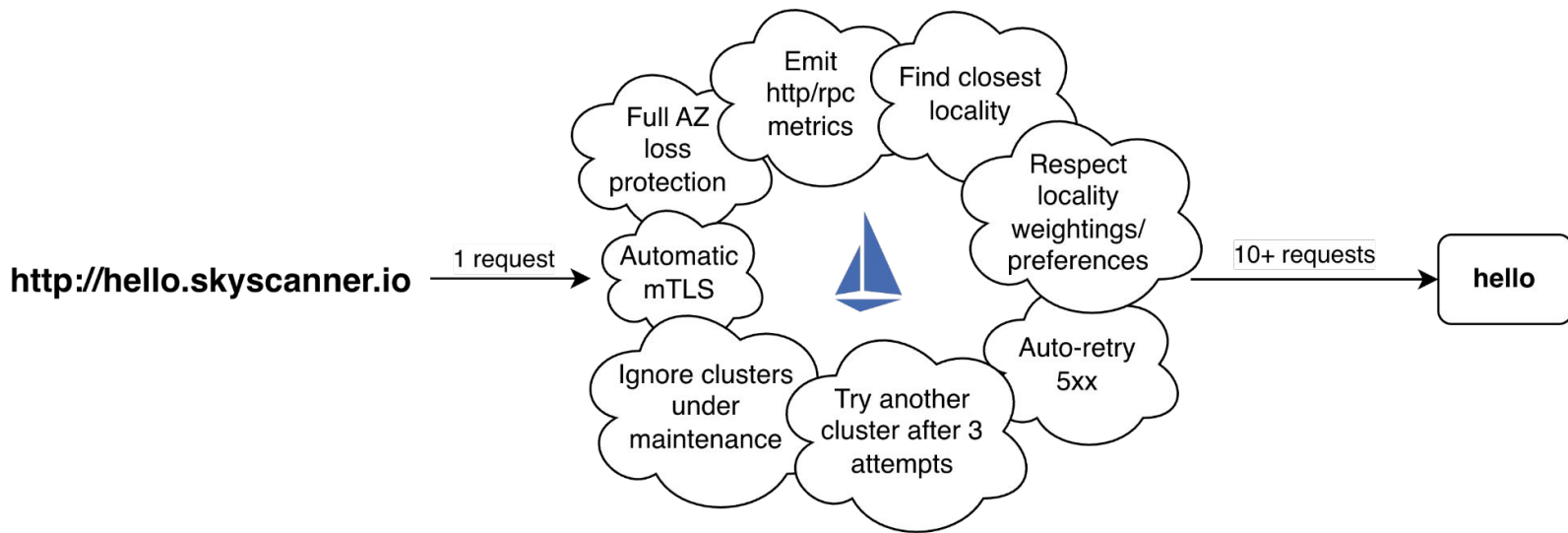


KubeCon



CloudNativeCon

Europe 2026



First Generation Mesh

2019-2024 - multicluster before multicluster

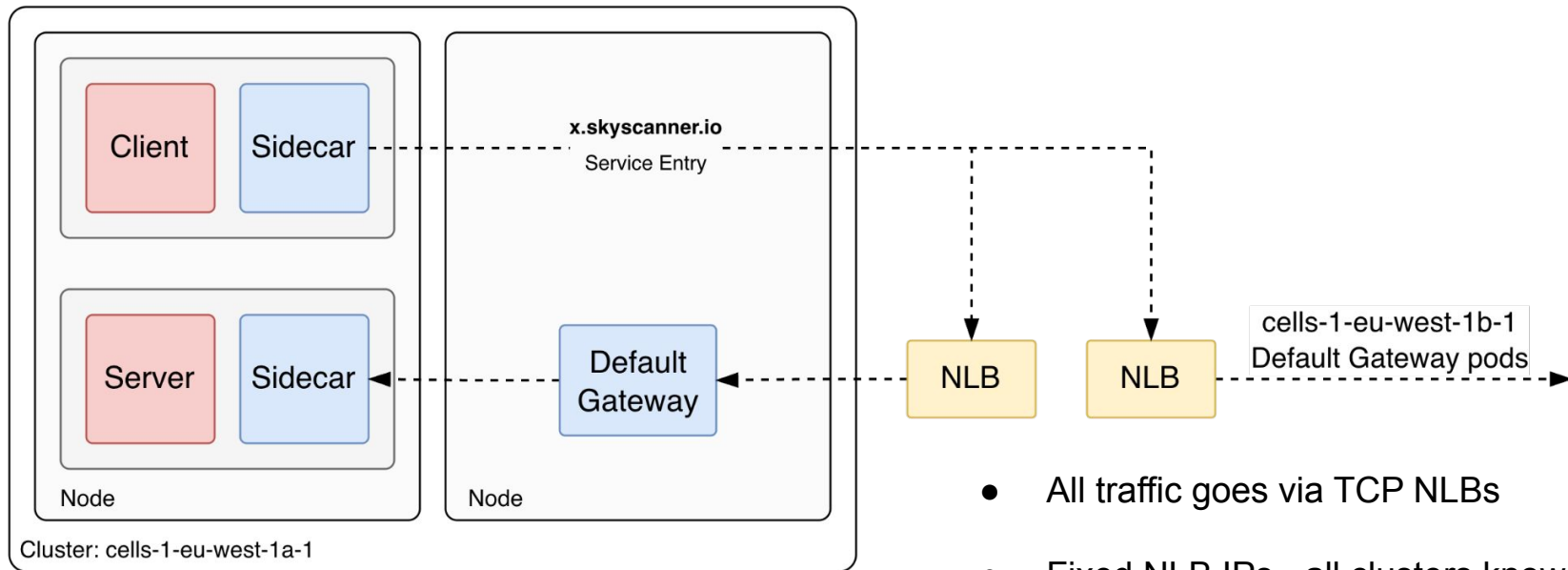


KubeCon



CloudNativeCon

Europe 2026



- All traffic goes via TCP NLBs
- Fixed NLB IPs - all clusters know them
- Costs scaled considerably

