# Feladatkiírás

A feladat egy WebRTC technológiára épülő, böngészőben futó csevegő alkalmazás elkészítése. A megoldásnak támogatnia kell a szöveges, videó és hang alapú csevegést is, a kliensek közötti (peer-to-peer) kapcsolatok segítségével. Az elkészült alkalmazásban, egy felhasználó kiválaszthat egy csoportot a többi online felhasználó közül, akikkel csevegni szeretne. Ekkor a kiválasztott módszertől függően, vagy csoportos üzenetet küld nekik, vagy csoportos hívást kezdeményez, esetleg videó konferenciát indít. A megoldás elkészítéséhez szükséges bizonyos központi architektúrális komponensek implementálása is, de a feladat fókusza a kliensek közötti kommunikáción van. A hallgató feladatának a következőkre kell kiterjednie:

* Ismerje meg és mutassa be a WebRTC szabvány fő komponenseit, azok funkcióit!
* Elemezze a különböző hálózati infrastruktúra elemek (tűzfalak, NAT) hatását a WebRTC-s peer-to-peer kapcsolatokra.
* Készítse el és tesztelje a rendszer implementációját a három (szöveg, hang, videó) adat típus átvitelére.



Budapesti Műszaki és Gazdaságtudományi Egyetem

Villamosmérnöki és Informatikai Kar

# Fedlap

Juhász Bálint

WebRTC alapú csevegő alkalmazás, hang és videó támogatással

**Konzulens**

Jánoky László Viktor

Budapest, 2018

# Tartalomjegyzék

Contents

[Feladatkiírás 1](#_Toc530047440)

[Fedlap 2](#_Toc530047441)

[Tartalomjegyzék 3](#_Toc530047442)

[Hallgatói Nyilatkozat 5](#_Toc530047443)

[Összefoglaló 6](#_Toc530047444)

[Abstract 7](#_Toc530047445)

[Valós idejű kommunikáció 8](#_Toc530047446)

[Bevezető 8](#_Toc530047447)

[Session Traversal Utilities for NAT (STUN) 8](#_Toc530047448)

[Network address translation (NAT) 8](#_Toc530047449)

[Traversal Using Relays around NAT (TURN) 8](#_Toc530047450)

[ICE 8](#_Toc530047451)

[Session Description Protocol (SDP) 11](#_Toc530047452)

[SDP fejléc 11](#_Toc530047453)

[WebRTC 14](#_Toc530047454)

[Lehetőségek 18](#_Toc530047455)

[Média 18](#_Toc530047456)

[Adatfolyam 18](#_Toc530047457)

[WebRTC a gyakorlatban 18](#_Toc530047458)

[Egy-egy 18](#_Toc530047459)

[Egy-több 18](#_Toc530047460)

[Csillag topológia 18](#_Toc530047461)

[Több-több 18](#_Toc530047462)

[Teljes gráf 19](#_Toc530047463)

[Selective Forwarding Unit (SFU) 19](#_Toc530047464)

[Multipoint Control Unit (MCU) 19](#_Toc530047465)

[SFU vs. MCU 19](#_Toc530047466)

[Scalable Video Coding 19](#_Toc530047467)

[Last N 19](#_Toc530047468)

[Önálló munka bemutatása 20](#_Toc530047469)

[Egy-egy kapcsolatú demó 20](#_Toc530047470)

[Megvalósítás 20](#_Toc530047471)

[WebRTC könyvtár 22](#_Toc530047472)

[Fájl küldés 23](#_Toc530047473)

[UI 23](#_Toc530047474)

[Összefoglaló 26](#_Toc530047475)

[Több-több kapcsolatú webapp 26](#_Toc530047476)

[Bevezető 26](#_Toc530047477)

[Első verzió 27](#_Toc530047478)

[Második verzió 30](#_Toc530047479)

[Önálló munka értékelése 31](#_Toc530047480)

[Hogyan csinálják mások? 31](#_Toc530047481)

[Összefoglaló 31](#_Toc530047482)

[Függelék 31](#_Toc530047483)

# Hallgatói Nyilatkozat

Alulírott Juhász Bálint, szigorló hallgató kijelentem, hogy ezt a szakdolgozatot/ diplomatervet meg nem engedett segítség nélkül, saját magam készítettem, csak a megadott forrásokat (szakirodalom, eszközök stb.) használtam fel. Minden olyan részt, melyet szó szerint, vagy azonos értelemben, de átfogalmazva más forrásból átvettem, egyértelműen, a forrás megadásával megjelöltem.

Hozzájárulok, hogy a jelen munkám alapadatait (szerző, cím, angol és magyar nyelvű tartalmi kivonat, készítés éve, konzulens(ek) neve) a BME VIK nyilvánosan hozzáférhető elektronikus formában, a munka teljes szövegét pedig az egyetem belső hálózatán keresztül (vagy hitelesített felhasználók számára) közzétegye. Kijelentem, hogy a benyújtott munka és annak elektronikus verziója megegyezik. Dékáni engedéllyel titkosított diplomatervek esetén a dolgozat szövege csak 3 év eltelte után válik hozzáférhetővé.

Kelt: Budapest, 2018. 05. 12.

........................................................

Juhász Bálint

# Összefoglaló

Az idő előrehaladtával egyre több funkciót támogatnak a böngészők. Ezek közé tartozik a WebRTC is, ami peer to peer hálózat kialakítását teszi lehetővé. De mi is az a WebRTC? Az RTC feloldása Real Time Communication, tehát valós idejű kommunikációt valósít meg. Ilyen megoldás a VOIP, illetve a WebRTC is. A kommunikáció UDP fölött zajlik, mivel a TCP nem rendelkezik a stream-eléshez alkalmas latency-vel. A WebRTC, ugyanúgy ahogy a VOIP is, SDP header megosztásával épít fel pont-pont kapcsolatot full duplex módon. Amennyiben mindkét kliens elérhető egymás felé saját ip címükön, a kapcsolat egyszerű SDP cserével épül fel, míg NAT hálózatokon át szükség van egymás publikus ip címére ugyanehhez. Ha ismerjük a saját publikus ip címünket, még mindig nincs gond, viszont ha nem, már szükség van egy publikus címen elérhető szerverre, ami meg tudja osztani a kliensekkel, hogy ők milyen publikus címen érhetőek el. Emiatt sajnos a WebRTC a gyakorlatban nem tekinthető teljesen decentralizáltnak, csak ha minden kliensnek egyedi címe van és ugyanazon a hálózaton vannak.

A WebRTC nem csak a böngésző forráskódjában, de API szinten is elég komplex, annak ellenére, hogy mivel kliens oldal, ezért JavaScriptben általában próbálnak minden API-t egyszerűvé tenni. Emiatt az első feladatom az volt, hogy írjak egy demó oldalt, ahol szerverek nélkül fel lehet építeni kapcsolatot két gép között. Ehhez egy egyszerűbb JS library-t is kellett írnom a WebRTC köré. Az SDP header-ök megosztását a kliensek Miután ezt sikerült elérnem, áttértem a fő feladatomra, egy Skype jellegű alkalmazásra, ahol több ember beszélgethet egymással minden kényelmetlenség nélkül. Ehhez már Signaling Server kell, ami az SDP header cserét elvégzi. Ahhoz, hogy a NAT mögötti gépek is kapcsolódni tudjanak egymáshoz, STUN vagy TURN szerverre van szükség. A STUN csak publikus ip-t szolgáltat vissza, a TURN pedig a kapcsolatfelvételtől kezdve minden üzenetet tükröz a másik félnek.

