Dimensioning IPTV VoD Services

Dimensioning IPTV VoD Services

By

Suman Pandey

Department of Computer Science and Engineering.
POSTECH

A thesis submitted to the faculty of POSTECH in partial fulfillment of the requirements for the degree of Master of Engineering in the Department of Computer Science and Engineering.

Pohang, Korea
May 27, 2009

Approved by

____

Major Advisor: James Won-Ki Hong

Dimensioning IPTV VoD Services

수먼 판데

위 논문은 포항공과대학교 컴퓨터공학과 석사 학위논문으로 학위논문 심사위원회를 통과하였음을 인정합니다.

2009년 5월 27일

학위논문심사 위원회 위원장 홍원기 (인)

위 원 서영주 (인)

위 원 송황준 (인)

DCSE 20079247 Suman Pandey, “A Methodology of Deployment Decision Making for IPTV VoD Services”, Department of Computer Science and Engineering, 2009, 41P, Advisor: James Won-Ki Hong, Text in English.

ABSTRACT

The IPTV subscription rate has increased steadily since the introduction of IPTV Video on Demand (VoD) services. We have developed a model to determine the optimum deployment strategy for IPTV delivery network from the IPTV service providers' perspective. The analysis technique in this paper helps us determine the best deployment scenarios to support a certain number of customers within a tolerant boundary of Quality of Experience (QoE) measures. The main QoE measures considered are server waiting time, one way minimum delay, and access network (AN) bandwidth consumption. All these models are made considering heterogeneous network conditions. The QoE models will help service providers make network deployment decisions including the number of servers, distance of servers from a community, and desirable AN bandwidth capacity. The methods of determining these network parameters and deployment strategy based on QoE measures proposed in this paper are very useful for the IPTV VoD service providers.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Overview

IPTV is a system capable of receiving and displaying a video stream using Internet Protocol. International Telecommunication Union (ITUT) defines IPTV as multimedia streaming over IP networks with reasonable quality assurance [1]. Triple play services are new windows of opportunity for telcos and cable providers. Some IPTV service providers such as AT&T and Verizon own full networking infrastructure. Others such as AOL, Apple and Google need not necessarily own the full infrastructure; they might lease some part of the infrastructure or rely on the current Internet. Software products such as Microsoft Mediaroom [2] are used as the IPTV service provisioning platform by these telcos to provide IPTV services.

IPTV video services can be broadly classified as live TV and video on demand (VoD). In the future, up to 90%, of IPTV traffic, may be due to VoD services [3], and VoD services are resource-intensive. However, the planning and deployment of VoD services with optimum cost has not been thoroughly studied.

This thesis focuses on determining the optimum network deployment strategy for VoD services in a heterogeneous networking environment. The purpose of dimensioning a network for VoD service is to determine the minimum capacity requirements that will still allow QoE requirements to be met. We define essential QoE measures, and build an analysis model for each QoE criterion.

1.2 Research Goal

The goal of this thesis is to provide an analysis method to enable potential IPTV VoD service providers to determine the optimum deployment conditions. While taking any deployment decision, the end user QoE must be considered. We can state the problem as follows: "*Using user QoE and user requirements help service providers determine the number of servers, access network bandwidth requirements, and distance from server to users to fulfill IPTV VoD service in a heterogeneous access network*".

The contributions of the thesis are the following:

- Our models will help service providers choose various network deployment strategies including bandwidth requirement in servers, and communities, and distance of server from community.
- Our QoE measures related to delay and network congestion cover major network performance aspects.
- The models are built considering heterogeneous network conditions.
- We developed comprehensive simulation models and analyzed the results.

1.3 Thesis Outline

The structure of the thesis is as follows:

Section 2 describes the standard IPTV architecture and classification of type of IPTV services and delivery mechanisms. This section also reviews related work in the field of network dimensioning of IPTV services. Section 3 describes our method and choice of QoE measures; it also provides each QoE description and presents the analysis model. Section 4 details the simulation scenario and the basic assumptions. Section 5 analyses the simulation results and their use in deployment decision making process. The concluding remarks and future work are given in Section 6.

2 Related Work

This section provides an overview of IPTV architecture and IPTV service classification. This also provides a survey about the studies done in dimensioning IPTV network in past.

2.1 IPTV Architecture

IPTVs have complex network architecture (Fig. 1) [3, 4]. A typical IPTV network is composed of a super headend (SHE) and video hub offices (VHOs) at the national network core. VHOs are connected to multiple video source servers (VSOs) which constitute the metropolitan area network (MAN) core, and VSOs are connected to central offices (COs) which are closer to the user communities. VSOs' edge routers route and aggregate local loop traffic from passive optical networks (PONs), or from digital subscriber line access multiplexers (DSLAMs). DSLAMs and PON support Gigabit Ethernet transport networks. There are video on demand (VoD) servers at each of these hierarchical levels. The contents are distributed over the VoD servers at each level in the hierarchy of VSO, VHO and CO, based on local preference and popularity, which usually follows a Zipf distribution [5]. At each level, advertisements are also added to the VoD contents. While making deployment decisions we need to consider how many servers should be deployed at each level of hierarchy, and what should be the server bandwidth capacity.

ANs are connected to COs and VSOs. ANs can have multiple outside plants (OSPs) connected to the CO. Each OSP may have a co-located DSLAM. These DSLAMs can be connected with remote DSLAMs which will be directly connected with the end user communities. Service providers need to analyze the AN bandwidth demand and make deployment decisions about DSLAMS or PON based on the bandwidth demand in the AN.

The IPTV infrastructure is logically divided into three main parts: client domain, network provider domain, and service provider domain [6]. IPTV service providers need to lease services from all of these domains; for example, they must obtain sustainable bandwidth within the client and network provider domains, and lease the optimum number of content servers. IPTV service providers lease services from multiple service providers, and provide the audio-video services to different kinds of users having different access network technologies including fiber to home (FTTH), fiber to building (FTTB), xDSL and cable networks. In such a heterogeneous environment, IPTV service providers need to consider heterogeneous network conditions while making optimum deployment strategies.

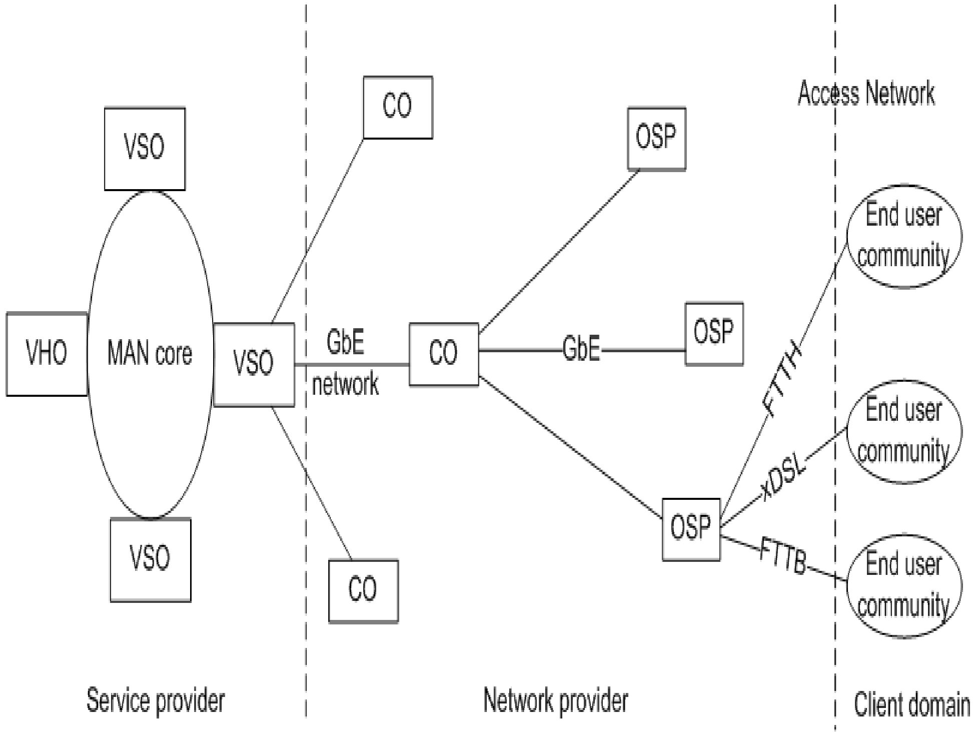

Figure 1. IPTV network architecture.

2.2 Network Dimensioning

Here we elaborate previous studies of deployment analysis in IPTV service provisioning. While dimensioning a network for IPTV services, the most important factors to consider are the type of services to deploy, and the protocols to use for delivery. We tried to analyze five IPTV service categories that fit well with the studies reported by previous researchers (Table 1). Several of these studies also explain network dimensioning aspects of deployment. The IPTV service can be categorize VoD and live

TV services and the delivery modes used to deliver these services, including unicast, P2P and multicast delivery mode. For this work, we concentrated on VoD services, and the unicast mode of delivery.

Table 1. Summary of types of IPTV services and delivery mechanisms. The IPTV services which are considered in the studies are marked √

Related work	IPTV Service Categories				
	Unicast VoD	**P2P VoD**	**Multicast VoD**	**Live TV**	**Deployment Analysis**
Agrawal et al. 2007 [6]	√			√	√
Wauters et al. 2005 [10]	√			√	√
El-Sayed et al. 2006 [4]	√				√
Thouin et al. 2007 [7]	√			√	√
Simsarian et al. 2007 [3]	√				
Han et al. 2008 [11]	√			√	√
Chen et al. 2009 [9]		√			√
Naor et al. 2007 [17]			√		
Ganjam et al. 2005 [18]			√		

Table 2. Studies related to deployment design.

