Converged Video Delivery over Heterogeneous Networks

by

Amit Limaye

Submitted to the Systems Design and Management Program in Partial Fulfillment of the Requirements for the Degree of Master of Science in Engineering and Management

Master of Science in Engineering and Management

at the

Massachusetts Institute of Technology

June 2012
[SEPTEMBER 2012]

Signature of Author ______________________________
Amit Limaye, System Design and Management Program
June 2012

Certified by ______________________________
Michael A M Davies
Sr. Lecturer, Engineering Systems Division
Thesis Supervisor

Accepted by ______________________________
Patrick Hale
Director, System Design and Management Program

This page is intentionally left blank

Converged Video Delivery over Heterogeneous Networks

By

Amit Limaye

Submitted to the Systems Design and Management Program in Partial Fulfillment of the Requirements for the Degree of Master of Science in Engineering and Management

Abstract

Mobile traffic has grown substantially over the last few years; a trend which is expected to continue. The chief reasons behind this phenomenon are the availability of better handsets, faster cellular networks and the variety of content available on the internet suitable for consumption on mobile devices. The nature of the traffic is also changing from pure web browsing with latency-tolerant traffic to video, which is becoming the major class of content consumed on mobile devices. This trend, combined with the trend of decreasing prices per GB of data, which constrains the amount of money an operator can spend upgrading its network and they see increasing value in alternative solutions to address this data deluge while managing costs and maintaining customer service.

A variety of solutions have been tried by operators based on enhanced charging, traffic engineering and backhaul infrastructure upgrades. Wi-Fi offload is one such promising solution as it addresses the congestion problem where it is most severe because of data consumption by users using streaming video. Cellular spectrum is a scarce and expensive resource for operators, and by allowing them to offload traffic to Wi-Fi networks in unlicensed spectrum they can free cellular spectrum for more valuable applications. Wi-Fi offload has, however, suffered from the incapability to manage seamless handovers and the required interaction of the user to select a Wi-Fi network. This made the process of attaching to a Wi-Fi network very complicated. These limitations have been addressed in recent standards and make the case for Wi-Fi offload more viable and attractive than earlier.

At the same time new video optimization techniques such as H.264/SVC which allow the use of multiple streams and channel will allow content providers or distributors to use multiple networks and to scale video seamlessly according to handset capabilities and network conditions.

The thesis proposes a solution, based on a set of new Wi-Fi standards and the new H.264/AVC codecs, which leverages a combination of low cost Wi-Fi and high reliability cellular networks to reduce the cost of video transmission while maintaining a comparable QOE for nomadic users. The thesis also enumerates some of the basic procedures that can be supported using the proposed architecture. This new architecture opens new opportunities for existing players in the mobile content ecosystem and adds new players to the ecosystem. The thesis identifies the needs and opportunities for each of the new player and also develops a cost model for streaming video using this solution.

Thesis Supervisor: Michael A M Davies
Title: Senior Lecturer, Engineering Systems Division

This page is intentionally left blank

Acknowledgements

My time at MIT has been an enriching experience and I wish to take this opportunity to extend my gratitude to everyone who made it possible, particularly to those who helped me with this thesis.

To Senior Lecturer Michael A M Davies, my Thesis advisor, thanks for guiding me through the thesis with insightful ideas and comments; most importantly helping me define the problem and helping me refine my problem statement. I honestly appreciate all the time and energy you put into this project. It has been great working with you.

To Pat Hale and the friendly SDM staff who have made the SDM such a great program. The friendly SDM staff has always made me feel at home.

To my classmates and friends at MIT and Boston for making this an unforgettable experience. I have been fortunate enough to meet wonderful people from diverse backgrounds. I have often used them to test my ideas. I have developed great friendships and I am confident that I will carry these through the rest of my life.

To my parents for their unconditional support and love; without their encouragement and support this would not have been possible. Their support has been crucial in every new challenge I have undertaken

Finally to my wife Sneha for her incredible support through the course when I have spent all my weekends busy in school. Her love and patient support while I worked my way through this thesis have been a source of great strength and have helped me enjoy my time through MIT. You made this possible.

This page is intentionally left blank

Contents

List of Tables

List of Figures

1. Introduction

1.1 Context

The thesis sits at the intersection of three technologies that have undergone rapid changes over the last decade: Mobile handsets, Wi-Fi/cellular networks and video codecs. Each of these technologies has changed drastically and has given rise to a lot of applications which were perhaps not feasible less than a decade ago. The thesis attempts to capture one such interaction and propose a solution to a problem which I think is represents one of the most important and interesting problem in this field.

Cellular networks have evolved from supporting just 114 kbps data during EDGE/GSM days to supporting more than 100 Mbps connections today using 4G/LTE networks. These networks along the way have moved from circuit switched to packet switched, mirroring a dominant trend today where data rather than voice is the most common application used on most cellphones. The evolution to packet switched networks has reduced the cost of sending/receiving data for both operators and users which have increased the popularity of these applications even more.

Cellular phones have also evolved majorly during this time. We have moved from simple feature phones which could support only one application, voice calling, to smartphones which are almost mini-computers in their own right. The capabilities of these smartphones mean they are being increasingly used as computers and the data volume they generate is mirroring computers in the connected world. This ideally should be great news for operators but the major problem they face is that high margin services like voice SMS are getting substituted by lower margin applications using VOIP or free instant messenger services. These services generate the same or even more data traffic, but don't generate the same amount of revenue for the operator.

While the cellular networks have increased both in terms of capacity and coverage the same can be said of Wi-Fi networks. It is predicted that by 2015 there will 2.1M public hotspots all over the world[1]. This coupled with at least 50%[2] of all users in the US owning smartphones with Wi-Fi network capability indicates a lot of this increased data can be offloaded onto the Wi-Fi. The IEEE standards body has been aware of these implications and has come up with standards to overcome the deficiencies of current Wi-Fi standards. The new standards like 802.11u/802.21 will help make logging into a Wi-Fi and mobile handoffs between different Wi-Fi networks. This is a promising development which will help address most that users have with Wi-Fi networks.

Video coding has also advanced rapidly over the last 5-7 years in comparison to the old MPEG-2 systems which were characterized by fixed spatial and temporal characteristics where the fixed natures of these codecs gave rise to a minimum network and compute resources required to support each encoded stream. Scalable video codecs which are now an extension of the standard H.264/AVC codec allow video streams to be dynamically scaled depending on device capabilities and thus provide an attractive option to modern video systems using Internet or other best effort IP networks with time varying capabilities to transmit video. The standardization of the SVC extension has enabled the encoding of a high quality video stream into multiple subset bit streams that can be themselves be decoded at the receiver with no additional complexity or loss of quality as compared to the original H.264/AVC codec. Critically the independent streams can also be independently sent over multiple paths and decoded at the receiver.

---

[1] Source: Informa Telecoms & Media and Wireless Broadband Alliance (WBA), "WBA Industry Report 2011: Global Developments in Public Wi-Fi," Nov 10, 2011

[2] Source: eMarketer, April 2012

These technologies combined together make video delivery over a combined Wi-Fi and cellular channel to a mobile handset a very interesting proposition. This thesis will integrate all these technologies and provide a cost effective architecture to deliver video over these networks.

1.2 Motivation

Improvements in cellular networks have solved the bandwidth problem for most consumers. Consumers are no longer constrained by bandwidth issues when choosing which applications to run on their mobile handsets. The capacity problem, which is primarily driven by scarcity of spectrum in public cellular networks, still persists however. The rapid increase in mobile data and more data hungry devices like laptops, tablets and portable gaming devices have brought the capacity problem to the fore. The problem of huge amounts of data has been faced before in the wired world but the problem in wireless/cellular networks is slightly different. The bottleneck in terms of bandwidth and capacity for wired networks was in the backbone networks and it was effectively solved using CDN (content delivery networks). The last mile did not have a meaningful capacity limitation and thus was never thought of as a bottleneck. Cellular networks of today as a virtue of arriving late get advantages of having little problems in the core backbone networks but capacity on the radio network/last mile is still limited and this problem is unique to cellular networks. Wi-Fi have increased in availability and have free unlicensed spectrum. They can be much more densely deployed as well due to low cost of access points.

Video as the killer app has been touted since the arrival of 3G networks. It did not take off when 3G was launched because the handsets were not capable enough to deliver high quality video and the network was not fast enough to support high frame rates or resolution. This led to a distinctly inferior experience as compared to laptops and TVs and video on wireless networks never took off. With the advent of LTE/4G networks we finally had enough bandwidth to support high quality video on a wireless network and the handsets had also improved drastically in the meantime. It seemed like the promise of video over wireless networks would be finally fulfilled but the capacity problem still remained and this was evident in the core network nodes that were being designed to support more users rather high bandwidth as they never expected the radio network to saturate the core. Wi-Fi which was available on most handsets provided a very attractive option to reduce the pressure on the radio network and save licensed spectrum for other applications.

Wi-Fi can provide very large amounts of capacity but has issues related to limited coverage area and seamless handovers between different access points. Cellular has successfully addressed coverage and handover problems but is hobbled by limited capacity and high cost. Any solution which could leverage the two network technologies effectively in combination even for just video would help the operator reduce the pressure on the cellular radio network.

A typical Wi-Fi offload solution lowers the quality of experience for a user who is not at home or work and attached to a single Wi-Fi network. This is where I started to look at ways and means of successfully offloading video on a Wi-Fi network without compromising user experience. The research along the way has changed the initial problem statement and new technologies like 802.11u/802.21 and H.264/SVC have made the job of doing this kind of thing a lot easier.

The architecture proposed in thesis combines the unlicensed Wi-Fi network and the licensed cellular network into a single virtual channel to deliver H.264/SVC encoded video to the handset. This requires integration of Wi-Fi and cellular technologies on the handset and a video streaming architecture which is network aware and can generate adaptable video streams. Support for mobility implemented using handover mechanism is also as essential component of such a system as Wi-Fi system have limited range and without a handoff feature the quality of experience will be bad due to frequent starts and stops in the video.

1.3 Thesis structure

I will break down the problem statement into different questions which I will attempt to answer in each of the section highlighted in the diagram below.

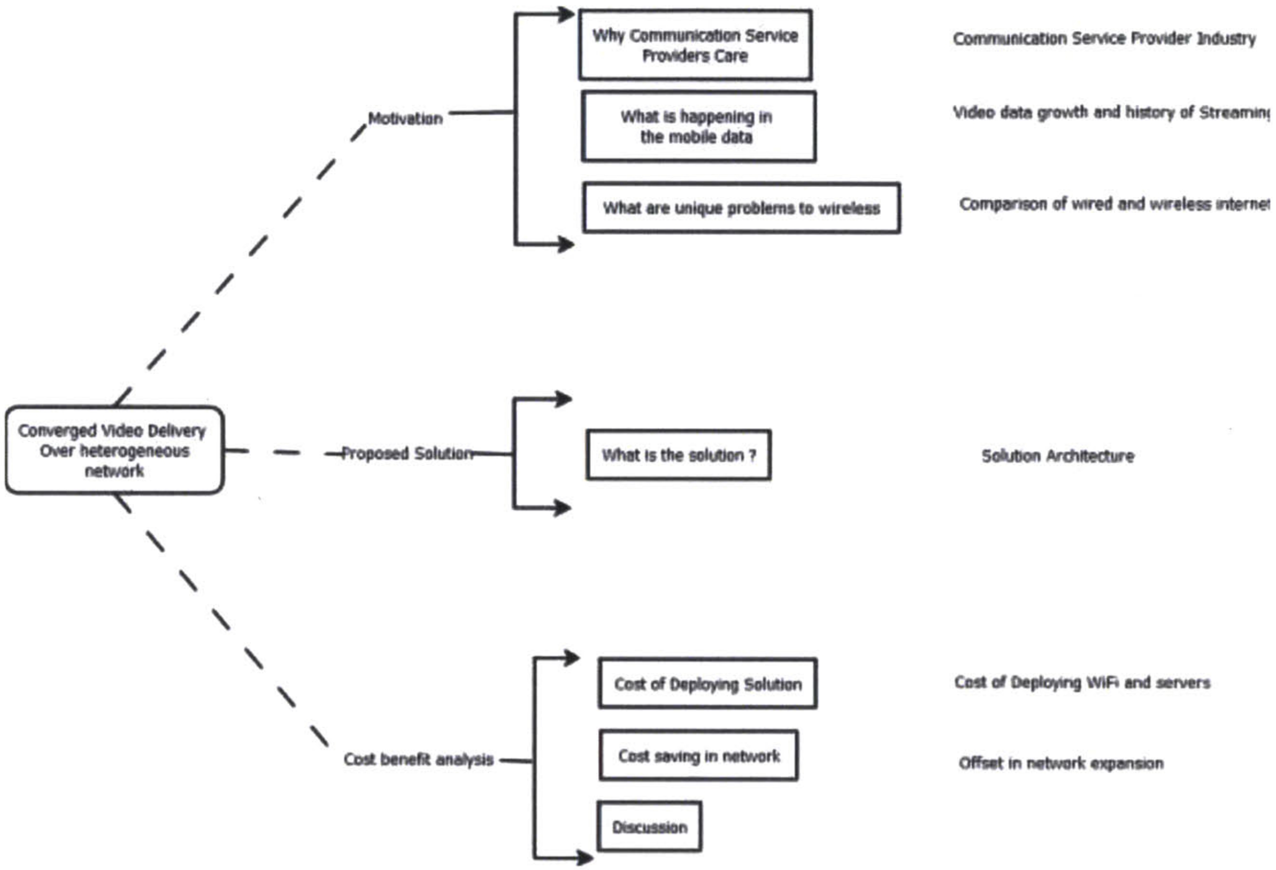


Figure 1: Thesis Structure

I will start with an overview of the mobile service operators and examine the trends of growing mobile data. Then we will try and identify why this growth a problem for mobile service providers. We will end this section by looking at the challenges in data transmission over wireless networks as compared to wired networks which has seen similar growth in data before.

The next section actually discusses the solution architecture and explains the various technology standards and components that are required to build the candidate architecture to successfully handle the video data. We also examine some typical sample data/message flows for information and data exchange between different elements of the sample architecture.

Finally we examine the cost of deploying such a solution as compared to just adding more capacity in the cellular network. We do an analysis and find situations in which this solution might make sense.

2 Mobile Service Providers

This chapter provides an overview of the current state of the mobile service providers with respect to mobile data and how is the traditional business model of the service providers being challenged by the tremendous increase in data as

2.1 Mobile data 1990s - 2010

In the early 1990s when mobile service operators started operating the first data services using EDGE/GSM networks the speed offered was 114 kbps or less, and the primary application for these initial services was email and text browsing. The speed supported by the networks was not fast enough for video and the cost of data was pretty high to make any high bandwidth application viable.

When 3G was initially launched mobile video was prominently advertised as a new service users would get with 3G networks. It was labeled as the "killer app" for the 3G network and the new network was supposed to finally make video calling and watching videos on the phone a reality. In anticipation of this uptake the standard body actually standardized the video in a range of ancillary services and video support was made mandatory for all devices confirming to the 3G specification. The standard also supported video streaming using the H.263 specification. Despite the hype and the proactive behavior of the standards body and the device manufacturers 3G video services never took off which could be possibly explained by the slow uptake of 3G itself but the services had some serious limitations like

- The bandwidth available in 3G networks was only sufficient for low frame rates and resolution which meant the video quality offered was noticeably less than what users could get on their TV and laptops; in addition the handsets were not capable enough to support complicated codecs to get over this limitations
- 3G video services were expensive to use and often fell back on using circuit switched video calls which added to the costs

- Given the limited uptake of 3G video services by users the content available for 3G devices was pretty limited which made the service even more unattractive.

This pretty much meant that 3G video services never took off the way they were supposed to and voice remained the primary source of traffic on 3G networks.

To overcome some of the limitations of this approach some operators and the standardization body tried to use various approaches like Mobile broadcast TV using DVB-H (Digital Video Broadcast Handheld) which successfully incorporated a number of features to support constrained screen sizes of handsets and limited battery life. This effort also failed due to either high cost which operators incurred since DVB-H worked on different frequencies which meant operators had to invest in new equipment. Moreover, in some case spectrum was not available to launch these services and the most popular handset models did not support this standard.

2.2 Mobile Video: History and Trends

The mobile telecommunication business is undergoing major changes driven by innovative technologies, globalization and deregulation. Changes in the consumption pattern of users has led this business to turn from being a mainly voice to mainly data business. Saturation in the developed market has led to almost no growth in the voice business while the data revenue though still not a majority has kept growing rapidly. This trend by most forecasts is going to keep on accelerating and the availability of powerful handsets, network bandwidth and large amount of content is going to shift towards a large percentage of data being video.

In the following sections we will first look at the history of mobile data and early attempts at mobile video and finally culminate the section with what are the present trends for mobile video.

2.2.1 Early Attempts

In the early 1990s with EDGE/GSM supported only 114 kbps data or slower connections which made video transmission over the early data networks all but impossible. These early mobile data connections were primarily used for email and text browsing. Thus capacity and cost pretty much eliminated the the option of video data transmission on these networks.

With the advent of 3G finally we had a network which was believed to be capable enough to support video content. All this meant operators genuinely believed that video would finally take off on the 3G networks and they could possibly have an additional revenue stream through data connections. Video was expected to be the "killer app" and a a lot of standards were specified around video calling using ITU-T H.324 standard, video streaming using the H.263 standard and also a new form of video sharing called MMS which was a service analogous to SMS service already existent on the 2G networks.

Although most people expected video to be the killer app video really never took off with only 6 million subsribers of mobile TV being reporting in 2007-2008[3]. There are many reasons for the apparent failure of video like

- The bandwidth available in 3G networks was only sufficient for low frame rates and resolution which meant the video quality offered was noticeably less than what users could get on their TV and laptops; in addition the handsets were not capable enough to support complicated codecs to get over this limitations
- 3G video services were expensive to use and often fell back on using circuit switched video calls which added to the costs

---

[3] Source: In-Stat, "3G Mobile TV Worldwide" as cited in press release, June 9, 2008

- Given the limited uptake of 3G video services by users the content available for 3G devices was pretty limited which made the service even more unattractive.

