# Chương I. TỔNG QUAN VỀ CÔNG NGHỆ STREAMING VIDEO TRÊN NỀN GIAO THỨC HTTP

## 1.1 Giới thiệu chương

Trong chương này đồ án sẽ giới thiệu tổng quan về công nghệ streaming video nói chung và công nghệ streaming video trên nền giao thức HTTP nói riêng. Trình bày các cơ sở lý thuyết nền tảng như các khái niệm, các bộ giao thức cốt lõi trong streamimg video. Đồng thời nêu lên những khó khăn hiện tại trong streaming video hiện này để từ đó đưa ra giải pháp thực hiện trong các chương sau.

## 1.2 Tổng quan về truyền thông đa phương tiện

### 1.2.1 Đa phương tiện và truyền thông đa phương tiện

**Đa phương tiện (multimedia)** là nội dung mà sử dụng kết hợp các hình thức nội dung khác nhau như văn bản, âm thanh, hình ảnh, hoạt hình, video và các nội dung có tính tương tác. Đa phương tiện có thể được ghi lại và được phát, hiển thị, tương tác, hoặc được truy cập bởi các thiết bị xử lý thông tin như các thiết bị điện tử, máy tính. Tuy nhiên, đa phương tiện cũng có thể là một phần của một cuộc trình diễn trực tiếp. Các thiết bị đa phương tiện là các thiết bị truyền thông điện tử được sử dụng để lưu trữ và giúp người dùng trải nghiệm nội dung đa phương tiện.

**Truyền thông đa phương tiện (streaming media)** là hình thức đa phương tiện được truyền và trình bày (chơi, phát, hiển thị) tới người dùng cuối một cách liên tục trong khi vẫn được truyền bởi một nhà cung cấp dịch vụ đa phương tiện. Ý nghĩa của động từ “stream” ở đây là nói về quy trình của việc truyền và nhận được đa phương tiện. Thuật ngữ này chỉ phương thức truyền và nhận của các loại đa phương tiện chứ không chỉ riêng mình các loại đa phương tiện hay hình thức tải toàn bộ nội dung của các tập tin đa phương tiện trước khi trình phát chúng. Người dùng cuối có thể dùng một chương trình để trình phát nội dung đa phương tiện trong khi một phần nội dung của đa phương tiện vẫn đang được truyền tải.

**Tốc độ bit mã hoá đa phương tiện (bitrate)** là số bit được mã hoá trong một đơn vị thời gian được tính bằng kích thước của tệp đa phương tiện tính bằng byte chia cho thời gian phát lại bản ghi tính bằng giây và nhân với tám.

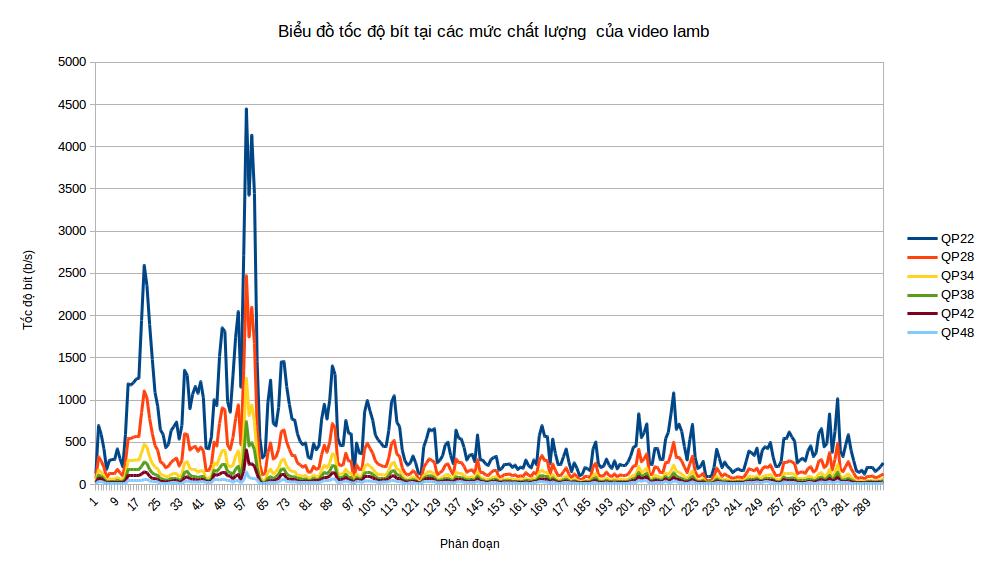
**Video** là một đa phương tiện điện tử để ghi, sao chép, phát lại, truyền tải và hiển thị một cách trực quan của chuyển động. Video có những đặc trưng như số khung hình trên giây, tỷ lệ khung hình, mô hình màu và độ sâu, chất lượng video.Video thường được lưu trữ và truyền đi sau khi mã hoá. Mục đích của việc mã hoá video là để giảm đi lượng dữ liệu dư thừa nhằm tiếc kiệm không gian lưu trữ cũng như băng thông khi truyền đi.

