1. Forward Proxy (Proxy de Encaminhamento)

🔎Definição

💡Um proxy de encaminhamento atua em nome de clientes (usuários ou dispositivos) para acessar recursos na internet. O cliente configura suas configurações de rede para rotear solicitações por meio do servidor proxy, que então as encaminha para os servidores de destino.

🔧Onde é utilizado

🧲Redes corporativas para controlar e monitorar o acesso de saída à internet.

🧲Para armazenar conteúdo em cache e melhorar os tempos de resposta.

🧲Para tornar o tráfego de usuários anônimo.

🧲Para filtrar solicitações por motivos de segurança ou conformidade.

2. Service Mesh (Malha de Serviço)

🔎Definição

💡Uma malha de serviço é uma camada de infraestrutura dedicada que gerencia a comunicação entre serviços em arquiteturas de microsserviços. Normalmente, ela usa proxies sidecar implantados junto com cada serviço para lidar com questões de rede como roteamento, novas tentativas, segurança e observabilidade.

🔧Onde é utilizado

🧲Em arquiteturas de microsserviços para gerenciar comunicações complexas entre serviços.

🧲Para descoberta de serviços, balanceamento de carga, segurança e observabilidade.

🧲Para implementar políticas de segurança de confiança zero entre serviços.

3. API Gateway

🔎Definição

💡Um Gateway de API é um servidor que atua como um único ponto de entrada para que aplicativos clientes acessem múltiplos serviços de backend. Ele lida com roteamento de solicitações, composição, autenticação, limitação de taxa e, às vezes, transformação.

🔧Onde é usado

🧲Em arquiteturas de microsserviços para expor APIs.

🧲Para agregar múltiplos serviços em uma única API.

🧲Para políticas de segurança, limitação de taxa e análises.

🧲Para ocultar a complexidade interna do serviço dos clientes.

4. Load Balancer (Balanceador de Carga)

🔎Definição

💡Um balanceador de carga distribui o tráfego de rede de entrada entre múltiplos servidores para garantir a confiabilidade e a utilização ideal dos recursos. Ele pode operar em várias camadas (Camada 4 — TCP/UDP, Camada 7 — HTTP/HTTPS).

🔧Onde é usado

🧲Para distribuir tráfego entre servidores web, servidores de aplicativos ou bancos de dados.

🧲Para melhorar a tolerância a falhas.

🧲Em arquiteturas de alta disponibilidade.

Para escalar horizontalmente.