The Core Capability - Technical

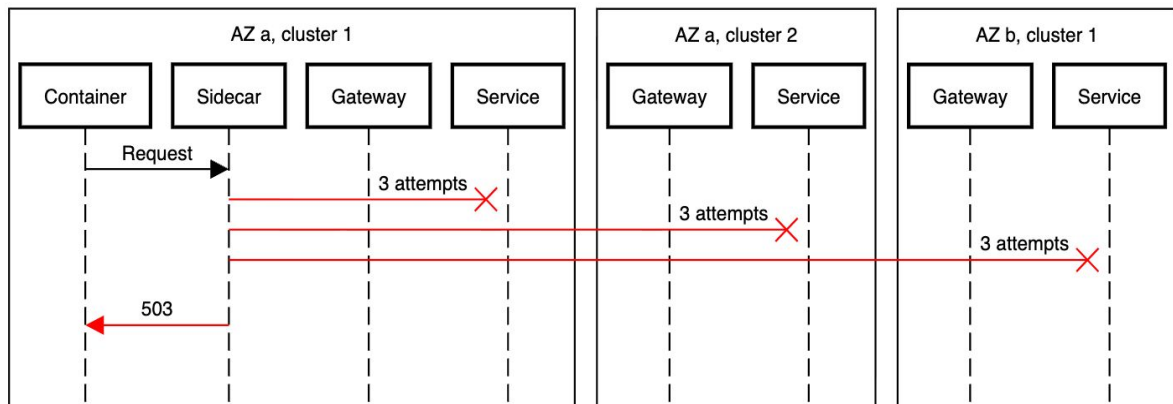


KubeCon



CloudNativeCon

Europe 2026



x.skyscanner.io resources

Service Entry

- Endpoint per target cluster
- Load Balancer static IPs
- Cluster names/AZ as localities

Virtual Service

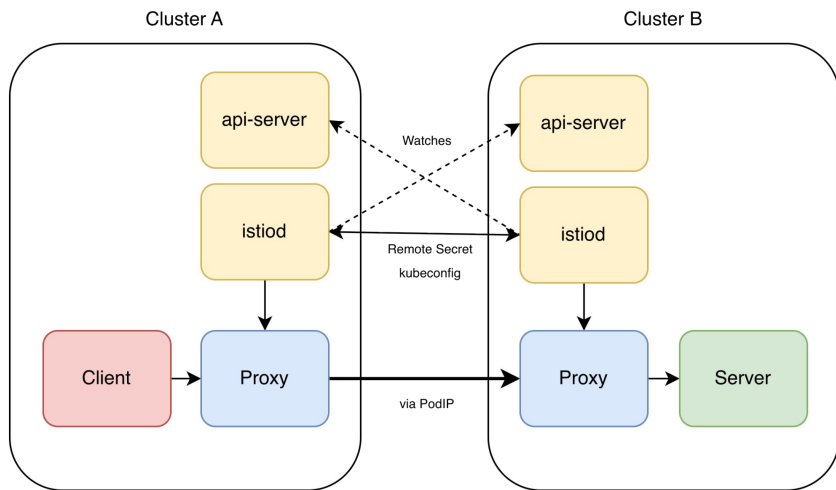
- Retries (e.g. 8 for 9 total)
- Subset selection

