Distributed ATM Switching Topologies in Tree-Structured Networks

Urs Thürmann[a], Martina Zitterbart[a], Thomas Meuser[b]

[a]Technical University of Braunschweig

[b]Philips Research Laboratories Aachen

ABSTRACT

HFC networks, today mostly used to broadcast traditional cable TV programs, will possibly play a major role in future broadband access network development. The structure of HFC networks is briefly described and bottlenecks in case of high load are identified. To improve performance and eliminate the bottleneck, the approach presented in this paper suggests the replacement of passive network elements like fiber nodes and signal amplifiers by ATM switches. This approach has been evaluated with respect to various network topologies and user scenarios. A simulation model used for the evaluation is presented within this paper and simulation results are discussed.

Keywords: HFC, access networks, ATM switching, QoS

1. INTRODUCTION

Today, HFC networks, are mostly used to broadcast traditional cable TV programs. Because of their ability to support high bandwidth communication and the large percentage of homes reached in many countries they will possibly play a major role in future broadband access network development. In[1] bottlenecks in the network in case of high load are pointed out. These bottlenecks can result in performance degradation and problems with QoS requirements of networked multimedia applications. To improve performance and eliminate the bottleneck, the approach presented in this paper suggests the replacement of passive network elements like fiber nodes and signal amplifiers by ATM switches. This approach has been evaluated with respect to various network topologies and user scenarios by simulation. This paper presents the simulation model that has been developed at the Technical University of Braunschweig and discusses the simulation results obtained.

The paper is structured as follows. Section 2 briefly outlines the structure of HFC networks and the usage of ATM switches within an HFC network. Some basic parameters for the simulation model are extracted with respect to the network structure and the different communication layers and components involved. The simulation model is outlined in section 3. Section 4 discusses traffic models and network parameters that has been used in simulation runs. Simulations results are discussed in section 5. Finally, section 6 summarizes the paper and gives an outlook on further activities.

2. BASIC MODELLING PARAMETERS

This section describes the structure and the MAC layer of HFC networks. Modelling parameters are extracted that are summarized in section 2.3.

2.1. Structure of an HFC network

HFC networks are based on a tree-structure as depicted in Figure 1. The root of the tree is formed by a so-called head end. The branches leaving the head end lead to fiber nodes. Fiber trunks are used for the connection between head end and fiber node. The fiber nodes can be roots of subtrees in the HFC network (cf., Figure 1). These subtrees are located in the coax domain, i.e., coax cables are used for interconnection inside the trees. The leaves of all trees are formed by so-called subscriber stations in the private home.

HFC networks are typically dimensioned as follows. About 10 to 50 fiber nodes are served by a single head end. A fiber node serves around 500 to 2000 subscriber stations. The maximum distance between the head end and any

Send correspondence to Urs Thürmann, thuerman@ibr.cs.tu-bs.de

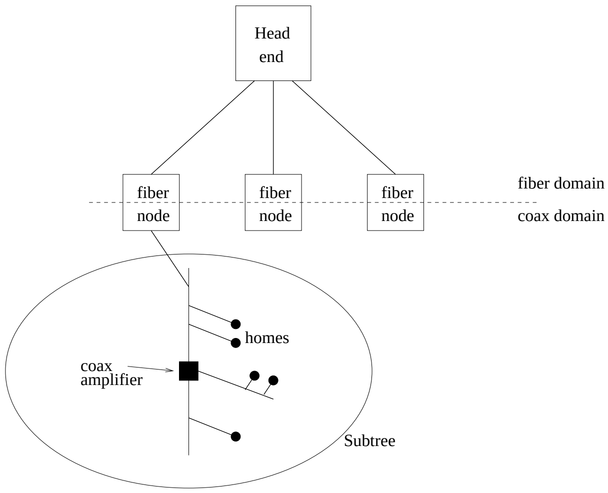


Figure 1. Basic Structure of an HFC Network

subscriber station is about 80 km. The maximum propagation delay between head end and subscriber station is 400 μs.*

With respect to a simulation model for an HFC network, the parameters n and m are important. With n reflecting the number of fiber nodes per head end and m representing the number of subscriber stations per fiber node. Moreover, the propagation delay d needs to be considered.

In a traditional HFC network all traffic flows through the head end. Therefore, the head end may easily become the performance bottleneck of the network since it concentrates all data flowing through the network. However, since the head end can listen to all the traffic, it is the potential candidate for collision detection and bandwidth control. The head end provides a bandwidth controller that schedules upstream transmission of the subscriber stations. Suited parameters and algorithms for the bandwidth controller are needed in order to simulate an HFC network. Furthermore, the collision detection algorithm needs to be included into the simulation model. Parameters for these components are derived below (cf., section 2.3).

In order to reduce the traffic flowing through the head end and, thus, to decrease the head end's load, ATM switching elements or ATM transceivers can be incorporated into HFC networks. They can replace fiber nodes and coax amplifiers.

The use of switches enables the forwarding of ATM cells among subscriber stations without passing the head end. Therefore, collisions are limited to the path between the subscriber station and the first switch in the tree. Moreover, the central queueing in the head end is distributed among the introduced ATM switches.

In contrast to ATM switches, which are connected by point-to-point links, ATM transceivers are connected through a ring. As a result, the transceivers together form a distributed ATM switch.

The use of switches in HFC networks leads to additional parameters that need to be addressed by the model. The depth of the tree is one parameter. The number of branches at each node is another parameter. Furthermore, the depth of the tree between the head end and the first switch needs to be considered. A simplification is to assume the same depth in all branches of the head end. The deployment of switches introduces further modelling parameters, namely the processing delay and the switch capacity.

