|  |
| --- |
| TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI  **VIỆN ĐIỆN TỬ - VIỄN THÔNG**  logo_128  BÁO CÁO  **ĐỒ ÁN TỐT NGHIỆP**  **Đề tài:**  **THIẾT KẾ, XÂY DỰNG**  **VIDEO STREAMING SERVER,**  **HỖ TRỢ VIDEO CALL ĐA NỀN TẢNG**  Sinh viên thực hiện: VŨ HỒNG VIỆT  Lớp ĐT01 – K59  Giảng viên hướng dẫn: TS. PHẠM DOÃN TĨNH  Hà Nội, 01 - 2021 |

|  |
| --- |
| TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI  **VIỆN ĐIỆN TỬ - VIỄN THÔNG**  logo_128  BÁO CÁO  **ĐỒ ÁN TỐT NGHIỆP**  **Đề tài:**  **THIẾT KẾ, XÂY DỰNG**  **VIDEO STREAMING SERVER,**  **HỖ TRỢ VIDEO CALL ĐA NỀN TẢNG**  Sinh viên thực hiện: VŨ HỒNG VIỆT  Lớp ĐT01 – K59  Giảng viên hướng dẫn: TS. PHẠM DOÃN TĨNH  Cán bộ phản biện:  Hà Nội, 01 - 2021 |

**ĐÁNH GIÁ ĐỒ ÁN TỐT NGHIỆP**

(Dùng cho giảng viên hướng dẫn)

Tên giảng viên đánh giá: **TS. Phạm Doãn Tĩnh**

Họ và tên sinh viên: **Vũ Hồng Việt** MSSV: **20145255**

Tên đề tài: **Thiết kế, xây dựng Video Streaming Server, hỗ trợ video call đa nền tảng**

**Chọn các mức điểm phù hợp cho sinh viên trình bày theo các tiêu chí dưới đây:**

Rất kém (1); Kém (2); Đạt (3); Giỏi (4); Xuất sắc (5)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Có sự kết hợp giữa lý thuyết và thực hành (20)** | | | | | | |
| 1 | Nêu rõ tính cấp thiết và quan trọng của đề tài, các vấn đề và các giả thuyết (bao gồm mục đích và tính phù hợp) cũng như phạm vi ứng dụng của đồ án | 1 | 2 | 3 | 4 | 5 |
| 2 | Cập nhật kết quả nghiên cứu gần đây nhất (trong nước/quốc tế) | 1 | 2 | 3 | 4 | 5 |
| 3 | Nêu rõ và chi tiết phương pháp nghiên cứu/giải quyết vấn đề | 1 | 2 | 3 | 4 | 5 |
| 4 | Có kết quả mô phỏng/thực nghiệm và trình bày rõ ràng kết quả đạt được | 1 | 2 | 3 | 4 | 5 |
| **Có khả năng phân tích và đánh giá kết quả (15)** | | | | | | |
| 5 | Kế hoạch làm việc rõ ràng bao gồm mục tiêu và phương pháp thực hiện dựa trên kết quả nghiên cứu lý thuyết một cách có hệ thống | 1 | 2 | 3 | 4 | 5 |
| 6 | Kết quả được trình bày một cách logic và dễ hiểu, tất cả kết quả đều được phân tích và đánh giá thỏa đáng | 1 | 2 | 3 | 4 | 5 |
| 7 | Trong phần kết luận, tác giả chỉ rõ sự khác biệt (nếu có) giữa kết quả đạt được và mục tiêu ban đầu đề ra đồng thời cung cấp lập luận để đề xuất hướng giải quyết có thể thực hiện trong tương lai | 1 | 2 | 3 | 4 | 5 |
| **Kỹ năng viết quyển đồ án (10)** | | | | | | |
| 8 | Đồ án trình bày đúng mẫu quy định với cấu trúc các chương logic và đẹp mắt (bảng biểu, hình ảnh rõ ràng, có tiêu đề, được đánh số thứ tự và được giải thích hay đề cập đến; căn lề thống nhất, có dấu cách sau dấu chấm, dấu phảy v.v.), có mở đầu chương và kết luận chương, có liệt kê tài liệu tham khảo và có trích dẫn đúng quy định | 1 | 2 | 3 | 4 | 5 |
| 9 | Kỹ năng viết xuất sắc (cấu trúc câu chuẩn, văn phong khoa học, lập luận logic và có cơ sở, từ vựng sử dụng phù hợp v.v.) | 1 | 2 | 3 | 4 | 5 |
| **Thành tựu nghiên cứu khoa học (5)** *(chọn 1 trong 3 trường hợp)* | | | | | | |
| 10a | Có bài báo khoa học được đăng hoặc chấp nhận đăng/Đạt giải SVNCKH giải 3 cấp Viện trở lên/Có giải thưởng khoa học (quốc tế hoặc trong nước) từ giải 3 trở lên/Có đăng ký bằng phát minh, sáng chế | 5 | | | | |
| 10b | Được báo cáo tại hội đồng cấp Viện trong hội nghị SVNCKH nhưng không đạt giải từ giải 3 trở lên/Đạt giải khuyến khích trong các kỳ thi quốc gia và quốc tế khác về chuyên ngành (VD: TI contest) | 2 | | | | |
| 10c | Không có thành tích về nghiên cứu khoa học | 0 | | | | |
| **Điểm tổng** | | **/50** | | | | |
| **Điểm tổng quy đổi về thang 10** | |  | | | | |

***Nhận xét khác*** *(về thái độ và tinh thần làm việc của sinh viên)*

|  |  |
| --- | --- |
|  | Ngày: … / 01 / 2021  **Người nhận xét**  (Ký và ghi rõ họ tên) |

**ĐÁNH GIÁ ĐỒ ÁN TỐT NGHIỆP**

(Dùng cho cán bộ phản biện)

Giảng viên đánh giá:

Họ và tên sinh viên: **Vũ Hồng Việt** MSSV: **20145255**

Tên đề tài: **Thiết kế, xây dựng Video Streaming Server, hỗ trợ video call đa nền tảng**

**Chọn các mức điểm phù hợp cho sinh viên trình bày theo các tiêu chí dưới đây:**

Rất kém (1); Kém (2); Đạt (3); Giỏi (4); Xuất sắc (5)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Có sự kết hợp giữa lý thuyết và thực hành (20)** | | | | | | |
| 1 | Nêu rõ tính cấp thiết và quan trọng của đề tài, các vấn đề và các giả thuyết (bao gồm mục đích và tính phù hợp) cũng như phạm vi ứng dụng của đồ án | 1 | 2 | 3 | 4 | 5 |
| 2 | Cập nhật kết quả nghiên cứu gần đây nhất (trong nước/quốc tế) | 1 | 2 | 3 | 4 | 5 |
| 3 | Nêu rõ và chi tiết phương pháp nghiên cứu/giải quyết vấn đề | 1 | 2 | 3 | 4 | 5 |
| 4 | Có kết quả mô phỏng/thực nghiệm và trình bày rõ ràng kết quả đạt được | 1 | 2 | 3 | 4 | 5 |
| **Có khả năng phân tích và đánh giá kết quả (15)** | | | | | | |
| 5 | Kế hoạch làm việc rõ ràng bao gồm mục tiêu và phương pháp thực hiện dựa trên kết quả nghiên cứu lý thuyết một cách có hệ thống | 1 | 2 | 3 | 4 | 5 |
| 6 | Kết quả được trình bày một cách logic và dễ hiểu, tất cả kết quả đều được phân tích và đánh giá thỏa đáng | 1 | 2 | 3 | 4 | 5 |
| 7 | Trong phần kết luận, tác giả chỉ rõ sự khác biệt (nếu có) giữa kết quả đạt được và mục tiêu ban đầu đề ra đồng thời cung cấp lập luận để đề xuất hướng giải quyết có thể thực hiện trong tương lai | 1 | 2 | 3 | 4 | 5 |
| **Kỹ năng viết quyển đồ án (10)** | | | | | | |
| 8 | Đồ án trình bày đúng mẫu quy định với cấu trúc các chương logic và đẹp mắt (bảng biểu, hình ảnh rõ ràng, có tiêu đề, được đánh số thứ tự và được giải thích hay đề cập đến; căn lề thống nhất, có dấu cách sau dấu chấm, dấu phảy v.v.), có mở đầu chương và kết luận chương, có liệt kê tài liệu tham khảo và có trích dẫn đúng quy định | 1 | 2 | 3 | 4 | 5 |
| 9 | Kỹ năng viết xuất sắc (cấu trúc câu chuẩn, văn phong khoa học, lập luận logic và có cơ sở, từ vựng sử dụng phù hợp v.v.) | 1 | 2 | 3 | 4 | 5 |
| **Thành tựu nghiên cứu khoa học (5)** *(chọn 1 trong 3 trường hợp)* | | | | | | |
| 10a | Có bài báo khoa học được đăng hoặc chấp nhận đăng/Đạt giải SVNCKH giải 3 cấp Viện trở lên/Có giải thưởng khoa học (quốc tế hoặc trong nước) từ giải 3 trở lên/Có đăng ký bằng phát minh, sáng chế | 5 | | | | |
| 10b | Được báo cáo tại hội đồng cấp Viện trong hội nghị SVNCKH nhưng không đạt giải từ giải 3 trở lên/Đạt giải khuyến khích trong các kỳ thi quốc gia và quốc tế khác về chuyên ngành (VD: TI contest) | 2 | | | | |
| 10c | Không có thành tích về nghiên cứu khoa học | 0 | | | | |
| **Điểm tổng** | | **/50** | | | | |
| **Điểm tổng quy đổi về thang 10** | |  | | | | |

***Nhận xét khác của cán bộ phản biện***

|  |  |
| --- | --- |
|  | Ngày: … / 01 / 2021  **Người nhận xét**  (Ký và ghi rõ họ tên) |

LỜI NÓI ĐẦU

Trong những thập niên cuối thế kỷ XX cho đến nay, sự bùng nổ về công nghệ đã đem lại cho đời sống của con người. Công nghệ đã góp phần không nhỏ trong việc cải thiện đời sống. Cùng với sự bùng nổ về công nghệ, việc chia sẻ hình ảnh qua internet là một nhu cầu không thể thiếu trong cuộc sống của con người hiện nay. Từ các trang mạng xã hội như Facebook, Twitter, Zalo, TikTok, Skype, Youtube cho đến các nền tảng xem phim trực tuyến như Netflix,… đều đưa công nghệ chia sẻ hình ảnh, video và đặc biệt là video streaming vào phục vụ người dùng.

Với mong muốn tìm hiểu và thiết kế việc kết nối con người trong thời gian thực, em đã lựa chọn đề tài *Thiết kế, xây dựng Video Streaming Server, hỗ trợ Video Call đa nền tảng* làm đề tài đồ án tốt nghiệp.

Em xin gửi lời cảm ơn đến TS. Phạm Doãn Tĩnh, người thầy đã trực tiếp hướng dẫn em thực hiện đề tài tốt nghiệp trong thời gian qua. Mặc dù nhiều lúc em có lơ là trong việc thực hiện đề tài nhưng thầy vẫn nhiệt tình động viên, hướng dẫn và đồng hành với em.

Tuy đã có gắng nhưng với lượng kiến thức hạn hẹp khó tránh khỏi sai sót mong thầy cô góp ý thêm để em hoàn thiện đề tài hơn. Em xin chân thành cảm ơn!

LỜI CAM ĐOAN

Tôi là Vũ Hồng Việt, mã số sinh viên 20145255, sinh viên lớp ĐT01, khóa K59. Giảng viên hướng dẫn là TS. Phạm Doãn Tính. Tôi xin khẳng định toàn bộ nội dung được trình bày trong báo cáo thực tập tốt nghiệp *Thiết kế, xây dựng Video Streaming Server, hỗ trợ Video Call đa nền tảng* là kết quả của quá trình tìm hiểu và nghiên cứu của tôi. Các dữ liệu được nêu trong báo cáo hoàn toàn là trung thực, phản ánh đúng kết quả do tôi nghiên cứu và tìm hiểu. Mọi thông tin trích dẫn đều tuân thủ các quy định về sở hữu trí tuệ, các tài liệu tham khảo được liệt kê rõ ràng. Tôi xin chịu hoàn toàn trách nhiệm về những nội dung được viết trong đề tài này.

Hà Nội, ngày … tháng 01 năm 2021

**Người cam đoan**

**Vũ Hồng Việt**

MỤC LỤC

[DANH MỤC KÝ TỰ VÀ CHỮ CÁI VIẾT TẮT i](#_Toc61966147)

[DANH MỤC HÌNH VẼ iii](#_Toc61966148)

[DANH MỤC BẢNG BIỂU v](#_Toc61966149)

[TÓM TẮT BÁO CÁO vi](#_Toc61966150)

[ABSTRACTION vii](#_Toc61966151)

[CHƯƠNG 1. TỔNG QUAN 1](#_Toc61966152)

[1.1 Giới thiệu vấn đề cần giải quyết 1](#_Toc61966153)

[1.2 Các phương pháp giải quyết vấn đề 2](#_Toc61966154)

[1.3 Mục đích 2](#_Toc61966155)

[1.4 Phạm vi thực hiện đề tài 3](#_Toc61966156)

[1.5 Tóm tắt cấu trúc của báo cáo 3](#_Toc61966157)

[1.6 Kết luận 3](#_Toc61966158)

[CHƯƠNG 2. CƠ SỞ LÝ THUYẾT 4](#_Toc61966159)

[2.1 Các khái niệm chung 4](#_Toc61966160)

[2.1.1 Streaming video 4](#_Toc61966161)

[2.1.2 Bitrate 4](#_Toc61966162)

[2.1.3 Adaptive Bitrate Streaming 4](#_Toc61966163)

[2.1.4 Codec 5](#_Toc61966164)

[2.2 Chuẩn nén video 6](#_Toc61966165)

[2.2.1 Cơ sở nén vieo 6](#_Toc61966166)

[2.2.2 Tiêu chuẩn nén JPEG 7](#_Toc61966167)

[2.2.3 Tiêu chuẩn nén MPEG 8](#_Toc61966168)

[2.2.4 Tiêu chuẩn nén H265/HEVC 10](#_Toc61966169)

[2.2.5 Tiêu chuẩn H.264 13](#_Toc61966170)

[2.2.6 Tiêu chuẩn VP8 15](#_Toc61966171)

[2.2.7 Giới thiệu Protocol 17](#_Toc61966172)

[2.2.8 Giao thức UDP và TCP 18](#_Toc61966173)

[2.2.9 Giao thức RTMP 19](#_Toc61966174)

[2.2.10 Giao thức RTSP/RTP 20](#_Toc61966175)

[2.2.11 HTTP 21](#_Toc61966176)

[2.2.12 WebRTC 22](#_Toc61966177)

[2.2.13 Độ trễ truyền phát giữa các giao thức 22](#_Toc61966178)

[2.3 NAT 23](#_Toc61966179)

[2.3.1 Giới thiệu 23](#_Toc61966180)

[2.3.2 Hoạt động 24](#_Toc61966181)

[2.3.3 Phân loại 25](#_Toc61966182)

[2.4 Thuật toán xác định loại NAT 27](#_Toc61966183)

[2.4.1 Gửi bản tin xác định địa chỉ IP 27](#_Toc61966184)

[2.4.2 Kiểm tra sự tồn tại của NAT 28](#_Toc61966185)

[2.4.3 Nhận biết loại NAT đang sử dụng 29](#_Toc61966186)

[2.5 Công nghệ WebRTC 29](#_Toc61966187)

[2.5.1 Giới thiệu 29](#_Toc61966188)

[2.5.2 Lịch sử ra đời 30](#_Toc61966189)

[2.5.3 Kiến trúc Web 31](#_Toc61966190)

[2.5.4 Kiến trúc WebRTC 31](#_Toc61966191)

[2.5.5 Đặc điểm của WebRTC 32](#_Toc61966192)

[2.5.6 STUN 33](#_Toc61966193)

[2.5.7 TURN 34](#_Toc61966194)

[2.5.8 SDP 36](#_Toc61966195)

[2.5.9 ICE 37](#_Toc61966196)

[2.6 Kết luận 37](#_Toc61966197)

[CHƯƠNG 3. PHÂN TÍCH THIẾT KẾ HỆ THỐNG 38](#_Toc61966198)

[3.1 Tổng quan hệ thống 38](#_Toc61966199)

[3.2 Yêu cầu hệ thống 39](#_Toc61966200)

[3.2.1 Yêu cầu chứ năng 39](#_Toc61966201)

[3.2.2 Yêu cầu phi chức năng 40](#_Toc61966202)

[3.3 Đặc tả và phân tích hệ thống 40](#_Toc61966203)

[3.3.1 Sơ đồ use case tổng quan 40](#_Toc61966204)

[3.3.2 Sơ đồ use case đăng nhập 41](#_Toc61966205)

[3.3.3 Use case danh sách bạn bè online 41](#_Toc61966206)

[3.3.4 Gọi video một người 42](#_Toc61966207)

[3.3.5 Gọi video nhóm nhiều người 43](#_Toc61966208)

[3.3.6 Class Diagram 44](#_Toc61966209)

[3.3.7 Sequence Diagram 46](#_Toc61966210)

[3.3.8 Cơ sở dữ liệu 47](#_Toc61966211)

[3.4 Ứng dụng 47](#_Toc61966212)

