

# 基于SRv6的可编排计算优先网络实现方法

## Implementation Method of Arrangeable Compute First Networking Based on SRv6

曹云飞,霍龙社,何涛(中国联通研究院,北京 100048)

Cao Yunfei, Huo Longshe, He Tao (China Unicom Research Institute, Beijing 100048, China)

### 摘要:

为了解决计算优先网络(CFN)缺乏算力编排能力的问题,基于CFN和SRv6的技术特点,提出了一种基于SRv6的可编排CFN实现方法。通过将多个服务编排成为服务功能链(SFC)来实现复合业务,从而提高CFN的业务创新能力。基于该方法设计并实现了一个应用实施例,体现了其可以在网络层解耦不同服务之间依赖关系的优点。

### 关键词:

算力网络;CFN;SRv6;IPv6+

doi: 10.12045/j.issn.1007-3043.2022.04.002

文章编号: 1007-3043(2022)04-0004-06

中图分类号: TN919

文献标识码: A

开放科学(资源服务)标识码(OSID):



### Abstract:

In order to solve the problem that compute first networking (CFN) lacks the ability of computing power arrangement, based on the technical characteristics of CFN and srv6, the implementation method of arrangeable CFN based on SRv6 is proposed. Multiple services are arranged into service function chain (SFC) to realize composite services, so as to improve the business innovation ability of CFN. Based on this method, an application example is designed and implemented, which reflects the advantage that it can decouple the dependencies between different services at the network layer.

### Keywords:

Computing power network; CFN; SRv6; IPv6+

引用格式: 曹云飞,霍龙社,何涛. 基于SRv6的可编排计算优先网络实现方法[J]. 邮电设计技术, 2022(4): 4-9.

## 0 引言

随着5G与人工智能的发展,车联网、远程医疗、智能工厂等新兴产业不断涌现,越来越多的数据需要就近处理和分析。为了满足这些需求,越来越多的计算下沉到接近数据源的网络边缘。随着边缘计算的诞生,多级计算节点遍布网络。通过网络将算力连接起来并进行统一管理和调度成为算力网络的目标。

算力网络是一种根据业务需求,在云、网、边之间按需分配和灵活调度计算资源、存储资源以及网络资

源的新型信息基础设施<sup>[1]</sup>。随着算力网络的关注度持续升高,该技术领域产生了不同的技术类型<sup>[2]</sup>。IRTF成立了在网计算研究组(COINRG)<sup>[3-6]</sup>,在网计算指网络设备的功能是“转发+计算”。算力网络领域的计算优先网络(Compute First Networking, CFN)以及基于SRv6和IPv6+的算力网络均属于在网计算。CFN将当前的计算能力状况和网络状况作为路由信息发布到网络,根据算力状况以及网络状况计算和执行路由。SRv6通过源路由实现报文转发路径的可编程。CFN为基于算力的路由提供了解决方案,但是目前还没有算力编排的能力。为了在网络层面实现算力编排,为CFN提供算力编排能力,本文提出了一种新的基于

收稿日期: 2022-03-07

SRv6的可编排计算优先网络实现方法。

## 1 研究背景

### 1.1 CFN

CFN将服务节点的计算能力状况和网络状况作为路由信息发布到网络,网络基于虚拟的服务ID将计算任务报文路由到最合适的计算节点,实现最优的用户体验、计算资源利用率以及网络效率<sup>[7]</sup>。CFN能够为应用提供的核心价值是基于算力和算法动态调度计算资源,保证业务的用户体验。

#### 1.1.1 CFN概述

如图1所示,CFN网络按角色分为服务器端、CFN域和客户端<sup>[3]</sup>。图1在参考文献[3]基础上进行了改动:边缘站点改为站点,服务节点增加了资源接口和服务接口,CFN层改为CFN域。由于算力网络不仅包括边缘站点,未来也会包括数据中心和端。算力服务提供者不仅限于边缘站点,所以本文将边缘站点改为站点。由于服务节点通过资源接口向CFN路由器提供资源信息,通过服务接口提供对外计算服务,所以服务节点增加了资源接口和服务接口。由于多个CFN路由器组成网状拓扑结构,而不是两两相连的一层,所以本文将CFN层改为CFN域。