**Chất lượng video** là một đặc trưng của video được đo thông qua một hệ thống truyền hoặc xử lý video. Một phép đo chính thức hoặc không chính thức sự suy giảm chất lượng của video so với video gốc. Một hệ thống xử lý, truyền nhận video có thể gây ra một lượng nhiễu hay méo nhất định đến tín hiệu video, sự biến động của thông lượng mạng gây nên sự gián đoạn tại phía người dùng cuối. Những điều này ảnh hưởng tiêu cực đến trải nghiệm của người dùng. Do vậy với các bên liên quan như nhà cung cấp nội dung, nhà cung cấp dịch vụ và nhà điều hành mạng việc đảm bảo chất lượng video làm một nhiệm vụ quan trọng.

Một trong những tham số quan trọng quyết định đến chất lượng video được mã hoá chính là tham số lượng tử QP. Khi QP rất nhỏ mọi thông tin của video đều được giữ lại, khi tăng QP một số thông tin được loại bỏ để giảm dung lượng và tốc độ bit cho video. Vì vậy, QP được coi là tham số đặc trưng cho mức chất lượng của video có giá trị nằm trong khoảng từ 0 đến 51 . Với mỗi mức chất lượng cụ thể, các video được chia thành nhiều phân đoạn. Độ dài của mỗi phân đoạn này thường nằm trong khoảng từ 2 đến 10 giây .

Có nhiều cách và tiêu chí để phân loại video như nội dung, độ dài, hay chất lượng, tuy nhiên đồ án chỉ tập trung vào phương pháp phân loại video theo tốc độ bit mã hoá của video. Theo tiêu chí này video có thể được chia làm hai loại như sau:

**Video VBR** là video có tốc độ bit biến đổi trong mỗi phân đoạn của video, hay nói cách khác lượng dữ liệu đầu ra của bộ mã hoá trên mỗi phân đoạn thời gian là khác nhau. Với cách mã hoá tốc độ bit biến đổi (VBR), mỗi mức chất lượng đặc trưng bởi tham số QP cố định, mỗi khung hình (hoặc phân đoạn) được mã hoá với kích thước khác nhau tuỳ thuộc vào nội dung video, làm cho chất lượng đầu ra không thay đổi trong khi tốc độ bit video thay đổi một cách tự nhiên. Điều này làm cho chất lượng của các khung hình được mã hoá sẽ tốt hơn khi sử dụng một tốc độ bit cố định. Hình 1.1 biểu diễn tốc độ bit của một video được mã hoá theo kiểu tốc độ bit biến đổi với các mức tham số lượng tử khác nhau.



Hình 1. 1 Tốc độ bit tại các phân đoạn một video VBR ứng với tham số lượng tử

**Video CBR** Một dạng mã hoá video khác được gọi là mã hoá với tốc độ bit cố định (CBR). Khác với VBR mã hoá video CBR với mỗi mức chất lượng thì tốc độ bit được giữ cố định trong suốt quá trình mã hoá hay tốc độ dữ liệu đầu ra của bộ mã hoá là không đổi. Mã hoá CBR rất hữu ích trong streaming các nội dung đa phương tiện trên kênh truyền có dung lượng hạn chế vì nó cho biết chính xác tốc độ bit thay vì chỉ biết tốc độ bit trung bình. Tuy nhiên CBR có nhược điểm là đối với những phân đoạn video nhiều cảnh tĩnh, ít chuyển động thì bị phân bố dư thừa dữ liệu. Và ngược lại những cảnh nhiều chuyển động lại không phân đủ dữ liệu dẫn đến chất lượng video đầu ra không ổn định.

### 1.2.2 Một số thuật ngữ trong truyền thông đa phương tiện

Băng thông và thông lượng là hai khái niệm thường đi với nhau khi nói về một đường truyền vật lý, trong đó:

**Băng thông (bandwidth)** tốc độ truyền dữ liệu tối đa theo lý thuyết trong một đơn vị thời gian của một đường truyền vật lý, đơn vị đo lường là bps (bit per secon). Băng thông được dùng đặc trưng như băng thông mạng, băng thông dữ liệu, băng thông số. Khái niệm băng thông này khác với khái niệm băng thông trong lĩnh vực xử lý tín hiệu được đo bằng đơn vị hertz có ý ngĩa là khoảng tần số giữa tần số thấp nhất và tần số cao nhất đạt được

**Thông lượng (throughput)** là lượng dữ liệu thực tế được gửi qua một kết nối đường truyền vật lý trong một đơn vị thời gian, đơn vị bps. Thông lượng luôn luôn nhỏ hơn hoặc bằng băng thông. Đối với video thông lượng được đo sau khi nhận được một phân đoạn video. Thông lượng được tính bằng tỉ số giữa kích thước dữ liệu của phân đoạn và khoảng thời gian tải phân đoạn đó. Khoảng thời gian tải phân đoạn đó được tính từ thời điểm máy khách gửi truy vấn cho đến khi nó nhận được byte cuối cùng của phân đoạn được truy vấn. Goodput là thông lượng tại tầng ứng dụng của một kết nối hay lượng dữ liệu thu được tại lớp ứng dụng trên một đơn vị thời gian. Goodput luôn nhỏ hơn hoặc bằng throughput và throughtput luôn nhỏ hơn hoặc bằng bandwidth.

**QoE (Quality of Experience)** là mức độ hài lòng hoặc khó chịu của người dùng một ứng dụng hoặc dịch vụ nào đó. Nó là kết quả của việc đáp ứng mong đợi của người đó đối với tiện ích và/hoặc sự thích thú đối với ứng dụng hoặc dịch vụ trong tính cách của người dùng và trạng thái hiện tại của họ. Định nghĩa này đã được Liên minh Viễn thông quốc tế (International Telecommunication Union - ITU) thông qua năm 2016 và đang được sử dụng rộng rãi trong cộng đồng để đánh giá chất lượng của một ứng dụng hoặc dịch vụ.