[3.4.1 Công nghệ sử dụng 47](#_Toc61966213)

[3.4.2 Màn hình login 48](#_Toc61966214)

[3.4.3 Danh sách bạn bè 49](#_Toc61966215)

[3.4.4 Màn hình gọi video 50](#_Toc61966216)

[CHƯƠNG 4. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 51](#_Toc61966217)

[4.1 Kết quả đạt được 51](#_Toc61966218)

[4.2 Phương hướng phát triển 51](#_Toc61966219)

[TÀI LIỆU THAM KHẢO 53](#_Toc61966220)

# DANH MỤC KÝ TỰ VÀ CHỮ CÁI VIẾT TẮT

|  |  |
| --- | --- |
| ABR | Adaptive Bitrate Streaming |
| API | Application programming interface |
| AVC | Advanced Video Coding |
| CMAF | Common Media Application Format |
| DASH | Dynamic Adaptive Streaming over HTTP |
| DCT | Discrete Cosine Transform |
| HDS | HTTP Dynamic Streaming |
| HEVC | High Efficiency Video Coding |
| HLS | HTTP Live Streaming |
| HTML | HyperText Markup Language |
| HTTP | HyperText Transfer Protocol |
| ICE | Interactive Connectivity Establishment |
| IETF | Internet Engineering Task Force |
| IP | Internet protocol |
| MPEG | Moving Pictures Expert Group |
| NAE | Network Adaptive Encoding |
| NAL | Network Abstaction Layer |
| NAT | Network address translation |
| OTT | Over-the-top |
| PC | Personal computer |
| RBSP | Raw Byte Sequence Payload |
| SDP | Session Description Protocol |
| SIP | Session Initiation Protocol |
| SRT | Secure Reliable Transport |
| STUN | Session Traversal Utilities for NAT |
| TCP | Transmission Control Protocol |
| TURN | Traversal Using Relays around NAT |
| URI | Uniform Resource Identifier |
| URL | Uniform Resource Locator |
| UDP | User Datagram Protocol |
| VCL | Video Coding Layer |
| WebRTC | Web Real-Time Communication |
| W3C | World Wide Web Consortium |
|  |  |

# DANH MỤC HÌNH VẼ

[Hình 1.1 Tỷ lệ thâm nhập điện thoại thông minh theo tỷ lệ dân số tại Việt Nam 1](#_Toc61966221)

[Hình 2.1 Kiến trúc mã hóa Quad-tree 12](#_Toc61966222)

[Hình 2.2 Cấu trúc bộ mã hóa video H.264 [4] 14](#_Toc61966223)

[Hình 2.3 NAL unit [4] 14](#_Toc61966224)

[Hình 2.4 Mô hình bảy tầng OSI [6] 17](#_Toc61966225)

[Hình 2.5 Giao thức TCP/IP [7] 18](#_Toc61966226)

[Hình 2.6 Cấu trúc gói tin của giao thức RTMP [8] 19](#_Toc61966227)

[Hình 2.7 Header của giao thức RTMP [8] 19](#_Toc61966228)

[Hình 2.8 Giao thức RTSP/RTP 20](#_Toc61966229)

[Hình 2.9 Mô tả hoạt động giao thức HTTP [9] 22](#_Toc61966230)

[Hình 2.10 Độ trễ truyền phát giữa các giao thức 23](#_Toc61966231)

[Hình 2.11 Mô hình hoạt động của NAT [7] 23](#_Toc61966232)

[Hình 2.12 NAT loại 1 25](#_Toc61966233)

[Hình 2.13 NAT loại 2 25](#_Toc61966234)

[Hình 2.14 NAT loại 3 26](#_Toc61966235)

[Hình 2.15 NAT loại 4 26](#_Toc61966236)

[Hình 2.16 Thuật toán nhận diện NAT 27](#_Toc61966237)

[Hình 2.17 Xác định danh tính client 28](#_Toc61966238)

[Hình 2.18 Mô hình kiến trúc WebRTC hình thang [11] 31](#_Toc61966239)

[Hình 2.19 Kiến trúc WebRTC hình tam giác [11] 32](#_Toc61966240)

[Hình 2.20 Tổng quan hoạt động của STUN 34](#_Toc61966241)

[Hình 2.21 Tổng quan hoạt động của TURN 35](#_Toc61966242)

[Hình 3.1 Sơ đồ tổng quan hệ thống 38](#_Toc61966243)

[Hình 3.2 Sơ đồ activity 39](#_Toc61966244)

[Hình 3.3 Sơ đồ use case tổng quan 40](#_Toc61966245)

[Hình 3.4 Biểu đồ class của Signaling server 44](#_Toc61966246)

[Hình 3.5 Biểu đồ class của Turn server 44](#_Toc61966247)

[Hình 3.6 Biểu đồ class của Client 45](#_Toc61966248)

[Hình 3.7 Biểu đồ Sequence của Signaling server 46](#_Toc61966249)

[Hình 3.8 Biểu đồ Sequence của Turn server 46](#_Toc61966250)

[Hình 3.9 Biểu đồ Sequence của Client 47](#_Toc61966251)

[Hình 3.10 Cơ sở dữ liệu 47](#_Toc61966252)

[Hình 3.11 Giao diện màn hình chính 49](#_Toc61966253)

[Hình 3.12 Giao diện danh sách bạn bè 49](#_Toc61966254)

[Hình 3.13 Giao diện màn hình cuộc gọi video 50](#_Toc61966255)

# DANH MỤC BẢNG BIỂU

[Bảng 2.1 Đặc điểm giao thức RTMP [6] 20](#_Toc61966256)

[Bảng 2.2 Đặc điểm giao thức RTSP/RTP [6] 21](#_Toc61966257)

[Bảng 2.3 Đặc điểm giao thức WebRTC [6] 22](#_Toc61966258)

[Bảng 3.1 Các use case 41](#_Toc61966259)

[Bảng 3.2 Use case đăng nhập 41](#_Toc61966260)

[Bảng 3.3 Use case danh sách bạn bè online 41](#_Toc61966261)

[Bảng 3.4 Use case gọi video một người 42](#_Toc61966262)

[Bảng 3.5 Use case gọi video nhiều người 43](#_Toc61966263)

# TÓM TẮT BÁO CÁO

Đề tài của đồ án tốt nghiệp là *Thiết kế, xây dựng Video Streaming Server, hỗ trợ Video Call đa nền tảng*. Mục đích của đề tài là nghiên cứu, tìm hiểu thiết kế và xây dựng cách người dùng trò chuyện video trực tiếp với nhau theo thời gian thực trên các thiết bị có kết nối internet.

Trong thời gian hai tháng, dựa trên những nền tảng công nghệ có sẵn, tôi đã tìm hiểu và xây dựng một hệ thống cho phép người dùng có thể trò chuyện trực tiếp với nhau trong thời gian thực.

# ABSTRACTION

The topic of the graduate project is *Designing, building Video Streaming Server, supporting multi-platform Video Call*. The purpose of the topic is to research, learn to design and build how users can live video chat with each other in real time on devices with internet connection.

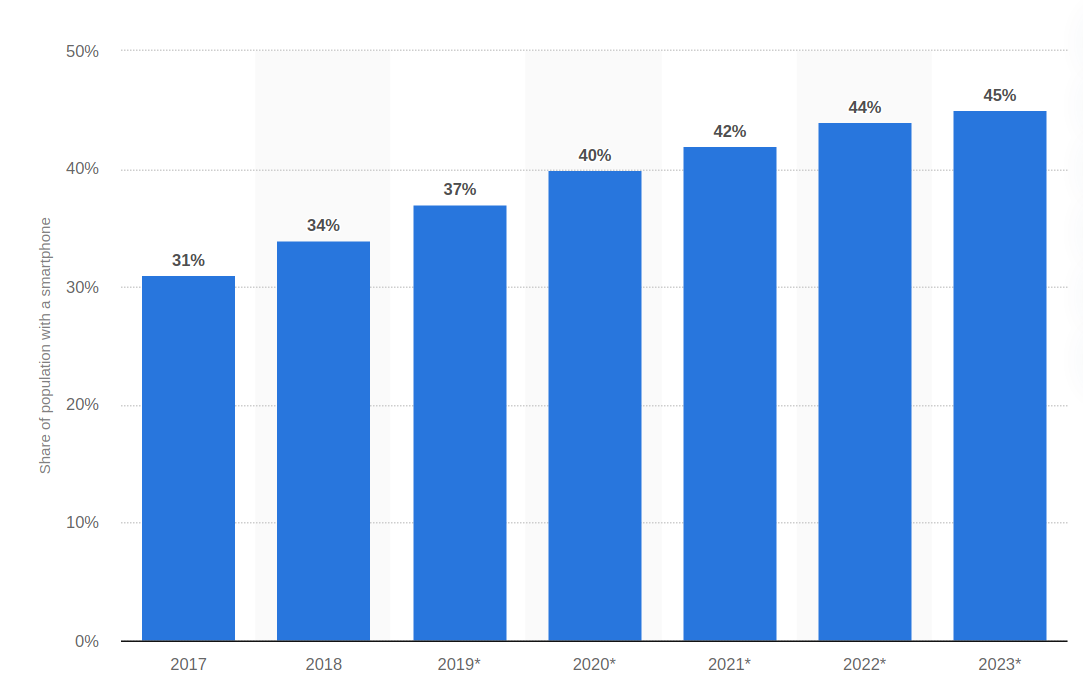
In two months' time, based on available technology platforms, I learned and built a system that allows users to chat directly with each other in real time.

# TỔNG QUAN

Chương 1 trình bày khái quát về đề tài, các phương pháp giải quyết vấn đề và phạm vi thực hiện của đề tài.

## Giới thiệu vấn đề cần giải quyết

Với sự bùng nổ công nghệ thông tin như hiện nay, yêu cầu về chia sẻ hình ảnh qua internet là một nhu cầu không thể thiếu trong đời sống. Hiện nay, số lượng người dùng điện thoại thông minh tại Việt Nam đang tăng trưởng nhanh chóng. Theo báo cáo của Statista [1] về tỷ lệ thâm nhập điện thoại thông minh theo tỷ lệ dân số tại Việt Nam kể từ năm 2017 như sau:



Hình . Tỷ lệ thâm nhập điện thoại thông minh theo tỷ lệ dân số tại Việt Nam

Theo số liệu từ Hình 1.1 cho thấy, tỷ lệ người dùng điện thoại thông minh ở Việt Nam năm 2017 chiếm 31%, năm 2018 là 34%, năm 2019 là 37%, năm 2020 là 40%... xu hướng tang tới 45% vào năm 2023. Từ số liệu thu được cho thấy tỷ lệ người dùng điện thoại thông minh ở Việt Nam đang tăng trưởng rất nhanh. Khi người dùng điện thoại thông minh tăng dẫn đến nhu cầu sử dụng các ứng dụng cũng tăng theo.

Video Streaming (hay còn được gọi là Streaming video) là một dạng truyền thông đa phương tiện được sử dụng rộng rãi trên các thiết bị công nghệ hiện nay. Đặc biệt là các ứng dụng áp dụng hệ thống nhúng đã dùng kỹ thuật streaming video để hoàn thiện ứng dụng mình hơn.

Vào năm 2009, lần đầu tiên công nghệ livestream được sử dụng tại Việt Nam trong một vài buổi biểu diễn ca nhạc của các nghệ sỹ lớn. Theo thống kê chưa chính thức, đến nửa cuối năm 2017 và năm 2018, công nghệ livestream bắt đầu có xu hướng phát triển mạnh, đặc biệt trong lĩnh vực bán hàng (sales) và tiếp thị (marketing). Theo ông Phạm Ngọc Duy Liêm, nhà đồng sáng lập Gostream, công ty cung cấp dịch vụ livestream chuyên nghiệp nhận định, năm 2019 chính là thời điểm để livestream tại Việt Nam thực sự khẳng định vị thế khi các thương hiệu ý thức rõ về sức ảnh hưởng và có sự đầu tư bài bản vào nền tảng này.

Với sự bùng nổ tỷ lệ người dùng điện thoại thông mình cùng với xu hướng mới nổi của nền công nghệ livestream, tôi đã lựa chọn đề tài *Thiết kế, xây dựng Video Streaming Server, hỗ trợ Video Call đa nền tảng* làm đề tài tốt nghiệp với mong muốn xây dựng được một hệ thống ứng dụng cho phép người dùng trao đổi video với nhau trong thời gian thực.

Với sự hiểu biết đôi nét về ngôn ngữ lập trình Android (Java), iOS (Swift) cùng với mong muốn tìm hiểu về cộng nghệ truyền thông đa phương tiện, tôi mong muốn sẽ xây dựng một hệ thống đã nền tảng cho phép người dùng thực hiện video call với nhau trong thời gian thực.

## Các phương pháp giải quyết vấn đề

Để giúp người dùng thực hiện video call với nhau trong thời gian thực có rất nhiều giải pháp. Một trong những nền tảng hỗ trợ là: website dành cho browser; ứng dụng dành cho các phiên bản trên mobile; ứng dụng desktop dành cho PC.

## Mục đích

Mục đích cuối cùng khi kết thúc khóa thực tập tốt nghiệp là mong muốn xây dựng một hệ thống cho phép nhiều người dùng thực hiện video call trực tiếp với một hoặc nhiều người qua mạng internet trong thời gian thực.

## Phạm vi thực hiện đề tài

Một hệ thống thực tế cho phép người dùng thực hiện cuộc gọi video qua internet là một hệ thống phức tạp với nhiều bài toán cần giải quyết. Để một hệ thống có thể phục số lượng người dùng lên đến hàng ngàn, hàng triệu người sẽ đặt ra một bài toán rất lớn về thiết bị hạ tầng của hệ thống.

Trong phạm vi khuôn khổ báo cáo đồ án tốt nghiệp, đề tài dừng lại ở việc xây dựng hệ thống server và các ứng dụng desktop, android, ios cho phép người dùng thực hiện video call trong thời gian thực đến một hoặc nhiều người qua internet với số lượng người dùng hạn chế.

## Tóm tắt cấu trúc của báo cáo

Cấu trúc báo cáo gồm 5 chương:

* Chương 1: Tổng quan. Trong chương này sẽ trình bày khái quát về đề tài, các phương pháp giải quyết vấn đề và phạm vi thực hiện của đề tài.
* Chương 2: Cơ sở lý thuyết. Chương này trình bày về các cơ sở lý thuyết trong việc xử lý tín hiệu đa phương tiện cùng các công nghệ sử dụng.
* Chương 3: Phân tích thiết kế hệ thống. Chương này thể hiện các phân tích nghiệp vụ, thiết kế xây dựng hệ thống và sản phẩm.
* Chương 4: Kết luận và hướng phát triển. Đây là chương rât quan trọng cho thấy các kết quả đật được và phương hướng phát triển cho ứng dụng trong tương lai

## Kết luận

Như vậy trong Chương 1 này tôi đã giới thiệu đôi nét về cái nhìn tổng quan của đề tài mà tôi thực hiện trong báo cáo này. Nội dung chính của Chương 1, tôi đã trình bày khái quát về mục đích, yêu cầu cũng như phạm vi thực hiện của đề tài.

# CƠ SỞ LÝ THUYẾT

Chương 2 trình bày về các khái niệm cơ bản trong kỹ thuật xử lý thông tin đa phương tiện và các cơ sở lý thuyết trong việc nén video, audio.

## Các khái niệm chung

### Streaming video

Streaming video (hay còn gọi là video streaming) là các quá trình luân chuyển liên tục luồng thông tin của một đoạn video đến người dùng cuối thông qua mạng Internet và được hiển thị trong khi dữ liệu vẫn đang trong quá trình tải.

Streaming video thực hiện chia nhỏ thông tin của video thành các frame và gửi các frame đến với thiết bị đầu cuối. Trong khi người dùng đang sử dụng dữ liệu của những frame này thì những frame khác tiếp tục được tải đến.

### Bitrate

Bitrate (hay còn gọi là tốc độ bit) là lượng dữ liệu được truyền và sử lý trong một đơn vị thời gian.

### Adaptive Bitrate Streaming

Truyền phát bitrate thích ứng (Adaptive Bitrate Streaming) hay con gọi là phát trực tuyến ABR, là một kỹ thuật để tự động điều chỉnh mức độ nén và chất lượng video sao cho phù hợp với băng thông mạng.

Phương pháp truyền phát video cũ dựa vào việc phân phối luồng video bitrate cố định. Nếu kết nối mạng của người dùng không thể hỗ trợ bitrate đó, người dùng hoàn toàn không thể xem video mà không có bộ đệm. Với ABR, giờ đây người dùng có thể truyền phát video trên Internet, với cả dịch vụ truyền phát trực tiếp tới nhiều thiết bị.

Đối với truyền phát point-to-point, ABR có thể điều chỉnh một luồng RTMP hoặc SRT duy nhất để phù hợp với băng thông có sẵn giữa hai thiết bị như bộ mã hóa và bộ giải mã. Để truyền phát video trực tiếp, một bộ mã hóa cần có khả năng điều chỉnh mức độ nén của luồng trong thời gian thực, vì băng thông có sẵn liên tục thay đổi. Điều này còn được gọi là mã hóa thích ứng mạng (NAE).

