Network Design for the Provision of Distributed Home Theatre services

N.L.S. Fonseca, C. M. R. Franco
F. Schaffa

Relatório Técnico IC 96-11

Agosto de 1996

I) Introduction

The advent of traffic integration opened avenues for countless multimedia applications. Among the most promising applications are video-conference and video-on-demand. Video-conferences extends current voice-conference by allowing participants to see their body language and to use visual information. Users of a video-on-demand system can select and watch films from video archives. In this paper, we consider Distributed Home Theatre (DHT), a hybrid application in which a film is simultaneously played for the participants of a video-conference. The basic idea is to allow a group of distributed users to discuss a film. In a DHT session, any participant can initiate a debate about a specific scene by performing VCR operations on the video. It is expected that distributed home theatre will have a great impact on distance-learning as well as on professional conferences [1].

Video applications will be the major bandwidth consumers of the future broadband integrated network. Take, for example, the deployment of video-on-demand service in the continental United State where there are approximately 77,000,000 viewing households during prime-time [2]. Using data rates of 6 Mb/s for MPEG-II NTSC, 10 Mb/s for JPEG NTSC, and 20 Mb/s for HDTV transmission, the total bandwidth requirements would be 462 Tb/s, 770 Tb/s and 1.54 Pb/s, respectively [3]. These requirements are far in excess of the current network infrastructure. Even with the deployment of high bandwidth switches and links, the huge bandwidth consume calls for proper network engineering.

In this paper, we investigate the interplay between bandwidth and replication of stored programs in network design for the provision of distributed home theatre services. We introduce a cost function which takes into account both server and bandwidth costs, and analyze different network scenarios by considering server replication and caching of programs. We also evaluate the influence of parameters such as: users geographical distribution, number of users per session and network topology on the total network cost. Moreover, we extend our results to study networks with both distributed home theatre and video-on-demand services.

This paper is organized as follows: Section II introduces distributed home theatre services. Section III describes a methodology to compute network cost. Sections III and IV analyzes different network scenarios considering server replication and caching of programs, respectively. In section V, we study networks with both distributed home theatre and video-on-demand services. Finally, conclusions are drawn in section VI.

II) Distributed Home Theatre Services

The basic idea of Distributed Home Theatre (DHT) is to allow a group of distributed users to discuss a film. DHT can be seen as a video/voice conference in which users are watching a film. Any participant of a DHT session can issue VCR operations and initiate a debate about a specific scene. Distributed Home Theatre is a promising tool for distance learning as well as for professional conferences.

In our analysis, we assume that there is a video stream for each individual user and that all the streams of a DHT session are distributed by the same server, which we call control server. Several architectural improvements could be envisioned. The first obvious one is to share a video stream up to the last common switch of two distribution paths. A more elaborated option would be to let the closest server to each user deliver the film. In this case, only control messages would be exchanged between servers. Although these alternatives could significantly reduce the bandwidth demands they require a server/switch complexity which might not be available for a first deployment of DHT services.

We also assume a hierarchical distribution network composed by a national ATM backbone, metropolitan ATM networks and local loops connected to the metropolitan network via head end ATM switches (Figure 1). In the current CATV infrastructure the number of users connected to a single tree must be less than a thousand in order to guarantee minimum Quality of Service [4]-[5]. This trend seems to continue in the future of ATM switched network in order to keep head end switches to a reasonable size. Video servers may be attached at any level of the distribution tree.

III) Network Costs

The trade-off between bandwidth and program replication guides the design of networks with video services. If bandwidth costs were negligible, we would have a central server (or servers) providing service to all network users. Conversely, if storage cost were close to zero, we would have a video archive at every user set top box. Obviously, none of these approaches are realistic. The current network resources are far behind the huge bandwidth demand generated by the central server solution, and video server architecture are much more complex and expensive than current personal computers. Consequently, a solution to ameliorate the bandwidth requirements consists in replicating a certain number of servers in some nodes of the distribution tree. In a real network design, the number and locations of servers are determined by regional demand, network topology and current cost of technology. In this section, we show a framework to compute network costs which can be used for engineering real networks. In our analysis we do not take into account voice and control signals given that video is the dominant component of the bandwidth consume.

III.1) Network Model

We assume that the distribution network is a balanced *d*-ary tree (Figure 2). Each node of the tree represents a switch and the leaves represent head ends switches which connect users to the network. Each switch may have a server (or a program cache) attached to it. The number of head end H is determined by the number of users with a maximum number of users per head end. The number of hierarchy levels is given by:

$$L(d) = \log_d(H)$$

and the number of switches $W(d)$ in such network is given by:

$$W(d) = \sum_{n=1}^{L(d)} d^n$$

Although in the numerical examples we use a binary balanced tree with 1024 head ends and 1000 users per head end, our methodology can also be applied to non-balanced d-ary trees.

III.2) User Distribution

The bandwidth cost of a DHT session and consequently the location of the control server are determined by the number of allocated links to it. Undoubtfully, user dispersion is a key parameter which impact bandwidth costs. However, user behavior might only be fully understood when service is deployed. Therefore, we evaluate our results considering three different distributions. We use: i) a uniform distribution, ii) a normal distribution, and iii) a normal distribution per head end to represent loosely, moderately and highly concentrated patterns of users per DHT session, respectively. The mean of the normal distribution is equal to the media of the number of head ends and each normal distribution per head end is centered in the associated head end. For instance, when considering a normal distribution per head end with mean equals to 525 and with standard deviation equals to 5, 68.26% of the participants of a DHT session are statistically located between head end 520 and head end 530.

III.3) Bandwidth Cost

The total bandwidth cost (C_b) is the sum of the cost of each allocated link, and the cost of an allocated link is proportional to the number of programs delivered through it. Thus,

$$C_b = \gamma_b \sum_{\Lambda} b(\lambda_i)$$

where:

Λ - is the set of links,

b (λ_i) - is the bandwidth on link λ_i,

γ_b - is a normalization constant.

C_b is normalized by the factor γ_b which means makes it easier to explore different cost scenarios.

In order to minimize the total bandwidth cost, we need to minimize the cost of each individual DHT session. The following theorem establishes the conditions for this optimization.

Theorem: The minimum bandwidth cost of a DHT session in a network with linear bandwidth/capacity costs is achieved by choosing a control server located at a node with fifty per cent or more of the session participants below the node whose descendants have individually less than fifty per cent of the session participants below themselves.

This theorem can be used in admission control policies to determine which among the network servers should be the control server of a DHT session. In the appendix, we give a proof of the theorem.

III.4) Storage Cost

The cost of a server depends on the number and on the access rate of the programs stored in it. The access rate of a program is related to the bandwidth needed to support the incoming requests. For example, a server with the top-ten most popular (higher access rate) movies may need to support more bandwidth (and consequently, costs more) than a large achieve of unpopular programs.