This pretty much meant that 3G video services never took off the way they were supposed to and voice remained the primary source of traffic on 3G networks.

The limitation of bandwidth could be addressed by using a parallel spectrum other than the mobile cellular network itself using technique like MediaFLO and other DVB-H standards. This was attempted in various countries when 3GPP standardized the mobile broadcast TV using DVB-H. The standard esentially allowed the transmission of broadcast video over another spectrum which was used exclusively for broadcast of linear TV. This addressed the limitation of network bandwidth but it was still constrained by the limited processing power and battery life of handheld devices. To compound this problem most early smart devices did not support the DVB-H standard. This was another failed attempt at trying to get video on the handsets.

The failure of these attempts and the proprietary nature of most hanset platform coupled with low power meant that voice remained the primary driver of revenue on the 3G networks and data service never really took off despite of multiple efforts by mobile service providers.

2.2.2 Present Scenario

In the early 2000s when 3G service providers were trying to push video content over wireless networks, Consumption of video over the internet or video consumption over any device except the television was not prevalent. This was partly driven by the limited capabilites of the end devices themselves and the unavailability of cheap broadband connections with speeds in Mbits range. This has changed dramatically over the last 5 years with broadband penetration near 70%[4] and around 50%[5] of the US population watching video online. This has led to people and broadcasters being more comfortable with

---

[4] Source: emarketeer 2012

[5] Source: emarketeer 2012

the idea of streaming and consuming video through the internet. Netflix and other top video website already consume close to 50% of all internet traffic today as shown in the table below.

Table 1: Percentage of bandwidth consumed (Source: Sandvine network demographics)

Service	Percentage
Netflix	27.6%
HTTP	17.8%
Youtube	10.0%

This popularity of video content over the wired internet has solved one important problem which was of availability of content on the internet. When the new 4G networks were launched people were already using the internet as a medium to consume video and content providers were finally comfortable enough to deliver video over the internet without fear of piracy. The premium content providers who had stayed away from the 3G online video movement were already actively participating in the online video market with premium content websites like Hulu or Netflix.

Handsets have also improved in the last 5 years in terms of better screen size and resolution, larger processing power, longer battery life and the capability to attach to multiple networks. This has increased the quality of experience for the end user which fills one more gap from the 3G era.

These different technologies coming together have made mobile video much more acceptable to content owners and users who are now familiar with video streaming over IP networks. Mobile video after having two false starts through the early 2000s seems to be finally taking off.

2.2.3 Mobile Data/Video Future

The Cisco visual networking index estimated that mobile data traffic double in the year 2012 and will double again this year. In fact as illustrated in the table below Cisco predicts that traffic will double every year till the year 2016

Global Mobile Traffic Data Growth	
2009	140%
2010	159%
2011	133%
2012	110%
2013	90%
2014	78%
2015	64%
2016	56%

Table 2: Project Mobile Data Growth (Source Cisco VNI)

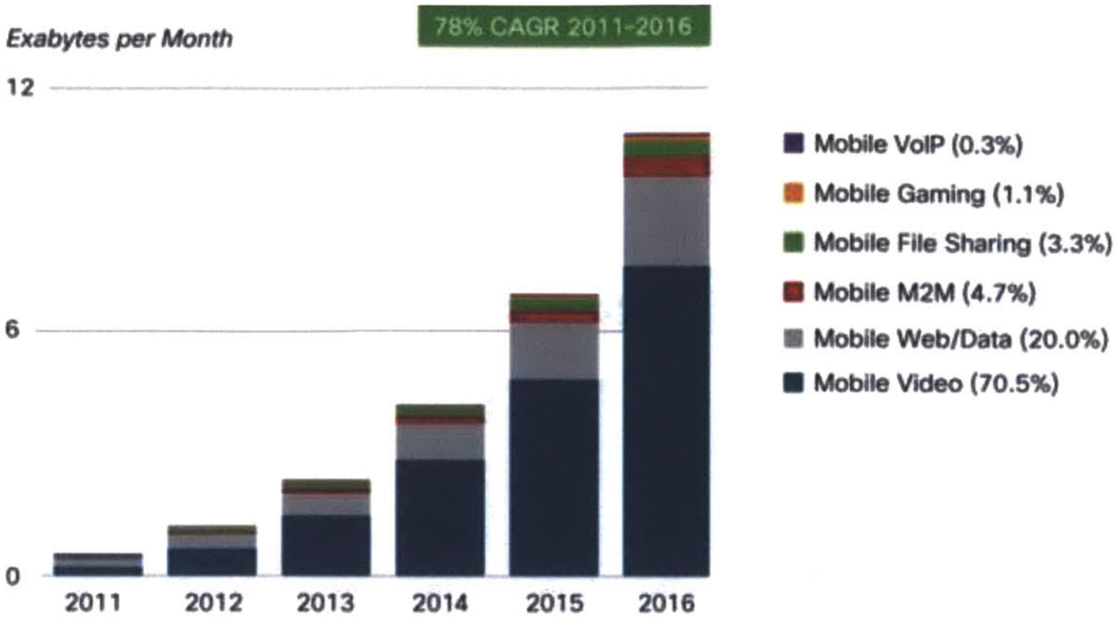
Growing at this rate we will hit 10.8 exabytes[6] by the year 2016 which is 78% CAGR over the next 5 years. Part of this increase can be explained by the increased diversification of devices attaching to the cellular network. As the percentage of smartphones and tablets increase the traffic is bound to increase. A single tablet can generate as much 121 times as much data as a basic feature phone and a smartphone generates as much 35 times more data than a feature phone[7].

The Cisco Visual networking index also gives a classification of the different traffic types that it expects to be present on the mobile network. The distribution is shown in the graph below

---

[6] Source: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html

[7] Source: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html



Figures in legend refer to traffic share in 2016.
Source: Cisco VNI Mobile, 2012

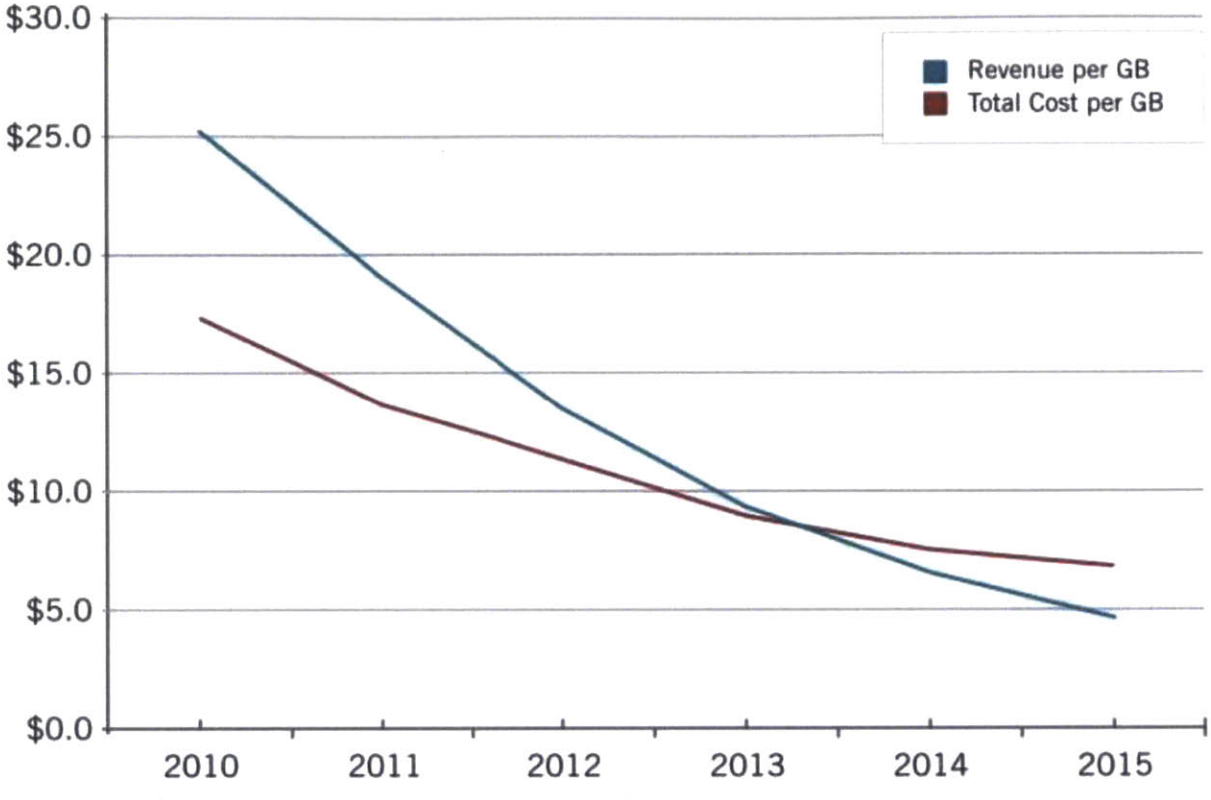
Figure 2: Exabytes per month

As we can see around 70% of the traffic by 2016 will be video. This compounds the problem of handling these large amounts of data as video traffic not only demands high traffic but also a very consistent bandwidth and latency of the connection.

Summarizing the trends discussed in the section above we can see that mobile data is going to increase exponentially driven by faster networks, cheaper data and more capable handsets. A majority of this new traffic is going to be video. We will discuss the effects of this exponential increase in data on mobile network operators in the next section.

2.3 Effects of data increase

The exponential growth in data is what operators have been trying to achieve for a long time as a replacement to declining or flattened voice revenues. The problem is that data revenues are increasing but not at the same rate as the cost for supporting this new growth of data. A Tellabs[8] study in 2010 paints a particularly bleak picture, predicting that by 2014 total cost of transmitting a GB of data could exceed the revenue generated for a GB of Data.



Source: Tellabs 2011

Figure 3: Revenue versus Cost trends(America, Kong, Zealand, & Korea, 2011)

The growth of data is causing an increase in CAPEX and OPEX of the operators as they try to keep up with the demand. Any adjustments in capacity enhancement will cause a degradation of experience for the consumer which will cause abandonment and increase the costs for the operator even more.

2.4 Why Won't Solutions from wired internet work

The kind of data growth we are seeing in mobile data was seen in wired networks a decade earlier and the broadband wired industry did not face the same survival challenge that the mobile industry faces

---

[8] http://www.tellabs.com/markets/tlab_end-of-profit_study.pdf

because of the video data they were transmitting. There are some key differences between wired and wireless video channels related to capacity, error rates and shared versus unique channels for transmission which were highlighted by Lehr & Chaplin(Lehr & Chapin, 2010) in their paper. Capacity is also an issue that Lehr& Chapin raise which we will discuss in detail here. Wired networks have a significant advantage over wireless networks in terms of capacity simply due to the greater frequency range that can be carried over wired networks. Though there have been improvements in modulation techniques used over wireless networks they cannot compensate for the massive frequency advantage that wired networks possess over wireless networks. The significance of limited capacity is shown in the high prices most operators pay for spectrum. This by far is the most important difference between wireless and wired networks. There are other characteristics of radio networks including higher bit error rates and rapidly changing bandwidth and latency due to environmental conditions all manifest in terms of limited capacity.

2.5 Conclusion

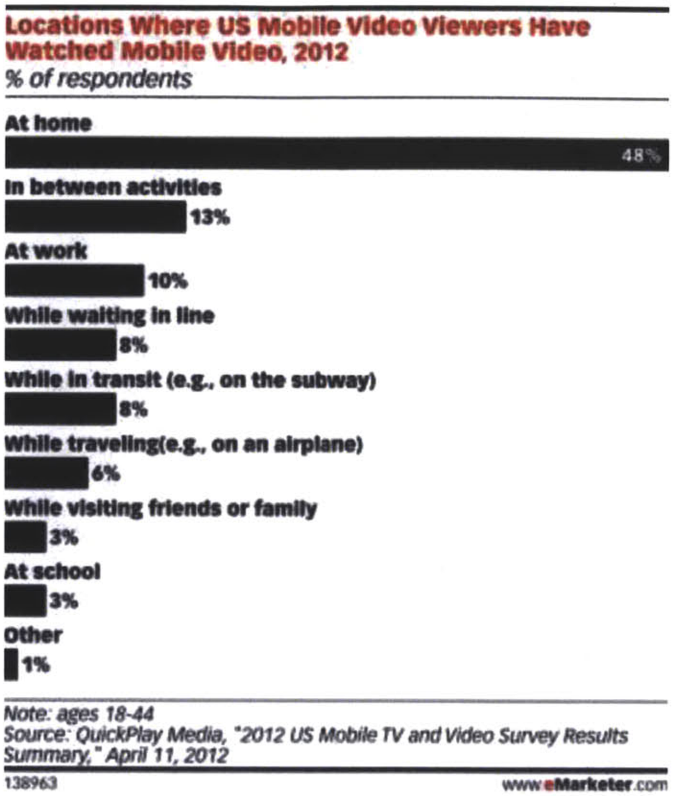
In summary we can easily characterize the problem that is facing mobile operators as one in which they are seeing demand increase exponentially while having a channel with limited capacity to serve them. This trend is bound to accelerate as the percentage of smartphones and tablets and other more capable devices increases. Also new spectrum is scarce or when available is very expensive. Given this problem various operators have been trying different techniques to increase capacity while managing cost during such expansion.

In the next section I will discuss the solution I propose to solve the problem of mobile video data based around Wi-Fi offload and video optimization using scalable video codecs.

3 Solution Architecture

As highlighted earlier the big challenge facing most operators today is the mobile tsunami they are facing. Data traffic keeps on increasing and is going to at least going to keep on like this for the next 5 years. This keeps on increasing the pressure on the network at the same time there is immense competition among operators to capture this new data revenue stream which is putting downward pressure on prices. This means operators will need to look for ways to reduce network congestion while keeping reducing costs and maintaining service levels so as to not lose customers. A variety of solutions exist to mitigate this problem like enhanced charging, backhaul and infrastructure upgrades and traffic management. But all these solutions still end up putting the same amount of data on the network and are temporary solution.

Wi-Fi offload promises to not only reduce the traffic on the operators licensed radio networks but also comes at a reduced cost as the operator doesn't have to pay to acquire spectrum. One more thing that is in favor of Wi-Fi offload is the place where most of the video consumption takes place. The graph below clearly highlights a majority of video consumption on mobile handsets does not happen when the users are mobile but typically happens when they are attached to single Access point.


Locations Where US Mobile Video Viewers Have Watched Mobile Video, 2012

% of respondents

Note: ages 18-44
Source: QuickPlay Media, "2012 US Mobile TV and Video Survey Results Summary," April 11, 2012

Figure 4: Locations where mobile video is consumed

A look at the graph tells us one interesting thing that while majority of the video consumption still happens at home or work where users have a steady Wi-Fi connections. A large percentage of video consumption also occurs while the users are moving slowly between connection points like in transit, travelling or while waiting in a line. The solution we propose addresses these usage scenarios where the user moves slowly between different Wi-Fi networks.

The number of public hotspots has also increased and the graph below shows a trend which clearly identifies the number of hotspots will increase making the percentage of video consumed within range of Wi-Fi AP even greater.

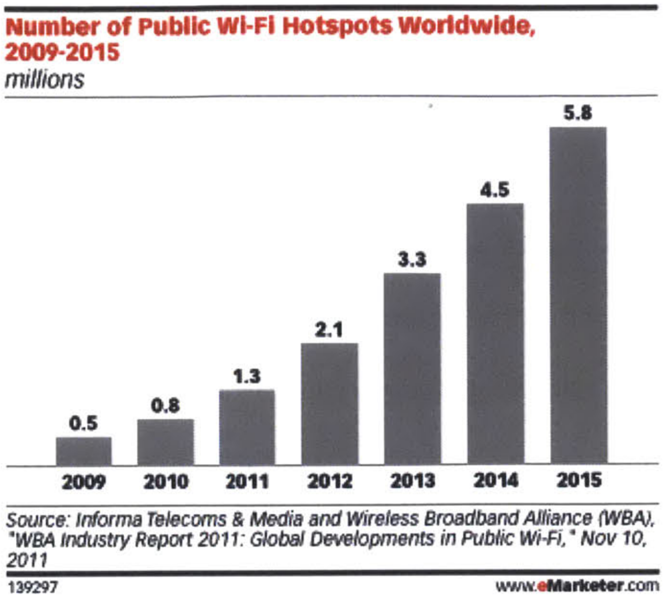


Figure 5: Number of hotspots worldwide

Given current video consumption patterns and the trend of increasing Wi-Fi AP density Wi-Fi offload for video seems pretty appealing but Wi-Fi offload approach has some limitations which I will highlight in the sections below.

3.1 Wi-Fi Offload and Video Streaming Limitation

With the increased density of hotspots and the viewing habits of consumers existing Wi-Fi standards need to overcome some limitations before they can be completely integrated solutions for offload offered by Wi-Fi carriers. Some of the limitations are discussed below

- **Manual Hotspot Discovery and selection**

  The current Wi-Fi methods rely on the user manually discovering the Wi-Fi networks in his vicinity and selecting the right network for accessing the service provider by the operator. This makes the process slow and too complicated for many users.

Also there is no method for the users to know what services the network provides before associating with the network. This wastes precious battery while the users manually associates with each network and discovers the network provides the services the user is looking for.

