

团 体 标 准

T/CCSA 521—2024

网络 5.0 骨干网流量整形控制协议

Traffic shaping control protocol for backbone of network 5.0

2024 - 04 - 01 发布

2024 - 06 - 01 实施

中国通信标准化协会 发布

版权声明

本技术文件的版权属于中国通信标准化协会，任何单位和个人未经许可，不得进行技术文件的纸质和电子等任何形式的复制、印刷、出版、翻译、传播、发行、合订和宣贯等，也不得引用其具体内容编制本协会以外各类标准和技术文件。如果有以上需要请与本协会联系。

邮箱：IPR@ccsa.org.cn

电话：62302847

目 次

前言	II
引言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	2
5 概述	2
5.1 总体概述	2
5.2 流量整形整体框架	2
6 骨干网流量整形控制协议	3
6.1 流量整形控制总体流程	3
6.2 流量整形参考报文	4
6.3 流量整形参数配置	5
6.4 流量整形流表	5
6.5 数据面流量控制	6
6.6 流量需求描述	7
6.7 节点状态请求消息格式	7
6.8 节点状态请求响应消息格式	8
6.9 流量测量报告	8
6.10 流量时隙映射	8

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件由中国通信标准化协会提出并归口。

请注意本文件的某些内容可能涉及专利，本文件的发布机构不承担识别这些专利的责任。

本文件起草单位：华中科技大学、中国信息通信研究院、江苏未来网络集团有限公司、中国移动通信有限公司、中国电信股份有限公司、华为技术有限公司、中兴通讯股份有限公司、腾讯、北京邮电大学、大连理工大学等。

本文件主要起草人：莫益军、聂秀英、魏亮、姚惠娟、雷波、周旭、郑秀丽、刘爱华、蒋胜、刘辰宇、李雪莹等。

本文件为首次发布。

引 言

为适应信息通信业发展对标准文件的需求，由中国通信标准化协会组织制定“中国通信标准化协会团体标准”，推荐有关方面采用。有关对本文件的建议和意见，向中国通信标准化协会反映。

面向算力与网络融合、产业互联网、移动业务承载、未来视频业务承载及全息通信等应用场景，网络5.0已经分析了业务指标和流量分类，明确了路由、时间和资源的确定性作为网络5.0网络确定性主要范围。流量整形作为网络确定性技术的一种，其通过为高优先级流量提供确定的传输时隙来保障确定的传输时延，避免流量突发造成重传和丢包。但目前网络中对于网络边缘处以及骨干网处的流量整形通信协议尚未规范。

为了针对网络5.0中大规模骨干网开启流量整形时的控制通信需求，对流量整形时的控制通信协议进行规范，在UNI处进行接纳控制、突发整流等的控制通信协议规范，在NNI处进行流量感知、突发通知等的控制通信协议规范，提供骨干网细粒度流量整形流程范式，以满足网络5.0对网络确定性的要求，提升网络5.0针对确定性时延、抖动等网络指标的保障能力。

网络 5.0 骨干网流量整形控制协议

1 范围

本文件规范网络5.0骨干网流量整形的协议与流程，规定了控制面和数据面间的流量整形参数和控制策略相关的协议流程和报文字段。本文件仅覆盖已通过的标准中包含的完成的工作。当通过新标准时，在本文件的下一版本中将增加支持新标准的内容。

本文件适用于网络5.0骨干网流量整形。

2 规范性引用文件

本文件没有规范性引用文件。

3 术语和定义

GB/T 20984-2007界定的术语和定义适用于本文件。

3.1

网络5.0 network 5.0 (Net 5)

网络5.0是第5代数据网络的统称，该网络面向满足2030年的应用场景（典型业务包括算力网络（也称ICT融合网络）、产业互联网、5G核心网络、全息通信）需求，基于现有IP网络协议基础，提供确定性、内生安全、可信、有限域（Limited Domain）四方面网络能力。

注：在本文件当前版本中，第1代至第4代数据网分别指公用交换电话网络（PSTN，其上加调制解调器）、公用分组交换数据网（PSPDN）、宽带综合业务数字网络（B-ISDN，ATM）、互联网络（IPv4和IPv6）。