Đối với các dịch vụ OTT, ABR thường dựa trên giao thức đóng gói ABR như HLS hoặc MPEG-DASH. Trong đó nhiều luồng được xác định bởi các cấu hình như chất lượng thấp, trung bình và chất lượng cao. Các luồng ABR được chia thành các đoạn video, trong khoảng 1 - 15 giây, để các thiết bị xem riêng lẻ có thể tự động chọn và chọn đoạn video phù hợp nhất với băng thông có sẵn tại một thời điểm nhất định. Truyền phát ABR cho OTT yêu cầu sử dụng bộ mã hóa hoặc bộ chuyển mã có thể mã hóa một nguồn video duy nhất ở nhiều bitrate. [10]

### Codec

Codec là mạch điện tử hoặc chương trình phần mềm giúp nén hoặc giải nén các tập tin đa phương tiện. Các định dạng dữ liệu nén của codec tuân thủ theo các thông số nén video tiêu chuẩn.

Quá trình nén các tập tin đa phương tiện thường bị mất dữ liệu. Hậu quả của việc này là video sau giải nén có chất lượng thấp hơn so với video gốc nhưng lại giúp giảm kích thước của các tập tin.

Hiện nay với sự bùng nổ của công nghệ đã cho ra đời rất nhiều codec sử dụng trong việc xử lý các tập tin đa phương tiện. Kéo theo đó là các chuẩn nén video ra đời.

#### Nén ảnh

Nén ảnh là kỹ thuật mã hóa ảnh số nhằm giảm số lượng bit dữ liệu cần thiết để biểu diễn hình ảnh. Mục đích của việc nén ảnh là giảm chi phí trong việc lưu trữ và truyền tải hình ảnh nhưng vẫn đảm bảo chát lượng hình ảnh.

Nén ảnh được thực hiện bằng cách loại bỏ các thông tin dư thừa trong hình ảnh và cho ra bức ảnh đã được mã hóa với kích thước nhỏ hơn ảnh gốc nhưng vẫn đảm bảo tính đầy đủ của thông tin. Dư thừa thông tin trong ảnh có thể là dư thừa về không gian, dư thừa về cấp xám, dư thừa về thời gian,...

#### Nén âm thanh

Mục đích của nén âm thanh cũng tương tự như nén ảnh, loại bỏ thông tin dư thừa giúp giảm chi phí trong việc lưu trữ và truyền tải âm thanh. Đối với thính giác của con người, âm thanh nghe được có tần số trong khoảng 20 Hz đến 20.000 Hz. Dư thừa thông tin trong âm thanh gồm có: dư thừa về tấn số, dư thừa về biên độ, dư thừa thời gian,...

## Chuẩn nén video

Cùng với sự phát triển không ngừng nghỉ của công nghệ, các nhà khoa học đã cho ra đời rất nhiều các chuẩn nén video nhằm giảm chi phí trong việc lưu trữ và truyền tải thông tin đa phương tiện nhưng vẫn đảm bảo chất lượng của các tập tin. Trong mục này, em xin trình bày về một số các chuẩn nén video thường gặp trong kỹ thuật xử lý đa phương tiện.

### Cơ sở nén vieo

Video có dải phổ tín hiệu nằm trong phạm vi từ 0 - 6MHz. Tuy nhiên mắt của con người chỉ nhạy cảm với phổ tín hiệu tập trung ở miền tần số thấp và có ít nhạy cảm hơn ở miền tần số cao. Từ đó dẫn đến lượng thông tin tập trung ở miền tần số cao sẽ bị mắt người bỏ qua. Đối với thông tin bị mắt người bỏ qua sẽ được gọi là thông tin dư thừa. Đối với tín hiệu video số, số lượng bit được sử dụng để truyền tải thông tin đối với mỗi miền tần số là khác nhau. Ở miền tần số thấp, nơi chứa đựng nhiều thông tin sẽ được sử dụng với số lượng bit dữ liệu lớn hơn và ngược lại với miền tần số cao, nơi chứa đựng ít thông tin sẽ sử dụng số lượng bit dữ liệu ít. Tổng số bit cần thiết để truyền tải thông tin về hình ảnh sẽ giảm một cách đáng kể và dòng dữ liệu được "nén" mà chất lượng hình ảnh vẫn đảm bảo.

Thực chất kỹ thuật "nén video số" là loại bỏ đi các thông tin dư thừa. Các thông tin dư thừa trong nén video số thường là: độ dư thừa không gian giữa các pixel; độ dư thừa thời gian do các ảnh liên tiếp trùng nhau; Độ dư thừa thống kê do các ký hiệu xuất hiện trong dòng bit với xác suất xuất hiện không đều nhau; độ dư thừa tâm lý thị giác của con người...

Từ đó có thể thấy được mục đích của nén video chính là:

* Giảm tốc độ bit của tín hiệu gốc
* Giảm dung lượng dữ liệu video
* Tiết kiệm chi phí trong việc lưu trữ và truyền dữ liệu

### Tiêu chuẩn nén JPEG

#### Giới thiệu

JPEG là tổ chức nghiên cứu các chuẩn nén ảnh được thành lập vào năm 1982. Tiêu chuẩn nén JPEG được phát triển cho việc nến ảnh tĩnh đa kênh màu và được ứng dụng trong nhiều lĩnh vực như: lưu trữ hình ảnh, truyền tải ảnh trong fax màu, truyền ảnh báo chí,...

Tiêu chuẩn nén JPEG được thực hiện qua bốn mode mã hóa:

* Mã hóa tuần tự: ảnh được mã hóa theo kiểu quét từ trái qua phải, từ trên xuống dưới dựa trên khối DCT.
* Mã hóa lũy tiến: ảnh được mã hóa bằng kiểu quét phức hợp theo chế độ phân giải không gian cho các ứng dụng trên các kiểu băng hẹp và do đó thời gian tryền dẫn có thể dài.
* Mã hóa không tổn thất: ảnh được đảm bảo khôi phục chính xác cho mỗi giá trị mẫu của nguồn. Thông tin không cần thiết sẽ bị cắt bỏ cho nên hiệu quả nén thấp hơn so với phương pháp có tổn thất.
* Mã hóa phân cấp: ảnh được mã hóa ở chế độ phân giải không gian phức hợp, để cho những ảnh óc phân giải thấp có thể được truy xuất và hiển thị mà không cần giải nén như những ảnh có độ phân giải trong không gian cao hơn.

Quá trình mã hóa và giải mã tiêu chuẩn JPEG là hai quá trình ngược nhau. Trong quá trình giải nén, bộ mã hóa entropy biến đổi dòng bit được nén thành một bảng zig-zag mới có các hệ số DCT. Các hệ số này được nhân với các hệ số giải lượng tử hóa và đưa đến quá trình biến đổi DCT. Đầu ra ta có một khối 8x8 pixels có thể không tạo lại một ảnh chính xác tín hiệu gốc vì thông tin bị mất trong quá trình mã. Do đó ảnh sau khi giải nén sẽ khác so với ảnh ban đầu.

Trên quan điểm coi hình ảnh động là chuỗi liên tiếp các hình ảnh tĩnh sắp xếp lại với nhau, tiêu chuẩn JPEG được áp dụng cho việc nén ảnh động và được gọi là M-JPEG. Tiêu chuẩn M-JPEG chỉ thực hiện trong mỗi ảnh, điều đó dẫn đến hiệu quả (tỉ số nén) thấp. [1]

#### Phân cấp cấu trúc số liệu video

Tiêu chuẩn JPEG bao gồm một phân cấp cấu trúc số liệu video nhằm tạo điều kiện thuận lời cho việc biến đổi các hình ảnh được mã hóa. Các thông số mã hóa, chẳng hạn như dạng làm việc của JPEG, kích thước và tần số ảnh, độ chi tiết điểm ảnh, độ chính xác của lượng tử, các bảng mã đều được cộng vào với dòng số dược truyền đi. Cấu trúc số leieuj video JPEG gồm 6 cấp khác nhau phục thuộc vào chế độ làm việc của JPEG.

* Đơn vị số liệu (DU): đơn vị số liệu bao gồm một khối 8x8 xác mẫu thành phần trong dạng nén mất thông tin.
* Đơn vị mã hóa nhỏ nhất (MCU): là nhóm nhỏ nhất các DU xen kẽ. Trong sử dụng nén DCT theo chuẩn CCIR-601, MCU bao gồm hai khối Y, một khối CR và một khối CB.
* Đoạn mã entropy (ECS): gồm một số các MCU. Đoãn mã entropy cho phép giảm kích thước cho khối phục từ giới hạn ngwats của số liệu mã entropy.
* Quét: tiêu chuẩn xác định phương pháp quét cho toàn bộ ảnh.
* Khung hình: có thể được tạo thành từ một hay nhiều quá trình quét.

Lớp ảnh: ảnh là cấp trên cùng của phân cấp số liệu nén, bao gồm khung lớn và các mã cho toàn bộ một bức ảnh. [1]

### Tiêu chuẩn nén MPEG

#### Giới thiệu

Chuẩn MPEG là chuỗi các chuẩn nén video với mục đích là mã hóa tín hiệu hình ảnh và âm thanh cho DSM ở tốc độ bit từ 1.5 – 50 Mbps và được biết đến như là MPEG-1, MPEG-2, MPEG-4,… Các chuẩn MPEG tiến tới tối ưu hóa cho những ứng dụng video động và các đặc điểm của nó cũng bao gồm một thuật toán cho việc nén dữ liệu audio với tỉ lệ vào khoảng từ 5:1 – 10:1.

* MPEG-1: tiêu chuẩn nén một ảnh động có kích thước 320x240 và tốc độ bit còn từ 1– 1.5 Mbps dùng cho ghi hình trên băng từ và đĩa quang (CD), đồng thời truyền dẫn trong các mạng (ví dụ như mạng máy tính).
* MPEG-2: tiêu chuẩn nén được sử dụng cho các ứng dụng cao hơn với tốc độ còn nhỏ hơn bằng 10 Mbps để truyền tín hiệu truyền hình số thương thường. Chuẩn MPEG-2 cũng cho phép mã hóa video với hàng loạt các ứng dụng đòi hỏi có thể phân tích ảnh theo các cách khác nhau như thông tin video trên mạng ISDN sử dụng ATM.
* MPEG-3: tiêu chuẩn nén tín hiệu số xuống còn dưới 50 Mbps để truyền tín hiệu truyền hình có độ phân giải cao. Năm 1992, tiêu chuẩn MPEG-3 được kết hợp với MPEG-2 dùng cho truyền hình thông thường và truyền hình có độ phân giải cao và có tên chung là MPEG-2.
* MPEG-4: chuẩn này dành cho nén hình ảnh video với ít khung hình và yêu cầu làm tươi chậm. Tốc độ dữ liệu yêu cần là 9 ÷ 40 Mbps. MPEG-4 hoàn thiện vào tháng 10/1998 nhằm mục đích phát triển các tiêu chuẩn mã hóa mới với tốc độ bit rất thấp.
* MPEG-7: chuẩn này được đề nghị vào tháng 10/1998 và kế hoạch trở thành chuẩn quốc tế vào tháng 9/2001. MPEG-7 sẽ alf chuẩn mô tả thông tin của rất nhiều loại đa phương tiện. Mô tả này sẽ kết hợp với chính nội dung của nó cho phép khả nưng tìm kiếm nhanh và hiệu quả theo yêu cầu người dùng. MPEG được gọi chính thức là “Giao thức mô tả nội dung đa phương tiện”.

Tiêu chuẩn MPEG là sự kết hợp giữa nén trong ảnh và nén liên ảnh. Tức là phương pháp nén có tổn hao dựa trên biến đổi DCT và bù chuyển động. Tiêu chuẩn MPEG khong biểu diễn cấu trúc bộ mã hóa một cách chính xác, chỉ đặc trưng chính xác các thuật toán nén và kích thước dòng số liệu. [1]

#### Cấu trúc ảnh

MPEG định nghĩa các loại ảnh khác nhau cho phép sự linh hoạt để cân nhắc giữa hiệu quả mã hóa và truy cập ngẫu nhiên. Các loại ảnh đó là:

* Ảnh loại I (Intra-picture): là ảnh được mã hóa riêng, tương tự như việc mã hóa ảnh tĩnh trong JPEG. Ảnh I chứa đựng dữ liệu để tái tạo lại toàn bộ hình ảnh vì chúng được tạo thành bằng thông tin của chỉ một ảnh. Ảnh I cho phép truy cập ngẫu nhiên, tuy nhiên đạt được tỉ lệ nén thấp nhất.
* Ảnh loại P (Predicted-picture): là ảnh đuộc mã hóa có bù chuyển dộng từ ảnh I hoặc P phía trước (ảnh dự đoán trước). Ảnh P cho hệ số nén cao hươn ảnh I và có thể sử dụng làm một ảnh so sánh cho việc bù chuyển động cho các ảnh P và B khác.
* Ảnh loại B (Bi-directional predicted picture): là ảnh ddowwjc mã hóa sẻ dụng bù chuyển động từ các ảnh I hawojc P ở phía trước và ở phía sau (ảnh dự đoán hai chiều). Ảnh B cho tỉ lệ nén cao nhất [1]

#### Nhóm ảnh (GOP)

Đối với chuẩn MPEG, chất lượng ảnh không những phụ thuộc vào tỉ lệ nén trong từng khuân hình mà còn phục thuộc vào độ dài của nhóm ảnh. Nhóm ảnh (GOP) là khái niệm cơ bản của MPEG. Nhóm ảnh là đơn vị mang thông tin độc lập của MPEG.

Công nghệ MPEG sử dụng ba loại ảnh: I, P, B. Trong đó ảnh P, B không phải là một ảnh hoàn chỉnh mà chỉ chứa sự khác biệt giữa ảnh đó và ảnh xuất hiện trước nó (đối với P) hay sự khác biệt đối với cả khuân hình xuất hiện trước và sau nó (đối với B). Để có một khuân hình hoàn chỉnh ảnh P và B cần có dữ liệu từ các ảnh lân cận. Chính vì vậy, đối với MPEG một khái niệm mới là GOP được sử dụng. Mỗi GOP bắt buộc phải bắt đầu bằng một ảnh hoàn chinhrI và tiếp sau đó là một loại các ảnh P và B. Nhóm ảnh có thể mở (Open) hoặc đóng (Close).

Tỷ lệ nén video của MPEG phục thuộc rất nhiều vào độ dài của GOP. Tuy nhiên GOP dài thường gây khó khăn cho quá trình tua, định vị, sửa lỗi,… Do đó tùy thuộc vào từng khâu mà chọ độ dài GOP thích hợp. [1]

### Tiêu chuẩn nén H265/HEVC

#### Giới thiệu

Tên gọi đầy đủ: High Efficiency Video Coding - HEVC Hiệu quả nén video cao

H.265 là công nghệ mã hóa mới nhất trong lĩnh vực Video. Với những thuật toán thông minh, giúp giảm bitrate cho video dựa trên công nghệ mã hóa tiêu chuẩn:

* Mã hóa dự đoán: giúp trích xuất đối tượng chuyển động từ một bối cảnh nền ổn định; mã hóa khác nhau cho đối tượng chuyển động và bối cảnh nền
* Triệt tiêu nhiễu: trích xuất và loại bỏ nhiễu
* Điều khiển bitrate: Bù bitrate cho các khoảng thời gian nhiều hoạt động trong dài hạn

Công nghệ mã hóa có băng thông rất thấp và hiệu quả tiết kiệm lưu trữ tốt

#### Nguyên lý hoạt động

Mục tiêu chính của việc nghiên cứu và phát triển tiêu chuẩn H.265/HEVC là cho phép cải thiện đáng kể hiệu suất nén tương đối so với cái tiêu chuẩn trước đó trong khoảng từ giảm tốc độ bit 50% và cho chất lượng video cảm nhận được bằng các tiêu chuẩn hiện có.

Bên cạnh đó HEVC đã được thiết kế để giải quyết về cơ bản tất cả các ứng dụng hiện có sử dụng H.264/MPEG-4 AVC và đặc biệt là tập trung vào hai vấn đề chính:

* Độ phân giải video tăng.
* Tăng cường sử dụng kiến trúc xử lý song song.

Vậy chúng ta sẽ đặt ra câu hỏi nếu tốc độ bit thực sự được giảm đi một nửa thì có thể sẽ có nhưng chuyện gì xẩy ra. Thực sự sẽ có những sự thay đổi lớn trong các ứng dụng và thiết bị đang sử dụng các chuẩn nén video trước đó.

* IPTV qua DSL: sự thay đổi lớn trong điều kiện IPTV
* Tạo điều kiện triển khai các OTT và màn hình đa dịch vụ
* Nhiều khách hang trên cơ sở hạ tầng như nhau: hầu hết các lưu lượng IP là video
* Nhiều phương tiện lưu trữ

Và H.265/HEVC sẽ đi tiên phong và chiếm lĩnh công nghệ trong các dịch vụ trong tương lai

* 1080p60/50 với bitrate so sánh với 1080i
* Trải nghiệm xem video siêu nét và siêu thực: Ultra-HD (4k-8k)
* Dịch vụ Premium
* HD TV 3D

Mã hóa hiệu quả (efficiency coding) là khả năng mã hóa video với tốc độ bit thấp nhất có thể trong khi duy trì một mức độ chất lượng video. Có hai phương pháp tiêu chuẩn để đo lường hiệu quả việc mã hóa của một video mã hóa tiêu chuẩn, trong đó có sử dụng một số liệu khách quan, chẳng hạn như tỷ số tín hiệu đỉnh trên nhiễu, hoặc sử dụng đánh giá chủ quan về chất lượng video. Đánh giá chủ quan của chất lượng video được gọi là cách quan trọng nhất để đo lường một tiêu chuẩn mã hóa video kể từ khi con người nhận thức chất lượng video chủ quan.