2.2. MAC layer

The physical structure and the cable lengths in an HFC network make standard media access methods used in IEEE LANs inappropriate for HFC networks. Subscriber stations send on upstream channels and listen on the downstream channel. They can not listen to upstream traffic generated by other stations. Therefore, a subscriber station is not able to detect a collision on the upstream channel. Thus, the head end has to be involved in collision detection. The head end recognizes collisions by checking the CRC field of upstream messages that are subject to collision. It

*The 802.14 draft specifies up to 200 μs which contradicts to cable lengths of 80 km

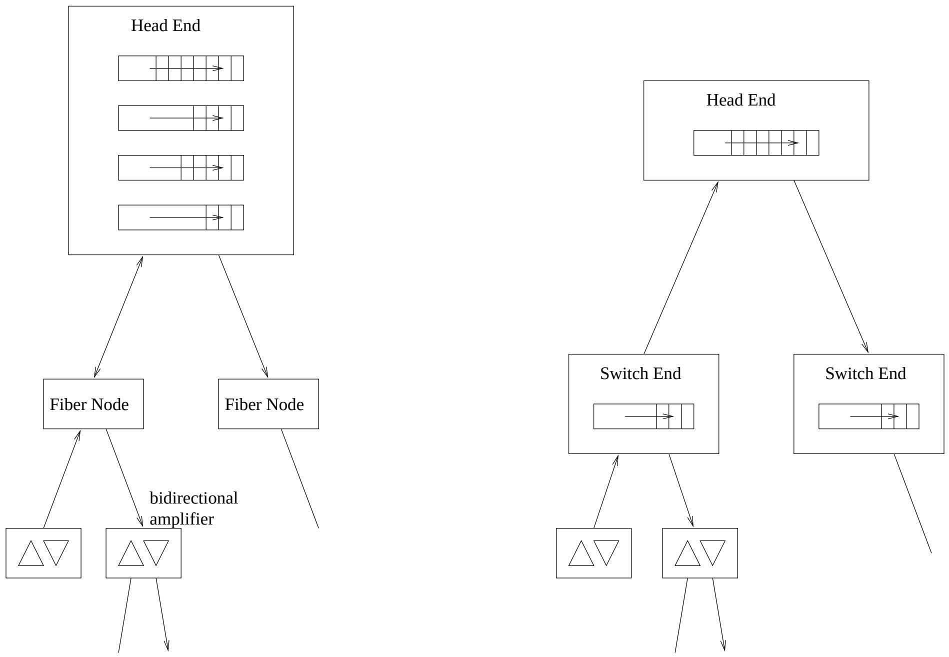


Figure 2. Modelling of an HFC network

forwards collision information on the downstream channel. The MAC layer, especially the mechanisms in the head end, need to be considered in appropriate detail in the simulation model. Important issues are discussed below.

2.2.1. Upstream PDUs

Data in upstream direction is sent in minislots. Each minislot contains 8 bytes of data. They are the basic transmission units. Based on the definition of minislots, three different types of PDUs are specified:

- RMPDU (Request Minislot PDU).

  RMPDUs are currently the only PDUs defined that consist of exactly one minislot. RMPDUs are used by a subscriber station to contend for upstream allocation.

- APDU (ATM PDU).

  An APDU consists of a number of minislots that is needed to carry an ATM cell. The format of the cell header is modified for the HFC network. It contains an Extended Bandwidth Request field (EBR) that can be used to piggyback upstream allocation requests.

- VLPDU (Variable length packet fragment PDU).

  Besides constant-length ATM cells, variable length fragments are supported. VLPDUs are optional. Since this document focusses on ATM switching they are not considered further.

RMPDUs are the only PDUs that are not allocated to a specific subscriber station and, therefore, are subject to contention.

Minislots marked as APDU (or VLPDU) minislots by the head end are assigned to a particular subscriber station. No station except the indicated station may send in such a slot, hence no collision can occur.

The head end contains a bandwidth controller that receives the transmission requests sent by subscriber stations. These requests can contain the information that a particular station has traffic pending and can, futhermore, inform the head end about the number of ATM cells in the station's waiting queue. Based on current traffic needs and negotiated QoS parameters, the bandwidth controller assigns APDU slots to the stations.