Destination Rule

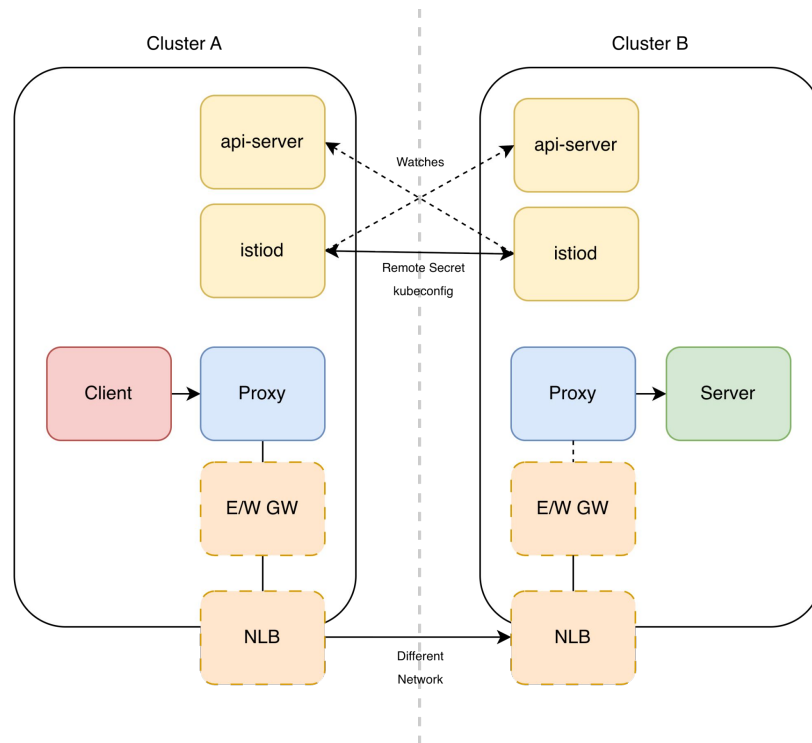
- Outlier Detection (e.g. 3 per cluster)
- Connection config, mTLS

Multi-Cluster Istio - Sidecar Mode

Flat Network



Multi-Network



How to Multicluster - Skyscanner-style?



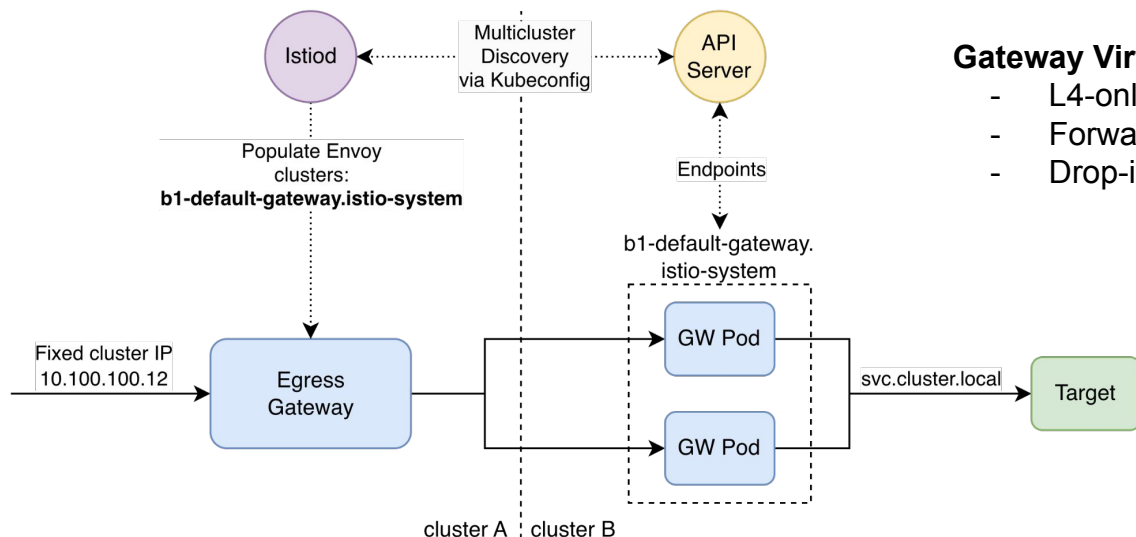
KubeCon



CloudNativeCon

Europe 2026

- Still need single IPs representing clusters
- Flat network - pod-to-pod
- Multicluster gives us all pod IPs by default
- ???



Configuration

Istio Control Plane

- Share only Default Gateway pod IPs
- Use "clusterLocal" config

Gateway Virtual Service

- L4-only - do not terminate
- Forward all traffic to remote Default Gateway
- Drop-in replacement for TCP NLB

