SOLO.IO

Unlocking Business Efficiency With

Service Mesh Updates in AWS

WHITE PAPER

Digital transformation is not just a buzzword, but a necessity, and ensuring the seamless and efficient operation of your applications is paramount. However, as organizations strive to embrace innovation, they often encounter challenges in managing the intricate web of application networking. The complexity of orchestrating and securing microservices, combined with the evolving demands of modern architectures, presents a formidable task for businesses aiming to unlock optimal efficiency.

Enter the world of service mesh, a crucial component in navigating the complexities of application networking. While Istio has emerged as the de facto standard for service mesh solutions, its power and sophistication come with a learning curve, often posing challenges for a streamlined and cost-effective approach. Recognizing this need for simplicity and cost-efficiency, Istio's ambient mode—a groundbreaking sidecar-less model—has been created within Istio. This new mode addresses onboarding challenges to reduce infrastructure costs associated with running sidecars, paving the way for a more accessible and resource-efficient service mesh deployment.

In this white paper, we will explore how ambient mode offers a seamless approach to managing service-to-service communication, enhancing flexibility and simplifying deployments in large scale complex environments regardless of hybrid or multi-cloud.

Managing the Expansion of Services

Modern applications in AWS are often built as distributed collections of microservices, each serving specific business functions. As companies grow, so do the number of services communicating with one another – and the amount of complexity. While AWS takes care of the virtualization layer and networking, customers assume responsibility for the management and security of their applications, resulting in additional tasks for already busy teams to handle.

These tasks can include managing discovery, load balancing, failure recovery, metrics, and monitoring. This is where a service mesh, like Istio, comes into play, offering transparently added capabilities such as observability, traffic management, and security. These capabilities decrease time to market, reduce risk, and identify failures quickly and accurately so companies can focus on providing best-in-class products and services.

Understanding Service Mesh

A service mesh is a dedicated infrastructure layer that can be added to applications – it allows customers to transparently add capabilities like observability, traffic management, and security without needing to alter their own code.

Istio, the industry-standard service mesh, simplifies complex operational requirements such as A/B testing, canary deployments, rate limiting, access control, encryption, and authentication. Istio empowers companies to reduce the complexity of their service-to-service communication.

However, Istio's traditional model deploys Envoy proxies as sidecars within application workloads. While sidecars offer advantages, they have limitations. They require modifications that lead to application disruptions, consume dedicated CPU and memory resources sometimes inefficiently, and can experience traffic breaking.

Introducing Ambient Mode in Istio

To address these limitations and reduce complexity, Istio introduced 'ambient mode'. This new data plane mode is designed for simplified operations, broader application compatibility, and reduced infrastructure costs – while also preserving Istio's core features, including zero trust security, telemetry, and traffic management.

Ambient mode eliminates the need for sidecar proxies in favor of a mesh data plane that's integrated into existing infrastructure. The ambient mode secure tunnel manages and monitors each application's incoming and outgoing network traffic, but it operates at the node level.

This improvement means companies can deliver and scale faster, reduce bottlenecks, and save significant costs in delivering applications and operational simplicity for troubleshooting.

'Ambient Mode' Features

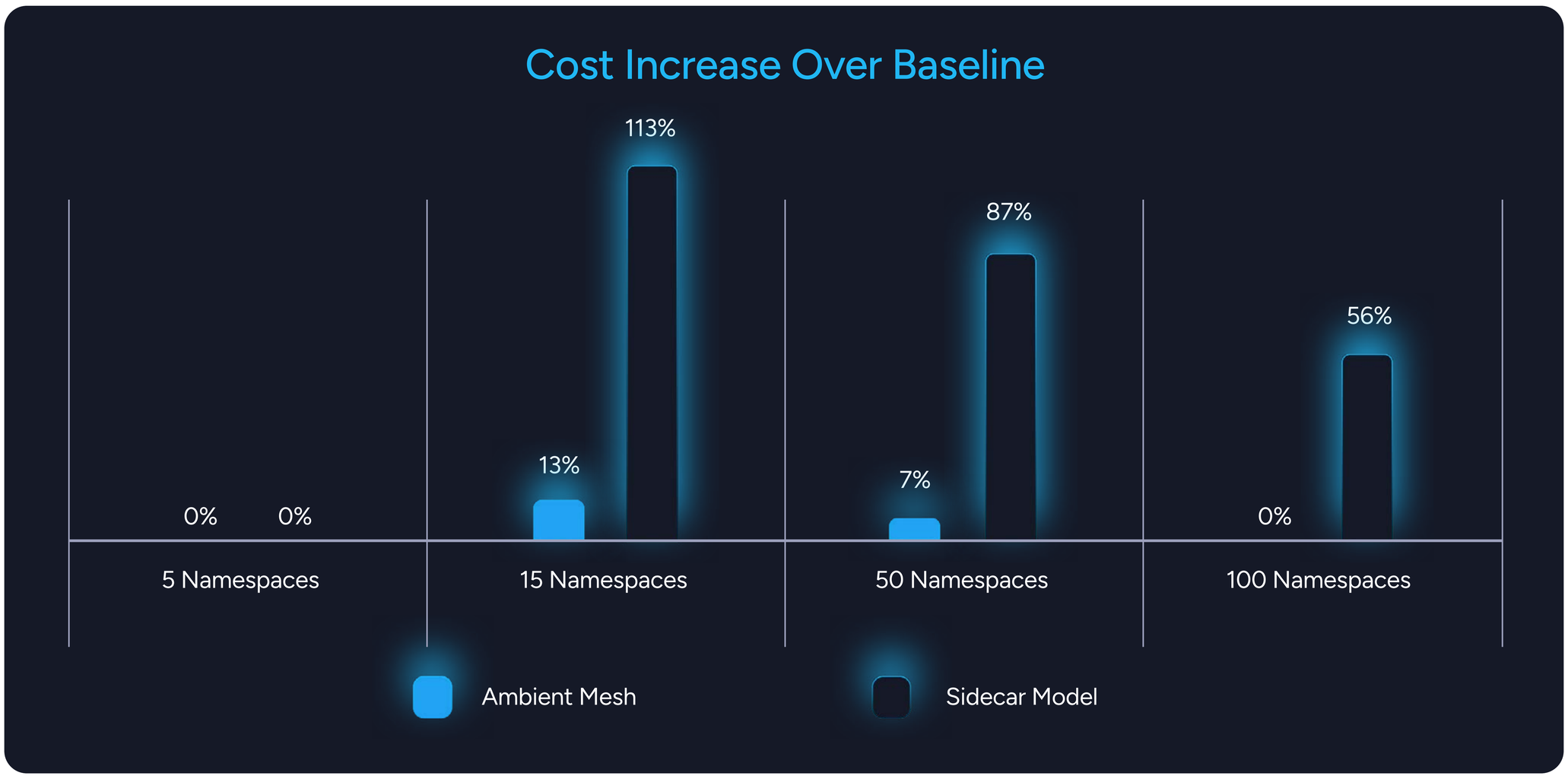
Ambient mode offers a sidecar-less approach to traffic management, splitting Istio's functionality into two layers: a secure overlay for routing and zero trust security, and optional L7 processing for advanced features.