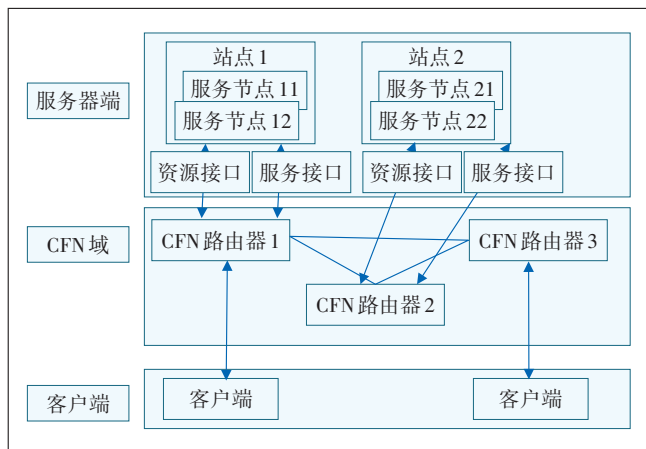


图1 CFN网络拓扑

#### 1.1.2 CFN控制面原理

如图2所示,CFN控制面完成算力信息的全网同步。站点的服务节点将本地服务状态注册到CFN路由器的本地注册表。本地服务状态包括Service ID——服务的唯一标识、服务IP地址和计算资源情况等。CFN路由器将本地服务状态基于CFN路由协议加载到报文中,将服务状态扩散到其他CFN路由器。

CFN路由器基于CFN路由协议将本地以及其他CFN路由器扩散的服务状态信息汇总生成服务信息路由表。

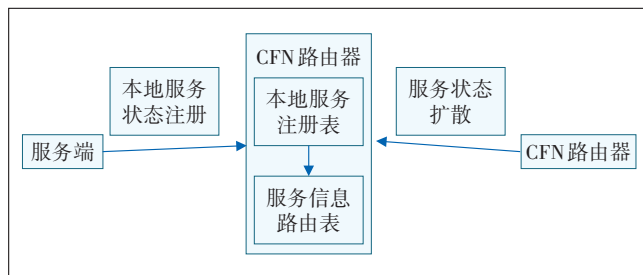


图2 CFN控制面原理

#### 1.1.3 CFN数据面原理

CFN数据面完成客户端对服务节点Service ID请求的路由转发。CFN Ingress节点收到请求后,根据网络情况、计算资源情况综合评估,选择一个服务节点以及相关联的CFN Egress节点,将原请求数据包封装并发送。CFN Egress节点收到数据包,根据Service ID查找对应服务节点IP,将数据封装并发送。外层数据包源地址为客户端IP,目的地址为服务节点IP。内层数据包源地址为客户端IP,目的地址为Service ID。服务节点收到数据包后在本地查询与Service ID绑定的服务地址,调用对应的服务,将结果返回给客户端。

### 1.2 SRv6

SR(Segment Routing)是源路由技术的一种,节点选择路径,并且引导数据包沿着该路径通过网络,其做法是在数据包头部插入带顺序的segment列表,以指示接收到这些数据包的节点怎么去处理和转发这些数据包。为了实现SRv6转发,需要向IPv6报文中插入一个段路由头(Segment Routing Header, SRH)的扩展头,存储IPv6的Segment List信息。报文转发时,依靠Segments Left和Segment List字段共同决定IPv6目的地址(IPv6 DA)信息,从而指导报文的转发路径和行为。

128位SRv6 Service ID主要由标识节点位置的LOC字段(IPv6前缀格式,可路由)、标识服务和功能的FUNC字段(本地识别)以及存储相关参数的ARG字段3部分组成。一个标准的SRv6 Service ID可以定义特定节点的路径信息、服务和功能信息。