Second Generation Mesh

2024-2026

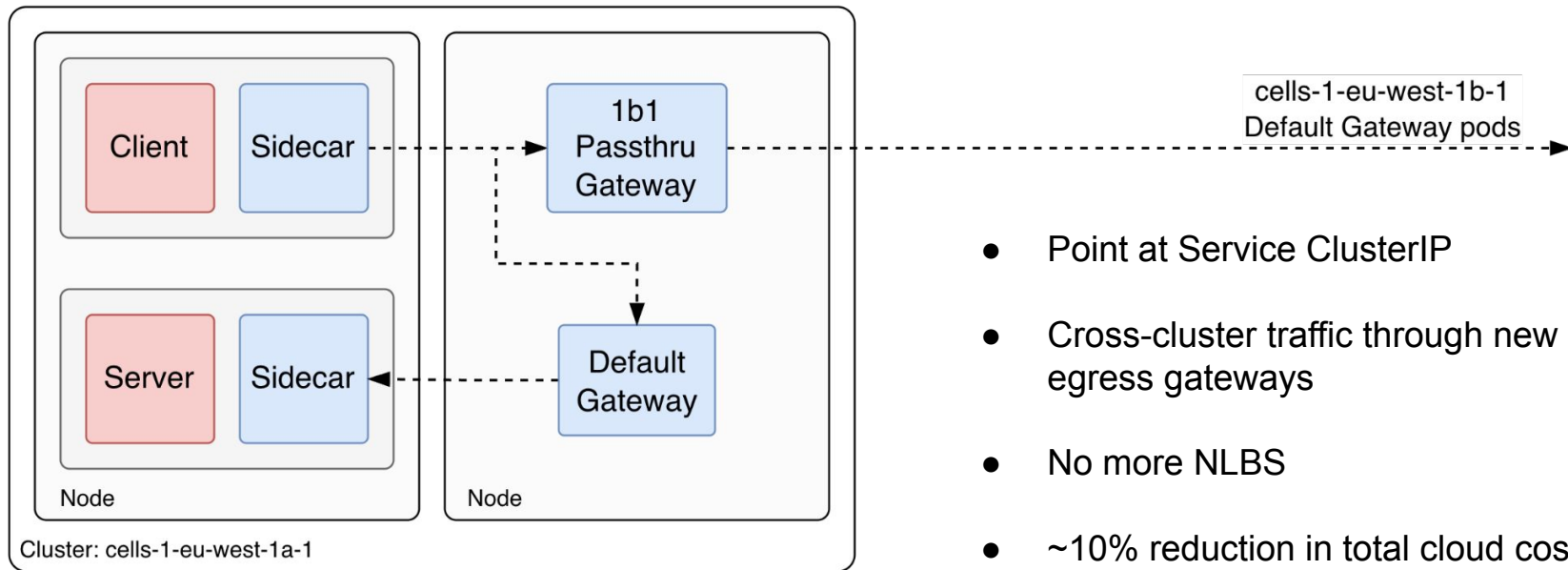


KubeCon



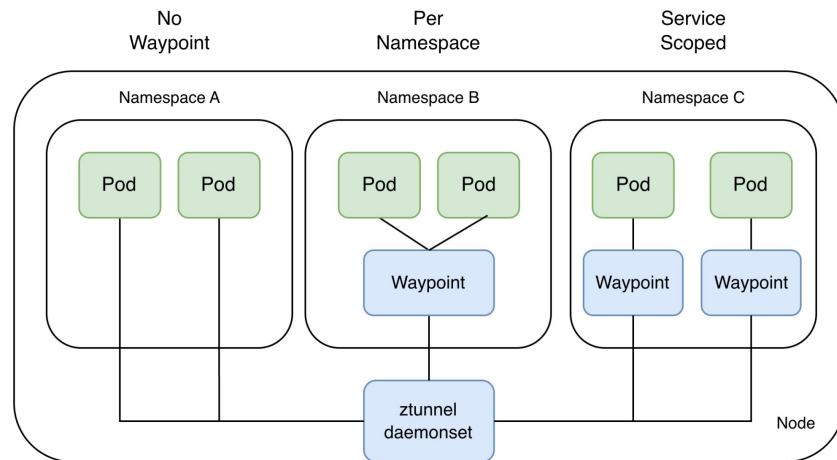
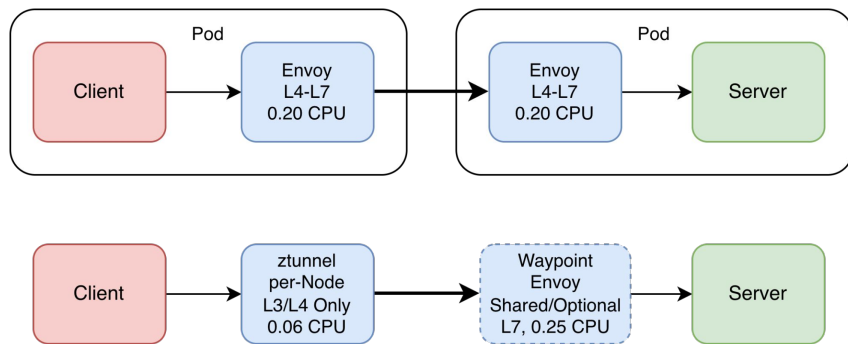
CloudNativeCon

Europe 2026



- Point at Service ClusterIP
- Cross-cluster traffic through new egress gateways
- No more NLBS
- ~10% reduction in total cloud cost
- Istiod 10x bigger for multicluster

