DMSC Working Group
Internet-Draft
Intended status: Standards Track
Expires: 5 January 2026

C. Lin
New H3C Technologies
W. Wang
X. Li
China Telecom
H. Zhang
New H3C Technologies

4 July 2025

Architecture of Content-Based Service Router
draft-lin-dmsc-content-based-service-router-01

Abstract

This document first describes an architecture of Content-based Service Router (CSR), which is used to exchange service prefixes and topology information based on distributed routing protocol, and optimize routing based on service prefixes and topology, as one important component of Distributed Micro Service Communication (DMSC) architecture.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 5 January 2026.

Table of Contents

1. Introduction

With the continuous emergence of various applications, Micro-services, as small and independent application segments, are becoming increasingly dense, making communication between Micro-services increasingly important.

A Service Mesh serves as a dedicated infrastructure for handling communication between Micro-services, providing functions such as Traffic Management and Secure Communication. The Service Mesh has evolved from general communication middleware functions, starting from monolithic Micro-services, evolving into integrated middleware, and finally developing into a sidecar pattern. However, the service network presents some challenging issues, as follows:

* Increased complexity: Implementing and deploying a Service Mesh requires adding additional components to the current application, significantly increasing the difficulty of mastering the Service Mesh.

* The performance overhead increases: The sidecar pattern requires deployment within the same Pod as the Micro-service, resulting in a tightly coupled mode in terms of resource usage, which significantly affects tenant resource usage efficiency.

* Maturity and acceptance are low: Although there have been several deployments of Service Mesh, it is still far from mature compared to traditional networks and related solutions.

To address the challenging issues of Service Meshes, a content-centric Micro-service communication architecture, called the Distributed Micro Service Communication (DMSC) architecture [I-D.li-dmsc-architecture], is proposed to enhance the efficiency and reliability of communication between Micro-services. DMSC has the following characteristics:

* Content-centered: Focus on content and services, not on business location.

* Decentralization: Registration, routing, and storage of content and services using distributed processing.

* Dynamic Resource Allocation: Optimization of resource allocation to enhance network efficiency.

* Scalability and flexibility: meets the needs of continuous network evolution and supports large-scale deployments comparable to current operator networks.

The DMSC architecture consists of four key parts: the Service Gateway (SG), the Service Router (SR), the Service Prefix Authentication (SPA) system, and the Service Mesh Communication Scheduling Center (SCSC) system [I-D.li-dmsc-architecture].

The SG is used for flexible adaptation of communication between various existing Micro-services. SRs are used to optimize routing based on service prefixes and topology, and to exchange service prefix and topology information based on distributed routing protocols. The SPA system is used to authenticate distributed service prefixes. The SCSC system provides auxiliary centralized optimization scheduling for Micro-service routing.

The SR holds a very important core position in the DMSC architecture. It serves as the main switching component for achieving routing reachability in distributed Micro-services, aiding in the large-scale deployment of Micro-service communication. Therefore, to implement the SR in the DMSC architecture, this document provides the first description of the structure of a Content-based Service Router (CSR).

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. CSR Architecture Overview

The CSR architecture is similar to the traditional router architecture, with the only difference being that traditional routers use address prefixes as routing and forwarding information, while CSR use content (name) prefixes as routing and forwarding information. The address prefix is a fixed-length IPv4 or IPv6 address, while the name prefix is a variable-length character string [RFC8569].

As Micro-services continue to increase, service names may become longer, which also causes the name prefixes in a request message to lengthen, thus increasing the package load. To reduce the package load, name prefixes may need to be compacted or optimized, and the method of compaction or optimization is beyond the scope of this document. For the implementation of the CSR, the most significant challenge is how to efficiently match the compressed or optimized name prefixes in the request message with the forwarding information base (FIB) for transmission through the network to a system that can issue a response. How to compact or optimize name prefix also is beyond the scope of this document.