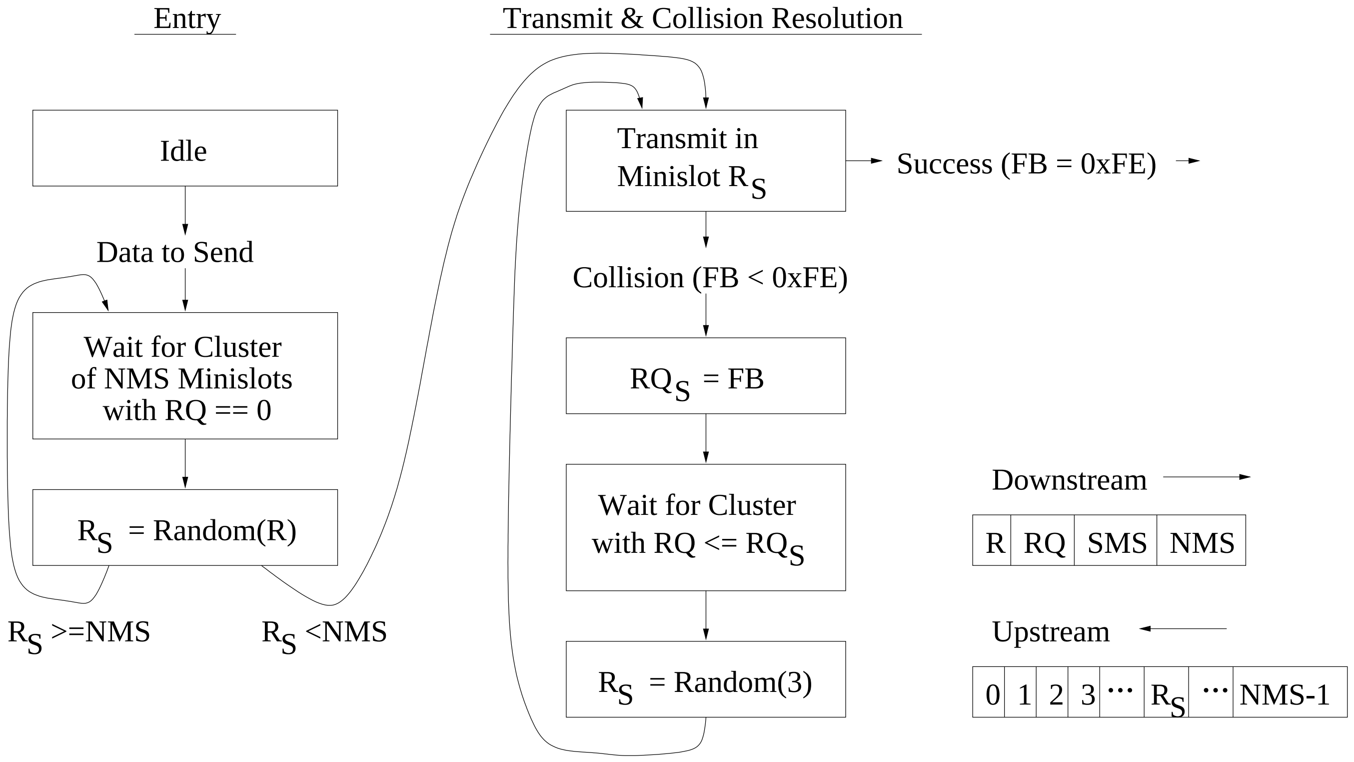


Figure 3. Collision resolution according to 802.14[2]

2.2.2. Downstream PDUs

The downstream channel is composed of data units of 6 bytes. These units can contain the following information:

- Idle pattern.

  The idle pattern is one unit in size, i.e, 6 bytes and is sent when no other traffic has to be send on the downstream channel.

- APDU.

  APDUs carry ATM cells from the head end to one or more subscriber stations. They have the same format as upstream APDUs except that the source local identifier (SLI) is replaced by a destination local identifier (DLI) and the EBR field is unused.

- VLPDU.

  VLPDUs are optional and are not considered further in this document.

MAC information from the head end to a station is encapsulated in an AAL5 frame and sent to the station using APDUs. The MAC information elements considered by the simulation model are:

- Request Minislot Allocation Elements specify upstream minislots that can be used for bandwidth requests.
- APDU grants that inform a station of an upstream APDU assigned to it.
- Contention Feedback messages to inform stations about the collision status of RMPDUs.

n	Number of fiber nodes
m	Number of subscriber stations per fiber node
h	Depth of the tree
h_S	Depth of the tree from head end to first switch
b	Number of branches at each node
D	Bandwidth per channel
q	Number of queues in subscriber stations
T_S, T_{SE}, T_{SS}	Processing time in switches, switch end, and subscriber stations
	Switch capacity
d_S	Propagation delay between the switches
d_T	Propagation delay within the subtree

Table 1. Parameters of an HFC network

2.2.3. Collision Resolution with a Ternary Tree Algorithm

The 802.14 draft describes a ternary tree algorithm to be used for collision resolution of upstream request minislots (Figure 3).

The head end regularly allocates clusters of minislots for the subscriber stations to send bandwidth requests. These slots are contention based slots and, therefore, some method of collision resolution is needed. Each cluster of minislots available for bandwidth request has assigned a request number (RQ). Clusters with $RQ = 0$ have an additional entry persistance R. By varying the value of R in a cluster with $RQ = 0$, the head end has some control over the probability of collisions in that cluster. Stations entering the contention process may only send in clusters with $RQ = 0$. They randomly select a value R_S with $0 \leq R_S < R$. If R_S is less than the number of minislots (NMS) in the cluster, the station sends its request in minislot number R_S. Otherwise, it has to wait for the next cluster with $RQ = 0$.

Request minislot feedback elements are sent by the head end to inform stations of the contention status of each request minislot processed. A feedback element has an 8 bit feedback field (FB) for each request minislot. If the transmission was successful ($FB = 0xFE$) the station waits for an APDU grant and transmits the data. In case of a collision the head end indicates this by sending $FB < 0xFE$. The station copies this feedback value FB into RQ_S and can then use any cluster with $RQ \leq RQ_S$ to retransmit its request. Clusters with $RQ \neq 0$ always contain 3 minislots that are open or contention. The station randomly selects one of these 3 minislots to send its request.

2.3. Summary

In this section the basic structure of an HFC network, the MAC layer and the contention algorithm in the subscriber station were described. Various modelling parameters were extracted that are summarized in table 1.

The components of the simulation model are identified:

- Head end
- ATM switches
- Switch ends
- Subscriber stations

3. SIMULATION MODEL

The simulation model described below is designed for event-oriented simulation. Figure 4 shows the structure of the model. The head end is connected to a switched network by fiber links. The top level switches are fiber nodes that convert from optical to electrical signal and vice versa. Beyond the fiber nodes switches are connected by coax cables. The switches can be built from ATM transceivers in which case they form distributed switches connected by rings. Different switch topologies will be considered.

