

一种多方通话动态控制技术^①

马跃¹, 柴安颖^{1,2}, 于碧辉¹, 于波¹

¹(中国科学院 沈阳计算技术研究所, 沈阳 110168)

²(中国科学院大学, 北京 100049)

摘要: 近些年多媒体技术和数据通信技术发展迅速, 多方音视频通话逐渐兴起. 多方音视频通话能够解决人们远距离通信问题, 但在媒体流控制方法上, 现存通话系统并未将集中式和分布式两种媒体流控制方式进行有效结合, 无法发挥两种控制方法的各自优势. 针对上述问题, 提出了一种多方通话动态控制技术, 以客户端网络带宽和用户数量作为变换条件, 通过对星状网络和网状网络两种拓扑结构的结合, 使多方通话系统能够自适应调节媒体流分发方式. 本系统以 SIP 会议控制为基础并运用 WebRTC 技术提高通话质量, 实验结果表明, 该技术可有效掌握系统内各客户端网络拥塞状况和资源占用率, 对系统媒体流控制方式能够适时进行调整, 使媒体流传输通畅高效, 提高通话质量.

关键词: 多方通话; 动态控制; 拓扑结构; SIP

引用格式: 马跃, 柴安颖, 于碧辉, 于波. 一种多方通话动态控制技术. 计算机系统应用, 2018, 27(3): 64-70. <http://www.c-s-a.org.cn/1003-3254/6276.html>

Multi-Party Call Dynamic Control Technology

MA Yue¹, CHAI An-Ying^{1,2}, YU Bi-Hui¹, YU Bo¹

¹(Shenyang Institute of Computer Technology, Chinese Academy of Sciences, Shenyang 110168, China)

²(University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: In recent years, multimedia technology and data communications technology has developed rapidly. Multi-party audio and video call is becoming popular. Multi-party audio and video call can solve the problem of long-distance communications. However, the existing call system does not combine the centralized and distributed media flow control modes effectively, which fails to give full play to the advantages of the two control methods. This paper presents a multi-party call dynamic control technology. It takes the client network bandwidth and the number of users as the transformation condition. And through the combination of star topology and mesh topology, the multi-call system can adaptively adjust the distribution method of media flow. Based on SIP conference control the system uses WebRTC technology to improve call quality. The experimental results show that the technology can effectively show the congestion and resource occupancy rate of each client in the system, and adjust the system media flow control in a timely manner, so that the media stream can be transmitted smoothly and call quality can be improved.

Key words: multi-party call; dynamic control; topology; SIP

随着全球经济的不断发展, 互联网不断融入人们的生活, 视频通话越来越受人们青睐, 多方音视频通信成为一种很受欢迎的通信方式. 在各通信系统的媒体

流传输中, 存在两种网络传输方式: 一种是星状网络拓扑, 另一种是网状网络拓扑. 但现有的通信方案往往只选取其中一种方法作为媒体流传输方式, 没有充分融

① 收稿时间: 2017-06-15; 修改时间: 2017-07-26; 采用时间: 2017-07-28; csa 在线出版时间: 2018-02-09

合两种方式的各自优势来提高系统对媒体流的处理能力。

实验室现有视频会议系统方案采用星状网络拓扑这种传统传输方式,系统客户端采用 WebRTC 技术^[1]进行音视频处理,服务端通过 SIP 信令集中模式对视频会议进行会议控制,并运用 WebRTC 与 SIP 的转换网关实现双方互通,使视频会议系统音视频通话质量有所保障,同时能够更容易对接到 IMS 网络^[2],满足融合通信的基本要求。本系统虽然能够正常提供音视频通话,但当会议组遇到网络拥塞或系统负载过重时,系统本身无法改变媒体流转发方式,更不能自适应调整媒体流控制方式,使部分通话无法正常进行^[3]。