3.2

流量整形 traffic shaping

流量整形属于网络流量工程机制，针对网络流量突发或网络流量过载情况，由网络中路由器交换机对流量进行排队，以控制网络流量传输速率，保证网络不拥塞，满足特定应用的带宽和时延需求，管理和平衡网络中流量，避免网络拥塞和资源过度竞争。

3.3

信用整形 credit-based shaping

信用整形是一种基于发送方和接收方之间的信用机制，用于限制发送方的传输速率。发送方在发送数据之前必须获取足够的信用，而接收方根据其可用的接收能力来授予信用。这种整形方法可以避免网络拥塞和数据丢失。

3.4

异步流整形 asynchronous traffic shaping

异步流整形是一种用于管理异步流量的整形方法。它通过限制发送方的传输速率来平滑流量，并确保不会超过接收方的处理能力。这种方法主要关注缓解网络拥塞和保证数据可靠性。

3.5

时间感知整形 time-aware shaping

时间感知整形是一种基于时间的整形方法，可以确保实时应用程序的数据在特定时间内到达接收方。它考虑到时序要求，将数据流整形为一系列符合脉冲规则的数据包。用于音视频通信、工业自动化等对实时性要求较高的场景。

3.6

循环排队转发 cycle queueing forward

循环排队转发是一种用于确定性网络转发机制。在这种机制中，多个接收方按照预定的顺序排列在循环队列中，每个接收方在其轮到时接收数据。这种转发方法可降低传输带宽需求，减少冗余。

3.7

流表 flow table

流表是实现流量控制和转发决策的跨层协议转发表,它可以根据数据包的特征,如源/目标IP地址、端口号、协议类型等,将数据包分类并采取相应的处理措施。

4 缩略语

下列缩略语适用于本文件。

CBS: 基于信用整形(credit based shaping)

CNC: 中央网络控制器(central network controller)

CQF: 循环排队转发(cyclic queuing and forward)

CSQF: 指定周期队列转发(cycle specified queuing and forward)

SID: 分段路由标识符(segment routing identifier)

TAS: 时间感知整形(time awareness shaping)

TSN: 时延敏感网络(time-sensitive networking)

5 概述

5.1 总体概述

网络5.0是全新的第五代数据网络,基于“以网络为中心、能力内生、全维可定义”的核心理念,实现对可信安全、确定性服务、资源感知管控等的内生支持。面向算力与网络融合、产业互联网、移动业务承载及全息通信等应用场景,网络5.0已经分析了业务指标和流量分类,明确了路由、时间和资源的确定性作为网络5.0网络确定性主要范围。

网络流量整形属于网络确定性技术的一种,为了实现确定延迟,TSN利用流量整形机制通过为高优先级流量提供确定的传输时隙来提供确定的传输时延,避免突发流量造成重传和丢包。目前,面向网络流量整形,业界在逐流过滤整形和队列转发控制方面进行了大量研究,实现了小规模以太网环境下轻载稳定流的微秒级时延和亚微秒级抖动保障。业界已提出的流量整形技术包括基于信用整形(CBS)、时间感知整形(TAS)、异步流量整形(ATS)和循环排队转发(CQF)等,其中TAS提供精确的确定性时延保障,具有最佳的低时延和低抖动性能,却具有最高的配置复杂性;CBS利用信用指标将传输时间分为高优先级流量和普通优先级流量两个传输时隙,时延性能最差,但却保证了流的QoS;ATS通过在每一跳重塑 TSN 流,提供无需严格时间同步的确定性时延,其时延性能介于两者之间,配置复杂性最低;CQF可以通过同步入口和出口的队列操作以降低TAS 配置的复杂性,实现与网络拓扑无关的零拥塞丢失和有界延迟,但对时间同步有较高要求。

对于大规模骨干网络,在面对长距离网络拓扑变化、传输时延波动和短时突发流量过载等需要进行流量整形的场景时,尚没有明确定义用户网络接口(UNI)以及网络网络接口(NNI)处使用的通信协议标准。本骨干网流量整形控制协议针对网络5.0中大规模骨干网开启流量整形时的控制通信需求,对流量整形时的控制通信协议进行规范,规定了控制面和数据面间的流量整形参数和控制策略相关的协议流程和报文字段,规定了数据面UNI处的接纳控制、突发流整形等的控制通信协议规范,在NNI处进行流量感知、突发通知等的控制通信协议规范,基于控制协议和交互报文,给出数据面交换机和路由的流量整形实现机制。本文提供骨干网细粒度时隙级流量整形流程,以满足网络5.0对网络确定性的要求,提升网络5.0针对确定性时延、抖动等网络指标的保障能力。