A több kliens közötti kommunikáció felvet egy toávbbi problémát. Ha N ember szeretne egymással beszélni, az egy böngészőre vetítve 2N darab stream, ami nem működőképes módszer. Emiatt médiaszervereket kell alkalmazni, amik ezt a számot limitálják. Amennyiben kapcsolatonként szeretnénk közbeavatkozni egy szerverrel, egy kliensnek elég N+1 stream, ha pedig egy eggyé renderelt képet szeretnénk kapni a többi klienstől, elég 2 csatorna, egy fel és egy le. Természetesen a szerver az utóbbi esetben sokkal több számítást végez.

Mint látható, egy WebRTC szolgáltatás létrehozása nagyon sok szinten kíván elmélyült tudást, főleg azért, mert akár 5 különböző szervernek is részt kell vennie, hogy gondtalanul működjön a szolgáltatás.

# Abstract

# Szakterület és probléma bemutatása

[Ebben a fejezetben fogod leírni hogy mi a WebRTC és miért nem egyszerű, ez kvázi a „Valós idejű kommunikáció” és „WebRTC a gyakorlatban” fejezetek összevonása]

## Valós idejű kommunikáció

### Bevezető

A valós idejű kommunikáció sok használati módot fed le. Mi most főként a media streaming-gel fogunk foglalkozni. Ha van valamilyen tartalmunk, legyen az videó, audió, nyers adat, különböző elvárásoknak kell megfelelnie a csatornának. Videó és audió esetén ez a késleltetés minimalizálása és a minőség megtartása, míg adat csatorna esetén az üzenet változatlansága. Ideális esetben az átviteli közeg ezt garantálja, hiszen végtelen sok adatot tud szállítani késleltetés nélkül, ezzel szemben a valóságban mindenhol kompromisszumokat kell kötni mely hol a késleltetés, hol a minőség, hol a hibátlan átvitel rovására megy. Egy ilyen környezetben kell használható megoldást találni a stream-elésre. Ez a valós idejő kommunikáció problémaköre.

Mielőtt rátérnék a WebRTC által alkalmazott legfontosabb protokollra, meg kell előtte említenem 3 technológiát, melyekre támaszkodik.

### Network address translation (NAT)

Elsőként a NAT-ot mutatom be, mivel ez a technológia okoz lényegében gondot a WebRTC-nek, ennek megoldására születtek a WebRTC-t segítő protokollok.

<https://en.wikipedia.org/wiki/Network_address_translation>  
https://www.think-like-a-computer.com/2011/09/16/types-of-nat/  
https://tools.ietf.org/id/draft-takeda-symmetric-nat-traversal-00.txt

A NAT-ot eredetileg ip címek egy-egy megfeleltetésű fordítására használták. Ha például egy iroda költözött, a gépek megtarthatták a címüket, a kimenő/bemenő kéréseket a router átfordította. A lokális címet a kiosztott publikus címre, illetve fordítva. Ezt hívják **Basic NAT**-nak.

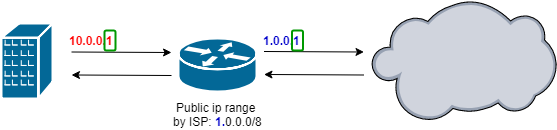


Figure . Basic NAT a gyakorlatban

Valójában a NAT-ot sok lokális ip kevés (legtöbbször 1) publikus ip címre való leképezésére használják. Ezt hívják **One-to-many NAT**-nak. Ilyenkor nincs kölcsönös megfeleltetés a két tartomány között, hanem a routernek más helyen kell eltárolnia azt az információt, hogy melyik lokális címről érkezett a kérés. TCP és UDP esetén ez a meta-adat a portszámban foglal helyet. Ha egy kliens a routeren keresztül el akar érni egy erőforrást, a router egy még nem foglalt source portszámhoz köti a klienst és erre a portra várja vissza is a választ. Így sikeresen vissza tud jutni a válasz a klienshez. *(megjegyzés: mivel a portszám 16 biten tárolódik a TCP és UDP header-ben, hiába van /8-as subnete egy hálózatnak, csak 65536 kliens kommunikálhat egyszerre kifelé, a maradék 16 millió addig nem éri el az internetet)*

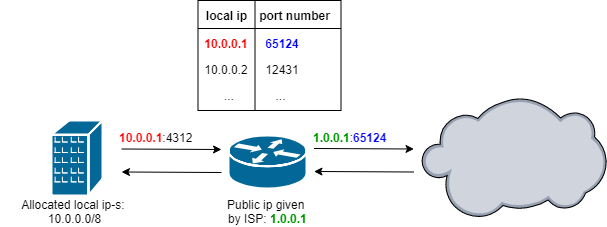


Figure . Many-to-one a gyakorlatban

Láthatjuk, hogy a One-to-many NAT-nak sokféle megoldása lehet, ha figyelembe vesszük, hogy nem csak request-response viselkedéső forgalom lehet a routeren keresztül, de például egy egy kívülről jövő kérést is le kell kezelni. A következő fordítási módszerek a legelterjedtebbek:

**Full-cone NAT (static NAT):** A basic NAT megfeleltetése úgy, hogy mivel csak egy publikus ip cím van, a lokális címeket portként azonosítja. Előnye, hogy amíg a kliens elérhető a belső hálózaton, kívülről bárki kommunikálhat vele. Hátránya, hogy csak egy porton keresztül lehet kommunikálni vele. *(megjegyzés: amikor kézzel állítunk be egy ilyen NAT-olást néhány gépre, azt hívják port forwardingnak. Így elérhetővé tehetünk lokális szervereket az interneten, azzal a hátránnyal, hogy egy portot csak az egyikük használhatja)*

*[Lásd:* Basic NAT a gyakorlatban*]*

**Address-restricted cone NAT:** Ugyanaz a működési elv mint a Full-cone-nál, viszont a NAT publikus oldalán lévő kliens csak akkor kapcsolódhat a lokális hálózaton lévő kliensre, ha az már forgalmazott felé. Viszont amennyiben már szabad kérést intéznie felé, a külső kliens bármely portjáról kapcsolódhat rá.

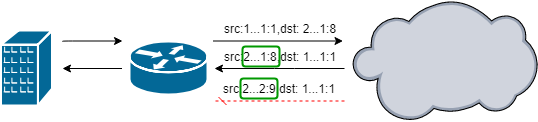


Figure . Az ARC NAT nem engedi be a kliensre a 2.0.0.2 címről érkező kérést

**Port-restricted cone NAT:** Ugyanaz a működési elv mint az address-restricted változatnál, de itt a külső kliens csak azon a porton indíthat kérést a lokális felé, amivel előtte a lokális kliens küldött adatot felé.

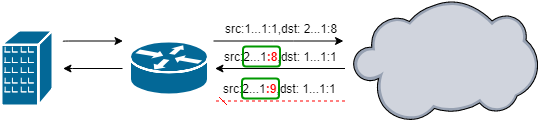


Figure . Azonos címről sem enged be, ha a port szám nem egyezik meg azzal, amire a lokális kliens kiküldte az adatot.

**Symmetric NAT:** Ugyanaz a működési elv mint a post restricted NAT-nál, viszont minden egyes TCP vagy UDP kapcsolat a lokális kliens felől a routeren külön portról indul, így csak az a publikus kliens küldhet vissza adatot, amelyiknek a felkérést címezték.