本文通过对当前音视频通话方案所存在的问题进行研究与分析,并基于实验室现有视频会议系统进行改进,提出了一种多方通话动态控制技术,充分运用会议组网络中星状拓扑和网状拓扑的各自优点,一方面在网络通畅并且用户数量较低情况下,发挥网状拓扑的便捷性,减轻服务器压力;另一方面当系统发生网络拥塞时,充分利用星状拓扑中服务器集中收发媒体流的优势,节省网络带宽并保持通话畅通。

1 相关技术介绍

1.1 SIP 协议分析

SIP(会话初始化协议)是 IETF 提出的一种支持多媒体会话的信令控制协议,在 IP 网络中以其简单、易扩展的特性被广泛应用,具体工作流程如图 1。

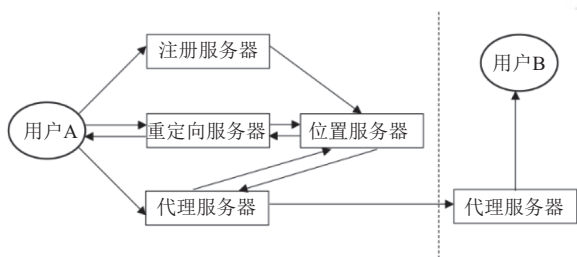


图 1 SIP 工作流程图

SIP 采用文本编码方式,其独立于传输层,属于应用层控制协议,主要由 SIP 服务器和 SIP 用户代理两个组件构成,SIP 服务器包括代理服务器、注册服务器以及重定向服务器^[4]。在消息会话过程中,各系统以 SDP(会话描述协议)传递消息内容,各部分协同工作,让会话更为流畅和高效^[5]。

1.2 MCU 简介

MCU 全称为 Multi Control Unit,即多点控制单元,主要功能为协调和控制多个终端间的音视频传输,为用户提供群组会议服务。如图 2。

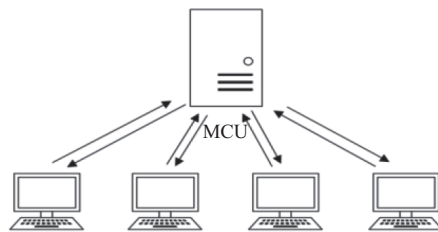


图 2 MCU 传输方式图

MCU 将各终端的信息流同步分离,分析并提取音视频数据和信令等信息,然后将信息送入各模块进行处理,完成相应的音视频混合与交换、路由选择和会议控制等过程,最后将各会议场所需的各种信息重新组合,发送给各客户端。

1.3 会议系统网络拓扑结构

SIP 会议系统模型^[6]分为松耦合模式和紧耦合模式。在紧耦合模式中存在着两种网络拓扑形式,分别为网状拓扑和星状拓扑。

在一个全联通的网状网络中,存在 $N*(N-1)$ 个连接,各端点之间相互连通能够直接进行数据通信,如图 3。这种网络的优点在于各端点之间不需要媒体中继服务器进行媒体分发与管理,即可实现多方通信。但缺点在于无法为大型会议系统提供通信服务,系统的音视频转码和混合在终端完成,且随着网络中参与者的增加,将大量消耗终端的 CPU、内存以及带宽等资源。

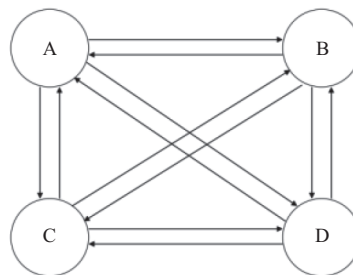


图 3 4 用户全联管网状网络图

星状网络最广泛的应用是通过一个多点控制单元集中整合各终端媒体流,然后发送到每个端点,如图 4。此类网络对于网络中任意客户端,只需传输单个流,节省了带宽并且减轻了 CPU 负载^[7]。但当更多客户端进

