Design and Analysis of a Bandwidth Management Framework for ATM-Based Broadband ISDN

Kunyan Liu, Hongbo Zhu, David W. Petr, Victor S. Frost, Cameron Braun, William L. Edwards**

Telecommunications & Information Sciences Laboratory
Department of Electrical Engineering and Computer Science
The University of Kansas
2291 Irving Hill Road
Lawrence, KS 66045

**Sprint Corporation, Overland Park, Kansas*

[1]

Abstract

When designing and configuring an ATM-based B-ISDN, it remains difficult to guarantee the Quality of Service (QoS) for different service classes, while still allowing enough statistical sharing of bandwidth so that the network is efficiently utilized. These two goals are often conflicting. Guaranteeing QoS requires traffic isolation, as well as allocation of enough network resource (e.g. buffer space and bandwidth) to each call. However, the statistical bandwidth sharing means the network resource should be occupied on-demand, leading to less traffic isolation and minimal resource allocation. We address this problem by proposing

[1] An abbreviated version of this paper was presented at IEEE International Conference on Communications, 1996

and evaluating a network-wide bandwidth management framework in which an appropriate compromise between the two conflicting goals is achieved. Specifically, the bandwidth management framework consists of a network model and a network-wide bandwidth allocation and sharing strategy. Implementation issues related to the framework are discussed. For real time applications, we obtain maximum queueing delay and queue length which are important in buffer design and VP (Virtual Path) routing.

1 Introduction

Broadband Integrated Service Digital Networks (B-ISDN) utilizing Asynchronous Transfer Mode (ATM) technology are expected to support a wide variety of services. As defined by ATM Forum, the different types of service supported by ATM are categorized into four service classes [2] [6]: Constant Bit Rate (CBR), Variable Bit Rate (VBR), Available Bit Rate (ABR), and Unspecified Bit Rate (UBR). According to the ATM Forum, CBR and real-time VBR connections have stringent delay and Cell Loss Ratio (CLR) requirements. Moreover, the CBR service class is designed for circuit switching emulation which requires a constant bandwidth capacity for each call. The traffic rate for a VBR connection may fluctuate around its average rate but not exceed its peak cell rate (PCR). The traffic rate for an ABR connection can be adjusted in real time, and its Minimum Cell Rate (MCR) is specified. An UBR source may send as fast as it desires (up to its PCR), but the network does not guarantee any QoS for it. Traffic parameters for a call are specified in its ATM traffic descriptor. An ATM traffic descriptor is defined as a generic list of traffic parameters that can be used to capture the traffic characteristics of an ATM connection [1]. The ATM Forum standard requires that *policers* be used at the User-Network Interface (UNI) to ensure all traffic entering the network conforms to the parameters specified in the traffic descriptor. The generic cell rate algorithm (GCRA) has been standardized as a general policing scheme [1].

Recent activities in ATM Forum indicate that one of the major challenges for the broad deployment of ATM-based B-ISDN is how to compromise between two conflicting goals, i.e., guaranteeing performance for each service class, while still allowing enough statistical multiplexing so that the network is efficiently utilized. In order to achieve these two goals, the ATM research community has proposed numerous control and management schemes [9] [12] [11] [5]. However, schemes

for different purposes are often treated independently and lack the capability of co-operating with each other. What is needed is, therefore, a bandwidth management framework under which the network can be efficiently utilized, and meanwhile, acceptable QoS can be achieved for all service classes.

Although it seems unlikely to optimize bandwidth management by taking into account all aspects of traffic behavior and performance requirements, it is possible to reach a good compromise between the goals of providing both performance guarantees and high network utilization. This paper is organized as follows. Section 2 and 3 constitute the proposed bandwidth management framework: Section 2 proposes a network model and the corresponding VP assignment policy; Section 3 presents the bandwidth allocation and sharing strategy and discusses corresponding cell-level schemes and switch architecture. Section 4 evaluates the maximum queueing delay and CLR performance for CBR and VBR service classes which are expected to mainly support real time applications. Section 5 draws a conclusion for the paper.

2 Network model

2.1 The partitioning of core and edge networks

We propose a network model in which the ATM-based B-ISDN is partitioned into core and edge networks as shown in Figure 1. The primary function of the edge networks is to provide broad-band access to the user through the UNI and to perform cell switching in the local area. The core network functions as the backbone network carrying concentrated traffic between edge networks. The interface between the core and edge network is provided by special edge nodes (gateways).