Mielőtt bemutatnám a STUN szerver feladatát, még meg kell említeni a NAT típusok közötti kapcsolat lehetőségeit.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Receiver | Open | Full cone | Restricted cone | Port restricted cone | Symmetric |
| Sender |
| Open | I. | II. | |  | III. |
| Full cone |
| Restricted cone |
| Port restricted cone |  |
| Symmetric | IV. | |

1. Mivel a fogadónak nincs NAT-ja, ezért a kapcsolat létre tud jönni.
2. A felek egymásra találhatnak ebben az esetben is, kivéve, ha a fogadó port restricted és a hívó fél szimetrikus NAT mögött van.
3. Egy invitation packet elküldése után felállhat a kapcsolat közöttük.
4. Általában nem megoldható a kapcsolat.

### Session Traversal Utilities for NAT (STUN)

<https://webrtchacks.com/symmetric-nat/>  
<https://tools.ietf.org/id/draft-takeda-symmetric-nat-traversal-00.txt>  
https://en.wikipedia.org/wiki/STUN

Most, hogy megismertük a NAT működését, bemutathatóvá válik a STUN, mint segédeszköz.

Röviden, egy STUN szerver segítségével a NAT mögötti kliens elég információt gyűjthet a rá érvényes NAT beállításokról, így ezt továbbíthatja magáról és kapcsolatot létesíthet más NAT mögött lévő klienssel.

A STUN szerver a TURN szerver egy egyszerű változata. Egyetlen feladata, hogy a kliensnek visszaadja, hogy a NAT-on kívül ő milyen címen és porton érhető el. Ehhez az egyszerű szolgáltatáshoz UDP-t használ, bár a TCP is lehetséges, csak ez megterhelőbb a szerverre nézve több ezer egyidejű lekérés esetén. Ezen kívül támogatott a TCP/TLS is, illetve a DNS szerver is tud szolgálni STUN címmel a megfelelő ’stun’ és ’stuns’ (TLS) rekord lekérése esetén.

Extra funkcionalitás a STUN-ban az authentikáció lehetősége.

Amennyiben a STUN szerver visszaszolgáltatta a kliens publikus ip címét és port számát, onnantól fogva az ICE protokolon múlik, hogy a két kliens egymásra talál-e.

### Traversal Using Relays around NAT (TURN)

A TURN arra az esetre készült, amikor nem lehet ügyeskedni úgy, mint a STUN-nal. Míg STUN esetén a kliensek egymásra találhatnak a publikus címeik alapján, addig szimmetrikus NAT esetén még ez sem lehetséges, az egyetlen kivitelezhető interakció ebben az esetben az, amikor a kliens egy publikus szerver felé kommunikál. Erre szolgál a TURN. Erőforrás igényes, mivel továbbítóként szolgál a két kliens között, azaz minden adat átfolyik rajta. Viszont meg tud bírkózni a szimmetrikus NAT-tal. (táblázat, IV. Rész)

Legelterjedtebb implementációja a coTurn, egy open-source STUN/TURN szerver implementáció. Minden szolgáltatás ezt használja, két okból is:

* C-ben íródott, gyors, alacsony fogyasztású
* A szabványokat erőteljesen betartja

### Interactive Connectivity Establishment (ICE)

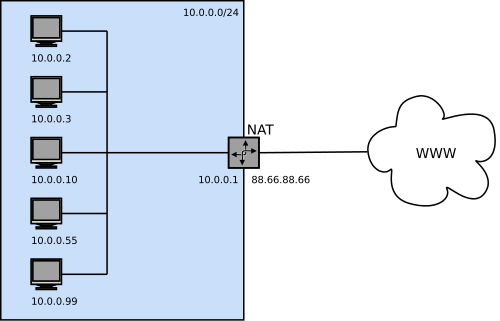
<https://www.slideshare.net/saghul/ice-4414037>

https://www.slideshare.net/rootkiskacsa/stun-turn-pocpilot?next\_slideshow=1

Mielőtt bemutatom az SDP fejlécet, amit a WebRTC kliensek generálnak, meg kell említenem pár fontos protokollt, ami képessé teszi a kommunikációt a lokális hálózaton kívül is.

A WebRTC lényege, hogy a kliensek között nyújtson összeköttetést. (Ennek előnyeiről és megoldásairól fogok írni a WebRTC fejezetben.) Viszont ez a való életben nem olyan egyszerű. Amennyiben lokális hálózaton próbálkoznánk a WebRTC fejlesztésével, nem jutna eszünkbe, hogy rendkívül kényelmes az IP címeket készpénznek venni, mint végpontok amik egy hop alatt eléretőek, de az esetek nagyrészében nem így néz ki az összeköttetés.

Vegyünk egy egyszerű példát. Alice T-Home-ot használ, Bob pedig UPC-t. Alice is NAT mögé van rejtve, Bob is. Nincs egyedi IP cím, ami alapján Alice eljutna Bob-hoz, és fordítva.



Amennyiben nem törődünk ezzel a problémával, legjobb esetben is csak egyirányú kapcsolatot kapnánk, hiszen csak aki nyitja a kapcsolatot, annak lenne nyitva a portja a válasz megkapásához.

Proxy használatával meg lehetne oldani a problémát, viszont a proxy szerverek általában fix portokon működnek, a WebRTC pedig random értékű kiosztást használ.

A problémát meg lehetne oldani kliens oldalon, de szimmetrikus NAT esetén még így se működne egymás megtalálása. Emiatt muszáj egy publikus hálózaton lévő szerver, ami segít a két kliensnek, őket hívják STUN/TURN szervereknek.

Ezek mind félig működő módszerek, azonban az ICE teljes egészében megoldja a NAT állította problémát. Tulajdonságai:

* STUN és TURN szerverek felhasználása
* TURN szerver a worst-case, hiszen ez a feladás kategória
* Csak akkor kezd adatot küldeni az adott session, amikor már teljesen felépült a kapcsolat és biztos nem lesz benne szakadás a NAT miatt
* Rendkívül komplex protokoll, 6 évbe telt, mire RFC fáziba jutott

Az ICE protokoll:

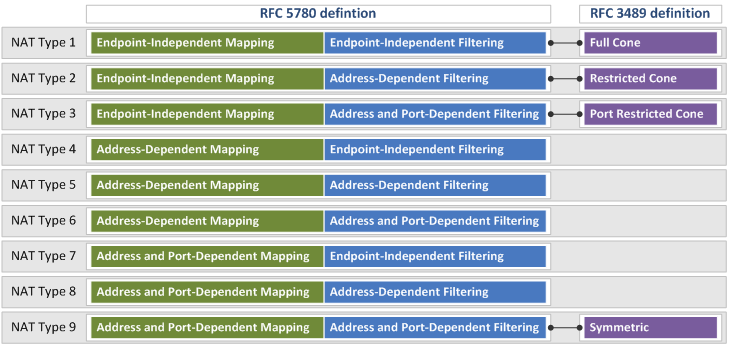
* Allokáció
  + Gyűjtsük össze az összes candidate címet. Ez lehet a gép hálózati interfészei és annak adatai, vagy a STUN/TURN szervertől tanult cím
* Priorizálás
  + Adott candidate milyen prioritást kap a lehetőségek közül
* SDP attribútum: a=candidate:2222700650 1 udp 2113937151 192.168.1.105 53352 typ host generation 0 network-cost 999
  + **2222700650:** unique id
  + **1:** RTP: 1, RTCP: 2
  + **Átvitel módja:** UDP
  + **Prioritás:** 
    - (2122260223>>24)=126 -> host candidate
    - 256-(2122260223&255)=1 -> rtp
  + **IP cím és port**
  + **Típus**
* Miután az offer (SDP header) létre jött, INVITE jelzésű üzenetként küldi el a kliens. A kliens akinek továbbítva lett a header, összegyűjti a saját candidate-jeit, rangsorolja őket és választ generál, ezt is meg kell kapnia a hívás indítójának
* Ellenőrzés
  + Mindkét fél birtokában van az SDP header.
  + Kapcsolat ellenőrőző üzeneteket küldenek egymásnak, ami abból áll, hogy a STUN szervert kérik, hogy nyisson portot nekik a másik gép felé, majd ezt ellenőrzik.