SRv6网络编程标准中,SRv6节点(Endpoint)通过本地定义的行为(Behavior)处理SRv6报文。SRv6定义了多种Endpoint Behavior,每个节点需要实例化它

们并分配 Service ID,同时通过路由协议公布,以便外部了解节点所能提供的 Behavior。常用的 Endpoint Behavior 有 END、END.X、END.DT4、END.DT6 等,实现 Underlay 选路、Overlay 业务承载等功能。

## 2 基于SRv6实现可编排计算优先网络

CFN 借助路由协议将获取的计算资源信息发布给相邻的 CFN 路由器,实现计算资源信息的全网扩散。SRv6 通过在 IPv6 报文中新增 SRH 报头来控制报文的转发路径和行为,从而提供网络可编程能力。通过 SID(segment ID)中的 Function 字段、Arguments 字段以及 Segment 序列之后的 Optional TLV (Type-Length-Value) 提供可扩展能力。

SRv6 的这些优点为算力编排以及参数传递提供了技术基础。CFN 为基于算力的路由提供了解决方案,但是没有提供算力编排的能力。为了在网络层面实现算力编排,为 CFN 提供算力编排能力,本文提出了一种新的基于 SRv6 的可编排计算优先网络实现方法。基于 SRv6 技术来实现 CFN 数据传输,实现报文转发路径的可编程,从而在网络层面实现算力编排。

为了基于 SRv6 技术实现 CFN 数据传输,所有的 CFN 路由器需要支持 SRv6。CFN 路由器需要支持 SRv6 报文的识别、SRv6 报文的处理以及 SRv6 报文的转发。

IS-IS 属于链路状态 IGP 路由协议,采用 TLV 携带链路状态信息,可以在不动协议主体的情况下,通过增加部分 TLV 类型就可以支持 IPv6、SRv6 等新协议特性。在 CFN 域中通过 IS-IS 协议来扩散 CFN 路由器的 SID 信息。支持 SRv6 的服务节点也需要通过 IS-IS 协议扩散自己的 SID 信息。

### 2.1 客户端、服务节点均为 IPv6 设备

随着 IPv6 技术的推广,越来越多的设备支持 IPv6。如果客户端和服务节点均支持 IPv6,而且服务节点支持 SRv6,则 CFN Ingress 节点作为 SRv6 源节点分类器(Service Classifier, SC),服务节点作为 Endpoint 节点(SRv6 Segment Endpoint Node),服务节点将 SRv6 的 SID 作为 Service ID。服务节点定义本地行为处理 SRv6 报文,根据 Service ID 调用服务。在客户端和服务节点均支持 IPv6 的情况下,本技术方案基于 SRv6 技术实现 CFN 数据传输,完成客户端对服务节点 Service ID 请求的路由转发。客户端和服务节点均支持 IPv6,而且服务节点支持 SRv6 的情况不需要 CFN Egress 节

点参与算力路由,具体网络拓扑如图 3 所示。

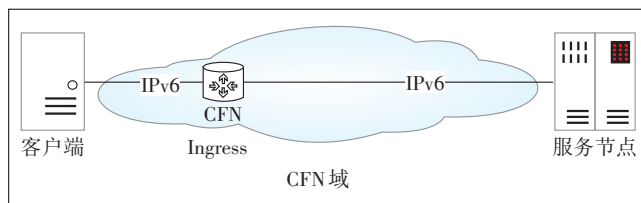


图3 客户端、服务节点均为IPv6设备的网络拓扑

客户端请求 Service ID 服务,源地址为客户端地址,目的地址为 Service ID。CFN Ingress 节点收到请求后,查找服务信息路由表,根据 Service ID 查找服务节点的 Service IP。CFN Ingress 节点作为 SRv6 源节点 SC 封装 SRv6 数据,数据源地址为客户端 IP,目的地址为 Service IP。在数据包头中插入 SRH (Segments Left=1),SRH 中的 Segment List 为 <Segment List[0]=Service ID,Segment List[1]=Service IP>。从 CFN Ingress 节点开始数据传输的第 1 跳为服务节点 Service IP,即服务节点的 IPv6 地址。第 2 跳为服务节点的 Service ID。