In order to capture the effect of both factors in the server cost, we use Zipf's law [6]. It was shown that Zipf's law accurately models the popularity of rented movies in United State by using data published in specialized magazines such as Billboard Magazine and Video Store Magazine [7]. To derive our results we assume that the popularity of programs in DHT services will be the same as their popularity in rental stores. Zipf's law says that the probability of choosing program i among N_p stored programs is given by:

$$z(i) = C / i$$

$$C = 1 \Big/ \sum_{i=1}^{N_p} (1/i)$$

We assume that the server at the root of the distribution tree cots one unit. Thus, the storage cost at the root is:

$$C_{root} = \sum_{i=1}^{N_p} z(i)$$

where $z(i)$ is the probability of choosing program i where programs are ordered according to their decreasing popularity.

The influence of the popularity can be seen by the fact that if we reduce by half the population with access to a certain server the server cost would be given by:

$$C_s = \sum_{i=1}^{N_p} z(2 \times i)$$

where $z(2 \text{ x } i)$ reflects the lower requirement of program i due to the lower request access rate and $z(j) = z(N_p)$, for $j > N_p$

Finally, the total server cost is given by the summation of the cost of each individual server:

$$C_s = \sum_{\Delta} C_i$$

where Δ is the network set of servers.

III.5) Total Cost

The total cost to provide DHT services is given by the summation of the bandwidth with the server cost.

$$C_t = C_b + C_s$$

Given that we consider normalized bandwidth and server costs we can define a weight factor, ρ, which reflects different cost ratios between bandwidth and server costs. Hence, the total network cost is given by:

$$C_t(\rho) = (2/(1+\rho)) \times (\rho \times C_b + C_s) \quad \text{for } \rho > 0$$

IV) Server Replication

In our attempt to understand the trade off between bandwidth and replication of storage program, we analyze a server replication strategy. By replicating servers, we decrease the bandwidth requirements on each server, and, consequently, we decrease individual server costs. We initially consider the cost of a single server located at the root of the distribution tree (level zero). We then place servers only at all nodes of level 1. We proceed by considering networks with servers located only at all nodes of the i^{th} level.

Results were obtained via simulation. We used the replication methods for generating confidence intervals with 95% of confidence level. The sample size used to compute each point of the curves was such that confidence intervals width were less than 5% of the mean. Given that results with a high degree of confidence were generated, we only show the mean value for the sake of visual interpretation.

To compare trends we normalize both bandwidth and server costs by the highest value in each curve. We also display results by normalized depth (level 0 corresponds to the root level and level 1 to the head end level). In our numerical examples, we display the impact of the number of users per session and the user distribution on the network cost.

When computing the cost of server replication at the i^{th} level of the distribution tree, we assign a server to be the control server of a DHT session if the server is the one with the majority of the session participants below it. Of course, this criterion produces the same assignment as our optimization theorem for a specific session and tree level.

The total bandwidth cost considering replication at l^{th} level is given by:

$$C_b(l) = \gamma_b \sum_{\beta} c_i(l)$$

where:

$c_i(l)$ - is the cost of the i^{th} DHT session when considering server replication at level l which is computed by the sum of links allocated to the session

β - is the set of all DHT sessions.

γ_b - is a normalization constant

In figure 3, we show the normalized bandwidth cost as a function of the normalized depth for different distributions. We note that when participants become more concentrated the optimum level for placing the server tends to move closer to the head end level. For highly dispersed users (uniform distribution - Figure 3.a) the optimum level is one level below the root. For moderately concentrated participants (normal distribution - Figure 3.b), the optimum level is down approximately a third of the tree height and for highly concentrated users (normal distribution per head end - Figure 3.c) the optimum level is two levels above the head end level. Figure 3 illustrates our bandwidth minimization principle which says that the optimum control server for a DHT session should contain fifty percent or more participants below it and none of its descendants should satisfy this condition. We note that as users are more dispersed, the minimization principle is achieved by nodes closer to the root.

It is interesting to note that the number of users per DHT session has less effect then the user distribution. Actually, we noticed that the number of users per session may impact the choice of control server only for moderately concentrated distributions. By increasing the number of participants in this type of distribution, for the optimal location, the control server moves towards the root. Figure 3.b shows that for a lower number of users per DHT session (3 and 4 users) the server is placed at 0.3 of the normalized depth of the distribution tree whereas for a higher number (5 and 6 users) the optimum level is at level 0.2. No such effect was observed for loosely (Figure 3.a) and for highly concentrated patterns (Figures 3.c)

The total server cost for replication at l^{th} level is given by:

$$C_s(l) = \gamma_s(l) \sum_{j=1}^{d^l} \sum_{i=1}^{N_p} z(i/x_j)$$

where:

x_j - is the j^{th} server demand, i.e., the ratio between the number of DHT session served by control server j and the total number of DHT session

$\gamma_s(l)$ - is a normalization constant for level l

The total normalized cost of server replication at level l is given by $C_b(l) + C_s(l)$. Figure 4 illustrates this computation. In this particular case, the cost is dominated by the server cost.

Figure 5 displays the normalized network cost as a function of the normalized depth. We observe that for loosely and moderately concentrated participants (Figures 5a and 5b) the normalized cost curve is almost flat until level 0.3. This is because the saving due to bandwidth reduction is not enough to compensate the high cost of servers close to the root. Servers close to the root are more expensive because they serve larger populations. For moderately concentrated distribution and a small number of users per session (3 and 4 users), we notice that the bandwidth saving counteracts the server cost with the minimum at level 0.3. For highly concentrated distributions as the concentration increases the bandwidth gain dominates the total cost pushing the optimum level closer to the head end level.

All the results presented up to this point considered that bandwidth and server costs have the same weight. In Figure 6, we show the impact of giving different weights for bandwidth and server costs (varying ρ). We notice that as the importance of bandwidth decreases so does the overall cost. In this way, the optimum location for the control server moves closer to the root. This is more noticeable for loosely concentrated users (Figure 6.a) than for highly concentrated user (Figure 6.b) due to the need of connecting dispersed user. The flattening of Figure 6.b around a normalized depth of 0.5 illustrates the influence of users distribution on the bandwidth costs.

In order to evaluate the dependence of our findings on the network topology, we analyzed different balanced n-ary trees by varying the node degree. Obviously, as we increase the node connectivity, we decrease the total (non-normalized) cost due to the reduction of the bandwidth requirements. In Figure 7, we show an example with 4096 head ends considering binary, 4-ary and 16-ary balanced trees. We notice that the topology does not impact significantly our results. All three curves have similar shape and equivalent behavior at the same normalized depth.

