

Performance Evaluation of Diffserv Driven HFC System

Giannis Pikrammenos, Helen-Catherine Leligou

National Technical University of Athens
Greece

Abstract. Shared access networks such as hybrid fiber/coaxial and passive optical networks have emerged as promising ways to reduce the cost of the transition to a broadband access infrastructure and provide a graceful upgrade path towards the photonization of the local loop. The MAC protocol as the only arbiter of the upstream bandwidth directly affects the Quality of Service (QoS) provided to each upstream traffic flow and must meet several constraints. Such constraints include the adequate speed of operation, flexibility to support efficiently the largest number of services and applications offering an adequate number of QoS classes, and independence of higher layers, protocols and future extensions to traffic management specifications. The implementation of a MAC mechanism targeting these goals and aligned to the emerging Differentiated Services Internet strategy is evaluated using computer simulation results in this paper. The benefits of the adopted prioritization scheme to provide and guarantee different QoS levels are illustrated.

Keywords: Interworking, HFC, ATM, IP, DiffServ, QoS, Access Network.

1. Introduction

Re-use of the existing infrastructure greatly reduces the initial investment outlay and provides a graceful upgrade path in step with service demand. Typical examples are Hybrid Fiber Coaxial (HFC) that use legacy CATV systems and their coaxial medium beyond the fiber node and Passive Optical Network (PON) systems that use the twisted pair based telephone networks enhanced with xDSL beyond the Optical Network Unit (ONU). The legacy topology is that of a broadcast tree-shaped access network. The attractiveness of HFC systems for the delivery of broadband services lies in the high reuse of existing infrastructure and a sound gradual upgrade strategy. The initial investment mainly consists in the incorporation of a return channel and fiber feeders to customer clusters of a few hundred homes. Sharing the common feeder and cable in the upstream direction (from stations to head-end) requires a MAC protocol to allocate slots to stations in a TDMA fashion [5], [6].

This sharing results in a distributed queuing system characterized by the long time required to pass control information from the queuing points to the service controller residing at the head-end. The allocation of upstream slots in a tree-topology access system is based on a reservation method which allows to dynamically adapt the bandwidth distribution to traffic fluctuations. The MAC controller works on the basis

of collected access requests to allocate the upstream slots by sending access permits. It is important to note that the service policy of the MAC governs the distributed multiplexing from a central point situated at the head-end. This has the important implication that, as regards delay control, the acquisition of arrival information and buffer fill levels includes a considerable delay element not found in the centralized multiplexing typically operating in a switch queuing point. In contrast, drop policies must be distributed over all network terminations (cable modems) where the flow identity and fill levels are known. Because of the much larger reservation delay and statistical behaviour of the aggregations from many customers, special care must be taken to safeguard QoS to sensitive traffic. This requires a prioritization scheme and the differentiated services strategy of IETF [1] is a very suitable approach to handle the problem.

The paper is partitioned in two sections. The concept of operation of the MAC protocol and the actual implementation is presented in section 1, while in section 2 the performance is evaluated using computer simulation.

2. Presentation of the MAC Concept – Integration Aspects

Tree-shaped access system exhibits a very different behaviour between the upstream and the downstream direction. In the downstream, replicas of the signal are created due to the broadcast nature of the medium, giving rise to privacy and security issues, which are typically dealt with by means of encryption. In the upstream, the MAC control function implies the multiplexing and concentration of the traffic from all active modems. However, in the upstream case the MAC is characterized by the distributed nature of the queuing points and the additional difficulties in the exchange of control information. To apply any scheduling or priority discipline requires the correlation of the traffic from all multiplexed sources going to the one common egress point of the system.

