

Design of Remote Network Audio and Video Communication System

Chen Lin
School of Communication
Engineering
Chengdu University of
Information Technology
Chengdu, China

Tianbao Wang
School of Communication
Engineering
Chengdu University of
Information Technology
Chengdu, China

Chenxi Yang
School of Communication
Engineering
Chengdu University of
Information Technology
Chengdu, China

Abstract: With the continuous development of information technology, real-time audio and video communication is becoming more and more important in people's daily life. At present, there are a lot of software with audio and video communication function in the market, such as Tencent QQ software, wechat, nailing, etc. these software are developed with C / S architecture, and can not achieve cross platform and cross terminal. Users need to download the software to carry out audio and video communication. In view of this, this paper uses webrtc technology based on ice framework to realize the design of remote audio and video communication system, completes the internal network penetration, establishes the communication process between clients, designs room server and signaling server for conversation, and completes the design of remote network audio and video communication system.

Keywords: Intranet penetrating, webrtc, remote audio and video

1. INTRODUCTION

At present, most of the audio and video real-time communication software in the market is developed by independent application program based on C / S architecture, which requires users to install software on the device, and the real-time communication of audio and video between software of different platforms is not allowed. Audio and video data coding and decoding methods, echo cancellation, communication protocol establishment and other issues are also the difficulties of traditional audio and video communication software. To solve this problem, this paper introduces and adopts webrtc real-time communication technology based on browser launched by Google, which provides audio and video acquisition, coding and decoding, audio / video synchronization, flow control (RTP / RTCP), network transmission and other functions [1]. It also supports cross platform communication based on IP network, which can realize real-time audio and video communication between communication terminals without platform control. The ice protocol stack, which integrates stun and turn advantages, breaks through the limitation of NAT / firewall and realizes intranet penetration. Using webrtc technology and Intranet penetration ice, users can realize audio and video communication through browser in heterogeneous network environment, reduce the load of client equipment, and meet the needs of cross platform and long-distance instant communication.

2. INTRODUCTION TO SYSTEM TECHNOLOGY

2.1 Webrtc Technology

Webrtc technology is a real-time multimedia communication technology opened by Google in 2011. Developers can complete the development of audio and video communication by calling the API interface provided by browser itself, and can realize real-time audio and video communication function based on browser without any third-party plug-in. As shown in the webrtc architecture diagram in Figure 1, the overall architecture

of webrtc is divided into three layers. The first layer of your web app is a web application developed by developers through a browser that provides webrtc support. The second layer mainly provides JavaScript API for developers to develop independently, For example: getUserMedia API provides webrtc with the function of acquiring video and audio data from the camera and microphone of the device; rtcpeerconnection API transfers stream data between browsers; rtcdata channel API establishes data channel for webrtc. The third layer is the core layer of webrtc technology, including voice engine, video engine and transport engine.^[2-3]The audio engine mainly encodes the audio data collected by the microphone for audio coding, noise suppression, echo cancellation, etc., and then transmits it to other clients through the network; the video engine encodes and processes the video data collected by the camera, enhances the color of the video, and improves the user experience, and then transmits it to other clients through the network. If the intranet penetration is realized, the audio and video data are encrypted and transmitted to different LAN clients.^[4]

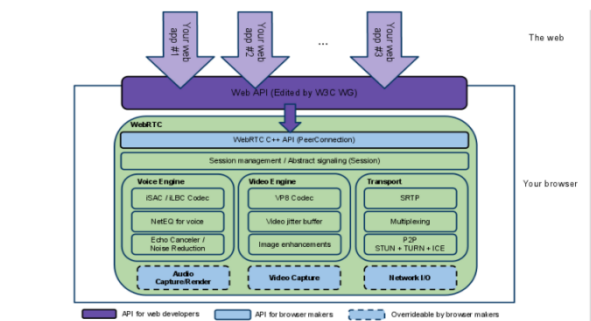


Figure.1 Webrtc architecture

2.2 ICE

At present, many schemes have been applied to solve NAT problems, such as stun (session transverse utilities for NAT),

turn (full name of turn is transverse using relays around NAT, mainly adding relay function), algs, symmetric RTP, RSIP, etc. However, when these technologies are applied to different network topologies, they have obvious advantages and disadvantages, so that different solutions can only be applied according to different access methods. Therefore, the problems of all NAT and efficiency can not be well solved, and many complex and fragile factors will be introduced into the system.