The hierarchy of a name is used for routing via the longest matching prefix in a CSR of Content-Centric Networking. The longest matching prefix is computed name segment by name segment in the hierarchical name, where each name segment must be exactly equal to match. For traditional routers based on IP addresses, the longest match prefix is performed by applying a bitwise AND operation between the destination IP address and the subnet mask in the forwarding table. The result is then compared with the network address of the forwarding entry. If they match, it is a match; otherwise, it is not a match. So, if the name prefix is long, calculating the longest match for the name prefix will be more complex than for the address prefix. How to efficiently execute the longest match of name prefix also is beyond the scope of this document.

Like traditional routers, The CSR architecture SHOULD be divided into the Control Plane and Data Plane, as shown in Figure 1. The Control Plane primarily exchanges Service Prefixes and topology information with Service Gateways (SG) or Service Routers (SR) through routing protocols, generates service routes, and delivers them to the Data Plane. The Data Plane mainly receives packets of service and performs look-up forwarding based on the Service Prefix (Service Name) carried in the data packet.

The CSR may optionally include a Service Plane to act as SG, which can be used to access online services and provide service response. The Service Plane also needs to send the service prefix to the Control Plane for publishing. The Service Plane is not a necessary

function component of CSR, when CSR do not act as SG. If CSR acts as SG to access online services, the Data Plane may provide security decryption and encryption to achieve several features of service mesh, such as traffic control, zero-trust network, and observability.

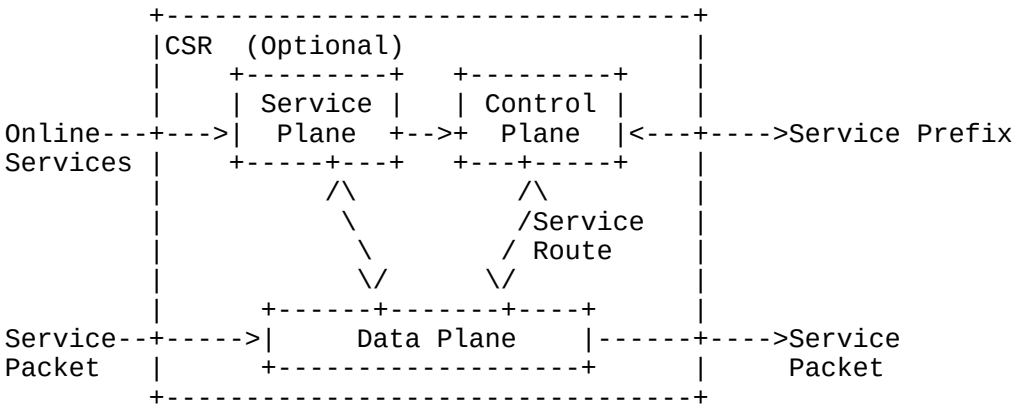


Figure 1: The CSR architecture

3. Control Plane Architecture for CSR

This section describes the Control Plane architecture of CSR. The Control Plane primarily includes Routing Protocols and Route Management, as shown in Figure 2.

Routing Protocols SHOULD be used for exchanging Service Prefixes and topology information, divided into Static Routing Protocol and Dynamic Routing Protocol. How to exchange Service Prefixes and topology information through routing protocols is beyond the scope of this document.

Route Management SHOULD be used to collect, integrate, and optimize Service Routing Information from Routing Protocols and distribute effective Service Routing Information to the Data Plane.

