Framework with Scaleable Error Control for Reliable Multipoint Services in ATM Networks

Georg Carle

Institute of Telematics
University of Karlsruhe, D-76128 Karlsruhe, Germany
Telephone: ++49 / 721 / 608-4027, Fax: ++49 / 721 / 388097
E-Mail: carle@telematik.informatik.uni-karlsruhe.de

Abstract

Distributed multimedia applications, as well as advanced applications for computer-supported cooperative work and distributed computing, depend on efficient communication subsystems providing a variety of services. This paper presents a novel framework for support of multipoint communication in ATM networks. Two adaptation layer protocols are presented that provide reliable multicast services. The first one, called RLMCP (Reliable Lightweight Multicast Protocol), uses a frame-based ARQ-scheme. The second one, called RMC-AAL (Reliable Multicast ATM Adaptation Layer), features cell-based ARQ in combination with FEC. A new network element, called the Group Communication Server (GCS), is presented for implementing the adaptation layer protocols in network nodes. It allows for hierarchical multicast error control and support of heterogeneous scenarios. The framework permits to select the combination of error control mechanisms most suitable for the application requirements in a specific communication scenario. The functionality of end systems and group communication server are described, and a basic implementation architecture is presented. The scalability properties of the different error control schemes is analysed, identifying the influence of group size, cell loss rate and path capacity.

1 Introduction

Upcoming applications, for example distributed multimedia systems, computer-supported co-operative work (CSCW) applications, and virtual shared memory systems require reliable high performance multipoint communication services. Quality of service (QoS) issues of importance are not only throughput, delay, and delay jitter, but also differences of delay and reliability within the group. A key problem that must be solved to provide a reliable multipoint service is the recovery from cell losses due to congestion in the switches. The probability for cell loss may vary over a wide range, depending on the strategy for usage parameter control (UPC) and call admission control which is applied. It is still an open question how low cell loss rates can be guaranteed for bursty multicast traffic, while using network resources efficiently. Cell losses caused by buffer overflows do not occur randomly distributed, but show a highly correlated characteristic [1]. If a reliable service in ATM networks is based on traditional transport protocols like TCP, severe performance degradations may be observed [2]. For the provision of a reliable multipoint service, the probability for losses increases for a growing number of receivers. However, there are still no convincing concepts for reliable high-performance group communications in ATM-networks. Therefore, the provision of reliable group communications requires the development of efficient protocols and of communication systems that achieve high performance even under conditions with high cell losses.

This paper focuses on design and assessment of error control mechanisms for correction of cell losses for group communication. Section 2 gives an overview on related protocols for error recovery. In section 3, the proposed framework for reliable multicast communication is presented. Section 4 presents performance results of different error control schemes.

2 Protocols for reliable services in ATM networks

2.1 Adaptation layer protocols

According to the B-ISDN protocol reference model, mechanisms for error recovery may be integrated into the Service Specific Convergence Sublayer (SSCS) of the adaptation layer for provision of an assured mode service [3]. Up to now, only two SSCS-Protocols that offer error control mechanisms are specified by ITU. The Service Specific Connection Oriented Protocol (SSCOP) [4] is subject of standardisation for an SSCS that offers assured mode service for signalling. The protocol provides end-to-end flow control and recovery of lost or corrupted data frames by selective retransmissions. However, SSCOP does not support assured mode multicast connections. As shown in [5], it is possible to extend SSCOP to allow for partial-frame retransmissions. For AAL1, an SSCS with FEC is proposed [3], based on a Reed-Solomon-Code applied on blocks of 128 cells that allows the regeneration of up to four missing cells. Additional FEC schemes for ATM were proposed and investigated in [6], [1], [7], and [8] but there are still a number of open questions concerning the combination of FEC and ARQ in ATM networks.

2.2 Transport protocols

Transport protocols that are suitable for a connectionless network layer, like TCP, TP4 and XTP, are not very well suited to an ATM environment. The error control mechanisms of these protocols are very general and not designed for ATM cells and AAL frames. These transport protocols need to tolerate packets delivered out of sequence by the network layer. An adaptation layer protocol may benefit from the in-sequence delivery of the ATM-layer service and may use sequence number gaps for error detection. XTP offers support for reliable multicasting by a list-based algorithm and the so-called bucket algorithm. However, error control based on the bucket algorithm has significant shortcomings, as shown in [9].

TP++ [10] is an example for a transport protocol that is suitable for ATM networks. It uses retransmissions in combination with FEC for error recovery (type I hybrid ARQ). However, it is only capable of unicast communication. Up to now, no protocol that combines ARQ and FEC was presented for multicast communication in ATM networks.

There are a number of additional approaches for integrating FEC into the transport layer (e.g., [11]). These schemes are capable of recovering lost packets. For transport protocols that reside above AAL5, the loss of a single cell leads to the discarding of a complete packet. In such cases, packet-based FEC within the transport layer has to recover not only the lost cell, but all cells of a corrupted frame. In ATM networks it is therefore possible to achieve a better performance with cell-based FEC schemes than with packet-based FEC schemes.

Another problem arises due to the fact that ATM signalling differs conceptually from signalling in traditional transport protocols. ATM is based on out-of band signalling, while conventional transport protocols are based on in-band signalling. If these protocols are to be used in ATM networks, mapping of transport layer connection control to ATM signalling needs to be performed [12].

2.3 Protocol Implementation

While transmission capacity was growing enormously over the last years, protocol processing and system functions in the transport component turned out to be a performance bottleneck. High performance communication subsystems, based on parallel protocol processing, and hybrid architectures with hardware components for time-critical operations [12] are required for the provision of a service with high throughput and low latency. For highest performance, complete VLSI implementations of transport subsystems are planned [14]. The performance bottleneck of the transport component that can be observed for point-to-point-communication is even more crucial for reliable multipoint connections. For a growing number of receivers, processing of a growing number of control packets (known as the implosion problem), and management of a large amount of status information needs to be performed.