- **No central login credentials:** Each Access point from a different service provider has a different set of user credentials which is the user is asked to enter when attaching to different networks. There is no standard way of using a single set of user credentials irrespective of who owns the access points. This makes the process more complicated than attaching to a cellular network and also does not allow an operator to use some other provider's access point effectively. In absence of these login credential implementing and managing roaming relationship between Network Providers (one who own Access points) and Service Providers (One's who have a relationship with the customer) difficult.
- **No method for protection from fraud:** There is no standard method to support verification of the Access point. Since the only the client is verified and the only thing that the client knows about the AP is the SSID it broadcasts. An attacker can easily masquerade a SSID and read data passing through an AP.
- **Absence of standard handover methods:** Wireless Access points are limited by range and thus a user walking will move out of range of one access point very fast. If there is no handoff mechanism between the current attach point and the next attach point the user loses his session and deteriorates his quality of experience. 802.21 Media Independent Handover will address this problem for us in the solution suggested below. It provides other advantages as well which will be discussed in another section in this chapter.
- **No standard method to utilize Wi-Fi and cellular as single channel:** The Wi-Fi offload methods currently proposed don't utilize the Wi-Fi and cellular networks simultaneously in the same session but rather make a policy decision to shift the entire session and its associated IP flows to a single

channel. This problem is further complicated by video which can't be effectively sent across multiple paths on the networks and mixed on the client only. This means that current Wi-Fi offload techniques effectively reduce the utilization of the cellular network to zero when they shift a session to a Wi-Fi AP which has smaller coverage area and prone to interference and loss of connectivity as the user moves. This results in worse QOE for the end user. The solution we propose addresses this problem by splitting the session across multiple channels and proposing the use of an encoding technology which is better adapted for multi-interface handsets.

The solution we propose addresses all the above problems using various enabling technologies and custom streaming server and client architectures which allow us to effectively balance the reliability and cost of transmitting of video using a combination of Wi-Fi and cellular networks. The next section highlights some of these enabling technologies before we discuss the architecture of the solution.

3.2 Enabling Technologies

In this section we will discuss the different enabling technologies which have been standardized and make the solution we propose feasible. The technologies we look at in this section are:

- **802.11u:** A new Wi-Fi standard which has been designed to support inter-network operation including vertical handoffs. It provides mechanisms to support new mechanisms for authentication and network selection.
- **802.21:** A new Media Independent Handover standard which has been designed to support handovers between different networks. The standard offers mechanisms to provide information to applications to make intelligent handover but does not participate in the handover process.
- **H.264/SVC:** The scalable video codec standard is designed to provide scalable video without too much added complexity. The standard offers the streaming server/client options to scale video streaming sessions independently along temporal, spatial or quality axes.

3.2.1 IEEE 802.11u

802.11u(Std, Committee, Society, & Board, 2007) adds interworking with external networks to the Wi-Fi standard. Interworking feature allows hotspots and handsets many of the limitations mentioned in the earlier section. Some of the new features supported are discussed in this section.

To understand how 802.11u addresses the limitations let us consider typical use cases and what would happen traditionally and how is it made easier with 802.11u

3.2.1.1 Network Discovery and Selection

When a user typically walks within range of a Wi-Fi network he has to manually scan the list of SSID for all available networks in the area and select one which he thinks is the most relevant based only on the SSID of the network. This means the operator has to ensure his SSID is unique and he cannot effectively use any other network without having them to transmit 2 SSID's.

This coupling between the SSID and network operator/owner is the problem and it has been solved in 802.11u using the concept of OI (Organization Identifiers). A 802.11u compliant access point in its beacon does not only send an SSID but also an OI. An AP beacon can have up to 3 OI's which can be used by the handset to automatically select a network. This means user intervention is no longer required to select which network to attach to as long as the handset is provisioned with the right OI beforehand. The handset also has the option of getting the same information using a probe request without waiting for a beacon from the Wi-Fi AP.

If the initial request does not contain the OI's the handset is capable of attaching to, the handset can request for a list of other OI's supported using the GAS protocol. This is another enhancement to the 802.11 specification. The GAS (Generic Advertisement Service) provides support for exchanging information between the OI and AP prior to authentication or association. Thus a handset can get more information about the AP without having to attach to it. This saves battery on the handset and also ensures that a more intelligent network selection can be made. The GAS acts as a container for different advertisement protocols including ANQP and 802.21 among others. The ANQP protocol is the basic version and supports information about

a.) Venue Type: - This gives the venue information about where the AP has been deployed.

b.) MIH command and event service capability: This is a part of the 802.21 specification and indicates the capabilities of the AP with respect to 802.21

c.) MIH information service: This information element provides the location of the MIH information service which can be queried by the handset using 802.21.

d.) Roaming Consortium List: This Information Element provides the list of OI supported by the AP. This list contains OI's which were not present in the beacon or the probe response.

e.) IP Type Address Availability: This information element is used to indicate the family of IP addresses IPv4/Ipv6 support on the AP.

f.) Network Address Identifier(NAI) List – This is the NAI realm information of the networks supported by the AP. Optionally each NAI is paired with the authentication methods supported by each realm.

g.) Network Authentication Type preferred : This information element indicates the type of authentication preferred by the AP. This could be a list of methods with the first one being preferred.

h.) AP geospatial information: This information element contains the location of the AP.

i.) Domain Name List: This information element is used to indicate the domain name to which the AP belongs.

Using this additional information the handset can now make a more informed decision about which network to attach to using more than just the SSID to arrive at the information.

3.2.1.2 *Centralized Login*

The AP with the interworking Information element and ANQP protocol extension can now support handsets belonging to multiple OI's to attach to the same Access point but we still haven't solved the problem of having separate authentication methods for each access point. The traditional standards don't support any standard method to handle authentication credentials from multiple service providers. Each Access point operator/owner deploys its own authentication methods which means though the association with the AP is seamless the user still has to remember different authentication credentials for each operator/provider. This can be solved if the standard supports a method for the AP to redirect the authentication request to the right AAA server depending on the OI selected by the handset. This is supported by 802.11u using the SSP (Subscription Service Provider) Interface.

The diagram below illustrates the typical architecture of a system using a SSPN interface.

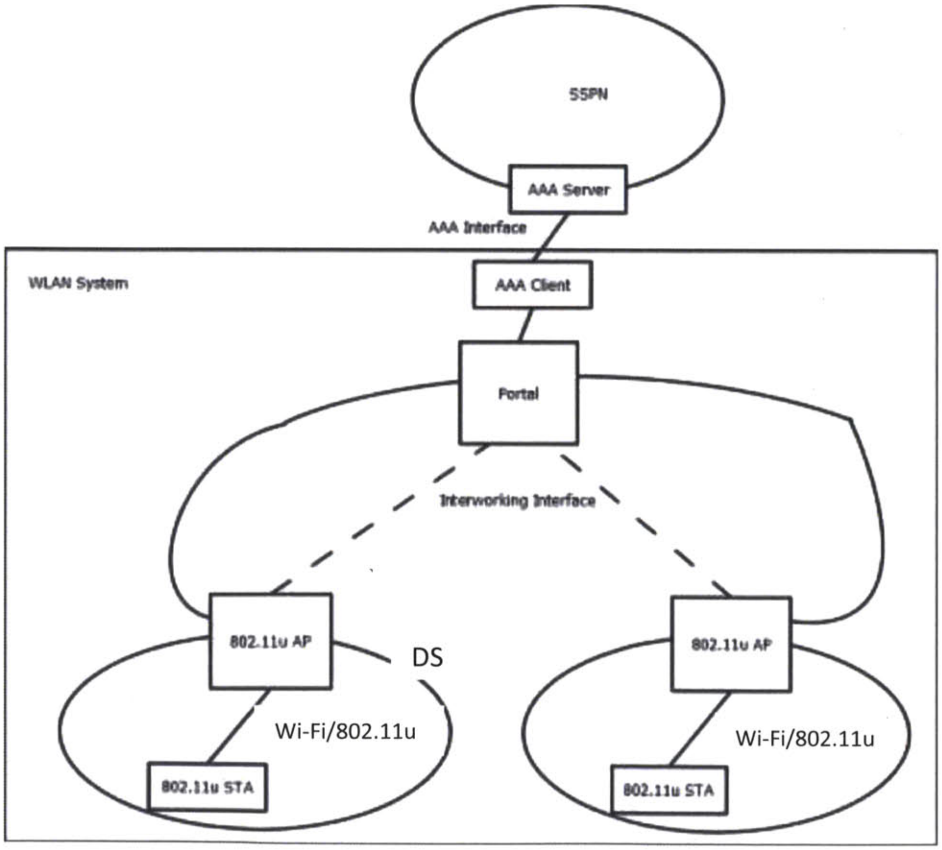


Figure 6: SSPN interface(Std et al., 2007)

The two AP use the SSPN interface to connect to the authentication portal of a service provider. The DS represents the distribution system and need not be owned by either the operator of the AP or the Service provider for the 802.11u stations. The SSPN (Subscription Service Provider network) is designed to be a virtual network over which the SSP interface is implemented. The actual protocol that is implemented between the AP and the portal is not specified and could most likely be RADIUS or Diameter protocol for exchange of AAA credentials.

In this case the user provides an OI it wants to authenticate with or the AP could use the users realm id from his authentication credentials to identify the right service provider to forward the credentials to.

The SSPN interface in 802.11u thus solves our second problem centralized login management for users irrespective of who actually owns the Wi-Fi AP.

Let us now look at a typical 802.11u attach procedure with the SSPN and OI's being used.

3.2.1.3 802.11u Attach procedure

1.) Mobile handset wanders into an 802.11u AP area

2.) Mobile handset receives a beacon from the AP containing

    - Network Type – Chargeable/Free: This indicates whether access to the services provided by the AP are free or chargeable
    - Internet Access Bit : If the bit is set the Access point provides internet access else it does not provide internet access
    - ASRA bit : If it is set means that the network needs additional authentication steps and the STA will be redirected to a portal after association.
    - The top 3 roaming consortium OI list which the AP supports. (One in the beacon + 2 additional ones)

3.) If the STA/handset recognizes any one of the 3 OI sent in the beacon it will attempt association using security credentials of that OI. Else it will send a native GAS (Generic advertisement service) request to get the list of the remaining OI's from the AP.

If the OI was recognized at this point the mobile will need to authenticate with the AP prior to association using security credentials for that OI.

   a. If the RSN element is present authentication method used is 802.1x
   b. If ASRA bit is set and no RSN element is received authentication is web-auth and the handset will perform a GAS query to check what kind of ASRA based authentication is supported.

4.) If the OI was not recognized the handset make a native GAS query. The response to the query contains :

   a. Roaming Consortium list which is a list of OI's not sent in the beacon.
   b. NAI Realm list is the list of realms (mobile service operators) or SSPN which are supported by the AP. The realm list also contains the authentication (EAP) methods supported by each realm.

The response also contains the other elements described in the ANQP response section earlier.

5.) The handset can now select the OI for which it has the credential and send an authentication request to the AP.

6.) On receipt of the authentication parameters from the handset the AP can forward these requests to the correct SSPN's AAA server to authenticate the user and get session parameters from the SSPN.

7.) Once an authentication successful message is sent to the handset it can start the association procedure and associate with the AP.

Looking at the attach procedure we can notice one thing easily that the user no longer needs to intervene to select the right network or enter the username/password or other authentication credentials when selecting a network. 802.11u makes the entire process automatic and seamless for the user and as easy to use as a cellular network. This addresses the first two limitations of Wi-Fi offload.

3.2.2 IEEE 802.21 Media Independent handover

So now we have an AP that is attached without user intervention to a 802.11u AP and the user decides to move to a new location. A Wi-Fi AP due to its low power level is limited to a range of maybe a 100 meters. This means that a nomadic user will move out of the range of the current AP and will eventually lose connection. To ensure that he stays connected even if he moves outside the ESS area we need to be able to execute a handover which is what the IEEE 802.21(Man, Committee, & Computer, 2009) standard supports.

The 802.21 standard implements the handover functionality by implementing a MIHF (Media Independent handover Function) between layer 2 and layer 3 of the standard IP stack. The protocol implements a SAP (Service Access Points) between the MIHF layer and the lower protocol layers to gather information and pass commands to the actual physical interfaces of different technologies and implements a standard IPC mechanism on the upper side to pass this information to the actual application in a technology independent way. The higher layers can send commands or register for events on all the interfaces without actually knowing the details of the technology involved. The MIHF layer translates all the technology specific events from the lower layer into 802.21 events for the higher layers. The typical architecture of a 802.21 implementation for the handset is shown below.

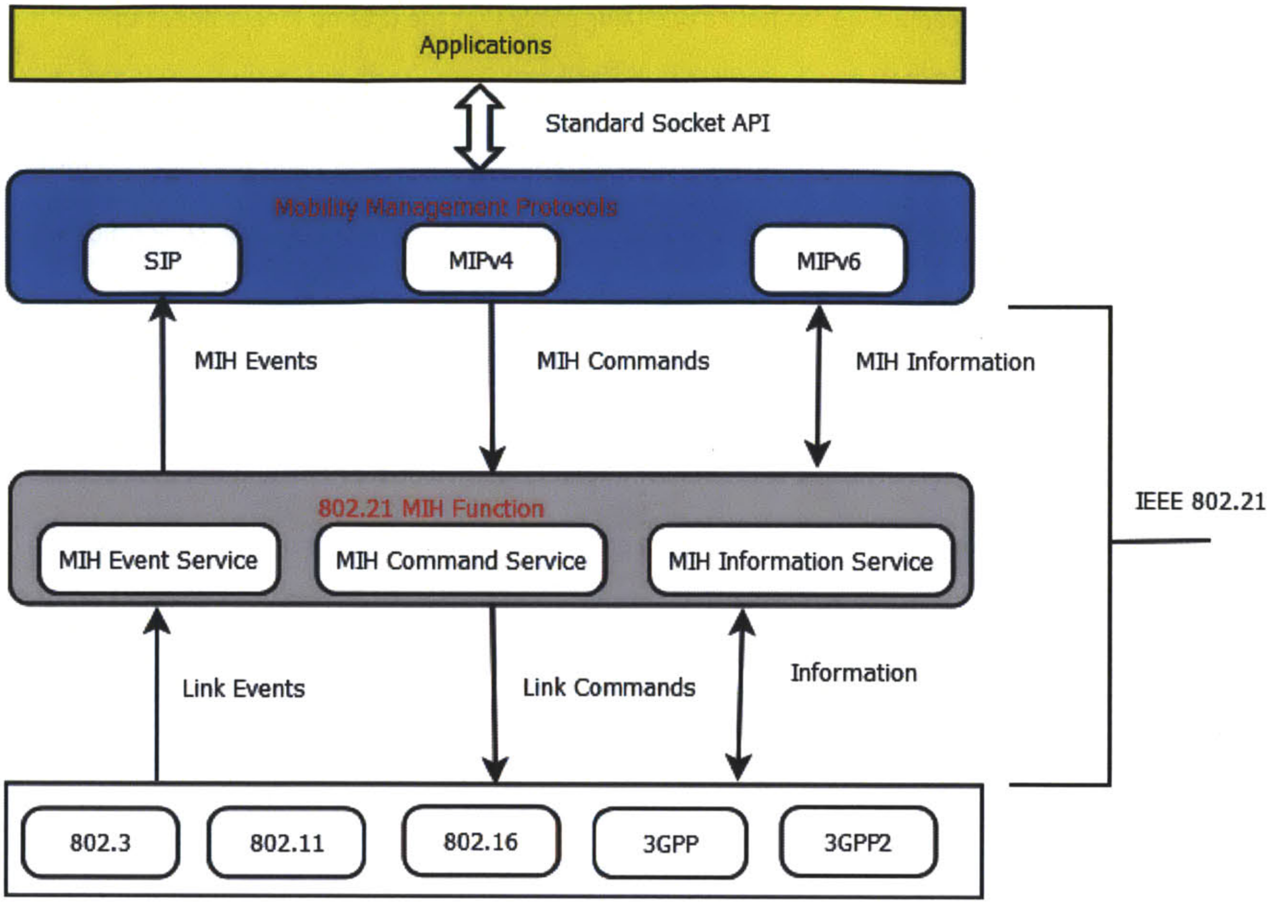


Figure 7: 802.21 architecture

As shown in the figure above the MIH layer can be further subdivided into three services. I will discuss the key functionalities and features of each of those services below.

- MIH Event Service

The MIH Event service acts as a sink for link events which are generated by the layers below the MIH and as a source for MIH events which are sent to the layer above. The upper layers have to subscribe to events they are interested in and the MIH function will translate lower layer events into MIH events for consumption based on subscription. The service can support the following event types MAC and PHY change events, Link Parameter events, Predictive Events, Link Synchronous Events and link transmission events. The 802.21 specification explains all the events in great detail.

- MIH command service

The MIH command Service provides function to gather status of links and invoke commands to control handover process. This is primarily used by the application layers to get the status of the links to make a handover decision.

- MIH Information Service

MIH information service provides access to a network database which contains a list of candidate networks that the handset could possibly attach to. This information is used by the applications during the handover process to select a network when multiple candidate networks are available.

IEEE 802.21 does not itself participate in handover at the higher layers or actually execute handover policy. The details of when and how to execute a handover are left to the higher layer applications There are different handover types suggested on top of 802.21 we will discuss some of these options later when we tackle handovers in the solution architecture.

Once we have solved the problems of network selection, discovery and handover we can now turn our attention to video optimization which we will discuss in the next section.

3.2.3 Scalable Video Codecs

Traditional video systems were based on the older H.320 for streaming and conferencing services. The codecs in these cases needed a network with consistent characteristics in terms of jitter and bandwidth to work correctly without unacceptable degradation in video. As long as majority of the video streaming was happening on comparatively reliable and predictable wired networks these codecs performed satisfactorily and were able to address most requirements of streaming applications. Modern video transmission happens over wireless networks where the latency or the bandwidth of the network can change rapidly dues to either changes in the environment or due to mobility of the receiving terminal. Also the range and capabilities of devices consuming video is pretty diverse in terms of computing power, screen sizes, resolutions they can render and the speed of the networks they can attach to.

Scalable video coding is an attractive solution to most problems faced by most operators wanting to transmit video over wireless networks. The scalable video coding standard which is an extension of the H.264/AVC standard allows scalability along different axis for a given video stream. This scalability allows the adaption of content to either heterogeneous network connections or device capabilities without re-encoding or transcoding of original content. The standard is also backwards compatible with older H.264/AVC decoders and a decoder can still render video of the baseline profile by decoding the base stream without knowing anything about SVC. A scalable codec can efficiently handle problems faced on wireless networks and increase robustness and QOE(Quality of experience) for the end user by supporting graceful degradation of video in case of changes in bandwidth or latency of radio network connections.