Note that the core and the edge networks are still part of a unified ATM network, and should be able to cooperate in terms of bandwidth management, congestion control, and other administration issues through network-network interfaces [7] [8].

The design of the core network will apply the Virtual Path (VP) concept in which ATM cells are processed based on Virtual Path Identifier (VPI) values (Figure 2). The VP concept [1] [3] has been developed to support semi-permanent connections in a large scale backbone network which transports a large number of simultaneous user calls carried by Virtual Circuits (VCs). A VP starts at an edge gateway and terminates at another edge gateway (see Figure 2). In the core network, available network resources, such as bandwidth and buffer space, can be managed simply and efficiently on a per-VP basis. On the other hand, the edge networks will carry a smaller number of simultaneous VCs, and will handle traffic on a call by call basis in order to process call arrivals and to setup and tear down individual VCs using Virtual Circuit Identifier (VCI) values (Figure 2).

As stated above, a VP is identified by its VPI values in the core network. There have been two kinds of VPI management methods [14]:

- Global VPI assignment, in which VPIs are managed centrally, each VPI has global significance, and each VPI corresponds to a route in the network. No VPI translation is needed at the core switches.

- Local VPI assignment, in which the VPIs have only local significance which is associated with each physical link and should be translated at each core switch . A VP is therefore identified by a series of physical link ID and VPI pairs.

Although the first method is simple and easy to implement, it imposes a limit on the total number of VPs within the network. With current 12-bit VPI definition [1], only 4096 VPs can

possibly exist. This is far from adequate for a large-scale network. Therefore, we prefer the local VPI assignment method in which, given the ability of identifying each port through its port ID, each core switch may support up to 4096 VPs on each input/output link. Now that VPI has only local importance (to identify a VP from other VPs on the same link), the same VPI can be reused in the network. As a result, the network is able to support any number of VPs, given a route layout such that no more than 4096 VPs exist on a single link. This gives the possibility to support fully-meshed VP network using current fiber network topology.

2.2 The VP assignment policy

Currently a number of VP layout and assignment schemes have been proposed [15] [16], which differ in the following ways:

- The connectivity of the VP network, i.e., should it be fully-meshed or sparsely connected, such as in a star or ring topology.

- How to map the various services to the VPs. One extreme is to use the same VP for all service classes, thus fewer VPs are needed. However, the task of guaranteeing QoS for all service classes in the VP could be difficult. The opposite is to have a separate VP for each service class, or even for each different QoS requirement within the same service class. Although QoS control is easier in this scheme, the total number of VPs needed can be very large.

We propose a fully-meshed scheme in which there should be at least two VPs assigned between each edge-node-pair (denoted as an Origin-Destination pair, or O-D pair), one for VBR and CBR

service, and the other for ABR and UBR service. Other VPs may also exist for alternative routing or other management considerations.

The VP assignment policy described above is based on the following considerations:

- In the fully-meshed VP network, pre-assigned VPs exists between all edge networks, and the core nodes can be easily implemented by ATM cross-connectors. No complicated VC level operations such as add/drop or rerouting are necessary. Meanwhile, even if the number of edge networks grows, the VP network can still scale well given the local VPI management scheme discussed in the last section.

- The mapping of service classes to VPs should be able to achieve a good balance between QoS achievement and complexity. Thus we need to carefully inspect the nature of service classes before determining how to map them into VPs. Real-time VBR [2] and CBR connections have similar performance parameters in terms of delay and CLR. On the other hand, ABR sources are expected to adapt their rates according to network states and do not require stringent delay performance. Separating ABR traffic from the VBR/CBR VP ensures that ABR rate changes do not affect the performance of CBR and VBR service classes. The nature of UBR services indicates that no network resources should be allocated to UBR connections, and consequently, allocating separate VPs to UBR connections is unnecessary. However, the network must provide the necessary isolation (described in the next section) between UBR and other service classes so that the traffic from UBR sources does not affect the performance of other users. Practically, once enough isolation is provided, UBR connec-

[2] A Non-real-time VBR connection can be viewed as a Real-time VBR with a large Cell Delay Variation Tolerance (CDVT) parameter. Therefore, Non-real-time VBR VCs can be integrated on VBR/CBR VPs.

tions may share the same VP with any other service classes. We choose to integrate UBR with ABR on the same VP because of the similar "best effort" nature for the two service classes.