5.2 流量整形整体框架

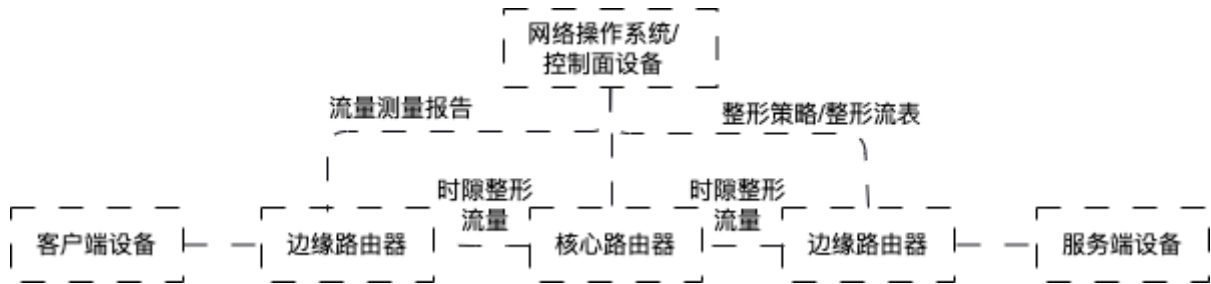


图1 流量整形框架

流量整形框架由控制面控制器、数据面边缘路由器和骨干路由器组成。端设备自身并不具备流量整形能力，端设备仅在业务流发起时，以主动或被动方式向网络声明其时延和抖动相关的服务质量等级需求。

当边缘路由器加入网络时，向网络控制器注册其流量整形能力。系统运行时，各路由器网络设备监测网络状态信息，在发生网络突发现象时通知上游设备网络突发产生，向网络控制器上报各设备采集的流量测量报告，使其开启流量整形减缓流量发送速率。开启时钟同步流量整形的相邻设备，获取交换相邻设备间队列模型（如CQF和CSQF模型）参数，包括队列数量、周期时隙大小、输出端口带宽、传输延迟等信息。以便及时动态调整流量整形参数。

网络控制器通过API接口发现并连接网络设备。获取当前网络域中的网络拓扑、链路容量和端口传输速率信息。获取注入网络域的时间敏感流量信息，包括端到端延迟界限要求、数据包发送频率、数据包大小和数量以及源/目的地地址。网络控制器综合考虑链路容量、端到端延迟边界等要求，结合设备流量整形能力和网络规模生成流量整形策略和配置参数，下发给边缘接入。根据各设备上报的流量测量报告，尤其是针对时延抖动业务流量指标偏离确定性需求，调整优化整形流表。