QoE nhằm mục đích xem xét mọi yếu tố đóng góp vào chất lượng cảm nhận của người dùng đối với hệ thống hoặc dịch vụ bao gồm các yếu tô về con người, hệ thống và bối cảnh. Các yếu tố ảnh hưởng được phân loại như sau:

* Yếu tố ảnh hưởng con người: tâm trạng, tính cách, độ tuổi, nền tảng văn hoá xã hội, nhu cầu, mục tiêu …
* Yếu tố ảnh hưởng hệ thống: nội dung, đa phương tiện (mã hoá, độ phân giải,...), mạng truyền thông (băng thông, độ trễ,...), thiết bị trình diễn (độ phân giải màn hình, tốc độ xử lý…).
* Yếu tố ảnh hưởng bối cảnh: bối cảnh xã hội, kinh tế, vị trí địa lý,...

Các nghiên cứu trong lĩnh vực QoE thường tập trung vào các yếu tố hệ thống. QoE được đánh giá và đo lường theo phương pháp khách quan và chủ quan. Phương pháp đánh giá khách quan là việc sử dụng các mô hình toán học nhằm mục đích xấp xỉ kết quả từ phương pháp đánh giá chất lượng chủ quan. Trong đó, mô hình toán học có thể là những mô hình thống kê hay những phương trình toán học dựa vào các tham số đặc trưng có thể đo lường được một cách khách quan như thông lượng mạng, tốc độ xử lý của các thiết bị đầu cuối, độ trễ truyền,… Khác với phương pháp đánh giá khách quan, phương pháp đánh giá chủ quan là dựa vào quan điểm của người dùng thông qua việc xem các video. Quy trình đo lường chất lượng video chủ quan được mô tả trong ITU-R recommendation BT.500 và ITU-T recommendation P.910. Trong quy trình này một tập các video được trình diễn cho một nhóm người xem. Các ý kiến của người xem về video được quy thành thang điểm từ một đến năm và tính trung bình thành điểm MOS (mean opinion scores) để đánh giá chất lượng của video. Tuy nhiên việc đánh giá chất lượng video theo phương pháp chủ quan phụ thuộc khá nhiều vào tâm trạng của người xem tại thời điểm xem và khả năng xử lý video của hệ thống sử dụng để chạy thí nghiệm.

## 1.3 Tổng quan về công nghệ streaming video

Video là một đa phương tiện quan trọng trong lĩnh vực truyền thông và giải trí. Ban đầu video được ghi và truyền đi ở dạng tương tự. Vào đầu những năm 1990, với sự phát triển mạnh mẽ của khoa học máy tính, công nghệ số hoá ra đời cho phép việc chuyển dữ liệu tương tự thành dữ liệu số đã tạo nên một cuộc cách mạng trong lĩnh vực nén và truyền video. Tại thời điểm đó các video được số hoá, mã hoá và lưu trữ tại các máy chủ dưới dạng các tệp đa phương tiện. Khi người dùng muốn xem một video cần phải tải tệp đó về lưu trữ trong máy tính cá nhân. Sau khi tải xong người dùng có thể sử dụng một trình phát đa phương tiện để chơi video. Nhược điểm lớn nhất của phương pháp này là việc người dùng phải đợi cho tới khi video được tải xong mới có thể xem, có khi mất một khoảng thời gian dài để tải video phụ thuộc vào băng thông mạng và kích thước của tệp video. Đồng thời việc tải về và lưu trữ trên máy sẽ làm tốn tài nguyên bộ nhớ.

Công nghệ streaming video ra đời ban đầu để giải quyết vấn đề này. Trong một hệ thống ứng dụng streaming video, máy khách bắt đầu chơi video trong vòng một vài giây sau thời điểm nó bắt đầu nhận được dữ liệu từ máy chủ. Hay nói cách khác, người dùng sẽ không cần phải đợi đến khi tải hết toàn bộ nội dung của video từ máy chủ mà vẫn có thể xem video trong khi vẫn tải nội dung video từ máy chủ. Công nghệ streaming video là công nghệ truyền video thời gian thực, yêu cầu một thời gian giới hạn. Ví dụ như khi truyền video, dữ liệu video phải được chơi liên tục tại phía người dùng cuối. Nếu dữ liệu không đến đúng thời điểm thì việc chơi video sẽ bị gián đoạn, điều này gây khó chịu cho người dùng.

### 1.3.1 Phân loại streaming video trên Internet

Streaming video trên Internet có thể được chia thành ba công nghệ chính như bảng 1.1 đó là unicast streaming, IP multicast, Application-layer multicast.