入会议后,服务器将对更多路媒体流进行整合与转发,大量消耗服务端资源,因此服务器承载能力会成为此类网络扩展的瓶颈。

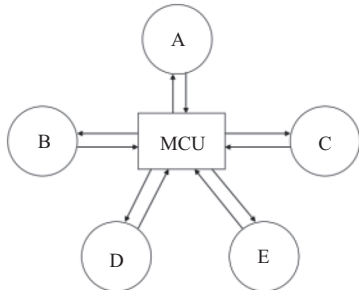


图4 5用户星状网络图

2 动态控制方法的设计与实现

2.1 系统总体结构设计

本系统采用 SIP 信令集中处理模式,实现视频会议的会议控制功能,并支持 SIP 和 WebRTC 两种类型客户端,其中 SIP 客户端可直接连接会议服务器进行会话控制,而 WebRTC 客户端在连接到 WebRTC 服务器后,通过 WebRTC 与 SIP 转换网关实现双方会议控制信息的转换,再与会议服务器进行交互^[8]。MCU 主要用来控制终端媒体流收发,建立流畅的会议通话。系统总体结构如图 5。

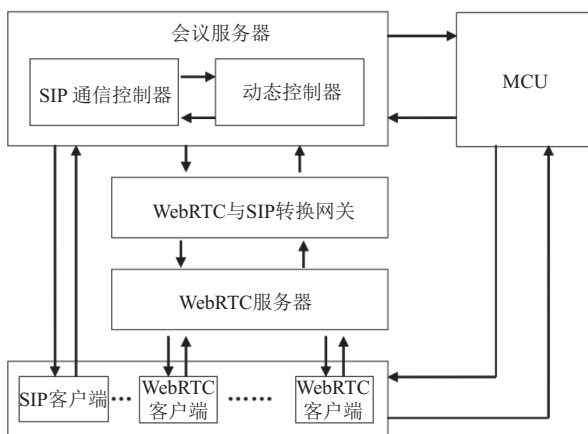


图5 系统总体结构图

系统利用终端采集本地网络数据以及终端资源占用情况,并对 SIP 消息进行扩展,以扩展后的 SIP 消息为载体向会议控制服务器发送各终端数据。

SIP 客户端直接与会议服务器进行连接,WebRTC 客户端则需要经由网关转换信息后才能上传到会议服

务器^[9]。

会议服务器包含 SIP 通信控制器和动态控制器两部分,SIP 通信控制器用来进行会议控制,同时解析各终端数据,并将其会话信息存入数据库中。动态控制器用来对终端提供的数据进行计算和处理,建立系统终端连接状态图,并对各会议组成员进行网络和资源评估,分析会议组媒体流控制方式的各种条件,决定会议组媒体流由 MCU 进行集中管理还是通过各终端自行分发。