The SVC works by encoding a high quality bit stream into multiple lower quality bit streams such that the decoder complexity is not increased at the receiver. The subset streams can represent either a lower spatial resolution, lower temporal resolution or a lower quality video signal compared to the original bit stream. The standard modes supported by scalable video codec are

- Temporal Scalability : Temporal scalability allows the bit stream to be rendered at lower frame rate or a higher frame rate depending on the augmentation streams available.
- Spatial Scalability: Spatial scalability allows larger or smaller picture size depending on the number of subset streams available.

- Quality/Fidelity Scalability: Quality Fidelity allows a higher/lower resolution rate depending on the number of substreams reaching the decoder.

The encoder can also use any combination of the above three modes and encode it to separate streams to support different combination depending on the capabilities of the networks and the handset consuming the video content.

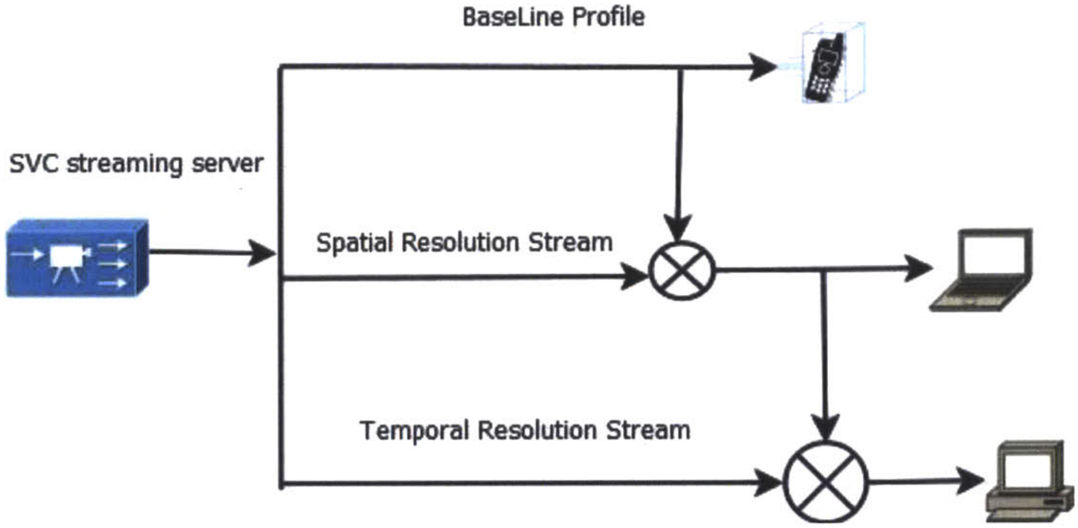


Figure 8: Scalable Video Codec

In the example scenario above the a SVC streaming server sends different streams to different receiving devices depending on the device capabilities and the networks they are attached to. In this case if we assume a simple handset with a 3G connection with a data rate of 384 kbps we send only the baseline profile video which has lower spatial quality reflecting the capability of the device. If the same video is received on a laptop attached to a Wi-Fi network with a higher connection rate and a more capable receiver we can send a higher resolution and spatial quality video. In the third case if it was set top box connected to high resolution HDTV box we would send all the streams to the receiver giving a HD picture on the TV. This kind of application of course requires the streaming server to be aware of the end capabilities of the receiver and the capabilities of the network it is attached to.

3.3 Streaming System Architecture

3.3.1 Introduction

In this section we will use the technologies discussed in the section 3.2 to build a solution architecture which allows us to balance cost versus QOE. Before we describe the solution architecture we will discuss the objective in light of the new technologies and highlight the key assumptions which the solution architecture makes. We will then discuss the high level architecture of a streaming system which can be divided into two different nodes namely the streaming server and the streaming client. The architecture proposed in this thesis adds another node to this architecture in terms of the 802.21 Information Server which enables seamless handover and network discovery for the mobile wireless client. In the next section I will discuss the high level architecture for each of these nodes discussing the protocols required to implement the solution we propose.

3.3.2. Objectives and Assumptions

The key objective of the proposed architecture is to use Wi-Fi and cellular networks to reduce the cost of video streaming to handsets without compromising the user experience. Cellular only solutions offer great QOE at a very high cost while Wi-Fi only solutions offer a lower cost but result in lower QOE for the end user as a result of either loss of coverage due to limited coverage area offered by a typical AP or due to interference since the spectrum is not owned by the service provider. The diagram below highlights these cases.

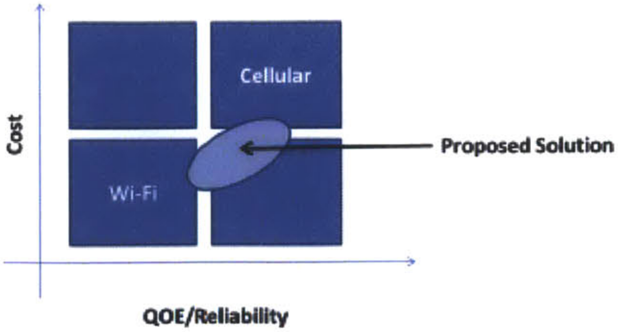


Figure 9: Cost versus QOE

The solution we propose uses a combination of Wi-Fi and cellular networks to increase offer almost similar QOE to a cellular only solution at a much lower cost. Studies like (Zinner) have shown that users are more receptive to continuous content at a low resolution rather than interrupted high quality content a metric which our solution addresses using layered multipath SVC streaming architecture. Using a combination of both Wi-Fi and cellular channels and the SVC codec we will be able to offer a continuous low quality guaranteed image whose image quality changes depending on network conditions. This helps us to achieve a comparable QOE while using the lower cost Wi-Fi network and reducing the total cost of video transmission.

The solution is based on a key assumption of the reliability of a cellular network. The architecture we propose assumes that the cellular network is very reliable and is available when we are starting a video streaming session. This allows us to eliminate the need for adding functional elements to support a Wi-Fi only architecture and also allows us to use the cellular network during handovers on Wi-Fi networks. The assumption of a cellular network streaming a base layer also greatly simplifies the handover process.

3.3.3. Streaming Server

The Streaming Server stores H.264/SVC encoded videos which it makes available for viewing over the network using the streaming infrastructure. It is designed to support multi homed clients capable of receiving and mixing streams for a single video session over multiple network paths. The streaming server also handles Authentication, Authorization and charging for users consuming video through this service. The architecture is geared towards mobile handsets and supports mobility for clients attaching to multiple networks.

The Streaming Server architecture represented as functional blocks is shown below. The role of each of these blocks will be explained later when we describe each of the high level functions of the system. Let us first look at what data/protocol do these blocks implement.

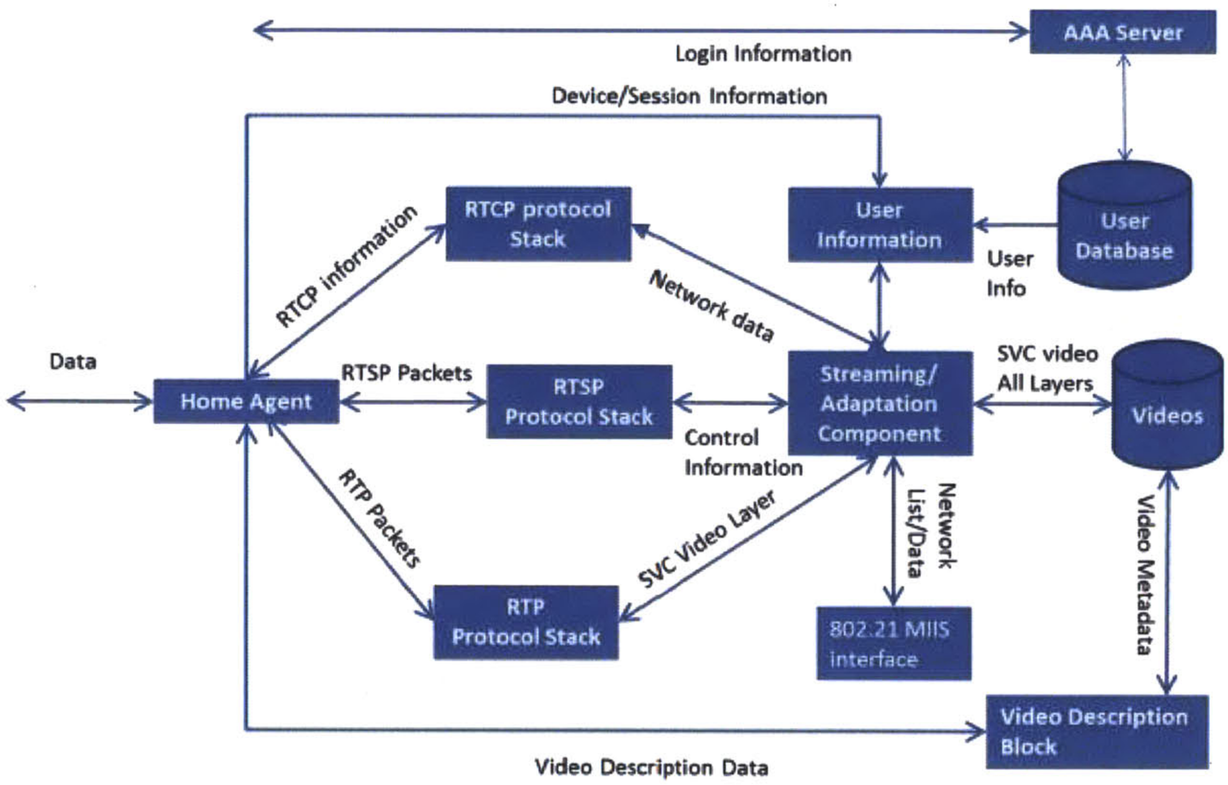


Figure 10: Streaming Server System Architecture

The streaming server uses multiple protocols to carry video data over the network, monitor the quality of the data delivered and provide interactive video capabilities to the client. The details of the protocols they implement and the associated data they carry is explained below. This is followed by the information stored in the remaining blocks and their use.

3.3.1.1 RTP protocol stack

The RTP protocol defines a standardized packet format for delivering audio and video over IP networks. It is used extensively in many streaming media applications like VOD (Video on Demand) application, IP telephony, video conferencing applications. The protocol provides a standard mechanism to deliver video data and adjust for jitter and out of delivery of packets by adding various identifiers to the packet containing the video data. The protocol by itself does not provide an Real Time Transport features and is typically run on top of IP/UDP network. IETF has published a standard RFC (RFC 3550) for specifying all the details of the protocol. We use RTP to carry H.264/SVC encoded video which is a little more complex than the original codecs for which RTP was designed and hence there is a new extension to the original RFC (RFC 6190) which is still in the proposed standard phase in IETF. (Wenger & Schierl, 2007)provides a good overview of the different features added to RTP to provide support for SVC streams.

3.3.1.2 RTCP Protocol Stack

RTCP is a sister protocol of RTP and is always deployed along with RTP. It provides mechanisms to monitor the quality of a RTP connection. In a typical RTP session if RTP service is available on even port than the corresponding RTCP session is on the next odd port. RTCP gathers statistics regarding total octets sent/received, number of packets sent/received, number of packets, number of packet lost, Inter packet arrival delay, jitter in packet arrival rate, Available application buffer, the kind of codec used in current session. The information is periodically sent by the receiver using Receiver Reports and senders using Sender Reports. This information can then be used by the transmitter or receiver to adjust session parameters. In our case this could lead to a lower resolution layer being sent or the complete stopping of the enhancement layers. RTCP in itself does not provide any mechanism to do this and provides only data so that applications can make informed decisions.

3.3.1.3 RTSP Protocol

RTSP(Real Time Streaming Protocol) is a protocol used to control streaming from media servers over the internet. The protocol is used for establishing and controlling media sessions between clients and servers over an IP network. It is a purely control protocol and is not used to carry data in a streaming session. The actual video data in a video session is typically carried by RTP which was discussed earlier. RTSP is a pure text based protocol similar to HTTP and has been standardized by IETF in RFC 2326. We use RTSP in the streaming server during session setup and later as a method for the client to control media streaming sessions.

3.3.1.4 SDP Protocol

The SDP (Session Description Protocol) is a format for describing streaming media initialization parameters. SDP is intended to be used during session initiation, announcement in changes of session parameters and parameter negotiation during session initiation. The protocol is not involved in either control of streaming sessions or actual transport of media information. The protocol in our case will be used to describe SVC layers in response to a RTSP describe message. The protocol has been standardized by the IETF in RFC 4566.

3.3.1.5 RADIUS

The RADIUS (Remote Authentication Dial in User Service) is a networking protocol that provides authentication, authorization and accounting services to computers that use a service by connecting over an IP network. RADIUS is a client server based protocol which runs on top of UDP. It provides standard message formats and message sequences to authenticate users and provide session parameters to these users. The RADIUS protocol has been standardized by the IETF in RFC 2865. The Streaming server uses RADIUS for AAA services and implements a RADIUS server to handle incoming RADIUS requests.

3.3.1.6 Home Agent (Mobile IP)

Mobile IP is an IETF standard (RFC 5944) which allows mobile terminals to move between different networks while keeping permanent address. The mobile IP standard defines a home agent which accepts all traffic from the mobile terminal and then tunnels the traffic to a care-of-address the terminal registers with the home agent every time it gets a new address from the network it attaches to. The protocol allows location independent routing of packets to the handset and decouples the notion of location/identity for the handset terminal. The care-of-address identifies the location of the terminal while the permanent address is used as the identity. The protocol defines registration and routing of packets from the home agent to the mobile IP terminal. The streaming server acts as the home agent for the mobile terminal tunneling packets to destined to the permanent address through tunnels it sets up.

The home agent functional block implements this functionality for the server. It tunnels all RTP/RTSP/RTCP traffic to and from the server. The home agent implements an additional functionality unique to our implementation in that it actually implements two tunnels to the client which terminate on the cellular and Wi-Fi network interface. The signaling required for this will be discussed in the authentication section.

3.3.1.7 User Database & User Information

The User database stores static information about the user connecting to the network. Typical fields in the User database relate to authentication and device information about the user. The User database also contains the service characteristics for the user, the amount of data the user has consumed this month, the user address, the user's cellular network provider and the plan he has subscribed to.

The user information process associated with the User Database stores an in memory copy of the information in the database for all active users. Along with this information it also stores other temporary information like the current location of the user, the network access speed, the current Wi-Fi attachment point, current screen size, and the device that the user is currently using. This information is used by the streaming server to identify the correct session parameters for the users and generate charging information for the user.

3.3.1.8 Video Storage and the Video Description

The video storage block represents storage on the system and it could be a simple RAID or a cloud based system. The details of which technology should actually be used is beyond the scope of this thesis. The only interface the system has with storage allows it to search and read the video files it needs to transmit. Each video file in video storage is also associated with a description of the video and the streams it contains. This is used in response to the RTSP DESCRIBE message during session setup to inform the client about the session parameters for the requested streaming session.

3.3.4. Streaming Client

The Streaming client implements a framework which allows the user to consume H.264/SVC video from the streaming server. The handset is a multi-homed device attaching a cellular 4G network and a Wi-Fi network whenever it is available. The framework implemented on the client allows reception of independent SVC layers over multiple networks. It implements the feature to allow for synchronization of such sessions and ensures that the SVC decoder is presented with packets in the correct order irrespective of the path they were received on. The client also supports a Quality Monitoring system which reports regular statistics of the session to the server which are used by the streaming server to adjust the parameters of the session with the client. The client also implements 802.11u and 802.21 IEEE standards which allows the client seamless authentication to Wi-Fi Access points and horizontal handover capabilities when the client moves from one Wi-Fi network coverage to another. These features are essential to the overall QOE of the session. The client framework also provides a player to render the video on the screen and provides a protocol layer to interact with the server in response to user commands and discover properties of the video and configure the right properties for the different protocol and decoder software in the system. The block diagram of the client framework is shown below. The role of each of these blocks in the Streaming Client will be explained later when we describe each of the high level functions of the system.