***Bảng 1. 1 Các phương thức streaming video trích dẫn luận án tiến sĩ anh Hưng***

|  |  |
| --- | --- |
| **Công nghệ streaming video** | **Đặc điểm** |
| Unicast streaming | Thông thường sử dụng các bộ giao thức RTP/UDP hoặc HTTP/TCP để truyền video. Hiện tại HTTP adaptive streaming (HAS) đang được sử dụng phổ biến. |
| IP multicast | Truyền phát cùng một video tới nhiều người dùng. Thiết lập nhiều nhóm người dùng dựa vào băng thông, truyền tải video có bitrate phù hợp với băng thông của từng nhóm người dùng. |
| Application-layer multicast | Sử dụng các giao thức riêng biệt (BitTorrent, …) để truyền video với các phương pháp Peer-to-Peer streaming. Tuy nhiên hình thức này gặp thách thức lớn trong việc bảo mật và chống các cuộc tấn công độc hại tới phía các máy khách vì khi sử dụng phương pháp P2P yêu cầu mở một số cổng kết nối mà hacker có thể lợi dụng các cổng mở này để tấn công người dùng. |

Ngoài cách phân loại theo công nghệ streaming như trên, streaming video còn được phân loại theo các đặc điểm kỹ thuật. Các kỹ thuật streaming video là rất đa dạng thường thì được phân loại theo năm yếu tố, đó là: yêu cầu về độ trễ, chế độ mã hoá video, các giao thức tầng giao vận, tính thích ứng và vị trí điều khiển thích ứng.

**Yêu cầu về độ trễ:** Dựa vào yêu cầu về độ trễ video streaming được chia thành ba loại: streaming tương tác, streaming trực tiếp và streaming theo yêu cầu. Đặc trưng của streaming tương tác và streaming trực tiếp là nội dung video được truyền tới người xem trong khi máy quay phía phát vẫn đang ghi nội dung tiếp theo của video. Hình thức streaming video này yêu cầu độ trễ rất thấp. Một ví dụ của streaming tương tác là hai người hoặc một nhóm sử dụng tính năng video call trên facebook để nói chuyện với nhau từ những vị trí địa lý khác nhau. Streaming trực tiếp là những trương trình trực tiếp như bóng đá U23 được phát trực tiếp trên website vtvgo của Đài truyền hình Việt Nam. Có thể thấy độ trễ trong các hình thức streaming này là rất quan trọng và ảnh hưởng trực tiếp đến người xem. Ngược lại với video theo nhu cầu, video được ghi trước và được lưu tại máy chủ nên nó không yêu cầu bất kỳ nào.

**Tốc độ bit mã hoá video:** Dựa vào tốc độ bit mã hoá video người ta chia ra làm hai loại là streaming video có tốc độ bit biến đổi (VBR) và streaming video có tốc độ bit không đổi (CBR). Việc dự đoán tốc độ bit cho một video dạng VBR là rất khó khăn vì vậy streaming video VBR đang còn nhiều thách thức. Với tốc độ bit cố định, streaming video CBR sẽ dễ dàng hơn so với video VBR.

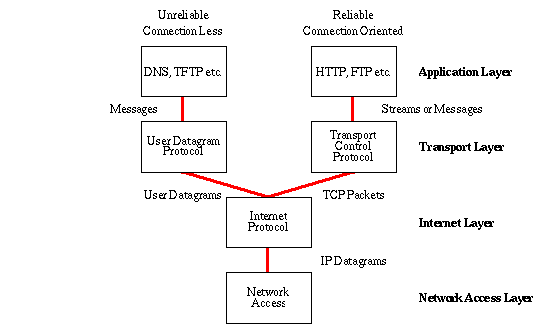
**Giao thức truyền:** Dựa vào các đặc điểm của các giao thức tầng giao vận UDP và TCP, streaming video ban đầu sử dụng giao thức nền là UDP vì có lợi thế về tốc độ. Tuy nhiên sau này TCP lại trở thành sự lựa chọn phổ biến vì cung cấp nhiều lợi thế hơn về chi phí kinh tế cho hệ thống cũng như tính an toàn cao hơn so với sử dụng giao thức UDP.

**Tính thích ứng:** Có hai loại đó là streaming thích ứng và streaming không thích ứng. Streaming thích ứng mang lại nhiều lợi thế khi tốc độ bit được lựa chọn phù hợp với thông lượng mạng của người dùng. Trong khi streaming không thích ứng người dùng có khả năng sẽ bị gián đoạn khi xem vì thông lượng mạng biến đổi.

**Vị trí điều khiển thích ứng:** Trong hệ thống streaming thích ứng, khối điều khiển thích ứng có thể được đặt tại hai vị trí đó là tại máy chủ hoặc máy khách. Trước đây bộ thích ứng được đặt tại máy chủ, tuy nhiên do có quá nhiều máy khách cùng truy cập gây phức tạp và quá tải cho máy chủ. Do đó, hầu hết các dịch vụ streaming (trừ streaming tương tác) đều đặt khối thích ứng tại máy khách.

### 1.3.2 Các giao thức tầng giao vận streaming video sử dụng

Các giao thức tầng giao vận quen thuộc với streaming video bao gồm UDP và TCP. Hai giao thức này cung cấp các chức năng truyền nhận cơ bản trong công nghệ streaming video. UDP và TCP là hai giao thức cốt lõi của bộ giao thức mạng Internet. Bộ giao thức mạng Internet cung cấp truyền thông dữ liệu đầu cuối xác định rõ cách dữ liệu được đóng gói, xử lý, truyền, định tuyến và nhận. Các chức năng này được tổ chức thành bốn tầng trừu tượng, mỗi tầng giải quyết một vấn đề có liên quan đến việc truyền dữ liệu, và cung cấp cho các giao thức ở tầng trên một dịch vụ được định nghĩa rõ ràng dựa trên việc sử dụng dịch vụ từ các dịch vụ của tầng phía dưới.