2.4 Selection of Protocol Mechanisms

In order to offer a wide range of services to the applications for various network parameters, several concepts of flexible communication subsystems are under development. The parallel transport system PATROCLOS [13] is a parallel implementation of a high performance transport system, offering a wide range of protocol mechanisms that may be selected according to the needs of an application. The Flexible Communication SubSystem (FCSS) [15] is a configurable, function-based transport system. It allows protocol configuration and the reservation of resources in order to provide applications with a specific service quality.

3 Framework for Reliable Multipoint Communication in ATM Networks

A conceptual framework was developed that allows to select the error control mechanisms most suited for a specific environment. It describes how frame-based ARQ, cell-based ARQ, and cell-based FEC may be integrated into the adaptation layer for the efficient provision of reliable multicast services. Additionally, it describes how these error control mechanisms may be integrated into dedicated servers, and how large groups may be supported by a hierarchy of servers for better scaling properties of reliable group communication.

The framework considers a number of different group communication services: a fully reliable multicast service with assured delivery to every receiver, a K-reliable multicast service with assured delivery to at least K receivers of a group, and a real-time service in which error control is performed subject to deadlines. Additionally, a multiplexing service is provided by the group communication server for multiplexing of AAL frames from different transmitters over a single virtual connection.

As lost retransmissions contribute significantly to the QoS offered by the AAL, it is frequently of high importance to decrease the probability of lost retransmissions. The cell-based retransmission scheme and the FEC scheme of RMC-AAL explained in section 3.3 allow to decrease this probability. Additionally, the capability of ATM to offer virtual channels (VCs) with different cell loss probabilities may be used. This is in contrast to conventional networks, where initial transmissions and retransmissions will generally observe identical loss rates.

3.1 Mapping to virtual connections

The framework allows to select one of the following alternatives for reliable group communication:

- In the simplest case, a single 1:N multicast VC from the transmitter to the receivers with same QoS for all cells will be used. This requires demultiplexing above the ATM layer to distinguish ordinary transmissions from retransmissions.
- Improved performance may be achieved by an 1:N VC with different cell loss priorities according to the cell loss priority (CLP) bit in the ATM cell header. The initial transmission of a frame uses lower priority cells (CLP = 1). Retransmitted frames are sent with higher priority cells (CLP = 0). However, when ATM traffic control uses its possibility of converting high priority cells to low priority cells, the QoS of the retransmissions will not be higher than the QoS of ordinary transmissions.
- For ensuring the QoS of retransmissions, two 1:N VCs from transmitter to receivers may be used. One VC has a QoS suitable for the initial transmission. The second VC is intended for retransmissions. Its QoS will be set to low delay and increased reliability.
- In order to support individual retransmissions to receivers that observed losses, one of the alternatives above may be combined with a number of VCs from the transmitter to single receivers or to a subset of receivers. This allows that retransmissions lead to a reduced network load, and prevents receivers from the need to filter out unwanted retransmissions.

3.2 Frame-based error control

For a simple and efficient provision of reliable multicasting, the Reliable Lightweight Multicast Protocol (RLMCP) was developed as a Service Specific Convergence Sublayer for AAL5. It provides assured mode point-to-point and point-to-multipoint services using frame-based ARQ. Retransmissions may be performed in selective repeat or go-back-N mode.

The RLMCP frame header has a length of 10 bytes and is identical to the SSCS header of data frames of RMC-AAL explained in the next section. The protocol is based on the following data format: the first byte of the header indicates the frame type (data frame, retransmission frame, or acknowledgement). It also contains a flag to request 'immediate acknowledgement' (I-Ack) and a flag indicating the 'last frame of burst’ (LastF). Frames carry frame sequence numbers of 24 bits, which is sufficient for high-speed WANs. Frames also carry a sequence number for the 'lower window edge' (LWE) of the transmitter, indicating the lowest sequence number a transmitter is prepared to repeat. This allows to avoid unnecessary retransmission requests in K-reliable and real-time services.

Receivers send acknowledgements periodically, after reception of a frame in which an 'immediate acknowledgement' bit is set, or after detection of a missing frame. Lost frames are detected by gaps of the frame sequence numbers, or by time-outs. The 'last frame of stream' bit allows the receivers to stop the loss detection timer. Receivers may use cumulative positive acknowledgements, sending a lower window edge, and selective positive or negative acknowledgements, using bitmaps with a length of 32 bytes (for a sequence of up to 256 frames). For flow control, acknowledgements contain an upper window edge for the highest sequence number a receiver is prepared to receive.

3.3 Cell-based error control

The Reliable Multicast ATM Adaptation Layer (RMC-AAL) features cell-based ARQ and FEC for an efficient provision of reliable multicast services under conditions of higher or varying cell loss rates, and for applications with strong delay requirements. Error recovery of RMC-AAL is based on three schemes: pure ARQ, type I hybrid ARQ, and pure FEC. A fully reliable service and a service that assures delivery to a subset of K receivers are offered.