Az ICE portokoll csak akkor alkalmas oda-vissza kapcsolatra, ha maximum az egyik fél van szimmetrikus NAT mögött.

<https://www.slideshare.net/rootkiskacsa/webrtc-hol-tartunk-ma>

https://www.netmanias.com/en/?m=view&id=techdocs&no=6065

Szabványos NAT-Tűzfal átjárás biztosítása



Hole punching:

### Session Description Protocol (SDP)

https://github.com/jbebe/webrtc-thesis/wiki/AllThingsUnsorted

https://webrtchacks.com/sdp-anatomy/

https://www.slideshare.net/saghul/ice-4414037

https://andrewjprokop.wordpress.com/2013/09/30/understanding-session-description-protocol-sdp/

https://en.wikipedia.org/wiki/Session\_Description\_Protocol

https://www.voip-info.org/sdp/

https://andrewjprokop.wordpress.com/2014/07/16/an-introduction-to-webrtc-and-signaling/

Az SDP (session description protocol) célja, hogy multimédia kommunikáció esetén információt nyújtson a másik fél számára a küldő lehetőségeiről, elérhetőségéről. Maga a protokoll nem egy folyamatos oda-vissza üzengetést ír le, sokkal inkább hasonlítható egy egyszeri HTTP fejléc request-response adatcseréjéhez, ahol megadjuk a saját meta-adatainkat illetve a küldött és várt adat típusát. Egy fél, aki előállítja saját SDP header-jét, felsorolja, hogy milyen módon lehet hozzá kapcsoldni és milyen módon tud fogadni, vagy stream-elni média tartalmat.

A fejléc nagyon egyszerűen néz ki, ebben is hasonlít egy kicsit a HTTP-hez. Teljesen plaintext alapú, a kulcsok karakterek, az értékek pedig új sort nem tartalmazó tetszőleges hosszú stringek.

Az SDP 3 fő részből áll; session, időzítés, media description. Session definícióból csak egy lehet, de időzítésből és media description-ből tetszőlegesen sok.

**SDP fejléc**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Az SDP fejléc első és legfontosabb része a session data. Ennek a következő kulcsai vannak:   |  |  | | --- | --- | | v=0 | Protokoll verziója. Jelenleg csak a 0 érték elfogadott. | | o=- 6037493519144602100 2 IN IP4 127.0.0.1 | Az első érték maga a session id. A 2-es szám a session verziójára utal, tehát ha később változtatunk a stream-en, például kikapcsoljuk a mikrofont, az újragenerált fejlécben 3-as érték fog szerepelni. Az IN az internetet jelenti, mint szállító eszköz, az IP4 természetesen az IPv4 protokollra, az utolsó ip cím pedig a gép ip címe. (Ez az utolsó három adat nem fog kelleni a WebRTC kapcsolat felépítéséhez, mivel ahhoz az ICE protokollt használjuk.) | | s=- | Az s kulcshoz a session szövegesen megadott neve tartozik, amit láthatóan nem használnak ki az implementációk. | | t=0 0 | A t a time-ra utal, ez a kulcs adja meg, hogy milyen időintervallumon valid a session. Láthatóan 0 időpillanattól 0-ig, ami jelenthetné azt, hogy soha nem érvényes, de jelen esetben ez azt jelenti, hogy tetszőleges időre szól a session. | | a=group:BUNDLE audio video data | Az a kulcs attribútumot jelent, ezen belül is a bundle group több stream együttes átvitelét jelzi, melyeket külön-külön később részletez a saját leírója. Mint látható audio, video és data stream-ek átvitelére való képességet jelez ami pont elegendő egy valós idejű chat-eléshez. | | a=msid-semantic: WMS lXRsYHOZTANYJUkp60Qfmeyc4eKkAwsaljaQ | Ez az attribútum egy egyedi azonosítót rendel a WMS-hez (WebRTC Media Stream). A stream-ek ezzel lesznek azonosítva. | |
| A következő (opcionális) rész a különböző médiatípusokat írja le:   |  |  | | --- | --- | | m=application 53350 DTLS/SCTP 5000 | Ez az attribútum írja le, hogy datastream is tartozik a session-höz, mégpedig az 53350-es porton DTLS titkosítással és SCTP protokollon keresztül. | | c=IN.... a=candidate... a=ice-... | Ezek az attribútumok hasonlóak az audio és video leírásokhoz. | |
| A következő (opcionális) rész az audio csatornát írja le:   |  |  | | --- | --- | | m=audio 58779 UDP/TLS/RTP/SAVPF 111 103 104 9 0 8 106 105 13 126 | Ez az attribútum írja le, hogy audio csatornát tud nyitni a session az adott porton. UDP-n keresztül, TLS titkosítással. Az RTP/SAVPF pedig az SRTP, SRTCP és RTCP protokollok használatát írja elő. | | c=IN IP4 217.130.243.155 | Ez a sor adja meg, hogy honnan várod az audio kapcsolatot. Mivel a pontos címet az ICE protokoll t | | a=rtcp:51472 IN IP4 217.130.243.155 | Ez a sor írja le az rtcp kapcsolatot. (A port megegyezik az SRTP porttal, mivel ezek multiplexálhatóak.) | | a=candidate:1467250027 1 udp 2122260223 192.168.0.196 46243 typ host generation 0  a=candidate:1467250027 2 udp 2122260222 192.168.0.196 56280 typ host generation 0 | Az ICE jelölt adatai. Rengeteg ilyen attribútum pár lehet az SDP fejlécben, mindegyik egy lehetséges csatornát ír le, melyen keresztül felépülhet a kapcsolat Az első komponens ahol 1-es szám van, az RTP protokollra vonatkozik, a 2-es szám pedig az RTCP-re. | | a=ice-ufrag:Oyef7uvBlwafI3hT  a=ice-pwd:T0teqPLNQQOf+5W+ls+P2p16 | E két attribútum a kapcsolat titkosságát hivatott erősíteni, tehát csak az adott jelszó birtokában kapcsolódhat a másik fél a hoszthoz. | | a=fingerprint:sha-256 49:66:12:17:0D:1C:91:AE:57:4C:C6:36:DD:D5:97:D2:7D:62:C9:9A:7F:B9:A3:F4:70:03:E7:43:91:73:23:5E  a=setup:actpass | Ezek az attribútumok a DTLS kapcsolathoz kellenek, a hash pedig a cert alapján készül. | | a=mid:audio | Ez az attribútum a bundle-ben megnevezett csatornára utal, tehát a média leírását köti a globális konfigban lévő hivatkozáshoz. | | a=extmap:1 urn:ietf:params:rtp-hdrext:ssrc-audio-level  a=extmap:3 http://www.webrtc.org/experiments/rtp-hdrext/abs-send-time | Ezek az attribútumok az RTP header-t egészítik ki a megadott módon metaadatokkal, tehát a másik félnek is támogatnia kell ezeket a kiegészítéseket. | | a=sendrecv | Fontos attribútum: ez jelzi, hogy a header kiállítója az adott csatornát küldi, fogadja, mindkettő, vagy egyik sem. Például egy előadás esetén a stream-et sendonly és recvonly módban érdemes használni. | | a=rtcp-mux | Ez a sor írja elő, hogy az RTCP és RTP kapcsolat egy porton keresztül fog menni. | | a=rtpmap:111 opus/48000/2 | Ez a sor az opus codec-et írja elő az átvitelhez. Mivel ingyenes és szabadon felhasználható, ezért ez az egyik legelterjedtebb codec. | |
| A video kapcsolat hasonló módon van előírva, csak ott természetesen más codec-eket használ a WebRTC. |

### WebRTC

<https://www.slideshare.net/rootkiskacsa/webrtc-hol-tartunk-ma>

Különböző definíciói vannak a WebRTC-nek.