V) Program Caching

We now investigate a program caching strategy. Instead of replicating the whole server, we replicate just the most popular programs. In this way, we try to reduce the overall cost by reducing individual server costs. However, we need to provide a full server at the root of the tree to handle the requests to non-popular programs (cache misses).

To evaluate the impact of adopting program caching, we use the same rationality used when studying server replication: we consider that there are caches only at all nodes of level *l* in addition to the full server at the root. For each DHT session, we assign a control cache at level *l*, i. e., a cache which will handle all the requests for that session.

The bandwidth cost for cache replication at the level *l* is given by:

$$C_b = \gamma_b \left(\alpha \sum_{\beta} c_b(l) + ((1 - \alpha) \times \mu \times h) \right)$$

where:

β - is the set of DHT sessions;

γ_b - is a normalization constant;

$c_b(l)$ - is the bandwidth cost when considering cache replication at the l^{th} level only;

μ - total number of participants;

h - tree height;

$\alpha = \sum_{i=1}^{M} z(i)$ - is the cache hit probability

and the cache cost for cache replication at level *l* is given by:

$$C_c(l) = \gamma_c(l) \sum_{j=1}^{d^l} z(1/x_j) + C_{root}$$

where:

$\gamma_c(l)$ - is a normalization constant;

x_j - is the j^{th} cache demand, i.e., the ratio between the number of DHT sessions served by the j^{th} cache and the total number of sessions

Figure 8 displays the normalized cache cost as a function of the normalized depth for several values of cache size. We also show the cost for server replication, i.e., considering (full) servers only at level l nodes. We notice the same trend observed when analyzing server replication. For loosely and moderately concentrated participants the cost curve is almost flat until level 0.3, and in this region of the distribution tree all caches cost the same. At the head end level, small caches are the most attractive and they can give a cost saving of approximately 10% of the server cost. One could expect that larger caches would be more attractive at the head end level because they would avoid cache misses. However, for loosely and moderately concentrated users, the price we pay for a cache miss is not a penalty. Actually, cache misses bring bandwidth savings since the level which minimizes the bandwidth demand is closer to the root. Hence, in a real network design if other video services demand that caches should be placed closer to the head end level, it is better (for DHT services) to adopt small caches. For highly concentrated users, the picture looks different. In this case, a cache miss is a real penalty because the optimum placement level is close to the head end. Consequently, we should try to satisfy the larger number of requests at the head end level. This principle can be observed by the cost saving as we increase the cache size, and by the fact that server replication always gives the best possible saving. Therefore, if by any design reason, we need to adopt program caching, large caches will give the least expensive solution. The same placement pattern for both server and cache replication can be understood by the high bandwidth cost to provide DHT services.

A compromise between server and cache replication is to use multi-level caches [8]. In a multi-level cache, caches are placed at different levels of the tree, and their content are non-overlapped and “contiguous”. Thus, if there is a cache miss at a certain cache level, the request can be satisfied at a higher cache level and not necessarily at the root level as in the single level case. The bandwidth and cache cost for multi-level caches are a generalization of the single level case. For instance, for a cache with two levels placed at levels l_1 and l_2 of the distribution tree are and is given by:

$$C_b(l_1, l_2) = \gamma_b\left(\alpha_1\sum_{\beta} c_b(l_1) + \alpha_2\sum_{\beta} c_b(l_2) + (1-\alpha_1-\alpha_2)\times u\times h\right)$$

$$C_c(l_1, l_2) = \gamma_c\left(\sum_{j=1}^{d^{l_1}}\sum_{i=1}^{M_1} z(1/x_j) + \sum_{k=1}^{d^{l_2}}\sum_{i=M_1+1}^{M_2} z(1/x_k) + C_{root}\right)$$

In our study, we found that multi-level caches are not worth adopting in network design for the provision of DHT services for the same reasons mentioned in the single level case.

VI) Networks with both DHT and VoD services

Video will be the major bandwidth consumer in the future broadband network. It is essential that we take into account the requirements of different video applications in a real network design. One of the most promising applications is video-on-demand. In a video-on-demand system, individual users can select movies to watch from a video server. Providing DHT services costs more than providing VoD services. This happens because users in a DHT system connect themselves to the server which minimizes the bandwidth cost of a DHT session whereas users in a VoD system connect themselves to the closest server. In other words, for the same number of users the total bandwidth cost of VoD sessions is always a lower bound of the bandwidth cost of a DHT session.

In Figure 9, we plot the overall cost of server replication in a network with DHT and VoD services. Video servers provide programs for both type of services. We vary the percentage of total number of users involved in DHT sessions from 10 to 50% of the population. We believe that in a real network it is most likely that the percentage of DHT users will be around 30% and will not exceed 50%. We know from [3] that the optimum placement of a server in a network with only VoD services is somewhere between 70% and 90% of the depth of the tree. Figures 9.a and 9.b shows that for loosely and moderately concentrated users, the VoD server placement trend is maintained only when we have a small percentage (10%) of DHT users. As the percentage of DHT users increases, the optimum locations moves towards the root of the tree. Moreover, for loosely concentrated users we have very little cost savings. In other words, the high bandwidth requirements of DHT services dominates the network costs even for percentage of DHT users as low as 30%. For highly concentrated users we have a different picture (Figure 9.c). Highly concentrated users present the same distribution pattern of VoD users. The optimum location is maintained even for percentages of DHT users as high as 50%. Actually, VoD users distribution can be considered an extreme case of DHT user pattern in which all the participants of DHT session are located below the same head end.

VII) Conclusions

Video services is both a major business driver and a bandwidth consumer for the future broadband integrated network. Understanding video applications requirements is of paramount importance for network design. In this paper, we explored the interplay between bandwidth consume and program replication for the provision of Distributed Home Theatre services. We analyzed server and cache replication strategies which aimed at reducing the high bandwidth demand. Moreover, we proved a bandwidth optimization principle which can be implemented in admission control policies. As user behavior can only be fully understood when service is deployed, we used three different distributions to derive our findings. In addition, we also investigated the influence of the number of users per session, and network topology. For loosely / moderately concentrated participants the optimum placement of server/cache is close to the root of the distribution tree whereas for highly concentrated users it is close to the head ends. This trend is dictated by the high penalties paid when the bandwidth minimization condition is not satisfied. We also show how these requirements impact significantly the design of networks with both DHT and VoD services even for a low percentage of DHT subscribers. It is essential, therefore, to combine program replication strategies with other network features. In line with that, we are currently comparing the provision of DHT services in networks with and without stream sharing.

VIII) References

[1] J. A. Schnepf, D. H. C. Du, E. Russell Ritenour and A. J. Fahrmann, Building Future: Medical education environments over ATM networks", *Communications of the ACM*, vol 38, n 2, pp 54-69, vol 38, n 2, Feb 1995.

