Architecture of Content-Based Service Router

draft-lin-dmsc-content-based-service-router-01

Changwang Lin (New H3C)
WeiWang (China Telecom)
Xueting Li (China Telecom)
Haiyang Zhang (New H3C)

IETF-123, July 2025

Background and Motivations

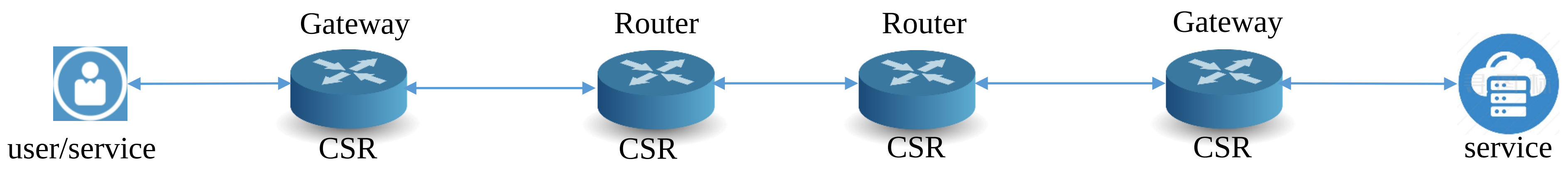
- **As a dedicated infrastructure for micro-service communication, Service Mesh has evolved from a single micro-service to a sidecar model as micro-services increase.**

  It faces many challenges, such as increasing complexity, significant performance overhead, and low maturity, which restrict its large-scale application.

- **Distributed Micro Service Communication (DMSC) was proposed to optimize the communication between micro-services through distributed processing, enhancing network efficiency and reliability.**

- **Content-based Service Router (CSR) is the main switching component of the DMSC architecture; CSR performs routing optimization based on service prefixes and topology information, and exchanges service prefixes and topologies based on distributed routing protocols, so CSR is crucial for the routing reachability of distributed micro-services.**

CSR Architecture*



CSR

Control Plane

Mainly includes **Routing Protocols** and **Routing Management**, and is responsible for the exchange of **service prefixes**, **topology information** and the generation and optimization of **routing information**.

Data Plane

Mainly performs **Routing Lookup** and **Data Forwarding** based on the service prefix to ensure the **effective transmission** of data packets.
*Or Data Plane may **provide Data Decryption and Encryption** to achieve several features of service mesh, such as **traffic control, zero-trust network, and observability**.