Related work	Deployment Features				
	Heterogeneous	**Bandwidth Model**	**Delay Model**	**Decision**	**Output Measure**
Agrawal et al. 2007 [6]	No		Zapping delay	For over provisioned network	QoE
Wauters et al. 2005 [10]	No – Ethernet WDM	No	No	No	Cost of traffic and content distribution
El-Sayed et al. 2006 [4]	No – EoF and EoS	Yes	No	No	Comparison of two AN
Thouin et al. 2007 [7]	No	No	No	No	Hit ratio
Han et al. 2008 [11]	No	No	No	No	Comparison of L1 technology
Chen et al. 2009 [9]	No	Yes – P2P	No	No	P2P traffic benefits

We also summarized the studies related to deployment analysis. The deployment features targeted by these studies are categorized and the outputs of these studies are also listed (Table 2). The studies related to network cost models for VoD services [7, 8] have focused on video distribution strategies for reducing network cost. Their cost models are based on the hit ratio and the cache size of the storage devices. The benefits of a P2P delivery mechanism for IPTV VoD servers have been explored [9], but the study proposed bandwidth modeling only for P2P.

Mathematical models for IPTV network deployment [6, 7, 8] have been proposed; however, they do not consider the heterogeneous aspect of network, or actual deployment strategies. Furthermore, some studies concentrate only on some specific technology deployment such as Ethernet-based WDM networks with ring topology [10], Ethernet over SONET, and Ethernet over Fiber technology [4]. The advantages and disadvantages of various layer-1 transport network alternatives have also been studied [11]. However, a generic model of deploying IPTV VoD services is needed.

None of the previous studies concentrate on heterogeneous network conditions while modeling delay and bandwidth requirement. Previous studies also do not concentrate on the decision-making process. Although a few studies concentrated on decision making for over-provisioned networks, and ongoing deployment conditions [6], none discusses deployment decision-making before deploying a new service.

Network-centric quality has been analyzed and discussed from the perspective of customers using real world commercial VoD IPTV network traffic traces in various user scenarios, and a bandwidth demand estimation method has been proposed [12]. In this study, the authors develop mathematical models to dimension IPTV networks from a service providers' perspective.

3 Proposed Method

To solve our problem, we analyzed the QoE measures of interest including server waiting time, one-way minimum delay and access network bandwidth consumption. To capture the input metrics for modeling these QoE we defined the notion of community which will provide us with user behavior. The metrics obtained from the community, from network providers, and from service providers are passed to the analysis models, which calculate essential QoE measures, and help in analyzing the best deployment strategies (Fig. 2). All the QoE simulation models are independent of each other. The desired QoE values are provided in the Service Level Agreement (SLA) between the IPTV service provider and the subscriber. For example the SLA can define that server waiting time is < 0.5 s, bandwidth occupancy is < 0.6 and one way delay is < 0.004 s. Service providers should meet minimum QoE conditions when making any deployment decisions, and also upgrade their networks whenever the QoE is not satisfied. The following subsections will justify our choice of the QoE metrics and explain the detailed models for each.

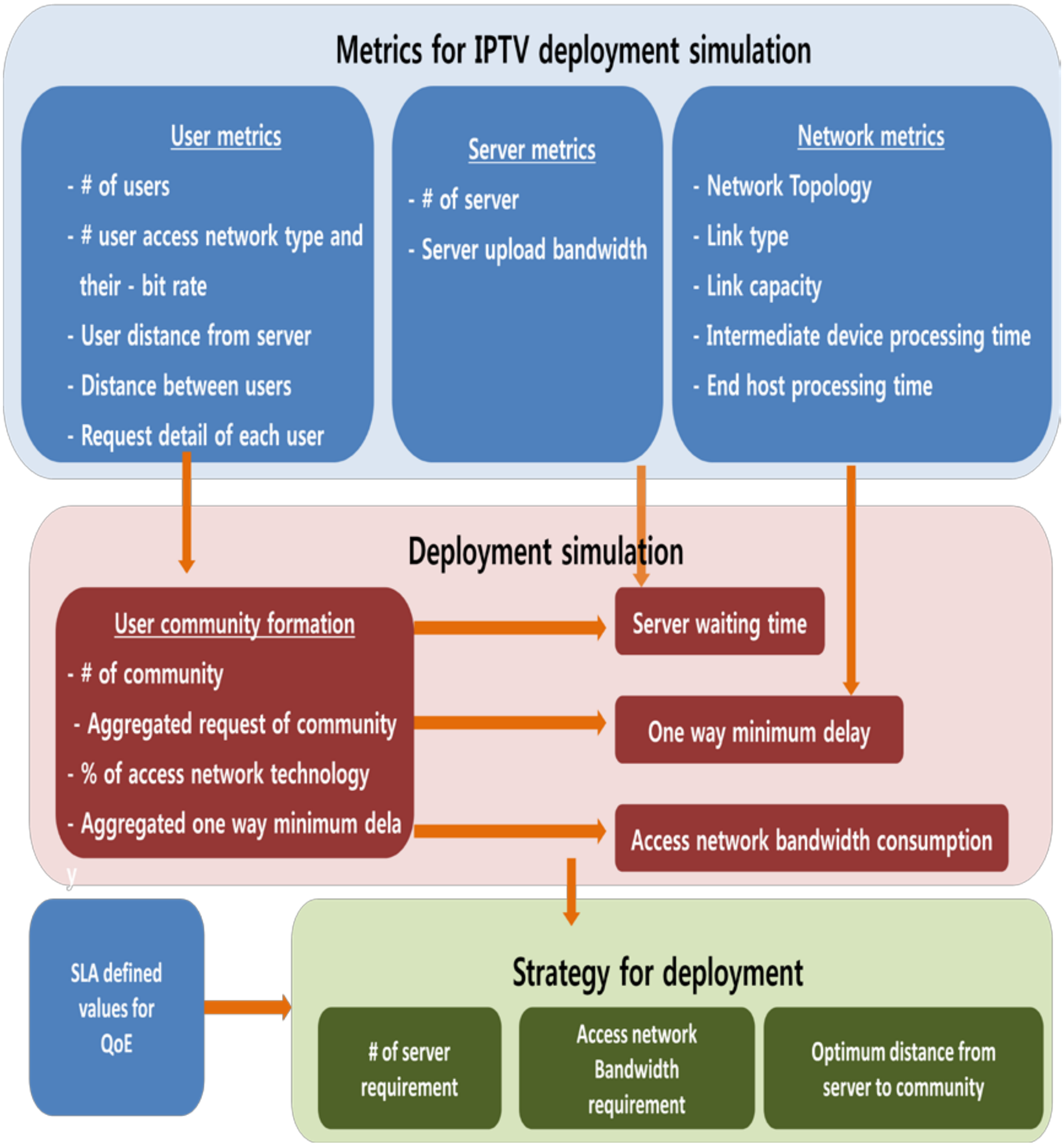

Figure 2. Overall design -Interaction of simulation metrics between the models and deployment optimization strategies.

3.1 Community Description

A community is a group of IPTV subscribers sharing similar characteristics, such as physical closeness, distance from the IPTV server, and type of network access technology. We can also model a community according to additional characteristics, such as viewing preferences and patterns. A clustering-based algorithm can be implemented to delimit a community. A community can be connected to one or more servers. Depending on the QoE measure, the community chooses the best servers for VoD delivery.

The community provides the following metrics to the IPTV service provider for modeling QoE:

- User viewing behavior
- Number of users in the community
- Mean link capacity
- Mean request rate during peak viewing hours
- Mean session duration
- Mean distance from the community to the server

The network and server properties and their related metrics are driven from the network service and storage providers (Fig. 3).

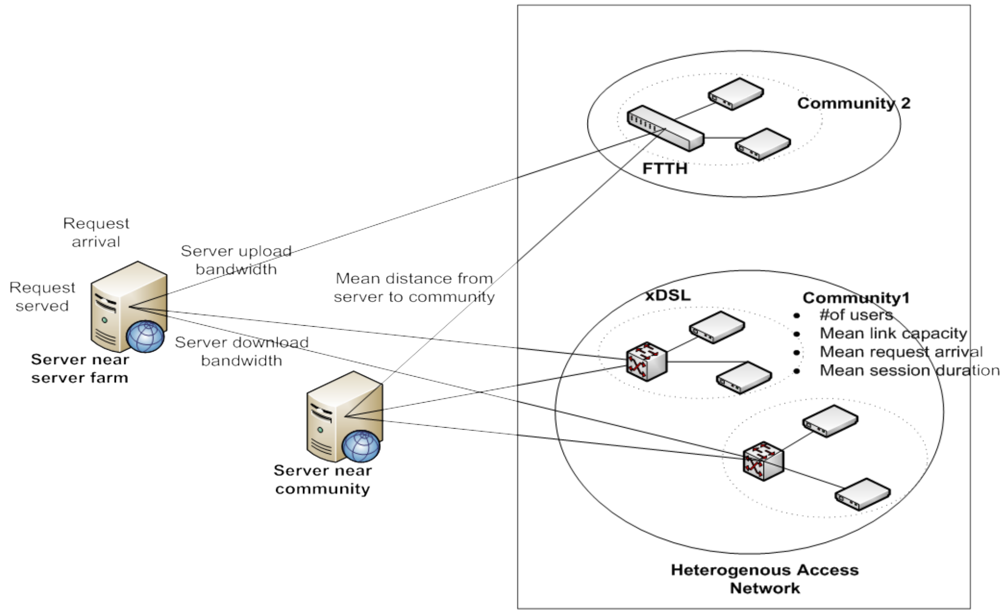

Figure 3. Metrics provided by community and network providers.