Ambient Mesh

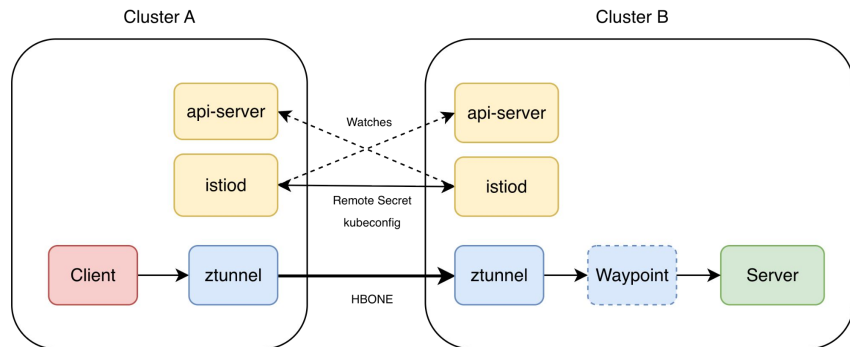


Sidecar: One proxy for **every pod** in the mesh

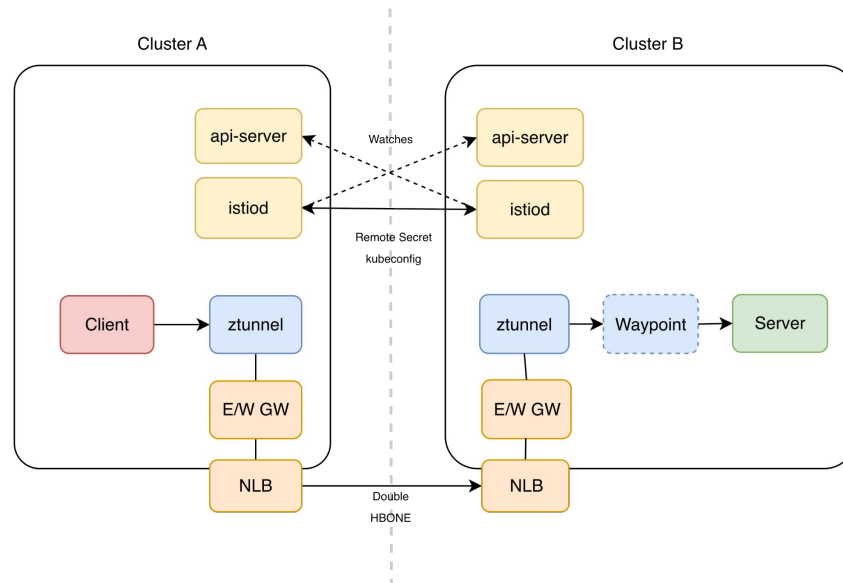
Ambient: One **per node** for L4, plus one **per NS/service/cluster when needed** for L7

Ambient Mesh - Multi-Cluster

Flat Network



Multi-Network



Ambient Mesh Primitives

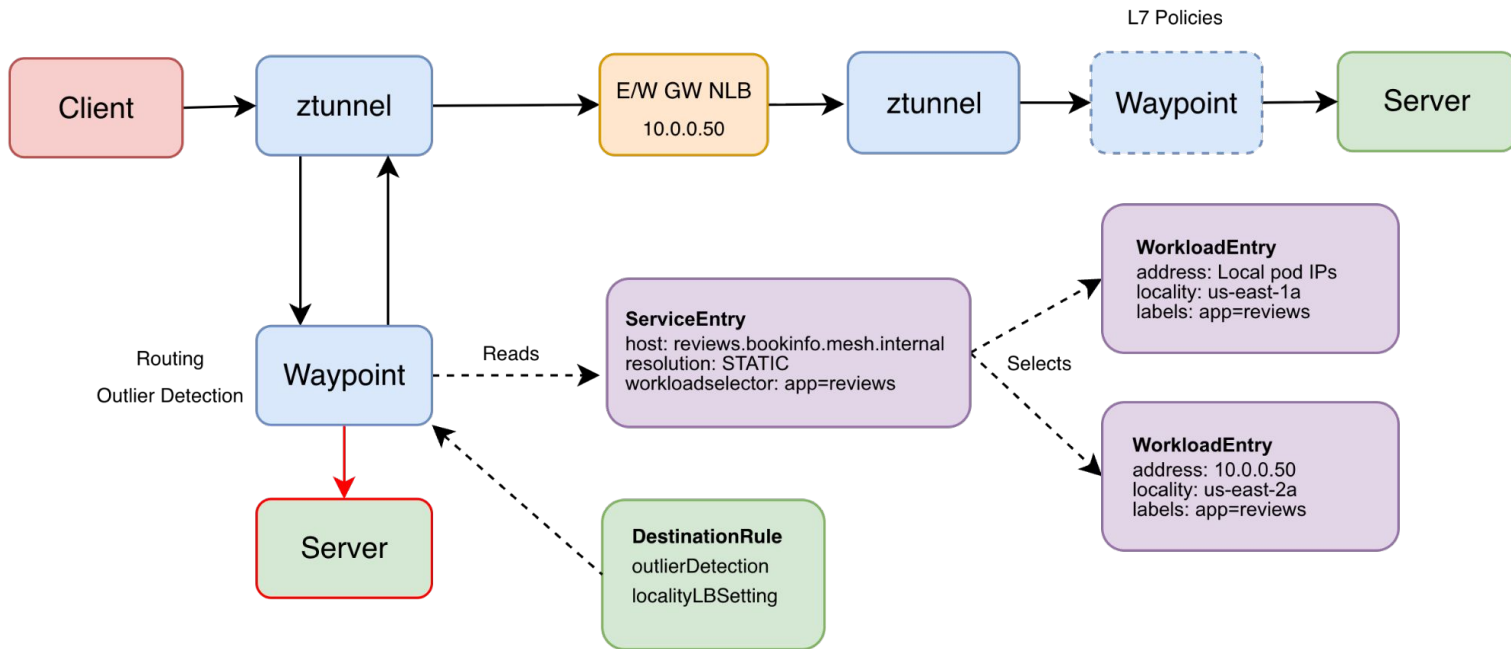


KubeCon



CloudNativeCon

Europe 2026



Moving to Ambient

- Complex existing architecture
- Moving from Sidecars to central ztunnel/Waypoints
- New multi cluster designs
- Replacing sidecar-generated telemetry
- Ambient primitives and where/how to use them