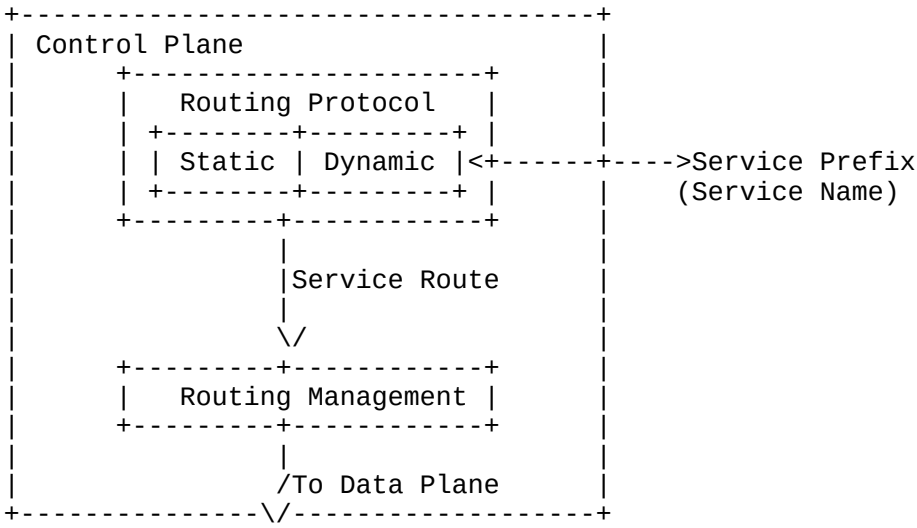


Figure 2: The Control Plane architecture

4. Data Plane Architecture for CSR

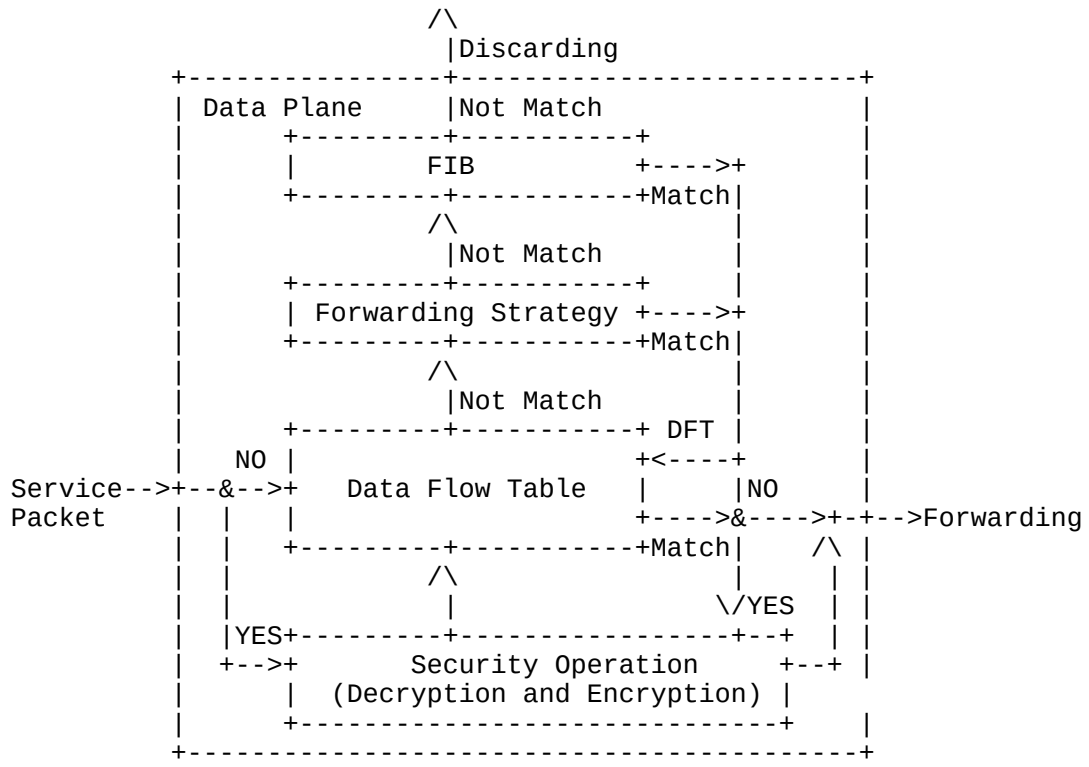
This section describes the Data Plane architecture of CSR, which mainly includes the Security Operation, the Data Flow Table (DFT), Forwarding Strategy (FS), and Forwarding Information Base (FIB), as shown in Figure 3.