Hình Mô hình bộ giao thức mạng Internet (chưa reference)

Từ thấp nhất là tầng liên kết dữ liệu (link layer) bao gồm các phương thức truyền dữ liệu có chức năng biến đổi dòng bit logic thành các tín hiệu vật lý phù hợp với đường truyền vật lý (ở bên phát) và điều chế giải điều chế, khôi phục tín hiệu (ở bên thu). Tầng internet (Internet layer) cung cấp liên kết giữa các mạng độc lập, tầng giao vận (transport layer) xử lý và giao tiếp giữa các máy chủ với nhau. Tầng ứng dụng (application layer) cung cấp trao đổi dữ liệu theo quy trình cho các ứng dụng.

Bộ giao thức mạng Internet có hai giao thức nền tảng chính ở tầng giao vận (transport layer), một là giao thức phi kết nối (connectionless protocol) và giao thức thứ hai là giao thức hướng kết nối (connection-oriented). Hai giao thức này bổ sung cho nhau. Giao thức phi kết nối là UDP, UDP không làm gì ngoài việc gửi các gói tin giữa các ứng dụng, cho phép các ứng dụng xây dựng giao thức của riêng chúng khi cần thiết. Giao thức hướng kết nối là TCP, TCP tạo các kết nối và là giao thức truyền tin tin cậy khi có cơ chế truyền lại, cùng với cơ chế điều khiển luồng, kiểm soát tắc nghẽn. Cả hai giao thức TCP và UDP đều hoạt động trên cơ sở của giao thức IP.

**Giao thức UDP – User Datagram Protocol**

Giao thức UDP được thiết kế bởi David P. Reed năm 1980 và chính thức được định nghĩa trong RFC 768. Với UDP các ứng dụng máy tính có thể gửi các tin nhắn, trong trường hợp này được xem như là một datagram tới các máy khác trong giao thức mạng Internet.

Giao thức UDP là giao thức đơn giản, phi liên kết và cung cấp dịch vụ trên tầng giao vận với tốc độ nhanh. UDP hỗ trợ kết nối một - nhiều và thường được sử dụng trong trong liên kết một nhiều bằng cách sử dụng các datagram multicast (cách thức truyền tin được gửi từ một điểm đến một tập hợp các điểm khác) và unicast (cách thức truyền tin được gửi từ một điểm tới một điểm khác).

Khi sử dụng UDP gói tin chỉ được gửi đến bên nhận, bên gửi sẽ không chờ đợi để xác nhận rằng bên nhận đã nhận được gói tin mà sẽ tiếp tục gửi các gói tin tiếp theo. Nếu bên nhận hủy bỏ một số gói tin, bên nhận sẽ không thể yêu cầu bên gửi gửi lại lần nữa. Không có gì để đảm bảo rằng bên nhận đã được tất cả các gói tin hay chưa nhưng bù vào đó các máy tính có thể giao tiếp với nhau một cách nhanh chóng hơn. UDP thường được sử dụng khi có nhu cầu về tốc độ và độ chính xác của thông tin không yêu cầu quá cao. Ví dụ UDP thường được sử dụng cho các chương trình phát sóng trực tiếp và trò chơi trực tuyến.

UDP không cần thiết lập liên kết. Không sử dụng các tín hiệu bắt tay nên có thể tránh được thời gian trễ. Do vậy DNS thường sử dụng giao thức UDP hơn là TCP.

UDP hỗ trợ các liên kết một – một, một – nhiều, trong khi TCP chỉ hỗ trợ liên kết một – một. Kích thước header 8 byte nhỏ hơn so với TCP là 20 byte giúp tiếc kiệm băng thông**.** Tuy nhiên UDP cũng có nhiều nhược điểm. UDP không đảm bảo dữ liệu được truyền đến đích hay chưa vì không có cơ chế thiết lập kênh truyền nên bên gửi không biết được bên nhận đã nhận được gói tin hay chưa chưa nhận được. UDP không hỗ trợ các chỉ số phiên (session ID) để duy trì các liên kết kết giữa 2 máy do bản chất phi liên kết của nó. UDP cũng không đảm bảo thứ tự các gói tin đến đích như đã được tạo ra ở nguồn. Một nhược điểm nữa của UDP là tính bảo mật không cao. Trong nhiều tổ chức, tường lửa và router cấm các gói tin UDP vì các hacker thường dò tìm các cổng UDP để biết được máy chủ đang chạy các dịch vụ mạng nào và thực hiện khai thác thông tin trái phép. UDP không có cơ chế kiểm soát luồng. Một ứng dụng UDP tồi có thể làm giảm băng thông của mạng.

**Giao thức TCP – Transport Control Protocol**

UDP là một giao thức đơn giản và có một số công dụng rất quan trọng như tương tác giữa các máy khách và máy chủ và truyền thông đa phương tiện. Tuy nhiên hầu hết các ứng dụng sử dụng Internet thì độ tin cậy, tuần tự là cần thiết, UDP không thể đáp ứng đặc điểm đó vì vậy một giao thức khác được phát triển là TCP để bổ sung cho những tính năng thiếu sót của UDP.