2.2 会议服务器设计

多方通话的动态控制技术以服务器为基础,实现媒体流控制变换,该系统的会议服务器由 SIP 通信控制器和动态控制器两部分组成,图 6 为服务器总体架构图。

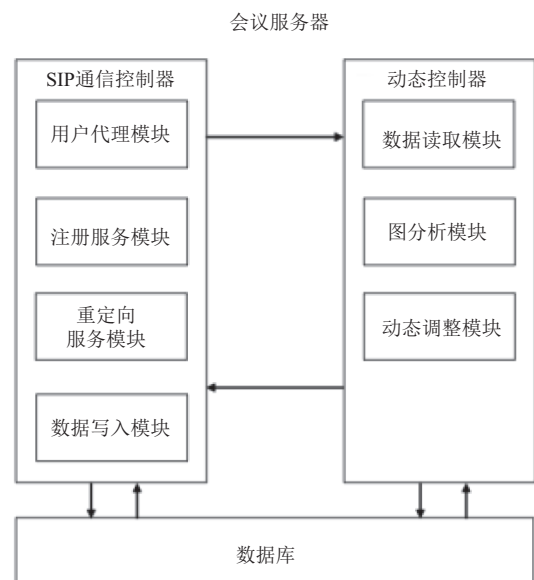


图6 服务器总体架构图

SIP 通信控制器分为用户代理模块^[10]、注册服务模块、重定向服务模块和数据写入模块: 1) 用户代理模块,主要作用是根据目标用户的位置信息将 SIP 信令代理转发给目标用户。2) 注册服务模块,实现对用户进行认证并登记用户的位置信息。3) 重定向服务模块,作用为向用户代理返回最新访问地址列表,让用户代理与新的访问地址进行通信。4) 数据写入模块,作用是将客户端传来的 SIP 消息进行提取,从中解析出 SWIT CHINFO 消息内容,并存入数据库。

动态控制器包含数据读取模块、图分析模块和动

态调整模块 3 部分: 1) 数据读取模块, 主要作用为提取数据库中存放的通话信息, 并传递给图分析模块. 2) 图分析模块, 将数据读取模块传来的信息分析处理, 进行系统会话连接图的还原操作, 掌握系统现存会话状态及各终端网络数据, 并结合终端 CPU 和内存资源使用状况, 采用基于带宽估计的拥塞检测方法实现对终端拥塞情况的计算与评估. 3) 动态调整模块, 根据终端网络状况以及 MCU 资源占用情况, 对系统现存会议组网络最终作出评估, 决策各会议组媒体流控制方式是否做出调整.

2.3 终端会议状态信息采集方案设计

多方通话动态控制技术通过终端向服务器发送各用户网络状况和资源占用情况等数据, 因此需要一种消息载体实现终端和服务器之间的数据传递. 针对以上问题, 系统对 SIP 消息进行扩展, 增加 SWITCHINFO 消息, 由终端每 5 s 采集数据, 通过 SWITCHINFO 消息为服务器传递实时数据.

SWITCHINFO 消息内包含 sip_id, conference_id, contact_points, conference_type, 和 network_condition 五种数据. sip_id 为各客户端在服务器中的唯一标识, conference_id 为多方通话中各会议唯一标识, contact_points 为各客户端媒体流分发时连接的所有端点, 内部数据存储为各连接端点的 sip_id. conference_type 为通话类别, 此类别分为“focus”和“dispersion”两种, 根据客户端之间媒体流连接状况, 当终端检测到没有其他端点与自己连接时, 将 conference_type 赋值为“focus”, 代表该会议组利用 MCU 为每个客户端收发媒体流, 选择了集中控制的方式; 当检测本终端与若干通话终端有媒体流交互时, conference_type 被赋值为“dispersion”, 表明该终端媒体流处理方式为基于网状网络的分散处理模式.

network_condition 则包含了各端点网络状况及 CPU、内存的使用情况, 此消息中所包含字符串较多, 由 SIP 通信服务器解析后会有以下参数存入数据库, 分别为 max_bandwidth、time_interval、consumption_send、consumption_receive、num_package、time_delay、cpu_alltime、cpu_freetime、ram_maxcapacity 和 ram_usedcapacity. 以上各参数会分别应用于图分析模块和动态调整模块中.

下面给出 SWITCHINFO 方法工作流程, User A 向 SIP Server 发送 SWITCHINFO 消息, 向服务器传

递各终端网络信息, SIP Server 响应并返回 200 OK, 表明系统服务器已经接收到了来自 User A 的 SWITCHINFO 消息, 如图 7.

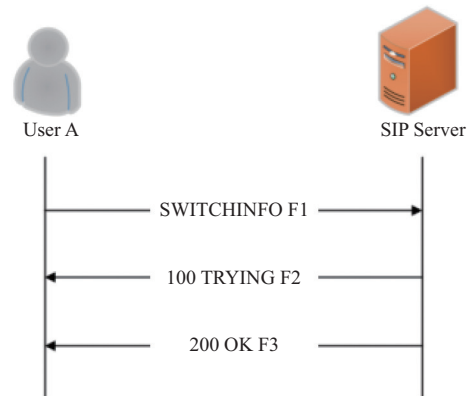


图 7 SWITCHINFO 工作流程图

2.4 图分析方案设计与实现

动态控制器每 5 s 向数据库读取数据, 进行会话分析及终端网络评估, 并以 30 s 为时间段检测终端和 MCU 资源占用情况, 在此时间段中得出 6 次处理结果, 并根据这 6 次统计结果综合判断会话媒体流控制方式是否做出调整.