* Protokoll nagy felbontású, sok képességgel rendelkező RTC alapú programok fejlesztéséhez minden platformra.
* Framework, protokollok és API-k összessége ami lehetőséget nyújt audio video és adatátvitelhez.
* Különböző platformok számára nyújt RTC képességet. A WebRTC ennek optimalizált megoldása.

Előnyei:

* Nem kellenek hozzá plugin-ek
* Nem kell hozzá telepítés
* Peer 2 peer működés
* Tűzfalon való átjárás
* WebAPI elrejti a komplexitást (A WebAPI-t pedig a lib-ek rejtik el)
* Authentikáció, Biztonság
* Implementáció elérhető nem csak a böngészőben, de mobilokra is, natívan is.
* Tetszőleges signaling protokoll
* Alacsony késleltetés (UDP)
* Torlódásvezérlés

Érdekességek:

* 2016-2017: 45% növekedés használatban
* 1300-nál több komoly projekt
* Böngészők 80%-a WebRTC ready

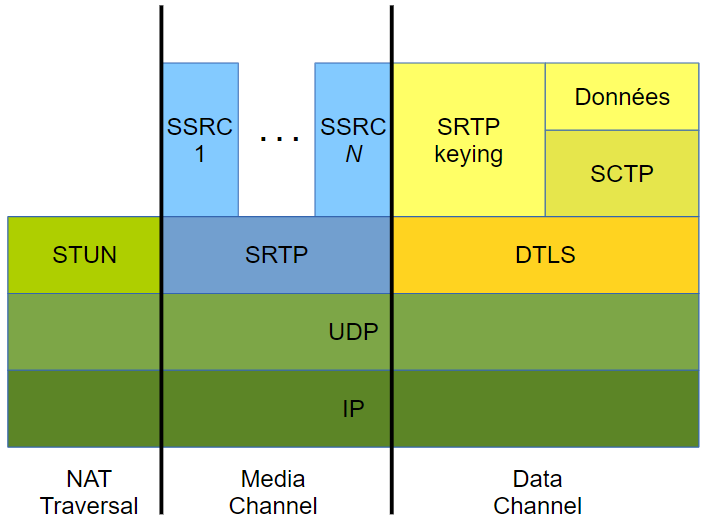


Figure 5. NAT-hoz tartozó protokoll, kép és hang küldéséhez használt stack, adat küldéshez használt stack



Figure 6. Mindez részletesebben

Idővonal:

* Google-nél megszületik az ölet
* 2010 nyarán Google Microsoft Apple Mozilla Skype Ericcson összeül
* 2011-ben már Chrome támogatás
* 2013 – Firefox támogatás
* 2015 – Edge
* 2017 – feature freeze, 1.0 szabvány

Kodekek:

* Opus és G711
* VP8 vs. H.264
* AfOM

Titkosítás

* SRTP-DTLS
* Személyazonosság

Gateway

* RTPEngine
* Janus
* Kurento
* FreeSwitch
* JITsi gateway to SIP

### Lehetőségek

#### Média

#### Adatfolyam

## WebRTC

https://www.slideshare.net/rootkiskacsa/webrtc-puzzle

Adapter.js

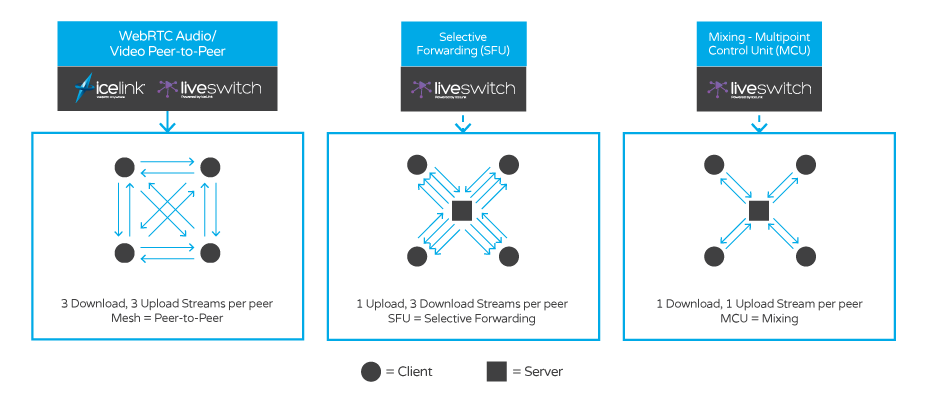
* Implementációs hézagok áthidalása
* Szabványosítás, implementáció

### Egy-egy

### Egy-több

#### Csillag topológia

### Több-több



#### Teljes gráf

Mivel a WebRTC egyéb segítség nélkül pont-pont kapcsolatot tud csak kiépíteni az SDP segítségével, ezért egy több szereplős kommunikáció esetén teljes gráfot kell kiépítenie a kliensek között. Ez 3 résztvevő esetén még nem túl nagy overhead, mivel csak 2 stream-et küldesz és fogadsz, de ennél több kliens kiszolgálása már komoly megterhelést jelent a hálózat számára.

#### Selective Forwarding Unit (SFU)

* JITSI video bridge
* Kurento Media Server
* Janus
* MediaSoup
* Licode Erizo

#### Multipoint Control Unit (MCU)

* FreeSwitch
* Kurento Media Server
* Medooze MCU Media Server
* Licode Erizo

https://www.slideshare.net/Dialogic/beyond-the-mcu

#### SFU vs. MCU

|  |  |
| --- | --- |
| * Nem kódol médiát * Gyorsabb reakcióidő * Simulcast ready | * Kódol * Nagyobb kompatibilitás |

https://bloggeek.me/webrtc-multiparty-video-alternatives/

#### Scalable Video Coding

<http://info.vidyo.com/rs/vidyo/images/WP-Vidyo-SVC-Video-Communications.pdf>

#### Last N

<https://jitsi.org/wp-content/uploads/2016/12/nossdav2015lastn.pdf>

## Létező megoldások

# Elkészült rendszer bemutatása

[Rövid szöveg ahol elmondod, hogy most hogy bemutattad a WebRT tehnológiát, kitrész arra amit te csinálsz ennek a felhasználásával.]

## Célkitűzések

[Ide írd le, hogy proof-of concepteket csinálsz hogy tesztelj dolgokat.]

## Tervezési szempontok

[Milyen szmepontokat vettél figyelembe, lehet bullshitellni, hogy „gyors iteráció”, megbízhatóság, támogatottság, stb-]

## Egy-egy kapcsolatú demó

### Megvalósítás

A ‘proof-of-concept’ 3 minimalista oldalt rejt. Egyrészt a főoldalt ahol választhatunk aközött, hogy mi akarjuk létrehozni a csatornát, vagy valaki már létrehozta, csak csatlakozni akarunk hozzá, másrészt e két aloldalt.