[2] *Statistical Abstract of United States*, U.S. Bureau of the Census, 113th ed., 1993.

[3] J. P. Nussbaummer, B. V. Patel, F. Schaffa and J.P.G. Sterbenz, "Networking requirements for interactive video on demand", *IEEE Journal on Selected Areas in Communications*, vol 13, n 5, pp 779-787, June 1995.

[4] J. P. Nussbaumer, B. V. Patel and F. Schaffa, "Multimedia delivery on demand: Capacity analysis and implication" in *Proc 1994 Local Computer Networks Conference*, 1994

[5] J. P. Nussbaumer and F. Schaffa, "Capacity analysis of CATV for on-demand multimedia distribution" in *Proc First ISMM Int Conf Distributed Multimedia Systems Applications*, Aug, 1994

[6] Y. S. Chen, "Mathematical modeling of empirical laws in computer application: a case study", *Comp. Math. Applicat.*, pp. 77-87, Oct. 1992.

[7] A.L. Cherenak, "Tertiary storage: An evaluation of new applications", Ph.D. dissertation, University of California, Berkeley, Dec. 1994.

[8] J.P. Nussbaumer and F. Schaffa, "On Bandwidth and storage trade-off in multimedia distribution networks", *Proc of INFOCOM'95*, pp. 1020-1026, 1995.

Appendix

We now prove the bandwidth minimization theorem.

Theorem: The minimum bandwidth cost of a DHT session in a network with linear bandwidth/capacity costs is achieved by choosing a control server located at a node with fifty per cent or more of the session participants below the node whose descendant have individually less than fifty per cent of the session participants below themselves.

This theorem can be used in admission control policies to determine which among the network servers should be the control server of a session, as well as in network design studies.

Proof: Let i be a node with fifty per cent or more of the session participants below the node whose descendant have individually less than fifty per cent of the session participants below themselves. Let C_i be the bandwidth cost of a DHT session when the control server is located at node i. We need to prove that the bandwidth cost of the same DHT session whose control server is located at node *j*, for $i \neq j$, C_j, is always greater than C_i. C_i and C_j are related by:

$$C_j = C_i + l_d + l_n$$

where

l_d - is the difference between the number of links allocated by the participants of the DHT session who are descendant of node *i* when the control server is located at node *j* and when the control server is located at node *i*

l_n - is the difference between the number of links allocated by the participants of the DHT session who are non-descendant of node *i* when the control server is located at node *j* and when the control server is located at node *i*

Thus, we need to show that C_j is always greater than C_i. In other words, we need to prove that $l_d + l_n > 0$. In order to prove it, we need to consider the following cases: i) node *j* is an ancestor of node *i*, ii) node *j* is a descendant of node *i*, and iii) node *j* is neither a descendant nor an ascendent of node *i*.

i) node *j* is an ancestor of node *i*

If node *i* is not a son of node *j*, we have:

$$l_d = d_{ij} \text{ x } p_i$$

$$l_n = -d_{ij} \times (P - p_j) + \sum_{k=0}^{d_{ij}-1} (d_{ij} - k) \times p_k$$

where:

d_{ij} - is the number of links connecting node *i* to node *j* (height difference)

p_i - is the number of participants who are descendants of node *i*

p_j - number of participants which are descendants of node *j*

P - number of participants of the DHT session

$K =$ {the ordered set of nodes of the path connecting node *j* and the father of node *i*, starting from node *j*}

p_k - the number of participants of the DHT session who are descendant of node $k \in K$ and who do not have any other ancestor $\tilde{k}$ which $\tilde{k} > k$ and $\tilde{k} \in K$.

The minimum value of l_n, $\tilde{l}_n$, is achieved when all node *i* non-descendant are concentrated below node *i* sibling. $\tilde{l}_n$ is given by:

$$\tilde{l}_n = -d_{ij} \times \left(P - p_j \right) - d_{ij} \times \left(p_j - p_i \right)$$

Thus if we prove that $\tilde{l}_n + l_d > 0 \Rightarrow l_n + l_d > 0$

We know that:

$$\tilde{l}_n = d_{ij} \times (p_i - P)$$

$$\tilde{l}_n + l_d = 2d_{ij} \times p_i - d_{ij} \times P$$

By definition we know that $p_i \geq 0.5P$. Thus,

$$\tilde{l}_n + l_d > 0 \Rightarrow l_n + l_d > 0 \Rightarrow C_j > C_i$$

If node *i* is node *j*'s son, we have that:

$$l_d = p_i$$

$$l_n = - (P - p_i)$$

$$l_n + l_d = 2p_i - P$$

By definition $2p_i \geq P \Rightarrow l_n + l_d \geq 0 \Rightarrow C_j \geq C_i$

ii) node *j* is a descendant of node *i*

If node *j* is not node *i*'s son:

$$l_n = (P - p_i) \ x \ d_{ij}$$

$$l_d = -d_{ij} \times p_j + \sum_{k=0}^{d_{ij}-1} (d_{ij} - 2k) \times p_k$$

where:

K = {the ordered set of nodes of the path connecting node *i* and the father of node *j*, starting from node *i*}

p_k - the number of participants of the DHT session who are descendant of node $k \in K$ and who do not have any other ancestor $\tilde{k}$ which $\tilde{k} > k$ and $\tilde{k} \in K$.

$\tilde{l}_d$ the minimum value of ld is obtained when the non-descendant of node *w* which is node *i*'s son are all concentrated below node *j*'s sibling. In this case, we have:

$$\tilde{l}_d = -d_{ij} \text{ x } p_j + d_{ij} \text{ x } \alpha P + (1 - d_{ij}) (p_i - \alpha P - p_j)$$

where:

α - is the fraction of the total number of the session participant which are descendant of node *i* and are not descendant of any other node $k \in K$

$$\tilde{l}_d = (2\alpha d_{ij} + d_{ij}) \text{ x } P - 2d_{ij} \text{ x } p_i + (p_i - \alpha P - p_j)$$

We can express p_i as:

$$p_i = (\alpha + \beta) P$$

where:

β - fraction of the total number of participants who are descendants of node *i* and whose ancestor different than node *i* are in *K*.

$$\tilde{l}_d = d_{ij} \times P - 2\beta \times d_{ij} \times P + (p_i - \alpha P - p_j) > 0$$