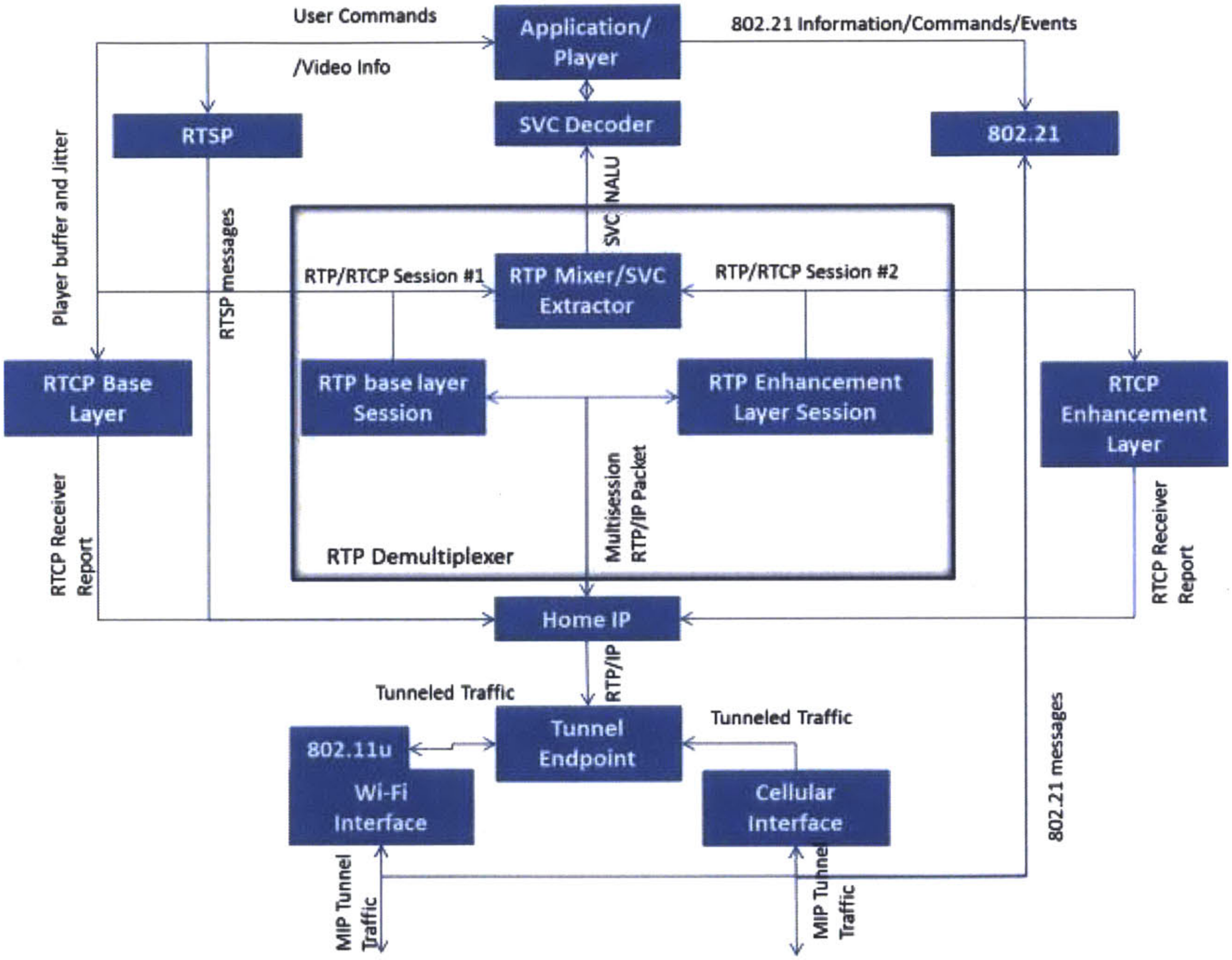


Figure 11: Streaming Client System Architecture

Let us first look at what data/protocol do these blocks implement. The streaming client uses multiple protocols to receive video data over the network, monitor the quality of the data delivered and provide interactive video capabilities to the video consumer. The details of the protocols they implement and the associated data they carry is explained below. This is followed by the information stored in the remaining blocks and their use. Some protocols like RTP, RTCP, SDP and RTSP have already been discussed in the server streaming section and we will not describe those protocols again. It is sufficient to know that protocol stacks in the client streamer act as clients to the streaming server. The RTSP protocol block for example provides API to the player/application to control server streaming or query the server for streaming parameters. RTP protocol block on the client is used to receive RTP packets and identify sessions and streams. RTCP monitors the parameters we had discussed in the streaming server and generates periodic Receiver Reports which are sent to the streaming server. The blocks which don't have a corresponding server block are now discussed below.

3.3.2.1 H.264/SVC Decoder

The SVC decoder is capable of decoding SVC encoded video in multiple layers. The decoder provides API which allows us to feed multiple layers in the decoder which can mix and render this video. This eliminates the need to implement a RTP mixer. H.264/SVC has been discussed in the earlier section and the decoder implements the H.264/SVC standard.

3.3.2.2 RTP Mixer/SVC Extractor

The RTP Mixer/SVC extractor is used to de-packetize the RTP streams and get access the NALU of the SVC encoded picture. Since the decoder allows us to feed NALU's into as different layers we don't need to mix the packets into a single stream. (Wenger & Schierl, 2007)mentions a method to ensure synchronization across different layers when the layers arrive in different streams. Using this we can ensure that the packets in different layers are fed to the decoder in the right decoding order. The RTP mixer implements the required functionality to ensure that the packets are fed to the decoder in the right order.

The RTP mixer also provides RTCP feedback sent to the server. The RTP mixer collects statistics like packet loss, inter-packet arrival delay, jitter and available buffers for each layer and reports it periodically to the server using RTCP Receiver Reports.

3.3.2.3 802.21 Client

The 802.21 Client implements the MIH Command Service, MIH Event Service and the MIH information service as described in the 802.21 section. The 802.21 Link Monitoring Function is implemented in the 802.21 client. It continually monitors the network interfaces on the client to confirm connectivity and availability of these interfaces. It provides the 802.21 event service interface to the application which helps the application know the status of the interfaces and take decisions based on such information. The 802.21 also provides a command service through which the application can control the interfaces and shut down or power up interfaces as required. These two interfaces are very useful when deciding a streaming strategy and the client may request a different layer distribution than the default depending on the network conditions and policy it implements. The layer also implements the client for the 802.21 information service which maintains a local database of adjacent networks and information about the 802.21 information server. The information collected about signal strengths, connection speeds and availability of certain networks by the handset can be synchronized with the information server at the end of a session thus providing more correct information about the availability of networks in a particular area.

3.3.2.4 Home IP and Tunnel Endpoint

The home IP is the permanent address assigned to the handset by the streaming server and it always treats the permanent home IP as the consumer of video. The tunnel endpoint implements the Mobile IP tunnel endpoint. The MIP tunnels extend from the home agent on the server to the Tunnel Endpoint on the client. The tunnel endpoint removes the external tunnel IP addresses and forwards the packet locally to the home IP. It also intercepts all outgoing packets from the home IP and adds tunnels headers to outgoing packets according to the interface it is configured to take.

3.3.2.5 802.11u and Wi-Fi Interface

802.11u was discussed as an enabler protocol in the earlier section and is implemented as a pure software module on the client. The protocol adds essential functionalities for seamless login and discovery of networks. The lower layers implemented in the Wi-Fi are standard Wi-Fi interfaces and provide radio connectivity between the handset and the Wi-Fi access point.

3.3.2.6 Cellular Interface

The cellular interface is a 3G/4G interface attached to a cellular network. The solution we propose assumes that the cellular network is available before the client makes a request for video to the streaming server. The initial authentication/authorization RADIUS requests travel over the cellular network.

3.3.3 802.21 Media Information Server

The 802.21 Media Information Server is a separate entity which provides essential support for mobility in the system. It provides upon request from the handset provides information about candidate networks and their relevant properties to the handset. This feature is used during the handoff procedure by the client to identify possible networks and saves time and battery power on the handset. The basic schema of information stored in the MIIS server is specified in the (Man et al., 2009) specification. The thesis proposes to add two more elements to the basic schema which make the selection of networks and handover process even more efficient as suggested by (Kim, Jung, Kim, & Kim, 2008) paper. The paper suggests storing the available frequencies/channels for each Access Point in the 802.21 MIIS server and sending this information as well in response to a query. Results in the paper show that this can considerably reduce scanning delays for the handset during handover.

The media information server (MIIS) IP address is configured in the handset. Another method could be discovery using a DNS query as specified in RFC 5679 (RFC 5679).

3.4 Functions of the streaming system

Now that we have an idea about the different functional blocks in each of the nodes in the streaming system architecture and the protocols available in the architecture let us look at the functionality needed to present a complete solution. The streaming system functionality can be neatly divided into three different functions which are

- AAA (Authentication, authorization, accounting) -- This function will implement access control, admission control, credential verification, provide infrastructure configuration for video streaming and finally record charging information to charge the user at the end of a session.
- Transporting Video – This function implements mechanisms to transport video to authenticated and authorized user attached to a IP network.
- Mobility – the streaming system proposed is supposed to deliver video to mobile handsets attached to multiple networks. It thus needs a method to not only identify the different networks the handset is attached to but also provide a method for seamless handoff when a handset wanders from one network coverage to another.

The rest of the section addresses each of these functions and demonstrates the delivery of each of these functions using the architecture defined in the earlier section.

3.4.1. AAA (Authentication, Authorization, Accounting)

The AAA function provides for authentication of users, authorization of specific services for these authenticated users and finally charging for the services provided. The streaming server implements the AAA server while the client implements the AAA client.

Authentication for the streaming server system is a two-step process in which the client first authenticates with the streaming server over its cellular interface and then over the Wi-Fi network to get access to the Wi-Fi network before we start delivering video.

The first authentication step is a simple RADIUS authentication request where the user sends username/password combination and its current location to the AAA server. The AAA server references the User Information database to check the validity of the pair and accepts or rejects the connection. On success the AAA server will return four things which are used by the client to establish the next phase of authentication.

1.) NAI/OI: This returns the list of OI the client should scan for when it brings up its Wi-Fi interface. This list could vary depending on different service levels or roaming agreements with different partners in different geographical areas.
2.) Session Key: This is a temporary identifier valid for the session assigned by the server during session setup. This session key is used while sending all accounting/charging requests for this session.
3.) 802.21 MIIS Server Address: The MIIS Server Address can either be a FQDN which the client can look up using DNS query. A method to do this based on a DNS address has been specified in IETF RFC 5679 (http://tools.ietf.org/rfc/rfc5679.txt). The MIIS server can be used to identify networks which provide the level of QOS and other parameters required by the handset framework to receive video
4.) Home IP Address: This is the address assigned by the streaming server to which the server will stream all traffic. This is attached to a virtual interface which gets packets forwarded from tunnel endpoints implemented on top of the Wi-Fi and cellular interface.

Once the user has been successfully authenticated on the cellular network and the handset has all the parameters necessary to make an intelligent network selection it will power up its Wi-Fi interface and start scanning for a Wi-Fi network to establish a connection.

This is where the 802.11u protocol we have discussed earlier plays a very important role since it allows the handset to look for OI and request a list of all roaming partners it supports before associating with the network. The 802.11u also provides an SSPN interface functionality which allows the handset to use a single set of credentials to authenticate itself on any network irrespective of the operator of the Access point or the network behind it by allowing the AP to delegate authentication of the user to a set of AAA server selected on the basis of the realm id/OI/NAI supplied by the user during authentication. The authentication process for 802.11u along with the SSPN interface was not discussed when the 802.11u protocol details were discussed in the earlier section. The section below discusses the network setup and the typical message sequence during a 802.11u authentication and association for a roaming handset.

The diagram below shows a typical 802.11u AP with SSP interfaces towards multiple SSPN's

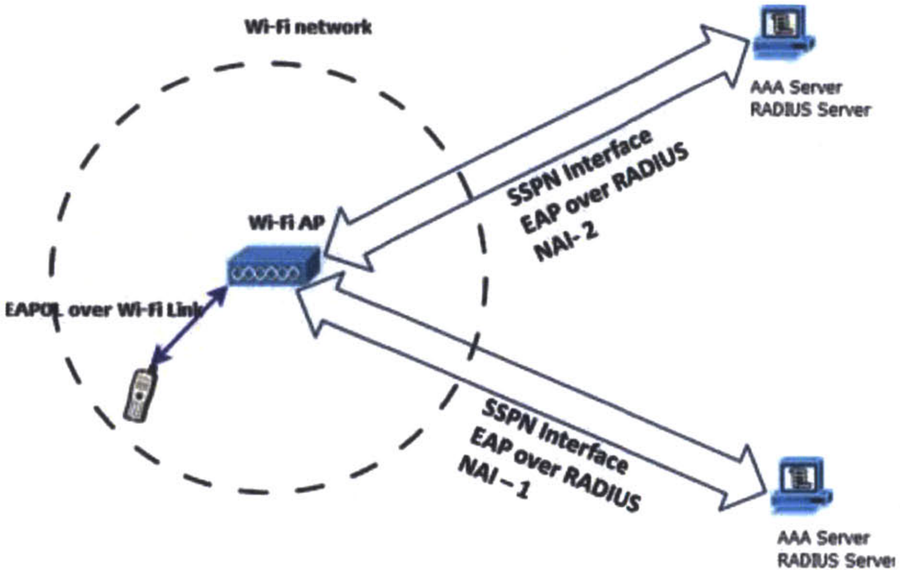


Figure 12: Single AP forwarding to multiple AAA servers

The diagram shows the protocols that I propose for implementation on the SSP interfaces. The RADIUS protocol is used between the AP and the AAA servers of the SSPN providers. The EAPOL protocol is used between the AP and the handset. The Wi-Fi authentication is more involved since the handset is not associated with the network already and the message sequence for authentication is shown below.

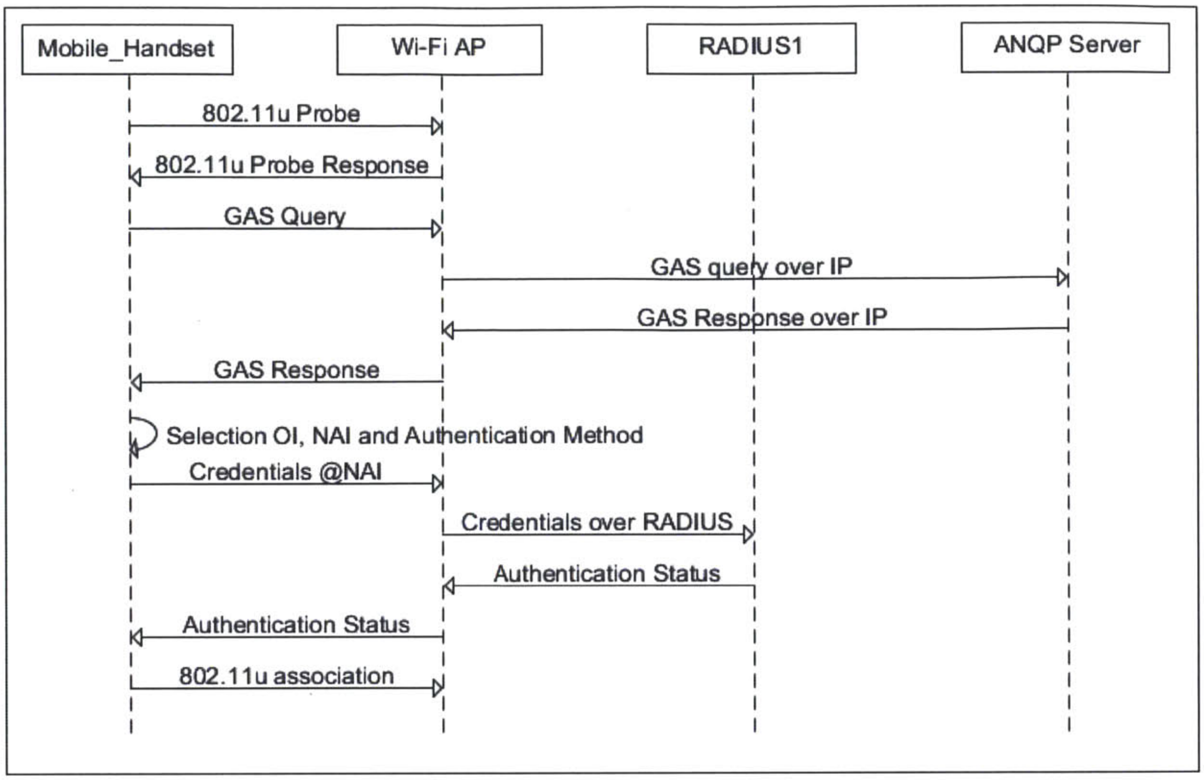


Figure 13: 802.11u authentication and attach

1.) A handset wanders into a Wi-Fi area covered by 802.11u compatible AP and sends a probe request to get the list of SSID available in the area.

2.) All AP's respond with the primary OI list. If the handset knows the OI and the authentication method support it can start authentication at this point using the known method.

3.) Let us assume that the handset doesn't know the authentication method in which case it will send an advertisement request for getting information about all the possible OI/NAI realms supported by the AP along with the supported authentication methods.

4.) The Access Point forwards this request to an advertisement server which can provide this information along with other information. The authentication information is what we are interested in at this point.

5.) The Access point forwards this to the handset as a list of NAI realms and their associated authentication methods. The NAI information element for 802.11u frame format is shown in the figure below.

Info ID	Length	NAI Realm Count	NAI Realm Data#1 (optional)	....	NAI Realm Data #n (optional)
2	2	2	Variable		Variable

Figure 14: NAI Information element frame format

Each NAI Realm Data element further optionally contains the authentication method supported for each realm. These are stored in the in the EAP method sub-field and contain the list of authentication methods that the NAI realm supports. The values are coded according to the EAP-IANA standard. (Table) shows the list of supported EAP methods.

The handset software can then choose the right NAI realm depending on configuration and use the right authentication method to authenticate with the realm. 802.11u AP can route the authentication request to the appropriate RADIUS server over the SSPN interface using the realm ID provided by the handset. The actual authentication protocol details are not discussed here.