3.2 QoE Selection

Audio and video services are compressed and have low entropy, which is why the network performance highly affects users' QoE for such services. Thus, while designing or dimensioning a network for video traffic, QoE is an important measure. ITU-T [1] divides QoE into subjective QoE and objective QoE. Subjective QoE includes emotions, linguistic background, attitude, and motivation. Objective QoE includes information loss and delay. Objective QoE is actually the Quality of service (QoS) and is therefore a measure of network performance; it includes service factors, transport factors and application factors. According to ITU-T, given a QoS measurement, one could predict the expected QoE for a user, and given a target QoE for users, one could deduce the net required service layer performance. In this thesis when I refer to QoE I mean objective QoE. We utilize following objective QoE for this work and deduce a model for these QoE metrics.

- Server waiting time.
- One way minimum delay in non traffic hours.
- Access network bandwidth consumption.

These QoE metrics will give an overall network performance measure for VoD data. Server waiting time will give stochastic delay caused by the server. One way minimum delay will give the deterministic delay caused by the path from server to community. Access network bandwidth consumption will provide us with a measure of congestion in the access network. Together, these QoE will give the performance measure related to the server, path from the server to the community and the access network of the community.

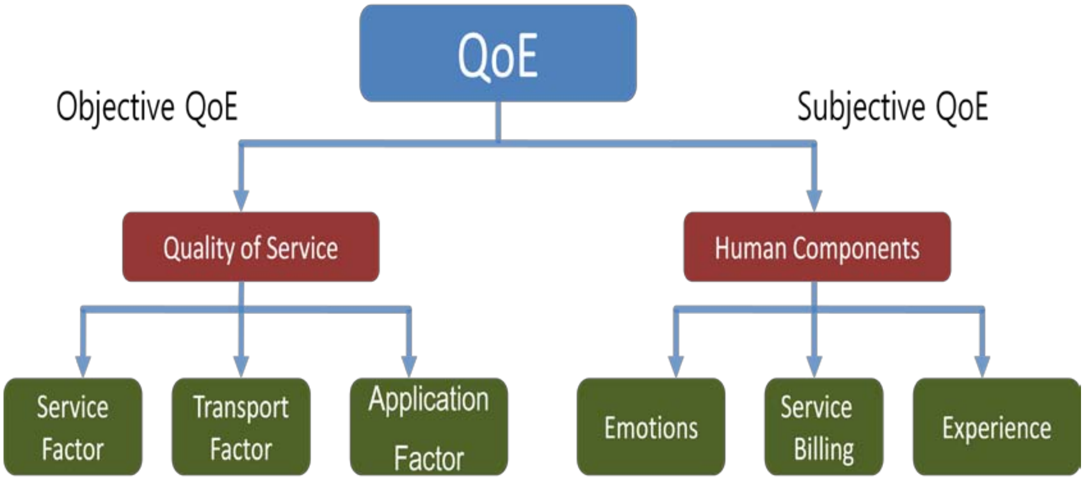

Figure 4. ITUT categorization of types of QoE measure.

3.3 Server Waiting Time Model

Server waiting delay is an important objective QoE measure. Communities generate service requests and servers fulfill these requests. As the load on a server increases the performance decreases; this causes delay in serving the requests. This delay is an important QoE metric for IPTV service providers to decide whether they need to increase the number of servers in the IPTV domain or to upgrade the server capacity. The server is modeled using queuing theory [6]. We utilized the Erlang C [13] model for modeling server waiting time. This model uses several variables (Table 3). We map the number of servers c in the Erlang C model to the upload bandwidth capacity of the server. A server can serve a number of simultaneous channels depending on its upload bandwidth capacity (Fig. 5). Each community connects with more than one server, and makes requests to the server that has the minimum load and that can serve with the minimum delay. The sum of the requests from all the communities to which a server is connected is called the mean request arrival λ. λ follows a Poisson distribution. The mean service time μ^{-1} follows an exponential distribution.

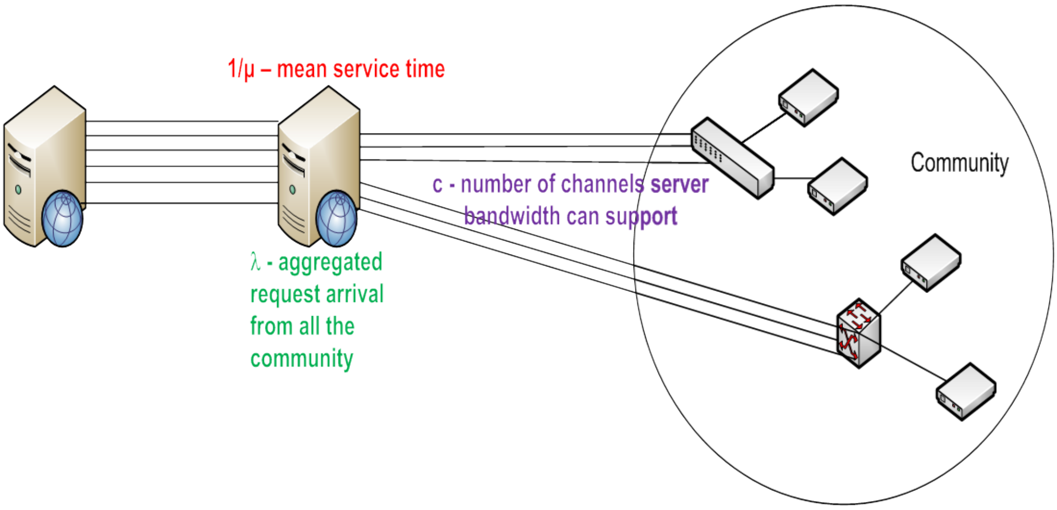

Figure 5. Server model and metrics.

Table 3. Parameters for server waiting time model.

Symbol	Description
λ	Mean request arrival from all communities (Poisson distribution)
μ^{-1}	Mean service time from all communities (exponential distribution)
c	Number of parallel identical servers (mapped to the server upload bandwidth capacity, if bandwidth capacity is higher number of server will be higher)

Using the Erlang C formula, the occupation rate ρ of a server is calculated as $\frac{\lambda}{c\mu}$. The value for ρ should always be less than one. If $\rho < 1$, the delay probability Π_w is:

$$\Pi_w = \frac{(c\rho)^c}{c!}\left((1-\rho)\sum_{n=1}^{c-1}\frac{(c\rho)^n}{n!} + \frac{(c\rho)^c}{c!}\right)^{-1}. \tag{1}$$

Using Little's law in queuing theory [13] the server's service waiting time *E(w)* is:

$$E(w) = \Pi_w \times \frac{1}{1-\rho} \times \frac{1}{c\mu}. \tag{2}$$

The overall delay is obtained by adding *E(w)* to the other delays such as serialization delay, propagation delay and constant delays such as encoding delay, packetization delay, and switching delay.

For server waiting time simulation, the aggregated requests from all the communities, mean request arrival of all the requests, and server capacity are passed to the server waiting time calculation module. The calculated value is compared with the predefined threshold value for server waiting time; if the calculated value is lesser than predefined QoE measure then we can deploy the server for the given requests, otherwise we increase the number of servers to match the threshold value (Fig. 6).

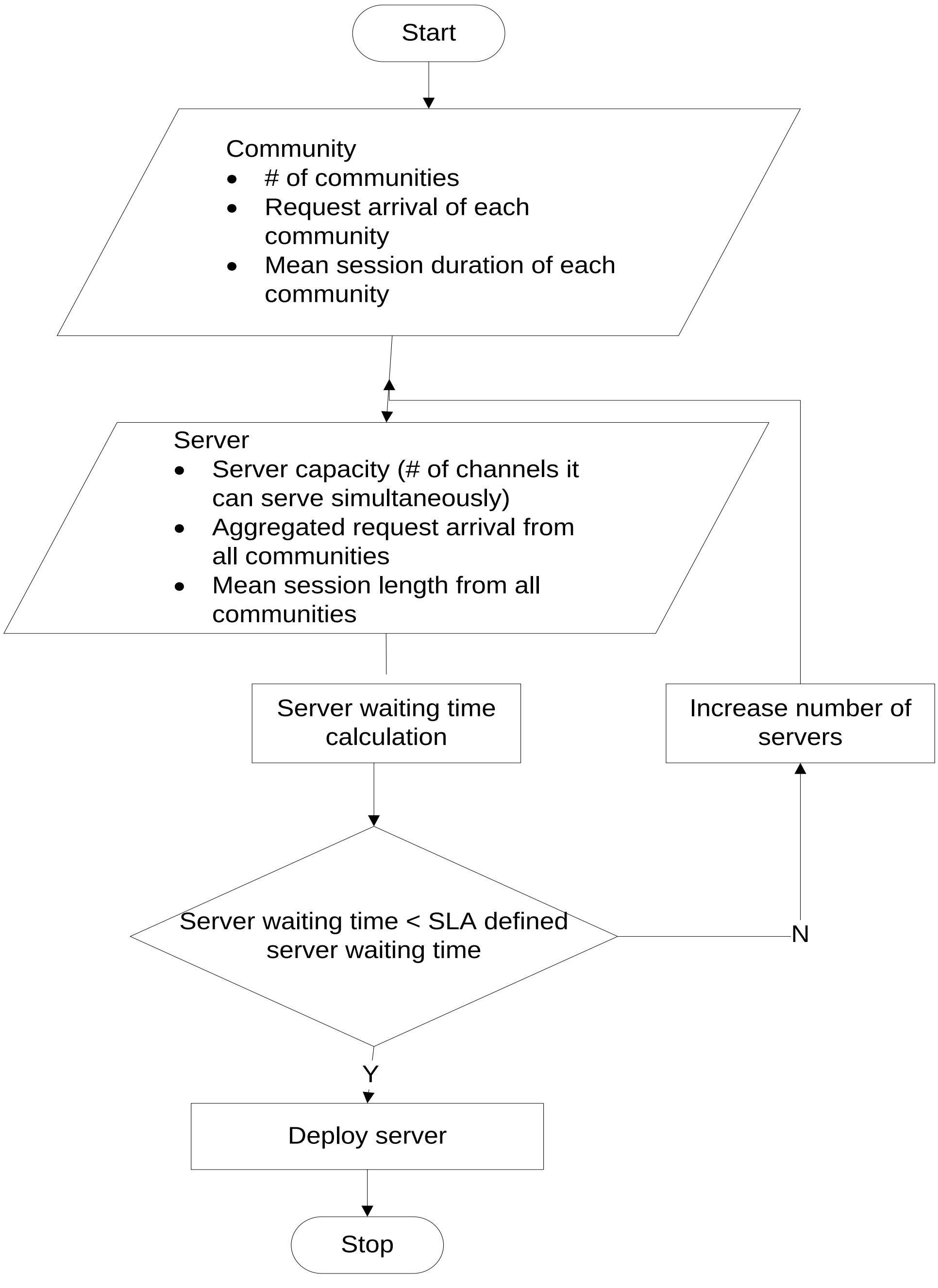

Figure 6. Flow diagram for server waiting time calculation.