TCP được định nghĩa chính thức trong tài liệu RFC 793 là một trong những bộ giao thức cốt lõi của bộ giao thức mạng Internet. TCP là giao thức hướng kết nối (connection-oriented) tức là khi muốn truyền nhận dữ liệu thì phải thiết lập kết nối trước. Sử dụng TCP các ứng dụng trên máy chủ được kết nối mạng có thể kết nối với nhau mà qua đó chúng có thể trao đổi dữ liệu là các gói tin và đảm bảo tính tin cậy, khả năng kiểm soát lỗi của dòng dữ liệu giữa các ứng dụng chạy trên các máy chủ đang kết nối với nhau thông qua giao thức IP. Các ứng dụng Internet đòi hỏi độ tin cậy cao của dữ liệu như là World Wide Web, email, quản trị từ xa, truyền dữ liệu đều dựa vào TCP.

TCP cung cấp dịch vụ truyền thông ở tầng trung gian giữa một trương trình ứng dụng và giao thức Internet chính là tầng giao vận của mô hình mạng Internet. Một ứng dụng không cần biết các cơ chế cụ thể để gửi dữ liệu qua một liên kết tới máy chủ khác ví dụ như xử lý phân mảnh gói tin để phù hợp với kích thước khung truyền dẫn. Tại tầng giao vận TCP xử lý tất cả các chi tiết bắt tay và truyền nhận giữa các máy chủ và tạo ra một kết nối mạng trừu tượng tới các ứng dụng thông qua giao diện socket. Socket là một đơn vị điểm đầu cuối được tạo ra bởi TCP để gửi và nhận dữ liệu. Mỗi socket có một địa chỉ IP của máy chủ và một số 16 bit dùng để định nghĩa port tương ứng với các dịch vụ mà máy chủ đang chạy. Như vậy một kết nối TCP phải được thiết lập rõ ràng giữa một socket trên một máy chủ và một socket trên một máy chủ khác.

TCP cung cấp một khả năng truyền dữ liệu an toàn giữa các máy chủ trong mạng với cơ chế thiết lập và huỷ bỏ kết nối. Cung cấp các chức năng kiểm tra chính xác dữ liệu khi đến đích và truyền lại dữ liệu khi có lỗi xảy ra. Sử dụng cơ chế báo nhận, đánh số thứ tự các gói dữ liệu và sắp xếp lại thứ tự dữ liệu tại bên nhận. Cung cấp khả năng đa kết nối cho các tiến trình khác nhau giữa hai máy chủ qua số hiệu cổng. Tất cả các kết nối TCP đều là song công (full-duplex). Song công nghĩa là dữ liệu có thể được truyền và nhận cùng một thời điểm. Đồng thời TCP chỉ hỗ trợ kết nối một – một (point – to – point), mỗi kết nối chỉ có đúng hai điểm đầu cuối. TCP không hỗ trợ kết nối multicast hay broadcast.

Với các đặc điểm của cả hai giao thức UDP và TCP công nghệ streaming video đã hình thành và phát triển dựa trên các đặc điểm của hai bộ giao thức này.

### 1.3.3 Sự phát triển của công nghệ streaming video

Từ khi được hình thành và phát triển công nghệ streaming được xây dựng sử dụng nhiều công nghệ khác nhau trải qua nhiều giai đoạn. Mỗi giai đoạn streaming video được cải thiện nhằm khắc phục những nhược điểm của các giai đoạn trước đó để đạt được mục đính đem lại chất lượng cao nhất cho người dùng cuối trải nghiệm. Streaming video có thể được chia thành ba giai đoạn phát triển: datagram streaming, streaming tải luỹ tiến, streaming thích ứng trên nền giao thức HTTP.

**Datagram streaming**

Trong đầu những năm 1990 khi mới hình thành công nghệ streaming video chủ yếu sử dụng bộ giao thức Real-time Transport Protocol (RTP) / Real-Time Control Protocol (RTCP). RTP là giao thức mạng được phát triển cho việc truyền nhận dữ liệu thời gian thực. RTP thường được sử dụng kết hợp với RTCP. RTP được sử dụng để truyền tải dữ liệu đa phương tiện trong khi RTCP được sử dụng để theo dõi số liệu thống kê truyền và chất lượng dịch vụ như nhận biết về trạng thái kết nối cũng như trạng thái trình phát video tại phía nhận. Giao thức RTP/RTCP là các giao thức tầng giao vận được xây dựng trên nền giao thức UDP. Sau này, một bộ giao thức mới quan trọng trong streaming video trên Internet được tổ chức Internet Engineering Task Force (IETF) xây dựng là bộ giao thức RTP/RTCP/RTSP (Real-Time Streaming Protocol). Ngoài RTP và RTCP, RTSP trong bộ giao thức được thiết kế để tạo và duy trì phiên video và cung cấp các chức năng điều khiển, cho phép người dùng có thể tạm dừng, hoặc phát lại hay tìm kiếm trong video streaming giống như chơi video được lưu trữ tại máy của người dùng [Two Decades of Internet Video Streaming:

A Retrospective View]

Tuy nhiên do các bộ giao thức RTP xây dựng trên nền UDP nên gặp một số vấn đề như sau: Do giao thức UDP không có cơ chế phát lại, chống tắc nghẽn nên để xây dựng các server đối phó với các vấn đề này là một công việc rất khó khăn. Hơn nữa hầu hết các kết nối của người dùng ra ngoài Internet thông qua tường lửa được cấu hình để chặn UDP [adaptive-streaming-a brief tutorial].