Service Plane

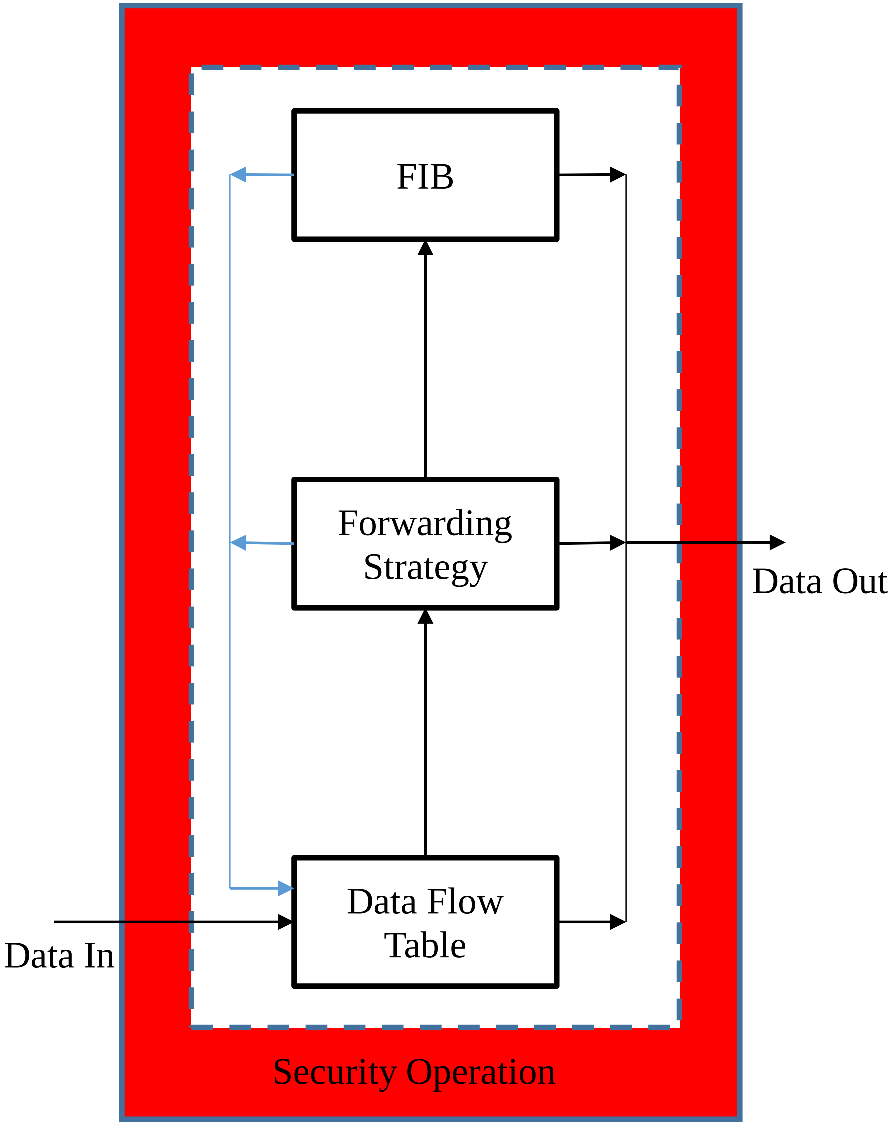
Be used to access **Online Services** and provide **Service Responses**, and to advertise service prefixes to the Control Plane for publication, when acting as Service Gateway.

Control Plane

- **Routing Protocols**
  - Includes **Static** and **Dynamic** Routing Protocols
  - for exchanging service prefixes and topology information.
  - Each sends its own optimal routes to the **Routing Management**.

- **Routing Management**
  - Responsible for integrating and optimizing routing information from all **Routing Protocols**
  - and distributing the best service routes to the **Data Plane**.

Data Plane*



◆ Forwarding Information Base (FIB)

- Receive the service routes from Control Plane.
- Used to guide the forwarding of service packets.
- Mainly consists of **Service Name** and **Outgoing Interface**.