The Differentiated Services (DS) strategy, recently adopted by IETF as a scalable and relatively simple methodology towards enriching with QoS the IP services, is applicable and quite appropriate in the case of such tree-shaped access systems, like HFC, where IP services are or could be dominant. To align such an access system to the concept of differentiated services requires the incorporation of priorities in the MAC function, for the appropriate handling of each flow aggregation with respect to its requirements. The strength of DS lies in the slow and graceful introduction of such complexities in line with revenues from a previous stage of introduction of such mechanisms. Dealing with behaviour aggregates and starting with static management based Service Level Agreements (SLAs) executed at slow time scales while keeping traffic conditioning at the edges of the network enables a low cost starting phase while smaller granularity levels can be sought out at later stages of deployment. Non-compliant intermediate nodes can be transparent but at risk of reduced overall performance should they become the bottleneck in the route of the flow. Slow distributed access multiplexer such as an HFC system residing at the network edge can not be relied upon to operate in a transparent fashion as regards the DS strategy since the MAC directly affects the temporal properties of the egress stream. This can

not be realized in an HFC system without embedding suitable differentiation mechanisms into the MAC control function.

The basis of the approach is the use of access priorities in the reservation system, which can be programmed to fit with required PHBs by means of the mapping of flows to priorities. [1] Logically separate queues for each priority are necessary for the proper operation of the prioritization scheme.

The characteristics of the four aggregation levels/priorities are the following:

1. The high priority is devoted to services with very strict delay requirements, which undergo strict traffic profile control (traffic conditioning) such as the EF (Expedited Forwarding) service [2].
2. The second priority level is devoted to real-time variable rate flows, such as video services or VoIP and it is provided with peak rate policing for guaranteed QoS. In the DS context it could be used for the top AF (Assured Forwarding) class [1].
3. The third priority is devoted to data services with higher requirements than best-effort. The traffic profile control assumed for this class aims at minimizing the loss of packets and the disturbance to other traffic. The 3rd priority mechanism is suited to the support of all four or the lower three AF classes [3]. (Drop policies can be independently applied at the modem queuing points).
4. The fourth priority is reserved for plain best-effort services which employ loss based flow control at the TCP level and can be very disruptive to the other classes when sharing the same queue.

The implemented system employs a TDMA slotting designed for ATM cells. We will simply consider the ATM size slots as a quantum of MAC assigned bandwidth allocation. The last three priorities employ reservation while the first unsolicited permits.

The reservations in the HFC MAC use a request field piggy-backed in the upstream slot and are usually assisted by contention on special reservation mini-slots [4]. The latter is needed for the announcement of the first arrival of a burst since the piggy-back mechanism is not self-starting but relies on the existence of previous traffic for the announcement of new arrivals [6]. However, the implemented MAC of the under discussion system departs from this approach in two points: first it provides a three simultaneous request field – one for each queue – which is embedded in every upstream slot irrespective of which queue provided the cell. Second, it employs round robin polling instead of contention. That was chosen for its simplicity and came without serious drawbacks in the targeted environment. In the AROMA system the one byte request field uses the place of the HEC field of the ATM cell which is not needed for cell delineation since an additional synchronization preamble is employed because of the burst mode operation. The multiple requests on the other hand are the strong point of the AROMA system and the tool for the higher QoS capabilities of this system. They are necessary if higher priority traffic is to be quickly made known to the head-end. This feature enables the algorithm in the MAC to offer precedence to the high priority cells when issuing access permits on a global basis and not just among the cells of the same termination which is of a very limited value.

The described MAC function was implemented on the Access Network Adaptation board of the under discussion system. Since the MAC considered a sub-layer of the physical layer and is closely related to the physical layer functions (TDMA, framing, etc.) the permit generation function was selected to be integrated with the downstream

framer component, which was responsible for the construction of the downstream frame including the transmission permits. The whole design was programmed and placed on a Field Programmable Gate Array (FPGA) chip enabling the re-design of the implementation to adjust to modified frame structures, rates etc. An embedded processor (On Board Controller) was used to calculate and modify all non-real time parameters such as the programmed (pre-allocated or provisioned) bandwidth distribution, which may vary with time due to the switch on/off of the modems. An external static RAM chip was used for buffering necessary protocol information in addition to the available on-chip memory provided by the FPGA, which was exploited to keep state information and speed-up the permit generation process.