The border of the switched network is at the so-called switch ends that connect the shared medium subtrees to the switched network (cf., figure 4). Access to the shared medium in the subtree is controlled by the IEEE 802.14 protocol. The subscriber stations are located at the leaves of the subtree.

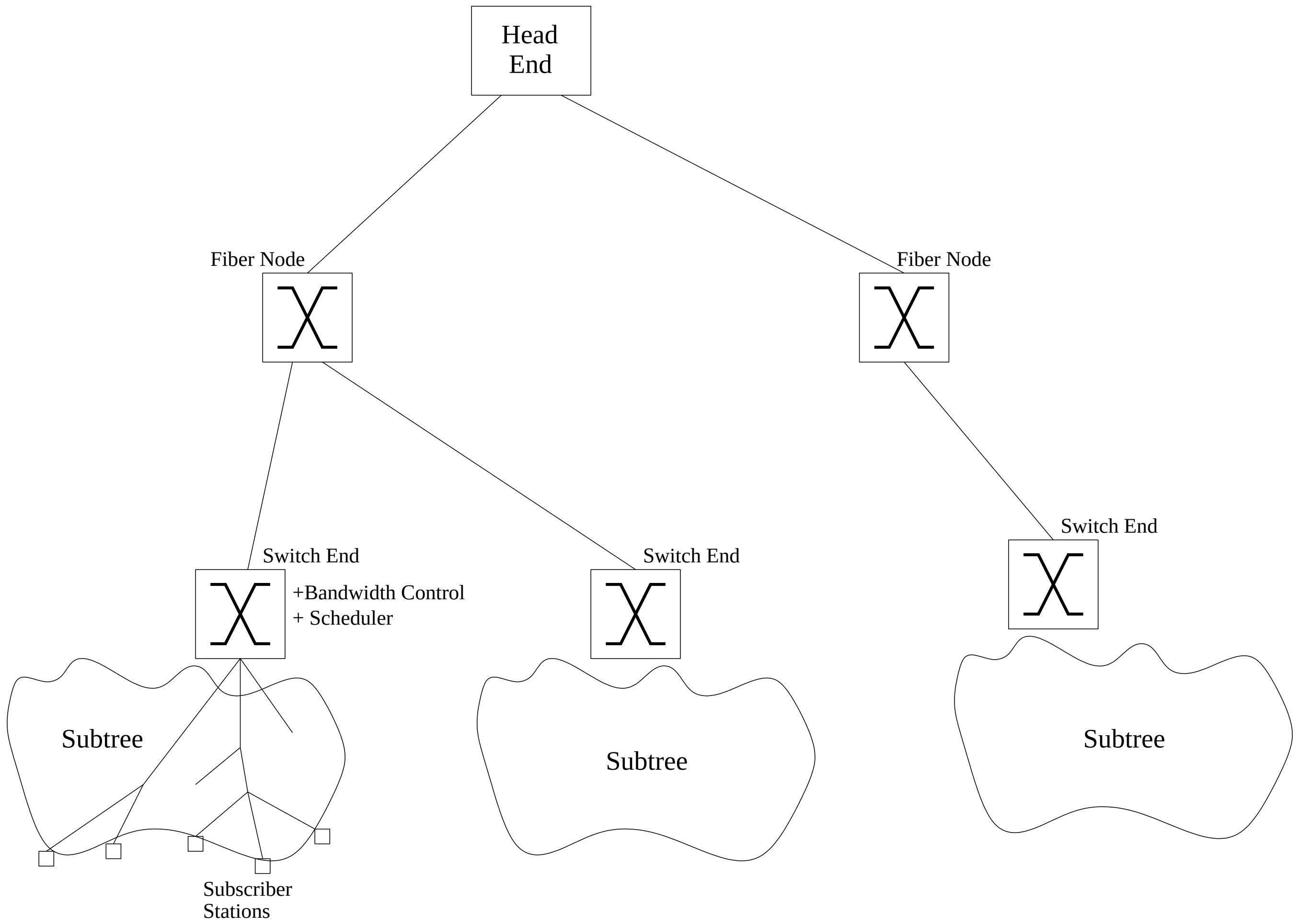


Figure 4. Basic structure of the simulation model

3.1. Subtree Module

The simulation model is modular, i.e., it consists of separate components that can be combined to form a complete model suitable for the simulation. The module described here models the subtree rooted at a switch end, in which no further switches are located. This module forms the components of figure 4 labeled "subtree". The internal structure of the module subtree is shown in figure 5. The main aspect of this module is collision and collision resolution.

A simplification introduced in the model is that different types of messages between the switch end and the subscriber stations are exchanged on separate lines. This avoids analyzing incoming PDUs in order to determine their type.

In the model, not all messages exchanged between the switch end and the subscriber station are subject to the propagation delay of the physical connection. Instead of sending RMPDU allocations and APDU grants in advance, specifying the minislots to be used for RMPDUs or APDUs, they are delayed. Allocations and grants are sent at the time the according upstream minislots appear. Furthermore, they only contain the number of minislots allocated but not the starting minislot number, i.e., the number of the first minislot allocated. This releases the subscriber station completely from handling minislot numbers. It does not have to store the allocation, count minislot time units and wait for the minislot allocated. Instead, the event that causes the station to send is generated in the switch end. The station responds immediately, i.e., in the same simulation time, with a bandwidth request or APDU, respectively. The overall behavior of the model remains compliant to the draft standard.