服务节点收到数据包,Segments Left 减 1,Segments Left=0,取出 Segment List[0] 对应的值 Service ID。服务节点在本地查询与 Service ID 绑定的计算服务,调用服务并返回结果。返回数据的数据源地址为 Service ID,目的地址为客户端 IP。

Service ID 可以是服务节点的某个端口或者某个 Loopback IP。服务节点调用 Service ID 对应的端口或者 Loopback IP。如果服务节点位于 gateway 的位置,则 Service ID 对应于内网某个服务器的内网 IP。服务节点调用该内网 IP 对应的服务。服务节点位于 gateway 位置的网络拓扑如图 4 所示。

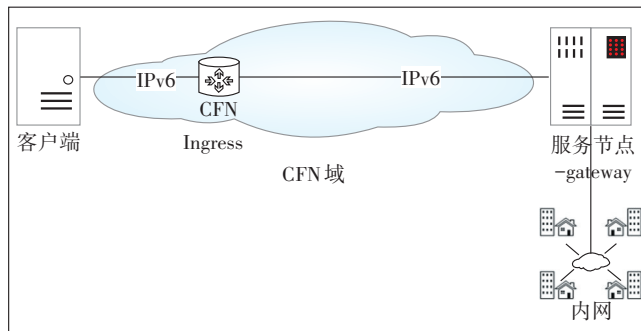


图4 服务节点为gateway的网络拓扑

客户端请求服务的数据传输过程如图 5 所示。

### 2.2 客户端、服务节点均为 IPv4 设备

如果客户端和服务节点均为 IPv4 设备,则 CFN In-

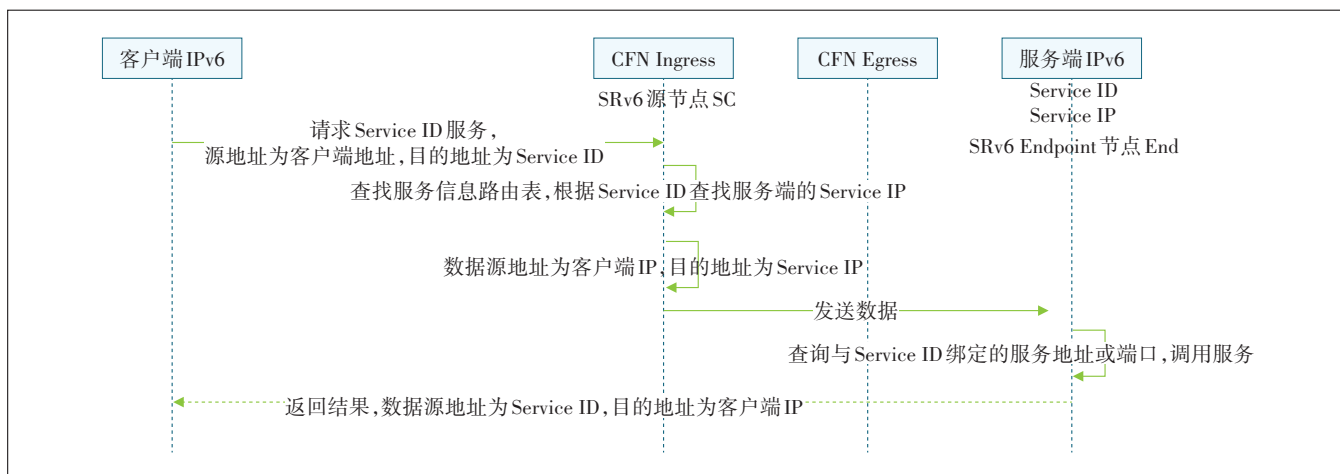


图5 客户端、服务节点均为IPv6设备时的数据流

gress节点作为SRv6源节点SC,CFN Egress节点作为Endpoint节点。CFN Egress节点定义本地行为处理SRv6报文,解封封装SRv6数据。CFN Egress节点作为服务节点的代理节点,将IPv6报文转换为IPv4报文并调用服务节点的服务。CFN Ingress节点和CFN Egress节点作为IPv6和IPv4协议转换节点需要支持IPv6和IPv4双栈。双协议栈技术就是指在1台设备上同时启用IPv4协议栈和IPv6协议栈。CFN Ingress节点和CFN Egress节点的不同接口上,分别配置了IPv4地址和IPv6地址,并分别连接了IPv4网络和IPv6网络。CFN域内为IPv6网络,CFN域外为IPv4网络,具体网络拓扑如图6所示。

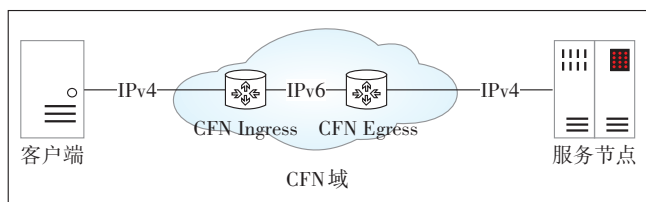


图6 客户端、服务节点均为IPv4设备时的网络拓扑