**Streaming video tải luỹ tiến**

Do những nhược điểm của streaming trên các bộ giao thức RTP, đồng thời với sự phát triển mạnh mẽ của mạng phân phối nội dung (Conten Distribution Networks – CDN), kỹ thuật streaming tải luỹ tiến ra đời. CDN là công nghệ phát triển các server lưu trữ đa dạng tại nhiều vùng địa lý khác nhau. Mạng phân phối nội dung CDN giúp nhà cung cấp dịch vụ nội dung nâng cao chất lượng và tăng khả năng đáp ứng truy cập dịch vụ nội dung số của người dùng. CDN cho phép người dùng streaming video gần họ, những yêu cầu của người dùng đều được điều hướng tới máy chủ gần họ nhất để mang lại lợi ích tốt nhất. Vì CDN chủ yếu được thiết kế tối ưu để phục vụ các nội dung web nên streaming video tải luỹ tiến được phát triển sử dụng giao thức HTTP để tải các phân đoạn video từ máy chủ web [Two Decades of Internet Video Streaming:A Retrospective View].

Phương pháp này mang lại một số lợi ích như: Chỉ yêu cầu các máy chủ web, không cần các máy chủ chuyên biệt và khó triển khai như bộ giao thức RTP. Tận dụng được các lợi thế mà công nghệ CDN mang lại. Đồng thời giao thức HTTP sử dụng TCP hoàn toàn thân thiện với tường lửa và các bộ định tuyến. Tuy nhiên HTTP streaming tải luỹ tiến lại có nhược điểm là phía người dùng phải chọn mức chất lượng video từ đầu khi mới bắt đầu phiên stream và không được chọn lại. Khi đó sẽ xảy ra các trường hợp như thông lượng mạng của người dùng tại thời điểm họ bắt đầu phiên stream đang thấp nếu chọn video chất lượng cao thì khả năng gián đoạn là lớn trong phiên truyền, nếu chọn video chất lượng thấp người dùng phải xem ở chất lượng thấp cả video trong khi thông lượng mạng có thể tăng lên trong lúc người dùng đang xem video. Hay ở một ngữ cảnh ngược lại khi thông lượng mạng của người dùng đang cao tại thời điểm bắt đầu phiên truyền, nếu người dùng chọn video có mức chất lượng cao để bắt đầu thì trong thời gian phiên truyền rất có thể thông lượng mạng sẽ biến đổi mạnh xuống thấp gây ra gián đoạn lớn trong phiên truyền. Sử dụng bộ đệm tại máy khách có thể giải quyết được một phần gián đoạn, tuy nhiên để đối phó với sự biến đổi của thông lượng mạng như vậy là chưa đủ.

Có thể thấy rằng, kỹ thuật streaming video tải luỹ tiến sử dụng HTTP không hỗ trợ thích ứng tốc độ bit theo sự biến đổi của thông lượng, dẫn đến chất lượng video không được cải thiện. Nhược điểm này dẫn đến sự phát triển của công nghệ streaming video thích ứng sử dụng giao thức HTTP (HTTP Adaptive Streaming – HAS).

**Công nghệ streaming video thích ứng sử dụng HTTP**

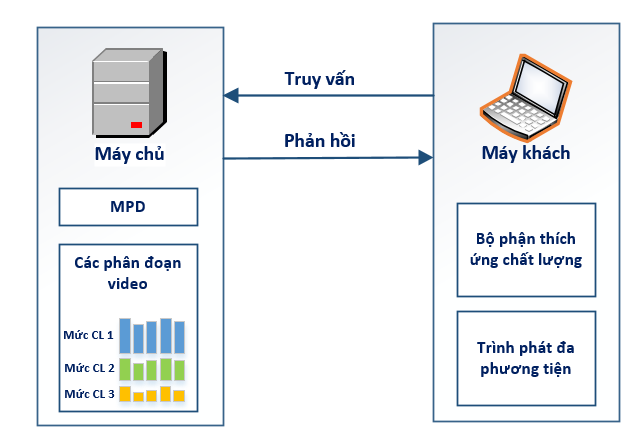
## 1.4 Công nghệ streaming video trên nền giao thức HTTP

### 1.4.1 Giao thức HTTP

HTTP là giao thức được dùng để liên hệ thông tin giữa máy cung cấp dịch vụ (web server) và máy sử dụng dịch vụ (web client) trong mô hình máy khách/máy chủ dùng cho World Wide Web. HTTP là một giao thức thuộc tầng ứng dụng nằm trên tầng giao vận trong mô hình mạng TCP/IP. Chuẩn HTTP đang được sử dụng phổ biến hiện nay là HTTP/2 được ra đời và công bố trong tài liệu RFC 7540. HTTP hoạt động trên nền giao thức TCP, máy khách gửi truy vấn tới máy chủ và đợi dữ liệu trả về từ máy chủ. Quá trình này là liên tục trong suốt phiên kết nối. HTTP mang lại một số lợi ích rõ ràng cho các dịch vụ streaming như chỉ cần sử dụng máy chủ Web thay vì một máy chủ riêng chuyên biệt để điều khiển truyền.