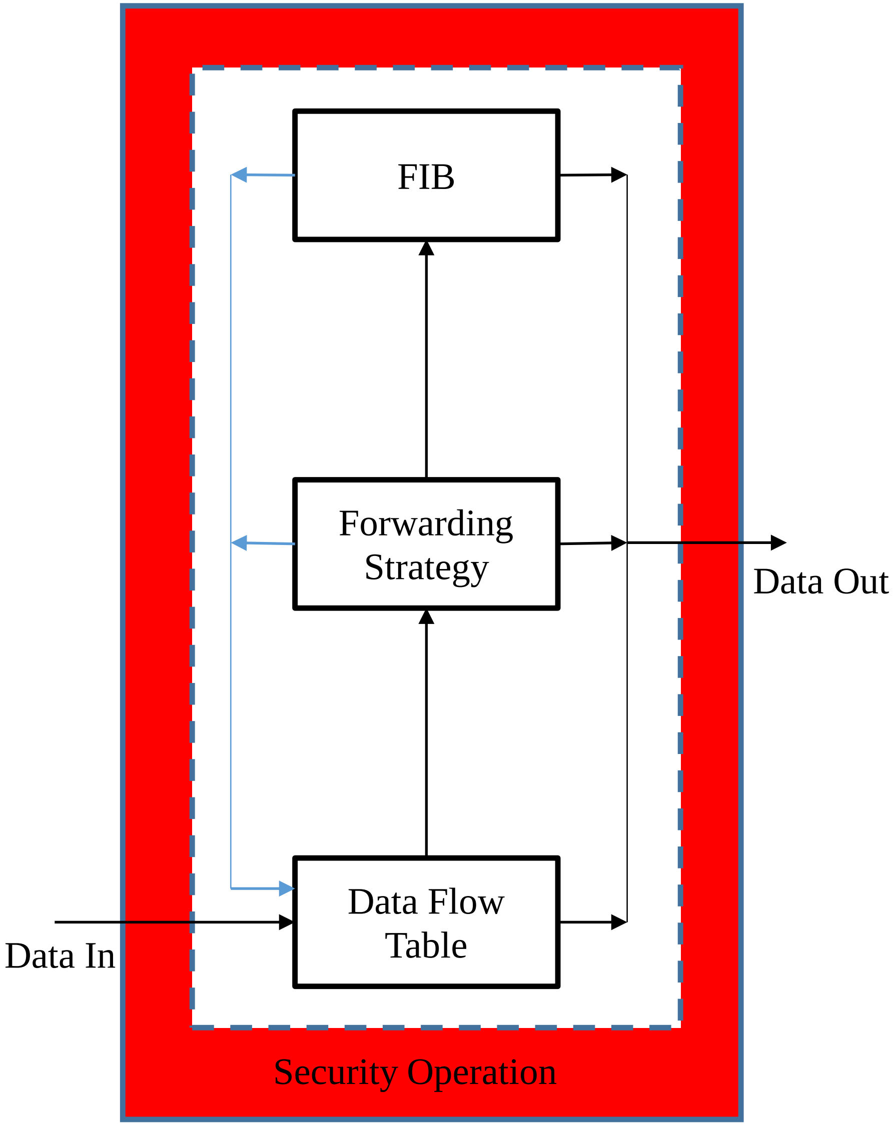
◆ Forwarding Strategy

- Used to implement special forwarding requirements through configuration, such as ACL and Flow Spec.
- Take precedence over normal **FIB**.

◆ Data Flow Table

- Maintain traffic **stickiness** and Improve Forwarding performance (No need to search FS and FIB for Non-first traffic).
- Record and monitor data packet information to help locate communication anomalies.
- May consist of the **Destination Service Name**, **Source Service Name**, and **Outgoing Interface**. Other Parameters (Future)

Data Plane*

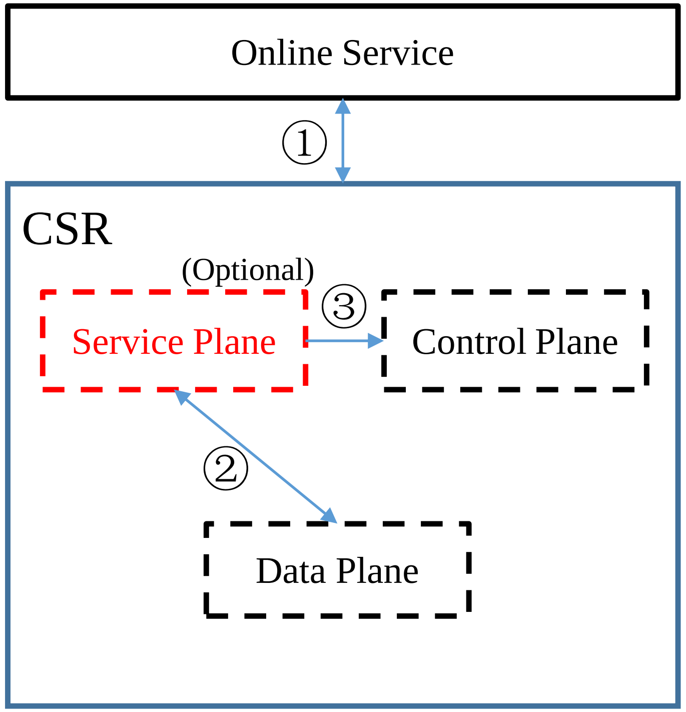


◆ **Security Operation**

- Check whether need security operation (data decryption) when receiving data (Data In).
- Check whether need security operation (data encryption) when sending data(Data Out).
- Do data decryption when receiving data.
- Do data encryption when sending data.