Cell sequence numbers (CSN, 6 bits) are provided for detection of missing cells. Frames are identified by a frame sequence number (FSN, 24 bit) in the frame header. Like RLMCP, RMC-AAL uses the trailer of AAL5-CPCS, protecting the payload of a frame by the cyclic redundancy check CRC-32. Data frames have a protocol overhead of 10 bytes in the frame header and 8 bytes in the frame trailer. In each cell, they have an additional overhead of one byte (2 bit for cell type CT, and 6 bit cell sequence number). Even for high speed VCs in WANs, no large cell numbering space is required, because every cell is identified by both FSN and CSN. The alternative solution of identifying cells entirely by their cell sequence numbers leads to a significantly higher overhead per cell. For example, the protocol BLINKBLT [16] which also offers cell-based retransmissions has a per-cell overhead of 4 bytes. The RMC-AAL frame header contains a transmitter identifier (Tx-id) and the length of the SSCS-PDU payload (Len-1). The frame header also contains the discriminator byte (Dis) with an identifier for the frame type (FrType), two flags (I-ACK and LastF), and the number of redundancy cells (#RedundCells) that follow the data frame. Redundancy cells use independent cell sequence numbers (RCSN). When FEC is used, h redundant cells are generated to protect the information cells of the frame. When FEC is used, h redundant cells are generated to protect the information cells of the frame. Encoding and decoding can be based on Reed-Solomon-Codes [17], or on simple XOR-operations and matrix interleaving [6]. Figure 1 shows a suitable FEC encoder, and Figure 2 a suitable decoder. Both are based on simple XOR-functions. Retransmissions may be sent by multicast or by unicast in selective repeat or go-back-N mode. It can be selected if retransmissions are frame-based (by retransmission of data frames) or cell-based (by retransmission of frame fragments). Frame fragments consist of a Fragment Header Cell, followed by a selection of original data cells of this frame. The fragment header cell contains the frame sequence number (FSN) of the original frame. This field is called 'Start of Bitmap' (SBM). A bitmap (BM) is used to indicate which cells of the original frame are retransmitted within the frame fragment. The field 'Length of Bitmap' (LBM) indicates the valid length of the bitmap, and the field 'Offset Bitmap' (OBM) indicates the cell number of the first bit of the bitmap.

Receivers send acknowledgements periodically, after reception of a frame in which an 'immediate acknowledgement' (I-ACK) bit is set, and after detection of cell loss. Acknowledgements contain a receiver identifier (Rx-id). An upper window edge (UWE) allows for window flow control. Receivers may use cumulative positive acknowledgements of frames by sending the frame sequence number of their lower window edge (LWE). Additionally, they may use bitmaps (BM) with a length of 32 bytes for negative acknowledgements of frames or individual cells. A frame sequence number (SBM), a field for the valid length of the bitmap (LBM), and a field for the offset of the bitmap identify the position of the bitmap within the window.

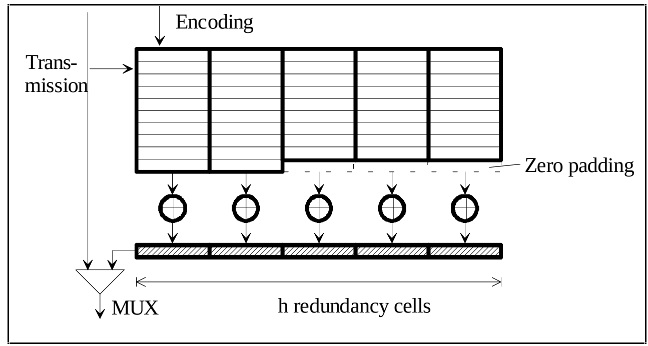

Figure 1: FEC Encoder

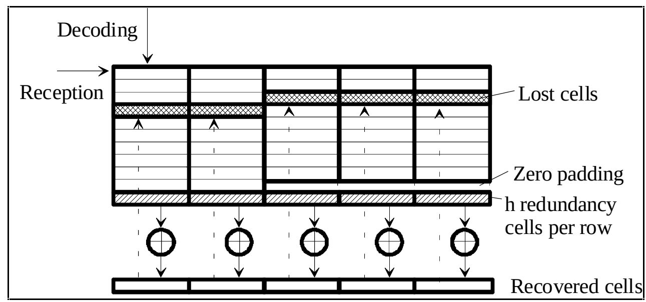

Figure 2: FEC decoder

3.4 Support of Real-time services

The presented framework may also be applied for real-time services, when the reliability of the network bearer service does not meet the reliability requirements of the application. In many scenarios, the delay requirements of real-time services can be met even in the case of retransmissions. This allows to trade off the disadvantages of additional complexity for error control in end systems and servers with the advantages of a simpler bearer service. Public network providers will be able to offer lower tariffs for a bearer service with higher loss probability. RLMCP and RMC-AAL can provide support for real-time services by managing deadlines in transmitters and receivers. After detecting that a deadline has passed for a specific frame, the transmitters will ignore negative acknowledgements, and receivers will stop sending negative acknowledgements.

3.5 Group Communication Server (GCS)

The presented reliable multicast adaptation layer represents an important step towards a high performance reliable multicast service. Further improvements of performance and efficiency may be achieved by the provision of dedicated servers in the network for hierarchical multicast error control. The proposed Group Communication Server (GCS) integrates a range of mechanisms that can be grouped into the following tasks:

- Provision of a high-quality multipoint service with efficient use of network resources;
- Provision of processing support for multicast transmitters;
- Support of heterogeneous hierarchical multicasting;
- Multiplexing support for groups with multiple transmitters.

For the first task, performing error control in the server permits to increase network efficiency and to reduce delays introduced by retransmissions. Allowing retransmissions originating from the server avoids unnecessary retransmissions over common branches of a multicast tree. The integration of FEC mechanisms into the GCS allows regeneration of lost cells and reinsertion of additional redundancy for adjusting the FEC coding scheme according to the needs of subsequent hops.

For the second task, the GCS releases the burden of a transmitter that deals with a large number of receivers, providing scalability. Instead of communicating with all receivers of a group simultaneously, it is possible for a sender to communicate with a small number of GCSs, where each of them provides reliable delivery to a subset of the receivers. Integrating support for reliable high performance multipoint communication in a server allows better use of such dedicated resources.