Once this authentication process has completed and the handset is now attached to the media server using two network interfaces. The client software will inform the server of its IP addresses on the cellular and the Wi-Fi interfaces. The server will configure tunnels to each of these interfaces from the home agent functional block on the server to the tunnel endpoint on the client. At this point the server-client network connection is complete and the client can now start requesting for video. The protocol layer diagram for the IP network at this point looks like

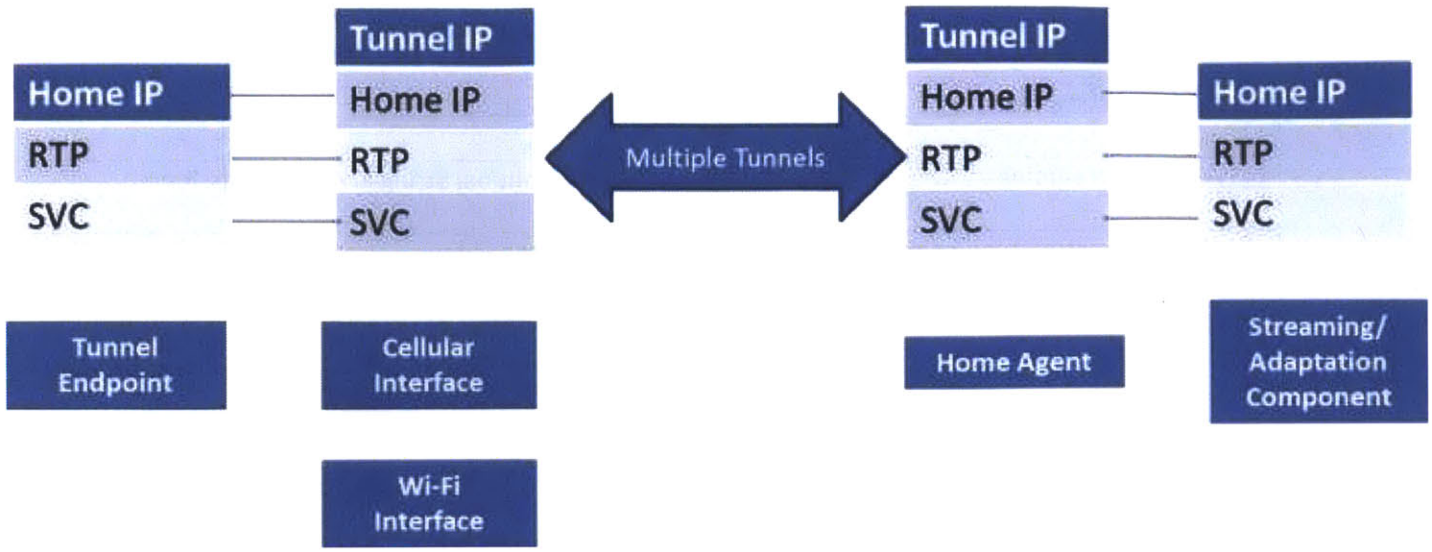


Figure 15: Protocol stack and packet headers

3.4.2 Transporting Video

Once the network has been setup the client can now request video. The interface and the protocols to actually search and finding the URI of the video is beyond the scope of this thesis. It can be implemented by a webserver which provides an interface to the video library available on the streaming server. Once the client has chosen the video it wants to view it can locate the actual streaming server which will stream the video content. The client needs to discover the properties of the video stream it is going to be receive and uses the RTSP Describe primitive to retrieve the properties of the stream. The server responds to the RTSP DESCRIBE message with a description of the video streaming session using the SDP protocol. A typical response in the SDP protocol format to a RTSP describe request is shown below.

v=0	Version Information
o=owner identifier, 2,3	Owner information and two session identifiers for

	2 RTP sessions
s= Video Consumption	Session Name
i=SVC video streaming	Informational String about the session
u=http://videostreaming.com/video	URI of the video
e= admin@videostreamng.com	Email address of video owner/administrator
c=xxx.xxx.xxx.xxx	IP address of media server
a=recvonly a=orientation:landscape	Attributes of the media stream
a=group:DDP 1 2	
m=video 27000 RTP/AVP 31	The details about the video session. This implies the session carries video from port 27000 and the protocol is RTP/AVP with format identification 31
a=rtpmap:31 H264/90000	This maps the payload type from m="" to a encoding name and the clock rate
a=mid:1	Group identification (Refer RFC: http://tools.ietf.org/html/rfc5583)
m=video 28000 RTP/AVP 33	The details about the video session. This implies the session carries video from port 28000 and the protocol is RTP/AVP with format identification 31
a=rtpmap:31 H264/90000	This maps the payload type from m="" to a encoding name and the clock rate
a=mid:2	Group identifier (Refer RFC : http://tools.ietf.org/html/rfc5583)

a=depend:lay 1	This layer depends on layer 1 (Refer RFC : http://tools.ietf.org/html/rfc5583)

Figure 16: RTSP Response in SDP format

The standard SDP protocol was not designed to support layered video and thus did not have a mechanism to define dependencies between layers. The draft specification 5583 addresses this shortcoming by adding the "depend" and "mid" attributes to the SDP protocol.

The client streaming application parses the SDP message received and triggers the RTSP stack to setup sessions corresponding to the number of streams reported in the SDP announcement messages. RTSP also doesn't understand the concept of layered media and thus needs two setup messages to setup a video streaming session and the RTSP PLAY command is also sent independently for each of the layers. A typical exchange including the RTSP DESCRIBE message and the application response to it are shown in the message diagram below.

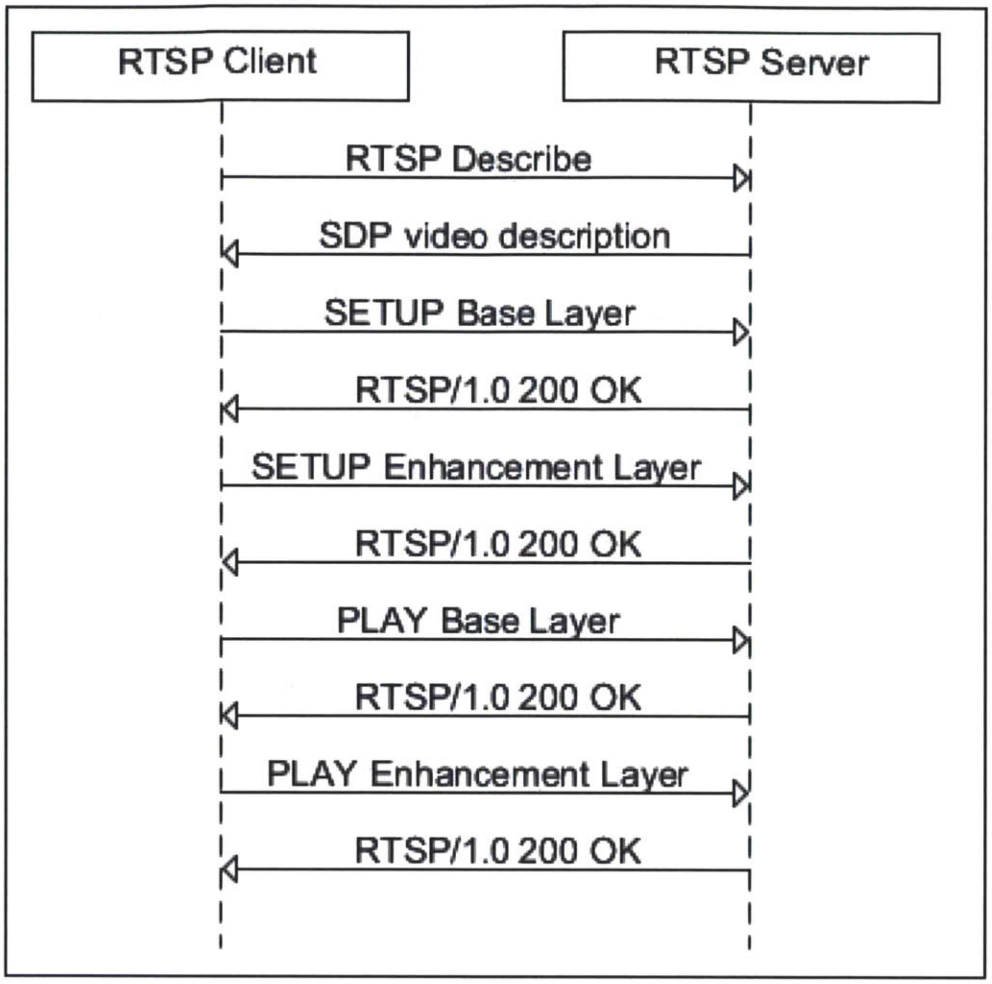


Figure 17: RTSP message to play layered video

A recent draft submitted to the IETF (https://datatracker.ietf.org/doc/draft-yue-mmusic-rtsp-substream-control-extension/) specifies a method of adding support for sub-stream control in the RTSP protocol. When a server receives a RTSP setup message for a particular stream it sets up a RTP session with the port mentioned in the SDP message and Source and Destination IP addresses. Thus at the end of the two RTSP setup messages we have RTP sessions created on the server and client for base layer and the enhancement layer of the RTP stream. Corresponding to each RTP session at both ends a RTCP session bound to next higher number port is established. RTCP is used to monitor the quality of the session.

The setup of the video session once involves interaction between multiple components between the server and client systems as shown in the diagram below

1.) Once the handset has authenticated successfully it receives a home IP from the server which it assigns to a local virtual interface. This IP is used by the streaming application.
2.) The handset also informs the streaming server using the cellular interface the Wi-Fi IP Address and the cellular IP address. This will be used later by the streaming server to configure the tunnels.
3.) The exact method to get the video URI is beyond the scope of this system design but a standard web server can be used as a front end to provide URI's for all videos accessible in the streaming servers media library. The client can get this URI using the http protocol if the front end is a http server.
4.) The client gets the streaming server IP address and the location of the video in the earlier step. It can use this information to get more information about the sessions it needs to setup tunnels later.
5.) The RTSP describe message is used to get the description of the video. The response is a SDP protocol message as described above. The message identifies the properties of the RTP session at the server and the number of layers the server is going to be sending. This is important for the client since we setup a RTP session for each SVC layer received on the client. The RTSP server gets the information from the streaming servers which reads the actual video information from the video server storage and the remaining parameters are configured by the streaming server depending on the information the streaming server obtains from the authentication

block. Information like screen size, device type and network type can be used by the streaming server to decide the appropriate parameters of the streaming session.

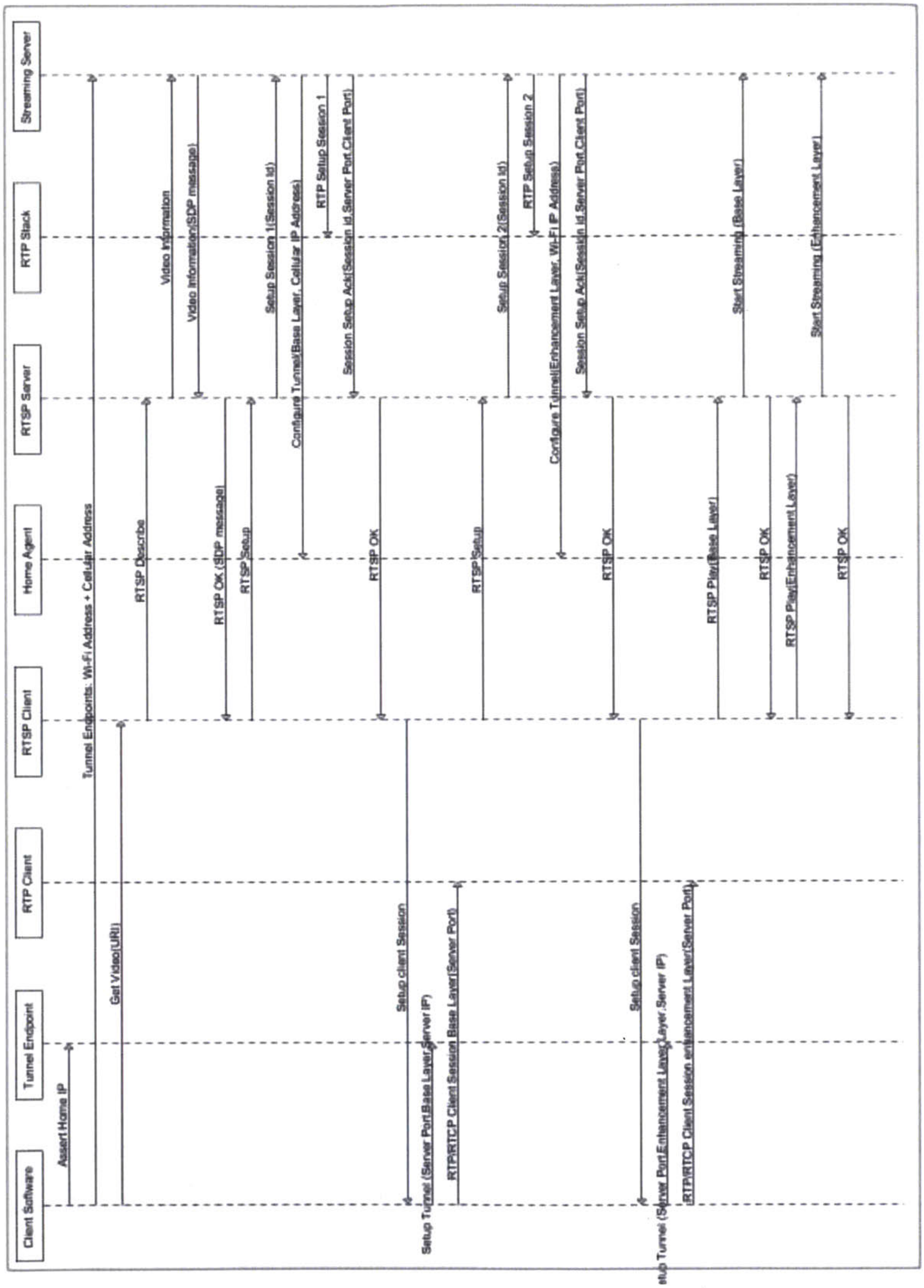


Figure 18: Video Session Setup message sequence

6.) The client in response to the SDP message is going to send a RTSP setup message for each of the layers communicating the local ports it wants to use for the connection and in response get the ports the server is going to use for each RTP session. This is also the point where the client and server communicate about which layers to send over which tunnels. The default behavior in the system at startup is to always use the tunnel over the cellular network to send a base layer while using the Wi-Fi tunnel to send enhancement layers. The client and the server both need to independent set rules on their ends to ensure that the right packets get routed over the right networks. This can be achieved by either using simple ip filtering rules or by building enough intelligence into the home agent/tunnel endpoint so can that they can look inside SVC layers to know which layer is carried in a particular packet.
7.) The RTSP Play is the final message that is sent across to play the video once the sessions have been setup. This will start the streaming of data from the video server. Currently RTSP requires that the play command be sent for each layer.

Once the video session has been setup RTCP is used to monitor the quality of the session. A RTCP session is run for each RTP session and the streaming server will get quality reports for each layer it sends separately this information can be used to alter the characteristics of the video stream in terms of the resolution, size or the frame rate of the video stream sent in the enhancement layer. The client can even request that the base layer be sent over the Wi-Fi network instead of the cellular network if it is in an area with bad cellular coverage.

The decisions that the streaming server has to continually evaluate during the scalable video streaming session can be broken down into a path selection problem and a video adaptation problem. There has been a lot of work around the selection of right paths and right adaptation layers. Some of the approaches discussed are not ideal for the solution proposed and some of these approaches address only a part of the problem which needs to be integrated with another approach to form a complete solution. In the paper (Singh, n.d.) the authors develop a model to deliver scalable video to a single client each attached to N networks. The paper presents an algorithm for deterministic packet scheduling for each client given the capacity of each network they are attached to and the error rates associated with each network. The algorithm works on a packet level where each NALU is scheduled to be delivered depending on what will cause the minimum distortion. It therefore assumes that all the layers are sent in a single RTP session. In the paper (Nightingale, Member, Wang, Grecos, & Member, 2010) authors propose another method to schedule packets independently over each network leg without having a strict rule about which layers to send over which network. This again implies a single RTP session spanning multiple paths.

(Sutinen & Rivas, 2011) uses multiple RTP sessions for streaming each SVC layer independently and also uses a fixed path for each layer of the system. The paper highlights one more important mechanism which we will implement in the client side software, ability of the client to request a mapping of each layer to specific interfaces is very important. When the cellular network becomes unreliable due to either the user moving indoors or reaching a hole in the coverage area the client can request that the base layer be also sent over the Wi-Fi network. I will elaborate on this in the handset framework section.

Since we are sending only two layers to the client over the two independent networks it is attached to we have a very limited capability of implementing scalability where the difference between the full picture and the picture missing the enhancement layer is pretty large. The SVC codec supports more than one enhancement layer and thus the server can adapt better if it mixes these layers prior to forming RTP sessions depending on RTCP feedback. One such algorithm is described in (Beaulieu, 2011) where the authors propose a scheme to add or remove enhancement layers from a video stream depending on the available bandwidth on a channel. The algorithm uses a combination of client screen size, client capabilities (computing power, supported codecs) and the capacity of network connections to calculate the number of enhancement layers and define a streaming strategy for a video streaming session with the client. RTCP feedback can be used to trade off packet loss/jitter with available bandwidth and use the algorithm to determine the number of layers to choose.

The way we use the layer selection algorithm means we need to independently be able to assess the packet error rates and jitter for the enhancement layers. A separate RTP session for this layer greatly reduces the complexity in software for monitoring this data.

In summary the architecture uses the algorithm defined in (Beaulieu, 2011) to dynamically adapt the bit rate of enhancement layers or drop the layer altogether if we lose Wi-Fi connectivity and use the method developed in (Sutinen & Rivas, 2011)for path selection.

3.4.3 Mobility

3.4.3.1 Introduction

Session continuity for a user viewing a video is very important for a good quality of experience in video streaming session. Wi-Fi AP's have a limited range typically limited to not more than 30 meters. This means that a mobile user will probably move out of the range of the current access point during a session and lose connectivity with the Wi-Fi networks unless and until the system in co-operation with the handset handover the connection to a new access point. The mobility procedure solves the problem of ensuring session continuity by defining a handover procedure which makes the process seamless for the end user.

3.4.3.2 Mobility procedure

The mobility function provided by the streaming system implements seamless handovers for the handset while the handset moves from one AP to another. The function is implemented different co-operating modules implemented in the client streaming framework, server streaming framework and the 802.21 MIIS server. The method suggested here has been investigated in (Kim et al., 2008). The message diagram has been reproduced from the paper with slight deviation given the time sensitive nature of video streaming. The buffering of the packets is not necessary since the handset still keeps on receiving the base layer video on the cellular network while the handset attaches to the new Wi-Fi acess point. These packets will not be useful when they arrive later since their presentation time might have already passed. It is much more useful for us to pause these packets at the streaming server itself, which is what the RTSP pause we propose achieves.

