Improved Support for Streaming Services in Vehicular Networks

Vincenzo Mancuso[1], Marco Gambardella[2], Giuseppe Bianchi[3]
Università di Palermo, Dipartimento di Ingegneria Elettrica
Viale delle Scienze, 90128 Palermo, Italy
[1]vincenzo.mancuso@tti.unipa.it, [2]marco@gambardella.net, [3]bianchi@elet.polimi.it

Abstract— [1] **This paper presents a resource management mechanism aimed at improving the effectiveness of streaming services in vehicular networks. The scenario considered in this paper is that of a group of customers located into a same public vehicle, e.g. a moving train connected to the network via a satellite link, and requesting either video-on-demand-like services, as well as real-time diffusive (broadcast) streaming services. We show that a proxy server, devised to introduce an elastic buffer aimed at decoupling the information retrieval download speed on the outer network from the natural play-out speed used in the vehicular network, results to be an extremely effective approach in reducing the outage probability given by link failure in the outer network (e.g. tunnel crossing). A resource management mechanism, called A2M, is applied to both video-on-demand and diffusive services, and its performance effectiveness is evaluated through simulation.**

I. INTRODUCTION

Mobile networking not only tackles the problem of providing a networking infrastructure for mobile customers, but also includes the ability of managing moving networks. This is the case of Vehicular Area Networks (VANs). These networks are formed by customers located on the same moving vehicle, e.g. a moving train or bus, and interconnected to the rest of the world via one or more wireless links (e.g. satellite or UMTS connectivity, but emerging wireless standards for metropolitan networks, such as IEEE 802.16, may become in the near future important technologies in the VAN arena).

The scenario considered in this paper is that of a VAN in which users connectivity to the rest of the network is managed by a specialized onboard gateway. For sake of simplicity, in this paper we assume that the outer network is implemented through a single high-capacity wireless link (e.g. a satellite link). The implementation details of internal network are out of the scopes of the present paper. In addition to interworking functions the gateway may act as proxy server [1] and it can implement mechanisms meant to support video and interactive stream [2], [3], as well as multimedia with quality of service [1], [4]. Effective caching and pre-fetching mechanisms for streaming services, which may be applied with minimal or null modification to the considered VAN scenario, have been thoroughly studied in [5], [6], [7].

A problem typical of the VAN scenario is the possible outage of the outer network link ("channel outage"). Such an

outage may occur while the vehicle crosses areas characterized by severe fading conditions, e.g. tunnels. This is a very critical issue when dealing with streaming services, which experience possibly long (order of several seconds) interruption and thus cause highly negative performance impairments (service disruption) in terms of the end customer point of view. In what follows we'll refer to the event of service interruption as "connection outage". We argue that, other than caching and pre-fetching, a further role of the proxy is to hide eventual outage periods to the final user, i.e. reduce the impact of channel outage in terms of resulting connection outage. As shown in this paper, this can be accomplished by decoupling, from a service level point of view, the inner network service from the resource management occurring in the outer network segment. This in turns has highly beneficial effects in terms of network efficiency, as shown in the old work [8] where the advantage of the adoption of local storage and decoupling mechanisms has been pointed out for wired networks.

More specifically, this paper deals with two types of streaming services: video-on-demand-like services, where each user may retrieve its favorite content, and diffusive services to access broadcast multimedia such as news and sports. We show that, by inserting an elastic buffer within the proxy, it is possible to decouple the information retrieval occurring on the outer network link from the natural play-out speed used by the streaming service in the inner network. The same idea can be adopted for diffusive services, provided that an initial play-out delay is artificially inserted in the inner network streaming process. A key contribute of this paper is the thorough performance evaluation of an effective resource management mechanism (called A2M) for the outer network link. This mechanism can be managed by a proxy, and is devised to minimize outage occurrences in streaming service experienced by customers in the inner network.

The rest of the paper is organized as follows: section II presents the working framework; section III affords the technical problem statement and introduces buffering algorithms and A2M. Section IV shows simulation results validating the A2M efficiency, while diffusive services are analyzed in section V. Finally section VI proposes conclusion and open issues.