For the third task, a GCS may diversify outgoing data streams, allowing conversion of different error schemes and support of different qualities of service for individual servers or subgroups. The group communication server will offer the full range of error control mechanisms provided by the reliable multicast adaptation layer. For end systems, it is not required to implement the full functionality of RMC-AAL. It will be sufficient to have access to a local GCS for participation in a high performance multipoint communication over long distances. The error control mechanisms of individual end systems have only negligible influence onto the overall performance, as simple error control mechanisms are sufficient for communication with a local GCS. If an additional priority field is used in the frame format, the server is able to distinguish packets of different importance. One example application would be hierarchically coded video. For information with different importance, different FEC codes may be applied inside one VC, or specific frames may be suppressed for certain outgoing links. The GCS also allows to support heterogeneous groups that use both RLMCP and RMC-AAL. For this purpose, functions for conversion between different frame formats are provided.

For the fourth task, the GCS provides support for multiplexing of frames onto a single point-to-multipoint connection. This allows to reduce the number of required VCs significantly for large groups with many transmitters [18].

3.6 Server modes

The Group Communication Server may operate in three different modes. In the forwarding mode, every frame is processed first by the GCS before being forwarded to the receivers. In case of simple 1:N multicasting, increased performance may be achieved in the bypass mode. In this mode, an ATM switch that supports multicasting will forward data directly to the server and the receivers, reducing the processing load of the server and the overall latency. In both modes, the GCS detects errors earlier than the receivers, and can report an error to the source with lower delay. Both modes also support processing of acknowledgments. For this purpose, every receiver may maintain an individual virtual channel to the GCS. The GCS will either perform the required retransmissions, or will forward retransmission requests to the source. If a window-based flow control scheme is enforced that includes the GCS, the GCS may guarantee to perform the retransmissions. However, buffer limitations in the GCS may limit performance in this case. The third mode is more complex, but allows the provision of a multipeer service with multiplexing of messages from different transmitters over a single virtual connection.

In all three modes, a hierarchy of servers allows for good scaling properties for large groups and high path capacities. For acknowledgements, the receivers maintain individual unicast VCs to the GCS.

3.7 Implementation of the GCS

Figure 3 shows a functional architecture for the GCS. The functionality is distributed to a number of modules. The ARQ manager processes acknowledgements and manages status information of individual receivers and the group. The send manger schedules between ordinary transmissions and retransmissions. The connection manager schedules between different multicast or multipeer connections.

The modules exchange control information with pointers to the payload of cells (indicated by thin interconnection lines), or control information together with the payload of a cell (indicated by thick interconnection lines). The modules are specified as individual finite state machines with local status information. Implementing the GCS in software, each module may be implemented by an individual thread. In a parallel hardware-based implementation, the state machines may be implemented by microprogrammable units [19]. For high-performance implementations, dedicated hardware support for acknowledgement processing may be provided for filtering and processing the bit maps of acknowledgements. Additional hardware components may be used for CRC, FEC, buffer management, list and timer management [20]. For demultiplexing, a content addressable memory (CAM) may be used to map the large VPI/VCI address space onto smaller internal identifiers.

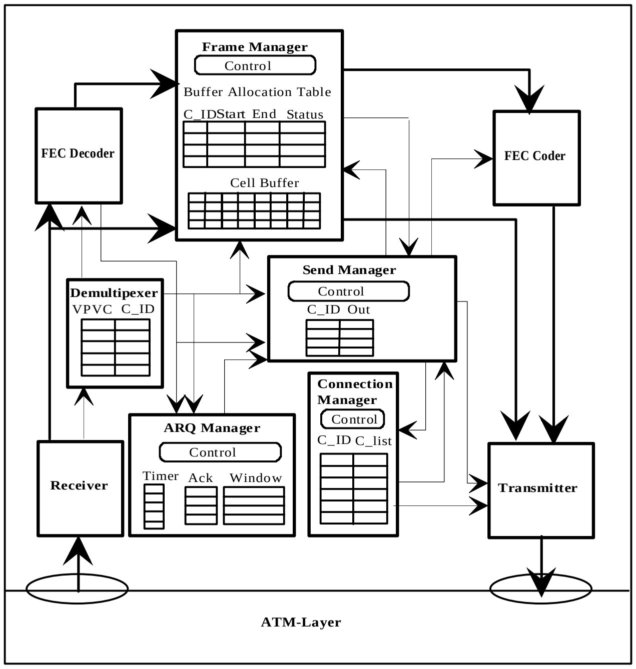

Figure 3: Architecture for implementation of the Group Communication Server

The functionality of GCSs is not necessarily restricted to pure servers inside the network. Instead, it is possible to combine an ATM end system with the functionality of a GCS.

Using the architecture of Figure 3 as a basis for the implementation of such a hybrid system, additional functionality is required to exchange AAL SDUs with higher layers. For this purpose, a host interface controller can be used to coordinate the communication between interface card and host memory via the processor bus.