Bridging the Gap - Single IP Multicluster

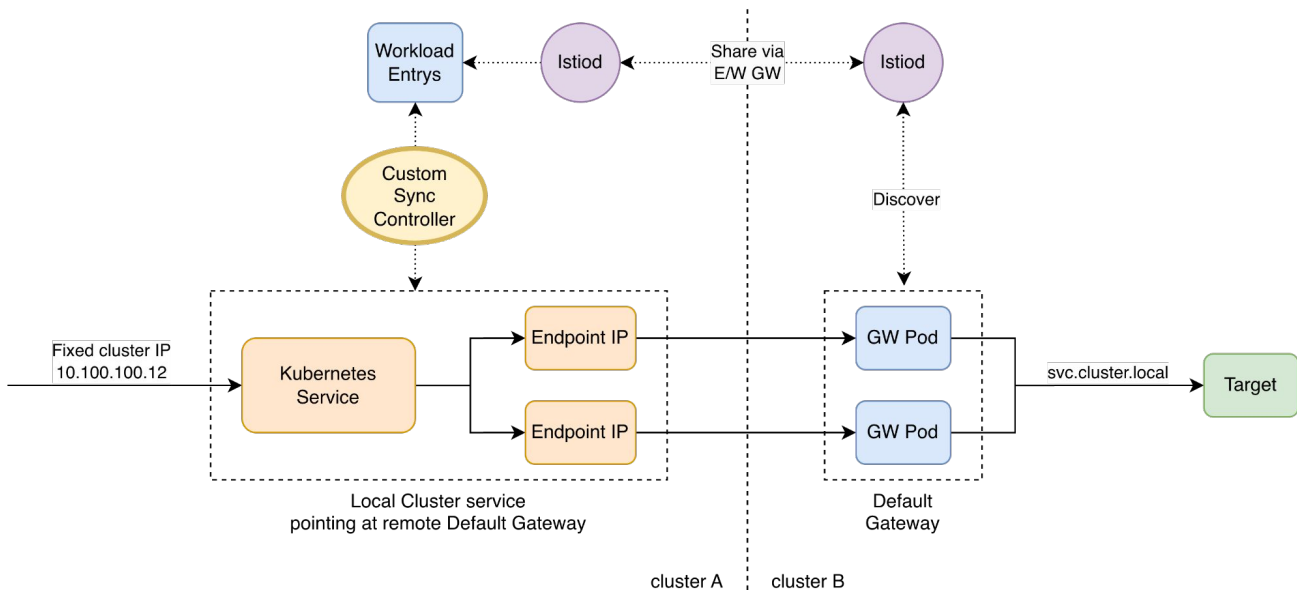


KubeCon



CloudNativeCon

Europe 2026



- Multicluster gives us workload entries per remote pod IP
- We still need single IPs representing a cluster
- Sync workload entries to k8s service Endpoints
- Remove passthrough gateways