In this paper, ice interactive connectivity is used.^[5]It is a solution of traversing NAT / firewall proposed by IETF. It is a NAT transmission protocol for offer / answer model. A medium based on UDP is established on this model. The SDP of offer and answer contains a variety of IP addresses and ports. The IP addresses in the local SDP and the remote SDP are paired, and then the connectivity test is carried out through P2P connectivity check. If the test is passed, then the test is conducted And indicates that the transmission address can establish a connection, It is suitable for use in heterogeneous network environment. As shown in Figure. 2 Ice framework, ice effectively integrates the advantages of stun and turn, and makes it work under the most suitable situation by comprehensively using stun and turn, so as to make up for the defects of using one of them alone, and provides the optimal solution for the system to penetrate NAT / firewall in various situations.

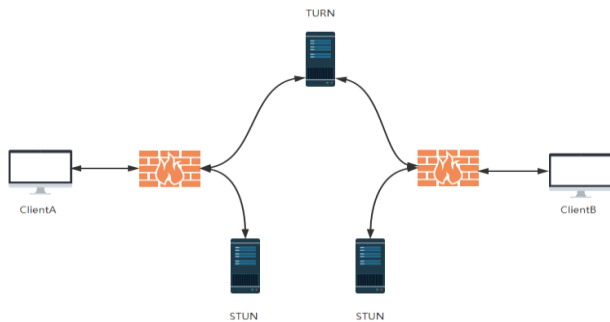


Figure.2 Ice framework

3. SYSTEM ARCHITECTURE

The system architecture of remote network audio and video communication system is shown in Figure 3. The system is composed of web communication client, stun and turn server based on ice framework, and signal server. Because the intranet IP can't communicate directly in the public network, it is necessary to borrow the public IP to realize the communication between clients not in the same LAN. Therefore, local and remote, the web communication client of remote network audio and video communication system, first obtains its communication candidate address by ice protocol including stun and turn protocol, that is, the browser sends test data package to stun server to obtain its public IP address (NAT address). In some cases, if stun server can not achieve traversal, it must In the public network, turn server is used to implement relay transfer, and then the candidate address obtained is communicated to the other party through the signaling server exchange. The signaling server is also responsible for the interaction of other signaling information, such as SDP message. After the signaling information exchange is completed, the transmission of media stream data can be realized^[6].In the process of media stream data transmission, the web communication client using webrtc technology will construct audio and video media controls in the web page, bind the audio and video data collected by local and remote cameras and microphones to the media controls for output, so as to realize remote network audio and video communication.^[7]

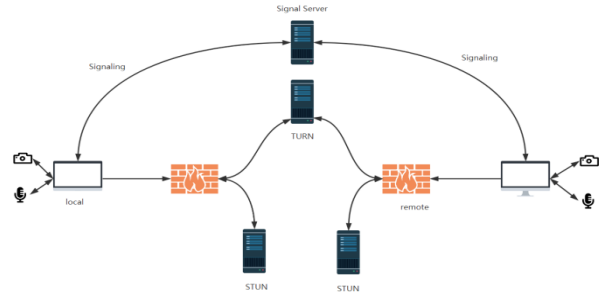


Figure.3 System architecture

4. TRANSMISSION MODE

The transmission mode of remote network audio and video communication system adopts ice framework, stun and turn are comprehensively used, so that client A and client B in heterogeneous network environment can realize communication under Wan, and meet the network transmission requirements of remote network audio and video communication system. As shown in Figure 4, symmetric NAT / firewall network topology based on iceframework.

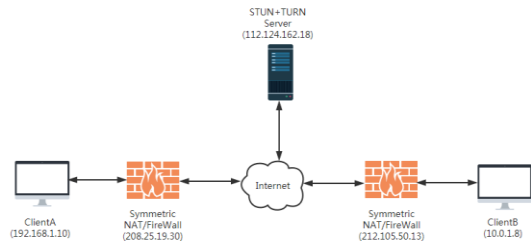


Figure.4 Symmetric NAT / firewall network topology

client A and client B realize the exchange process through NAT / firewall through ice framework^[9]as follows:

- (1) client A collects IP address and finds out the traffic address that can be received from stun server and turn server;
- (2) client A sends an address list to stun server, and then sends start information to client B according to the sorted address list. The purpose is to realize the communication between client A and client B;
- (3) Client B sends a stun request to each address in the startup message;
- (4) client A sends the reply message of the first stun request received to client B;
- (5) After receiving stun's reply, client B finds out the addresses that can realize communication between client A and client B;
- (6)The highest address in the list is used to further complete the communication between client B and client A.