Service Plane

1. **Provide access to Online Service**
   - If online service need direct access to CSR when acting as Service Gateway.



2. **Provide fast service response**
   - Store the service data within a certain time;
   - If the request service is local, the service data can be obtained directly.

3. **Publish Online Service prefix**
   - Send online service prefix to Control Plane .
   - Publish online service prefix through the Control Plane.

Compared with Traditional Router

The architecture of CSR is similar to that of Traditional Router !

Classification	Prefix Characteristics	Forwarding Packet Characteristics
CSR	Service Name, variable length and hierarchical structure (example: /example/video/segment1)	The DMSC Data message should contain a Source Service Name and a Destination Service Name.
Traditional Router	IP Address, fixed length (4 or 16 bytes) (example: 1.1.1.1 or 100:1::1)	Traditional Data message contains Source IP Address and Destination IP Address.

Challenges in CSR:

- ✓ How to **control** and **optimize** CSR prefix length to **reduce** Packet Payload?
- ✓ How to **achieve high-performance** Service Name routing lookup? In other words, How to **store** the Service Name route? Hardware and Software.
- ✓ When the CSR need the Security Operation (Data Decryption/Encryption) , the Forwarding performance need to be considered.

Implementation Considerations

- Control Plane

Routing Protocol

Choose dynamic Routing Protocols such as BGP, OSPF, ISIS, etc., and perform protocol extensions on them to publish Service Name Prefixes.

BGP： Add new address family and new NLRI to support the publishing of Service Name prefixes;

OSPF, ISIS： Add new type in LSDB to support for publishing Service Name Prefixes;

Static: Add new command to configure service name prefix routing.

Routing Management

Added a new Service Name Routing Table to store Service Name Routes from various Routing Protocols.

Service Name Routing includes :

- Service Name Prefix (example: /example/video/part1)
- Outgoing Interface (interface1)
- Priority
- …