3 Bandwidth management strategy

In order to achieve a successful bandwidth management framework, it is necessary to incorporate efforts at both the cell level and the network design level. In this section we will first introduce the basic concepts of the proposed framework and the cell-level schemes to support them, then look into the network design level issues. Furthermore, we also present a sketch of a possible implementation of the proposed framework.

3.1 Basic ideas: bandwidth allocation vs. reservation

In traditional telecommunication networks, usually a certain amount of bandwidth is *reserved* for all connections, i.e., each connection will always be given and only be able to use the portion of bandwidth explicitly assigned to it. For example, in a TDM system, each connection has (and pays for) its own digital channel and the associated fixed bandwidth. No connection may use the bandwidth on any other channels, even if there is no traffic on them at the time. Since the majority of traffic in those networks is CBR (voice connections), the *reservation-based* scheme is sufficient. However, VBR and "best effort" traffic (ABR and UBR) will play very important roles in ATM networks, and it is difficult, if not impossible, to support these services efficiently by reservation-based scheme.

In order to achieve both QoS guarantee and high network utilization for all service classes in ATM networks, a new kind of bandwidth management , which we call bandwidth *allocation*, must be introduced. In a bandwidth allocation-based scheme, each connection is allocated a certain amount of bandwidth (which could be zero), and

- Each connection is *guaranteed* to have access to its allocated bandwidth, whenever it has something to send.
- Unused bandwidth is available to other connections.
- Consequently, a connection sometimes can be using bandwidth exceeding its allocation, but *only* when other connections are *not* using their allocation.

Note that different services emphasize different aspects of the allocation-based scheme. For example, since the QoS requirements for CBR/VBR connections need to be guaranteed all the time once they are admitted into the network, it is necessary to allocate them an amount of bandwidth that will guarantee that the QoS requirement will be met. On the other hand, ABR connections would be allocated only enough bandwidth to guarantee their minimum cell rate (MCR), since they are supposed to utilize the bandwidth "spared" by CBR and VBR connections. Similarly, UBR connections would likely not be allocated any bandwidth.

3.2 Cell-level supporting schemes

3.2.1 Cell scheduling schemes

Various cell scheduling schemes, such as Weighted Fair Queueing [17], Round-Robin, and Virtual Clock [18] have been proposed for ATM networks. Among them, Weighted Round Robin(WRR)

[19], [20], [21] seems to be the most promising algorithm to support allocation-based bandwidth management.

The basic idea of WRR can be described as follows: there are multiple incoming connections, each of them with a separate queue. One output link is shared among all connections, and the access to it is controlled by a server. The server serves all the queues in the order decided by a circular schedule, in which each queue has a certain number of entries. If the current queue is "inactive", i.e., it does not have any cell to be served (transmitted), the server will then move to poll the next queue on schedule, until it finds an active connection. Hence the cell slot will not be wasted unless all connections are inactive.

The WRR algorithm has several notable features that make it ideal for our purpose:

Guaranteed allocated bandwidth: Given the following parameters:

- CS: one cell slot, i.e., the time to serve one cell
- M: the total number of cell slot entries in a schedule
- W: number of schedule entries (allocated slots) for a particular queue

The allocated bandwidth BW (in Cells/Sec) for the target queue can be obtained as:

$$BW = \frac{W}{M} \times \frac{1}{CS} \tag{1}$$

Automatic sharing of unused bandwidth: If a connection does not have enough cells in its queue to use up all its schedule entries during a serving cycle, the WRR server will use these "left-over" cell slots to serve other active connections. Thus bandwidth sharing is achieved.

Intrinsic fairness: In the WRR algorithm, the "left-over" cell slots will be automatically given to the active connections in proportion to their allocated weight. In this sense, it provides a means to distribute the spared bandwidth fairly.

One version of WRR server is the *distributed WRR server*. Here, *distributed* means the schedule entries for a connection are evenly distributed within the schedule. Besides the common WRR characteristics, the distributed WRR also helps to smooth the traffic in the multiplexing/demultiplexing procedure. Therefore, we will assume the distributed WRR in the remainder of this paper.