The following variables are defined that describe the model's state at time t:

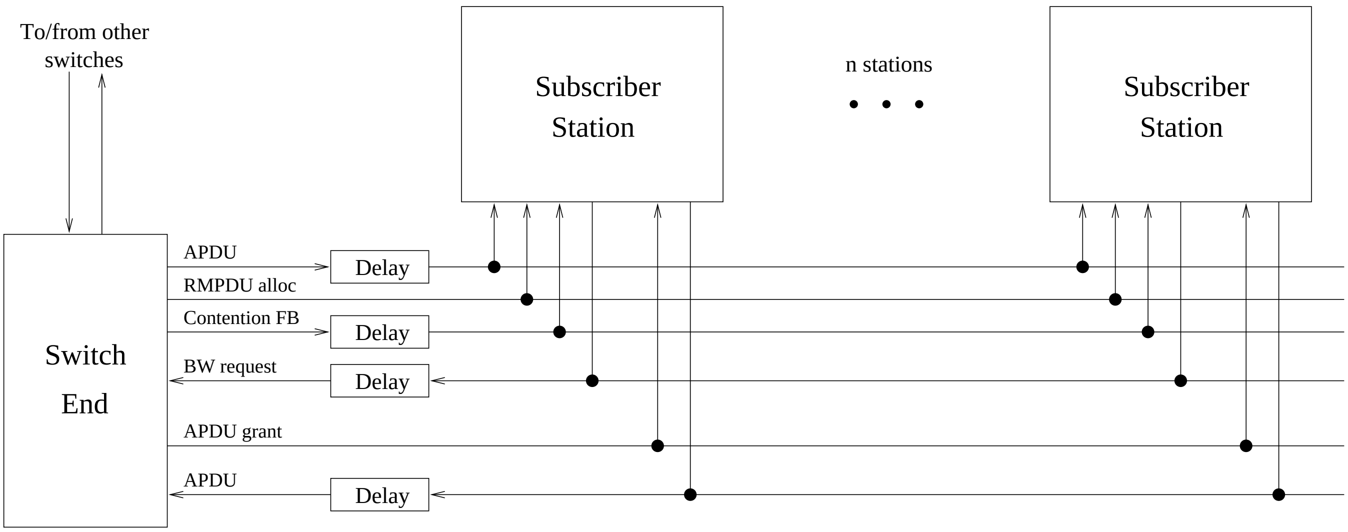


Figure 5. Model of the subtree

$E(t)$	Ordered list of events
$L(s,t)$	Length of the output queue of station s
$R(s,t)$	Number of already requested APDUs (as committed by the switch end)
$C(t)$	Set of stations that have entered the contention process
$RQ(s,t)$	RQ register of station s.

The following events can occur:

- A station s generates new data for transmission. In this case the length of the queue is incremented by one, and, if it was empty, the station is added to the set of contending stations.

  $L(s,t+) = L(s,t) + 1$
  $RQ(s,t+) = 0$
  $C(t+) = C(t) \cup \{s\}$ if $L(s,t) = 0$
  $E(t+) = E(t) + \text{arrival of next packet}$

- The switch end schedules a cluster of request minislots with values RQ and R.

  $S = \{s \in C(t) : RQ \leq RQ(s,t)\}$ Set of stations that are allowed to use these minislots

  A random map r and sets S_i are created

  $$r : S \to \mathbf{N}_0 \text{ with } r(s) = \begin{cases} \text{Random}(R), & \text{if } RQ(s,t) = 0 \\ \text{Random}(3), & \text{otherwise} \end{cases}$$
  $S_i = \{s \in S : r(s) = i\}$

  For all $i = 0, \ldots, n-1$

  1. if $|S_i| = 0$.
     No station tried to send in this request slot. Nothing to be done.
  2. if $|S_i| = 1$.
     Exactly one station did send its bandwidth request in this minislot so it was successfully received at the switch end.
     $R(s,t+) = R(s,t) + \text{number of requested APDUs}$
     $C(t+) = C(t) - \{s\}$
  3. if $|S_i| > 1$.
     A collision occured. The switch end sends a feedback value FB that the station uses for the next cluster of request slots.
     $RQ(s,t+) = FB$