II. SERVICE FRAMEWORK

We first discuss the case of video-on-demand (VoD) service support via proxy. We recall that the scenario considered is

that of a Vehicular Area Network (VAN) where connectivity to the external network is managed through a proxy server. For convenience of presentation, we refer to the (single) wireless link connecting the VAN to the terrestrial network as "satellite link", though we recall that the proposed model does not depend on the implementation technology adopted for such link. The goal of our approach is to make the system robust to the uncertain behavior of the satellite link, requiring a line-of-sight connection, which cannot be guaranteed.

The adoption of an onboard proxy is proposed to minimize outage periods. The idea is to split the video-on-demand connection into two separate segments. In the inner network, a streaming session is set-up between the end user and the proxy server. In the outer network, the video information is downloaded at a rate higher than the natural play-out speed of the video. This is of course possible provided that the sum of the play-out rates for the concurrent streaming sessions is lower than the satellite link capacity, i.e. provided that extra bandwidth is available on the satellite link. The excess information downloaded from the satellite is then dynamically buffered in a suitable storage area made available at the proxy. When the satellite link is in outage, the proxy stops receiving data from server. However, clients connected to the proxy continue to receive data. Connection outage occurs only when the satellite link is in outage and the buffered data terminates. The described operation gives raise to an "elastic" buffer, which is filled during the periods in which the satellite link is active, and whose buffered data is consumed at a fixed rate given by the play-out speed for each streaming session, multiplied by the number of concurrent streaming sessions.

We'll show in the next section III that a very important role is played by the strategy adopted in managing the satellite link capacity, and specifically we will show that a uniform allocation of the extra available bandwidth to the concurrent downloads is a highly sub-optimal strategy.

In addition to VoD-like services, delivery of broadcast information (Diffusive Services - DS) may also take significant advantage of the availability of a proxy server. To compensate for satellite link outage, it suffices to introduce a delay in the onboard play-out of the DS. In other words, if an event is scheduled to be broadcast at a given time t, onboard transmission will start at time $t + D$, and the proxy will buffer all the information related to the D seconds of delay introduced. Clearly, when outage occurs on the satellite link, recovery of the missing information must be provided via a dedicated proxy to network download connection, which will use the extra bandwidth available on the satellite link. Technical details will be presented in section V.

III. Resource Management algorithm

This section gives an insight of the relation between buffering schemes and bandwidth assignment strategies in a VoD scenario. Extensions to DS will be tackled in section V.

Let us consider the eventuality that a customer request a file. If resources are available the proxy firstly checks if the requested file is locally stored, and then eventually retrieves
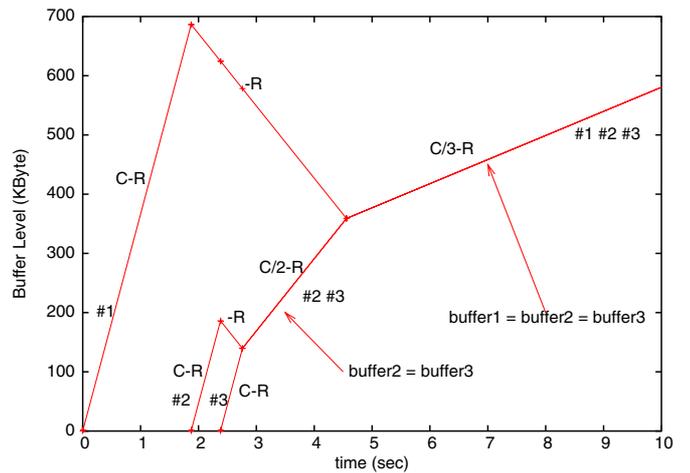

Fig. 1. A2M: bandwidth is shared between connection with lower buffer

the file from server at the available rate while delivering the content to the client. The entire wireless channel is shared between all active connections: according to the specific user data rate, the proxy reserves a proportional satellite bandwidth to each download. If a new connection is accepted, the available bandwidth is fairly reallocated and redistributed to all connections. We call this proxy operation bandwidth "Equal-Distributed", namely ED.