3.2.2 Other cell-level schemes

Besides the cell scheduling schemes such as WRR, there are other type of schemes that can be used as either supplementary or alternative means in the bandwidth management framework, especially the following:

Traffic policing schemes: Policing, or Usage Parameter Control (UPC) has long been recognized as an effective way to enforce the user-network traffic contract. At present the Generic Cell Rate Algorithm (GCRA), which is based on a leaky-bucket algorithm, has been chosen by ATM Forum [1] as the definition of traffic conformance. The policing function at the UNI determines if the individual cells are conforming to the traffic contract and either drops violating cells, or marks them with Cell Loss priority (CLP) = 1 (Conforming cells carry CLP = 0). Since the GCRA includes bandwidth-related parameters such as Peak Cell Rate (PCR) and Sustainable Cell Rate (SCR), it can also be used in bandwidth management.

Buffer management schemes: In the case that multiple connections share the same physical buffer, buffer management schemes (also known as space priority schemes) [22] are necessary to ensure the proper buffer access priority of different services. The most commonly used space priority schemes are *partial buffer sharing* (also known as *nested threshold cell discarding*) [23] and *push-out queue* [24]. Although the implementation of the two schemes are different, both of them support selective discarding of individual cells. Therefore if the cells are marked as CLP = 0 or CLP = 1 by policing functions, these schemes can be used to protect CLP = 0 traffic from CLP = 1 traffic by giving higher buffer access priority to CLP = 0 cells.

3.3 Network design level issues

Given the appropriate cell-level schemes, the next question is how to structure the network based on them. Again, CBR/VBR traffic and ABR/UBR traffic need to be treated differently because of their different nature.

3.3.1 Bandwidth management for CBR/VBR traffic

As we have mentioned before, CBR/VBR services generally require a worst-case QoS guarantee, and the primary way to achieve this is to allocate enough bandwidth to each connection (sufficient buffer space should also be allocated). Thus the admissible traffic load on a VP is determined by the total amount of bandwidth allocated to that VP. From the traffic engineering point of view, this amount should be determined by relatively long-term considerations, such as physical link capacity, traffic forecast and estimation methods, and may be updated in a time scale such as hours

or days, rather than on a call-by call basis.

The main advantage of this long-term allocation of VP bandwidth is that it simplifies the VC-level connection admission control (CAC) and offers traffic isolation to provide performance guarantees for each VP. The CAC is simplified because the decision of whether to accept a CBR or a VBR call can be made at the corresponding source edge gateway by comparing the bandwidth requirement of the new call and the available amount of allocated bandwidth on the VP which is to carry the new call. In our framework, an incoming CBR call is admitted if its PCR can be accommodated by the VP, and an incoming VBR call is admitted if its SCR can be accommodated by the VP. Determination of an appropriate value for SCR is a challenging, ongoing research topic. One possibility is to use equivalent bandwidth [12]; another is explored in [27]. Also notable is that under this strategy, the optimization of VP routes becomes possible using mathematical programming techniques.

Although ABR/UBR VPs should be able to use spare bandwidth from CBR/VBR VPs, bandwidth sharing among CBR/VBR VPs is undesirable. The traffic entering a CBR/VBR VP should be restricted to the allocated VP bandwidth to ensure that the VBR rate fluctuation does not degrade the performance of CBR VCs which are integrated on the same VP. To clearly understand this, note that the spare cell slots from other VPs at one node may not be available at the downstream nodes. Consequently, the extra VBR cells transmitted using spare cell slots from other VPs may be throttled at a downstream node, causing CBR connections sharing the same VP queue to incur more delay variation and even cell loss. Note, since CAC decision for CBR and VBR should always be based on the allocated VP bandwidth even if there is spared bandwidth in the network, the above restriction will not impact the network capability to accommodate CBR/VBR service.

However, we believe that in order to fully exploit the possibility of statistical multiplexing, it is still desirable to have VC-level bandwidth sharing inside each CBR/VBR VP.

3.3.2 Bandwidth management for ABR/UBR traffic

Through the allocation-based cell scheduling schemes, ABR/UBR VPs will be able to utilize the spared bandwidth from CBR/VBR VPs in the network. As a result, only the small amount of bandwidth corresponding to MCR of ABR service is necessary for ABR/UBR VPs. The available bandwidth for ABR/UBR VPs beyond that allocated for MCR is determined by the traffic load on the CBR/VBR VPs in the network, which is changing constantly. Consequently, in order for ABR sources to use this bandwidth and still achieve low CLR, a mechanism is necessary to feed back the bandwidth information to the ABR traffic sources. The mechanism could be the Resource Management (RM) cell procedure currently being developed by the ATM Forum [2]. Note that the RM cells are needed at both VP and VC level. Generally, the VP-level RM cells carry the available VP bandwidth information collected in the core network to the edge gateways, where the bandwidth is further allocated to individual VCs and sent to the ABR sources by VC-level RM cells. Recent research in this area indicates that efficient algorithms can be developed to control the ABR source rate in order to achieve both low cell loss and good bandwidth utilization [26] [25].