3.4 One Way Minimum Delay Model

One-way minimum delay during low-traffic hours is an important QoE measure for VoD. For video download from the server we need to consider only one-way delay, to examine the time taken for a response to come from the server to the community. Low-traffic hours are considered to avoid the effect of network congestion. Later we can add this deterministic one way delay to the stochastic delay caused by congestion to obtain an overall delay. One way minimum delay can help determine the behavior of underlying network infrastructure. One way minimum delay for a path in the network will not change unless the underlying infrastructure changes, so we need not measure it frequently. This QoE measure can be obtained in two ways. The first way is to model one way delay, where we need various inputs from network providers (Table 4). The second way is to measure one-way delay using round trip time (RTT).

Table 4. Parameters for the one-way delay model.

Symbol	Description
N	Number of hops (intermediate device) to reach each server
D_h	Mean processing delay of each hop
D_e	Processing delay at end host
l	Types of links in the path
d_n	Physical length of link of type n
v_n	Propagation velocity of the link type n
$RTT(s,i)$	RTT measured by sender for packet i
$RTT(r,i)$	RTT measured by receiver of packet i

First, we discuss the method to model one-way delay. Propagation delay D_p depends on the physical length of the communication path d and the propagation velocity of the physical medium v. If the path uses different types of physical media l, then D_p is the sum of all propagation delays:

$$D_p = \sum_{n=1}^{l} \frac{d_n}{v_n} \ . \tag{3}$$

Apart from D_p the processing power of each intermediate node D_h and the end host D_e also cause delay. So the overall deterministic delay D can be calculated as:

$$D = \sum_{n=1}^{l} \frac{d_n}{v_n} + N * D_h + D_e \ . \tag{4}$$

Keeping the processing power of the intermediate nodes and end host constant, the one way minimum delay depends primarily on the distance from the server to the community. Now we discuss the method to measure one way delay RTT_{min} [14]. RTT_{min} of the nth packet from the server to the community is

$$RTT_{min\,(n)} = RTT_{min\,(0)} - \sum_{i=1}^{n} [RTT(s,i) - RTT(r,i)] \ . \tag{5}$$

Modeling one way delay requires many network inputs. Especially, modeling propagation delay can be very challenging if network topology is complex. Given the correct input parameters the values of D and RTT_{min} will be the same,

$$RTT_{min\,(n)} = D\ . \tag{6}$$

If D depends on the distance between the community and the server, then RTT_{min} also depends on this distance. If the distance between the community and the server increases, RTT_{min} also increases.

For one-way minimum delay the number of intermediate devices, their processing times, length of the path, percentage of communication link types, and end host processing delay are sent to the one way delay calculation module. The calculated value is then compared with the predefined threshold value for one-way minimum delay. If the calculated value is less than the predefined QoE measure then the deployment decisions are correct, otherwise service provider need to deploy the server closer to the community, or deploy better links or intermediate devices (Fig. 7).

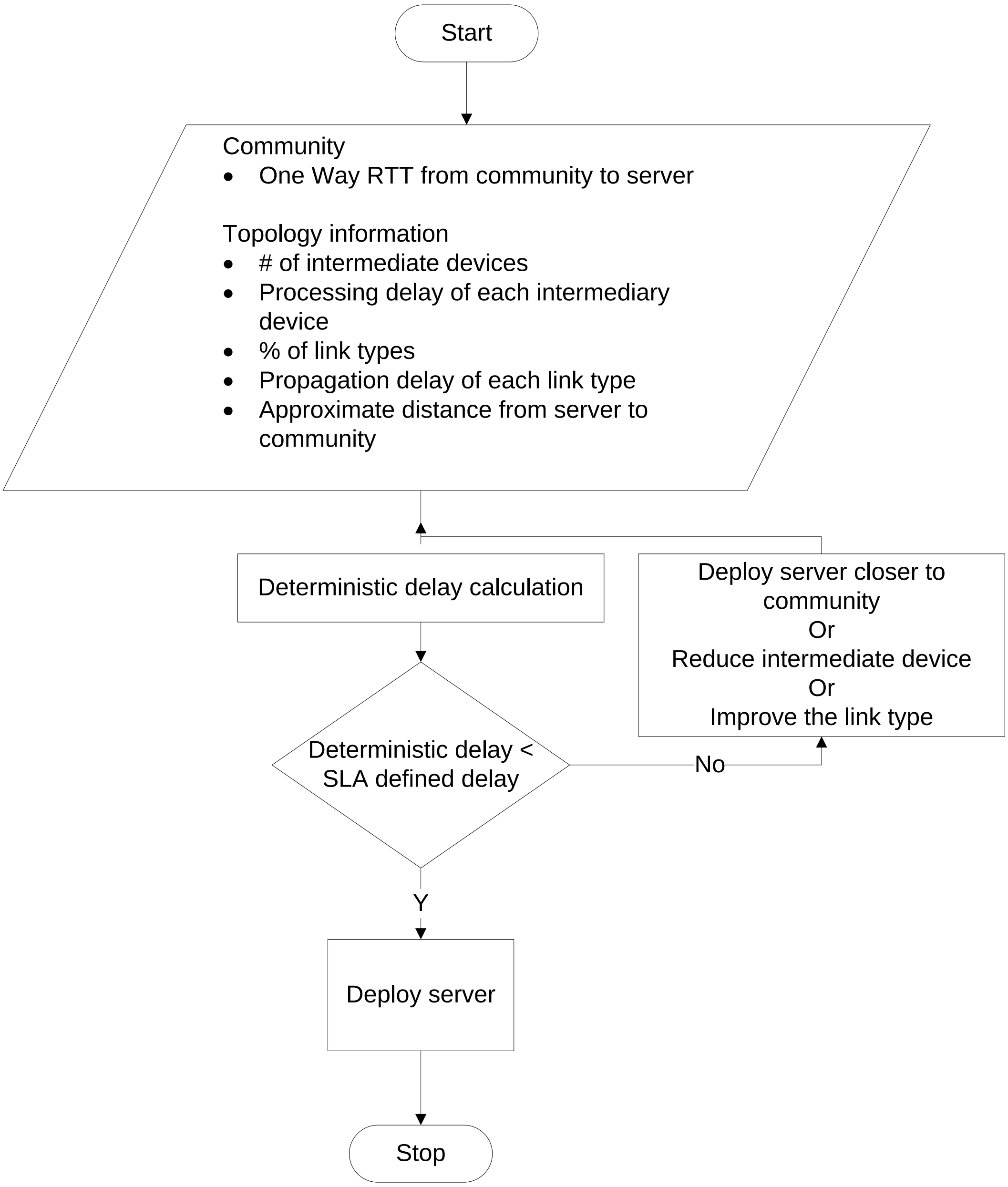

Figure 7. Flow chart for one way minimum delay calculation.

3.5 AN Bandwidth Consumption Model

The access network bandwidth consumption also affects user viewing quality. High bandwidth consumption causes an increase in latency and poor application performance at the end user; conversely, low bandwidth consumption means that IPTV service providers need to pay for unnecessary bandwidth. In an IPTV infrastructure, even if server performance is good and servers are placed very close to the community networks, the video quality degrades significantly if bandwidth consumption in a community network is too high. Therefore, we determined that a bandwidth consumption model is an important objective QoE measure.