Những lợi ích của HEVC đến từ việc sử dụng cây mã hóa đơn vị CTU (Coding Tree Unit) cỡ lớn. điều này đã được chứng minh trong các thử nghiệm PSNR với HM-8.0 mã hóa HEVC nơi nó bị buộc phải sử dụng kích thước CTU dần dần nhỏ hơn.

#### Cấu trúc mã hóa (Coding Structure)

HEVC đã thay thế macroblocks, được sử dụng với tiêu chuẩn trước đó bằng cấu trúc mã hóa cây đơn vị (Coding Tree Unit - CTUs) và cấu trúc mã hóa khối cây (Coding Tree Block - CTB). Việc sử dụng cấu trúc CTUs đã là cho HEVC gia tang đáng kể hiệu quả mã hóa. HEVC ban đầu chia hình ảnh thành 3CTB:

* 1 luma CTB
* 2 chroma CTB

Kích thước L x L của một luma CTB có thể được chọn là L = 16, 32 hoặc 64 mẫu, với các kích thước lớn hơn thường cho phép nén tốt hơn. HEVC còn hỗ trợ phân vùng của CTBs thành các khối nhỏ hơn.

Chart

Description automatically generated

Hình . Kiến trúc mã hóa Quad-tree

Trong công nghệ HEVC thì CUs là đơn vị dự đoán cơ bản. CUs có thể là 64x64, 32x32, 16x16.

### Tiêu chuẩn H.264

#### Giới thiệu

H.264 phát triển nhằm đưa ra video có chất lượng tốt hơn các chuẩn nén video trước nhờ sự kế thừa các ưu điểm của các chuẩn nén tiền nhiệm.

* Chuẩn H264 phân chia mỗi hình ảnh thành các Block gồm nhiều các điểm ảnh. Điều đó giúp quá trình xử lý dữ liệu được diễn ra nhanh chóng thông qua các Block.
* H.264 loại bỏ dư thừa về mặt không tồn tại giữa các hình ảnh liên tiếp thông qua dự đón về không gian, phép biến đổi, quá trình lượng tử hóa và mã hóa Entropy.
* H.264 khai thác phụ thuộc tạm thời của các Block của hình ảnh liên tiếp, do đó chỉ cần mã hóa những chi tiết thay đổi giữa các ảnh liên tiếp. Việc này được thực hiện thông qua dự đoán và bù chuyển động.
* H.264 khai thác tất cả sự dư thừa về không gian còn lại trong ảnh bằng việc giaiar mã các Block dư thừa. [1]

#### Lớp trừu tượng mạng NAL

Lớp trừu tượng mạng được xác định để định dạng dữ liệu và cung cấp thông tin Header trong việc truyền tài video. Tất cả dữ liệu được chưa trong các khối NAL, mỗi khối chứa một số nguyên byte nhất định các định dạng truyền thông đa phương tiện. Mỗi khối NAL xác định định dạng chung cho việc sử dụng trong cả hệ thống định hướng gói và hệ thống định hướng dòng bit.

Lớp trừu tượng mạng (NAL) là một phần của bộ quy chuẩn H.264/AVC và HEVC về mã hóa video. H.264 phân biệt rõ rệt giữa lớp mã hóa (VCL) và lớp trừu tượng hóa mạng. Kết xuất của quá trình mã hóa là dữ liệu VCL, được gói trong các gói NAL (NAL units) trước khi lưu trữ hay truyền đi. NAL được thiết kế với mục tiêu "thân thiện với đường truyền", cho phép sử dụng đơn giản và hiệu quả các dữ liệu VCL trong nhiều hệ thống đa dạng như:

* RTP/IP cho mọi loại giao tiếp thời gian thực có dây hoặc không dây (đàm thoại hoặc streaming).
* Các định dạng File (ISO MP4 để lưu trữ và MMS...)
* H.32x cho các dịch vụ hội thoại có dây và không dây.
* Các hệ thống MPEG-2 cho các dịch vụ quảng bá (truyền hình,...)

Diagram

Description automatically generated

Hình . Cấu trúc bộ mã hóa video H.264 [4]

Việc chỉ định riêng rẽ lớp mã hóa video và lớp trừu tượng hóa mạng nhằm mục đích phân biệt giữa các đặc trưng của riêng mã hóa (tại lớp VCL) và vận chuyển (tại lớp NAL).

Diagram

Description automatically generated

Hình . NAL unit [4]

Thành phần quan trọng nhất của NAL đó là NAL unit (gói NAL), mọi dữ liệu mã hóa video đều được đua vào các gói NAL unit, mỗi gói được tổ chức sao cho chúng chứa một số nguyên các byte. Dữ liệu payload trong gói NAL được lồng ghép một số các byte chống giả lập (emulation prevention bytes) nếu cần thiết, được chèn vào với một giá trị được định trước để ngăn ngừa việc một mẫu dữ liệu cụ thể gọi là start code prefix (tiếp đầu mã) bị lặp lại một cách tình cờ bên trong dữ liệu payload.

Mỗi gói NAL chứa một chuỗi byte thô (RBSP), chứa dữ liệu video mã hóa hoặc thông tin trong header. Các đơn vị RBSP được truyền đi như các gói NAL. Một gói NAL chứa 1 byte header cho biết loại đơn vị RBSP và dữ liệu RBSP nào được gói trong NAL unit.[3]

### Tiêu chuẩn VP8

#### Giới thiệu

Tháng 5 năm 2010, Google đã công bố bắt đầu thực hiện một dự án truyền thông đa phương tiện mã nguồn mở mới “WebM”, đây là dự án được dành riêng để phát triển một định dạng đa phương tiện nguồn mở, chất lượng cao cho dành cho các nền tảng web. Cốt lõi của dự án là một định dạng nén video mã nguồn mở mới, VP8. Định dạng VP8 ban đầu được phát triển bởi một nhóm nghiên cứu nhỏ tại On2 Technologies, Inc.asa và được kế thừa từ phiên bản mã hóa video VPx. So với các định dạng mã hóa video khác, VP8 có nhiều tính năng kỹ thuật đặc biệt giúp nó đạt được hiệu quả nén cao và độ phức tạp tính toán thấp. Kể từ khi công bố WebM, VP8 không chỉ nhận được sự ủng hộ mạnh mẽ từ rất nhiều công ty lớn trong ngành mà còn bắt đầu thu hút sự quan tâm rộng rãi trong cộng đồng nghiên cứu mã hóa video từ cả ngành công nghiệp và lĩnh vực nghiên cứu. [5]

#### Thiết kế và tính năng nổi bật

Ngay từ khi bắt đầu phát triển VP8, các nhà phát triển đã tập trung vào các ứng dụng video dựa trên ứng dụng web. Sự tập trung này đã dẫn đến một số giả định cơ bản trong thiết kế tổng thể VP8 như sau [5]:

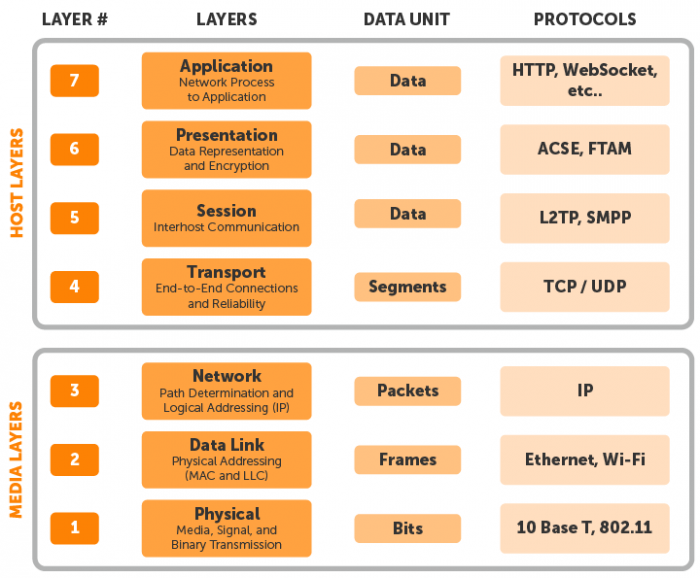
* Yêu cầu băng thông thấp: một trong những giả định thiết kế cơ bản là trong tương lai gần, băng thông mạng khả dụng sẽ bị hạn chế. Với giả định này, VP8 được thiết kế đặc biệt để hoạt động chủ yếu trong một phạm vi chất lượng video có thể xem được của Cameron (~ 30dB) cho đến chất lượng trực quan không mất dữ liệu (~ 45dB).
* Phần cứng thiết bị người dùng không đồng nhất: đây là một yếu tố quan trọng góp phần Có phổ rộng phần cứng máy khách được kết nối với web, từ thiết bị di động và thiết bị nhúng công suất thấp đến máy tính để bàn tiên tiến nhất có nhiều lõi xử lý. Do đó, nó phải có khả năng tạo ra các triển khai hiệu quả cho nhiều loại khách hàng.
* Định dạng video trên nền tảng web: VP8 được thiết kế để xử lý định dạng hình ảnh được sử dụng bởi phần lớn các video web: lấy mẫu 420 màu, độ sâu màu perchannel 8 bit, quét liên tục (không xen kẽ) và kích thước hình ảnh lên tới tối đa 16383x16383pixels.

Việc thúc đẩy hiệu quả nén và đơn giản hóa bộ giải mã theo các giả định thiết kế này đã dẫn đến một số tính năng đặc biệt trong VP8 liên quan đến các định dạng nén video đã biết khác, như MPEG-2, H.263 và H.264/AVC. Sau đây là những cải tiến kỹ thuật trong VP8 [5]:

* Biến đổi lai với lượng tử hóa thích ứng: VP8 sử dụng biến đổi cosine rời rạc (DCT) 4x4.
* Khung tham chiếu linh hoạt: VP8 sử dụng ba khung tham chiếu để dự đoán liên ảnh, nhưng sơ đồ này hơi khác so với sơ đồ tham chiếu bù chuyển động trong các định dạng khác. VP8 giới hạn yêu cầu kích thước bộ đệm đối với ba bộ đệm khung tham chiếu nhưng vẫn đạt được hiệu quả khử tương quan trong bù chuyển động.
* Dự đoán nội bộ và dự đoán liên ảnh: VP8 sử dụng hiệu quả dự đoán nội bộ và dự đoán liên ảnh. Truyền tin trực tiếp của VP8 có một chế độ mới “TM\_PRED”. Đây là một trong nhiều phương pháp dự đoán nội bộ đơn giản và hiệu quả. Để diễn giải, VP8 có chế độ Linh hoạt SPLITMV, có khả năng mã hóa các mẫu khối tùy ý trong một macroblock.
* Nội suy sub-pixel hiệu suất cao: Bù chuyển động của VP8 sử dụng các vectơ chuyển động chính xác tới một phần tư pixel cho các pixel luma và các vectơ chuyển động chính xác đến một phần tám pixel cho các pixel màu. Nội suy sub-pixel của VP8 có quy trình nội suy một giai đoạn và một bộ các bộ lọc nội suy với hiệu suất cao.
* Lọc gỡ lỗi trong vòng lặp thích ứng: VP8 có bộ lọc gỡ lỗi trong vòng lặp thích ứng. Kiểu lọc và cường độ của bộ lọc có thể điều chỉnh cho các chế độ dự đoán khác nhau và các loại khung tham chiếu.
* Mã hóa entropy thích ứng ở mức khung: VP8 sử dụng mã hóa nhị phân cho hầu hết các giá trị dữ liệu ngoại trừ một vài bit tiêu đề. Sự thích ứng của entropy ở mức khung hình đã tạo ra sự cân bằng giữa hiệu quả nén và độ phức tạp tính toán.
* Xử lý song song phân vùng dữ liệu: VP8 có thể đóng gói các hệ số biến đổi mã hóa entropy thành nhiều phân vùng giúp tạo điều kiện thuận lợi cho việc xử lý song song trong bộ giải mã. Thiết kế này đã giúp cải thiện hiệu suất giải mã trên các bộ xử lý đa lõi nhưng không làm ảnh hưởng đến hiệu suất nén và giải mã.

### Giới thiệu Protocol

Giao thức (Protocol) là tập hợp các quy tắc truyền tải dữ liệu di chuyển từ hệ thống truyền thông này sang hệ thống khác. Protocol được xếp chồng lên nhau để tạo thành một chồng giao thức. Bằng cách đó, các giao thức ở mỗi lớp có thể tập trung vào một chức năng cụ thể và hợp tác với nhau. Lớp thấp nhất hoạt động như một nền tảng và mỗi lớp bổ sung thêm độ phức tạp.



Hình . Mô hình bảy tầng OSI [6]

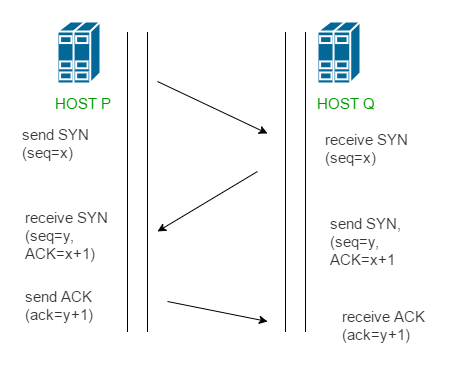
IP là viết tắt của Protocol Internet. Cấu trúc của giao thức này cho biết các thiết bị sử dụng internet để giao tiếp như thế. IP nằm ở lớp mạng và thường được phủ bởi TCP ở lớp vận chuyển, cũng như HTTP ở lớp ứng dụng.

Mô hình bảy tầng OSI bao gồm: tầng vật lý (Physical), tầng liên kết dữ liệu (Data link), tầng mạng (Network), tầng vận chuyển (Transport), tầng phiên (Session), tầng trình bày (Presentation) và tầng ứng dụng (Application) - được xác định theo mô hình Liên kết hệ thống mở của Tổ chức tiêu chuẩn hóa (IS0,) như mô tả ở trên. [6]

### Giao thức UDP và TCP

Giao thức UDP và giao thức TCP đều là các thành phần cốt lõi của bộ giao thức IP. Cả hai giao thức này đều nằm trên tầng Transport. UDP và TCP khác nhau về chất lượng và tốc độ.

Sự khác biệt chính giữa UDP và TCP là TCP yêu cầu bắt tay ba bước khi vận chuyển dữ liệu: Client (HOST P) gửi một yêu cầu truy cập đến server (HOST Q); server trả lại phản hồi và client tiếp nhận đồng thời duy trì phiên giữa hai đầu. Vì lý do này mà TCP khá đáng tin cậy và có thể giải quyết mất gói và đặt hàng.



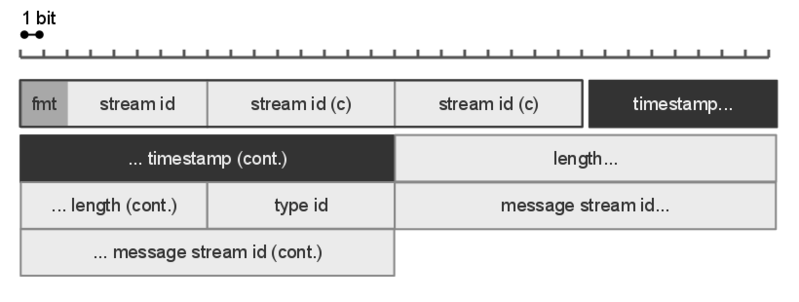
Hình . Giao thức TCP/IP [7]

Mặt khác, UDP vận chuyển dữ liệu mà không yêu cầu bước bắt tay. UDP vận chuyển dữ liệu bất kể ràng buộc băng thông, làm cho nó nhanh hơn nhưng tồn tại nhiều rủi ro hơn. Vì UDP không hỗ trợ truyền lại, đặt hàng gói hoặc kiểm tra lỗi, nên có thể sẽ bị mất dữ liệu trên đường truyền.

Các giao thức như Giao thông tin cậy an toàn (SRT) thường sử dụng UDP, trong khi các giao thức xây dựng dựa trên HTTP sử dụng TCP. [6]

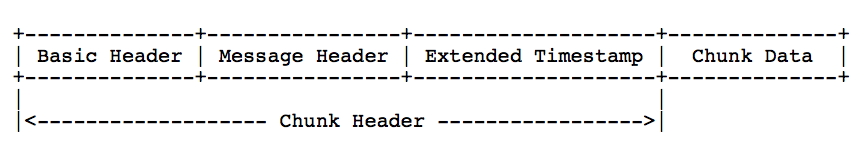
### Giao thức RTMP

RTMP là giao thức không công khai do Adobe phát triển và giữ bản quyền. RTMP được thiết kế cho ứng dụng thời gian thực, cho phép ứng dụng sử dụng video và âm thanh với tốc độ nhanh, hạn chế bị đứng hình hoặc méo tiếng.