Since the allowable transmission rate of ABR sources is subject to flow control, the ABR UPC parameters need to be updated accordingly. That means *dynamic* UPC functions rather than the *static* UPC in the CBR/VBR case. As to the UBR cells, since the network will not provide any QoS guarantee to them, it is reasonable to mark all of them as $CLP = 1$.

As previously indicated, by utilizing space priority schemes, it is safe to let ABR and UBR connections share the same buffer, and hence the same portion of bandwidth.

3.3.3 Summary on strategy

The conclusions from the above discussion can be summarized as follows:

1. VP bandwidth for CBR/VBR VPs should be determined semi-permanently (update intervals measured in hours or days).

2. Once the VP bandwidth is determined, the traffic entering CBR/VBR VPs should be throttled to the VP bandwidth at the ingress edge gateway.

3. VC-level bandwidth sharing should still be supported within each CBR/VBR VP

4. High network utilization can be achieved by letting ABR/UBR VPs “fill-in” the bandwidth gap left by CBR/VBR VPs on the link, through VP-level allocation-based cell scheduling schemes.

5. Dynamic UPC function is necessary for ABR traffic.

6. By introducing space priority schemes, ABR and UBR traffic may safely share the same buffer in the network. However, since many ABR control algorithms [25] rely on queue-fill information, some kind of meachnism, such as a separate ABR cell counter, might be necessary to keep track of the number of ABR cells in the shared buffer.

3.4 An implementation sketch

To further illustrate how the ideas discussed above can be incorporated into a bandwidth management framework, we here present a sketch of a network implementation based on them. The implementation consists of three parts: *the ingress function of edge gateway*, *the core switch*, and *the egress function of edge gateway*.

3.4.1 The ingress function of edge gateway

As shown in Figure 3, each VC has a GCRA policer [1] at the UNI to ensure that the incoming traffic is conforming. The conforming cells are given high cell loss priority (CLP = 0) and the non-conforming cells are marked as CLP = 1. The GCRA parameters of ABR VCs should be dynamically adjustable to accommodate the fluctuation of available bandwidth in the network. In addition, all UBR cells are marked as CLP = 1 in order to provide necessary isolation for ABR cells.

At the ingress of the edge gateway, there are two stages of distributed WRR servers. At the first stage, CBR/VBR VCs are multiplexed into CBR/VBR VPs by using a VC-based distributed WRR server for each CBR/VBR VP. Each CBR (VBR) VC has a separate queue and is allocated a W corresponding to its PCR (SCR) bandwidth. The output rate of the distributed WRR is fixed at the allocated VP bandwidth; thus the traffic entering a CBR/VBR VP is limited according to the policy specified previously. Since each CBR/VBR VP is throttled to the allocated VP bandwidth at ingress, and the core switches provide at least the allocated VP bandwidth, core switch buffers for VBR/CBR VPs can be very small (see numerical example in section 4). The ABR/UBR VCs do not have per-VC queueing and VC-based WRRs at the first stage. The ABR and UBR VCs on

the same VP are simply mixed into a single ABR/UBR VP queue. At the second stage, the VPs are routed into the core network.

3.4.2 The core switch

The core switches are now simply ATM cross-connectors performing VP multiplexing/switching by using allocation-based cell scheduling schemes such as WRR.

3.4.3 The egress function of edge gateway

As shown in Figure 4, the VCs are demultiplexed from VPs at the egress edge switch, and routed to their destination UNI. There is a VC-based WRR server at each UNI. Again, each CBR (VBR) VC has a separate queue, and allocated a W corresponding to its PCR (SCR). The ABR/UBR VCs on the same UNI share the same queue.

All queues are implemented as push-out queues ([24]), i.e, the arriving CLP=0 cells can "push out" those CLP = 1 cells in the queue if the queue is full. Therefore, the throughput of CLP=0 cells will be essentially unaffected by CLP=1 cells. Note, the CLP=1 cells may still enter the network given any available bandwidth, but without any performance guarantee.