Miután kiválasztottuk a chat létrehozása opciót, az oldalon több lehetőségünk is van. A logger ablakban láthatjuk az összes WebRTC-vel kapcsolatos eseményt, ez az oldal felső részén helyezkedik el. Mellette kissé balrább a nyílra kattintva visszaléphetünk a főoldalra. Miközben az oldal nyitva van, a log alatt a bal panel láthatóan tölt. Ennek az az oka, hogy az SDP header még nincsen kész. Ahhoz, hogy a proof of concept oldalon elég legyen egy header-t átküldeni a másiknak, megvárom míg a candidate keresés - a kapcsolat felépítésének legtöbb időt igénylő része – befejeződik. Mellesleg a leghamarább akkor tudunk arról, hogy több candidate nem lesz, ha kapunk egy olyan eseményt OnNewICECandidate esetén, ahol a paraméter null. Tehát amikor már az összes candiate rendelkezésre áll, az SDP header megjelenik a bal panelen. Ahhoz, hogy felhasználjuk, nem kell kijelölni, elég ha egyszerűen a “Copy SDP Header” gombra kattintunk. Ekkor a teljes tartalma a vágólapra másolódik, azzal a kiegészítéssel, hogy ez a kimásolt változat már json formátumú adat. A headert amit kaptunk most el kell küldenünk egy másik kliensnek aki szintén megnyitotta az oldalt, csak a “Join chat” gombra kattintott. Neki ugyanezen oldalon nem töltődik be megnyitás esetén a bal panelen a header, mert előbb szüksége van az offer típusú headerre, amit éppen most generáltunk. Tehát valamilyen médiumon keresztül eljuttatjuk az információt a másik félnek, aki az “Add client’s SDP Header” gombra kattintva beimportálja azt. Lokális hálózaton STUN szerver nélkül ez a header egészen kicsi, ezért nem érdemes külön fájlként kezelni, ezért döntöttem a szöveges másolás mellet. Természetesen több hálózati kártyával illetve STUN szerver segítségével az SDP header egészen nagy lehet amit már nem lehet könnyen kezelni, például átküldeni egy chat programon.

Miután beimportálta a fogadó fél a headert, az ő WebRTC kliense kezd el ICE candidate-eket keresni. Ez is eltarthat egy ideig, végül megszületik az “ANSWER” típusú SDP header. Ezt szintén kimásoljuk, ahogy előbb is tettük, és átadjuk a chatet indító félnek. Ezzel elvégeztük a ránk eső feladatot. Ekkor az indító fél az answer headert az “Add client’s SDP Header” gombra kattintva véglegesíti az információcserét (signaling). Ekkor az indító fél oldaláról indulva kiépül a csatorna a két fél között. A második webkamera kép is megjelenik és a felek tudnak chatelni.

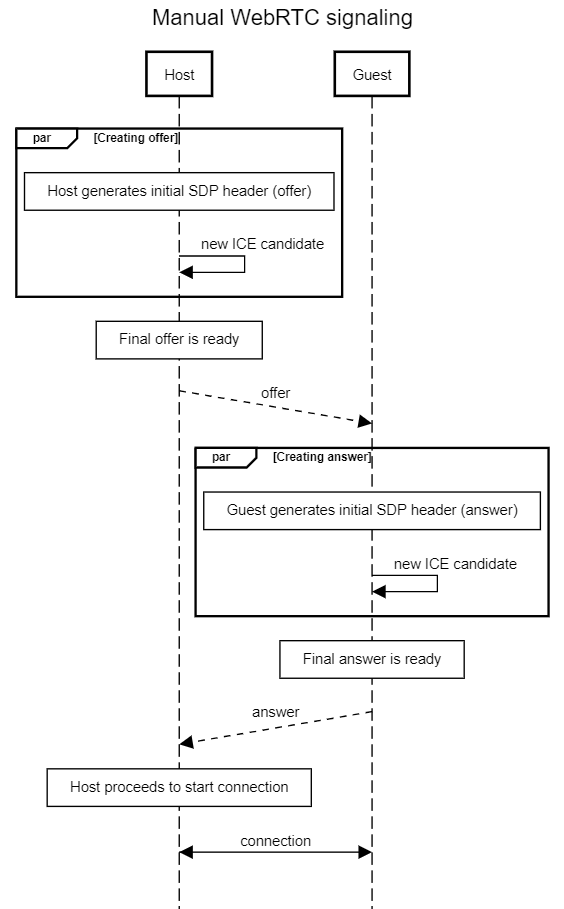


Figure 7. Kapcsolat kiépülése a PoC demóban

### WebRTC könyvtár

A WebRTC könyvtár három részből áll. Van egy base és külön a host és client, aszerint hogy ki hív és ki válaszol.

A Base gondoskodik arról, hogy a datachannel, videó, audio kapcsolatok létrejöhessenek, illetve kezeli az ICE candidate keresést. A Host és Client osztályok már csak a base-ből származnak le és használják fel a képességeit, hogy a saját specifikus függvényeiket ezzel együtt hajthassák végre. A Host a datachannel létrehozásáért felel, illetve az offert állítja elő, míg a Client már csak event-ként kapja meg a datachannelt és az answer-t is csak az offert megléte esetén állíthatja elő.

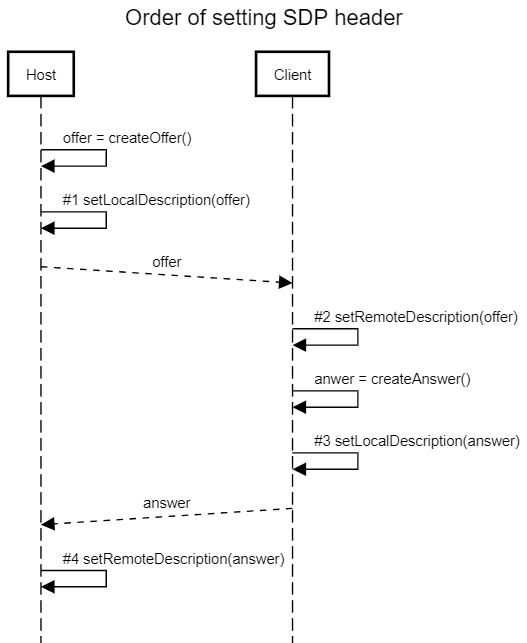


Figure 8. az SDP header beállításának sorrendje

Ennek a könyvtárnak a megírásánál még nem figyeltem a komplexitásból eredő problémákra, így ES5 nyelvi készlettel írtam, nem használva ki teljesen az újabb JS verziók adta lehetőségeket. Például nem megfelelően használtam a Promise-t, nem await-eltem nagyon semmit, helyette mindig a .then() metódust használtam ami nem segített érthetőbbé tenni a kódot.

### Fájl küldés

Egyetlen módja a fájlok küldésének WebRTC esetén a datachannel használata. Datachannelt használhatunk bármire, hiszen a fogadott paramétertípusok a következőek: string, Blob, ArrayBuffer, ArrayBufferView. Én a PoC esetén egy ennél egyszerűbb módszert használtam, mégpedig azt, hogy egy bizonyos jelzést hagytam az üzenet elején ami alapján el lehet dönteni, hogy az éppen fájl küldés lesz, vagy csak szöveges üzenet. Természetesen használhattan volna akkor már sorosított json-t, de nem tettem. Későbbi tervként szerepel ennek implementálása.