Hình . Cấu trúc gói tin của giao thức RTMP [8]

Adobe đã thiết kế đặc tả RTMP để truyền dữ liệu âm thanh và video giữa các công nghệ như máy chủ phát trực tuyến chuyên dụng và Adobe Flash Player. Giao thức có độ đáng tin cậy cao và độ trễ thấp, điều này phù hợp với yêu cầu truyền tải video trực tiếp. [6]



Hình . Header của giao thức RTMP [8]

Theo thông báo do Adobe công bố Flash sẽ bị “khai tử” dự kiến ​​vào năm nay, năm 2020. Mặc dù Flash sẽ không được sử dụng trong tương lai nhưng giao thức RTMP vẫn có thể dược giữ lại. Các bộ mã hóa RTMP vẫn là một lựa chọn phù hợp cho nhiều nhà phát triển mặc dù giao thức độc quyền đã không còn được ưa chuộng đối với các nhà phát triển phần mềm. [6]

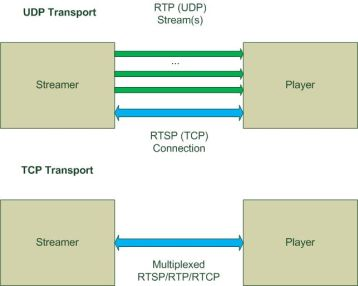
Đặc điểm giao thức RTMP:

Bảng . Đặc điểm giao thức RTMP [6]

|  |  |
| --- | --- |
| **Đặc điểm** | **Thông tin** |
| Chuẩn nén audio hỗ trợ | AAC, AAC-LC, HE-AAC+ v1 & v2, MP3, Speex, Opus, Vorbis |
| Chuẩn nén video hỗ trợ | H.264, VP8, VP6, Sorenson Spark®, Screen Video v1 & v2 |
| Độ trễ | 5s |
| Ưu điểm | độ trễ thấp, không yêu cầu bộ đệm |
| Nhược điểm | không được tối ưu hóa cho chất lượng trải nghiệm hoặc khả năng mở rộng |
| Khả năng tương thích | không được hỗ trợ rộng rãi, chỉ hỗ trợ Flash Player, Adobe AIR, RTMP-compatible players |

### Giao thức RTSP/RTP

Giống như RTMP, RTSP/RTP là giao thức trạng thái được sử dụng để truyền tải video đến nhiều thiết bị. Mặc dù RTMP là giao thức lớp presentation cho phép người dùng cuối điều khiển máy chủ phương tiện thông qua khả năng dừng và phát.



Hình . Giao thức RTSP/RTP

RTP là giao thức vận chuyển thời gian thực đặc tả một tiêu chuẩn định dạng gói tin sử dụng trong việc truyền tải âm thanh và hình ảnh qua internet. Các thiết bị Android và iOS không có player tương thích với RTSP làm cho giao thức này khó sử dụng khi phát lại dữ liệu đa phương tiện. [6]

RTSP là giao thức ở tầng ứng dụng trong bộ các giao thức Internet để kiếm soát việc truyền dữ liệu theo thời gian thực. RTSP cung cấp một nền tảng mở rộng cho phép kiểm soát, truyền theo yêu cầu của dữ liệu thời gian thực. RTSP được sử dụng để thiết lập và quản lý các phiên làm việc giữa các điểm truyền, phát tin đa phương tiện

Đặc điểm của giao thức RTSP/RTP:

Bảng . Đặc điểm giao thức RTSP/RTP [6]

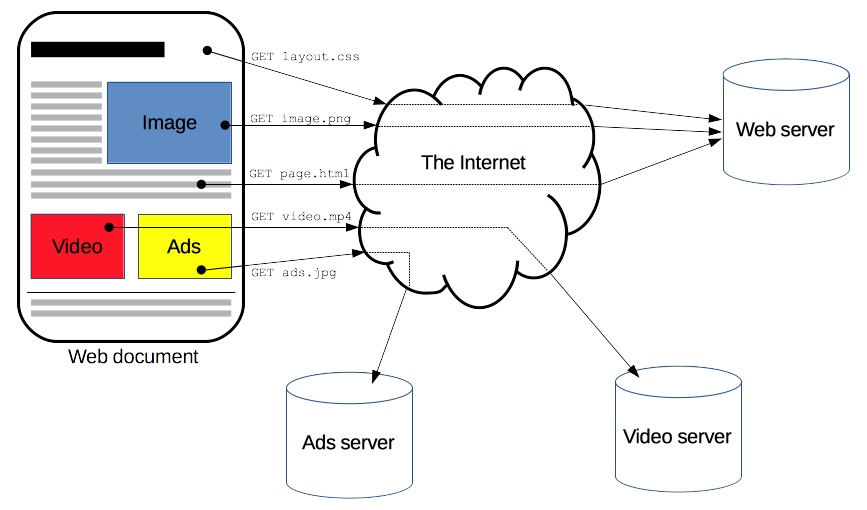
|  |  |
| --- | --- |
| **Đặc điểm** | **Thông tin** |
| Chuẩn nén audio hỗ trợ | AAC, AAC-LC, HE-AAC+ v1 & v2, MP3, Speex, Opus, Vorbis |
| Chuẩn nén video hỗ trợ | H.265 (preview), H.264, VP9, VP8 |
| Độ trễ | 2s |
| Ưu điểm | độ trễ thấp, không yêu cầu bộ đệm |
| Nhược điểm | không được tối ưu hóa cho chất lượng trải nghiệm và khả năng mở rộng |
| Khả năng tương thích | không được hỗ trợ rộng rãi (Quicktime Player and other RTSP/RTP-compliant players, VideoLAN VLC media player, 3Gpp-compatible mobile devices) |

### HTTP

HTTP là một giao thức cho phép tìm nạp các tài nguyên, như các tài liệu HTML. HTTP là nền tảng của bất kỳ trao đổi dữ liệu nào trên Web và là giao thức client-server, có nghĩa là các yêu cầu được bắt đầu bởi người nhận, thường là trình duyệt Web. HTTP được sử dụng để truyền tải các tập tin văn bản, mô tả bố cục, hình ảnh, video, tập lệnh,…

Client và server liên lạc bằng cách trao đổi các tin nhắn riêng lẻ (trái ngược với luồng dữ liệu). Các tin nhắn được gửi bởi client, thường là trình duyệt Web, được gọi là các yêu cầu (request) và các tin nhắn được gửi bởi máy chủ dưới dạng câu trả lời được gọi là phản hồi (responses). [9]

Một số giao thức được xây dựng dựa trên giao thức HTTP phổ biến hiện nay gồm có: MPEG-DASH và HLS của Apple.



Hình . Mô tả hoạt động giao thức HTTP [9]

### WebRTC

Một trong số các công nghệ mới nổi trong lĩnh vực truyền thông đa phương tiện thời gian thực là công nghệ WebRTC. Công nghệ mới WebRTC là sự kết hợp của các tiêu chuẩn, giao thức và API JavaScript cho phép giao tiếp thời gian thực.

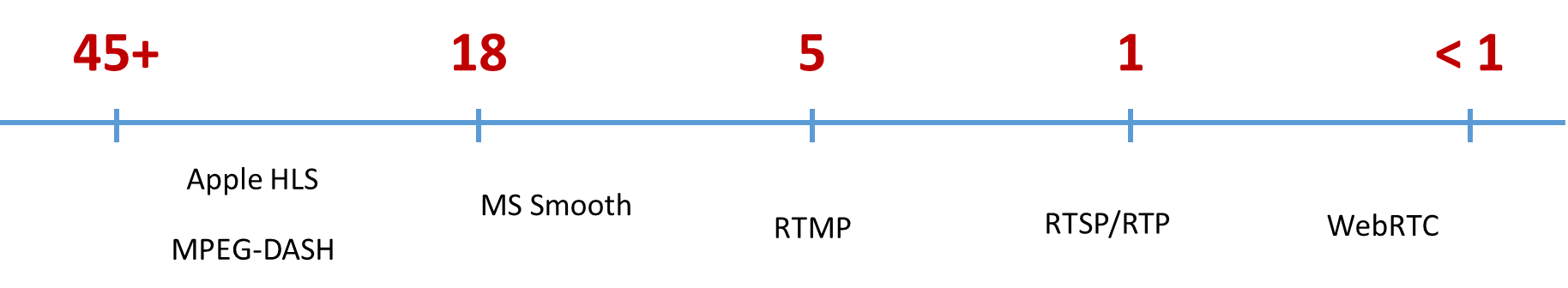
Bảng . Đặc điểm giao thức WebRTC [6]

|  |  |
| --- | --- |
| **Đặc điểm** | **Thông tin** |
| Chuẩn nén audio hỗ trợ | Opus, iSAC, iLBC |
| Chuẩn nén video hỗ trợ | VP8, VP9 |
| Độ trễ | < 1s |
| Ưu điểm | Độ trễ thấp, có hỗ trợ trình duyệt và các nền tảng mobile |
| Nhược điểm | Được thiết kế theo mô hình hội nghị (room) |
| Khả năng tương thích | Chrome, Firefox và Safari hỗ trợ WebRTC mà không cần bất kỳ plugin nào |

Nội dung chi tiết về công nghệ WebRTC sẽ được trình bày trong mục 2.4.

### Độ trễ truyền phát giữa các giao thức

Theo thống kê của Wowza [6] cho thấy các giao thức truyền thông đa phương tiện khác nhau sẽ có độ trễ khác nhau (Hình 2.10). Theo đó, RTMP truyền tải video với tốc độ gần bằng tốc độ truyền hình cáp (chỉ trong hơn 5s). RTSP/RTP thậm chí còn nhanh hơn với độ trễ khoảng hơn 1s. Với độ trễ thấp nhất phải kể đến công nghệ mới nổi hiện nay là WebRTC với độ trễ nhỏ hơn 1s. Các giao thức này đạt được thanh tựu to lớn như thế điều sử dụng phương pháp firehose thay vì yêu cầu tải xuống cục bộ hoặc bộ đệm.



Hình . Độ trễ truyền phát giữa các giao thức

## NAT

### Giới thiệu

Dịch địa chỉ mạng (NAT) là một phương pháp ánh xạ lại một không gian địa chỉ IP sang một không gian địa chỉ IP khác bằng cách sửa đổi thông tin địa chỉ mạng trong tiêu đề IP của các gói tin. Việc sửa đổi này được thực hiện khi các gói tin đang được truyền qua một thiết bị định tuyến. NAT đã trở thành một công cụ phổ biến và thiết yếu trong việc bảo tồn không gian địa chỉ toàn cầu trước tình trạng cạn kiệt địa chỉ IPv4. Một địa chỉ IP có thể định tuyến Internet của cổng NAT có thể được sử dụng cho toàn bộ mạng riêng.

C:\Users\LTS\Pictures\nat.png

Hình . Mô hình hoạt động của NAT [7]

Khi các gói tin được gửi ra ngoài internet, NAT giúp chuyển đổi một địa chỉ IP private trong mạng LAN thành địa chỉ IP publish. Do sự phổ biến của kỹ thuật này để bảo tồn không gian địa chỉ IPv4, thuật ngữ NAT gần như đồng nghĩa với việc giả mạo IP.

Khi dịch địa chỉ mạng sửa đổi thông tin địa chỉ IP trong các gói, việc triển khai NAT có thể thay đổi hành vi cụ thể của chúng trong các trường hợp địa chỉ khác nhau và ảnh hưởng của chúng đối với lưu lượng mạng. Các chi tiết cụ thể về hành vi của NAT thường không được ghi nhận bởi các nhà cung cấp thiết bị có chứa các triển khai NAT.

NAT cung cấp nhiều lợi ích nhưng cũng đi kèm với nhiều nhược điểm. Rắc rối nhất trong những nhược điểm đó là việc chúng phá vỡ nhiều ứng dụng IP hiện có và gây khó khăn cho việc triển khai các ứng dụng mới. Các hướng dẫn đã được phát triển mô tả cách xây dựng các giao thức "NAT thân thiện", nhưng nhiều giao thức đơn giản là không thể được xây dựng theo các hướng dẫn đó. Ví dụ về các giao thức như vậy bao gồm hầu hết tất cả các giao thức ngang hàng, như giao tiếp đa phương tiện, chia sẻ tệp và trò chơi.[13]

### Hoạt động

Phần lớn các NAT ánh xạ nhiều máy chủ riêng tới một địa chỉ IP được hiển thị công khai. Trong cấu hình thông thường, mạng cục bộ sử dụng một trong các mạng con địa chỉ IP riêng. Một bộ định tuyến trên mạng đó có một địa chỉ riêng trong không gian địa chỉ đó. Bộ định tuyến cũng được kết nối với Internet với một địa chỉ công cộng được chỉ định bởi nhà cung cấp dịch vụ Internet. Khi lưu lượng truyền từ mạng cục bộ sang Internet, địa chỉ nguồn trong mỗi gói được dịch nhanh chóng từ một địa chỉ riêng sang địa chỉ công cộng. Bộ định tuyến theo dõi dữ liệu cơ bản về từng kết nối hoạt động (đặc biệt là địa chỉ đích và cổng). Khi trả lời trở lại bộ định tuyến, nó sử dụng dữ liệu theo dõi kết nối mà nó được lưu trữ trong giai đoạn gửi đi để xác định địa chỉ riêng trên mạng nội bộ để chuyển tiếp trả lời.

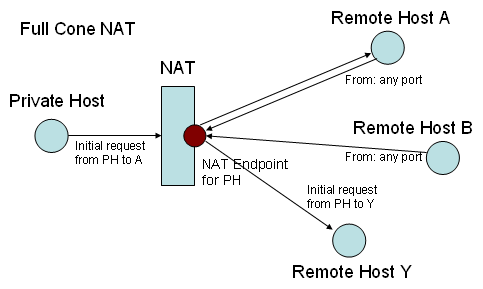
Tất cả các gói IP đều có địa chỉ IP nguồn và địa chỉ IP đích. Thông thường các gói truyền từ mạng riêng sang mạng công cộng sẽ được sửa đổi địa chỉ nguồn của chúng, trong khi các gói truyền từ mạng công cộng trở lại mạng riêng sẽ được sửa đổi địa chỉ đích. Để tránh sự mơ hồ trong cách trả lời được dịch, cần phải sửa đổi thêm cho các gói. Phần lớn lưu lượng truy cập Internet sử dụng giao thức TCP hoặc giáo thứcUDP. Đối với các giao thức này, số cổng được thay đổi để sự kết hợp giữa địa chỉ IP và thông tin cổng trên gói được trả về có thể được ánh xạ rõ ràng đến đích mạng riêng tương ứng.

Phương pháp này chỉ cho phép giao tiếp qua bộ định tuyến khi cuộc hội thoại bắt nguồn từ mạng riêng vì việc truyền phát ban đầu là điều thiết lập thông tin cần thiết trong các bảng dịch.

### Phân loại

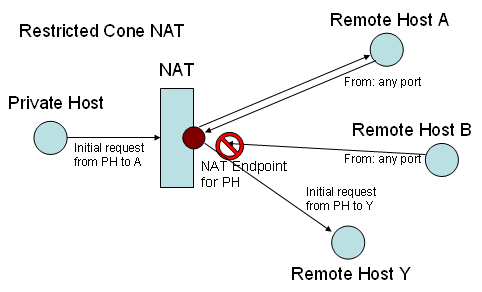
Có tổng cộng 4 loại NAT khác nhau:

* NAT loại 1 (Full Cone): Khi một địa chỉ nội bộ (iAddr: iPort) được ánh xạ tới một địa chỉ bên ngoài (eAddr: ePort), mọi gói tin từ iAddr: iPort đều được gửi qua eAddr: ePort. Bất kỳ máy chủ bên ngoài nào cũng có thể gửi các gói đến iAddr: iPort bằng cách gửi các gói đến eAddr: ePort.



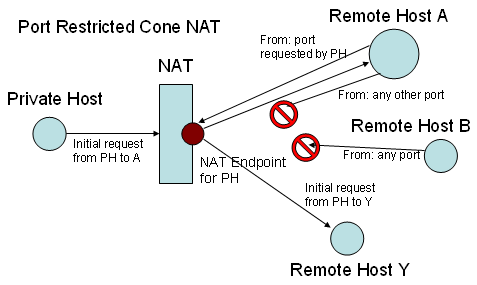
Hình . NAT loại 1

* NAT loại 2 (Restricted Cone): Khi một địa chỉ nội bộ (iAddr: iPort) được ánh xạ tới một địa chỉ bên ngoài (eAddr: ePort), mọi gói tin từ iAddr: iPort đều được gửi qua eAddr: ePort. Một máy chủ bên ngoài (hAddr: any) có thể gửi các gói đến iAddr: iPort bằng cách gửi các gói tới eAddr: ePort chỉ khi iAddr: iPort trước đó đã gửi một gói đến hAddr: any. "Bất kỳ" có nghĩa là số cổng không thành vấn đề.



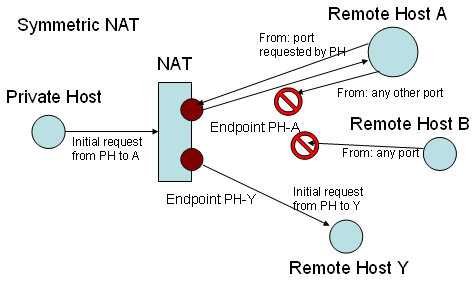
Hình . NAT loại 2

* NAT loại 3: Giống như một địa chỉ giới hạn hình nón NAT, nhưng hạn chế bao gồm số cổng. Khi một địa chỉ nội bộ (iAddr: iPort) được ánh xạ tới một địa chỉ bên ngoài (eAddr: ePort), mọi gói tin từ iAddr: iPort đều được gửi qua eAddr: ePort. Một máy chủ bên ngoài (hAddr: hPort) có thể gửi các gói đến iAddr: iPort bằng cách gửi các gói đến eAddr: ePort chỉ khi iAddr: iPort trước đó đã gửi một gói đến hAddr: hPort.