4 Maximum queueing delay and maximum queue length for CBR/VBR services

Real-time applications carried by CBR/VBR VCs will have stringent delay and cell loss requirements. We evaluate maximum queueing delay and maximum queue length for CBR/VBR services

under the proposed framework and implementation. These worst case performance evaluations will have important impacts in buffer size design and VP routing decisions.

On the other hand, ABR services only have a CLR objective which will be determined by the specific rate-based flow control deployed. The UBR service does not have any QoS requirement. For these reasons, a detailed performance evaluation for ABR/UBR services will not be included here.

4.1 Maximum queueing delay

According to the proposed network model, the cell transfer delay consists of three elements: the propagation delay on transmission links, cell routing delay in switches, and queueing delay at WRR servers and VP multiplexers/switches.

However, the first two will remain relatively fixed after the connection is established. Therefore we only focus on *maximum queueing delay* performance.

As shown in Figure 5, the end-to-end connection (A-E) consists of a VC-based WRR between (A-B) and a VP Multiplexer (VP MUX) between (B-C) at the ingress edge, N VP MUXes between (C-D) in the core , and a VC-based WRR between (D-E) at the egress edge. The end-to-end queueing delay $Delay_{A-E}$ is the sum of queueing delays incurred at ingress edge VC-based WRR, egress edge VC-based WRR, and all VP MUXes:

$$Delay_{A-E} = Delay_{A-B} + Delay_{B-D} + Delay_{D-E} \tag{2}$$

For a particular CBR or VBR VC between $\{$A,E$\}$, the worst case end-to-end delay performance occurs under the following conditions:

1. At both ingress edge and egress edge VC-based WRRs, the target VC can only be served according to its allocated bandwidth (PCR for CBR VC and SCR for VBR VC).

2. All cell slots of the ingress edge VC-based WRR are used, i.e., the rate of traffic entering a target VP reaches the upper limit.

3. The target VP can obtain only its allocated bandwidth at each VP MUX it passes.

With these conditions in mind, the maximum queueing delay for a VBR connection is evaluated as follows.

Suppose a target VBR VC conforms to $GCRA(T, 0)$ (T is $\frac{1}{PCR}$, CDVT is zero) and $GCRA(T_s, \tau_s)$ (T_s is $\frac{1}{SCR}$, τ_s is the burst tolerance) [1]. Accordingly, the size of a maximum conforming burst [1] is: $MBS = 1 + \frac{\tau_s}{T_s - T}$.

At the ingress edge VC-based WRR server, the W for the target VBR VC is, according to its SCR, $W = M \times CS \times \frac{1}{T_s}$. As shown in Figure 6, when the last cell (C_{MBS}) of a maximum burst arrives, the queue will reach its maximum length; consequently, this cell will incur the longest queueing delay. The maximum queueing delay at the ingress edge is therefore:

$$MD_{VBR}^{A-B} = t_3 - t_2 = (MBS \times T_s + CS_{in}) - [(MBS - 1) \times T] = T_s + \tau_s + CS_{in}, \quad (3)$$

where CS_{in} is the cell slot at the ingress edge WRR. Given condition 3, the traffic on the corresponding VP conforms to $GCRA(\frac{1}{BW_{VP}}, 0)$. Accordingly, the maximum queueing delay at the *i*th VP MUX is: $MD_i = \frac{1}{BW_{VP}} + CS_i, \ \ where \ \ i = 1, \ldots, N+1$, and CS_i is the cell slot at the *i*th VP MUX. Thus the maximum queueing delay between $\{$B,D$\}$ is:

$$MD_{VBR}^{B-D} = \sum_{i=1}^{N+1} MD_i = \frac{N+1}{BW_{VP}} + \sum_{i=1}^{N+1} CS_i \quad (4)$$

Viewed by the egress edge VC-based WRR given condition 1, 2, and 3, the target VBR VC is $GCRA(T_s, 0)$ conforming, i.e., the "bursty" VBR VC becomes constant-bit-rate after it passes the ingress edge VC-based WRR. Therefore, the maximum queueing delay at the egress edge is:

$$MD_{VBR}^{D-E} = T_s + CS_{out}, \tag{5}$$

where CS_{out} is the cell slot at the egress edge WRR. Therefore, the maximum end-to-end queueing delay for the target VBR VC is:

$$MD_{VBR}^{A-E} = MD_{A-B} + MD_{B-D} + MD_{D-E} = 2T_s + \tau_s + CS_{in} + CS_{out} + \frac{N+1}{BW_{VP}} + \sum_{i=1}^{N+1} CS_i \tag{6}$$