在媒体流控制方式调整过程中, 图分析模块获得最新数据后, 立即执行系统图分析及还原操作, 并采用基于带宽估计的拥塞检测方法计算各终端的可用带宽及资源占用情况, 得出其拥塞程度指标.

数据读取模块将 SIP 会议控制器存储在数据库中的各项数据进行提取, 发送给图分析模块进行系统的图还原操作. 图分析模块以 conference_id 作为会议群组区分, 如果 contact_points 不为空且 conference_type 值为“dispersion”, 控制器将各端点 contact_points 内包含的端点解析出来并存储在邻接链表中, 以无向图构造算法还原各点连接状况, 并计算正在进行的各会议组内终端数量 participant_count, 以 4 端点连接为例, 相关结构如图 8 和图 9.

图分析模块对各终端 network_condition 解析出的 10 种参数进行转换和计算, 利用基于带宽估计的拥塞检测方法, 求取拥塞程度指标 Q . 采用基于带宽估计的拥塞检测方法可实时估计终端带宽使用情况, 并将终端 CPU 使用率和内存占用率一起纳入终端拥塞判定标准, 以加权形式得到最终计算值 Q , 具体原理如下.

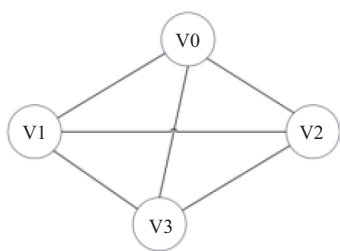


图8 4端点连接图

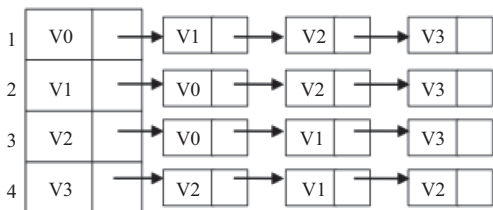


图9 4端点邻接链表图

检测终端最大带宽表示为 B , 终端由于排队延迟所占用的带宽表示为 B_i , 外部应用占用网络带宽为 B_e , 则终端可用带宽为:

B_e 所表示占用带宽为接收占用带宽和发送占用带宽的总和, 设 T 为采样时间间隔, T_r 为间隔内接收数据消耗时间, T_s 为间隔内发送数据消耗时间, 由公式 (1) 可以得出一个采样间隔内网络占用带宽计算值。

$$B_c = \left(\frac{T_r + T_s}{T} \right) B \quad (1)$$

同时系统采用指数加权平均算法对实时网络占用带宽 B_e 进行计算 $B_e(k) = (1 - \alpha)B_e(k - 1) + \alpha B_c(k)$, α 为常数且 $\alpha \in (0, 1)$ 。

排队延迟造成的占用带宽 B_i 可通过 $B_i = m \frac{\sum_{i=1}^n T(i)}{n} B$ 来表示, 其中 m 代表队列存在的数据包, n 代表采样数据包数量, T_i 代表第 i 个包的传输延迟, B 代表终端最大带宽。已知 B_i 与 B_e 便可计算出 B_a , 进一步求得终端带宽使用率 $P_b = 1 - \frac{B_a}{B}$ 。