Table 5. Parameters for access network bandwidth consumption model

Symbol	Descriptions
λ_a	Request arrival in access network every time tick (Poisson distribution)
μ_a	Request leaving in access network every time tick (exponential distribution)
br_l	Bit rate between VOD server and viewer via link *l*. the link can be any heterogeneous network such as xDSL, Cable, FTTH, FTTB
p_l	Percentage of generated traffic from different type of access link network *l*, such as xDSL, Cable, FTTH, FTTB
K	Number of access network types
B	Total outgoing bandwidth from access network to server

The bandwidth consumption model uses several variables (Table 5). If request arrival rate λ_a and bit rate of the video stream *br* are given, then the total traffic generated by the users in the community is $\lambda_a \times br$. Because we are considering a heterogeneous environment, the bit rate of each type of broadband network (e.g., FTTH, FTTB, xDSL or Cable) will be different. To accommodate this scenario, the IPTV service provider needs to know the bit rates of the multiple types of AN br_l, and the percentage of traffic p_l that each type of link will generate in an AN. This gives the total traffic generated by the users in heterogeneous environment as:

$$T(t)= \lambda_a \left(\sum_{l=1}^{k} br_l \times p_l \right). \quad (7)$$

Also some requests μ_a will be satisfied every time tick, so we need to subtract the traffic generated by those requests from the total traffic. This will give the total traffic generated by the AN *T (t)*:

$$T(t) = \lambda_a \left(\sum_{l=1}^{k} br_l \times p_l \right) - \mu_a \left(\sum_{l=1}^{k} br_l \times p_l \right). \quad (8)$$

The bandwidth consumption $B_{consump}(t)$ of the AN at time *t* is then derived by dividing *T(t)* by the total outgoing bandwidth *B* from AN to server.

$$B_{consump}\ (t) = \frac{\lambda_a \left(\sum_{l=1}^{k} br_l \times p_l \right) - \mu_a \left(\sum_{l=1}^{k} br_l \times p_l \right)}{B}. \quad (9)$$

For AN bandwidth consumption, the penetration ratio of different type of access technology, their bit rates for download, aggregated request arrival from all the communities in AN, and mean session duration are passed to the access network bandwidth consumption calculation module. If the calculated value is less than the predefined QoE measure then the deployment decisions are correct; otherwise service provider may need to increase the AN's aggregation switch bandwidth capacity (Fig. 8).

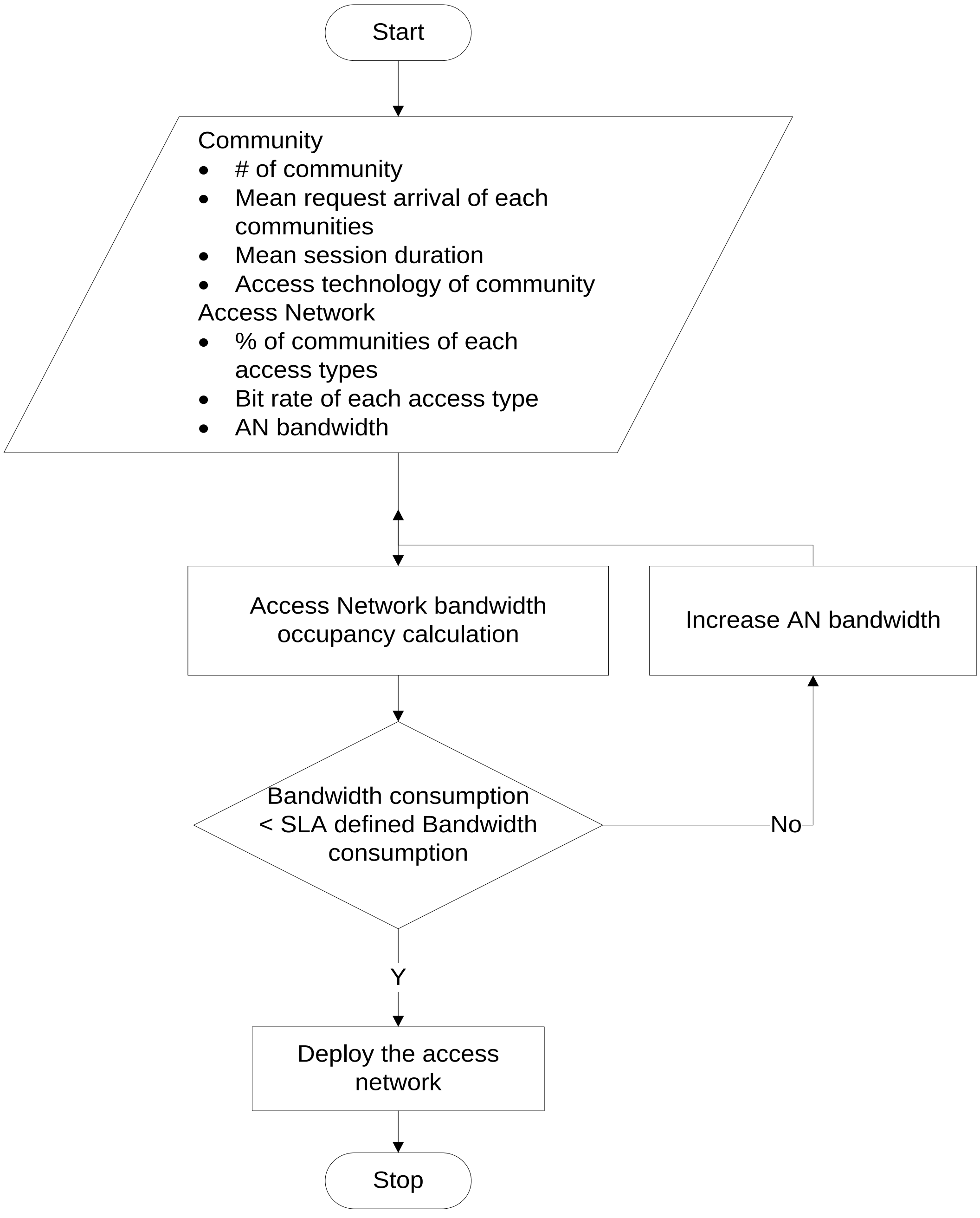

Figure 8. Flow chart for access network bandwidth consumption calculation.

4 Simulation Scenario: Assumptions and Preset Parameters

For simulation we used several preset parameters (Table 6). The details of various other assumptions related to content popularity, type of VoD service, user behavior, server properties and subscribers access technologies are explained in consequent subsections. All these assumptions and preset parameters are provided as input to the simulation model.

Table 6. Assumptions and preset parameters in simulation.

Assumptions for server waiting time	
Mean request arrival time (*l*)	3600 to 10800 users/hour (total requests)
Mean session duration (*m*)	20 min
Simultaneous requests per server (*c*)	50 to 500(1 to 10 GbE)
Assumptions for one way minimum delay	
Intermediate device processing time (D_h)	224 ms
End host processing time (D_e)	155 ms
Links (d_n) and propagation delay (v_n)	coaxial(0.195m/µs), fiber(0.198 km/µs)
Assumptions for access network bandwidth consumption	
% of xDSL, FTTH, Cable in AN (p_l)	xDSL(60%), FTTH(20%), Cable(20%)
Bit rate of xDSL, FTTH, Cable in AN (b_{rl})	xDSL(6), FTTH(11), Cable(7)
Total available bandwidth of AN (*B*)	100 Mbps to 1 Gbps

4.1 Content Format and Popularity

ITUT defines two standard formats for video traffic standard-definition television (SDTV) and high-definition television (HDTV); however, demand for HDTV is increasing and it is highly likely that HDTV will replace SDTV in the future. We assumed that all the requested VoD services are delivered in HDTV format. The average bit rate for HDTV is considered to be 10 Mbps for the simulation.

In addition, the content popularity usually fallows Zipf distributions, but for simplicity of the model, we assumed that the probability of having the user-requested content on the VoD server is always 1.

4.2 User Behavior Model

The user behavior model is important for modeling server waiting time and AN bandwidth consumption. We follow models mentioned below from queuing theory as basic assumptions for performing the simulation.

- Request arrivals on the server and requests generated in the community follow a Poisson distribution over time.
- Session duration always follows an exponential distribution over time.
- The formulas presented in this paper represent steady state simulations.

China Telecom is the biggest IPTV service provider in China; we utilized their analysis of user behavior [15] for our models. If a city has 36,000 users, and 10% of users are active during the highest traffic hour for 1 hour, then mean request arrival on the server will be 1 request/s. While modeling we used $1 \leq \lambda \leq 5$ depending on the population of the city [15].

Session length of the users varies from 30 min to 60 min. The probability is 0.37 that a user will browse a channel for 5 min only, 0.52 that they will browse up to 10 min, 0.75 that they will browse up to 25 min and 0.94 that they will browse up to 60 min. Considering this data and the exponential distribution of the session length, we chose 120 seconds as the mean session length.

4.3 Network Properties

Here we will explain some of the preset parameters and assumptions related to the server, and access technologies. For simulating server waiting time we consider that each CO can have multiple VoD servers. A VoD server is connected to a network with 1 – 10 GbE capacity. Considering that half of the bandwidth is reserved for other kinds of services such as VoIP and best effort services, and only half is available for serving VoD requests, a server with 10 GbE capacity can serve 500 simultaneous channels.

To simulate the one-way minimum network delay, we consider that the mean processing delay D_h per intermediate device is 224 μs. Using traceroute, we can determine the number of intermediate devices *N*. Distance between server and community d, kind of cable l and propagation speed of the different cable type v_l (Table 7) are known in advance. The processing delay D_e at the end node is 155 μs.

Table 7. Propagation speed of various cable medium.

Medium	Propagation Speed
Thick Coax	0.77c (231,000 km/sec)
Thin Coax	0.65c (195,000 km/sec)
Twisted Pair	0.59c (177,000 km/sec)
Fiber	0.66c (198,000 km/sec)
AUI Cable	0.65c (195,000 km/sec)

To simulate access network bandwidth consumption we consider that an AN is composed of communities having xDSL, FTTH and a cable networks with penetration ratios of 60%, 20% and 20% respectively. In customer premises the bit rate for the video will differ from customer to customer depending on their broadband access technology. The bit rate *br* of xDSL is 6 Mbps, of FTTH is 11 Mbps and of cable is 7 Mbps. The outgoing bandwidth from AN to server *B* is considered 1 GbE or 100 Mbps for our simulation.

5 Simulation Analysis

This section explains our simulation results and analysis based on the simulation scenario explained in Section 4. An example of the deployment decision is also illustrated according to our simulation analysis.