4 Performance Assessment

4.1 Processing delay

In order to study processing delay and implementation complexity, the implementation of a GCS on a network adapter module with the following properties was investigated: An embedded controller (32 bit RISC processor with an average performance of 100 Mips), hardware support for segmentation and reassembly, hardware for CRC32 and hardware for FEC processing were assumed. It was found that such a system could handle a 150Mbit/s ATM interface at full speed. Based on a assembly level pseudocode specification of the modules, the processing delay of the first cell of a frame is 2.15µs when cell-based ARQ is selected in combination with FEC. The processing delay of a cell in the middle of a frame is 1.3µs, while the processing delay of the last cell of a frame is 1.47µs. The study also revealed that a implementation with FEC processing performed in software, using a 64 bit RISC Processor with 150 MHz clock frequency (Digital Alpha processor), will support the same line speed. Memory required for administration of the buffers are 64 bytes per frame for multiplexing and frame-based selective repeat. In order to support cell-based selective repeat, additional 32 bytes are required for storing the bitmaps of every frame with lost cells. For systems that observe a frame-loss rate of 10^{-2} for a time period in the order of one to ten round trip times, only 1% of the frames buffered in the system need the allocation of memory blocks for bitmaps. A software implementation which does not use FEC will reduce processing requirements by approximately 30%, while the transition from cell-based to frame-based ARQ will reduce the processing requirements by another 25%.

4.2 Assessment of error control schemes

It is important to know which error control scheme is best suited for a given situation. Analytical methods were applied in order to evaluate the achievable performance of the proposed error control schemes. The following four alternatives were compared: frame-based ARQ, frame-based ARQ combined with FEC, cell-based ARQ, cell-based ARQ combined with FEC.

For the following analysis, a system model was applied based on time slots T for the transmission of a single ATM cell. For a 155Mbit/s SDH link with a transfer rate of 149,76Mbit/s at the ATM layer, the mean cell interarrival time is T=2.831µs. The normalised path capacity S represents the number cells that may be stored on the links and in buffers of multiplexing equipment between the transmitter and a receiver. With RTT denoting the mean round trip time, S may be expressed as

$$S = \frac{RTT}{T} \quad . \tag{1}$$

In the following, a multicast scenario with a common link, an ATM-switch in which copying of cells is performed, and individual links to the receivers will be investigated. The path capacity of the common link will be called S_C, and the path capacity of the individual link to receiver number i will be called S_i. Similarly, the cell loss probability of the common link will be called q_C, and the cell loss probability of the individual links will be called q_i. It is assumed that cell losses due to multiplexing within ATM switches occur on these links. In order to take into account the queueing delays in the multiplexer buffers, the maximum queueing delay is used for an upper bound of the normalised path capacity.

Without FEC, the probability for successful delivery of a frame with a length of k cells to all receivers is given by

$$1 - Q = (1 - q_C)^k \cdot \prod_{i=1}^{N} (1 - q_i)^k \quad . \tag{2}$$

with Q denoting the frame loss probability. For identical cell loss probabilities on all links, (2) can be simplified to

$$1 - Q = (1 - q)^{k \cdot (N+1)} \quad . \tag{3}$$

where q denotes the cell loss probability and N denotes the number of receivers. Formula (2) uses the assumption of statistically independent cell losses. For schemes in which complete frames are discarded after the loss of a single cell, such losses lead to a higher frame loss probability Q than statistically dependent cell losses [21]. Therefore, (2) may also be used as a conservative approximation in the case of correlated cell losses.

In the next step, an upper bound and a lower bound for the achievable efficiency will be derived. A worst-case approximation (Weldon approximation of [22]) is used to determine a proper receiver buffer size for the selective repeat protocol. Shacham [23] shows that the mean receiver buffer occupancy is finite for the case when the probability to successfully transmit a frame approaches zero. The size of a finite window sufficiently large to achieve an efficiency close to the efficiency in the ideal case of infinite receiver buffers will be derived in the following. If m denotes the number of trials for successful transmission of a frame, the efficiency is in inverse proportion to the mean number of transmission trials per frame. Hence the normalised throughput efficiency η is

$$\eta = \frac{1}{\overline{m}} \quad . \tag{4}$$

4.3 Influence of the receiver buffer

For infinite receiver buffers, the probability that the receivers get a frame in the m = i-th transmission trial is geometrically distributed [24]:

$$P(m = i) = Q \cdot (1 - Q)^{i-1}, \; i \geq 1 \quad . \tag{5}$$

Then, a lower bound for the mean of the random variable m is given by

$$\overline{m}_{min} = \sum_{i=1}^{\infty} i \cdot P(m=i) = (1\text{-}Q) \sum_{i=1}^{\infty} i \cdot Q^{i-1} = 1\text{-}Q \cdot \frac{1}{(1\text{-}Q)^2} = \frac{1}{(1\text{-}Q)} \; . \tag{6}$$

Using (3) and (4), the maximum achievable efficiency is

$$\eta_{max} = (1 - Q) = (1 - q)^{k(N+1)} \quad . \tag{7}$$

4.4 Finite Receiver Buffer

To take into account the influence of limited receiver buffers on the achievable efficiency, Weldon's approximation [25] may be applied for receiver buffers which are capable of storing L path capacities (i.e., the receiver buffer holds L·S frames). It is assumed that all receiver buffers are empty at the beginning of the transmission, and that the probability of a frame being successfully received on the first L+1 transmission attempts is given by

$$P(m = i) = (1 - Q) \cdot Q^{i-1} \quad , \quad 1 \leq i \leq L + 1 \quad . \tag{8}$$

For every retransmission, S new frames have been sent before the next retransmission takes place. If more than L+1 transmission attempts are necessary, the frames that are meanwhile correctly received from the previous transmission attempts will have filled up the receiver buffers. Further frames have to be discarded due to receiver buffer overflow, so they also have to be retransmitted. Hence,

$$P(m = i + (i - (L + 1)) \cdot S) = (1 - Q) \cdot Q^{i-1}, \; L + 2 \leq i \leq \infty \; . \tag{9}$$