Hình . NAT loại 3

* NAT loại 4: Mỗi yêu cầu từ cùng một địa chỉ IP nội bộ và cổng đến một địa chỉ IP đích và cổng cụ thể được ánh xạ tới một địa chỉ IP và cổng nguồn bên ngoài duy nhất; nếu cùng một máy chủ nội bộ gửi một gói ngay cả với cùng một địa chỉ nguồn và cổng nhưng đến một đích khác, thì ánh xạ khác sẽ được sử dụng. Chỉ có máy chủ bên ngoài nhận gói từ máy chủ nội bộ mới có thể gửi lại gói.



Hình . NAT loại 4

## Thuật toán xác định loại NAT

Hệ thống nhận diện NAT phía người dùng gồm có hệ thống máy chủ 2 card mạng riêng biệt đóng vai trò như 2 quan sát viên (observer) hỗ trợ nhận diện cổng mạng (port) và địa chỉ mạng (IP). Cách thức hoạt động của hệ thống xác định loại NAT như sau:

Diagram

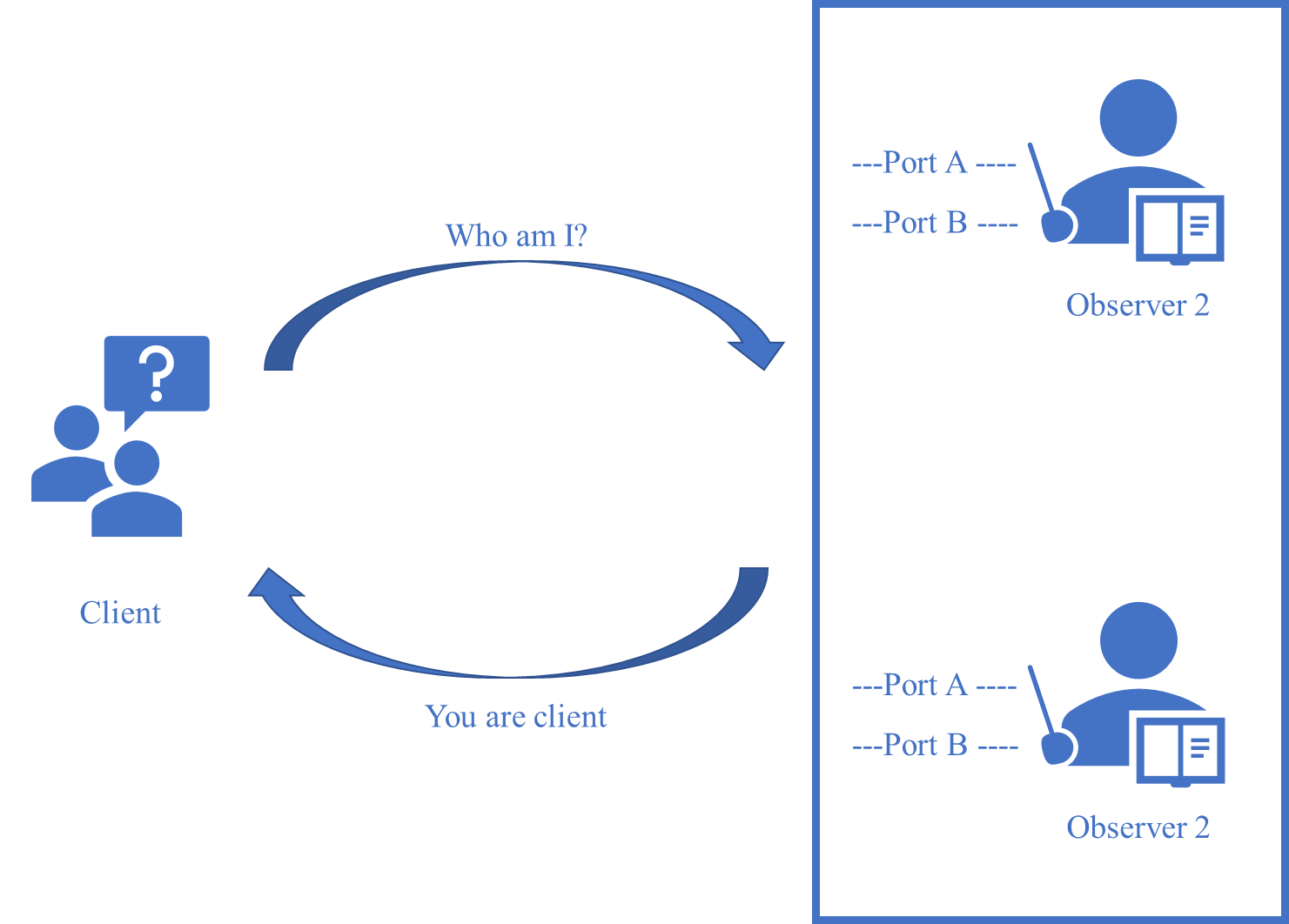
Description automatically generated

Hình . Thuật toán nhận diện NAT

### Gửi bản tin xác định địa chỉ IP

Để nhận biết được thiết bị của người dùng có đứng sau NAT hay NAT đó thuộc loại nào thì client chủ động xác minh “danh tính” với một hệ thống chứa hai observer (quan sát viên). Observer ở đây được hiểu là một hệ thống server giúp client nhận biết được “tôi là ai”. Hệ thống server giúp xác định

Bản tin xác định danh tính là bản tin UDP yêu cầu observer cung cấp địa chỉ IP và địa chỉ port công khai phía client.



Hình . Xác định danh tính client

Để xác minh được địa chỉ IP, client chủ động gửi đến port A của observer 1 bản tin UDP yêu cầu trả về thông tin địa chỉ IP công khai (public IP) của client. Khi server nhận được gói tin yêu cầu xác định địa chỉ IP, phía server sẽ nhận diện và phản hồi lại cho client một bản tin UDP chứa thông tin IP client yêu cầu.

Lúc này tại phía client sẽ xuất hiện hai tình huống như sau:

* Trường hợp 1: không nhận được gói tin phản hồi
* Trường hợp 2: nhận được gói tin phản hồi

Trong trường hợp 1, gói tin UDP đã bị chặn và không thể kết nối do đó không thể xác định được loại NAT phía người dùng. Đối với trường hợp 2, phía client sẽ tiến hành bóc tách bản tin để xác định địa chỉ IP.

### Kiểm tra sự tồn tại của NAT

Sau khi xác định được địa chỉ IP trong gói tin phản hồi, phía client thực hiện so sánh địa chỉ IP hiện tại (local IP) và địa chỉ IP nhận được từ phía server (public IP). Có hai trường hợp xảy ra khi so sánh 2 địa chỉ IP này:

* Trường hợp 1: local IP trùng public IP
* Trường hợp 2: local IP khác public IP

Đối với trường hợp 1, điều đó có nghĩa rằng phía client *không có NAT*. Khi trường hợp này xảy ra, phía client tiếp tục gửi một bản tin UDP như trên đến port A của observer 2. Nếu client nhận được phản hồi từ observer 2 thì điều đó có nghĩa rằng client không bị chặn bởi tường lửa (firewall) và ngược lại.

Đối với trường hợp 2, điều đó có nghĩa rằng client đang đứng sau NAT.

### Nhận biết loại NAT đang sử dụng

Với mỗi loại NAT khác nhau sẽ có phương thức truyền tải video khác nhau. Bởi vậy, khi client đã xác định được sự tồn tại của NAT cần xác định thêm NAT đó thuộc loại NAT nào trong 4 loại NAT đã được trình bày ở mục 2.3.

Để xác định được NAT đó thuộc loại NAT nào, client tiếp tục gửi một bản tin UDP xác định danh tính đến port B của observer 1. Nếu client nhận được phản hồi từ observer 1 thì điều đó có nghĩa rằng client đứng sau *NAT loại 1*.

Ngược lạ, nếu client không nhận đươc phản hồi từ phía observer 1 sẽ tiếp tục gửi một bản tin xác định danh tính đến port B của observer 2. Nếu tiếp tục không nhận được phản hồi từ observer 2, client đứng sau *NAT loại 4*.

Để xác định 2 NAT còn lại, client tiếp tục gửi bản tin đến port B của observer 1, nếu nhận được phản hồi, đó là *NAT loại 2* và trường hợp còn lại là *NAT loại 3*.

## Công nghệ WebRTC

Webrtc là một giải pháp đa nền tảng với khả năng RTC. Người ta có thể truyền phát luồng video của riêng mình từ camera hoặc màn hình hoặc bất kỳ video nào khác đến bất kỳ bạn bè nào qua webrtc. Nếu một người đang phát triển một trò chơi ngang hàng theo thời gian thực với chia sẻ dữ liệu thời gian thực giữa các đồng nghiệp, thì WebRTC là một trong những lựa chọn tốt nhất.

### Giới thiệu

Web Real-Time Communication (WebRTC) là một tiêu chuẩn mới và là nỗ lực của ngành công nghệ trong việc mở rộng mô hình trình duyệt web. Lần đầu tiên, các trình duyệt có thể trao đổi các tập tin đa phương tiện trực tiếp qua trong thời gian thực với các trình duyệt web khác theo kiểu mạng ngang hàng.

W3C và IETF đang cùng nhau xây dựng API JavaScript, thẻ HTML5 tiêu chuẩn và giao thức truyền thông cơ bản để thiết lập và quản lý kênh liên lạc đáng tin cậy giữa bất kỳ cặp trình duyệt web nào.

Mục tiêu tiêu chuẩn hóa là xác định API WebRTC cho phép ứng dụng web chạy trên mọi thiết bị, thông qua kết nối an toàn tới các thiết bị ngoại vi đầu vào (như webcam, micro,…), để trao đổi các tập tin đa phương tiện và dữ liệu thời gian thực. [11]

### Lịch sử ra đời

Ý tưởng phát triển WebRTC được nhóm kỹ sư chịu trách nhiệm cho Google Hangouts đưa ra từ năm 2009. Vào thời gian đó, để truyền tải video, hình ảnh trên web thìngười ta thường phải dùng Flash. Nhóm kỹ sư Hangouts lại không muốn sử dụng công nghệ này, nên họ bắt đầu tự làm một chuẩn riêng cho mình.

Vào tháng 5/2011, Google ra mắt một dự án nguồn mở dành cho việc giao tiếp thời gian thực giữa trình duyệt với nhau, và từ lúc này dự án mang tên WebRTC. Song song đó, Hiệp hội W3C và Tổ chức IETF cũng đang phát triển một số giao thức để dùng cho việc việc kết nối thời gian thực, vì vậy họ bắt tay nhau tiếp tục hoàn thiện và sau đó kết hợp chung vào WebRTC.

Ngày 27/10/2011, Hiệp hội W3C ra mắt bản “nháp” đầu tiên của WebRTC. Tháng 11/2011, Chrome23 ra mắt, trở thành trình duyệt đầu tiên có tích hợp WebRTC ngay từ bên trong. Công nghệ WebRTC vẫn đang tiếp tục được phát triển và liên tục cập nhật.

Đến nay, vào ngày 22/3/2016, Google đã phát hành phiên bản Chrome 50 (trong bản Beta Channel). Theo đó, trình duyệt Chrome 50 có sự hỗ trợ chuẩn video H.264 (codec video) của WebRTC.

Tiếp đó, ngày 29/8/2016, Mozilla đã phát hành phiên bản trình duyệt Firefox 49. Firefox 49 đã được tăng cường hỗ trợ tiêu chuẩn WebRTC trong kiểm soát băng thông. Các cải thiện quản lý băng thông bao gồm: cho phép cho phương tiện truyền thông một chiều và xử lý băng thông tốt hơn. Đồng thời, trình duyệt Firefox 49 cho phép kiểm tra NAT traversal trong các topo NAT khác nhau.

Hiện nay, WebRTC vẫn còn đang tiếp tục được phát triển. Mặc dù chưa hoàn thiện nhưng các nhà nghiên cứu và phát triển ứng dụng bắt đầu giới thiệu về công nghệ WebRTC dành cho các thiết bị di động chạy hệ điều hành IOS và Android. [12]

### Kiến trúc Web

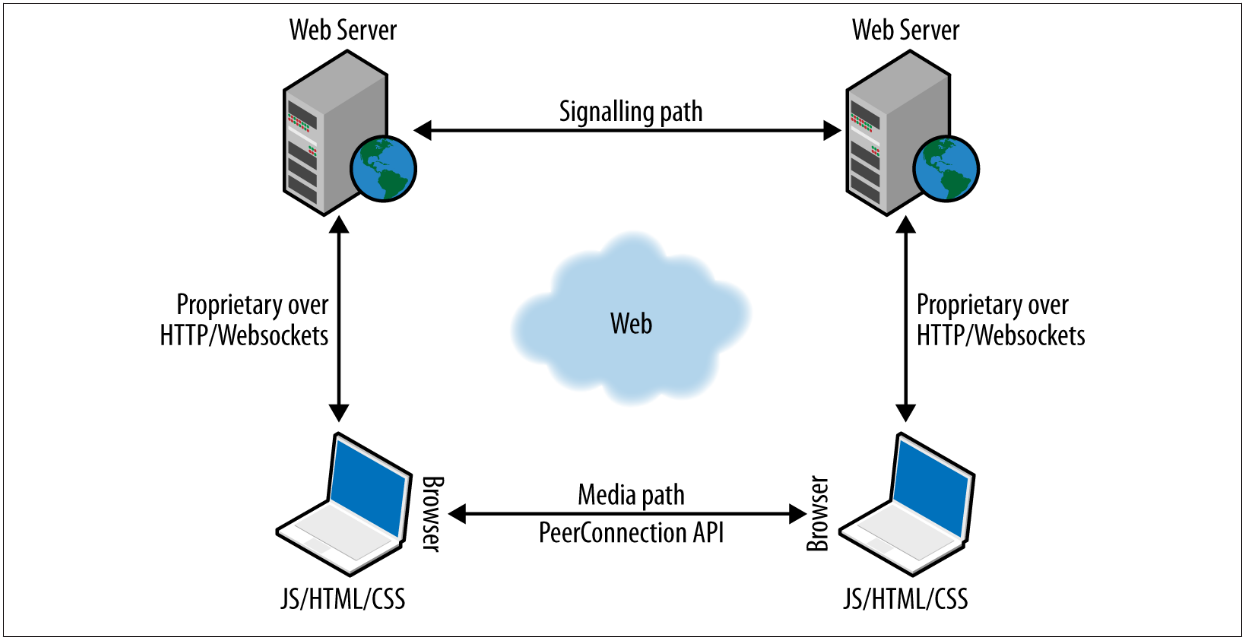
Kiến trúc web (web architecture) cổ điển là dựa trên mô hình client-server, nơi các trình duyệt gửi nội dung yêu cầu qua giao thức HTTP đến máy chủ và máy chủ web trả lời với phản hồi có chứa thông tin được yêu cầu.

Các tài nguyên được cung cấp bởi một máy chủ được liên kết chặt chẽ với một URI hoặc URL.

Trong kịch bản ứng dụng web, máy chủ nhúng một số mã JavaScript vào trang HTML và gửi lại cho máy khách. Mã như vậy có thể tương tác với các trình duyệt thông qua API JavaScript tiêu chuẩn và với người dùng thông qua giao diện người dùng. [11]

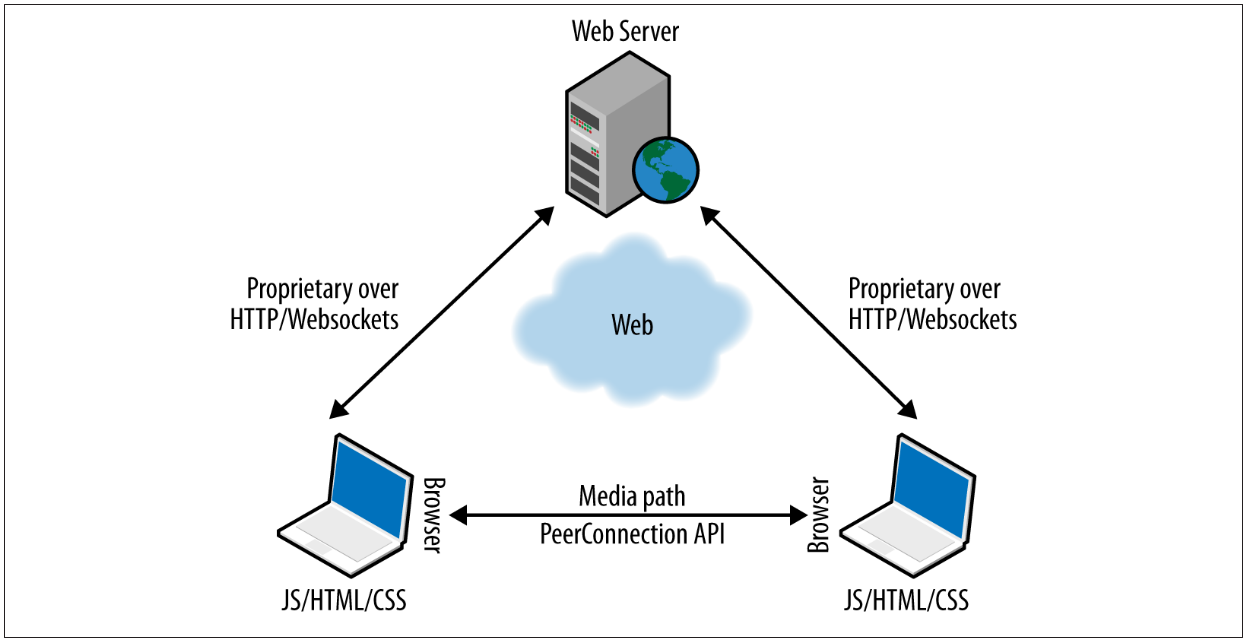
### Kiến trúc WebRTC

WebRTC mở rộng ngữ nghĩa của mô hình client-server bằng cách đưa ra mô hình giao tiếp mạng ngang hàng (peer-to-peer) giữa các trình duyệt. Mô hình kiến trúc WebRTC tổng quát nhất chính là mô hình WebRTC hình thang (Hình 2.18) được lấy cảm hứng từ giao thức khởi tạo phiên (SIP).