A CBR VC can be thought as a VBR VC with $T_s = T$,$\tau_s = 0$. So for CBR VCs, the end-to-end maximum queueing delay can be obtained by simplifying equation (6):

$$MD_{CBR}^{A-E} = 2T + CS_{in} + CS_{out} + \frac{N+1}{BW_{VP}} + \sum_{i=1}^{N+1} CS_i \tag{7}$$

4.2 Maximum queue length

The following evaluation concentrates on the maximum queue length for CBR and VBR connections at each WRR server. If each queue size is designed to be at least the maximum queue length, there will be no buffer overflow for conforming CBR/VBR traffic.

For a $GCRA(T, \tau)$ conforming cell stream served by a distributed WRR with allocated $W = \frac{1}{T} \times M \times CS$, the relationship between the maximum queueing delay D_{max} and the maximum queue length MQL is: $D_{max} = MQL \times T + CS$. Combined with the result of equation 3, the maximum queueing length can be obtained as:

$$MQL = 1 + \frac{\tau}{T} \tag{8}$$

Using Equation 8, the maximum queue length at each stage of a CBR or VBR connection can be obtained as follows:

At ingress VC-based WRR: MQL for CBR VCs is 1, while MQL for VBR VCs is $1 + \frac{\tau_s}{T_s}$.

At ith VP MUX: Because the output rate of the ingress edge VC-based WRR is limited at the corresponding allocated VP bandwidth, the traffic entering the VP at reference point B must be $GCRA(\frac{1}{BW_{VP}}, 0)$ conforming. Note between B and the *i*th VP MUX, the maximum queueing delay is $\sum_{j=1}^{i-1} CS_j + \frac{i-1}{BW_{VP}}$, and the minimum queueing delay is $\sum_{j=1}^{i-1} CS_j$. Therefore, the cell delay variation (CDV) [3] introduced between B and the *i*th VP MUX is:

$$CDV_{B-i} = \frac{i-1}{BW_{VP}} \qquad i = 1, \ldots, N+1 \tag{9}$$

Therefore the total CDV at the *i*th VP MUX is:

$$CDV_i = CDVT_B + CDV_{B-i} = 0 + CDV_{B-i} = \frac{i-1}{BW_{VP}} \qquad i = 1, \ldots, N+1 \tag{10}$$

Consequently, the maximum queue length at the *i*th VP MUX is obtained as:

$$MQL_i = 1 + CDV_i \times BW_{VP} = i \qquad i = 1, \ldots, N+1 \tag{11}$$

Note that equation 11 is true for both CBR and VBR VCs.

At egress edge: For a particular VBR VC, the maximum queueing delay at the egress edge denoted by D_{max}^{egress} is the end-to-end maximum queueing delay (equation 6) subtracting the minimum ingress edge delay and minimum core delay. Therefore:

$$D_{max}^{egress} = 2T_s + \tau_s + CS_{out} \tag{12}$$

[3] This notion of CDV is defined by the ATM Forum [2] as peak-to-peak CDV

Noting that $D_{max}^{egress} = MQL_{VBR}^{egress} \times T_s + CS_{out}$, the maximum queue length for VBR VCs at egress edge should be:

$$MQL_{VBR}^{egress} = 2 + \frac{\tau_s}{T_s} \tag{13}$$

Consequently, for CBR VCs, $MQL_{CBR}^{egress} = 2$.

4.3 Numerical example

The ACTS (Advanced Communications Technology Satellite) ATM Internetwork (AAI) is an ARPA research network providing wide area ATM connectivity . Initially, AAI consists of three core switches and seven edge networks (see Figure 7). All seven edge networks use a FORE System's local area ATM switch as edge gateways. Initially, all links (including access links) have DS3 (nominal 45Mb/s) capacity. The AAI is supporting research in the areas of network signaling, congestion management, multicast, gateways to non-ATM LANs, etc.