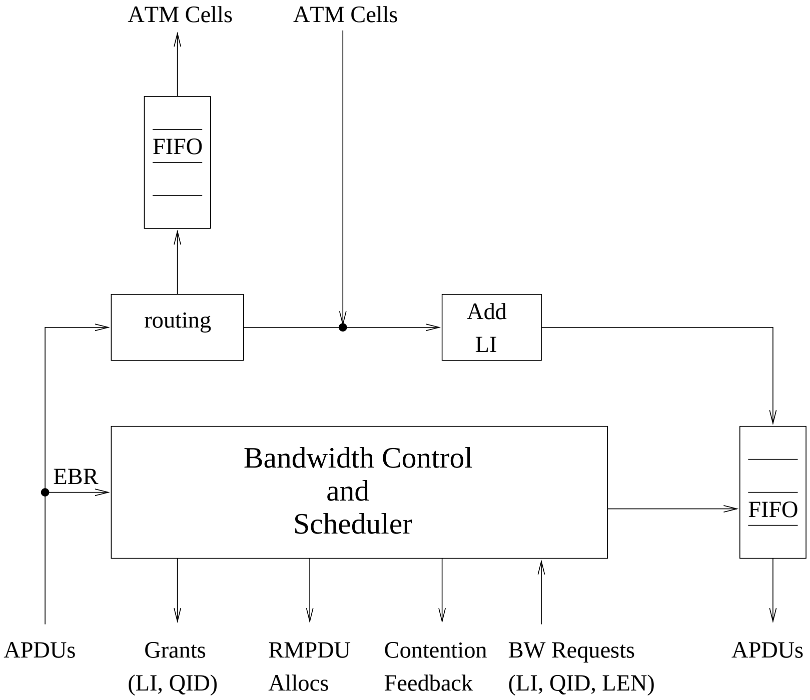


Figure 6. Model of the switch end

- The switch end schedules an APDU for station s to transmit one ATM cell. If the station has cells in its output queue for which it has not yet sent bandwidth requests to the switch end, i.e., $R(s,t) < L(s,t)$, it does so using the EBR field of the APDU.
  $L(s,t+) = L(s,t) - 1$
  $R(s,t+) = R(s,t) + (L(s,t) - R(s,t)) = L(s,t)$
  $E(t+) = E(t) +$ arrival of APDU at the switch end after delay d_T
- The switch end receives an APDU from a station containing an ATM cell. The switch end determines the destination of the cell using the SLI and VCI fields in the APDU. The destination can be another station in the same HFC network or an ATM switch connecting the switch end to other networks.
  $E(t+) = E(t) +$ arrival of APDU at destination after delay $T_{SE} + d_T$
- A station s receives an APDU from the switch end, containing an ATM cell. The cell is forwarded to the destination application where it is processed.

The simulation model described above is being implemented at the Technical University of Braunschweig using the Bones/Designer simulation software. The next two subsections briefly describe the implementation of the head end model the the subscriber station model.

3.2. Switch End

The simulation model of the switch end is shown in figure 6. ATM cells arriving at the switch end from some other ATM switch have to be forwarded to a station in the HFC network. The switch end determines the local identifier (LI) for the destination station and creates an APDU containing the cell. This APDU is then placed into a FIFO queue to be sent downstream to the station. Since the switch end has exclusive access to the downstream channel no collisions can occur and cells can be sent from the FIFO without having to wait for the medium.

On a separate connection the switch end informs stations about minislots that can be used for contention based bandwidth requests. For each cluster a message is sent. The stations decide if and in which minislot they request bandwidth. Requests are sent on the connenction labeled "BW Requests". On the line labeled "Contention Feedback" the switch end informs the station if their requests have succeeded or collided with other requests. In response to

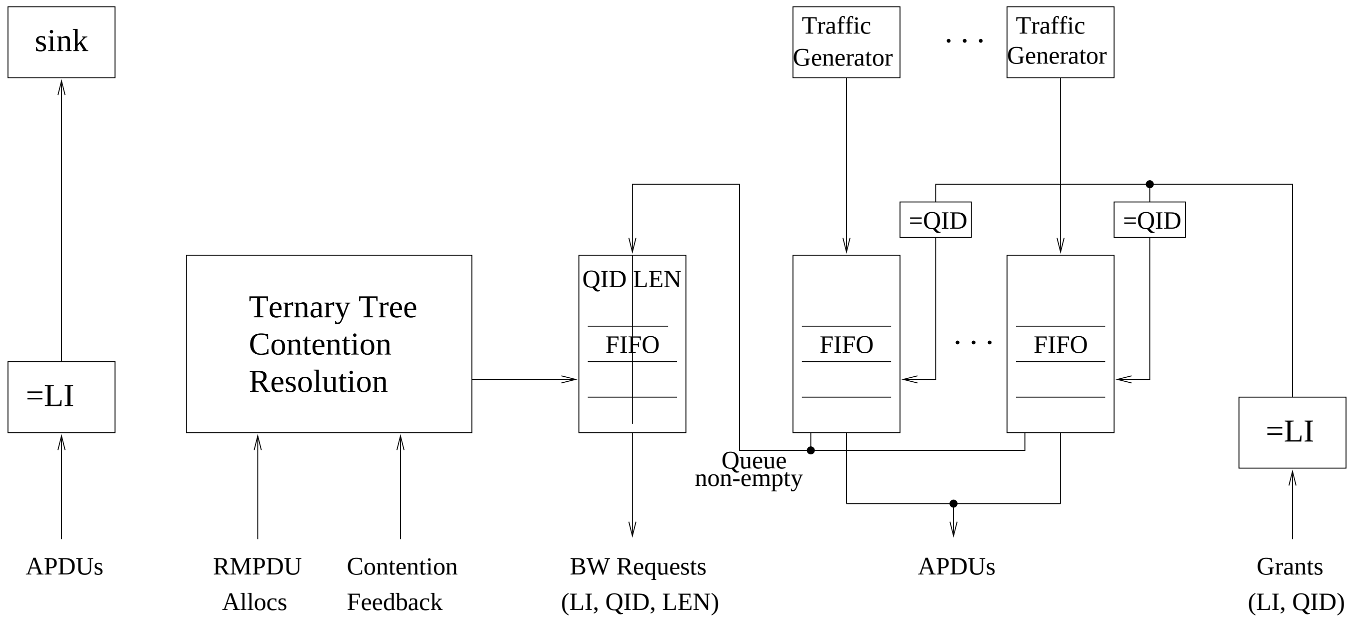


Figure 7. Model of a subscriber station

successful bandwidth requests the switch end sends an APDU grant message to the station. It is transmitted on the connection named "Grants". The station in return immediately sends a pending APDU to the switch end that forwards the cell appropriately.