We also propose a different approach that consists in the dynamical adaptation of the bandwidth allocated to each connection between server and proxy. In particular, we monitor buffers at the proxy side and assign all the available bandwidth to the flow, or the group of flows, suffering the lower buffer level. This is the A2M operational-mode, in which the bandwidth is reserved "All to Minimum" buffer stream. Thus the A2M operation can be summarized as follows: when a client requests a file, the proxy checks if the requested file is locally stored, and eventually starts retrieving it from the server as fast as possible; while delivering the file, the proxy checks for flows with lower buffer level and makes the total server bandwidth shared among these flows. The temporal behavior of proxy buffer level is plotted in 1. In that figure, the proxy starts from an idle state and then progressively accepts three incoming VoD connections requesting non-pre-fetched files, and requiring the same data rate. The figure plots the buffer level for each connection vs. the simulation time. Firstly all the available bandwidth C is reserved to the first incoming connection, so that the buffer level grows with a rate $C - R$, where R is the proxy-to-client rate.

When a second connection is accepted, all the bandwidth C is switched to this new stream, so that the former stream buffer decreases with a rate R. The same occurs when the third connections is accepted. After the buffer level of the third stream ($buffer3$) reaches the above one ($buffer2$), A2M algorithm operates a re-distribution of the bandwidth. Streams #2 and #3 are now using half the full server rate and their buffer levels grow until $buffer1$ level is reached. From that instant on, three buffers are served at the same rate, that is $C/3$. A similar dynamic bandwidth allocation mechanism for

circuit-switched satellite network, proposed in [9], shows the improvements due to the adoption of such a scheme in terms of bandwidth utilization and QoS guaranties.

Nonetheless, a stream could stop while running, due to the probability that satellite and mobile proxy are not in line-of-sight (P_{ch-out}) for a long time. For what it concerns the effects of A2M on the connection outage probability ($P_{conn-out}$), it is shown in section IV that a beneficial impact is obtained. In fact, using A2M each new connection is timely made able to accumulate a certain buffer level on the proxy. Since the higher is the buffer level, the lower is ($P_{conn-out}$), A2M tries to fill proxy buffers as fast as possible, preventing outages.

What described can be easily extended to a framework with flows characterized by different data rate. The only difference is that buffer level has to be metered in terms of play-out time. Note that, for a VoD flow, in the case of channel outage, the buffer level represent the margin before an outage occurs. For this reason we will use also the term "outage margin".

IV. VOD SERVICES USING A2M

The behavior of VoD flows and of A2M algorithm has been described in the previous sections. Now we want to investigate about the effectiveness of the proposed solutions.