By definition $\beta < 0.5$. Thus,

$$l_n + \tilde{l}_d > l_n + l_d > 0 \Rightarrow C_j > C_i$$

If node *j* is node *i*'s son:

$$l_d = -p_j + (p_i - p_j)$$

$$l_n = P - p_i$$

$$l_n + l_d = P - 2p_j$$

By definition: $P > 2p_j$. Thus, we have:

$$l_n + l_d \geq 0 \Rightarrow C_j \geq C_i$$

If *j* is a son of *i*

$$l_d = - p_j + (p_i - p_j)$$

$$l_n = P - p_i$$

$$l_n + l_d = P - 2P_j > 0 \Rightarrow C_j > C_i$$

iii) Node *j* is neither an ancestor nor a descendet of node *i*

$$l_d = d_{ij} \text{ x } p_i$$

$$l_n = -d_{ij} \times p_j + \sum_H s_h \times p_h$$

where:

d_{ij} - is the (minimum) distance between node *i* and node *j*

p_h - is the number participants below headend *h* which is neither a node *i* descendat nor a node *j* descendant

s_h - diferenece between the link demand of the participants below head end *h* when the control server is located at node *j* and when it is located at node *i.*

$$l_n + l_d = -d_{ij} \times p_j + \sum_H s_h \times p_h + d_{ij} \times p_i$$

We know that $\sum_H s_h \times p_h$ is minimized, (i.e. maximize link saving) when all the participants neither below node *i* nor below node *j* are concentrated below a node *j* sibiling. Thus, we have:

$$l_n + l_d = d_{ij} \times p_i - d_{ij} \times p_i + (2 - d_{ij}) \times (P - p_j - p_i)$$

$$l_n + l_d = 2d_{ij} \times p_i - d_{ij} \times P_i + 2\,(P - p_j - p_i)$$

by definition $2P_i > P$ and $P - p_j - p_i > 0$. Thus,

$$l_n + l_d \geq 0 \Rightarrow C_j \geq C_i$$

We notice that the theorem above is also valid for symetrical trees, i. e., trees in which links at the same level have the same cost.

Acknowledgements

We would like to thank CNPq and FAPESP for their financial support and CENAPAD/SP for their computational support.

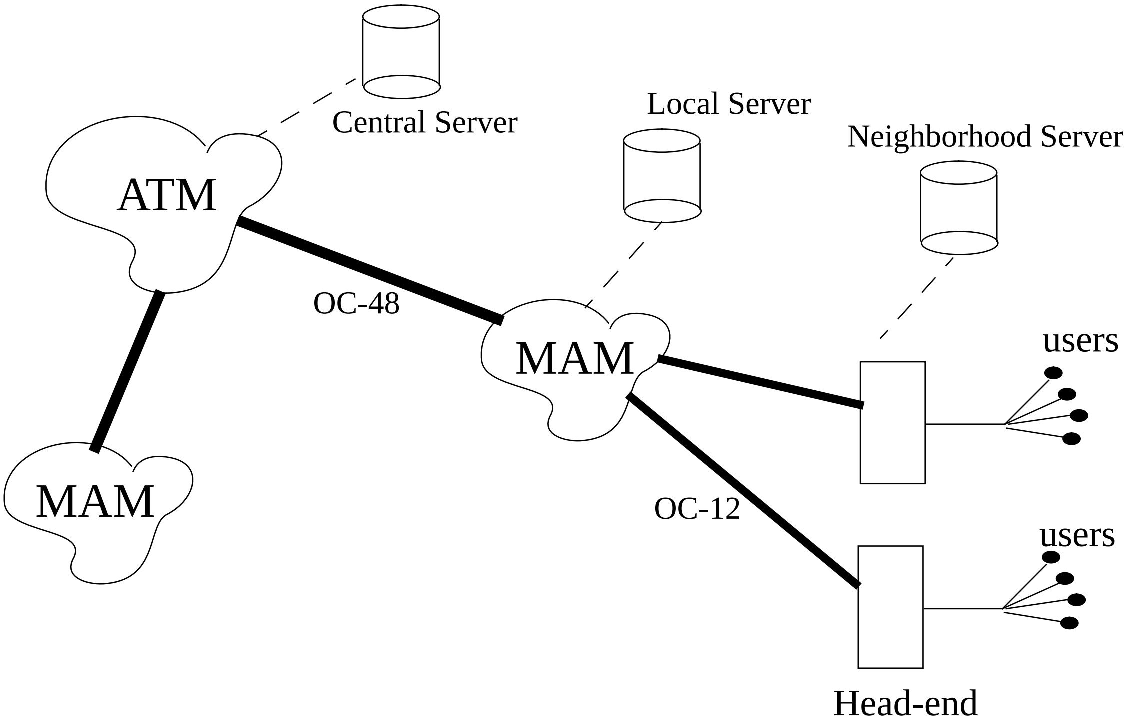


Figure 1: Network model.

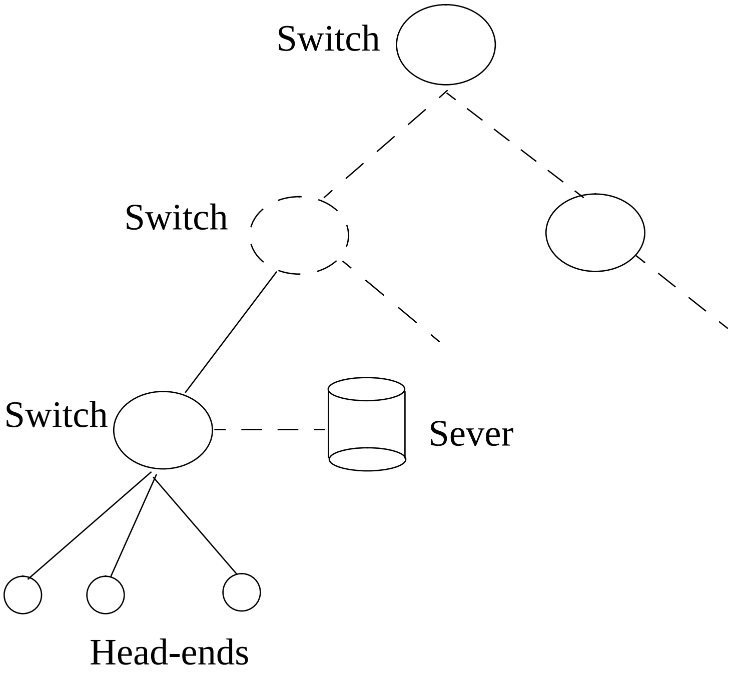


Figure 2: The Distribution tree.