在客户端和服务节点均支持IPv4的情况下,本技术方案基于SRv6技术实现CFN数据传输,完成客户端对服务节点Service ID请求的路由转发。客户端发送的数据由CFN Ingress节点封装为SRv6报文,数据源地址为CFN Ingress IPv6地址,目的地址为CFN Egress IPv6地址。CFN Ingress节点生成SRH:(Segments Left=0);<Segment List[0]=Egress IPv6地址>。内层数据为IPv4数据,源地址为客户端IP,目的地址为Service ID。数据传输的第1跳为CFN Egress节点的IPv6地址。CFN Egress节点收到数据报文后,解封封装SRv6数

据,取出内层数据,查找本地服务注册表,根据Service ID查找对应Service IP。CFN Egress节点将内层数据进行封装并发送。封装数据报文的外层源地址为CFN Egress IPv4地址,目的地址为Service IP。内层数据报文源地址为客户端IP,目的地址为Service ID。服务节点收到数据报文后,解封封装数据报文,取出内层数据,在本地查询与Service ID绑定的服务,调用服务并返回结果。返回数据报文的数据源地址为Service ID,目的地址为客户端IP。数据传输过程如图7所示。

对于客户端和服务节点支持IPv6的情况以及客户端和服务节点支持IPv4的情况,本文详细说明了基于SRv6的可编排计算优先网络实现方法。

### 3 应用实施例

为了验证本文所提方法的实际效果,特设计并实现了下述应用实施例,对于视频服务提供附加的鉴别服务。

本应用实施例包含2个服务节点,A提供视频服务,B提供鉴别服务。客户端向A服务节点请求视频服务,B服务节点对客户端所请求的视频内容进行鉴别。如果视频内容符合规定,则视频正常返回给客户端。如果视频内容不符合规定,则返回错误信息给客户端。客户端只知道视频服务即A节点的Service1 ID,不需要知道鉴别服务即B节点的Service2 ID。CFN Ingress节点、CFN Egress1节点、CFN Egress2节点、视频服务节点A和鉴别服务节点B组成了功能业务链SFC。CFN Ingress作为分类器SC位于SFC域的边界入口,报文进入SFC域后会首先进行流量分类。视频服务节点A和鉴别服务节点B作为服务功能节点