检测终端 CPU 在 T_1 到 T_2 时间段内, 总使用时间表示为 T , 空闲使用时间表示为 T_a , 可估计出 CPU 即时利用率 $P_r = 1 - \frac{T_a}{T}$ 。根据从数据库提取的参数可知终端最大内存空间为 R , 已使用内存空间为 R_e , 可计算出内存使用率 $P_r = \frac{R_e}{R}$ 。由公式 (2) 计算得出终端拥塞程度指标 Q , β 和 γ 均为加权系数, $\beta \in (0, 1)$, $\gamma \in (0, 1 - \beta)$ 。

$$Q = \beta P_b + \gamma P_r + (1 - \beta - \gamma) P_r \quad (2)$$

图分析模块以 5 s 为时间单位对数据读取模块传来的更新数据进行处理, 重构系统端点连接图并重新

对各终端网络进行评估, 并将计算出的各终端拥塞程度指标 Q 连同系统连接图一并交给动态调整模块进行下一步操作。

2.5 动态调整方案设计与实现

动态调整模块接收由图分析模块传来的数据, 同时向 MCU 发出消息请求, 获取 MCU 的 CPU 使用率、内存使用率和通路路数等数据。并通过 MCU 资源占用的综合评估法, 得出 MCU 资源占用指标 M , 具体方法如下。

动态控制器将 MCU 所传递数据进行解析, 获得 CPU 即时利用率 E_c , 内存使用率 E_r 和通路路数 H_e 。已知 MCU 最大支持通路路数 H , 已存在通路路数为 H_e 可得通路线路使用率 $E_h = \frac{H_e}{H}$ 。系统为更全面衡量 MCU 资源占用情况, 将 3 个参数进行加权处理, 综合评估当前 MCU 负载能力, 得出 MCU 资源占用指标 M , 具体公式为 $M = \alpha E_c + \beta E_r + \gamma E_h$, $\alpha \in (0, 1)$, $\beta \in (0, 1)$, $\gamma \in (0, 1)$ 。

动态调整模块以 conference_id 作为通话组区分, 利用各通话组连接图寻找各组成员, 并以各终端拥塞程度指标 Q 和 MCU 资源占用指标 M 作为标准, 对各会议群组的媒体流控制方式进行评估, 采用公式 $Z = \alpha Q + (1 - \alpha) M$, $\alpha \in (0, 1)$, 得出综合评判指标 Z 。

当 conference_type 值为“dispersion”时, 表明会议组网络为网状网络, 若会议组内所有终端存在 participant_count/2 以上数目综合评判指标 $Z > 0.8$, 表明会议组内多数终端出现网络拥塞状况, 则该会议组必须做出调整。若会议组内存在 participant_count/2 以上数目终端综合评判指标 $0.6 < Z < 0.8$, 判断会议组网络将要发生拥塞, 此时系统对本会议组进行标记, 进入观察状态。若会议组内存在 participant_count/2 以上数目终端综合评判指标 $Z < 0.6$, 则表明会议组无需进行调整。

当 conference_type 值为“focus”时, 若会议组 participant_count ≤ 5 且 MCU 资源占用指标 $M > 0.6$, 则适合会议组媒体流控制方式进行调整。