A very basic scenario was tested, in order to clearly show the influence of the A2M scheme. Nonetheless, a great amount of parameters has to be considered for a deeper analysis. Firstly, in the evaluation process it is relevant the statistical characterization of the way a client requests an object. A very common approach consists in assuming the file request process as a Poisson arrival process. Since we want to stress the A2M mechanism, in our simulation we adopted an infinite arrival rate for the Poisson process, that is a "continuous load" is offered to the proxy. Thus, performance results obtained have to be considered as a worst case performance. As regards the file size distribution, files downloaded from web fit a heavy tailed distribution function [6]. Since we considered a video retrieval scenario, a uniform file size distribution was adopted. Admission control schemes and transport protocol selection affect QoS performance in terms of reliability and download times. According to the selected admission control scheme it is possible to prevent congestions and optimize the channel utilization, while the transport protocol determines the effective file download rate. For instance, a standard TCP protocol could severely affect the transmission over a satellite link, while a UDP protocol should maximize the data rate at the expenses of reliability, since no delivery checks are performed. In the envisioned scenario, the client rate is almost fixed, while the proxy can store bytes at a variable rate. This implies that different protocol could be used on the server-proxy connection and on the proxy-client one. As to channel behavior, the satellite transmits data over a reliable channel. Anyway, the satellite transmission requires a line-of-sight between sender and receiver; otherwise a channel outage event occurs. The statistic of such event is related to the VAN mobility. In particular, a channel outage event occurs when the train, with proxy on board, enters a tunnel. Thus, the
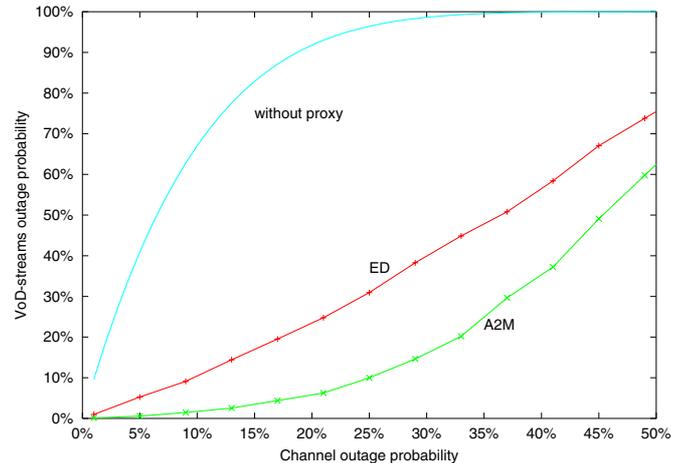

Fig. 2. Performance comparison between A2M and ED schemes

channel outage probability is route-dependent. For the sake of simplicity, we consider satellite link as an on/off function with both on and off periods exponentially distributed. Finally, the proxy cache strategy is responsible for channel usage optimization and for clients' connections outage prevention, as previously shown in section III.

The proposed scenario has been tested by means of a fluidic C++ event-driven simulator, whose parameters are reported in table I. First of all, in figure 2 we briefly show that performance obtained using A2M is better than that obtained with ED. In terms of outage probability, the figure 2 reports a typical behavior of the outage probability per VoD flow vs. the outage probability of the satellite link. Performance obtained in a scenario without proxy is also reported. First of all it is possible to argue that using a proxy has an intrinsic and remarkable benefic impact over system performance. As to the comparison between ED and A2M, figure 2 shows that the ED curve is well above the A2M curve for a wide range of P_{ch-out}, and using A2M allows us to further improve network performance of about 10-15%.

Figure 3 only deals with A2M, and it is related to the VoD flow performance in terms of additional of time needed to complete the download of a video, in comparison with the flow duration itself. Three curves are depicted, corresponding to different number of parallel flows. It is possible to say that good performance can be obtained only if the channel outage probability P_{ch-out} is smaller than 15%, with a substantial reduction if outages are rare.

| Simulation Parameter	Value
proxy download rate C	$32 Mbps$
client download rate R	$2 Mbps$
proxy buffer fill rate F	$0 \leq F \leq C$
proxy max conns N	$0 \leq N \leq 16$
proxy buffer emptying rate E	$E = R = 2 Mbps)$
requests arrival	continuous load
admission ctrl: max clients M	$0 < M \leq N$
wireless channel	on-off (exp. distrib.)
channel outage prob P_{ch-out}	$0 \leq P_{ch-out} \leq 1$
channel outage mean time T_{off}	180 s
channel mean *on* time T_{on}	$T_{on} = T_{off} \frac{1 - P_{ch-out}}{P_{ch-out}}$

TABLE I
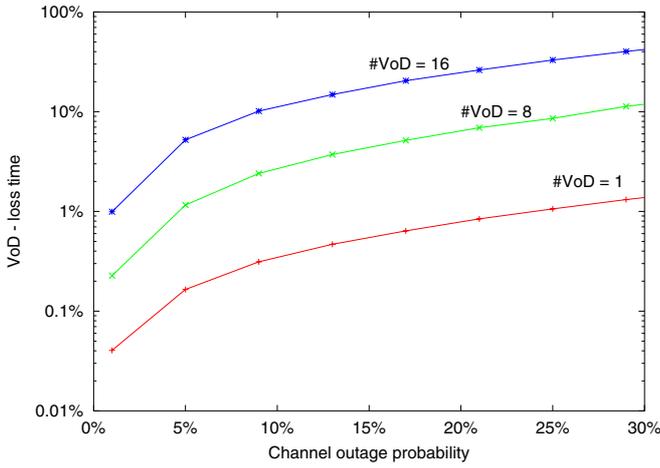
Fig. 3.   VoD flows performance using A2M