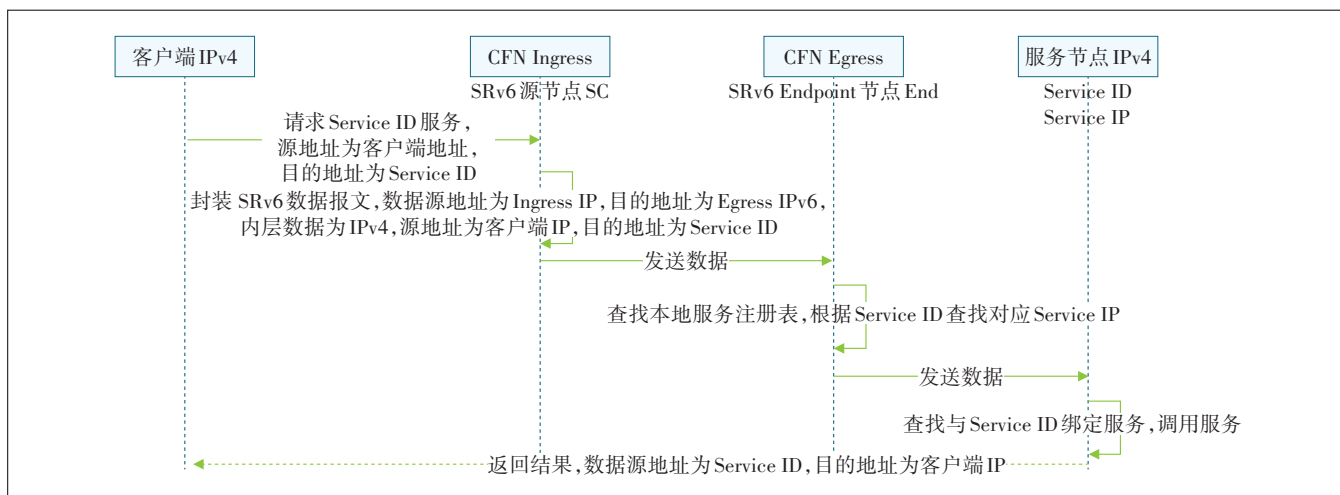


图7 客户端、服务节点均为IPv4设备时的数据流

SF(Service Function)提供增值服务功能。由于SF本身不支持处理SRv6报文,所以由CFN Egress1节点和CFN Egress2节点作为业务链转发节点SFF(Service Function Forwarder)代表SF处理SRv6报文,删除SRv6封装信息,通过本地逻辑组件将报文发给SF,接收从SF发回的报文,重新为之添加SRv6封装信息。本应用实施例网络拓扑如图8所示。

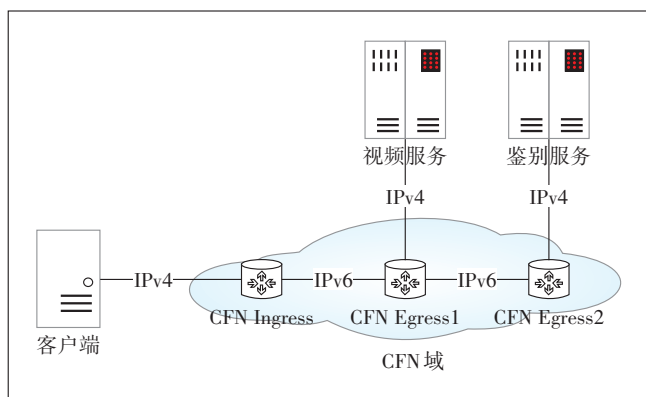


图8 应用实施例网络拓扑

客户端请求视频服务的流程如下:客户端请求视频服务即Service1 ID服务,CFN Ingress节点收到请求,查找服务信息路由表,根据Service1 ID查找服务端对应的Egress节点Egress1。根据事先编排的业务逻辑,CFN Ingress查找Service2 ID对应的Egress节点Egress2。CFN Ingress节点封装SRv6数据报文,数据源地址为Ingress IP,目的地址为Egress1 IPv6地址,生成SRH:(Segments Left=1):<Segment List[0]=Egress2 IPv6,Segment List[1]=Egress1 IPv6>。内层数据为