6 骨干网流量整形控制协议

6.1 流量整形控制总体流程

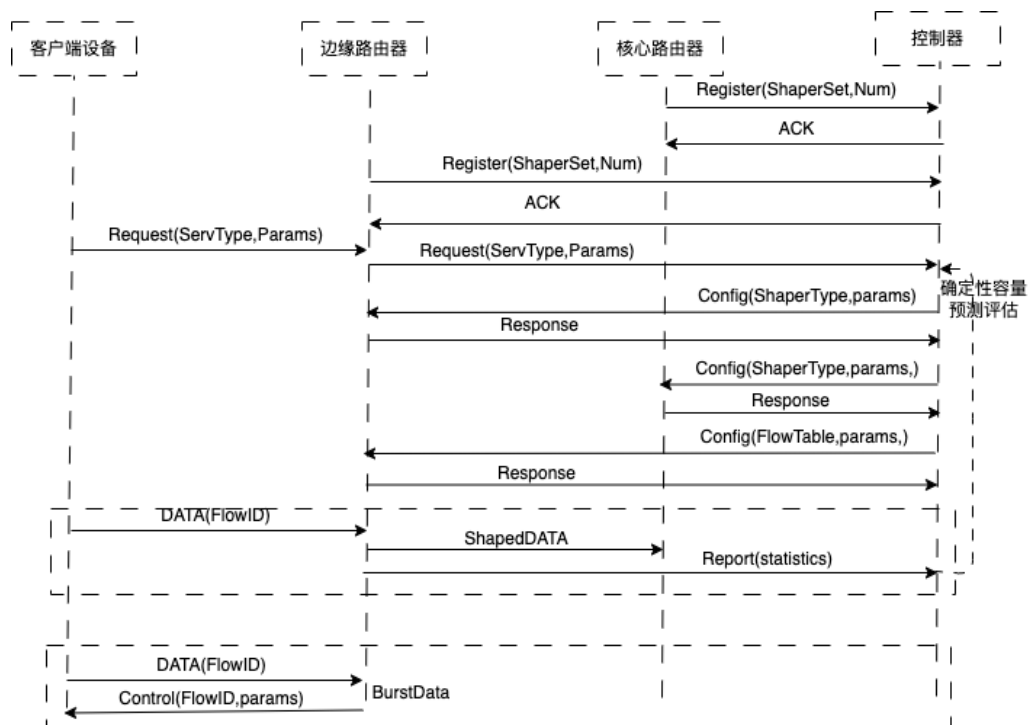


图2 流量整形控制总体流程示意图

流量整形控制总体上包括能力注册、整形配置、转发整形和整形调优四个阶段。

a) 能力注册阶段

数据面设备入网时，向控制面设备或网络控制器发起注册请求，完成入网许可和整形能力上报。数据面设备可具备CBS、ATS、TAS、CQF和时隙流整形等多种整形能力，根据其资源能力，上报其支持整形队列数量、长度和时隙周期等参数。控制器接收到注册请求后存储整形能力库。但数据面设备整形能力集并不代表最终运时加载的整形队列。

b) 整形配置阶段

该阶段是控制面控制器根据数据面设备上报的整形能力及用户应用类型的统计分布，按需为各设备生成时隙周期、最迟发送时间、发送速率、报文大小、整形队列类型和队列长度等整形策略相关参数配置，并下发至数据面设备。数据面设备调用相关API进行配置，若配置成功则响应配置成果，若配置不成功则响应带有错误码的失败报告。控制器根据失败代码进行定位，并重新生成配置策略。

c) 转发整形阶段

该阶段控制面控制设备根据应用需求，针对应用流的源目的地址和端口设置相关的整形流表。数据面交换路由设备根据流表匹配的转发动作进行应用流转发，同时对确定性指标和流量指标进行策略，对能正常满足要求的进行整形队列处理，对于流量微突发超出整形能力范围的，则向源端设备或上游设备返回数据报文时携带整形资源余量，实现源端口流量抑制，以保证整体网络的服务质量。

d) 整形调优阶段

整形调优阶段则是当数据面流量整形测量报告上传至数据面控制器后，控制器根据各应用流的QoS指标偏离度、服务保障容量、时延抖动裕度调整数据面设备的整形参数，满足确定性应用接入能力需求。

6.2 流量整形参考报文

数据面设备入网时注册的流量整形流程如图1所示，为兼容网络管理需求，流量整形能力请求响应报文采用Json格式的YangModel模型，其参考报文如下：

注册报文：

Register /traffic_shaping_capabilities HTTP/1.1 //注册消息

Host: hostname //控制器名

Authorization: QWxhZGRpbjpvcmVudHNlc2FtZQ== //认证鉴权

Content-Type: application/json

```
{ "device_id": "12345", //设备标识
  "capabilities": //支持整形能力
  [
    { "ShaperType": "CBS", //整形器类型
      "rule": "limit_bandwidth", //整形规则
      "parameters": //整形参数
      { "bandwidth": "100 Mbps" }
    },
    { "ShaperType": "ATS", //整形器类型
      "rule": "rate_control", //整形规则
      "parameters": //整形参数
      { "rate": "10 Mbps",
        "Delay": "10ms",
        "Jitter": "100us"
      }
    }
  ]
}
```

响应报文：

HTTP/1.1 200 OK

Content-Type: application/json

```
{ "status": "success",
  "message": "Traffic shaping capabilities registered successfully." }
```