However, the Weldon approximation is a pessimistic approximation: the maximum number of S frames is lost only if S frames have been successfully transmitted between two transmission attempts of the considered frame. While this assumption works well for low frame loss probabilities, there is a growing deviation from the real efficiency at higher frame loss probabilities. For (8) and (9), an upper bound for the mean of the random variable m can be derived (see [22]) to

$$\overline{m}_{max} = \frac{1 + S \cdot Q^{L+1}}{1 - Q} \quad , \tag{10}$$

and a lower bound of the efficiency is given by

$$\eta_{min} = \frac{1 - Q}{1 + S \cdot Q^{L+1}} = \frac{(1 - q)^{k(N+1)}}{1 + S \cdot (1 - (1 - q)^{k(N+1)})^{L+1}} \quad . \tag{11}$$

Hence, the achievable efficiency is bounded by

$$\eta_{min} \leq \eta \leq \eta_{max} \quad . \tag{12}$$

4.5 Scenarios with Group Communication Servers

After investigating the influence of frame-based selective repeat schemes, it is of interest to compare them with the simpler Go-back-N-schemes, and to investigate the influence of Group Communication Servers. Go-back-N may be modelled by G/G/1 queues, with the service time being the virtual transmission time [24]. According to [26], the efficiency of a memoryless multicast go-back-N protocol is

$$\eta_{1:N,GBN} = \frac{1}{\overline{m}} = \frac{1-Q}{1+sQ} \tag{13}$$

This result can be used as a lower bound for the efficiency of frame-based Go-back-N-schemes employed in end systems and GCSs. According to a lower bound of the efficiency (11), the efficiency of a selective-repeat protocol with receiver buffers of one path capacity is

$$\eta_{1:N,SR} = \frac{1}{\overline{m}} = \frac{1-Q}{1+sQ^2} \tag{14}$$

The efficiency of the two retransmission modes in three different scenarios will be presented. Scenario 1 represents a basic 1:N multicast without GCS. Scenario 2 represents 1:N multicasting with a GCS that performs retransmissions as multicast. In scenario 3, the GCS uses individual VCs for retransmission. Figure 4 gives the results for a group of 100 receivers and a data rate of 622 Mbit/s. Two cases are distinguished. The upper diagram of figure 7 shows the efficiency for an overall distance of 1000 km (distance of 500 km from GCS to the receivers), and the lower diagram shows an overall distance of 505 km (distance of 5 km from GCS to the receivers). The figure shows that in all cases, the efficiency is increased significantly by the GCS. Highest efficiency may be achieved for scenario 3 and selective repeat. Scenario 2 improves significantly for a shorter distance between GCS and the receivers. Go-back-N retransmissions show acceptable performance only for moderate bandwidth-delay products as they occur in LANs. Regarding efficiency, scenario 3 and selective repeat should be selected. However, this solution requires the highest implementation complexity for end systems and GCS. Therefore, an approach that uses powerful error control mechanisms for WAN connections between GCSs and simple error control mechanisms for LAN connections between end systems and GCSs allows the efficient provision of high-performance multipoint services. At the same time, such a solution has a relatively low implementation complexity.

4.6 Bandwidth overhead of cell-based vs. frame-based selective repeat

This section describes how to identify in which cases the frame-based ARQ scheme and in which cases the cell-based scheme shows better performance, taking into account the influence of protocol overhead in frames and cells. This is the first time that the performance of cell-

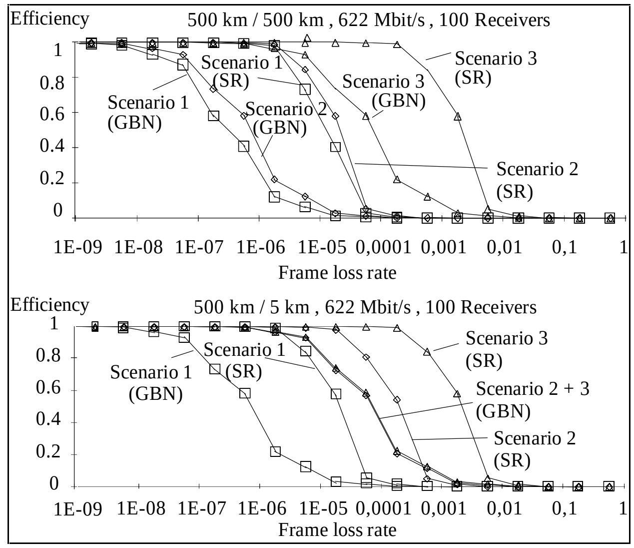

Figure 4: Efficiency in Scenarios with and without GCSs

based vs. frame-based selective repeat is investigated for reliable multicasting. In [21], a comparison of cell-based vs. frame-based selective repeat is presented for point-to-point-communication and for the assumption of unlimited receiver buffers. Here, the influence of a limited receiver buffer is also taken into account. The derivation uses the term overhead ratio as fraction of (protocol overhead + payload)/payload. In the following formulas, k_F denotes the overhead ratio of a frame-based scheme (e.g., RLMCP), in which the protocol overhead for the initial transmission attempt and the first retransmission are identical. For the cell-based scheme of RMC-AAL, the protocol overhead of the initial transmission is different from the protocol overhead of subsequent transmissions. Therefore, the overhead ratio k_{C1} is defined for the initial transmission, and the overhead ratio k_{C2} is defined for subsequent transmissions. Considering the protocol overhead of a frame-based selective repeat scheme in the derivation of (11), a lower bound for the efficiency for receiver buffers of one path capacity is given by

$$\eta_F = \frac{1 - Q}{k_F + SQ^2} \tag{15}$$

For the cell-based scheme of RMC-AAL and receiver buffers of one path capacity, a lower bound for the efficiency is given by

$$\eta_C = \frac{(1 - q)^N}{k_{C1} + (k_{C2} - k_{C1})(1-(1-q)^N) + S(1-(1-q)^N)^2} \tag{16}$$