Fig. 4.   Connection outage after multiple link failures

V. DIFFUSIVE SERVICES USING A2M

Each active Real Time channel has a dedicated buffer on board of proxy, storing data received by the proxy. Thus data can be delivered to users after a fixed time delay (we will call it "play-out"). The behavior of DS imposes the use of fixed buffers; the buffer size is exactly equal to the one needed to store video data during the play-out time. Our studies focuses on the proxy parameters choice and on the coexistence of DS and VoD services on the same link, sharing the same bandwidth (here referred to as "recovery bandwidth").

The presence of a buffer and the use of a play-out delay of the stream does not resolve the problem of temporary channel outages. In fact, at the re-establishment of the link (i.e. immediately after the end of a tunnel), a transient period will be necessary to refill the buffer. The difference with the VoD scenario is that during the channel outage, the satellite broadcast transmission goes on. As a consequence, the proxy buffer stores data and leaves empty spaces during channel outages. In what follows, we will refer to these spaces as "unreceived windows". Unreceived windows are filled after the link is re-established, using a "recovery procedure", lasting a transient period. The duration of the transient period depends on many factors. This kind of application requires the logical division of satellite bandwidth into two parts: a fixed rate will be statically assigned to the retrieval of the latest data (common to all VANs in the same coverage area), while the remaining bandwidth (recovery bandwidth) will be used by proxies to fill their "unreceived windows".

The buffer will be characterized by the presence of a number of unreceived windows, that corresponds to the number of outages occurred in the last M seconds, where M is the size of the buffer (i.e. the play-out delay), minus those windows that the system has been able to recover meanwhile. An important parameter associated to the unreceived window is the "outage margin", that is the difference between the position (in the buffer) of the older unreceived window and the last data segment delivered to clients. The meaning of the outage margin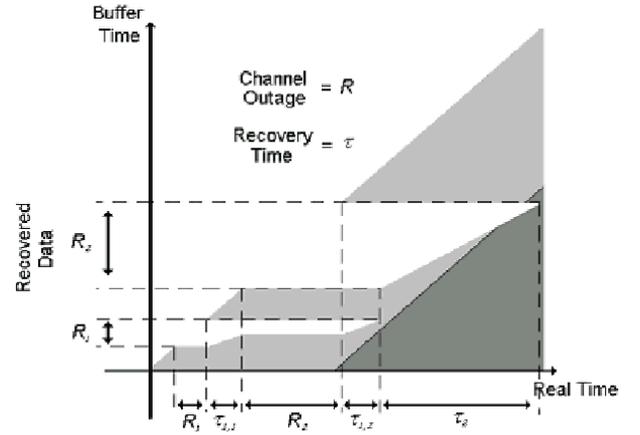 is: in absence of recovery procedures, a connection outage would occur in a time equal to the outage margin. Our analysis is based on the integration between DS and VoD services. To apply A2M algorithm to DS means that at any time all available bandwidth is assigned to the flows (VoD or RT) whose buffers have the lowest outage margin.

An important parameter is the ratio between the bandwidth assigned to the recovery of RT data and the sum of bit-rates of RT flows. This parameter can be intended as the amount of video seconds that will be recovered in a second (in the future we will call it K).