Hình . Mô hình kiến trúc WebRTC hình thang [11]

Trong mô hình WebRTC hình thang, cả hai trình duyệt đang chạy một ứng dụng web, được tải xuống từ hai máy chủ web. Tin nhắn báo hiệu (messages) được sử dụng để thiết lập các thông tin của các máy khách và kết thúc liên lạc. Tin nhắn được vận chuyển bởi giao thức HTTP hoặc WebSocket thông qua các máy chủ web (có thể sửa đổi, dịch chuyển hoặc quản lý chúng khi cần). Điều đáng chú ý là tín hiệu giữa trình duyệt và máy chủ không được chuẩn hóa trong WebRTC, vì nó được coi là một phần của ứng dụng. Đối với đường dẫn dữ liệu, PeerConnection cho phép các tập tin phương tiện truyền trực tiếp giữa các trình duyệt mà không cần bất kỳ máy chủ can thiệp nào. Hai máy chủ web có thể giao tiếp bằng giao thức báo hiệu tiêu chuẩn như SIP hoặc Jingle (XEP-0166). Nếu không, các máy chủ có thể sử dụng một giao thức báo hiệu độc quyền.



Hình . Kiến trúc WebRTC hình tam giác [11]

Kịch bản WebRTC phổ biến nhất có thể nói đến là kịch bản mà hai trình duyệt đang chạy cùng một ứng dụng web, được tải xuống từ cùng một máy chủ web. Trong trường hợp này, mô hình hình thang trở thành mô hình hình tam giác (Hình 2.19).

### Đặc điểm của WebRTC

Để sử dụng WebRTC, các nhà phát triển có thể sử dụng rất nhiều loại ngôn ngữ lập trình quen thuộc như: JavaScript (dành cho browser), Java (dành cho ứng dụng Android), Objective-C (dành cho ứng dụng IOS), C++ (dành cho ứng dụng Windows).

Tuy nhiên, có sự khác biệt về số lượng hàm API WebRTC được hỗ trợ trong các trình duyệt Web. Điều này làm giảm đi khả năng hoạt động của các ứng dụng WebRTC. Do đó, các nhà phát triển phần mềm sẽ phải dành nhiều công sức hơn để tinh chỉnh lại trang Web hoặc App của mình cho phù hợp với từng trình duyệt.

WebRTC chưa có chuẩn video thống nhất. Đặc điểm nổi bật nhất của WebRTC là khả năng kết nối video từ trình duyệt web của người sử dụng. Nhưng hiện nay, các hãng trình duyệt cũng chưa thống nhất với nhau là chuẩn video nào sẽ được dùng cho WebRTC

Google và Mozilla muốn dùng VP8 hoặc VP9, một codec video do chính Google phát triển theo mô hình mã nguồn mở. Thực tế cho thấy, đối với trình duyệt web Chrome 50 trở lên, codec VP9 tốt hơn. Nó hỗ trợ cho chất lượng tốt hơn trong truyền tải phương tiện truyền thông nhờ khả năng dự phòng (RED), cơ chế FEC tương tác với ước lượng băng thông và kiểm soát các thông số (NACK, POI, RTX, remb,...) trong trình duyệt, khả năng tương tác với các trình duyệt khác tốt hơn VP8. Do đó, VP9 cho độ phân giải cao hơn,giảm thiểu các lỗi (nhờ hỗ trợ cho Error Correction Forward).

Trong khi đó, Microsoft và một số công ty khác muốn đề xuất sử dụng H.264 hoặc H.265 cho WebRTC, vốn đang là codec được dùng phổ biến nhất hiện nay trên Internet. H.264 lại thuộc quyền sở hữu của Hiệp hội MPEG LA và phải trả phí bản quyền để sử dụng.

Hiện nay, nhóm làm việc của Google gồm Niklas Blum, Justin Uberti và Per Ahgren đã thực hiện được những cải tiến về hiệu suất, đó là trình duyệt Chrome đã có thể bớt thời gian kết nối trung bình từ 2 giây xuống còn 650ms cho một kết nối 1Mbps. Họ cũng đang làm việc để mang lại VP9 – codec video đến các thiết bị di động, nhờ đó sẽ cho phép đạt chất lượng video cao với yêu cầu băng thông ít hơn. [12]

### STUN

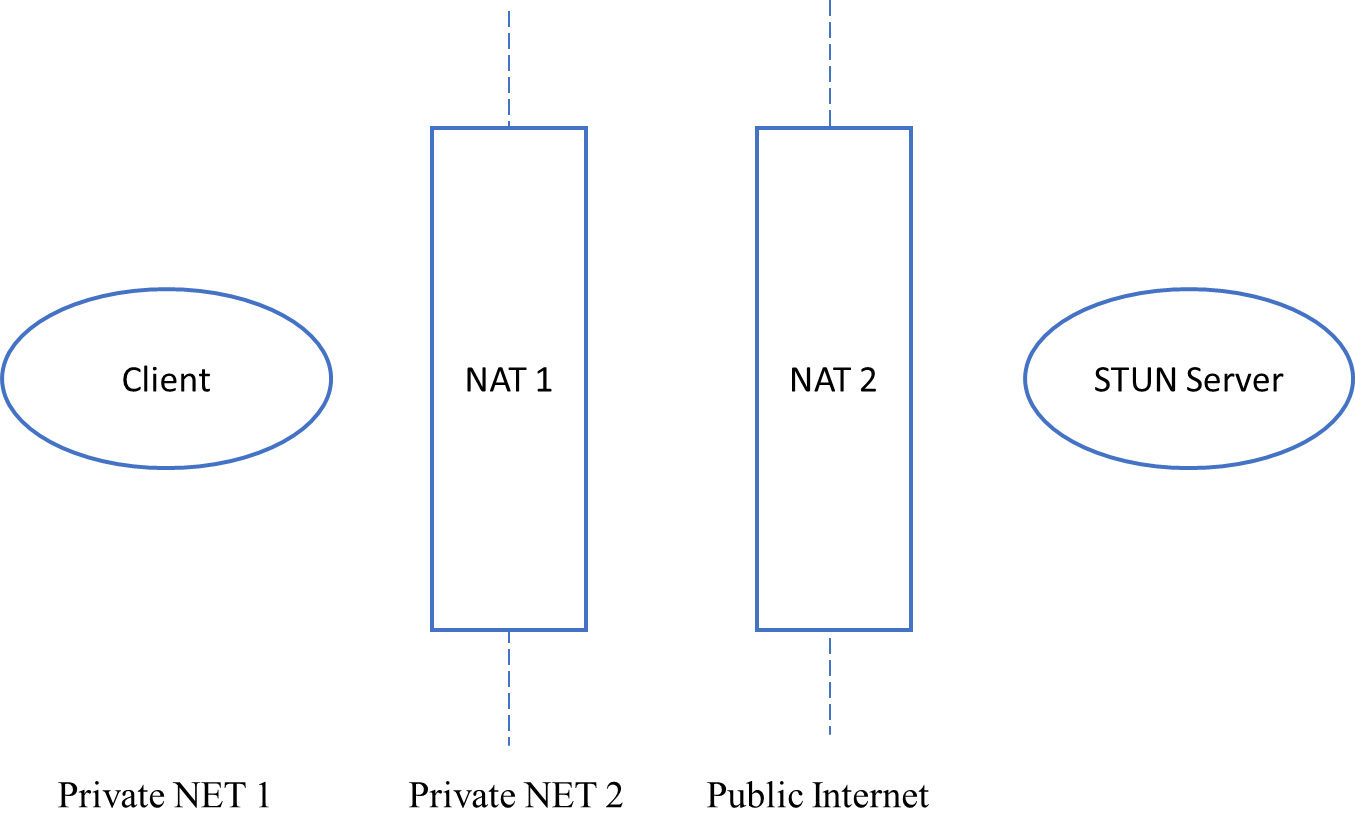
#### Giới thiệu

STUN là giao thức cho phép ứng dụng máy chủ phát hiện ra sự hiện diện của trình dịch địa chỉ mạng trên mạng internet. Sự phát hiện này cho phép máy chủ có được địa chỉ IP công cộng (public IP) và cổng kết nối (port) của máy khách. Để làm như vậy, giao thức cần có sự trợ giúp từ máy chủ STUN của bên thứ ba được công khai trên mạng internet. [11]

#### Hoạt động

Cấu hình STUN điển hình được hiển thị trong Hình 2.20. Máy khách được kết nối với mạng riêng 1 (private NET 1). Mạng này kết nối với mạng riêng 2 (private NET 2) qua NAT 1. Mạng riêng 2 kết nối với Internet công cộng thông qua NAT 2. Máy chủ STUN nằm trên mạng Internet công cộng.

STUN là một giao thức client-server đơn giản. Một khách hàng gửi yêu cầu đến một máy chủ và máy chủ trả về phản hồi. Có hai loại yêu cầu là yêu cầu liên kết, được gửi qua UDP và Yêu cầu bí mật chung, được gửi qua TLS qua TCP. Yêu cầu bí mật được chia sẻ yêu cầu máy chủ trả lại tên người dùng và mật khẩu tạm thời. Tên người dùng và mật khẩu này được sử dụng trong yêu cầu liên kết và phản hồi ràng buộc tiếp theo được sử dụng cho mục đích xác thực và toàn vẹn thông điệp.



Hình . Tổng quan hoạt động của STUN

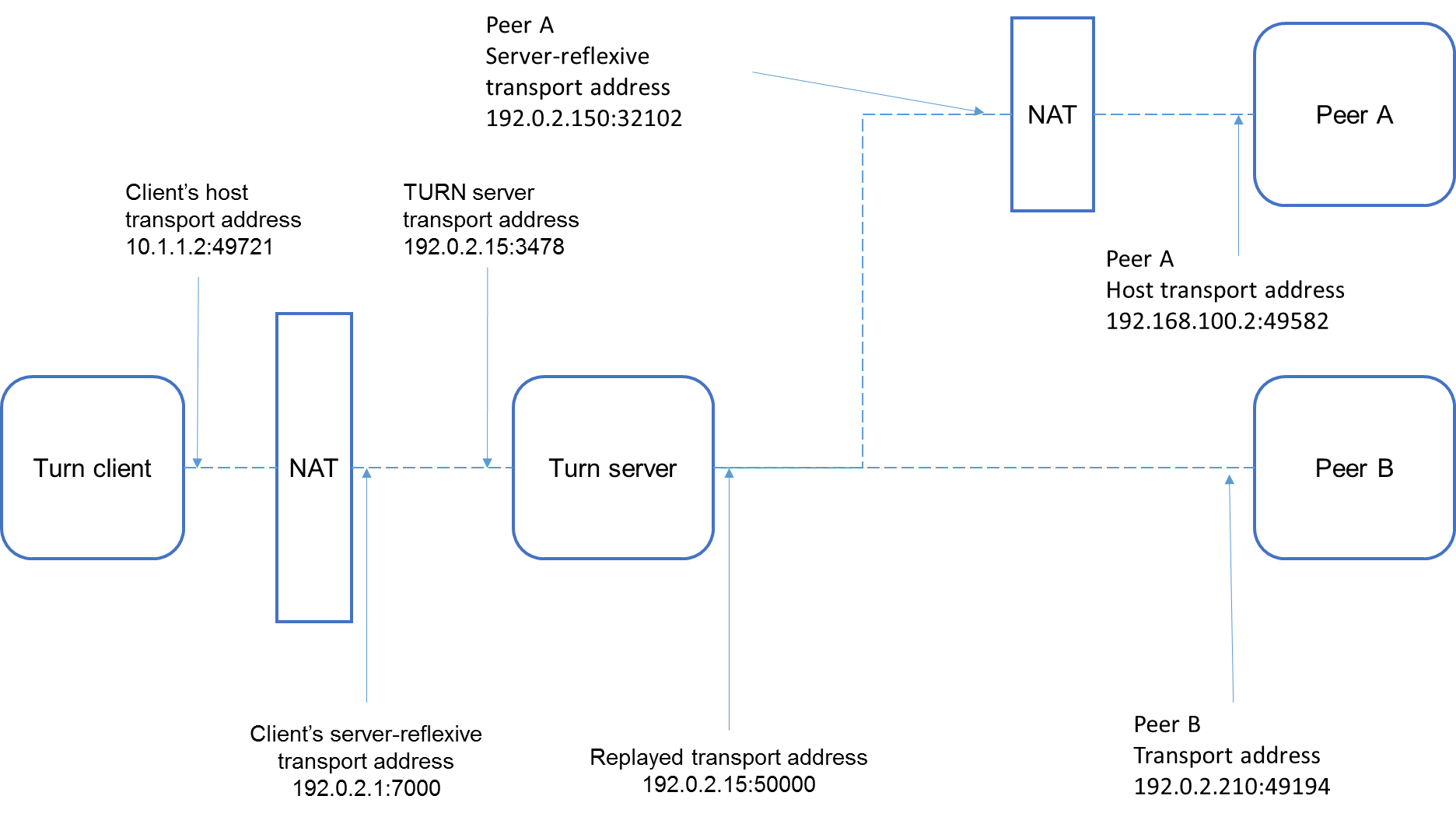
Yêu cầu liên kết STUN được sử dụng để phát hiện lại NAT được sử dụng, địa chỉ IP công cộng và cổng liên kết được ánh xạ. Yêu cầu liên kết được gửi đến máy chủ STUN bằng UDP. Yêu cầu liên kết trước khi đến máy chủ STUN có thể đã được chuyển qua một hoặc nhiều NAT giữa máy khách và máy chủ STUN. Do đó, địa chỉ nguồn của yêu cầu mà máy chủ nhận được sẽ là địa chỉ được ánh xạ được tạo bởi NAT gần nhất. Máy chủ STUN sao chép địa chỉ IP nguồn và địa chỉ cổng liên kết vào phản hồi ràng buộc STUN và gửi lại địa chỉ IP nguồn và cổng của yêu cầu STUN. Đối với tất cả các loại NAT ở trên, phản hồi này sẽ được gửi đến máy khách. [13]

### TURN

TURN là giao thức cho phép máy chủ phía sau NAT lấy địa chỉ IP công cộng và cổng mạng từ máy chủ chuyển tiếp trên Internet. Nhờ địa chỉ truyền tải được chuyển tiếp, máy chủ có thể nhận diện được bất kỳ phương tiện từ máy ngang hàng nào có thể gửi gói đến mạng Internet công cộng. [11]

Trong cấu hình thông thường, máy khách TURN được kết nối với mạng riêng và thông qua một hoặc nhiều NAT với Internet công cộng. Trên Internet công cộng là một máy chủ TURN. Ở những nơi khác trên Internet là một hoặc nhiều peer mà máy khách TURN muốn liên lạc. Những peer này có thể hoặc không thể đứng sau một hoặc nhiều NAT. Máy khách sử dụng máy chủ làm rơle để gửi và nhận các gói từ các peer.

Hình 2.21 là một mô hình điển hình mô tả hoạt động của TURN. Trong hình này, máy khách TURN và máy chủ TURN được phân tách bằng NAT. Máy khách nói chuyện với máy chủ từ tập hợp các (địa chỉ IP, cổng) được gọi là HOST TRANSPORT ADDRESS của máy khách. Máy khách sẽ gửi các tin nhắn TURN từ địa chỉ vận chuyển máy chủ của nó đến một địa chỉ vận chuyển trên máy chủ TURN được gọi là TURN SERVER TRANSPORT ADDRESS. Máy khách tìm kiếm địa chỉ máy chủ TURN thông qua một số phương tiện không xác định (ví dụ: cấu hình) và địa chỉ này thường được nhiều khách hàng sử dụng.



Hình . Tổng quan hoạt động của TURN

Vì máy khách đứng sau NAT, nên máy chủ sẽ thấy các gói từ máy khách gửi đến từ địa chỉ NAT. Địa chỉ này được gọi là địa chỉ vận chuyển SERVER-REFLEXIVE của máy khách; các gói được gửi bởi máy chủ đến địa chỉ vận chuyển phản xạ của máy chủ sẽ được NAT chuyển tiếp đến địa chỉ vận chuyển của máy khách.