The efficiency equilibrium q_{cf} (see Figure 5) can be obtained by solving the equation $\eta_F = \eta_C$, using

$$Q = (1-q)^{kN} \tag{17}$$

for the frame loss rate and substituting $q \cdot N$ by x results in a cubic equation

$$x^3 + Ax^2 + Bx + C = 0 \tag{18}$$

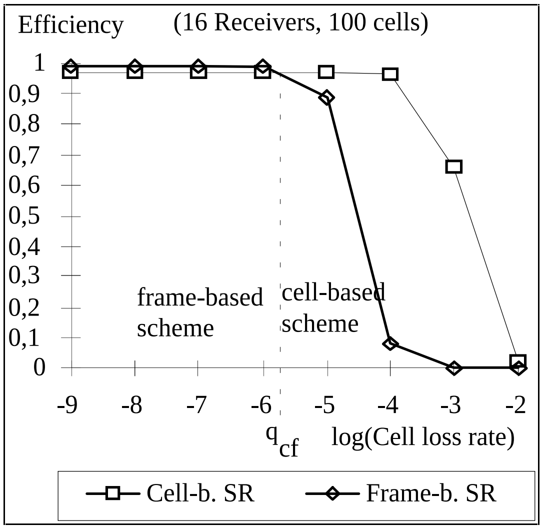

Figure 5: Efficiency equilibrium q_{cf} of Frame-based and cell-based SR

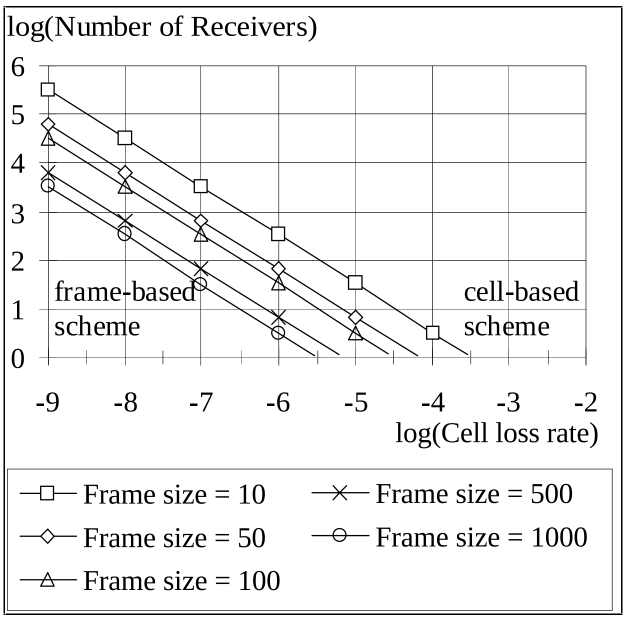

Figure 6: Frame-based vs. cell-based SR

with $A = \frac{S - Sk^2 - (k_{C2} - k_{C1})k}{Sk(k - 1)}$;

$B = \frac{k_F - k_{C1}n + k_{C2} - k_{C1}}{Sk(k - 1)}$; $C = \frac{k_{C1} - k_F}{Sk(k - 1)}$.

The analytical solution of a cubic equation is fairly cumbersome. It is not presented here, but can be found in literature (e.g., see [27]). After identifying the correct one of three possible solutions of (18), Figure 6 shows the resulting threshold for the efficiency equilibrium for a path capacity of 1750 cells (e.g., 150 Mbit/s over 1000 km).

For every frame size, a line of the figure separates the region with lower cell loss rates, where a frame-based scheme results in better efficiency, from the region with higher cell loss rates, where a cell-based scheme leads to better efficiency.

5 Conclusion

It was pointed out that a large number of alternatives exist for the provision of a reliable multicast service in ATM networks. Existing approaches have significant shortcomings, however there does not exist a single approach best suited for all scenarios that need to be considered. A new framework is presented which has the potential to fulfil many requirements. For small groups and low cell loss rates, a frame-based end-to-end error control is most appropriate. In this case, RLMCP or another lightweight protocol for reliable multicast can be used as SSCS for AAL5. In case of significant cell loss, large group sizes and higher path capacities, and also for applications with stringent real-time requirements, the new adaptation layer type called the Reliable Multicast ATM Adaptation Layer (RMC-AAL) is proposed. It is the first AAL protocol offering cell-based ARQ and FEC for reliable multicasting and has only one byte protocol overhead per cell. For better scalability and support of heterogeneous scenarios, the deployment of a new network element called the Group Communication Server (GCS) is proposed. It allows an hierarchical approach for multicast error control and the combination of different error control schemes. Investigating the implementation complexity revealed that that cell based error control schemes contribute very little to the processing load for error-free transmissions. By analysis, it was shown how the efficiency is affected by cell loss, path capacity, and number of receivers. A cell loss threshold was determined for which a cell-based scheme outperforms a frame-based scheme.

Acknowledgement

The author would like to thank Martina Zitterbart and the other members of the High Performance Networking Group for valuable discussions. The provision of the simulation tool by the Alta Group of Cadence Design Systems, and the support by the Graduiertenkolleg „Controllability of Complex Systems“ (DFG Vo287/5-2) is also gratefully acknowledged.