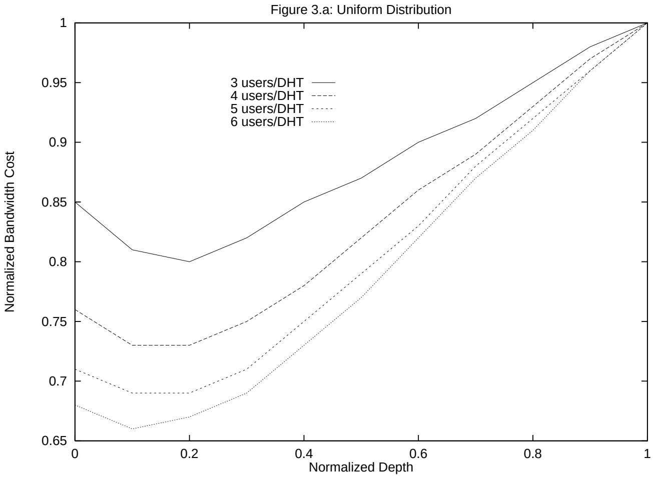


Figure 3.a: Uniform Distribution

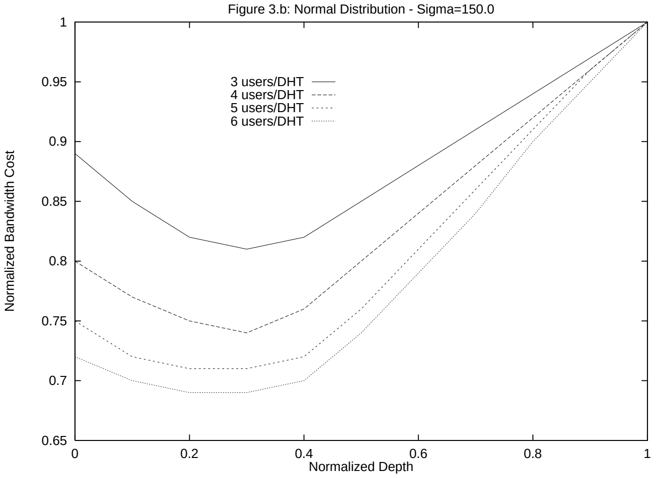


Figure 3.b: Normal Distribution - Sigma=150.0

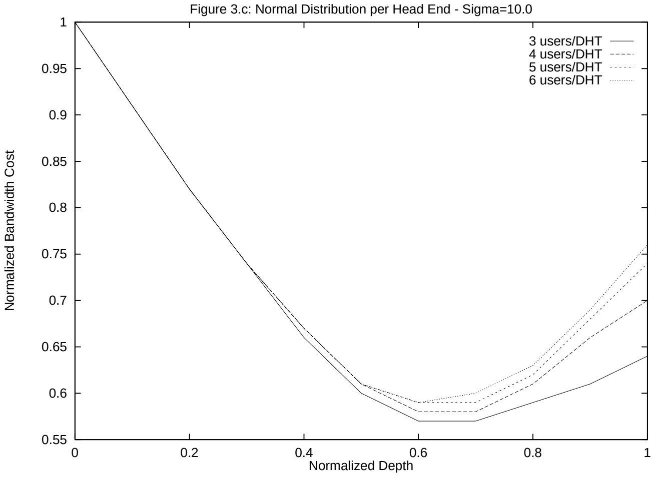


Figure 3.c: Normal Distribution per Head End - Sigma=10.0

Figure 3: Normalized bandwidth cost x normalized depth for: i) different number of users per DHT session, ii) a uniform distribution (Figure 3.a), iii) a normal distribution with $\sigma = 150$ (Figure 3.b) and iv) a normal distribution per head end with $\sigma = 10$ (Figure 3.c).

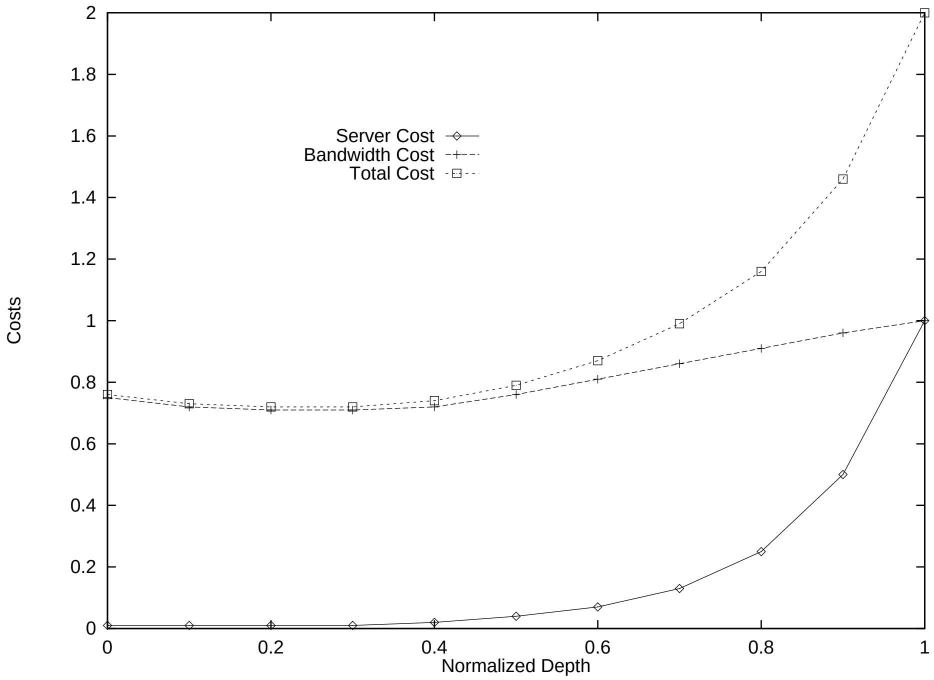
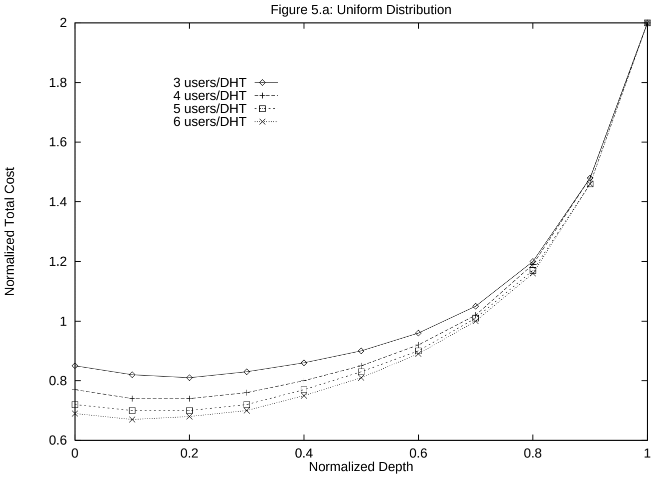
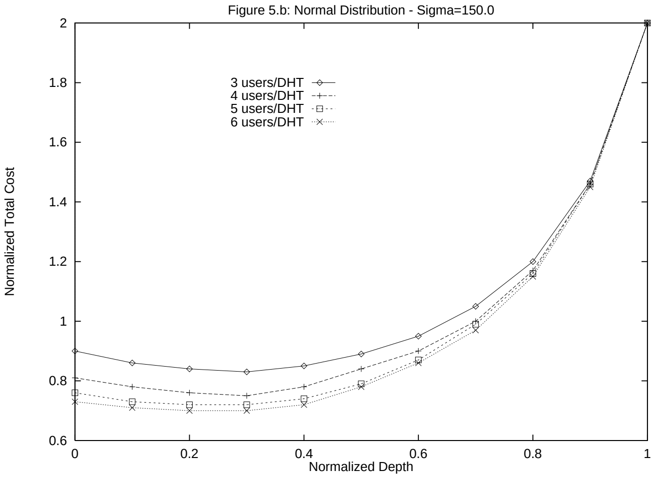