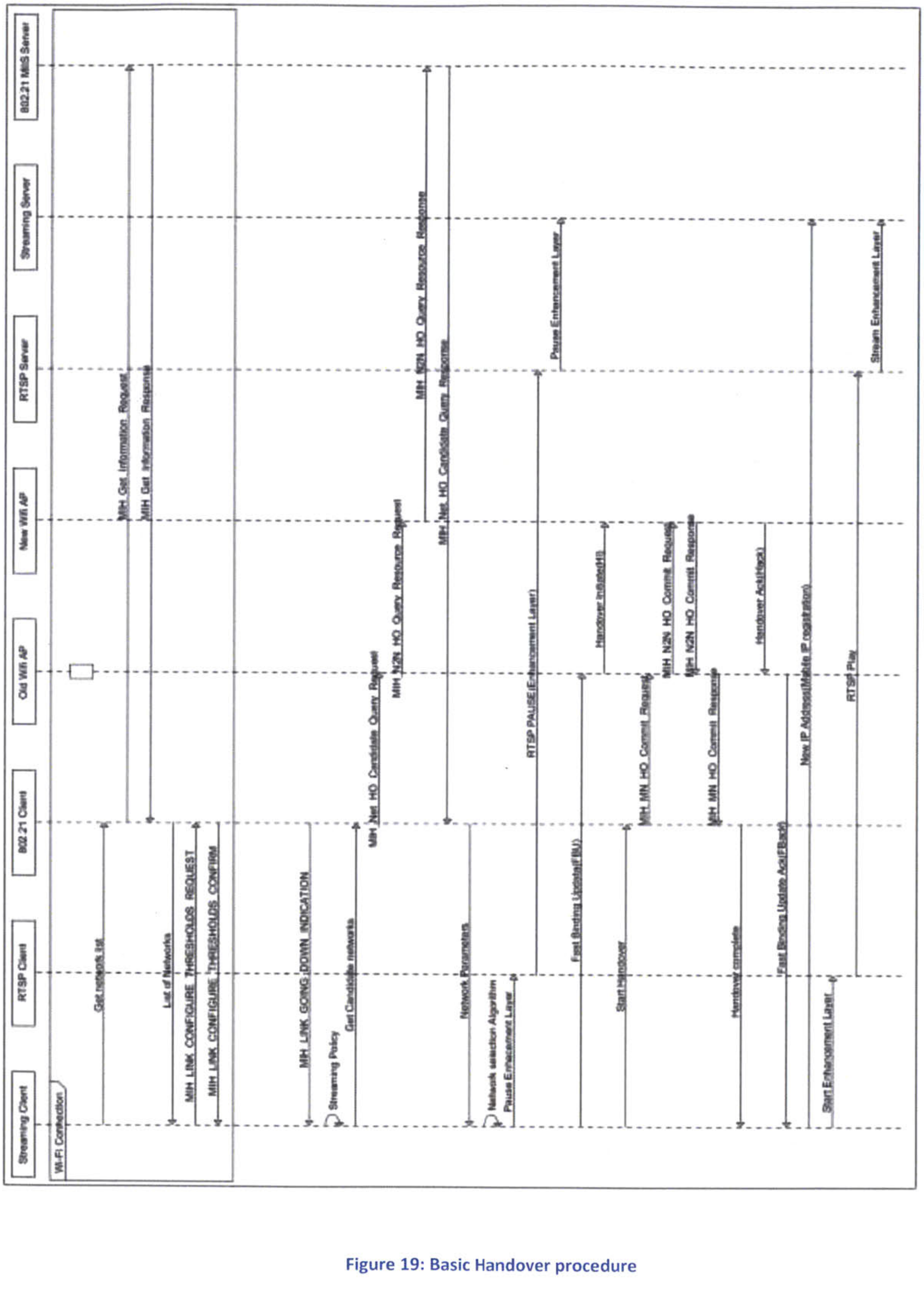


Figure 19: Basic Handover procedure

1.) The handset is already attached to a Wi-Fi network and during the initial attach it queries the MIIS and gets information about all the available networks around its current location. It also configures the threshold levels for either signal strength/packet loss on the Wi-Fi interface on the 802.21 stack so that the 802.21 MIH stack will signal the streaming client in case any of the threshold conditions are met.

2.) As the handset moves through the area and reaches the edge of the current AP's coverage area the 802.21 stack generates a link going down message. This prompts the client application to start the handover process.

3.) The client application in a simple case looks up the initial list of networks it has and sends a request to 802.21 to get candidate network properties for the networks selected by the algorithm. The 802.21 will in turn forward it to the currently connected AP which will gather the required statistics for each of candidate networks. This is achieved by an 802.21 implementation on each of the participating AP's. The properties that get exchange could be subnet id's provided by each AP, the QOS, the availability of resources, the number of free channels.

4.) This list of networks along with the associated properties are sent back to the handset which will now select the target network it wants to attach to.

5.) At this point we send a RTSP PAUSE request to the streaming server to stop streaming the enhancement layer. The client still receives the base layer from the tunnel on the cellular network. This approach requires that the client actually have a buffer to store received video streams and play it after an initial delay. This buffer can be used when the handset loses enhancement layers when it does a handoff from one AP to another. Streaming the enhancement layer over the cellular network while we do a Wi-Fi handover could be another option we can explore.

6.) After Pausing the enhancement layer we can now execute a Mobile IP handover and a L2 handover. The Fast Binding update is just to ensure that packets in transit get forwarded to the new Access Point while the handset gets IP address and authenticates with the new AP. By pausing the stream the number of packets reduces substantially.

7.) The client next executes a L2 handover using the 802.21 commands the MIH_MN_Commit_Request informs the current Access Point that the handset is going to terminate the current L2 Session and associate with the new AP.

8.) The client can retrieve the active channels to scan for the new AP as a result of having retrieved it with the MIH_Net_HO_Candidate_Query_Response message. This reduces the scanning time when associating with the new AP.

9.) Once the handset has associated with the new AP the client does the MIP signaling required to register the new IP address with the home agent.

10.) The client has the Wi-Fi interface up again with the tunnel established and it sends RTSP Play to start streaming the enhancement layer again.

3.5 Summary

In this section I have discussed the solution architecture and the different network elements and protocols needed to implement this solution. We have also defined a set of basic procedures using these elements for streaming video, charging users and managing handovers between networks. The methods to stop video or cases where only one network is available can be addressed using subsets of the same procedures discussed above. The bit rates for the enhancement layer and the base layer are completely configurable in the system and can be adjusted according to the quality of experience the user might desire. The next section will look at the stakeholders for a Wi-Fi offload. We will focus on identifying different players within the mobile business ecosystem.

4 Stakeholders

4.1 Introduction

The previous chapter discussed the technologies and the proposed architecture for streaming video over heterogeneous networks. The solution involves new software at various points in the network and requires co-operation from various players in the mobile content value chain.

This chapter will focus on the different players in the value chain what are their needs, how strong is their influence on the value chain, what service do they provide and ultimately what opportunities does the new architecture proposed provide to these players.

4.2 Stakeholder Analysis

Mobile content value chain has multiple stakeholders starting from the content producers , aggregators, distributors, telcos, and Wi-Fi operators and finally the consumer. A high level stakeholder diagram is shown below. We will discuss the needs, technological capability, Influence they have and opportunities our solution might create for each stakeholder.

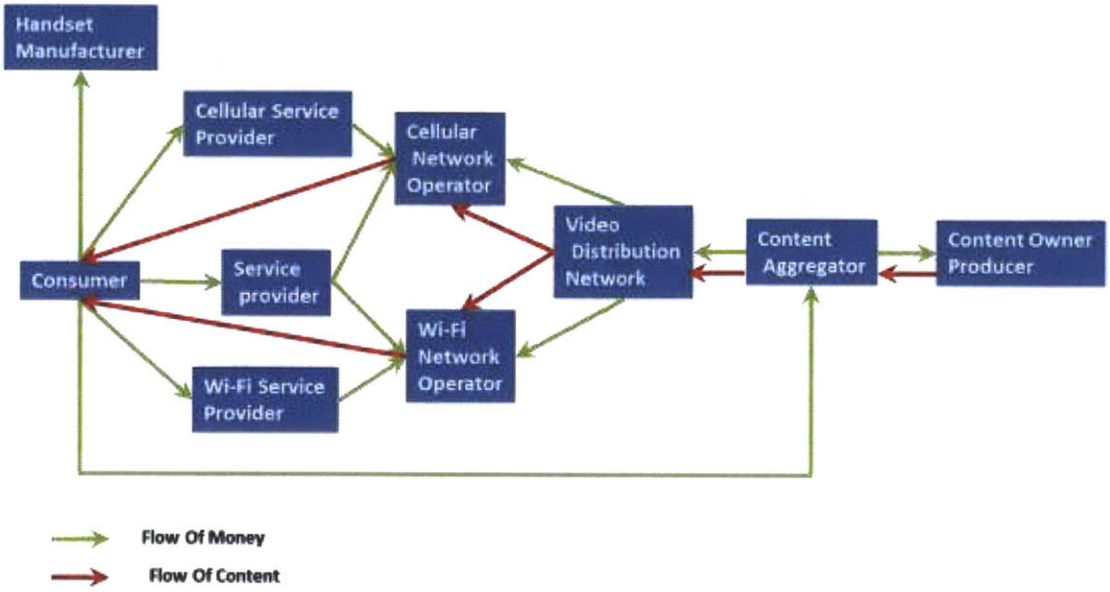


Figure 20: Stakeholder diagram

4.2.1 Content Producer

The content producer is typically the generator of original content for consumption by the consumer. The typical content producers are movie studios or TV production houses like CBS, NBC.

Need: Content producer typically need channels to distribute their content. They charge advertisers to air advertisements embedded in their content. Thus the primary need of a content producer is to have more channels so it can be viewed by many users and it can thus get more advertisement revenue from advertisers or direct revenue from consumer. One more important need for the content producers is the assurance that their content will be protected from unauthorized copying or transmission.

Technology: Content producers primary technology asset is knowledge of content production which is interesting to consumers. Their primary knowledge is in content authoring, encoding so that it can be consumed by the largest variety of devices. In our solution architecture the content producers need to support H.264/SVC encoded videos for the solution to work without the added cost of transcoding videos at the aggregator.

Influence: Content producer are the owners of content and thus control the channels and formats in which their video is delivered to consumers. They also sometimes have direct relationships to customers and understand customer choices well. They also have a strong relationship with advertisers and thus control the revenue from both ends. They exert a very strong influence on the value chain in mobile content and their support is critical to the success of the solution.

Opportunities: The new solution architecture proposed allows the content producers to provider a better quality of experience to the end users at a lower cost to the users. This allows them to reach mobile users who would not be reachable with the traditional mobile content delivery model. This provides typically more users who can view the content and is a new opportunity for the content producers.

4.2.2 Content Aggregator

The content aggregators aggregate content from different producers and provides a package to consumers which they can subscribe to. On the internet sites like Netflix, Hulu provide the content aggregation service. Cable operators like Comcast, Time Warner also provide content aggregation services in the wired world and are now moving to providing similar services online.

Need: Content aggregators need to aggregate the right variety of content for consumers so as to make their packages attractive to the consumers. They are responsible for enforcing the digital rights of the content producer and have relationships with content distribution companies so that they can distribute this content to the consumer. Their primary need is to have a large number of content producers who support their platform and lower cost of content distribution.

Technology: Content aggregators do not have a unique technological strength and don't use technology to differentiate themselves. Their primary strength lies in the relationship they have with multiple content producers and video distribution companies.

Influence: Content aggregators have little influence on the mobile content value chain unless and until they are integrated with service providers who own the networks which is used to relay content to end consumers. In absence of such integration the content aggregators have little or no influence since they neither own the content nor the networks over which the content is delivered.

The primary strength of the content aggregators is their knowledge of how to bundle different sources into an attractive package. They build relationships with multiple content owners for this and provide a service to content producers.

Opportunities: Content aggregators currently can only transmit to consumers through networks owned by big mobile operators and thus cannot completely leverage their customer relationship. The solution we propose might allow the content aggregators to launch a service in partnership with smaller Wi-Fi hotspot operators and provide a OTT service where they would own part of the network. This will need them to develop new technology skills and relationships with other vendors but will provide them the opportunity to become equal players along with the mobile network operators in the mobile content value chain.

4.2.3 Video Distribution Network

The video distribution network is typically a content distribution networks owned by a private company which rents it out to content aggregators to transmit video content. The network implemented by the CDN Company typically is more intelligent than ordinary networks. Along with packet transport it provides caching, intelligent delivery of content and fault tolerance to the video aggregators streaming service.

Need: Video distribution network operators need to increase the utilization of their network and be able to deliver video from the edge of their networks or as close to the edge as possible. They also need to develop relationships with content aggregators who use their networks to deliver video to consumers.

Technology: Video Distribution Network operators have great technology skills in distributing video to different geographic regions at a very low cost. They have the capability to build and maintain such networks and are essential in being able to support the new streaming architecture we propose. They need to add intelligence to their network to be able to deliver multiple layers of video.

Influence: Video Distribution Network operators suffer from not having a direct relationship with the customers and though their technology plays a critical role in mobile content delivery they have a limited influence in the value chain.

Opportunities: No new opportunities will become available to the video network distribution operator in terms of new business arrangements or relationships with new stakeholders.

4.2.4 Cellular Network Operator

The cellular network operators provide last mile connectivity using a cellular network. They own the spectrum and edge infrastructure and mobile license. They also typically run a service provider business which provides billing and other services to the end consumer. If they are not integrated with service provider network operators provide their network access to various entities like service providers, private corporations on a wholesale basis where they charge for the traffic passing through their networks.

Need: The cellular network operator needs to increase the number of users attaching to the network and consuming content from their network. They are currently faced with increasing traffic and falling revenue per GB and need to find a way to reduce the cost of infrastructure to provide service to their customers.

Technology: The cellular network operators have capabilities in building wireless and edge networks. They also have capabilities to maintain these networks and manage additions effectively. They will play an important role in the delivery of video as they own the last leg of the content delivery chain and it is the most constrained.

Influence: The cellular operators own the relationship with the customer if they are integrated with a service provider. They also own the spectrum over which the video is delivered. They either provide wholesale network access to service providers or billing services to individual customers. They don't have a relationship with content producers/owners who are an important part of the value chain.

Cellular operators when integrated with service providers own the relationship with the customers and have strong relationships with handset manufacturers/developers which allows them to exert a substantial influence over the mobile content value chain. Even when they act as just network operators their technology choices affect the viability of the solution.

Opportunities: The solution architecture we propose satisfies the most important need the operators currently have of reducing cost of data delivery. This presents the operators an opportunity to get in more customers who consume video content on their networks. They could also roll out their own Wi-Fi network and become intelligent participants in the video content delivery ecosystem by partnering with content providers.

4.2.5 Cellular Service Provider

Cellular Service Providers is a wireless communication service provider that does not own the network, spectrum or the mobile license. It typically gets these services from a mobile network operator on a wholesale basis and sells it along with other value added services to consumers. Such virtual service providers often target segments ignored by the traditional mobile network operator or try to service customers in niche areas with products customized for them.

The Service provider's needs are very similar to the mobile network operator's needs. They don't possess the expertise to manage or roll out networks and are typically focused on providing superior customer service or niche solutions to specific market areas. They are bound by the infrastructure support that the mobile network operators can provide them and sometimes because of their small size exert very little influence on the mobile value chain.

4.2.6 Wi-Fi network operator and hotspot service provider

A Wi-Fi network operator and hotspot service provider owns the Wi-Fi infrastructure and the customer relationship. It provides Wi-Fi access in public hotspots for users in return for a fee.

Need: The Wi-Fi network operator and hotspot service provider needs to increase the utilization of its network by increasing the number of users attaching to its network. They also need to make the process of logging on their network as simple and automatic as a cellular network.

Technology: The Wi-Fi network operator and service provider have the technology capability to rollout Wi-Fi networks and manage the backend infrastructure. They also have developer knowledge of billing systems.

Influence: Most Wi-Fi network operators are relatively small and don't have relationships with any other players except the end consumers in the mobile value chain and thus exert limited influence on the value chain. There are some major Wi-Fi players who can exert considerable influence and perhaps be implementers for this solution if they can build relationships with handset manufacturers and content providers in the value chain.

Opportunity: They have a great opportunity to partner with existing cellular operators as they try to offload traffic on to Wi-Fi networks. This increases the utilization of their network and reduces the cost of sending data for the cellular operator and is great option for most Wi-Fi network operators.

4.2.7 Handset Manufacturers/Developers

Handset Manufacturers manufacture the devices and develop the basic software on the end devices used by consumers.

Need: Handset manufacturers need to sell as many handsets as possible and contribute towards improving the user experience. They also need to develop software which supports the widest possible formats for content consumption.

Technology: Handset manufacturers primarily build handsets and building consumer electronic devices is their primary technological skill. Most handset manufacturers also customize either a publicly available Operating system or build a custom Operating system for their devices.

Influence: Handset manufacturers directly interact with customers and have a relationship with most operators. The requirements for the device are driven by the operator and in many cases handset makes have little or no influence on the mobile value chain. This is changing with handsets like iPhone where the company that makes the handset also provides the content.

Opportunities: The solution architecture we propose provides the handsets a new opportunity to become an even more integral part of the mobile content delivery chain as it requires custom software to be built on the handset.

4.2.8 Consumers

Consumers in the value chain are individuals who want to view video on their handsets and pay the operator and the content aggregator a fixed fee.

Need: Consumers want to be able to watch video with the best QOE (Quality of experience) at the lowest price that can be offered to them. They are not particular about which network the video is served over as long as it is simple to use and doesn't result in worse quality than existing systems.

Technology: Consumers don't have technical skills to contribute towards developing the solution architecture and typically don't want to be exposed to the technical details of how video gets delivered to their handsets.

Influence: Consumers have tremendous influence on the mobile content value chain and are the drivers of demand. Any solution architecture has to involve a close relationship with the consumers.

Opportunity: The new solution architecture proposed provides an opportunity for consumers to lower their online entertainment expenses by reducing their cellular bills.

4.3 Conclusion

The table below summarizes the different parameters we evaluated each of the stakeholders on with respect to their importance and relevance to the solution architecture we propose. Need in the table indicates the relevance of the problems the solution addresses to the needs of each stakeholder. Technology indicates the technology capability of the stakeholder to actually implement the solution. The influence is the capability of the stakeholder to influence choices made in the mobile value chain. Opportunity indicates the new opportunities that would be created with the proposed solution architecture as it would directly map to the desire of the stakeholder to implement the solution.