### UI

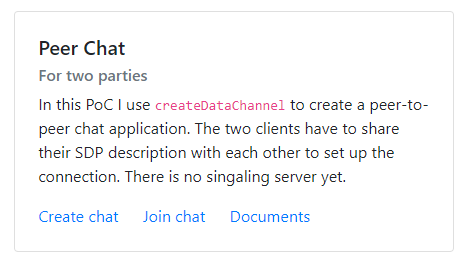


Figure 9. Kezdő képernyő

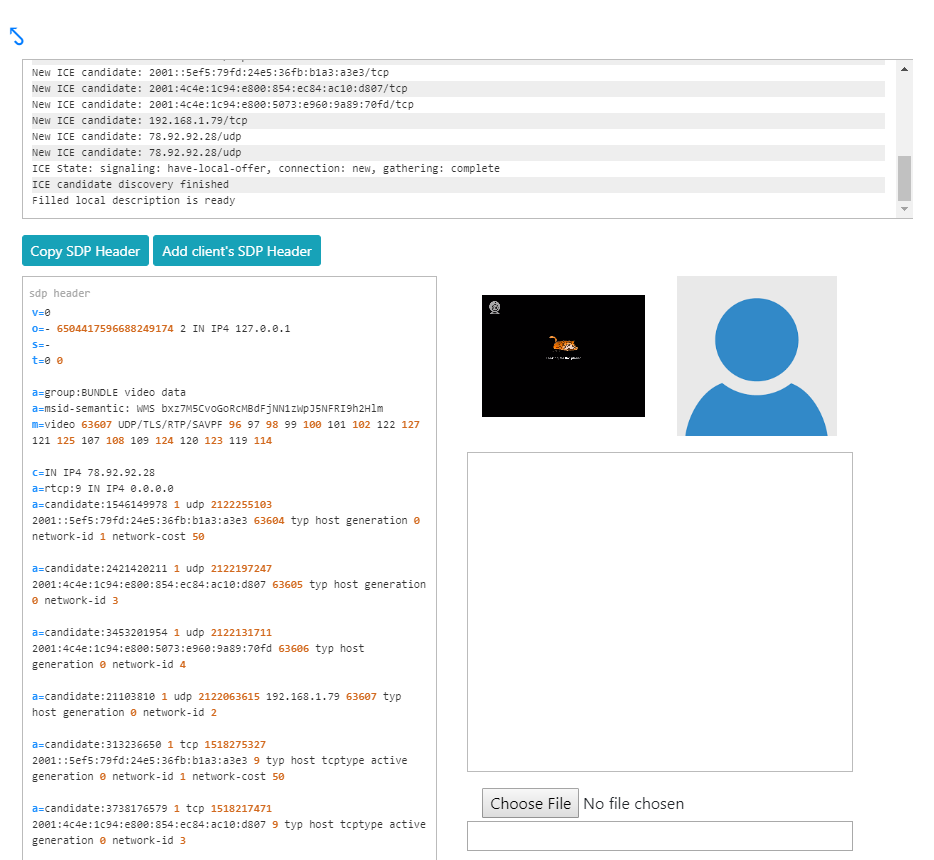


Figure 10. Host SDP létrehozása

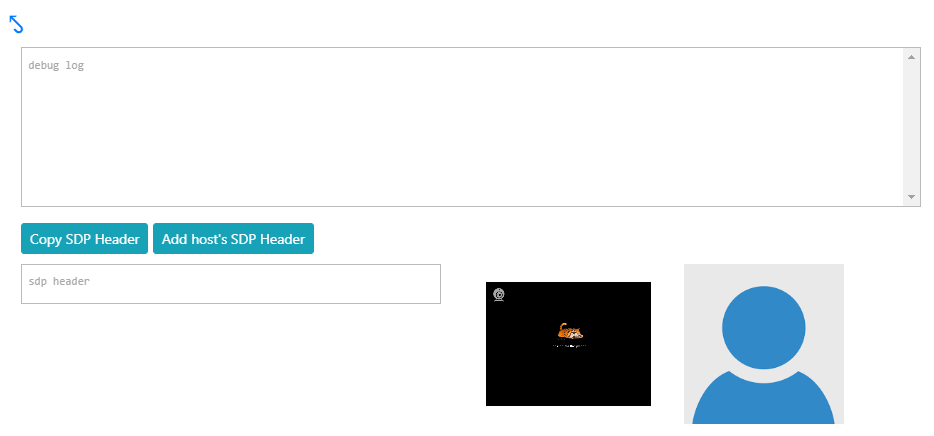


Figure 11. Guest oldal mielőtt hozzáadnánk az offert

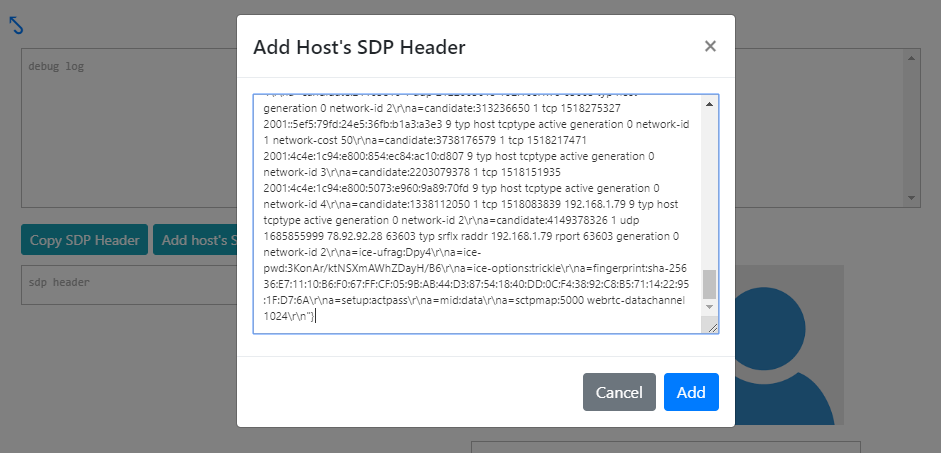


Figure 12. SDP header hozzáadása guest oldalon



Figure 13. Létrejött kapcsolat

### Összefoglaló

A proof of concept megírása segített megérteni a WebRTC működését. Az volt a tervem, hogy ezután már előre megírt WebRTC könyvtárat fogok használni, de később kiderült, hogy a sajátomat kellett volna folytatnom, nem nyerek máséval, sőt, nehezebb lesz kiegészíteni extra funkciókkal.

## Több-több kapcsolatú webapp

### Bevezető

Ahhoz, hogy a megírt ‘proof of concept’ demó működhessen a kényelmetlen SDP header manuális cseréje nélkül, egy szerverre van szükség. Mint azt a WebRTC bemutatásánál is említettem, a WebRTC ránk bízza a signaling-ot, ami annyit tesz, hogy az SDP cserét nekünk kell lebonyolítanunk. Ezen kívül ha már egy szerver is képbe került, session-öket is tarthatunk számon, különböző felhasználókat, azok authentikációját, stb. Magyarán egy szerverrel az előbbi demó lehetőségeit szabadon kibővíthetjük kényelmi funkciókkal, melyek egy akár eladható termék főbb feature-ei is lehetnek.

### Első verzió

#### Szerver

Az első változat egy Node.js alapú websocket szerver volt TypeScript nyelven írva. Két feladata van, egyrészt az SDP cserét jelzi a kapcsolódó felek felé, másrészt a felhasználók státuszát, elérhető felhasználók listáját küldi vissza.

Az architektúra nagyon egyszerű. Egyrészt van egy statikus fájl kiszolgáló, ami az SPA-t tölti be, illetve maga a signaling server ami websocketre épül. Mind a fájl kiszolgáló, mind a WS könyvtár az express.js http szerver könyvtárat használja. Websocket-hez a socket.io könyvtárat használtam.