In the MAC operation algorithm, a list of 512 permits is employed, which has been prepared by the OBC on the basis of subscription data for the scheduling of all pre-arbitrated permits. The OBC executes the Call Acceptance Control (CAC) scheme (or Service Level Agreement in the DS terminology) and writes periodically permits for the 1st and 2nd priority connections at the peak and the minimum guaranteed rate accordingly (to support the expedited and assured forwarding classes). Permits are also inserted for polling requests by modems that have not established a 1st or 2nd priority connection. This list is cyclically read out by the MAC controller H/W and the permits are sent downstream. Permits are spaced in the list according with techniques given in [6]. This list is stored on the external SRAM chip. At the end of each cycle, the embedded controller updates also a list of credits for the 2nd and 3rd priority connections. These credits are used for policing 2nd priority connections at the contracted peak rate and guaranteeing the minimum calculated bandwidth of 3rd priority connections as described above. Since ATM signaling is also supported in AROMA, the permit and credit lists can be updated dynamically to add new connections using a second copy.

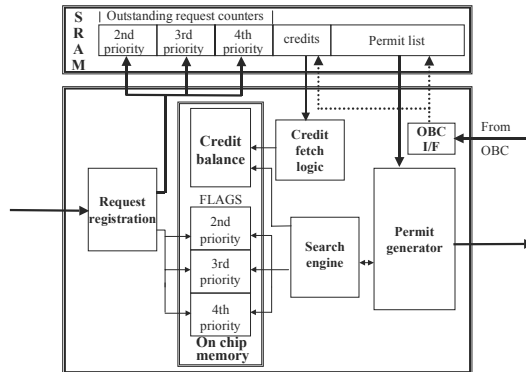


Fig. 1. Operation of the permit scheduler

The three lower priorities are serviced dynamically, on the basis of requests, by filling in empty locations in the permit list (representing unallocated bandwidth). The requests per queue, for each modem, arrive at the MAC controller in the head-end in a format of one byte, with every cell transmitted from the modem (as mentioned above), partitioned into 2, 3 and 3 bit areas for the 2nd, 3rd and 4th priorities respectively. This information is used to update the 3 outstanding request counters per

modem, one for each priority. These counters are holding the total request number for each modem, and are incremented by the priority relevant part of the request byte content for each new cell arrival.

At the same time, to replace the unallocated permits, a search engine scan in a round robin fashion the outstanding request counters and reduces them by one for each permit scheduled. The higher priority counters are inspected first and only if all are empty the same process is repeated for the immediately lower priority. To expedite the process, flags are used to quickly detect and skip all empty locations. The 2nd and 3rd priority dynamically assigned permits are issued provided that the credits are not exhausted. The MAC H/W subtracts the credits as it issues permits and stops serving any modem queue that exceeded its allocated apportionment. This policing action guarantees that 2nd priority malicious users can not disturb complying traffic. As far as for the 3rd priority, the credits guarantee that this class will achieve a better than best-effort service. When credits are exhausted, 3rd priority left-over requests are added to the 4th priority ones, thus a minimum rate is guaranteed but any excess is considered plain best-effort, in accordance with Assured Forwarding rules. The operation of the permit scheduler and the execution of the MAC algorithm are depicted in Figure 1. In the modem side, arrive the permit the cell waiting in the higher priority queue is transmitted.

3. Evaluation of the MAC Algorithm

The behaviour of the above MAC mechanisms was studied with the help of a computer simulation model created with the PTOLEMY tool. PTOLEMY is a graphical simulation environment, where different software entities are represented with graphical objects. The derived topology of the AROMA system is further processed in order to generate executable code, which is to be finally simulated.