6.3 流量整形参数配置

控制平面控制器向数据平面发送各种整形器的配置参数，包括整形速率、流量限制、整形容量、优先级和队列长度等。不同整形器的配置参数根据其实现机理和控制参数各有不同，对于ATS来说，主要参数包括流量限制、优先级、队列长度等。其配置命令示例如下。

Config /shapercnf HTTP/1.1 //配置消息

```
{
  "ShaperType": "ATS", //整形器类型
  "parameters": //整形参数
  { "rate": "10 Mbps",
    "Delay": "10ms",
    "Jitter": "100us",
    "queue_length": 100,
    "rate_limit": 1000,
    "priority": 1
  }
}
```

响应报文:

HTTP/1.1 200 OK

```
{
  "status": "success",
  "message": "Configuration applied successfully."
}
```

6.4 流量整形流表

数据面流量整形最终实现需依赖于具体的流表，控制平面控制器向数据平面发送各种应用的流表转发控制策略，其流表下发报文采用OPENFLOW协议格式，如下例：

下发命令报文:

OFPT_FLOW_MOD (Flow Modification Message)

Header:

Version: 1.3

Type: FLOW_MOD

Transaction ID: 12345

Body:

Cookie: 0x12345678

Command: OFPFC_ADD

Idle Timeout: 0

Hard Timeout: 0

Priority: 100

Buffer ID: OFP_NO_BUFFER

Out Port: OFPP_ANY

Flags: 0

Match Fields:

- Ethernet Source Address: 00:11:22:33:44:55

- Ethernet Destination Address: 66:77:88:99:AA:BB

- IP Source Address: 192.168.0.1

- IP Destination Address: 192.168.0.2
- IP Protocol: TCP
- Instructions:
 - BandWidth Upbound: 100Mbps
 - BandWidth Bottombound: 10Mbps
 - Maxallowed Delay: 10ms
 - Maxallowd Jitter: 100us
- Action:
 - SourceControl

下发响应报文:

```
OFPT_FLOW_MOD (Flow Modification Message)
Header:
  Version: 1.3
  Type: FLOW_MOD
  Transaction ID: 12345
Body:
  Cookie: 0x12345678
  Table ID: 0
  Command: OFFPFC_ADD
  Idle Timeout: 0
  Hard Timeout: 0
  Priority: 100
  Buffer ID: OFP_NO_BUFFER
  Out Port: OFFP_ANY
  Flags: 0
  Match Fields:
    - Ethernet Source Address: 00:11:22:33:44:55
    - Ethernet Destination Address: 66:77:88:99:AA:BB
    - IP Source Address: 192.168.0.1
    - IP Destination Address: 192.168.0.2
    - IP Protocol: TCP
  Instructions:
    - Apply Meter: 1
```

6.5 数据面流量控制

数据面流量控制是在数据面交换机或路由器发生流量整形能力越界时，在数据报文中携带的对源端设备和上游设备进行抑制的指示字段，避免流量时延和抖动进一步劣化。

```
POST /api/flows HTTP/1.1
Host: controller.example.com
Content-Type: application/json
{
  "switch": "00:00:00:00:00:00:00:01",
  "priority": 100,
  "match":
  {
    "in_port": 1,
    "eth_type": 0x0800,
```

```

"ipv4_dst": "192.168.0.0/24"
},
"actions":
[
  {
    "type": "OUTPUT",
    "port": 2
  },
  {
    "type": "RATE_LIMIT",
    "rate": "10 Mbps"
  }
]
}

```

6.6 流量需求描述

流量需求描述是指应用发起时，主动显示方式向边缘路由器发起的需求描述请求，通告网络即将入网的应用类型、流量类型、带宽需求、突发容限、时延需求和抖动需求等。该需求描述在端侧为可选项，当应用端不显式告知需求时，由边缘路由器在流量首次进行网络时进行流量识别和测量，自动生成流量需求描述，并将该需求描述报告给网络控制器，供网络控制器生成或修改相关流表。

POST /configure_traffic_requirements HTTP/1.1

Host: management_system.example.com

Authorization: Basic QWxhZGRpbjpvGvuIHNIc2FtZQ==

Content-Type: application/yang-data+json

```

{ "device_id": "12345",
  "application": "voice_call",
  "requirements":
  {
    "latency":
    {
      "max": "100 ms"
    },
    "bandwidth":
    {
      "min": "100 kbps",
      "max": "1 Mbps"
    }
  }
}