A különböző WS kéréseket egy egyszerű dispatcher kezeli:

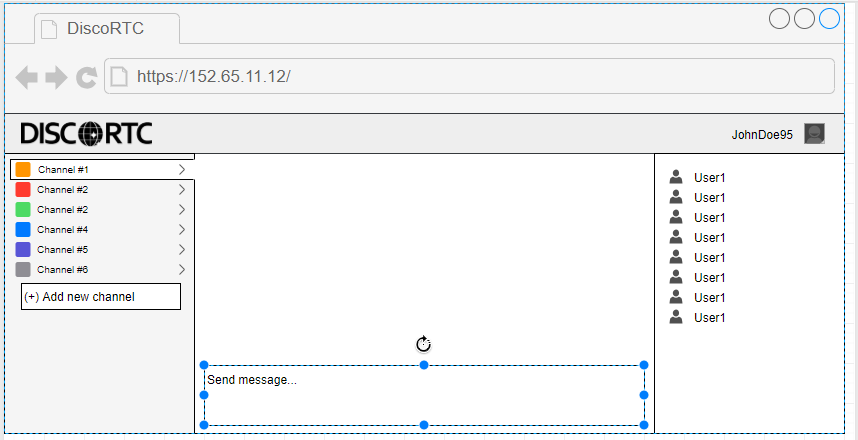
|  |
| --- |
| (router[clientMessage.type] || onError).call(router); |

Ha létezik a bejövő üzenet típusában megadott metódusnév, akkor az fut le, ha nincs ilyen, akkor pedig az onError kezeli le. Az éppen aktív felhasználókat egy User[] member tárolja, ezért a szerver leállásával elveszik minden eddigi adat.

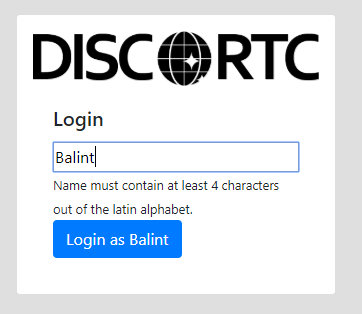
Jó döntésnek tekinthető a Node.js abban a tekintetben, hogy ugyanazon a nyelven íródott, mint a kliens, ezért nem olyan nehéz a context switch. Ezen kívül egy közös fájlban tárolható minden típus amit a szerver és a kliens használ ami megkönnyíti a fejlesztést és kisebb az esélye a hibázásnak. És végül azért is jó, mert a Node.js jól kezeli a gyors aszinkron műveleteket, amiből a WS szerver miatt éppen sok van.

#### Kliens

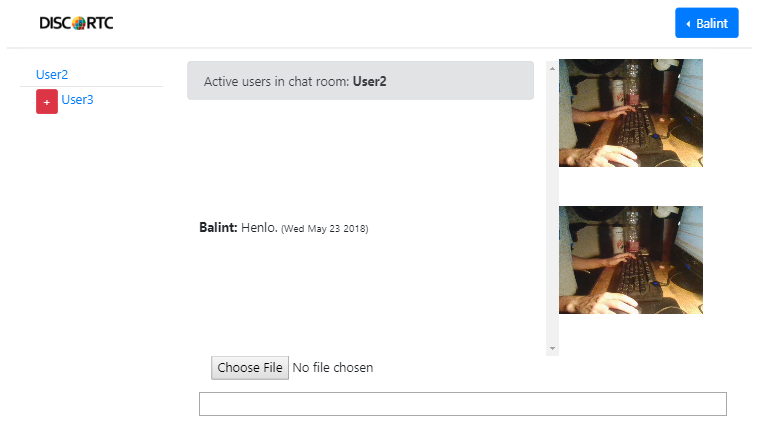
Az első kliens terv kürülbelül ez volt:



A terv szerint első közelítésben mindenki egyedi felhasználóként beszélhet a másikkal, tehát csatornák létrehozására nincs lehetőség. Miután kiválasztunk egy aktív felhasználót beszélgetéshez, opcionálisan behívhatunk a csoportba másokat is. Mivel ez az implementáció nem használja fel az SFU/MCU nyújtotta lehetőségeket, ezért minél több embert hívunk be a beszélgetésbe, annál jobban nő az adatforgalom, ami már 3 emberrel való beszélgetés esetén is leterhel egy egyszerű számítógépet.



Kliens oldalon az Angular 5 framework-öt használtam, így nem is volt lehetőség dönteni, hogy TypesScript-ben, vagy anélkül fogjak neki a kódoláshoz.



Két oldal van, a login és a chat. A login oldalon adódik, hogy az egyedi felhasználónevet ellenőrzi a kliens, majd regisztrál minket az aktív chat-elők között, míg a chat oldalon a felhasználókat választhatjuk ki és beszélgethetünk velük.

Ahogy azt az Angular-nél megszokhattunk, az üzleti logikát service-ek implementálják. Itt sincs másképp, egy service kezeli a WebRTC kapcsolatokat, illetve egy service kezeli a WebSocket kapcsolatot. A WebSocket-et a socket.io-client könyvtár segíti, míg a WebRTC kapcsolatot a simple-peer.

#### Összegzés

A WebRTC hátránya, hogy nagyon sok esemény keletkezik és dolgozódik fel egy event handler-ben. Ezeket pedig összefogni nagyon átgondolt struktúrát követel meg. Az első verzió elvetésének oka nem is a szerver hiányosságai, hanem inkább a kliens oldal kezelhetetlensége, annak refaktorálásának reménytelensége.

### Második verzió

#### Szerver

Újra terveztem a teljes backend architektúrát. Először is az általam – szintén – jól ismert C# nyelvre váltottam, illetve komolyabban felülvizsgáltam, hogy mi kellhet szerver oldalon. Először is több különálló szervert hoztam létre. Ezek teljesen állapot mentesek, leszámítva a WebSocket framework-öt, mivel a SignalR eltárolja az adott kliens adatait azonosításhoz. A signaling külön lett választva az API-tól, így most már két külön szerverként futnak. A userekről az adatokat egy Redis cache tárolja, így végre perzisztens az információ róluk. A WebSocket framework-öt a SignalR nyújtja, ami sokkal egyszerűbbé teszi a kommunikációt mind szerver, mind kliens oldalon. Tehát röviden a változás az ASP.NET stack, SignalR, Redis és decoupling.

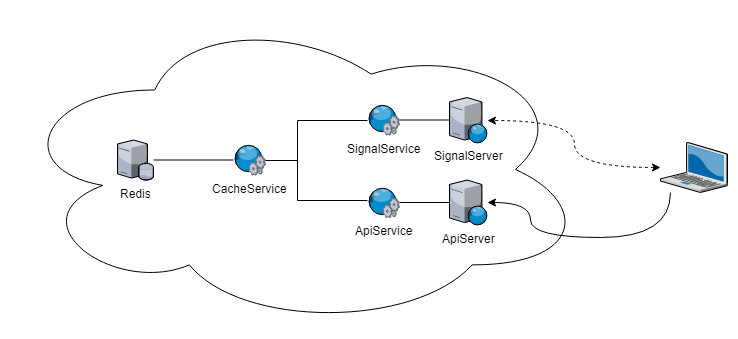


Figure 14. Robosztusabb szerver architektúra

#### Kliens

A kliens az előző változathoz hasonlóan Angular 6-ban lett megírva. Nem foglalkoztam a megjelenítéssel, csak azt a célt tűztem ki magam elé, hogy olyan klienst készítsek, ami megbízhatóan működik,

# Elkészült rendszer értékelése

## Ellenőrzés

## Teljesítmény

## Szerzett tapasztalatok

# Összefoglaló

# Függelék