The scenarios used 10 cable modems loaded with uniform traffic for each priority. Each source used a common ON-OFF model, generating traffic at the slot/cell level, with on and off periods exponentially distributed. The peak rate was equal to the system rate. For the simulations, IP packets with geometrically distributed lengths were considered and were segmented to fit into slots. There was no limit set to credits, since the policing function would only be meaningful to demonstrate with malicious sources that was out of the scope of these tests, since the results are determinist in this case. The duration of the runs was 1.5 million slots because the PTOLEMY model was heavy and slow to run in the Solaris OS available. The upstream slot duration was used as a time unit for the simulation (170.6 μ s).

Figure 2 depicts the probability distribution function (pdf) of the access delay for the three lower priorities (2nd, 3rd, and 4th) under a total load of 85%. Only the variable part of the delay is shown, as the fixed round trip time (about 4 slots) required for the modem request (upstream) to reach the head-end and the permit response (downstream) to reach the modem is not counted. The 10% of the load were devoted to CBR traffic (or virtual leased lines by EF in the DS model of usage) through the 1st priority queue of the modems, while the other 75% were equally distributed among the other 3 priorities. The 1st priority is not presented since it exhibits deterministic

behaviour, as the transmission permits are pre-programmed in the list [7], and thus, the delay never exceeds the fixed permit distance.

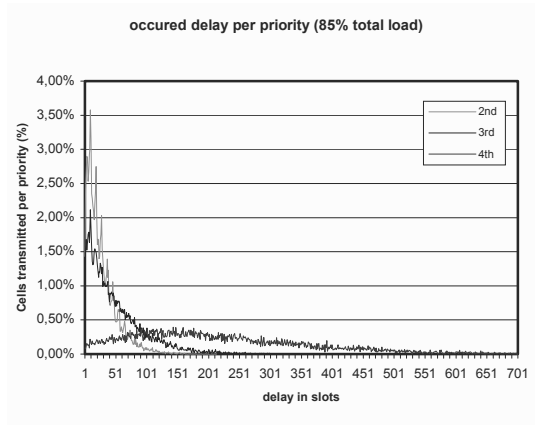


Fig. 2. Pdf of access delay

The most important result caused by the prioritization scheme is the delay advantage provided to the higher priorities. The transmission of almost all the 2nd priority cells delay is less than 250 slots (i.e. 43ms) while those of the 3rd less than 350 slots (60ms). Of course there is no bound for the 4th priority which can exceed any limit depending on the total loading.

Under a higher offered load of 110%, the probability distribution function of delay is shown in Figure 3. The 1st priority occupied 10% of the load while the 2nd and 3rd priority sources were offered a load equal to 30% of capacity each, and 40% for the 4th. As it can be observed from the figure, 2nd and 3rd priorities do not exhibit any significant behaviour difference with the previous scenario, since their service up to the contracted rate is guaranteed. The effect of prioritization is exactly to hide the

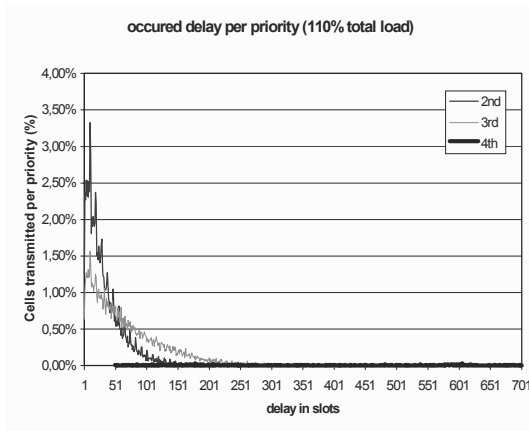


Fig. 3. Pdf of access delay

presence of any lower priority traffic, which is prevented from competing against the protected sensitive traffic. The effect on the 4th priority is very strong since it is throttled down to the 30% of capacity left over from the other three. The occurred delay distribution is thinly spread till theoretically infinite delay. This can not be shown in the simulation results since buffers overflow no matter how much memory is allocated. The average delay is not bounded for the 4th priority, as shown in Figure 4.