The Security Operation SHOULD be used for decryption and encryption of data. The service packets into the Security Operation is determined by a Security Switch. In Service Mesh network, the Service packets are encrypted to protect user privacy. If the operator needs the CSR with traffic control and observability, the Service packets need to be decrypted by the CSR. And for achieving zero-trust network, the decrypted service packets need to be encrypted again before the service packets are forwarded. If enabling the Security Switch allows service packets to enter Security Operation, the forwarding performance of service packets will inevitably be affected. Therefore, it is recommended not to enable the Security Switch unless necessary.

The DFT SHOULD be used to maintain traffic stickiness, ensuring the output orientation remains consistent when the source and destination information are the same. The DFT also records and monitors the forwarding information of data packets, which is very useful for locating communication anomalies. The DFT SHOULD consist of the source service name, destination service name, and output interface information.

The FS SHOULD be used to meet special forwarding needs, having higher precedence than the FIB. The FS can specify the forwarding of data packages based on the configured output orientation when matching the source service name or destination service name.

The FIB SHOULD be used for regular guidance in forwarding data packages, mainly consisting of service names (prefixes) and output interfaces (next hops).



&: Security Switch to decide whether to enter Security Operation

Figure 3: The Data Plane architecture

Upon receiving a Service Packet, first check whether the Security Switch is open or not. If the Security Switch is not open, directly send service packet to the DFT. Else send service packet to the Security Operation. After the service packet is decrypted in Security Operation, it is also sent to the DFT.

In the DFT, if there’s a match, forward directly to the Security Switch. If not matched, send the data to check the FS. If it matches a FS, generate flow table information to DFT and forward to the Security Switch. If there is no match with a FS, send the data to check the FIB. If it matches the FIB, generate flow table information to DFT, and forward to the Security Switch. If there’s no match in the FIB, discard it.

When matched, the service packet is finally sent to the Security Switch. If the Security Switch is open, this means the service packet is not encrypted, and should be sent to the Security Operation for encryption. If the Security Switch is close, or the service packet is encrypted by the Security Operation, the service packet can then be forwarded.

5. Security Considerations

Security issues are not discussed in this document.

6. IANA Considerations

TBD.

7. Summary

The basic architecture of a content-based service router is described. The operation principles of the control plane and data plane of the service router are explained. Further investigation and discussion SHALL be necessary to execute the design of control protocols which MAY depend on the requirements by users.

8. References

8.1. Normative References

[I-D.li-dmsc-architecture]
Li, X., Wang, A., Wang, W., and D. KUTSCHER, "Distributed Micro Service Communication architecture based on Content Semantic", Work in Progress, Internet-Draft, draft-li-dmsc-architecture-00, 2 January 2025, <https://datatracker.ietf.org/doc/html/draft-li-dmsc-architecture-00>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

8.2. Informative References

[RFC8569] Mosko, M., Solis, I., and C. Wood, "Content-Centric Networking (CCNx) Semantics", RFC 8569, DOI 10.17487/RFC8569, July 2019, <https://www.rfc-editor.org/info/rfc8569>.

Authors' Addresses

Changwang Lin
New H3C Technologies
Beijing
China
Email: linchangwang.04414@h3c.com

Wei Wang
China Telecom
Beijing
Beiqijia Town, Changping District, 102209
China
Email: weiwang94@foxmail.com

Xueting Li
China Telecom
Beijing
Beiqijia Town, Changping District, 102209
China
Email: lixt2@foxmail.com

Haiyang Zhang
New H3C Technologies
Beijing
China
Email: zhang.haiyangA@h3c.com