This layered approach allows gradual adoption, seamlessly transitioning from no mesh to a full L7 processing, and coexisting with sidecar-based workloads. Istio ambient mode allows companies to more easily evolve their operations as their needs change. Benefits of ambient mode include reduced costs and resource consumption, enhanced observability and security, and efficient performance and scalability.

Reduced Costs and Resource Consumption

Ambient mode significantly reduces infrastructure costs by minimizing the number of proxies to manage. This is particularly advantageous at scale, as it reduces the compute and memory requirements per node. It also allows applications to join the mesh without requiring restarts or reconfigurations, ensuring operational continuity.

By replacing traditional sidecar models with ambient mode, companies can achieve substantial cost savings. In some cases, ambient mode implementation incurs no additional cost, making it a cost-effective choice.



Monthly Cost Estimates

A number of tests were run to discover more information related to costs, and after completion, some very insightful results were gathered. The process to activate various configurations is straightforward; it involves setting the application namespace's label to the value required by the configuration type. Subsequently, deployments within that namespace are restarted to allow the load to stabilize. The Cluster Autoscaler identifies bottlenecks, predominantly CPU-related for the specified instance type.

The accompanying table illustrates a significant increase in memory demand in the sidecar model. The number of EKS Nodes is automatically adjusted using ASG, aligning with Cluster Autoscaler recommendations. Post adjustments, metrics are collected and analyzed using the AWS Pricing Calculator. The detailed analysis can be found in the table below.

Size (Namespaces)	Configuration (Mesh Model)	CPU (Cores)	Memory (Gib)	Nodes (EC2 Instances)	Monthly Cost (USD)	Surplus (%)
5	Baseline (No Mesh)	1.76	5.36	1	$71.67	---
	Istio Ambient Mode	1.96	5.61	1	$71.67	0.00%
	Sidecar Model	2.06	5.71	1	$71.67	0.00%
15	Baseline (No Mesh)	30.7	43.81	8	$573.37	---
	Istio Ambient Mode	33	47.93	9	$645.04	12.50%
	Sidecar Model	60.5	90.74	17	$1,218.41	112.50%
50	Baseline (No Mesh)	53.7	76.41	15	$1,075.07	---
	Istio Ambient Mode	59.6	84.97	16	$1,146.74	6.67%
	Sidecar Model	105	156.1	28	$2,006.80	86.67%
100	Baseline (No Mesh)	138	199.56	43	$3,081.87	---
	Istio Ambient Mode	148	212.06	43	$3,081.87	0.00%
	Sidecar Model	259	379.43	67	$4,801.98	55.81%

As these graphs and charts demonstrate, utilizing ambient mode over a traditional sidecar approach results in a much smaller cost increase per namespace.

Configuration practices also remain consistent, so there's no need for changes in your operational pipelines, ensuring a seamless transition and maintaining operational efficiency while integrating new technologies.

Enhanced Obervability and Security

Ambient mode enhances observability because it is node-based, allows for network traffic tracing, and supports various traffic types due to its L4-based approach. Setting these tools up without Istio would require deep expertise and significant development time.

Ambient mode offers security features such as mutual TLS and access control, helping companies ensure compliance with security standards and regulations.

Efficient Performance and Scalability

Istio ambient mode optimizations and its node-based approach mitigate performance concerns. Ambient mode becomes more efficient as applications scale, so growing businesses don't incur significant infrastructure costs as they grow.

Start Using Ambient Mode in Istio

Ambient mode in Istio represents a strategic choice for companies seeking to reduce operational costs and complexity in AWS. Its ability to streamline microservices, reduce infrastructure costs, and provide enhanced monitoring and security make it an attractive option for modern applications.

By embracing ambient mode, companies can harness the full potential of Istio while optimizing their operations and enhancing their bottom line. Get started with the Solo and AWS ambient mode EKS add-on today.

solo.io

The Gateway to AI Innovation