### 1.4.2 Nguyên lý hoạt động của hệ thống streaming thích ứng trên nền giao thức HTTP

Kiến trúc phổ biến của một hệ thống HAS bao gồm máy chủ, mạng truyền tải và máy khách được trình bày trên hình 1.2. Các phân đoạn video và tập mô tả trình diễn đa phương tiện (MPD) được lưu tại máy chủ và sẽ được truy vấn bởi máy khách. Máy khách khởi tạo một phiên streaming bằng việc tải một tệp MPD bao gồm một mô tả các phiên bản chất lượng khác nhau của video và các phân đoạn. Sau đó một bộ phận thích ứng được triển khai tại máy khách chịu trách nhiệm cho việc quyết định phân đoạn nào được yêu cầu, dựa vào tệp MPD và trạng thái hiện tại của mạng. Video sau đó sẽ được truyền tới máy khách thông qua một chuỗi truy vấn – phản hồi HTTP.



***Hình 1. 2 Sơ đồ hệ thống HAS***

**Các thành phần quan trọng trong hệ thống HAS**

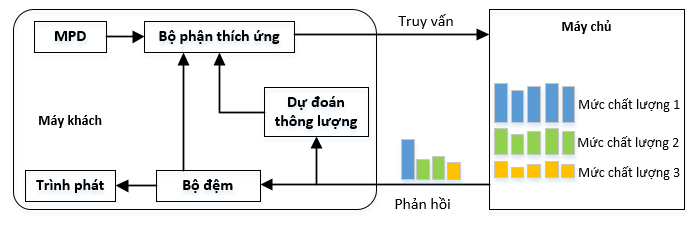
**Tệp MPD** là một tệp đuôi .xml chứa thông tin siêu dữ liệu bao gồm các thông tin như sau:

* Thông tin truyền tải chung như tên tệp, thông tin mật mã, truyền tải theo nhu cầu hay trực tiếp.
* Loại luồng như âm thanh, hình ảnh, phụ đề.
* Các mức chất lượng như thông tin về dạng mã hoá, tốc độ bit trung bình
* Chỉ số của các phân đoạn, đường dẫn tới các phân đoạn.
* Tốc độ đỉnh đối với video VBR

**Truyền video** trong hệ thống HAS, đồ án chỉ xét đến truyền các phân đoạn video tại tầng ứng dụng. Sau khi nhận được giữ liệu là tập MPD máy khách bắt đầu đưa ra một chuỗi các truy vấn HTTP để tải các phân đoạn video. Máy khách sẽ tuỳ vào thông lượng mạng mà chọn phân đoạn video có chất lượng phù hợp.

**Máy khách trong hệ thống HAS**

Hình 1.3 chỉ ra sơ đồ khối của máy khách trong hệ thống HAS. Tuỳ thuộc vào truy vấn của máy khách và phiên bản của giao thức HTTP, máy chủ sẽ phản hồi một hay nhiều phân đoạn với cùng một tốc độ bit. Tất các các phân đoạn được tải từ máy chủ sẽ được lưu ở bộ đệm máy khách trước khi chuyển tới trình chơi video.



***Hình 1. 3 Máy khách trong hệ thống HAS***

Bộ phận quan đóng vai trò quan trọng là bộ thích ứng, khi một phương pháp được phát triển để quyết định xem phân đoạn có chất lượng như thế nào sẽ được truy vấn. Một quyết định tốt cho một truy vấn cần dựa vào mức bộ đệm hiện tại, thông lượng được ước tính hiện tại và dữ liệu từ MPD. Tệp MPD là được nhận từ lúc bắt đầu phiên streaming, máy khách phải liên tục ức tính thông lượng và đo được mức sử dụng bộ đệm trong toàn bộ phiên streaming.

Tuỳ vào từng trường hợp cụ thể sẽ có những cách dự đoán thông lượng khác nhau. Một số cách dự đoán thông lượng điển hình như đo thông lượng của một phân đoạn hiện tại sau đó sử dụng một mô hình tuyến tính để dự đoán thông lượng cho phân đoạn kế tiếp. Ngoài ra có thể đo thông lượng của nhiều phân đoạn gần với phân đoạn hiện tại nhất và tính trung bình để dự đoán cho thông lượng kế tiếp.

Bộ đệm trong máy khách đóng vai trò quan trọng trong việc đối phó với biến động của thông lượng mạng. Bộ đệm giúp cho việc streaming video không bị gián đoạn. Bộ đệm với kích thước lớn, khả năng xảy ra gián đoạn sẽ nhỏ hơn bộ đệm có kích thước nhỏ. Kích thước và mức sử dụng bộ đệm được tính theo đơn vị thời gian. Khi bắt đầu một phiên streaming máy khách sẽ nạp vào bộ đệm một lượng dữ liệu nhất định trước khi giải mã và hiển thị dữ liệu ra màn hình. Đây là giai đoạn nạp bộ đệm ban đầu, sau đó máy khách chuyển qua giai đoạn streaming nhận dữ liệu và hiển thị đồng thời. Trong giai đoạn này nếu mức sử dụng bộ đệm đủ lớn người xem sẽ không bị gián đoạn bởi thông lượng mạng biến đổi gây ra.

## 1.5 Kết luận chương

Trong chương I, đồ án đã trình bày các cơ sở lý thuyết nền tảng liên quan và công nghệ streaming video nói chung, đặc biệt là công nghệ HAS nói riêng. Sau đó là các yếu tố tác động đến QoE trong HAS. Đây sẽ là kiến thức nền tảng để em có thể thực hiện các chương tiếp theo.

# Chương II. TRIỂN KHAI GIẢI PHÁP THÍCH CHẤT LƯỢNG CHO VIDEO VBR