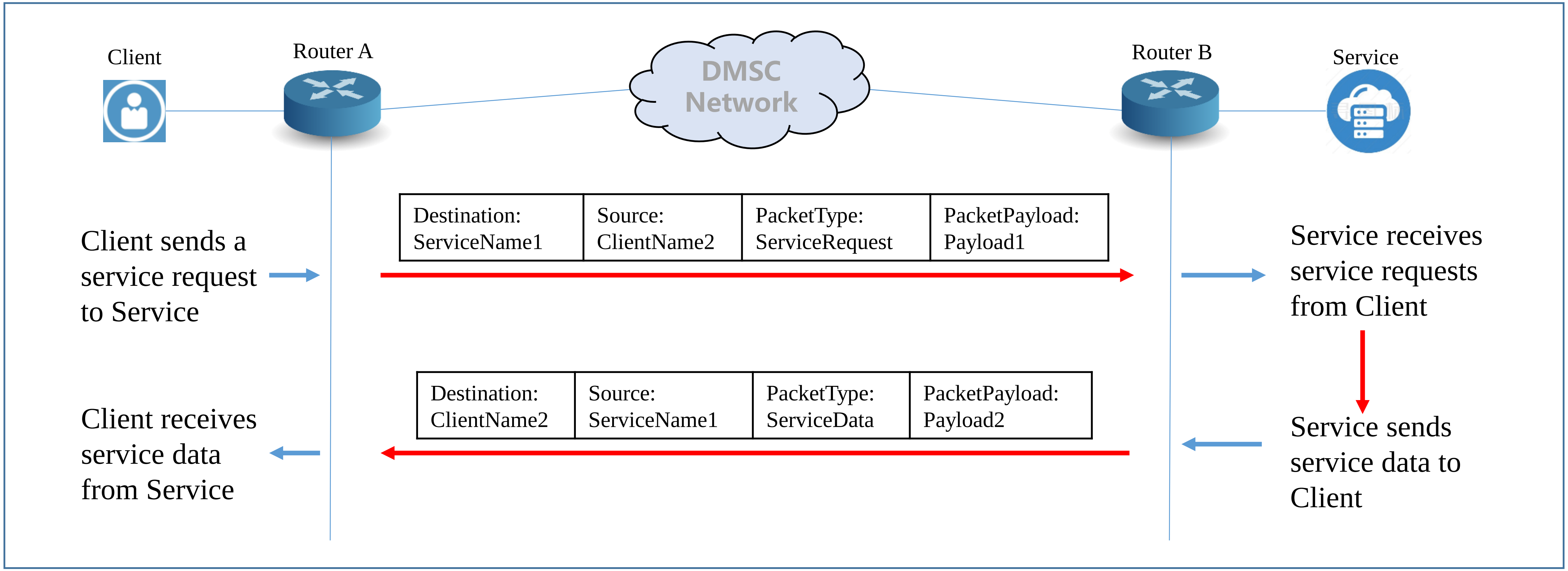
Implementation Considerations

- Data Plane

Refer to IPv4 or IPv6 packet format :

DMSC packet format (New IP version)

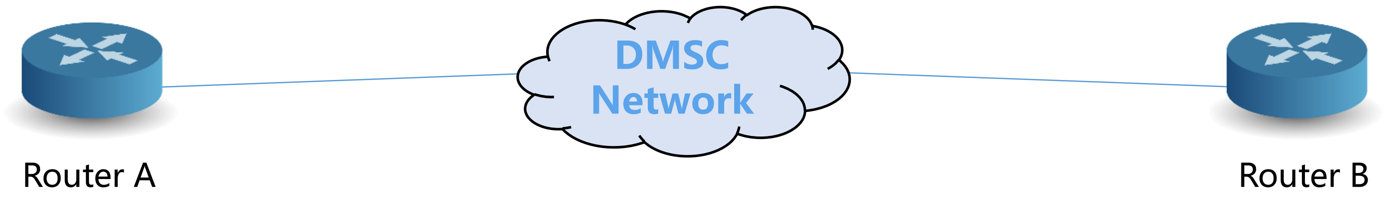
Destination Service Name	Source Service Name	Packet Type	Packet Payload



Implementation Considerations

- OAM (DMSC PING)

DMSC PING packet : similar to IPv4 or IPv6 ping packet for detecting the reachability between devices.



- Record the sending timestamp T1 of the **EchoRequest** packet;
- Record the received timestamp T2 of the **EchoReply** packet from the destination service;
- Round-Trip Time（RTT）= T2 – T1;
- If the sending of the **EchoRequest** packet times out for 1s, the PING is considered to have failed.

Destination: Router B	Source: Router A	PacketType: **EchoRequest**	PacketPayload: xxxx

Sending Timestamp T1

Destination: Router B	Source: Router A	PacketType: **EchoReply**	PacketPayload: xxxx

Received timestamp T2

- Receive **EchoRequest** Packet；
- If the destination is not the local service (internal service of the router), continue forwarding until the server is found;
- If the destination is the local service, then send an **EchoReply** packet in the reverse direction;

Next Steps

- **Continuously improve the CSR framework.**
- **Explore the service name routing lookup method (variable length prefix lookup method).**
- **Consider the proper format of DMSC packet (Maybe refer to IP Packet).**
- **Consider how to improve forwarding performance when encrypting and decrypting data.**
- **Any comments welcomed. Or can you give me your suggestions?**

Thanks!