Figure 4: An example of total normalized cost computation considering server replication for a normal distribution with $\sigma = 150$ and 5 users per DHT session.





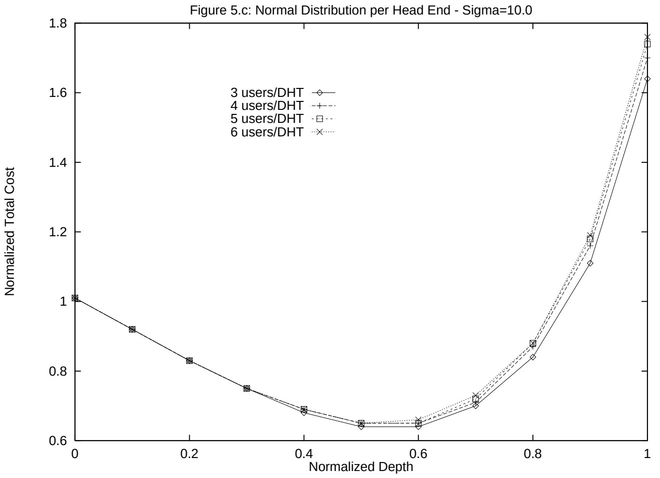


Figure 5: Normalized network cost x normalized depth for: i) different number of users per DHT session, ii) a uniform distribution (Figure 5.a), iii) a normal distribution with $\sigma = 150$ (Figure 5.b) and iv) a normal distribution per head end with $\sigma = 10$ (Figure 5.c).

Figure 6.a: Uniform Distribution

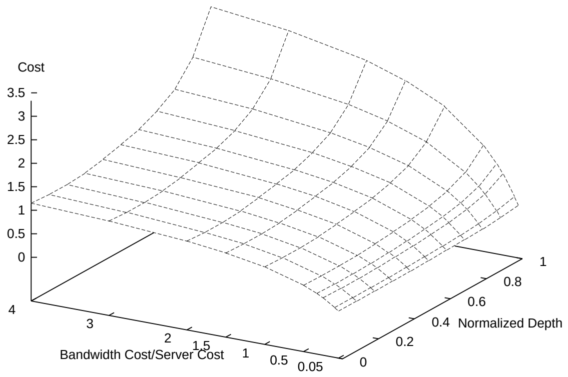


Figure 6.b: Normal Distribution per Head End - Sigma=5.0

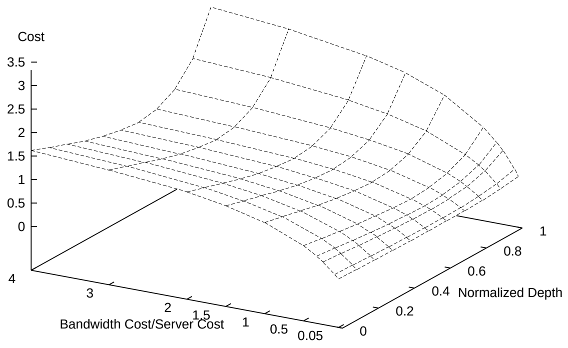


Figure 6: Normalized network cost x the normalized depth x ρ for a uniform distribution and 5 users per DHT session.

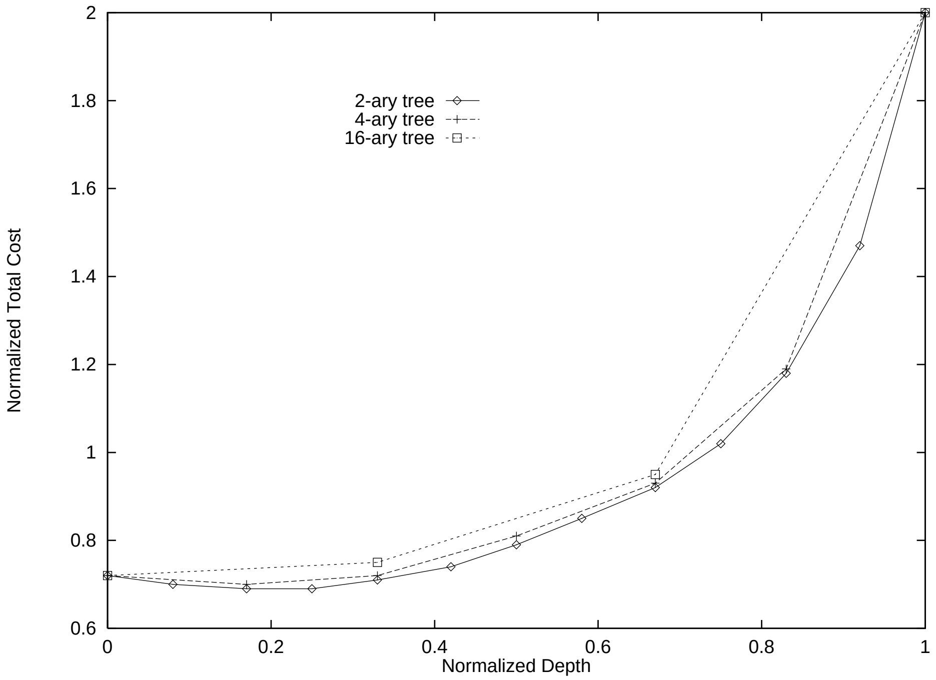


Figure 7: Normalized network cost x normalized depth for different topologies.