3.3. Subscriber Station

Figure 7 shows the module of the subscriber station. The subscriber station connects to one or more traffic generators. Traffic of different classes with different QoS requirements can be placed into the same or into separate output queues. For the time being, FIFOs are applied in the model.

If the queue was empty before the new entry was queued a bandwidth request is placed into the request queue. Bandwidth requests in this queue consist of the queue identifier (QID) and the number of APDUs requested (LEN). If the station has entries in the request queue it waits for an appropriate RMPDU allocation message from the switch end, randomly selects a minislot, and causes the FIFO to emit its first entry without removing it. If the contention feedback from the switch end indicates a successful request, the first entry is removed from the FIFO. Otherwise, the station waits for the next RMPDU allocation.

If the station receives an APDU grant it compares the LI field of the grant with its own local identifier. If they match it forwards the grant to one of the output queues according to the queue identifier (QI) field in the grant. An APDU is created from the first entry in the output queue with the EBR field set to the number of entries that have not yet been requested bandwidth for.

Incoming APDUs that are addressed to the station are forwarded to a traffic sink.

4. TRAFFIC MODELS AND SIMULATION PARAMETERS

4.1. Traffic Parameters

Several categories of applications which require different types of QoS from the underlying network can be identified, dependent on their requirements with respect to bandwidth, delay and delay jitter. Some applications, such as telephony, require constant, i.e., isochronous data streams, while other applications, e.g., video games, typically produce very bursty traffic patterns.

While WWW accounts for the majority of the traffic in todays residential environment future networks scenarios will be quite different. New services will partly replace services as they are seen in the residential environment today.

In[3] a forecast of new services and their ranking is given. The numbers are obtained by a Delphi survey. Nearly 400 questions were submitted to experts in two rounds. Table 2 shows the results of the question which applications are expected to have the greatest number of subscribers. The numbers indicate how often the application was selected by an expert.

Application	Times selected
Video on demand	41
Home office, simple	34
Videotelephony	27
Remote education, home	11
Multimedia telegrams	10
Home office, advanced	6
Interactive TV/specialized channels	6
Electronic newspapers	4
Advertising and marketing	1
Telecommunity/telemedicine	–
Remote education, studio	–

Table 2. Ranking of services on wideband connections

Many of the services shown in table 2 heavily depend on real-time multimedia communication. The QoS of audio and video data streams achievable in HFC networks is to be investigated. In addition, in the near future WWW will still play a major role in network traffic patterns.

The simulation has, therefore, been performed with a mix of different traffic generators, each one representing a traffic source of a specific kind.

Results of traffic measurements are described in[4] for WWW traffic and in[5] for MPEG video traffic. Audio sources coded with PCM, ADPCM or GSM produce constant bit rates. The following traffic generators have been implemented and used in the simulations:

- CBR
  This traffic generator is parameterized by the rate r of the CBR traffic flow to be generated. It produces cells at fixed intervalls of $\frac{48*8}{r}$.
- WWW
  The WWW traffic generator models WWW requests as described in.[4] It produces bursts of packets where each burst is described by its duration and the interarrival time of the burst's packets. Both, duration and interarrival time obey a Weibull distribution. The OFF period between the bursts is Pareto distributed. The parameters for the Weibull and Pareto distributed random generators are fixed and set according to.[4] The WWW traffic generator has only one parameter specifying the number of clients producing WWW requests, that are to be modelled.

In the simulations traffic flows of different kinds have been generated by the subscriber stations. For each kind of traffic the number of generators have been varied to test the networks under different load conditions.

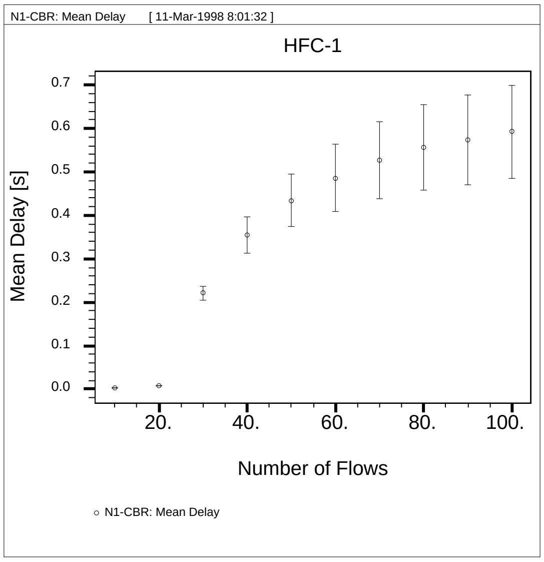
4.2. Network Parameters

Simulation experiments have been conducted with different configurations of HFC networks. All configurations included the same number of subscriber stations.

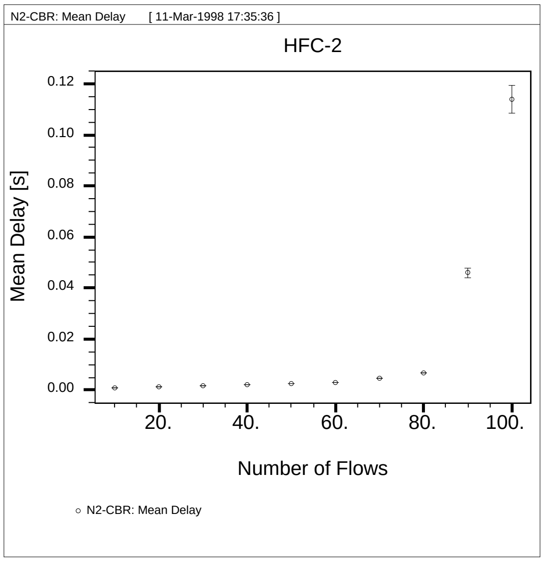
The configurations differ in the size of subtrees and the topology of the ATM switch network.