In figures 4 and 5 a scenario dealing with a multiple unreceived windows is depicted. In what shown in figure 4, the recovery time of the first unreceived window, R_1, is split into two parts: $\tau_{1,1}$ and $\tau_{1,2}$. The time $\tau_{1,1}$ is employed to recover the first unreceived windows, but a second channel outage event occurs before the windows is fully recovered. At the link re-establishment, after the second tunnel, all available bandwidth is initially used to fill the former unreceived window ($\tau_{1,2}$). Then all available bandwidth is used to fill the latter unreceived window (τ_2). Unfortunately τ_2 is too long, and the slope K results inappropriate to avoid a connection outage. When K>1, connection outages can occur only when the train is in a tunnel. But K>1 means that more than 50% of total bandwidth is allotted to recovery; this is a poor efficiency in link utilization. When K<1, the time needed to recover a flow interval is longer than the interval itself: thus an "unreceived window" cannot be recovered while delivering it.

The simplest way to analyze the problem is to observe the temporal behavior of the outage margin as depicted in fig 5. The connection outage occurs when outage margin becomes negative. During a connection outage the outage margin decreases with a slope equal to -1. During the recovery of an unreceived window it has a slope of $K-1$ (that can be positive or negative). Immediately after the recovery of an "unreceived window" the outage margin presents a discontinuity, constituted by the block of data following the unreceived window. The outage margin can remain constant (and equal to the play-out time) when no unreceived window exists and the satellite link is active.
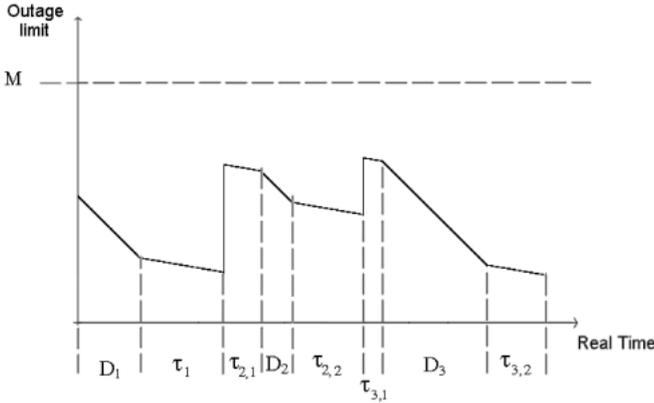
Fig. 5. Temporal behavior of outage margin. τ_i is the recovery time (divided in sub-intervals $\tau_{i,j}$), D_i the tunnel duration, and M is the play-out time.
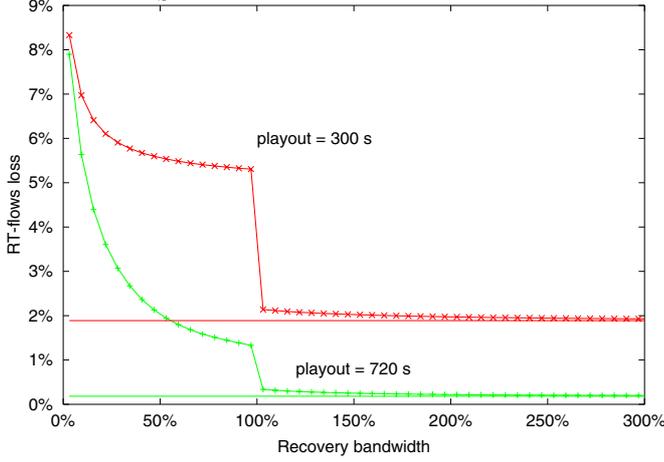

Fig. 6. Percentage of Real Time flow lost.

The last figure (fig 6) represents the percentage of Real Time flow lost as a consequence of connection outages (unrecovered data). Results was obtained via simulation using a single never-ending broadcast flow, with its related recovery channel. A dependence on the Recovery Bandwidth (normalized to the RT flow bandwidth) is given: when it grows, RT-flow loss decrease. It shows an asymptote that depends of the play-out time. In fact the connection outage probability (with infinite recovery bandwidth) is equal to the probability of a channel outage longer than the play-out time. Using exponential distribution with mean $\mu = T_{off}$, if P_d is the play-out delay adopted, T the overall simulated time, the mean number of tunnel per simulation is:

$$N = \frac{T}{T_{off} + T_{on}} = \frac{T}{\mu + \frac{1 - P_{ch-out}}{P_{ch-out}}\mu} = P_{ch-out}\frac{T}{\mu}$$

while the mean outage time per tunnel is:

$$L = \int_{P_d}^{\infty}(x - P_d)e^{-\frac{x}{\mu}}dx = \mu e^{-\frac{P_d}{\mu}}$$

