

Service Mesh

What is it?

An infrastructure layer that facilitates *communication*, as well as *observability*, *security*, and *load balancing* between discrete components of a microservice application.

How does it work in the context of distributed systems?

A service mesh consists of two key components:

1. Proxy instances called **sidecars**:
 - Attached to each service instance
 - Handle communication to and from the service instance
 - Also take care of monitoring and security
 - Service instances, sidecars, and their interactions form the “**data plane**” – this refers to the fact that this is the part of the architecture that carries user traffic
2. The “**control plane**”:
 - Oversees creation of service instances, and implementation of network security and management protocols
 - Often connected to a GUI for application management

Why is it beneficial to a technology organization?

- Abstracts inter-service communication from business logic
- Simplifies resolution of communication-related bugs
- ↑ # of microservices == ↑ benefits of using a service mesh
- Automates much of the process of scaling

What are its use cases?

- **Highly regulated industries:** Service meshes allow consistent and uniform implementation of security protocols.
- **Optimizing application performance (and debugging):** Increased observability can help highlight a whole host of problems.
- **Meeting rapidly growing demand:** A service mesh automates away the creation of new service instances, allowing an application to seamlessly handle increasing traffic.

Drawbacks

- A developer is required to set up and configure a service mesh
- Service meshes can be very costly
- The control plane could act as a single point of failure
- Extra step involved in communication (passing through side car) could result in increased latency