动态控制器每 5 s 对各会议组进行评估并记录, 以 30 s 为时间段所得到的 6 次评估结果作为最终决策标准。

若会议组 6 次评估结果中存在 3 次及以上符合调整条件, 动态调整模块向 SIP 会议控制器发送调整消息, 由 SIP 会议控制器对该会议组进行媒体流控制方

式的调整,实现网状网络与星状网络之间的切换,如图10为媒体流控制方式切换过程图。

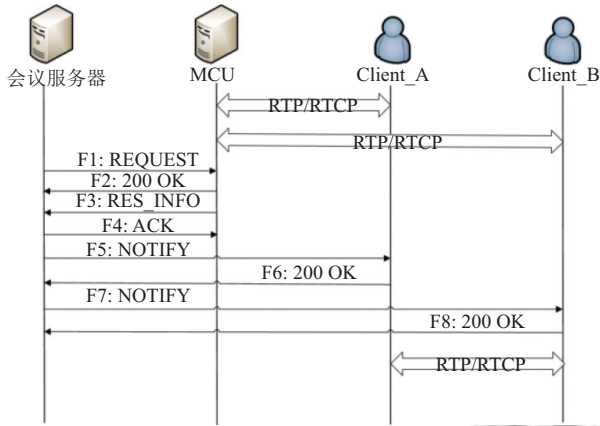


图10 媒体流控制方式切换过程图

3 测试

本次实验采用5台相同配置WebRTC客户端进行测试,选用10Mbps网络带宽进行数据流传输,客户端配置如表1。

表1 客户端配置列表

设备名称	设备参数
CPU	Inter(R) Core(TM) i3 2.27 GHz
内存	2 G
操作系统	Windows 7

多方通话建立之初,系统采用分散控制方式对媒体流在客户端之间进行收发,网络中则形成20条进行数据交互的数据通路。5台客户端音视频效果均能够满足通话需求,系统运行流畅。

然后在各客户端运行笔者编写的Python程序,提高客户端内存和CPU占用比,将客户端两种资源占用率分别提高到70%和80%,同时使用WANem网络模拟工具自定义网络带宽,将各客户端可用带宽比率控制在20%以下,此时发现客户端音视频通话出现卡顿现象,媒体流丢包严重。后台管理系统从图分析模块获取实时数据,经过处理得到当前会议组网络状况图,如图11。

● 拥塞 ● 观察态 ● 正常

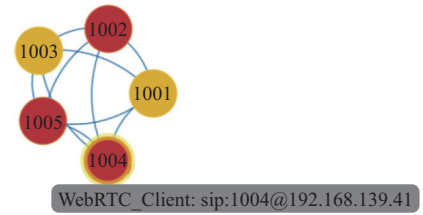


图11 调整前媒体流传输图

经过30秒后,系统服务器做出响应和决策,对会议组媒体流控制方式进行切换,将会议组网络由网状网络切换为星状网络,控制台输出信息如图12。

```

conference_type: dispersion
contact_point: sip:1001@192.168.139.122;sip:1002@192.168.139.41;sip:1003@192.168.139.73;sip:1004@192.168.139.132
network_condition: 2.0.0.0.2079;0.6962;1179;0.17894;0.6962;0.624;3946;1142.06
Client information:
sip_id: sip:1004@192.168.139.73
user_part: 1004
port: 5060
sip_part: 1004
sip_part: 192.168.139.73
conference_id: 472579051732
conference_type: dispersion
contact_point: sip:1001@192.168.139.122;sip:1002@192.168.139.41;sip:1003@192.168.139.156;sip:1004@192.168.139.132
network_condition: 2.0.0.0.2163;0.0731;1065;0.15729;0.0759;0.0316;3048;1247.38
Client information:
sip_id: sip:1005@192.168.139.132
user_part: 1005
port: 5060
sip_part: 1005
sip_part: 192.168.139.132
conference_id: 472579051732
conference_type: dispersion
contact_point: sip:1001@192.168.139.122;sip:1002@192.168.139.41;sip:1003@192.168.139.156;sip:1004@192.168.139.132
network_condition: 2.0.0.0.5401;0.0661;1492;0.14385;0.1029;0.0426;3048;1248.45
The number of conference groups: 1;Meeting numbers: 5; Media stream connection: decentralized;Statistics Result: 6
Conference id is 472579051732, the clients sip : 1001@192.168.139.122, sip:1002@192.168.139.41, sip:1003@192.168.139.156, sip:1004@192.168.139.73, sip:1005@192.168.139.132 belong to the WebRTC category, through the gateway and the conference server to communicate, is currently distributed media flow control, the 3 and 3 clients more comprehensive evaluation of the terminal index 2> 0.5 in the 0 statistical results, Web resource occupancy indicators 1< 0.5, meet the switching conditions.
08:15:25.034 Sending adjustment message to SIP call controller...
08:15:27.442 SIP call controller receives message...
08:15:28.223 Sending NOTIFY message to clients...
08:15:29.247 sip:1001@192.168.139.122 return 200 OK, successful reception
08:15:30.170 sip:1002@192.168.139.41 return 200 OK, successful reception
08:15:30.623 sip:1003@192.168.139.156 return 200 OK, successful reception
08:15:31.260 sip:1004@192.168.139.73 return 200 OK, successful reception
08:15:31.894 sip:1005@192.168.139.132 return 200 OK, successful reception
08:15:32.313 Sending NOTIFY message to clients...
08:15:36.732 Switch successful, media flow control mode has been switched