References

[1] Ohta, H., Kitami, T.: „A Cell Loss Recovery Method Using FEC in ATM Networks,“ IEEE Journal on Selected Areas in Communications, Vol. 9, No. 9, December 1991, pp.1471-1483

[2] Romanov, A.: „Some Results on the Performance of TCP over ATM,“ Second IEEE Workshop on the Architecture and Implementation of High Performance Communication Subsystems HPCS’93, Williamsburg, Virginia, U.S.A., September 1993

[3] ITU-TS Draft Recommendation I.363: „BISDN ATM Adaptation Layer (AAL) Specification,“ Geneva, 1993

[4] ITU-TS Draft Recommendation Q.2110: „BISDN ATM Adaptation Layer (AAL) - Service specific connection oriented protocol (SSCOP),“ Geneva, 1994

[5] Inoue, I., Morita, N.: „Partial-Frame Retransmission Scheme for Data Communication Error Recovery in B-ISDN“, Proceedings of 4th IFIP Workshop on Protocols for High-Speed Networks, pp. 328-348, Vancouver, Canada, August 10-12, 1994

[6] Shacham, N.; McKenny, P.: „Packet recovery in high-speed networks using coding,“ in Proceedings of IEEE INFOCOM '90, San Francisco, California, USA, June 1990, pp. 124-131

[7] Biersack, E. W.: „Performance Evaluation of Forward Error Correction in an ATM Environment,“ IEEE Journal on Selected Areas in Communication, Volume 11, Number 4, pp. 631-640, May 1993

[8] Ayanoglu, E. ; Gitlin, R.D. ; Oguz, N.C.: „Performance Improvement in Broadband Networks Using Forward Error Correction,“ Journal of High Speed Networks, vol. 2, 1993, pp. 287-304

[9] Santoso, H.; Fdida, S.: „Transport Layer Multicast: An Enhancement for XTP Bucket Error Control,“ in: Danthine, A.; Spaniol, O. (eds.): High Performance Networking, IV, IFIP, North-Holland, 1993

[10] Feldmaier, D.: „An Overview of the TP++ Transport Protocol Project,“ Chapter 8 in ‘High Performance Networks - Frontiers and Experience,’ Ahmed Tantawy (Ed.), Kluwer Academic Publishers, 1993

[11] Schatzmaier, R; Popescu-Zeletin; R..: „Providing Support for Data Transfer in a New Networking Environment,“ in D. Hutchison, A. Danthine, H. Leopold, G. Colson (Eds.): Multimedia Transport and Teleservices; International COST 237 Workshop, Vienna, Austria, November 13-15, 1994, Lecture Notes in Computer Science 882, Springer-Verlag, pp. 241-255

[12] Kure, O.; Sorteberg, I.: „XTP over ATM,“ Computer Networks and ISDN Systems, 26 (1993). pp. 253-262

[13] Braun, T.; Zitterbart, M.: „Parallel Transport System Design,“ 4th IFIP Conference on High Performance Networking HPN’92, Liège, Belgium, December 1992

[14] Schiller, J.; Braun, T.: „VLSI-Implementation Architecture for Parallel Transport Protocols,“ IEEE Workshop on VLSI in Communications, Stanford Sierra Camp, Lake Tahoe, California, U.S.A., September 1993

[15] Zitterbart, M.; Stiller, B.; Tantawy, A.: „A Model for Flexible High Performance Communication Subsystems,“ IEEE Journal on Selected Areas in Communications, Volume 11, Number 4, pp. 507-517, May 1993

[16] Goldstein, F.: „Compatibility of BLINKBLT with the ATM Adaptation Layer,“ ANSI Technical Subcommittee T1S1 Contribution T1S1.5/90-009, February 1990

[17] McAuley, A.: "Reliable Broadband Communication Using a Burst Erasure Correcting Code," Presented at ACM SIGCOMM '90, Philadelphia, PA, U.S.A., September 1990

[18] Wei, L.; Liaw, F.; Estrin, D.; Romanow, A.; Lyon, T.: „Analysis of a Resequencer Model for Multicast over ATM Networks“, Third International Workshop on Network and Operating Systems Support for Digital Audio and Video, San Diego, U.S.A., November 1992

[19] Carle, G.; Schiller, J.; Schmidt, C.: „Support for High-Performance Multipoint Multimedia Services,“ in D. Hutchison, A. Danthine, H. Leopold, G. Colson (Eds.): Multimedia Transport and Teleservices; International COST 237 Workshop, Vienna, Austria, November 13 - 15, 1994, Lecture Notes in Computer Science 868, Springer-Verlag, pp. 219 - 240

[20] Carle, G.; Siegel, M.; Design and Assessment of a Parallel High Performance Transport System; in Proceedings of Euro-ARCH '93; October 1993; Munich, Germany, pp. 407 - 418

[21] Bondi, A.; Lai, W.-S.: „The influence of cell loss patterns and overheads on retransmission choices in broadband ISDN,“ Computer Networks and ISDN Systems 26, pp. 585-598, 1994

[22] Weldon, E.: „An Improved Selective-Repeat ARQ Strategy,“ IEEE Trans. Comm., vol. COM-30, pp. 480-486, March 1982

[23] Shacham, N.: „Queuing Analysis of a Selective-Repeat ARQ-Receiver“, in Proceedings IEEE INFOCOM '87, San Francisco, California, February 1987, pp. 521-526

[24] Sabnani, K.: "Multidestination Protocols for Satellite Broadcast Channels", Ph. D. Dissertation, Columbia University, NY, U.S.A., 1982

[25] Aghadavoodi, J.: „Multicopy ARQ Strategies for Heterogeneous Networks,“ 4th IFIP International Workshop on Protocols for High Speed Networks, Vancouver, B.C, Canada, August 1994, pp.

[26] Gopal, I., Jaffe, J.: „Point-to-Multipoint Communication Over Broadcast Links,“ IEEE Trans. Commun., Vol. Com-32, No. 9, pp. 1034-1044, September 1984

[27] Bronstein, I., Semendjajew, K.: „Taschenbuch der Mathematik,“ (in German) Verlag Harri Deutsch, Thun und Frankfurt/Main 1989.