IPv4数据,源地址为客户端IP,目的地址为Service1 ID。从CFN Ingress节点开始数据传输的第1跳为Egress1 IPv6地址,第2跳为Egress2 IPv6地址。CFN Egress1节点解封装SRv6数据报文,根据SID在本地定义的动作,取出内层数据,查找本地服务注册表,根据Service1 ID查找对应Service1 IP,封装数据并发送。外层数据源地址为Egress1 IPv4地址,目的地址为Service1 IP,内层数据源地址为客户端IP,目的地址为Service1 ID。视频服务节点收到数据解封装数据报文,查询与Service1 ID绑定服务,调用视频服务并返回结果。数据源地址为Service1 ID,目的地址为Egress1 IPv4地址。

CFN Egress1节点将Segments Left减1,则Segments Left=0,将Egress2 IPv6地址拷贝到目的地址。Egress1封装数据,外层数据源地址为Egress1 IPv6地址,目的地址为Egress2 IPv6地址。内层数据源地址为客户端地址,目的地址为Service1 ID。CFN Egress2节点解封装SRv6数据报文,取出内层数据得到Service1 ID,根据事先编排的业务逻辑由Service1 ID查找到Service2 ID,查找本地服务注册表,根据Service2 ID查找对应Service2 IP。CFN Egress2节点封装数据并发送。封装的数据报文外层数据源地址为Egress2 IPv4地址,目的地址为Service2 IP。内层数据源地址为客户端IP,目的地址为Service1 ID。鉴别服务收到数据后调用鉴别服务并返回结果。返回的数据报文外层数据源地址为Service2 ID,目的地址为Egress2 IPv4地址。内层数据源地址为客户端IP,目的地址为Service1 ID。CFN Egress2节点收到鉴别服务返回的数据