Máy khách sử dụng các lệnh TURN để tạo và thao tác ALLOCATION trên máy chủ. Allocation là một cấu trúc dữ liệu trên máy chủ. Cấu trúc dữ liệu này chứa các thông tin của RELAYED TRANSPORT ADDRESS. Địa chỉ vận chuyển được chuyển tiếp là địa chỉ vận chuyển trên máy chủ mà các máy peer có thể sử dụng để có dữ liệu chuyển tiếp từ máy chủ đến máy khách. Một phân bổ được xác định duy nhất bởi địa chỉ vận chuyển chuyển tiếp của nó.

Khi allocation được tạo, máy khách có thể gửi dữ liệu ứng dụng đến máy chủ cùng với dấu hiệu cho biết dữ liệu nào được gửi ngang hàng và máy chủ sẽ chuyển tiếp dữ liệu này đến máy peer thích hợp. Máy khách sẽ gửi dữ liệu ứng dụng đến máy chủ bên trong tin nhắn TURN. Tại máy chủ, dữ liệu được trích xuất từ tin nhắn TURN và được gửi đến máy peer trong một bản tin UDP. Theo hướng ngược lại, một máy peer có thể gửi dữ liệu ứng dụng trong một bản tin UDP đến địa chỉ vận chuyển được chuyển tiếp để phân bổ. Sau đó, máy chủ sẽ đóng gói dữ liệu này trong một tin nhắn TURN và gửi nó đến máy khách cùng với một dấu hiệu cho thấy peer nào đã gửi dữ liệu. Vì thông báo TURN luôn chứa một dấu hiệu cho biết máy khách đang liên lạc với ai, nên máy khách có thể sử dụng một phân bổ duy nhất để liên lạc với nhiều peer.

Khi peer đứng sau NAT, thì máy khách phải xác định peer bằng cách sử dụng địa chỉ vận chuyển theo phản xạ của máy chủ thay vì địa chỉ vận chuyển của máy chủ. Ví dụ: để gửi dữ liệu ứng dụng đến Peer A trong ví dụ trên, khách hàng phải chỉ định 192.0.2.150:32102 thay vì 192.168.100.2: 49582.

Mỗi phân bổ trên máy chủ thuộc về một máy khách và có chính xác một địa chỉ vận chuyển được chuyển tiếp chỉ được sử dụng bởi phân bổ đó. Do đó, khi một gói đến địa chỉ truyền tải được chuyển tiếp trên máy chủ, máy chủ sẽ biết dữ liệu nào được gửi cho máy khách. Tại một thời điểm máy khách có thể có nhiều phân bổ trên một máy chủ.[13]

### SDP

SDP được WebRTC sử dụng để đàm phán các thông số của session. Do không có tín hiệu trong WebRTC, SDP được tạo và sử dụng bởi WebRTC được coi là được truyền thông bởi ứng dụng chứ không phải bởi chính WebRTC.

### ICE

ICE là một framework giúp kết nối các peer lại với nhau. ICE được sử dụng như người chỉ dẫn đường giúp tìm đường dẫn tốt nhất cho mỗi cuộc gọi video. Đồng thời ICE cũng giúp thu thập tất cả các kết nối đã có sẵn như: địa chỉ IP cục bộ, địa chỉ phản xạ - địa chỉ STUN và địa chỉ được chuyển tiếp – TACK...

## Kết luận

Dựa trên những kiến thức tìm hiểu về các công nghệ trong lĩnh vực truyền thông đa phương tiện, tôi đã lựa chọn ra được cho mình công nghệ phù hợp để giải quyết yêu cầu đặt ra của đề tài. Cùng với các kiến thức đã tìm hiểu được, tôi đã trình bày ngắn gọn về các công nghệ liên quan cũng như công nghệ sử dụng trong đề tài.

# PHÂN TÍCH THIẾT KẾ HỆ THỐNG

Chương 4 trình bày về phân tích yêu cầu nghiệp vụ, đưa ra các yêu cầu chức năng, phi chứ năng và quy trình hoạt động của hệ thống.

## Tổng quan hệ thống

Diagram

Description automatically generated

Hình . Sơ đồ tổng quan hệ thống

Hệ thống gồm có:

* Mobile app (iOS, Android), Desktop app (Java app): Là máy khách được cài đặt sẵn các ứng dụng có chức năng thực hiện các cuộc gọi video.
* Server: Hệ thống máy chủ (server) gồm hai thành phần chính là Signaling server và Turn server. Signaling server cho phép máy khách kết nối đến máy chủ để lấy các thông tin về cuộc gọi. Signaling server là thành phần chính trong máy chủ giúp thiết lập sự bắt tay giữa các người dùng và đóng vai là là “trọng tài” trung gian để truyền tải video trong thời gian thực. Turn server đóng vài trò chung gian chuyền tải media data giữa các client trong trường họp client ko thể kết nối trực tiếp với nhau (có thể bị chặn công Nat, Firewall, …).

Luồng hoạt động chính

Diagram

Description automatically generated

Hình . Sơ đồ activity

## Yêu cầu hệ thống

### Yêu cầu chứ năng

* Xác thực người dùng
* Danh sách bạn bè online
* Gọi video một người: chọn 1 bạn bè từ danh sách online và không đang trong cuộc gọi khác
* Gọi video nhiều người: chọn 1 hoặc nhiều bạn bè từ danh sách online và không đang trong cuộc gọi khác

### Yêu cầu phi chức năng

* Giao diện chương trình thân thiện với người dùng.
* Cho phép người dùng tương tác video với người dùng khác trên internet.
* Hỗ trợ hệ điều hành Android phiên bản 5.0 trở lên.
* Hỗ trợ hệ điều hành IOS phiên bản 12 trở lên.
* Hỗ trợ hệ điều hành Mac, Windows, Linux có môi trường Java 8.
* Hệ thống hỗ trợ nhiều người dùng.
* Hệ thống cho phép nhiều người dùng cùng chat video cùng lúc
* Ứng dụng hỗ trợ đa ngôn ngữ.

## Đặc tả và phân tích hệ thống

### Sơ đồ use case tổng quan

Diagram

Description automatically generated

Hình . Sơ đồ use case tổng quan

Bảng . Các use case

|  |  |  |
| --- | --- | --- |
| **Tác nhân** | **ID** | **Tên use case** |
| User | UC-US-01 | Đăng nhập |
| UC-US-02 | Xem danh sách bạn bè online |
| UC-US-03 | Gọi video một người |
| UC-US-04 | Gọi video nhiều người |

### Sơ đồ use case đăng nhập

Bảng . Use case đăng nhập

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tên Use-case**: Đăng nhập | | **Id**: UC-US-01 | | **Mức quan trọng**: cao |
| **Tác nhân chính**: Người dùng | | | **Loại use-case**: chi tiết, cần thiết | |
| **Mô tả ngắn** | Người dùng thực hiện thao tác đăng nhập | | | |
| **Ràng buộc** | Người dùng chưa đăng nhập | | | |
| **Các mối quan hệ** | Không có | | | |
| **Luồng sự kiện chính** | Người dùng đăng nhập hệ thống  Hiển thị giao diện đăng nhập  Nhập thông tin đăng nhập  Kiểm tra thông tin đăng nhập  Hiển thị giao diện người dùng | | | |
| **Luồng sự kiện con** | Không có | | | |
| **Luồng sự kiện ngoại lệ** | Không thể kết nối đến hệ thống máy chú  Thông tin đăng nhập không hợp lệ | | | |

### Use case danh sách bạn bè online

Bảng . Use case danh sách bạn bè online

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tên Use-case**: Danh sách online | | **Id**: UC-US-02 | | **Mức quan trọng**: cao |
| **Tác nhân chính**: Người dùng | | | **Loại use-case**: chi tiết, cần thiết | |
| **Mô tả ngắn** | Người dùng thực hiện thao tác vào màn hình danh sách bạn bè đang online | | | |
| **Ràng buộc** | Người dùng đã đăng nhập | | | |
| **Các mối quan hệ** | Không có | | | |
| **Luồng sự kiện chính** | Lấy danh sách bạn bè onlnie  Hiển thị giao diện danh sách bạn bè online | | | |
| **Luồng sự kiện con** | Không có | | | |
| **Luồng sự kiện ngoại lệ** | Không thể kết nối đến hệ thống máy chú | | | |

### Gọi video một người

Bảng . Use case gọi video một người

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tên Use-case**: Gọi video một người | | **Id**: UC-US-03 | | **Mức quan trọng**: cao |
| **Tác nhân chính**: Người dùng | | | **Loại use-case**: chi tiết, cần thiết | |
| **Mô tả ngắn** | Người dùng thực hiện thao tác gọi video với một người khác | | | |
| **Ràng buộc** | Người dùng đã đăng nhập | | | |
| **Các mối quan hệ** | Không có | | | |
| **Luồng sự kiện chính** | Người dùng họn một người dùng để gọi video  Gửi yêu cầu đến người được gọi  Hiển thị màn hình cuộc gọi | | | |
| **Luồng sự kiện con** | Không có | | | |
| **Luồng sự kiện ngoại lệ** | Không thể kết nối đến hệ thống máy chú  Người được gọi không phản hồi  Người được gọi từ chối cuộc gọi | | | |

### Gọi video nhóm nhiều người

Bảng . Use case gọi video nhiều người

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tên Use-case**: Gọi video nhiều người | | **Id**: UC-US-04 | | **Mức quan trọng**: cao |
| **Tác nhân chính**: Người dùng | | | **Loại use-case**: chi tiết, cần thiết | |
| **Mô tả ngắn** | Người dùng thực hiện thao tác đăng nhập | | | |
| **Ràng buộc** | Người dùng đã đăng nhập | | | |
| **Các mối quan hệ** | Không có | | | |
| **Luồng sự kiện chính** | Người dùng chọn gọi nhóm nhiều người  Gửi yêu cầu đến những người trong nhóm  Hiển thị màn hình cuộc gọi | | | |
| **Luồng sự kiện con** | Không có | | | |
| **Luồng sự kiện ngoại lệ** | Không thể kết nối đến hệ thống máy chú  Người được gọi không phản hồi  Người được gọi từ chối cuộc gọi | | | |

### Class Diagram

#### Signaling

Diagram

Description automatically generated

Hình . Biểu đồ class của Signaling server

#### Turn

Diagram

Description automatically generated

Hình . Biểu đồ class của Turn server

#### Client

Diagram

Description automatically generated

Hình . Biểu đồ class của Client

### Sequence Diagram

#### Signaling

Diagram, engineering drawing

Description automatically generated

Hình . Biểu đồ Sequence của Signaling server

#### Turn

Diagram, schematic

Description automatically generated

Hình . Biểu đồ Sequence của Turn server

#### Client

Diagram

Description automatically generated

Hình . Biểu đồ Sequence của Client

### Cơ sở dữ liệu

Diagram

Description automatically generated

Hình . Cơ sở dữ liệu

## Ứng dụng

### Công nghệ sử dụng

* Server Signaling: sử dụng Java Socket TCP làm hệ thống định danh, gửi các yêu cầu và phản hồi giữa máy khách.
* Server Turn: sử dụng Java Socket UDP dùng để chung truyển media giữa các máy khách.
* Ứng dụng desktop: sử dụng ngôn ngữ lập trình Java Swing lập trình giao UI Desktop.
* Ứng dụng trên nền tảng Android: xây dựng trên ngôn ngữ lập trình Java-Android.
* Ứng dụng trên nền tảng IOS: xây dựng trên ngôn ngữ lập trình Swift.
* Quản lý mã nguồn: sử dụng GitHub là nơi lưu trữ mã nguồn.
* Google Cloud Platform: được cung cấp bởi Google, là một bộ dịch vụ điện toán đám mây chạy trên cùng một cơ sở hạ tầng.
* Trình soạn thảo mã nguồn (IDE) - Xcode: bộ phát triển phần mềm tích hợp được Apple phát triển chạy trên hệ điều hành Mac để các lập trình viên có thể phát triển phần mềm chạy trên hệ điều hành Mac và iOS
* Trình soạn thảo mã nguồn (IDE) – IntelliJ: IDE Java để phát triển các phần mềm máy tính. Nó được phát triển bởi JetBrains, nó được cấp phép Apache 2 cho phiên bản cộng đồng, và một phiên bản thương mại độc quyền.Cả hai có thể được sử dụng cho phát triển thương mại.

### Màn hình login

Màn hình hiển thị chức năng đăng nhập và trạng thái đăng nhập của người dùng.

Khi người dùng đăng nhập xong có thể thực hiện các tính năng gọi một người hoặc gọi nhóm với danh sách bạn bè đang online.

Graphical user interface, application

Description automatically generated

Hình . Giao diện màn hình chính

### Danh sách bạn bè

Chọn 1 hoặc nhiều bạn bè trong danh sách để bắt đầu cuộc gọi.

Graphical user interface, application

Description automatically generated

Hình . Giao diện danh sách bạn bè

### Màn hình gọi video

Hiển thị giao diện video call một người hoặc nhiều người.



Hình . Giao diện màn hình cuộc gọi video

# KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

## Kết quả đạt được

Sau khoảng thời gian tìm hiểu và nghiên cứu tôi đã xây dựng được một hệ thống cho phép người dùng thực hiện video call hai chiều qua mạng internet.

Trong quá trình thực hiện đề tài, tôi đã đạt được một số các kết quả như sau:

* Nâng cao kỹ năng lập trình đặc biệt là kỹ năng lập trình bằng ngôn ngữ Android (Java), iOS (Swift), Java Swing.
* Tìm hiểu các công nghệ phục vụ trong lĩnh vực truyền thông đa phương tiện thời gian thực.
* Xây dựng các ứng dụng cho phép người dùng thực hiện video call hai chiều với một hoặc nhiều người dùng thông qua mạng internet công cộng.
* Triển khai các ứng dụng trên thiết bị thật.

Bên cạnh những kết quả trên thì vẫn còn đó những mặt hạn chế trong việc xây dựng hệ thống. Các khó khăn đấy là:

* Trải nghiệm người dùng của ứng dụng còn thấp, chưa đáp ứng được yêu cầu ban đầu đề ra.
* Ứng dụng chưa thể đáp ứng được số lượng lớn người dùng cùng thực hiện tương tác với nhau trong cùng một phòng chat video.
* Ứng dụng vẫn còn bị lỗi trong khi thực hiện thao tác kết thúc cuộc gọi.

## Phương hướng phát triển

Với các thành tựu đạt được hiện nay, sản phẩm vẫn chưa sẵn sàng đáp ứng các yêu cầu của người dùng. Do đó sản phẩm cần được cải thiện và phát triển hơn nữa. Sau đây là những định hướng phát triển sản phẩm tương tác video hai chiều hỗ trợ nhiều người dùng trong tương lai:

* Tìm hiểu nguyên nhân và tiến hành vá lỗi treo chương trình khi thoát khỏi cửa sổ video call.
* Phát triển ứng dụng tương tác video hai chiều hỗ trợ nhiều người dùng với các thiết bị sử dụng Web.
* Phát triển thêm các chức năng mới như: nhắn tin trực tiếp, thực hiện video call có thêm các hiệu úng giữa hai máy khách với nhau,…
* Cải thiện giao diện và tăng tương tác trải nghiệm người dùng bằng các hình ảnh trực quan, các hiệu ứng chuyển động đẹp, âm thanh và cảm ứng rung khi tương tác với ứng dụng.
* Cải thiện hiệu năng của chương trình bằng cách giảm sự tiêu tốn tài nguyên phần cứng như CPU, RAM, băng thông mạng sử dụng xuống mức tối thiểu.
* Triển khai hệ thống đến tay người dùng.

# TÀI LIỆU THAM KHẢO

1. <https://www.statista.com> truy nhập cuối cùng ngày 10/1/2021
2. Đỗ Hoàng Tiến và Dương Thanh Phương*, Giáo trình kỹ thuật truyền hình*. Nhà xuất bản khoa học và kỹ thuật, 2004.
3. Mai Thị Lan Oanh. (2005) *Các chuẩn nén và ứng dụng truyền video trên mạng internet* [Online]. Có sẵn: <https://www.academia.edu/>
4. Yoss Cohen. (2010) *H.264 nal and RTP* [Online]. Có sẵn: <https://slideshare.net/>
5. Jim Bankoski, Paul Wilkins and Yaowu Xu. (2012) *Technical overview of VP8, an open source video codec for the web* [Online]. <https://static.googleusercontent.com/>

1. <https://www.wowza.com/> truy cập cuối cùng ngày 10/1/2021

1. <https://www.geeksforgeeks.org/> truy nhập cuối cùng ngày 10/1/2021

1. <https://www.globaldots.com/> truy nhập cuối cùng ngày 10/1/2021

1. <https://developer.mozilla.org/> truy nhập cuối cùng ngày 10/1/2021

1. <https://www.haivision.com/> truy nhập cuối cùng ngày 10/1/2021
2. Salvatore Loreto and Simon Pietro Romano, “Introduction”, in *Real-Time Communication with WebRTC,* Simon St.Laurent and Allyson MacDonald. O’Reilly Media, May 2014, pp. 1-10.
3. ThS. Hoàng Thị Kim Chi. (2017). *WebRTC - Công nghệ web truyền thông thời gian thực* [Online]. Có sẵn: <https://www1.napa.vn>/

1. <https://ietf.org/> truy nhập cuối cùng ngày 10/1/2021