Figure 8.a: Uniform Distribution

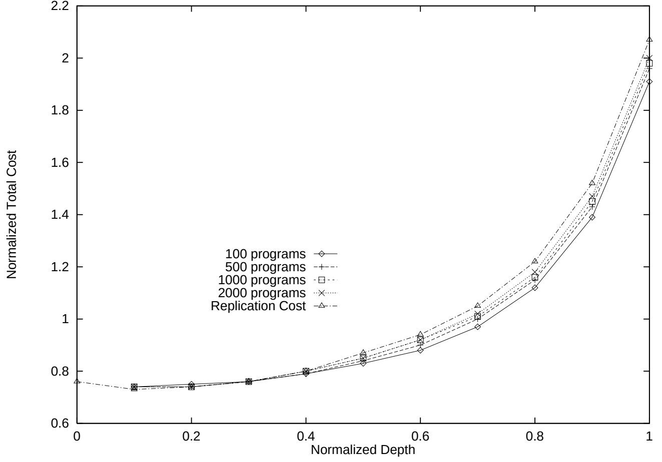


Figure 8.b: Normal Distribution - Sigma=110.0

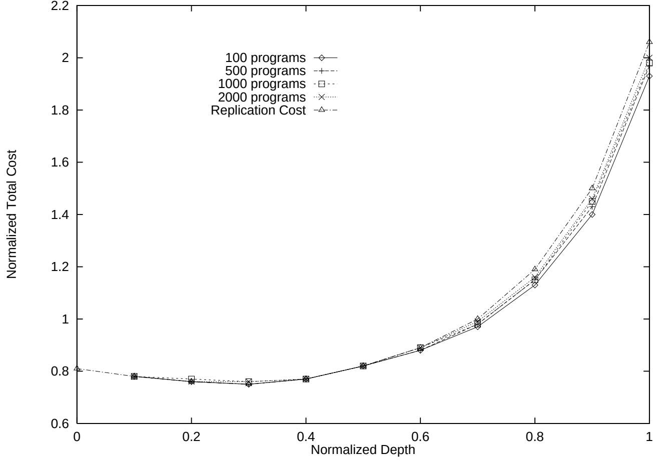


Figure 8.c: Normal Distribution per Head End - Sigma=5.0

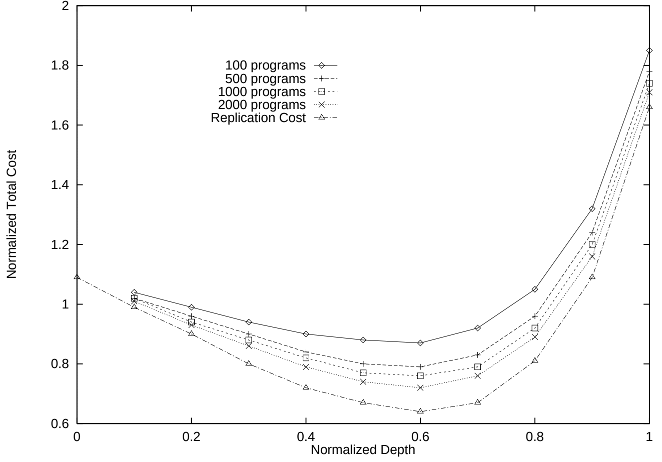


Figure 8: Normalized cache cost x normalized depth for: i) different cache sizes, ii) 5 participants per DHT session, iii) a uniform distribution (Figure 8.a), iv) a normal distribution with σ = 110 (Figure 8.b) and v) a normal distribution per head end with σ = 5 (Figure 8.c).

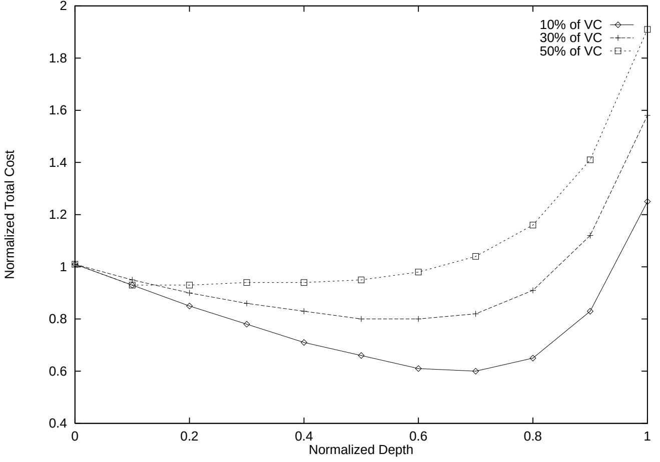


Figure 9.a: Uniform Distribution

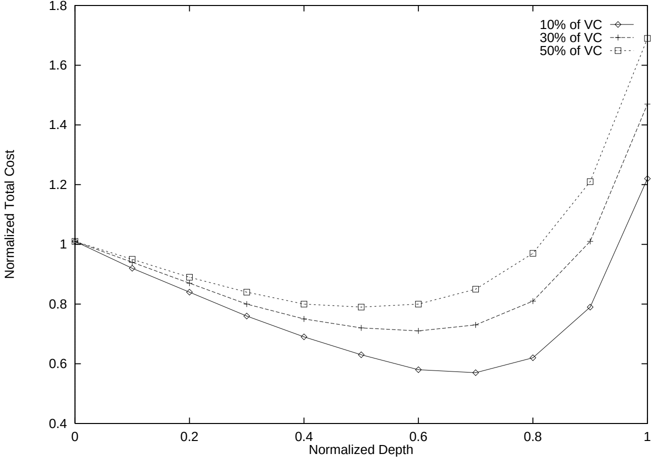


Figure 9.b: Normal Distribution - Sigma=110.0

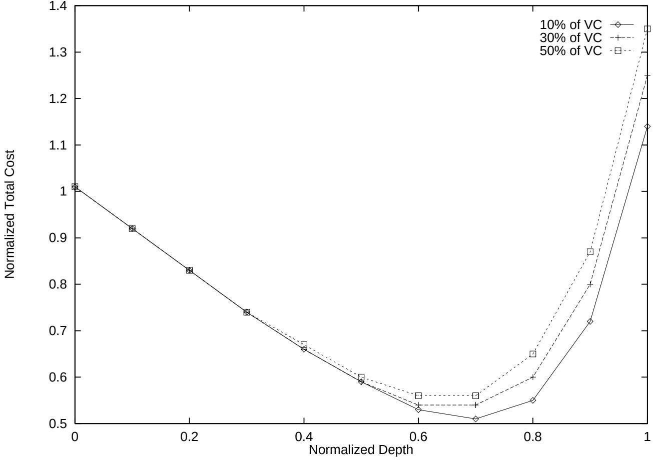


Figure 9.c: Normal Distribution per Head End - Sigma=5.0

Figure 9: Normalized network cost x normalized depth considering server replication in a network with both DHT and VoD services for: i) 5 users per DHT session, ii) uniform distribution (Figure 9.a), iii) a normal distribution with $\sigma = 110$ (Figure 9.b) and iv) a normal distribution per head end with $\sigma = 5$ (Figure 9.c).