5. AUDIO AND VIDEO COMMUNICATION DESIGN

5.1 Communication process establishment

The design requirements of remote network audio and video communication system: 1. Be able to collect real-time audio and video data from local client A and remote client B; 2. The collected images and audio have certain clarity, which will not affect the identification between communication personnel; 3. Send audio and video packets in the same step; 4. Video

window can receive the transmitted video stream and display real-time pictures; 5. Local area client A and remote client B enter the same room server and can exit the room server at any time. According to the design requirements, webrtc technology is used to establish the communication process between local and remote personnel, the room server for audio and video conversation between local client A and remote client B, and the signaling server for signaling exchange between local client A and remote client B.

The communication flow between local client A and remote client B is established as shown in Figure 5.

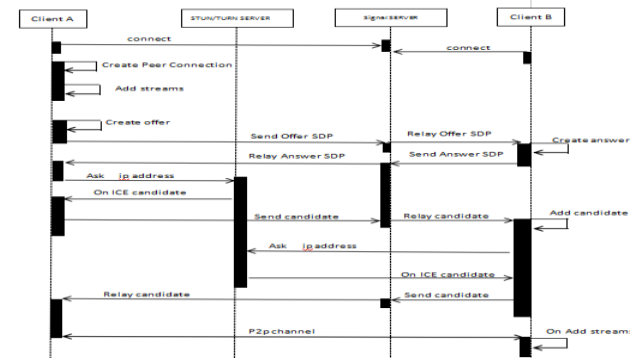


Figure.5 Communication flow chart

(1) Local client A first creates a peerconnection object, then opens the local audio and video device, encapsulates the audio and video data as mediastream and adds it to peerconnection.

(2) Local client A calls the createoffer method of peerconnection to create an SDP object for offer. The SDP object stores the relevant parameters of the current audio and video. Local client A saves the SDP object through the setlocaldescription method of peerconnection, and sends it to remote client B through signal server.

(3) Remote client B receives an offer sent by local client A The SDP object is saved by the setremotedescription method of peerconnection, and a response SDP object is created by calling the createanswer method of peerconnection. The SDP object is saved by setlocaldescription method of peerconnection and sent to local client A through signal server.

(4) Local client A receives the response SDP object sent by remote client B and saves it through the setremotedescription method of peerconnection.

(5) In the offer / answer process of SDP information, local client A and remote client B have created corresponding audio channels and video channels according to SDP information, and started the collection of candidate data. Candidate data can be simply interpreted into IP address information of communication terminal (local IP address, public IP address, address assigned by relay server).

(6) When local client A collects candidate information, peerconnection will send notification to local client A through onicecandidate interface. Local client A will send the received candidate information to remote client B through signal server, and remote client B will save it through addicecandidate method of peerconnection. Do the same for remote client B and local client A again.

(7) In this way, the local client A and the remote client B have established a P2P channel for audio and video transmission. After receiving the audio and video stream from local client A,

remote client B will return a mediastream object identifying the audio and video stream of local client A through the onaddstream callback interface of peerconnection, which can be rendered in remote client B. The same operation also adapts to the transmission of audio and video stream from remote client B to local client A.