```

图12 控制台输出信息图

此时,后台管理系统的会议组网络状况图发生变化,具体如图13,表明该会议组媒体流控制方式已经改变,并且通过分析节点颜色含义,说明各客户端网络状况良好,系统音视频通信运行流畅,达到用户通话的满意效果。

● 拥塞 ● 观察态 ● 正常

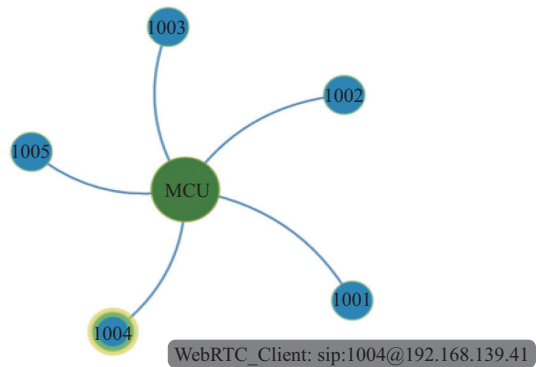


图13 调整后媒体流传输图

4 结语

本文通过对 SIP 会议模型中网络拓扑结构的深入研究,设计了一种基于多方通话的动态控制技术,将集中式和分布式两种媒体流控制方式进行有效结合,并在实验室现有视频会议系统中予以实现.通过最后的测试结果,验证了该技术的可行性以及优越性,当终端发生网络拥塞时系统能够合理调整媒体流传输方式,使用户获得流畅通话体验.

参考文献

- 1 Chiang CY, Chen YL, Tsai PS, *et al.* A video conferencing system based on webRTC for seniors. Proceedings of 2014 International Conference on Trustworthy Systems and Their Applications (TSA). Taichung, China. 2014. 51–56.
- 2 Chen XQ, Jones HM, Jayalath ADS. Congestion-aware routing protocol for mobile ad hoc networks. Proceedings of the 66th Vehicular Technology Conference. Baltimore, MD, USA. 2007. 21–25.
- 3 Rosenberg J. RFC 4353, A Framework for Conferencing with the Session Initiation Protocol (SIP). New York: IETF, 2006.
- 4 Rosenberg J, Schulzrinne H, Camarillo G, *et al.* RFC 3261, SIP: Session Initiation Protocol. New York: IETF, 2002.
- 5 竹洪涛. 一种基于 SIP 和 WebRTC 的实时可视对讲方案设计[硕士学位论文]. 成都: 西南交通大学, 2013.
- 6 黎忠侃. 基于 SIP 的视频会议控制服务器的设计与实现[硕士学位论文]. 广州: 华南理工大学, 2010.
- 7 王森. SIP 多方视频会议应用服务器设计与实现[硕士学位论文]. 沈阳: 中国科学院研究生院(沈阳计算技术研究所), 2008.
- 8 李冬. 一种基于 WebRTC 的视频会议系统的架构. 智能计算机与应用, 2015, 5(4): 39–41.
- 9 胡敏, 刘六程, 刘鹏. 基于 WebRTC 的视频会议系统的设计与实现. 电视技术, 2013, 37(1): 141–143, 150.
- 10 曲冠桦. 基于 SIP 协议的视频会议系统实现[硕士学位论文]. 成都: 电子科技大学, 2012.