5.1 Server Waiting Time Analysis

We simulated the decrease in the server waiting time (seconds) with increase in the number of servers (Fig. 9). Each server is capable of serving 500 simultaneous channels. Request arrival follows a Poisson distribution with mean request arrivals of 1, 2 and 3 request per second or 3600, 7200 and 10800 per hour. The session length is exponentially distributed with a mean session length of 120 seconds. Server waiting time decreased exponentially with the increase of the number of servers. With the SLA defined Log (server waiting time) <-300 or server waiting time < 0.5 second, the acceptable range of number of servers for 3600 users / hour is 3 to 8 (Fig. 10). Having more than eight servers would be a waste of resources.

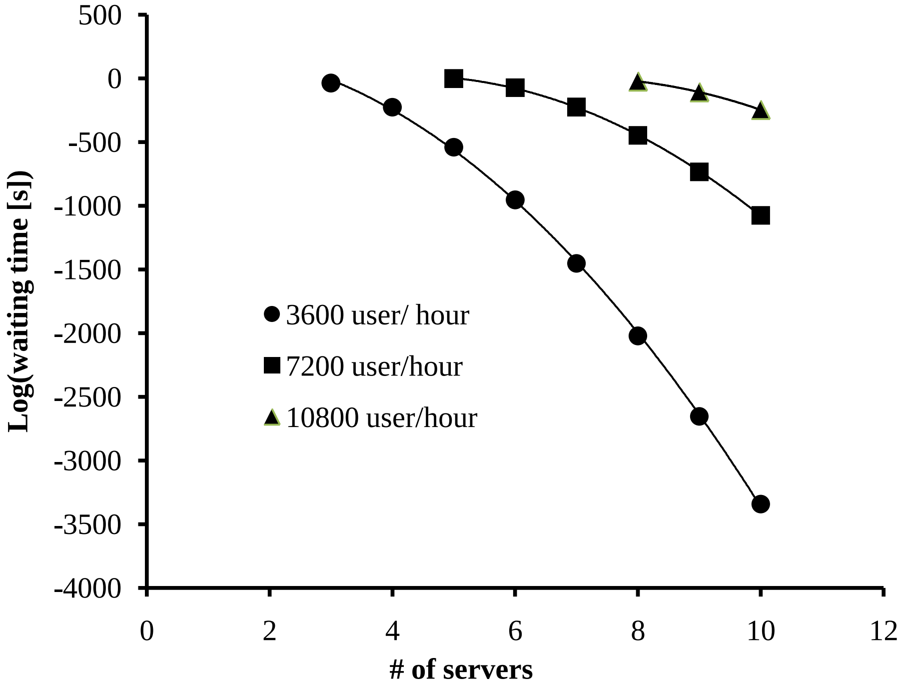

Figure 9. Decrease in Log of server waiting time with increase in number of servers for different number of users per hour.

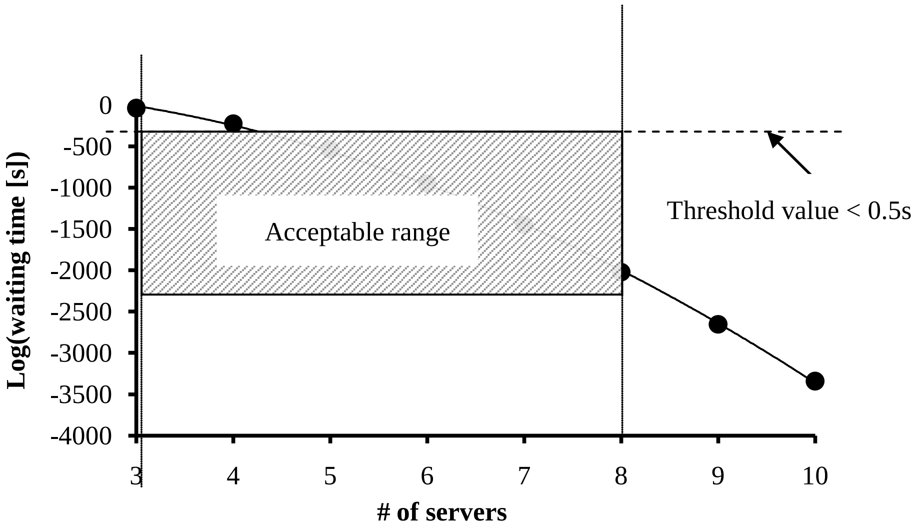

Figure 10. Decrease in Log of server waiting time with increase in number of servers, with 3600 users per hour.

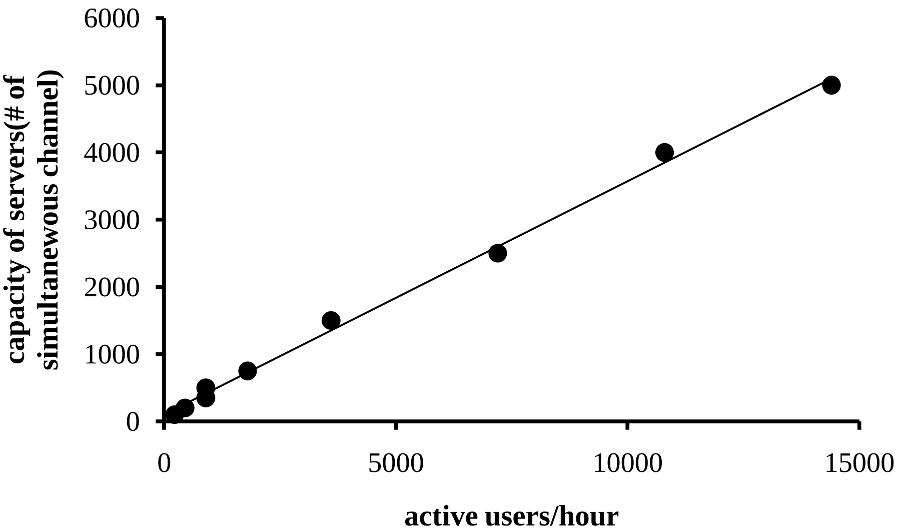

Figure 11. Linear increase in server capacity with increase in number of users per hour.

Table 8. Required number of servers with respect to the server capacity for active user counts in MAN.

Active user/hour	Number of servers	Server capacity	Waiting time (s)	Total server capacity
225	2	50 (1 GbE)	$1.77 x10^{-1}$	100
450	4	50 (1 GbE)	$1.44 x10^{-3}$	200
900	7	50 (1 GbE)	$6.88 x10^{-2}$	350
900	2	250 (5 GbE)	$2.30 x10^{-25}$	500
1800	3	250 (5 GbE)	$1.69 x10^{-8}$	750
3600	3	500 (10 GbE)	$1.73 x10^{-16}$	1500
7200	5	500 (10 GbE)	$3.05 x10^{-1}$	2500
10800	8	500 (10 GbE)	$9.22 x10^{-11}$	4000
14400	10	500 (10 GbE)	$1.39 x10^{-2}$	5000

With the increase in the number of users the minimum required number of servers increases linearly (Fig. 11). For a given number of users the required number of servers with certain server capacity can be calculated (Table 8). We conclude that serving a population of 36,000 users in a city where 10 % of users are active in peak traffic hours, requires at least three servers, of which has 10 GbE capacity and is capable of serving 500 requests simultaneously. The order of magnitude in the results is high because each server can handle multiple simultaneous requests depending on the server capacity. Agrawal et al. [6] used a similar approach to calculate server deployment; however, they calculate blocking probability instead of server waiting time, and their assumptions about user behavior and number of users in the community are not clear. They proposed a server blocking probability of 4.237×10^{-17} for a server capable of serving 500 simultaneous requests.

5.2 One Way Minimum Delay Analysis

We simulate the overall delay simulation according to the increase in distance from server to community and the number of intermediate devices (Fig. 12). We could not observe the clear trend of increase in delay with the increase in distance because the one way delay depends on many other parameters, such as the number of intermediate devices, distance, link type, and propagation delay of the link.

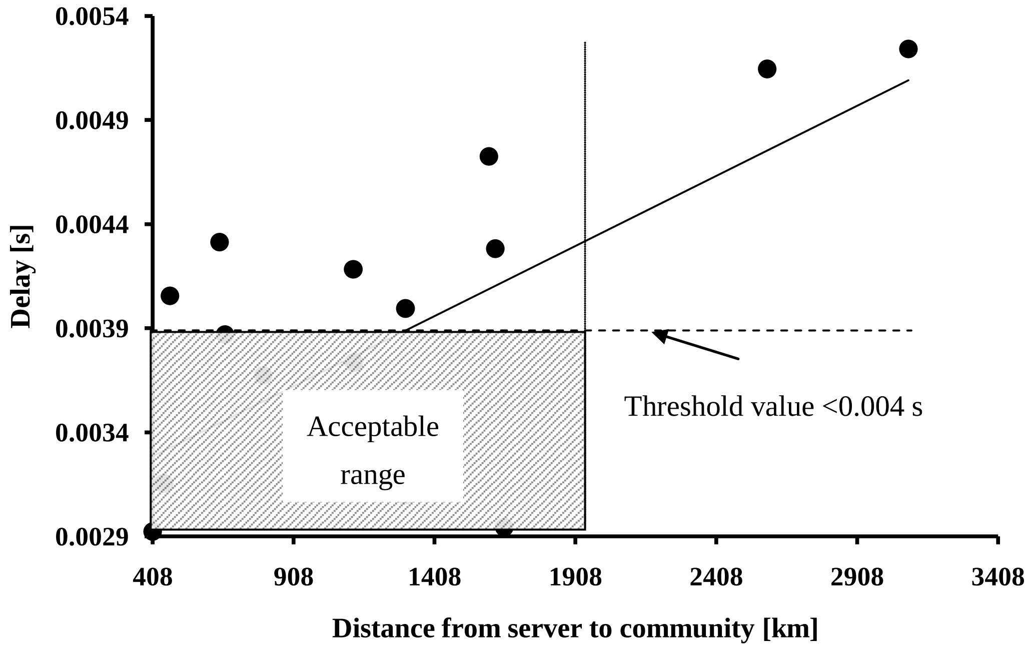

Figure 12. Linear increase in delay with increase in distance from server to community.