The average percentage loss is:

$$L_\% = 100\frac{LN}{T} = 100 \cdot P_{ch-out} \cdot e^{-\frac{P_d}{\mu}}$$

In figure 6 horizontal asymptotes are computed according to the previous formula, and they only depends on the adopted play-out, being $\mu = 180s$ and $P_{ch-out} = 10\%$. It is possible to see that very good performance can be experienced with a reduced recovery bandwidth amount; moreover it is not useful to increase the recovery bandwidth far above the natural flow stream (the "broadcast" data rate).

Another important detail is the discontinuity showed in correspondence to a Recovery Bandwidth of 100%. This band is equal to the total flow bandwidth, so it corresponds to $K = 1$. The meaning of the discontinuity is: when $K > 1$ and a connection outage occurs inside a tunnel, the loss time is equal to the remaining fraction of the tunnel; connection outage cannot occur outside a tunnel. When $K < 1$ and a connection outage occurs inside a tunnel, the loss time is equal to the remaining fraction of unreceived window that has not been recovered; moreover, and mostly important, connection outages can occur outside a tunnel (see also figure 4).

VI. CONCLUSIONS AND OPEN ISSUES

We considered VAN users attempting to download multimedia contents from the outer network, relying on a proxy. Data flows is split into a secure and reliable proxy-client connection, and a time-variant server-proxy connection. VoD and DS services have been taken into account, and a suitable bandwidth sharing algorithm, A2M, has been proposed to confer reliability to both kinds of service. Simulation results show the validity of A2M, minimizing the probability that a client's download is stopped due to proxy buffer emptying, when satellite server is out-of-sight. A2M can be considered a well suited bandwidth sharing algorithm, able to improve the robustness of the proxy system; it is also a scalable mechanism, as no complex state information has to be managed, but the number of downloads. Next steps in research mainly relate to the transport protocol influence, and the implementation of an admission control scheme, taking advantage from the amount of pre-fetched data in the proxy.

REFERENCES

[1] R.Rejaie, H.Yu, M.Handley, D.Estrin "Multimedia proxy caching mechanisms for quality adaptive streaming applications in the internet", proceeding of Infocom 2000, March 2000
[2] S.Sent, J.Rexford, D.Towsley, "Proxy prefix caching for multimedia streams", Proceeding of Infocom 99, April 1999
[3] M.Reissline, F.Hartanto, K.W.Ross, "Interactive video streaming with proxy servers", proceeding of IMMCN, February 2000
[4] B.Wang, S.Sen, M.Adler, D.Towsley "Optimal Proxy Cache Allocation for Efficient Streaming Media Distribution", proceedings of Infocom 2002, 2002
[5] Y.Wang, Z.L.Wang, D.H.Du, D.Su, "A network-conscious approach of end-to-end video delivery over wide area networks using proxy servers", Proceeding of Infocom 98, 1998
[6] M.Crovella, P.Barford,"The network effects of prefetching", Proceedings of IEEE Infocom'98, 1998
[7] S.Jin, A.Bestavros, A.Iyengar, "Accelerating Internet Streaming Media Delivery using Network-Aware Partial Caching" proceeding of the 22th International Conference on Compuyting Systems - IEEE ICDCS'02, 2002
[8] G.Bianchi, R.Melen, "The role of the local storage in supporting video retrieval services on ATM networks", IEEE/ACM Transaction in networking, vol.5, no.6, December 1997
[9] N.Blefari-Melazzi, G.Reali, "Improving the efficiency of circuit-switched satellite networks by means of dynamic bandwidth allocation capabilities", IEEE journal on selected areas in communications, vol.18, no.11, November 2000