	Need	Technology	Influence	Opportunity
Content Producer	Weak	Weak	Strong	Medium
Content Aggregator	Weak	Weak	Weak	Weak
Video Distribution Networks	Weak	Strong	Medium	Weak
Mobile Network Operators	Strong	Strong	Strong	Strong
Wi-Fi network Operators	Strong	Strong	Weak	Strong
Handset Manufacturers	Weak	Weak	Strong	Weak
Consumers	Strong	Weak	Strong	Weak

Table 3: Relative scoring of different stakeholders

The table above shows that the mobile network operators have the strongest need in that they are faced with increasing traffic and dropping revenue/GB. A solution which offloads a large percentage of their traffic onto Wi-Fi networks saves them tremendous costs in infrastructure while still keeping the users tied to their networks. They also have the technology to build the network infrastructure required to implement the solution and the relationships with handset vendors to co-ordinate custom software development on the handset required to implement the solution. They own the relationship with the consumers and have a very strong influence on the value chain. This solution architecture also allows the operator an opportunity to reduce the price of video delivery and perhaps get more customers on its network where they can be offered other services. What the operators currently lack is a relationship with content producers or aggregators which will allow them access to content and hopefully implement an end-to-end solution.

5 Economics of Wi-Fi Offload

The previous chapters discussed the solution architecture and the stakeholders involved in the proposed solution. We have identified the different motivations and advantages for each of the stakeholders in the system. We still need to understand the economics of delivering data over Wi-Fi which we will investigate in this chapter.

In this chapter we will build the model for investigating the economics of Wi-Fi offload and try to model the CAPEX/Mbps provided by these technologies. We will compare cases where Wi-Fi is used to provide continuous coverage in high demand density areas like transport hubs, shopping malls, urban town centers, stadiums and point coverage in the areas as well where there are localized small areas of high density.

5.1 Model Assumptions

We have built the model assuming LTE and Wi-Fi coverage as options and in this section we will discuss the assumptions for each of the technologies associated with the solution.

5.1.1 LTE assumptions

For the purpose of this model we assumed a LTE base station includes three sectors in addition to antennas. We assume for comparison purposes that a typical coverage area of 1KM2 giving us a radius of a circular cell to be 600 meters. The total capacity of the cell site is assumed to be 35 Mbps per sector. Each node costs $50,000 and needs $1,000 of labor to install. The controller and base station costs $50,000 again with an installation cost of $1,000. Pole to install the radio costs approximately $100,000 with identical labor cost. Additional costs include wiring costs of approximately $20,000 and site acquisition costs of $22,000. Adding up all the components gives a total cost of $233,000 in CAPEX for coverage of 1KM2.

5.1.2 Wi-Fi Assumptions

For the purpose of this model we assume a Wi-Fi AP which has 75 meters range which gives us a coverage area of 0.017 KM2. The total capacity for the Wi-Fi access Point is 100 Mbps. The cost of each node is assumed to be $1000 with an installation cost of $500. The actual smart controller for a Wi-Fi point is assumed to cost $9000. The pole on which the radio is installed costs $3000. We cover a considerably smaller area with the Wi-Fi access point and thus have a smaller pole. The wiring and the site acquisition costs come to around $3100 dollars. The typical Wi-Fi system needs one controller coupled with around 7 Wi-Fi access points which gives us a total cost of $65,700 giving us a capacity of 100 Mbps and coverage over an area of 0.017KM2.

The table below summarizes all the assumed costs in the model.

Technology	LTE	Wi-Fi
Radius (m)	600	75
Capacity (Mbps)	35	100
Node		
Cost	$ 50,000	$ 1,000
Install labor cost	$ 1,000	$ 500
Controller/base station		
Cost	$ 50,000	$ 9,000
Nodes per controller / base station	1	100
Redundant controllers (1=no, 1=yes)	1	2
Install labor cost	$ 1,000	$ 500
Pole		

Cost	$ 100,000	$ 3,000
Install labor cost	$ 1,000	$ 500
Wiring		
Cost (per meter)	$ 20	$ 2
Distance to each node	1,000	300
Site		
Acquisiton cost (CapEx)	$ 5,000	$ 1,000
Survey cost (CapEx)	$ 5,000	$ 1,000
Real estate (OpEx)	$ 12,000	$ 500
Total Capex	233000	$ 65,700

Table 4: Model Inputs

5.2 CAPEX/Mbps

Before we look at demand density and total coverage by Wi-Fi scenarios let us take a quick look at CAPEX/Mbps cost comparisons for Wi-Fi and cellular for point coverage. Point coverage is sufficient in small areas with very high demand like shopping malls, coffee shops and transportation hubs. Most of these places provide a Wi-Fi network connection but attaching to the network is not seamless. Our solution addresses this gap and makes Wi-Fi and cellular combination possible in these areas. The table below shows the CAPEX/Mbps for Wi-Fi and LTE.

	Wi-Fi	**LTE**
CAPEX/Mbps	$6657	$657

Table 5: CAPEX/Mbps

As we can see from the table above the CAPEX to provision access using Wi-Fi networks is almost 1/10 the cost of cellular access. This means we should ideally be moving as much traffic on the Wi-Fi network whenever one is available and the user is not mobile. But we have to consider the cost of contiguous coverage which the next section addresses using demand density as the factor compare the cost effectiveness of Wi-Fi versus cellular.

5.3 CAPEX versus demand density

Let us now look at the relationship between demand density and the CAPEX required for each of the technologies. To calculate the let us look at the Mbps/m^2 available for different densities given a base case of 35Mbps/KM^2. The table below shows the different areas we can serve for different demand per m^2.

Technology	**Mbps/m2**							
	100	10	1	0.1	0.01	0.001	0.0001	0.00001
LTE Macro	0.35	3.5	35	350	3500	35000	350000	1130973.35
Wi-Fi	1	10	100	1000	10000	17671.4	17671.4	17671.4

Table 6: Area Covered versus Demand Density

As we can see from the table above as the demand density increases the area effectively served by the area effectively served by the respective technologies. We will use this data to calculate the CAPEX/m^2 required for each of the technologies. The table below shows the values of CAPEX/m^2 for both LTE and

Technology	**Mbps/m^2**							
	100	10	1	0.1	0.01	0.001	0.0001	0.00001
LTE Macro	$665,714	$66,571	$6,657	$665.71	$66.57	$6.66	$0.67	$0.21
Wi-Fi	$65,700	$6,570	$657.0	$65.70	$6.57	$3.72	$3.72	$3.72

Table 7: CAPEX/m2

Again from the table above we can see that the LTE macro technology is cheaper when the demand density is lower than 1Kbps but Wi-Fi becomes increasingly attractive when the demand density increases beyond 1Kbps. This is because the Wi-Fi AP provides point coverage and can thus be more effective in regions of high traffic density.

5.3.1 Scenarios

A typical video streaming session requires 900 Kbps streaming HD video to a 4" screen. Given the overhead of SVC we add 10% to the original video stream which now have 990Kbps video stream. To be able to support 990 Kbps session we can either stream the entire session on a cellular network or stream it entirely on a Wi-Fi network or on a combination of both Wi-Fi and cellular network. Let us look at the cost of each of these scenarios which are shown in the table below.

Scenario	**Access Networks**		**Total CAPEX**
	Wi-Fi	**LTE**	
990 Kbps on LTE (Scenario 1)	$0	0.99*$6657 = $6590	$6590
990Kbps on Wi-Fi (Scenario 2)	$657 * 0.99 = $650	0	$650
792Kbps on Wi-Fi + 198Kbps Cellular (Scenario 3)	$657*0.792	0.198*$6657	$1838

Table 8: Video Streaming Cost

We can clearly see from Table 8 that in the Wi-Fi only scenario the CAPEX for delivering video over a Wi-Fi is only $650. This is the cheapest option but suffers from coverage limitation. The solution works for a static user who is already in the coverage area of the particular AP but won't work for a nomadic user who will move out of the coverage area and lose connectivity. The resulting jitter as the application switches from Wi-Fi to cellular results in a reduced quality of experience. On the other end of the spectrum cellular provides wide area coverage and a defined handoff mechanism which will result in a good QOE but is almost 10x ($6590) more expensive than the Wi-Fi solution.

The architecture we have proposed splits the video stream across both cellular and Wi-Fi channels using the SVC codec and delivers only 20% of the data on the cellular network while delivering the rest on Wi-Fi. This results in a net CAPEX cost for providing capacity for a single video session of only $1838 which is approximately 27% of the cellular only option but gives comparable QOE for the user.

5.4 Conclusion

We can clearly see from the two sections above that Wi-Fi works better when the area coverage required is very small; that is we are providing point coverage and the demand density as measured Mbps/m^2. Table 8 also clearly helps us identify that in a perfect coverage scenario the cost of providing a capacity only $1838 given our solution with a comparable QOE to existing solutions.

The solution we propose does not lend itself for implementation at home or at work places where the Wi-Fi coverage is excellent and the user is static. In such situations the cost calculation shown in Table 8 would clearly dictate the obvious choice to be a pure Wi-Fi network used for video delivery. The solution we propose also doesn't lend itself really well for a fast moving user who might move out of coverage areas very fast and thus would never get a stable Wi-Fi connection. Trying to cover areas with sparse demand is also not cost effective as seen from Table 7. Fig 4 highlights the different places where video is typically consumed and we can clearly see that about 58% of the video is consumed at home or at work where the Wi-Fi connection is very stable and our solution is not appealing. But around 21% of video is consumed while waiting in line or in between activities where the user is mobile and is covered by a public Wi-Fi network. By making authentication and handoff easier our solution makes it much easier to attach to a Wi-Fi network. These are the areas as well which typically have high demand density. In such cases the cost of providing capacity for a video streaming session can be reduced using our solution. Fig 5 shows the trend of increasing number of public Wi-Fi hotspots which will make a Wi-Fi network in such places more frequently available than today. If we can offload 50% of these sessions we can reduce the pressure on the cellular network in these areas and also reduce the cost of streaming video for the operators.

Based on the economic analysis, trends in growth of mobile data, growth in number of hotspots and the location based pattern of video consumption. We can easily conclude that the solution offers cost savings in scenarios where there is localized high demand which can be addressed by a Wi-Fi network and the video streaming consumers are not moving very fast and are not in areas with familiar and stable Wi-Fi networks. As the number of hotspots and the coverage of Wi-Fi networks increases this solution becomes even more appealing as it can address an even greater number of users and reduce the cost of supporting video streaming for the operator.

Bibliography

802.11u bootstrap procedure with 802.21. (n.d.). Retrieved January 12, 2012, from http://www.docstoc.com/docs/55700989/80211u-Bootstrap-Procedure-with-80221

America, N., Kong, H., Zealand, N., & Korea, S. (2011). Tellabs ¯End of Profit _study executive summary, (January).

Arkko, J., & Haverinen, H. (n.d.). Extensible Authentication Protocol Method for 3rd Generation Authentication and Key Agreement (EAP-AKA). Retrieved from http://tools.ietf.org/html/rfc4187#page-9

Beaulieu, C. (2011). Scalable Video Coding (SVC) for multipath video streaming over Video Distribution Networks (VDN), 206-211.

Choi, H., Won, J., Member, K., & Kim, J.-gon. (2007). Dynamic and Interoperable Adaptation of SVC for QoS-Enabled Streaming, *53*(2), 384-389.

Develder, C., Lambert, P., Lancker, W. V., Moens, S., Walle, R. V. D., Nelis, J., Verslype, D., et al. (2012). Delivering scalable video with QoS to the home, 129-148. doi:10.1007/s11235-010-9358-3

Diego, S. (2011). Future of Heterogeneous Networks – Opportunities and Challenges through insights that anticipate the future , creative new offers , and building business ecosystems, (December).

Houston, E. T. (2011). Wi-Fi Wholesale Service Opportunity Presentation to Executive Team Today we will discuss why CCI should become a Wi-Fi wholesaler and the economics of an offer, (September).

Jeong, J. (n.d.). Statistical Multiplexing using Scalable Video Coding for Layered Multicast. *Engineering*.

Kim, B.-K., Jung, Y.-C., Kim, I., & Kim, Y.-T. (2008). Enhanced FMIPv4 Horizontal Handover with Minimized Channel Scanning Time Based on Media Independent Handover (MIH). *NOMS Workshops 2008 - IEEE Network Operations and Management Symposium Workshops*, 52-55. Ieee. doi:10.1109/NOMSW.2007.11

Lehr, W. H., & Chapin, J. M. (n.d.). On the Convergence of Wired and Wireless Access Network Architectures, 1-28.

Lopez, Y., & Robert, E. (2010). Handover Implementation and Its Application to Proactive pre-Authentication, 14-25.

Man, L. A. N., Committee, S., & Computer, I. (2009). IEEE Standard for Local and metropolitan area networks--Media Independent Handover Services, (January).

Man, L. A. N., Committee, S., & Computer, I. (2011). *Part 11 : Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 9 : Interworking with External Networks IEEE Computer Society Sponsored by the. Networks* (Vol. 2011).

Nightingale, J., Member, S., Wang, Q., Grecos, C., & Member, S. (2010). Optimised Transmission of H . 264 Scalable Video Streams over Multiple Paths in Mobile Networks, 2161-2169.

Palau, C. E., Mares, J., Molina, B., & Esteve, M. (2010). Wireless CDN video streaming architecture for IPTV. *Multimedia Tools and Applications*, *53*(3), 591-613. doi:10.1007/s11042-010-0516-0

Piri, E., & Pentikousis, K. (2010). Towards a GNU / Linux IEEE 802 . 21 Implementation. *Design*.

RATE CONTROL AND STREAM ADAPTATION FOR SCALABLE VIDEO STREAMING OVER MULTIPLE ACCESS NETWORKS Deutsche Telekom R & D Laboratories USA 5050 El Camino Real 221 Los Altos , CA 94022 Cisco Systems , Inc . 425 East Tasman Drive San Jose , CA 95134. (n.d.)., (i).

Ramzan, N., & Izquierdo, E. (2011). Scalable and Adaptable Media Coding Techniques for Future Internet During the last decade a noteworthy amount of research has been devoted to scalable, 381-389.

Sarkar, C., Rein, S., & Wolisz, A. (2011). Scalable Video Streaming with Utilization of Multiple Radio Interfaces : A Customized Method for Signaling and Bandwidth Estimation ∗, 829-833.

Zinner, T., Abboud, O., Hohfeld, O., Tran-gia, P., Zinner, T., Abboud, O., Hohlfeld, O., et al. (2010). Towards QoE Management for Scalable Video Streaming, (March).

Schierl, T., Hellge, C., Mirta, S., Grüneberg, K., & Wiegand, T. (2007). Using H . 264 / AVC-based Scalable Video Coding (SVC) for Real Time Streaming in Wireless IP Networks, 3455-3458.

Schierl, T., Stockhammer, T., & Wiegand, T. (2007). Scalable Video Coding. *IEEE Transactions on Circuits and Systems*, *17*(9), 1204-1217.

Schwarz, H., Marpe, D., & Wiegand, T. (2007). Overview of the Scalable Video Coding Extension of the H.264/AVC Standard. *IEEE Transactions on Circuits and Systems for Video Technology*, *17*(9), 1103-1120. doi:10.1109/TCSVT.2007.905532

Std, I., Committee, M. A. N., Society, I. C., & Board, I.-sa S. (2007). *IEEE Standard for Information Technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements Part 11 : Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications* (Vol. 2007).

Sun, H., Vetro, A., & Xin, J. (2007). An overview of scalable video streaming. *Wireless Communications and Mobile Computing*, *7*(2), 159-172. doi:10.1002/wcm.471

Taleb, T., Nakamura, T., & Hashimoto, K. (2008). On Supporting Handoff Management for Multi-Source Video Streaming in Mobile Communication Systems.

Tavares, J. (2011). Service Provider Wi-Fi.

Technology, N.-generation. (2006). A Ready Market: Introducing H.264-SVC.

Tellab End Of Profit . (n.d.). Retrieved April 28, 2012, from http://www.tellabs.com/markets/tlab_end-of-profit_study.pdf

Vogeleer, K. D., Ickin, S., & Erman, D. (n.d.). A decentralized Information Service for IEEE 802 . 21 - Media Independent Handover (MIH).

Wenger, S., Wang, Y.-kui, & Schierl, T. (2007). Transport and Signaling of SVC in IP Networks, *17*(9), 1164-1173.

Wien, M., Cazoulat, R., Graffunder, a., Hutter, a., & Amon, P. (2007). Real-Time System for Adaptive Video Streaming Based on SVC. *IEEE Transactions on Circuits and Systems for Video Technology*, *17*(9), 1227-1237. doi:10.1109/TCSVT.2007.905519

Wierenga, K. (2009). 802 . 11u Executive Summary. *Networks*, 1-10.

Wu, M.-hong, Chen, Y.-mu, Chung, T.-yaw, & Hsu, C.-hung. (n.d.). A Profile-Based Network Selection with MIH Information Service User Profile-Based Network Selection.

Yoo, S.-J., Cypher, D., & Golmie, N. (2008). Timely Effective Handover Mechanism in Heterogeneous Wireless Networks. *Wireless Personal Communications*, *52*(3), 449-475. doi:10.1007/s11277-008-9633-8

Zhang, Y., Liu, Y., Xia, Y., & Huang, Q. (2007). LeapFrog: Fast, Timely WiFi Handoff. *IEEE GLOBECOM 2007-2007 IEEE Global Telecommunications Conference*, 5170-5174. Ieee. doi:10.1109/GLOCOM.2007.980

cisco visual networking index: global mobile data traffic forecast update, 2011–2016&nbsp; [visual networking index]. (n.d.). Retrieved April 18, 2012, from http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html

ieee 802.21 - the internet protocol journal, volume 12, no.2. (n.d.-a). Retrieved April 27, 2012, a from http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_12-2/122_ieee.html