Table 9. One way minimum delay with respect to given number of intermediate devices, distance from server and ratio of fiber and coaxial cable.

Serial Number	Number of hops	Distance	Ratio of fiber and coaxial cable	Delay (s)
1	12	408	0.6,0.4	0.002923
2	12	408	0.6,0.4	0.002923
3	13	444	0.6,0.4	0.003154
4	13	448	0.5,0.5	0.003155
5	17	469	0.3,0.7	0.004055
6	18	645	0.5,0.5	**0.004314**
7	16	665	0.6,0.4	0.003869
8	15	800	0.3,0.7	0.003673
9	14	985	0.5,0.5	0.003485
10	17	1120	0.5,0.5	**0.004183**
11	15	1122	0.6,0.4	0.003735
12	12	1301	0.6,0.4	0.003098
13	16	1305	0.6,0.4	0.003995
14	12	1305	0.5,0.5	0.003099
15	19	1601	0.6,0.4	**0.004725**
16	17	1624	0.5,0.5	**0.004282**
17	11	1656	0.5,0.5	0.002944
18	20	2589	0.3,0.7	**0.005145**
19	20	3090	0.6,0.4	**0.005241**

Considering that the SLA defined one-way minimum delay to be < 0.004 s, many values do not satisfy this threshold value. For records 5, 6, 15, and 16 (Table 9), reducing the number of intermediate devices or deploying higher ratio of fiber cable can reduce the overall delay, but for records 18 and 19, reducing the distance from server to community by deploying extra servers closer to the community can reduce the overall delay. After analyzing the results we conclude that the acceptable range of total distance is 400 km to 2000 km (Fig. 12), and the number of intermediate devices should be 12 to 15 (Table 9).

5.3 AN Bandwidth Consumption Analysis

We simulated the bandwidth consumption of the AN over time. The AN is heterogeneous, and has xDSL, FTTH and a cable network with penetration ratios of 60%, 20% and 20% respectively. The bit rate of xDSL is 6 Mbps, of FTTH is 11 Mbps and of cable is 7 Mbps. The request arrival rate varies from 225 to 1800 per hour, and mean session duration is 20 min. The maximum bandwidth of AN is considered to be 100 Mbps. Bandwidth occupancy and bandwidth demand increase linearly with the increase in the number of users in the access network (Fig. 14). The SLA defined value of bandwidth occupancy is < 0.6, and 100 Mbps is sufficient to serve a maximum of 900 users per hour (Table 10). If the number of users of the AN increases, then better outgoing links in AN must be deployed. Ma et al. [16] shows an average 1,000 Mbps bandwidth requirement for 200 active viewers per second in a multicast live TV environment, although that research cannot be directly compared with our simulation because they did not consider heterogeneous access network conditions.

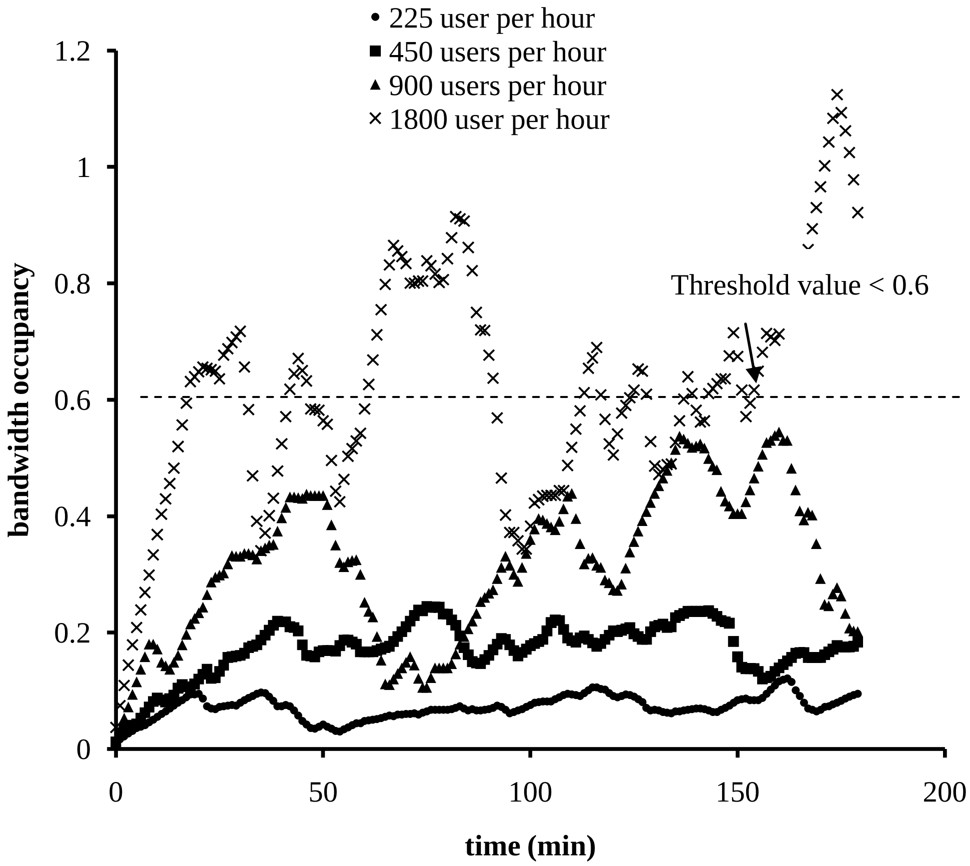

Figure 13. Increase in bandwidth occupancy with time for mean request arrival per min, and mean session duration of 20 min per session.

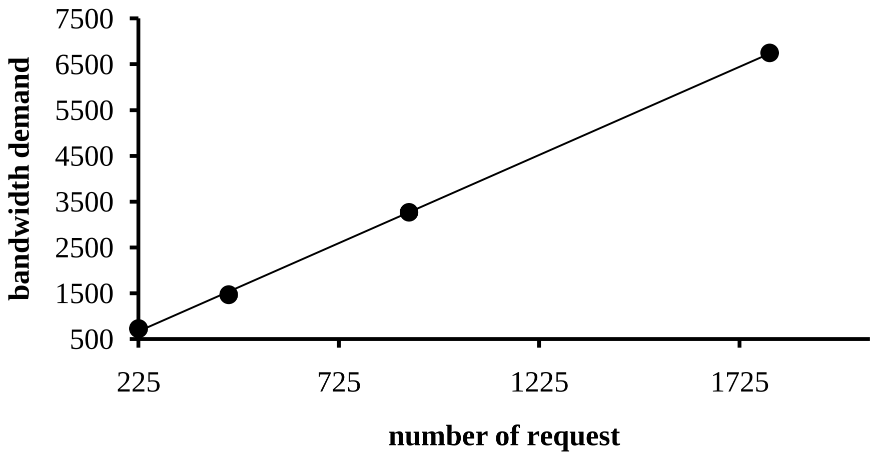

Figure 14. Linear increase in bandwidth demand for access network with increase in number of users in access network

Table 10. Bandwidth consumption for 100 Mbps link from AN to server, and overall bandwidth demand.

Number of request / hour	Bandwidth occupancy for 100 Mbps link	Bandwidth demand / hour
225	0.1212	727.2
450	0.2448	1468.8
900	0.5448	3268.8
1800	1.1244	6746.4

5.4 Overall Analysis: Deployment Decision

Now we consider a complete planning scenario for providing VoD services, and decision making for the scenario based on the analysis explained in previous sections. With the given metrics, assumptions explained in Section 4, and given range of QoE measures, we can calculate the number of servers required, distance from server to community and access network bandwidth capacity requirements (Table 11).

Table 11. Deployment decision according to the given scenario.

Metrics provided by network and service providers	QoE measures on mutual agreement of customers and service provider	Deployment decisions made by service providers
• Users - 36,000 • Active users/hour - 10% • Three AN types (%Community in AN) xDSL-60% FTTH-20% Cable-20%	Server waiting time < 0.5 sec	Server requirement - 3 servers each capable of handling 500 simultaneous request
	One way delay < 0.004 sec	Distance of community from server - 400 to 2000 km with 12 to 15 intermediate device
	Access Network bandwidth consumption < 0.6	Access network - 100 Mbps or more bandwidth

6 Conclusions

6.1 Summary & Contributions

This study developed a model to dimension the network for IPTV service providers that offer VoD services to their customers in heterogeneous environments. The proposed modeling and simulation technique allows determination of the optimum deployment conditions for a given number of potential IPTV users while satisfying pre-defined QoE measures. Our analysis model made the following observations:

- As the number of users increases in an MAN, server waiting time increases exponentially.
- The distance of the server from the user community is the main cause for one way minimum delay.
- As the number of users increases in an AN, the network bandwidth consumption increases.

By analyzing user requirements and network configurations, we calculated substantial objective QoEs such as server waiting time, one way minimum delay, and AN bandwidth consumption. The expected QoE conditions are compared, and used to guide deployment decisions. The deployment decisions can be made considering the following parameters: the number of VoD servers' requirement (Table 8), the minimum physical distance of server from community (Table 9), and AN bandwidth capacity requirement (Table 10).

The models are built considering heterogeneous network conditions. We developed a comprehensive simulation models and analyzed the results. Because the overall delay and bandwidth consumption are important objective QoE measure, determining network parameters and deployment strategy based on these QoE measure are very helpful for the IPTV VoD service providers.