5.2 Building room server

The room server of audio and video conversation between remote client B and local client a adopts loose coupling mode [10], which makes local client a and remote client B directly enter the room for interactive audio and video communication. The loosely coupled room server is connected with the server of signaling server through websocket through WS module in nodejs. The algorithm flow of the whole staff joining the room is as follows:

```

(1) this.rooms = {}; // initialize room
(2) socket.send(JSON.stringify({
(3)   "eventName": "join",
(4)   "data": {
(5)     "room": "roomname"
(6)   } }, //send join signaling containing the name
of the room and join the room
(7) current.send(JSON.stringify({
(8)   "eventName": "new_people",
(9)   "data": {
(10)    "socketId": socket.id
(11)  } }, //send the information of new employees to
all the people in the current room
(12) socket.send(JSON.stringify({
(13)   "eventName": "others",
(14)   "data": {
(15)     "connections": ids,
(16)     "you": socket.id
(17)   } }, //send information about other people in
the room to the new person for establishing a connection
with each other
(18) room[i].send(JSON.stringify({
(19)   "eventName": "removeroom",
(20)   "data": {
(21)     "socketId": socket.id

```

```
(22)    }); //send exit room signaling to inform  
other personnel
```

```
(23) that.removeSocket(socket); //personnel exit the  
room
```

5.3 Establishment of signaling server

Before the connection between remote client B and local client a is established, data cannot be transmitted between the two ends. Therefore, we need to realize the signaling data exchange between the two ends through the signaling server, such as the exchange of media description information SDP [11], the exchange of connection address, the exchange of connection message to control the opening or closing of communication, and the message exchange informing each other when an error occurs. Establish a point-to-point connection at both ends. Nodejs is used to design the signaling server this time. At present, nodejs is a very mature web server. It uses V8 engine on the server side to parse the written JavaScript application. After the application is parsed by V8, the C / C + + API at the bottom of nodejs is called to start the server. First load the client of the server, that is, the room server where remote client B interacts with local client a for audio and video interaction. The server listens to port 3000 to process different messages. For example, when the server receives a message message, it will broadcast directly, and all clients connected to the server will receive the broadcast message; the server receives "create or" if there is no one in the room, it will send a "created" message; if there is a person in the room, it will send a "join" message; if you exit the room, it will send a "removeroom" message, and then through the node server.js The command starts the signaling server.

6. CONCLUDING REMARKS

This audio and video part uses webrtc technology to establish the local and remote communication process, completes the design of the local and remote room server and signaling server used to exchange signaling information, and completes the design of remote audio and video communication system under heterogeneous network environment combined with ice framework, which is conducive to users' real-time long-distance audio and video communication on cross platform web pages, and reduces the cost of communication software Load, reduce the hardware load. On the basis of this design, more functions can be developed and designed, which is conducive to more research on remote network audio and video system in the future.

7. REFERENCES

- [1] Li Yuxuan. Design and implementation of instant messaging video system based on webrtc [D]. 2016
- [2] Hu Xun, Zhou Yuanping. Wireless video transmission onAndroid mobile terminal based on webrtc [J]. Microcomputer and applications, 2015 (19): 83-85 DOI:10.3969/j.issn.1674-7720.2015.
- [3] Lin Hong, Wang song, Yang Xin, et al. Application andplatform technology development and design based on webrtc technology [J] . Telecommunication science, 2013, 29 (9): 20-25
- [4] Cao Shenhui. Research and implementation of NAT traversal technology [D]. Nanjing: Nanjing University of Posts and telecommunications, 2013

- [5] Du Derong, Liu Jianming, Li Hongzhou, et al. Research on SIP traversing NAT in mobile Internet environment [J]. Microelectronics and computer, 2013,30 (7): 11 - 113
- [6] Zhu Guang, Zhang Yunhua, Lu Juan. Research on the scheme of VoIP traversing NAT based on ice 00. Computer applications and software, 2011,28 (10): 223-224 NAT and firewall traversal with STUN/TURN/ICE[Z/OL]. Perreault S. <http://viagenie.qc.ca/publications/2008-09-24-astricon-stun-turn-ice.pdf> .
- [7] Peng Yongchao, Fei Jinghao. Interactive live video system based on webrtc and PWA [J] . Computer programming skills and maintenance, 2017 (2): 74-76
- [8] NAT and firewall traversal with STUN/TURN/ICE[Z/OL]. Perreault S. <http://viagenie.qc.ca/publications/2008-09-24-astricon-stun-turn-ice.pdf> .
- [9] media server technology for convergent WWW/mobile real—timemultimedia communications supporting WebRTC [C] / / World of Wireless, Mobile & Multimedia Networks. IEEE, 2013.
- [10] Liang Yan. Analysis of webrtc technology based on HTML5 [J] . Information and communication technology, 2014 (2): 52- 56