However in Figure 3 only the cells that managed to get through are included. Those values are showing a low probability to depart at any delay value as they reflect the seeking from the 4th priority traffic of “holes” left out in the permit generation of the higher priority traffic. In addition to this phenomenon, Figure 5, the line for the 4th priority extends, theoretically, up to infinity, but only a small section of the very large values that occurred in the simulation run is included in the figure. This explosion of delay values is better illustrated in Figure 4, where the buffer fill levels for each priority are shown. The tendency to an almost linear long-term increase of the 4th buffer is clear due to the steady long-term average cell birth generation rate by the ON-OFF source model. The buffer size increases without limit.

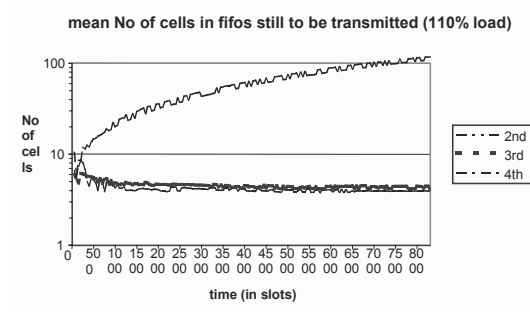


Fig. 4. Buffer size evolution with time.

Except these two scenarios, a set of simulations were performed where the occupied bandwidth in comparison to the total available is within the range from 60% to 110%, where the 10% of the upstream bandwidth was pre-allocated to the 1st priority (by means of programmed permits in the permit list as necessary for the polling of new bursts) and the rest equally distributed among the priorities. Figure 4 depicts the results from these simulations where the delay for the 2nd and 3rd priorities increases very smoothly with the load, leaving the 4th priority to suffer the congestion. The 4th priority delay increases asymptotically towards the 100% line, as expected in any queuing system.

At the same figure we can observe the behavior of the system when no priorities are available except the unsolicited permits of the 1st priority, i.e. the sources which were used for the three other classes are all feeding the same queue, marked “all” in figure 4. The total load is the same with 10% used for the 1st priority and the rest for “all”. The benefits of prioritization are clear since no traffic class can enjoy bounded delay in this case. In contrast, when priorities are enforced the 2nd and the 3rd classes enjoy a seemingly lightly loaded medium. Only the last priority sees a performance reduction, which however is well equipped to handle using the TCP layer congestion

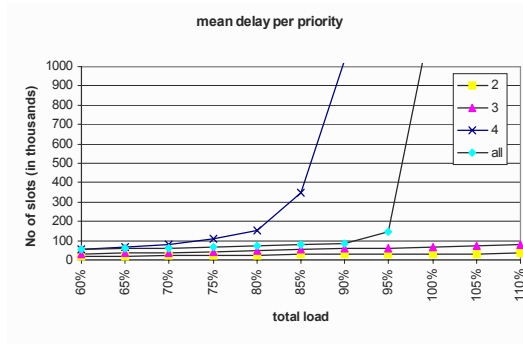


Fig. 5. Mean delay v. total load

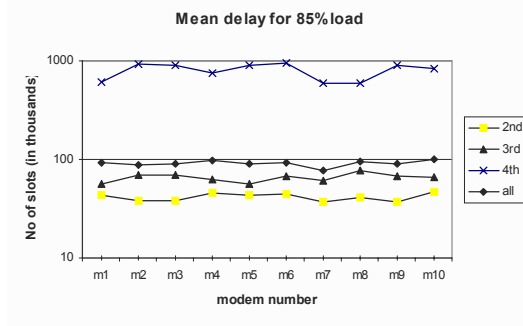


Fig. 6. Mean delay v. modem number

control tools. The delay increases linearly with the total load until the 90% point, where the congestion appears. The “all” line seems to increase asymptotically to the 100% limit.

Finally, as it can be witnessed from Figure 6, the system shares its resources fairly to all the modems involved. For the following instance, it is 25% for 2nd, 3rd and 4th and 75% for the “all”, totaling a 85% load when the 10% for the 1st is added. There is an almost linear behaviour of the various modems, for each priority, as their traffic profiles are alike for occupying exactly the same amount of load percentage. The variance from the linear behaviour is justifiable from the random traffic profiles that each modem had.