后,取出内层数据并返回给客户端。数据源地址为 Service1 ID,目的地址为客户端地址。数据传输过程

如图9所示。

本应用实施例为用户提供了视频服务和附加的

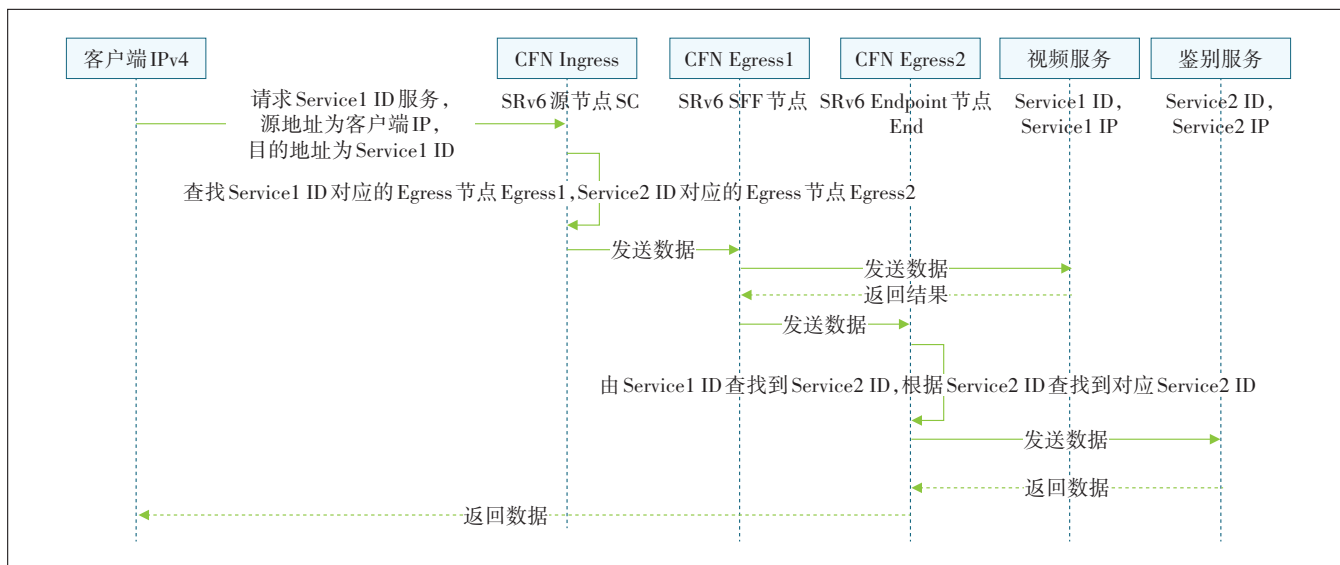


图9 应用实施例数据流

鉴别服务,对客户端暴露的只有视频服务(节点A),鉴别服务(节点B)对于客户端是隐藏的。基于本技术方案的应用实施例在网络层实现了业务编排,在视频服务基础上增加了鉴别服务,该方案不需要修改视频服务的业务逻辑。如果不在网络层做业务编排,在应用层做业务编排需要改变视频服务现有业务逻辑。视频服务在处理用户请求后,不是将结果返回给客户端,而是将结果发送给鉴别服务,最后将鉴别结果返回客户端。本应用实施例在网络层解耦了视频服务和鉴别服务,从而消除了视频服务和鉴别服务之间调用的工作量。本应用实施例不需要客户端感知鉴别服务,从而减少了客户端的业务复杂度。

#### 4 结束语

CFN可以实现用户的就近接入和服务的负载均衡,适应服务的动态性,但目前缺乏算力编排能力。本文提出了一个新的基于SRv6的可编排计算优先网络,使用SRv6实现CFN数据传输,通过实现报文转发路径可编程,在网络层面实现算力编排,将多个服务编排成为SFC来实现复合业务,从而提高了CFN的业务创新能力。基于该方法,本文设计并实现了一个可编排计算优先网络的应用实施例,可在网络层解耦2个在业务逻辑上有依赖关系的服务。2个服务不需要感知到对方,消除了2个服务之间调用的工作量,减少

了客户端的业务复杂度。

#### 参考文献:

- [1] 雷波,刘增义,王旭亮,等. 基于云、网、边融合的边缘计算新方案: 算力网络[J]. 电信科学,2019,35(9):44-51.
- [2] 中国通信标准化协会. 面向全网算力的算力感知网络关键技术研究[R]. 2020.
- [3] IETF. Framework of compute first networking(CFN): draft-li-rtgwg-cfn-framework-00[DB/OL]. [2021-11-04]. <https://datatracker.ietf.org/doc/html/draft-li-rtgwg-cfn-framework-00>.
- [4] IETF. A report on compute first networking (CFN) field trial: draft-gu-rtgwg-cfn-field-trial-00[DB/OL]. [2021-11-04]. <https://datatracker.ietf.org/doc/html/draft-gu-rtgwg-cfn-field-trial-00>.
- [5] SAPIO A, ABDELAZIZ I, ALDILAIJAN A, et al. In-network computation is a dumb idea whose time has come [C]//Proceedings of the 16th ACM Workshop on Hot Topics in Networks. New York: Association for Computing Machinery, 2017: 150-156.
- [6] IETF. Segment routing architecture: RFC 8402[S/OL]. [2022-01-01]. <https://datatracker.ietf.org/doc/html/rfc8402>.
- [7] 网络5.0产业联盟. 网络5.0产业联盟CFN特设组倡议与筹备汇报[Z]. 2019.
- [8] 何涛,曹畅,唐雄燕,等. 面向6G需求的算力网络技术[J]. 移动通信,2020,44(6):131-135.

#### 作者简介:

曹云飞,毕业于西安交通大学,硕士,主要研究方向为SDN、云计算、新一代网络编排技术等;霍龙社,教授级高级工程师,博士,主要从事软件研发工作;何涛,毕业于北京邮电大学,高级工程师,硕士,主要从事云化网络及数据通信网络相关技术研究工作。