Spanmetrics - Dashboard

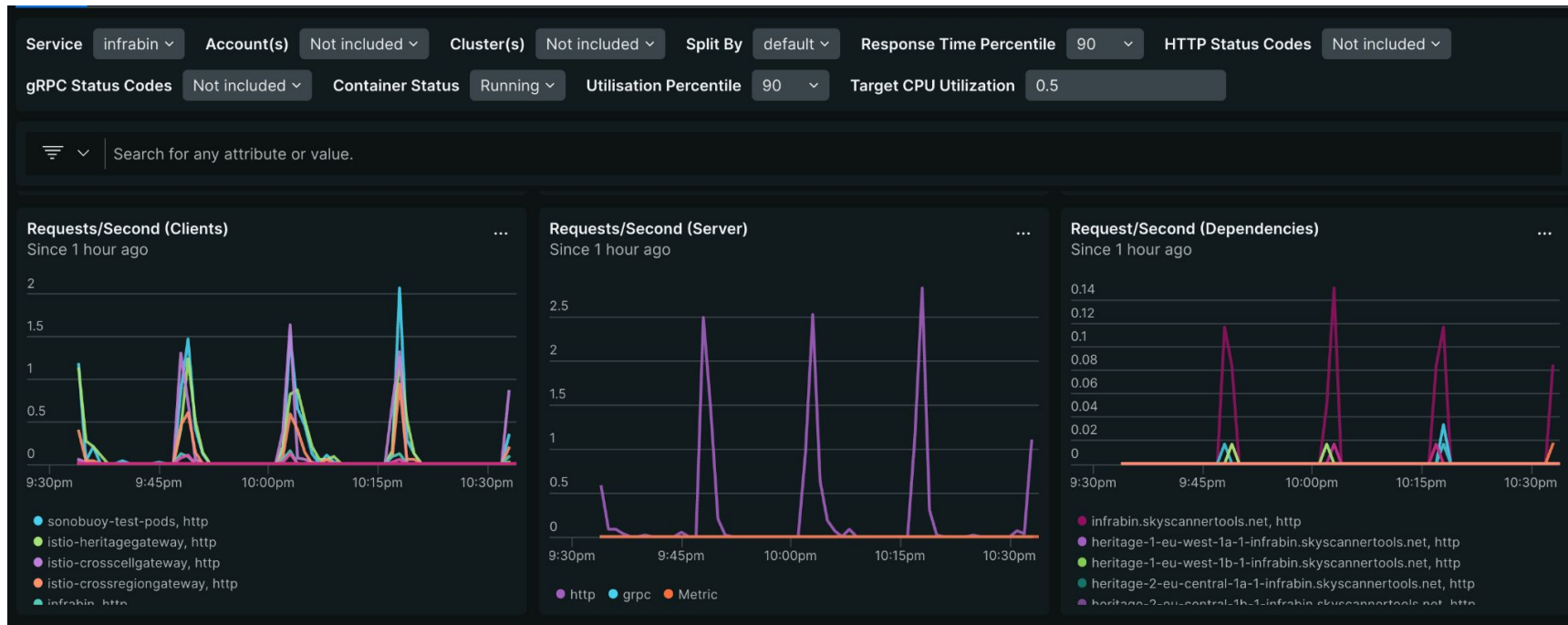


KubeCon



CloudNativeCon

Europe 2026



Spanmetrics - Sidecar

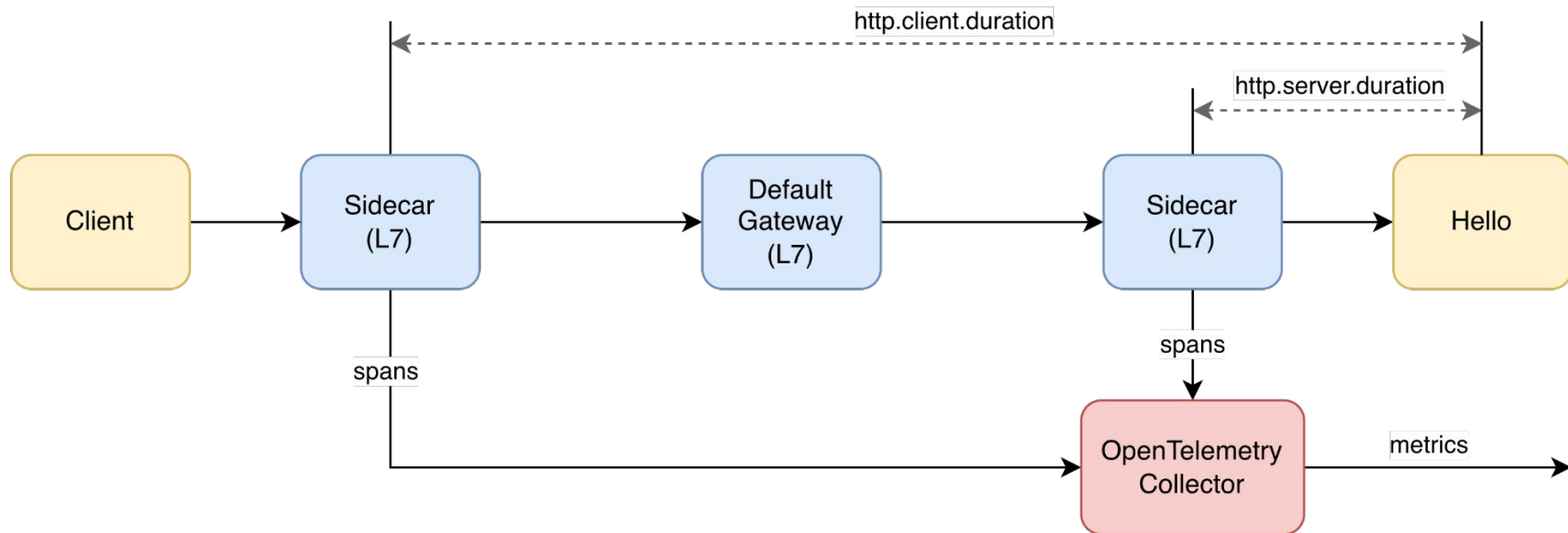


KubeCon

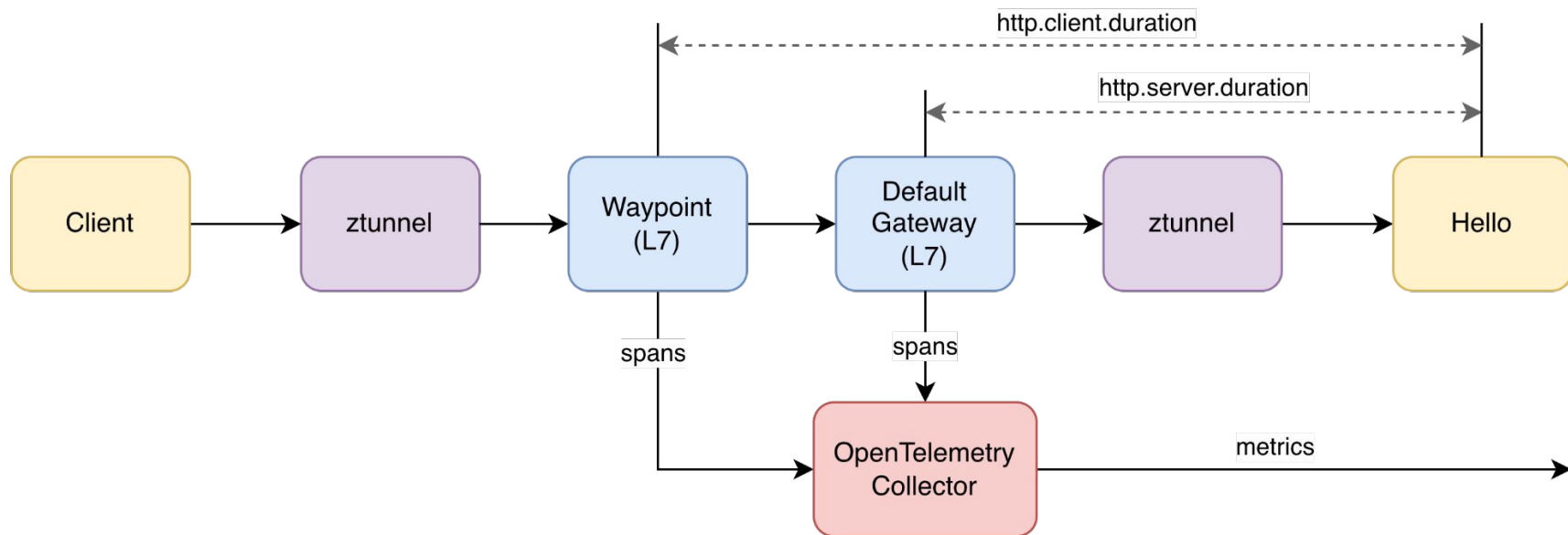


CloudNativeCon

Europe 2026



Spanmetrics - Ambient



Third Generation Mesh

2026 (in progress)

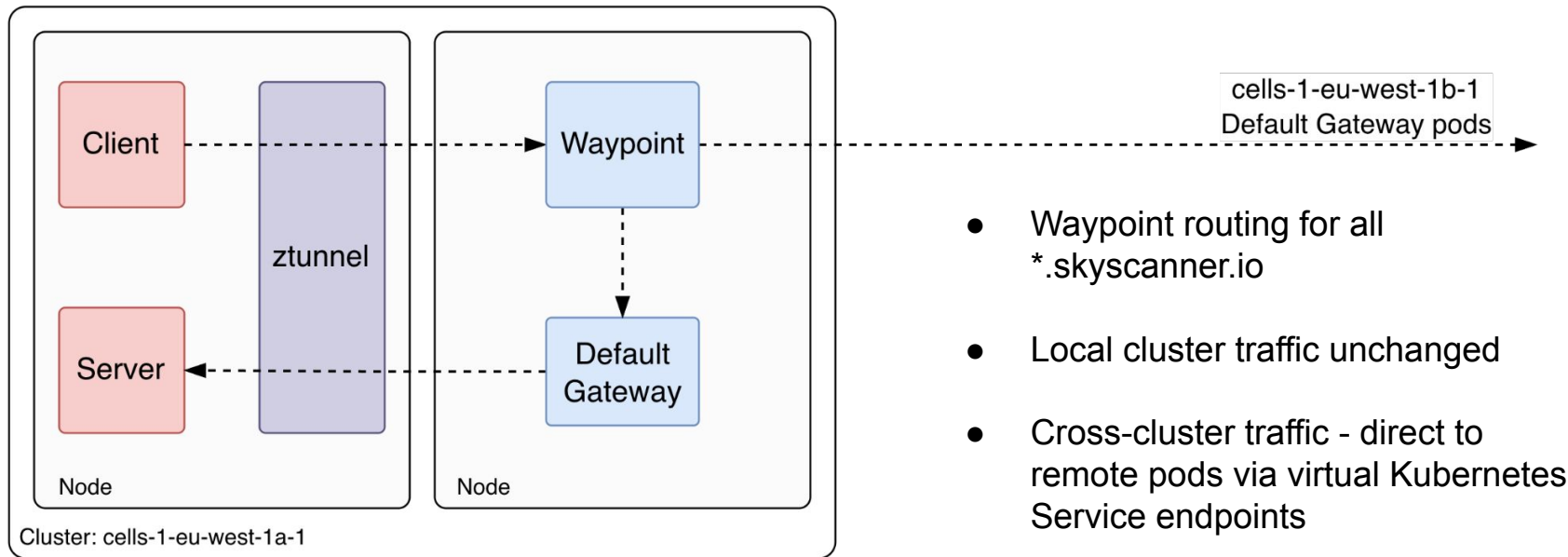


KubeCon



CloudNativeCon

Europe 2026



- Waypoint routing for all *.skyscanner.io
- Local cluster traffic unchanged
- Cross-cluster traffic - direct to remote pods via virtual Kubernetes Service endpoints
- E/W discovery/sharing, Istiod scaled right back down



- Great for greenfield platforms
- Lots of new useful abstractions and primitives
- Centralised routing decisions helps costs and service teams
- Most things work out of the box
- Difficult to fit into existing complex architectures



KubeCon



CloudNativeCon

Europe 2026

Thank You!