4. Conclusions

Tree-shaped topologies present attractive cost advantages for broadband access networks by allowing many customers to share the expensive head-end equipment and the feeder section. In addition they offer reuse of the copper last drops to the customer at least during the crucial introductory phase and probably for many years to come. Tree-shaped shared medium access networks such as PONs and HFC effect a

distributed multiplexing function which concentrates traffic from many users and many services with diverse requirements. To be able to guarantee that the QoS of sensitive traffic will not be disturbed by best-effort data traffic requires embedding differentiated support for flow aggregates with common requirements.

The under evaluation system thanks to prioritization can guarantee the QoS required by sensitive traffic while exploiting any unreserved bandwidth for the support of best-effort traffic. At higher layers, where the TCP resides, congestion avoidance would exercise flow control in response to losses, while re-transmitting lost packets, and reduces the rate, leading to a quite efficient overall system. The high priority traffic represents high revenue services, enjoying satisfactory performance, while the accommodation of 4th priority best-effort traffic exploiting the spare capacity, as it becomes available, guarantees a high system utilization. This is achieved with no harm to the TCP-based user applications, since in real-life instead of a tendency to infinite buffer overflows and unbounded delays, the rates of the sources would adapt to the possible bottleneck sharing the spare capacity equally among them.

A method to guarantee resource availability for the services with demanding quality, such as simple over-provisioning or fully dynamic SLAs (e.g. RSVP) would be of course necessary. Any such scheme should of course allow for some spare bandwidth to protect plain best-effort service users from the frustration of bandwidth starvation. However, as it would only involve longer term averages and not rule out brief intervals with the full capacity taken up by high priority traffic, (which can not use closed loop congestion control), the bandwidth devoted to the 4th and partly the 3rd priority would provide the necessary leeway for the statistical gain of the preventive control based services. In other words, if we did not have the lower class we would have either to accept less demanding- quality traffic or leave the excess bandwidth unutilized in an effort to avoid performance degradation during statistical extremes. It is the availability of closed loop congestion control of TCP that allows us to offer both guaranteed performance to demanding traffic and good system resource utilization, at the same time.

Acknowledgements

The work presented in this paper was partially funded by the EU ACTS project AC327 "AROMA". The opinions appearing here are those of the authors and not necessarily of the other members of the AROMA consortium. We thank also Prof. J. D. Angelopoulos for his fruitful guidance as and Th. Orfanoudakis and J. Sifnaios for their valuable help.

References

1. IETF, Differentiated Services Working Group, RFC 2475 "Architecture for Differentiated Services", December 1998
2. Van Jacobson, Kathleen Nichols, Kedarnath Poduri, Internet Draft, draft-ietf-diffserv-phb-ef-02.txt, "An Expedited Forwarding PHB", February, 1999

3. Juha Heinanen, Fred Baker, John Wroclawski, Internet Draft, draft-ietf-diffserv-af-06.txt "Assured Forwarding PHB Group", February, 1999
4. J.D.Angelopoulos, G.C. Boukis, I.S.Venieris, "Delay priorities enhance utilization of ATM PON Access Systems", *Computer Communications Journal*, Elsevier, Vol. 20, No , December 1997, pp. 937-949
5. J. D. Angelopoulos, Th. Orphanoudakis, "An ATM-friendly MAC for traffic concentration in HFC systems", *Computer Communications Journal*, Elsevier, Vol. 21, No. 6, 25 May 1998, pp. 516-529
6. J. D. Angelopoulos, N. I. Lepidas, E. K. Fragoulopoulos, I.S. Venieris, "TDMA multiplexing of ATM cells in a residential access SuperPON", *16 Journal on Selected Areas in Comm.*, Special issue on high capacity optical transport networks, Vol. 16, No. 7, September, 1998
7. ATM Forum Technical Committee, Traffic Management Group, "ATM Forum Traffic Management Specification Version 4.0" Doc. 95-013R11, March 1996