All networks investigated consist of 1024 subscriber stations. The following configurations of networks have been simulated:

1. HFC-1
   HFC network containing a single subtree of 1024 station, i.e., no ATM switches and transceivers are involved.



(a) HFC-1: 1024 stations



(b) HFC-2: 4 × 256 stations

Figure 8. Mean Delay of CBR traffic in HFC networks

2. HFC-2
   HFC network consisting of 4 subtrees, each one containing 256 stations. The 4 subtrees are connected by a ring of ATM transceivers.

3. HFC-3
   HFC network, similar to the previous one, consisting of 4 subtrees, but the subtrees a connected by a tree structured ATM network following the traditional structure of HFC networks.

Parameters that had to be selected but were not varied between these configurations, include link lengths, upstream and downstream data rates of the HFC subtrees, and processing delays.

5. DISCUSSION OF SIMULATION RESULTS

Figures 8(a)-8(b) compare the configurations 1 and 2 listed above, i.e., a plain HFC network without any switching and an HFC network in which fiber nodes have been replaced by ATM transceivers. The figures show the mean delay of cells sent between two subscriber stations for a number of 64 kbps CBR traffic flows. The X-axis shows the number of traffic flows generated and thus corresponds to the network load.

As can be clearly seen the introduction of ATM transceivers into the HFC network improves its performance significantly. Up to a number of about 20 CBR traffic flows, the HFC-1 network without switches yields a mean delay of a few milliseconds with little variations. Increasing the number of traffic generators raises the delay to 600 ms with 100 ms jitter for 100 traffic flows. This is unacceptable for many applications and results in poor quality for bidirectional audio conversations, for example. The maximum number of CBR traffic flows of 64 kbps than can be sent through the 2 Mbps upstream channel without collisions is 31.

Figure 8(b) shows that in the case of the HFC-2 network 80 CBR flows can be supported, before the mean delay rises drastically, i.e., about 4 times as many as in the HFC-1 network.

The effect of the ATM switching topology is small compared to performance gain by replacing fiber nodes with switches (cf. figure 9). Delay times are of the same order of magnitude. However, with a tree structured ATM switch network, increase of delay and jitter starts at a lower number of CBR connections.

This can be explained by observing that the switch at the root of the tree is passed by one half of the traffic flows, assuming uniformly distributed source and destination stations, and this traffic has to be forwared onto 2 links. In

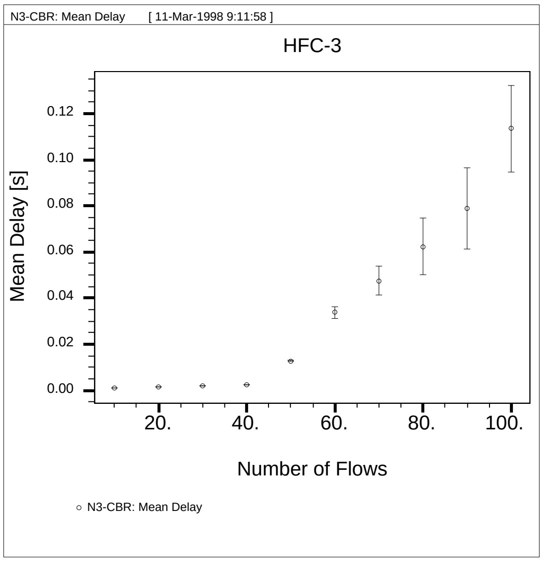


Figure 9. Mean Delay of CBR traffic for HFC-3

the case of HFC-2 each switch is passed by one half of the traffic, too, but these flows are distributed over 3 links. Also, in HFC-3, switches next to the HFC switch ends experience a higher load than switches in configuration HFC-2. Again, assuming uniformly distributed source and destination nodes, these switches are passed by 5/8 of the traffic.

6. SUMMARY AND OUTLOOK

Compared to other access network technologies, HFC provides high bandwidth in both, the upstream and downstream channels. However, the bandwidth available has to be shared among all active subscriber stations. Furthermore, as a consequence of physical characteristics and the MAC layer protocol, collisions can occur in the upstream channel. The head end concentrates all data flowing through the network and may easily become a bottleneck in case of high load. Introduction of ATM switching elements into network, replacing fiber nodes and coax amplifiers, can eliminate this single point of traffic concentration. The simulation results show that achievable quality of service can be improved with this approach.

Ongoing research will involve other access network technologies. ADSL is one of HFC's competitors in this market. Comparisons between HFC and ADSL, concerning the QoS they can deliver to applications, will be focus the of our near future's work.

REFERENCES

1. U. Thürmann, M. Zitterbart, and T. Meuser, "ATM-Switching in HFC-Verteilnetzen," in *Proc. Workshop Kommunikationstechnik*, Technische Universität Hamburg-Harburg, 10.–12. September 1997.
2. *IEEE Std. 802.14 Draft1R3*, May 1 1997.
3. K. Stohrdahl and E. Murphy, "Forecasting long-term demand for services in the residential market," *IEEE Communications Magazine* **February**, 1995.
4. S. Deng, "Empirical model of WWW document arrivals at access link," in *IEEE ICC'96*, 1996.
5. O. Rose, "Simple and efficient models for variable bit rate MPEG video traffic," *Performance Evaluation* **30**, pp. 69–85, 1997.