This thesis makes the following contributions to the field of network dimensioning and deployment of IPTV VoD services.

- Providing survey and categorization of IPTV services
- Selecting important QoE measures for dimensioning networks
- Building models for simulating QoE measures
- Considering heterogeneous network conditions for modeling QoE measures
- Providing analysis for network deployment decisions

6.2 Future Work

In this study we selected important QoE measures to dimension the network for IPTV service provisioning. The work proposed in this thesis can be extended in four directions including 1) Content popularity model integration, 2) Bandwidth reduction, 3) Delivery mechanisms, and 4) Live TV delivery models. In the following paragraphs we will explain each of these future works in detail.

The popularity of the content is an important aspect of IPTV deployment. Integrating a popularity and distribution model with the existing QoE metrics can be very useful for dimensioning networks. We can extend and enhance our community model to integrate viewers' preference and patterns. Viewers' preference and popularity of content can help telcos choose the distribution of the content in the optimum way, and can improve QoE and consequently reduce various network loads, including reduction of server load, reduction of access time and bandwidth utilization.

Reduction of bandwidth can be achieved in multiple other ways as well. The other direction of our future work can include finding ways for reducing bandwidth utilization. Multiplexing is an important method of reducing bandwidth. In the future we can explore WDM, which can expand 10 GbE to 1.6 Tbit/s, as well as some other multiplexing protocols like SONET, SDH (TDM) to find ways of reducing bandwidth consumption and improving the QoE of users. Bandwidth consumption can also be reduced by utilizing various batching and patching algorithms. Although this is a well-studied area, we can explore suitable algorithms for different IPTV services and different scenarios of network deployment.

In this study we explored only the unicast mode of video delivery. However, P2P and multicast are also popular ways of video delivery. An IPTV service provider can choose to deliver video using P2P and multicast for reducing network load. Aspects of network deployment and service provisioning will change if models consider set top boxes having

P2P support, and intermediate routers having multicast support. In the future we can enhance our QoE model to support P2P and multicast delivery mechanisms. P2P can reduce the load on the backbone network, but it increases the load on the AN. Analysis of such tradeoffs is also important and beneficial for service providers because the deployment decision will be affected by such analysis. Our future work can include such tradeoff analysis, for different delivery mechanisms.

We concentrated only on VoD services in this thesis, however live TV may become popular in the future as well. Live TV has different characteristics than VoD, so QoE metrics for live TV might need further analysis; for example, channel zapping delay can be considered as an important QoE for live TV delivery. Internet Group Management Protocol (IGMP) protocols are used for delivering live TV, so viewers' preferences and group formation are important aspects of consideration in improving live TV delivery. In the future we will focus on selecting important QoE for live TV and developing models for those QoE measures. Further integration of P2P-assisted live TV and multicast live TV will also be considered in future work. Going further we will compare all delivery mechanisms and determine the tradeoffs for deploying these different type of services and give guidelines for optimum strategy of deployment for given scenarios.

References

[1] ITUT- IPTV Focus group proceedings 2008, http:// www.itu.int/ITU-T/IPTV/.

[2] Microsoft media room, http://www.microsoft.com/mediaroom/.

[3] J. E. Simsarian and M. Duelk, "IPTV Bandwidth Demands in Metropolitan Area Networks," 15th IEEE Workshop on Local and Metropolitan Area Networks, NJ, USA, June 10-13, 2007, pp. 31-36.

[4] M. El-Sayed, Y. Hu, S. Kulkarni and N. Wilson, "Access Transport Network for IPTV Video Distribution," Optical Fiber Communication Conference, March 5-10, 2006.

[5] L. Guo, E. Tan, S. Chen, Z. Xiao and X. Zhang, "Does Internet media traffic really follow Zipf-like distribution?" International Conference on Measurement and Modeling of Computer Systems, California, USA, June 12-16, 2007, pp. 359-360.

[6] D. Agrawal, M. S. Beigi, C. Bisdikian and K. W. Lee, "Planning and Managing the IPTV Service Deployment," 10th IFIP/IEEE International Symposium on Integrated Network Management, Munich, Germany, May 21-25, 2007, pp. 353-362.

[7] F. Thouin and M. Coates, "Video-on-Demand Server Selection and Placement," 20th International Teletraffic Congress, LNCS 4516, Ottawa, Canada, June 17-21, 2007, pp. 18-29.

[8] F. Thouin, M. Coates and D. Goodwill, “Video-on-demand Equipment allocation,” the 5th IEEE International Symposium on Network Computing and Applications, MA, USA, July 24-26, 2006, pp. 103-110.

[9] Y. F. Chen, Y. Huang, R. Jana, H. Jiang, M. Rabinovich, J. Rahe, B. Wei and Z. Xiao, “Towards capacity and profit optimization of video-on-demand services in a peer-assisted IPTV,” Multimedia Systems, vol. 15, no 1, February, 2009, pp. 19-32.

[10] T. Wauters, D. Colle, M. Pickavet, B. Dhoedt and P. Demeester, “Optical Network Design for video on demand services,” Optical Network Design and Modeling Conference, Milan, Italy, February 2005, pp. 250-259.

[11] S. Han, S. Lisle, and G. Nehib, “IPTV Transport Architecture Alternatives and Economic Considerations”, IEEE Communications Magazine, vol. 46, issue 2, February 2008, pp. 70-77.

[12] Y. J. Won, M. J. Choi, J. W. K. Hong, C. K. Hwang and J. H. Yoo, “Measurement of Download and Play and Streaming IPTV Traffic,” IEEE Communications Magazine, vol. 46, issue 10, October 2008, pp. 154-161.

[13] I. Adan, and J. Resing, “Queueing Theory,” Department of Mathematics and Computing Science, Eindhoven University of Technology, 2002.

[14] J. H. Choi and C. Yoo, “Analytical derivation of one-way delay,” IEEE, Electronic Letters, vol. 39, issue 25, December 11, 2003, pp. 1871-1872.

[15] H. Yu, D. Zheng, B.Y. Zhao and W. Zheng, “Understanding user behavior in large-scale video-on-demand systems,” ACM SIGOPS Operating Systems Review, vol. 40, issue 4, October 2006, pp. 333-344.

[16] H. Ma, K. G. Shin, “Multicast Video-on-Demand Services,” ACM Computer Communication Review, vol. 32, No. 1, 2002, pp. 31-43.

[17] Z. Naor, “Multicast Content Distribution over IP Networks,” IEEE GLOBECOM, Nov. 26-30 2007, pp. 2081-2085.

[18] A. Ganjam and H. Zhang, “Internet Multicast Video Delivery,” Proceedings of the IEEE, vol. 93, no. 1, Jan 2005, pp. 159-170.

Acknowledgements

I take this opportunity to express deep sense of gratitude towards Prof. James Won Ki Hong for his invaluable guidance, constant support and encouragement. His broad knowledge in the subject and his humanitarian approach has always been a constant source of inspiration to me. I am grateful to POSTECH for giving an opportunity and Scholarship to carry out my studies and research.

My sincere thanks are due to Young Joon for helping me at various stages of this work. This work wouldn’t have succeeded without his support and immense knowledge of the subject. My special thanks to Prof Derek J. Lactin for editing my thesis and guiding me to improve my English writing. Thanks are also due to my friends Prof Mi-Jung Choi, Eun-Hee, Changkeun, Sung-Ju, Sung-Chul, Joon-Myung, Byungchul, Sung-Su, Kim Jin, Sin-Seok, Hyeok Soo and Jae Yoon who gave me tremendous moral support and help.

This acknowledgement would remain incomplete without expressing my sincere feelings towards my husband Dr. Pawan Kumar Tiwari, who changed my life forever, for his love and encouragement to pursue my studies and achieve higher goals in life.

Suman Pandey

Bio

Birthday:	Jan. 4, 1981
Address:	Information Research Laboratories, Rm. 353,
	POSTECH, Pohang, Korea, 790-784
E-mail:	suman@postech.ac.kr
	suman17july@gmail.com
Tel:	+82-54-279-5654
Fax:	+82-54-279-5699

Education

2007-2009	MS, Dept. of computer Science and Engineering, POSTECH, Korea
1999- 2002	B.C.A, M.C.R.P., Bhopal, India

Publication

Published

1. Suman Pandey, Mi-Jung Choi, Sung-Joo Lee, James Won-Ki Hong, “IP Network Topology Discovery using SNMP,” The International Conference on Information Networking (ICOIN 2009), Chiang Mai, Thailand, Jan 2009
2. 이성주, Suman Pandy, 최미정, 홍원기, "SNMP 기반의 엔터프라이즈 IP 네트워크 연결 정보 관리 시스템", Proc. of KNOM 2008 Conference, Changwon, Korea, April 24-25, 2008.

3. Hyountaek Yong, Jin-ha kim, Suman Pandey, Jeongkyu Shin, Seung-won Hwang, "Blog rank for Maximal Advertisement Coverage" The 8th AEARU Web Technology and Computer Science Workshop, KAIST, Daejeon, Republic of Korea, 2008.

Submitted

1. Suman Pandey, Mi-Jung Choi, Young Jun. Won, James Won-Ki Hong ,"SNMP-based Enterprise IP Network Topology Discovery," International Journal of Network Management, submitted on 8th April 2009.

2. Suman Pandey, Young J. Won, Hong-Taek Ju, and James W. Hong ,"Deployment Analysis of IPTV VoD Service in Heterogeneous Network," 12th Asia Pacific Network Operations and Management Symposium, September 23 ~ 25, 2009, Jeju, KOREA.

Work Experience

2004-2007	Senior software engineer, Business Objects (currently overtaken by SAP)
2003-2004	Software engineer, Dharma Systems
2002-2002	Trainee engineer, vReach solutions