```

6.7 节点状态请求消息格式

IP	TCP/UDP	MsgID	MaxSize	Request ID	Request Type
----	---------	-------	---------	------------	--------------

图3 请求消息格式

控制器向网络设备发送请求报文，查询队列详细信息，主要包括队列缓存状态、队列拥塞丢包信息、相邻节点时延信息等，等待设备响应。当有微突发流量信息产生时，需通过控制器对网络进行调整，此时控制器向网络设备发送控制请求报文，开启网络设备流量整形机制，并等待设备发送请求响应报文。

请求消息格式字段含义如下：

MsgID: 请求报文的序列号。

MaxSize: 消息发送者所能够容纳的消息最大字节,同时也表明了发送者能够接收到的最大字节数。

Request ID: 用于匹配请求和响应。

Request Type: 指明请求类型, 01查询请求, 02控制请求。

6.8 节点状态请求响应消息格式

IP	TCP/UDP	MsgID	MaxSize	Response ID	Response Type	MD 1	...	MD N
----	---------	-------	---------	-------------	---------------	------	-----	------

图4 请求响应消息格式

网络设备收到设备发送的请求报文, 查询队列详细信息, 主要包括队列缓存状态、队列拥塞丢包信息、相邻节点时延信息等, 响应该查询请求报文, 将相应查询信息放入元数据并发送至控制器。当收到控制器发送的控制请求报文时, 网络设备执行相关操作, 并根据操作完成情况生成请求响应报文, 告知控制器该请求执行成功或失败。

请求响应消息格式字段含义如下：

MsgID: 请求报文的序列号。

MaxSize: 消息发送者所能够容纳的消息最大字节,同时也表明了发送者能够接收到的最大字节数。

Response ID: 用于匹配请求和响应。

Response Type: 指明请求响应类型, 01查询请求, 02控制请求。

MD 1-N: 元数据, 存储查询信息。

6.9 流量测量报告

网络设备通过监测传入的流量, 识别流量类别, 并记录流量统计信息, 如带宽使用情况和延迟等。

IP	TCP/UDP	MsgID	MaxSize	Response ID	Response Type	MD 1	...	MD N
----	---------	-------	---------	-------------	---------------	------	-----	------

图5 时延检测消息格式

如图5所示, 遥测数据包通过将INT标头插入原始数据包的TCP或UDP报文头中进行封装, 其中, INT HDR是INT的固有报头, 设备通过此字段表示数据包。MD 1-N是元数据, 用于存储INT的监测信息。

网络遥测技术INT可用于检测网络上的端到端延迟, 监测报文转发路径上每台设备的入出端口和队列信息, 入出设备的时间戳信息, 队列的拥塞信息等。INT可有效感知网络链路状态信息, 当监测到流量突发时控制器即开启流量整形。

监测具体流程为源节点定期向目标节点发送INT数据包, 目标节点接收探测数据包后再向源节点发送响应包, 源节点根据响应包接收时间计算数据包时延, 并将时延信息报告给控制器。在大规模网络场景下, 设备只在相邻设备间执行监测并报告给控制器, 控制器收集信息并评估当前网络链路状态信息, 判断是否需要开启流量整形。

6.10 流量时隙映射

网络5.0骨干网流量整形控制协议作用于网络边缘处和骨干网内部。在网络边缘处通过控制实验敏感流量注入时间实现网络边缘处的访问控制, 提前规划队列资源, 缓解域间流量聚合和突发, 保障时延确定性; 在骨干网内建立相邻节点一对一周周期映射关系, 通过控制器保障确定性时延上下界。若网络不支持时钟同步, 则采用异步时间戳方式, 采用分时调度算法来对分配给不同队列的数据包进行整形和调度。将时间划分为不同的时隙, 然后以异步的方式为每个时隙分配带宽。

带外方式时隙映射报文

POST /timeslot_mapping HTTP/1.1

Host: 1234

Content-Type: application/json

```
{  
  "mapping_rule": "round-robin",  
  "timeslot_size": 100,  
  "send_time": "10:00:00",  
  "payload": {  
    "timeslot_1": ["packet_1", "packet_2"],  
    "timeslot_2": ["packet_3", "packet_4"],  
    ...  
  }  
}
```