Taking the AAI network configuration of figure 7 as an example, consider a hypothetical VP traversing two core switches from NRL (Naval Research Lab) at Washington, D.C. to TIOC (Technology Integration and Operation Center) at Sprint Corporation, Overland Park, Kansas. Suppose at a certain moment, the VP is carrying ten 64kb/s CBR voice channels, one VBR MPEG video channel, and one VBR non-MPEG video channel. Table 1 shows the PCR, SCR and τ_s parameters for each type of call, where all parameters were selected based on by measurements from real traffic trace data. It is assumed that no traffic shaping function is used by the video sources, so the PCR of video sources is the access link rate. The SCR and τ_s for video calls are obtained from real trace data by using tools we have developed [27]. Also, assume the VP is allocated a bandwidth of 19.31Mb/s (45544 cells/sec) which is the sum of ten voice PCRs and two video SCRs (see Table

1).

Based on the analysis conducted in the previous section, the maximum queueing delay and queue length performance is presented in Table 2. The results show that under the proposed bandwidth management framework, these services require reasonably small buffer sizes and can obtain satisfactory queueing delay performance on the current AAI network topology. It should be noted that all these figures are derived from worst-case analysis; in reality the performance could be better due to VC-level bandwidth sharing inside the target VP.

5 Conclusion

We have proposed a bandwidth management framework for ATM-based B-ISDN. The bandwidth management framework consists of a network model and a bandwidth allocation strategy. Here the network is partitioned into core and edge networks. The advantage of this partitioning has been discussed. The network bandwidth is allocated in such a way that each VP is semi-permanently allocated a certain amount of bandwidth, while statistical bandwidth sharing may still be allowed among different VPs and VCs. The VP routes can be optimized using existing optimization techniques.

Cell scheduling and queueing implementations were discussed. The major elements of our framework related to implementation are the use of distributed WRR servers, push-out queues, and GCRA policers. Under the proposed implementation, maximum end-to-end queueing delay and cell loss performance have been evaluated for CBR and VBR connections. We conclude that based on the proposed bandwidth management framework, all ATM service classes can be served with reasonable QoS guarantees, the CAC procedures can be easily implemented, and potential

rate-based ABR congestion control [25] [6] can be easily incorporated.

References

[1] The ATM Forum Technical Committee, *User-Network Interface (UNI) Specification Version 3.1*, 1994.

[2] The ATM Forum Technical Committee, *Traffic Management Specification Version 4.0*, AF-TM 0056.000, April, 1996.

[3] Martin de Prycker, *Asynchronous Transfer Mode - Solution for Broadband ISDN*, 2nd edition, Ellis Horwood, 1993.

[4] D. V. Batorsky, D. R. Spears and A. R. Tedesco, "The Evolution of Broadband Network Architectures," *Proc. IEEE Globecom'88* pp. 367-373, 1988.

[5] A. E. Eckberg, "B-ISDN/ATM Traffic and Congestion Control," *IEEE Network Magazine*, Vol. 6 No. 5, pp. 28, September 1992.

[6] Raj Jain, "Congestion Control and Traffic Management in ATM Networks: Recent Advances and A Survey," *Invited submission to Computer Networks and ISDN Systems*, Draft Version, January 26, 1995.

[7] ATM Forum/95-0221R2, *Draft PNNI Signaling*, 1995.

[8] J. Totzke and J. Welscher, "A prototyped Implementation of B-ISDN Signaling at the Network Node Interface," *Proc. IEEE Globecom'95*, Vol. 1, pp. 252-257, 1995.

[9] C. J. Chang, R. G. Cheng, "Traffic Control in an ATM Network Using Fuzzy Set Theory," *Proceedings of IEEE Infocom'94*, Vol. 3, pp. 1200, 1994.

[10] A. Charney, D.D. Clark, R. Jain, "Congestion Control with Explicit Rate Indication," *Proc. IEEE ICC'95*, June 1995.

[11] The ATM Forum Technical Committee, *Flow Controlled Virtual Connections - Proposal for ATM Traffic Management*, September 1994.

[12] R. Guérin, H. Ahmadi, and M. Naghshineh, "Equivalent Capacity and Its Application to Bandwidth Allocation in High-Speed Networks" *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 7, September 1991.

[13] V. S. Frost, M. T. Mullen, "Dynamic Bandwidth Allocation for B-ISDN Based on End-to-End Delay Estimates,"*Proc. IEEE ICC'92*, Volume 1, pp. 225, 1992.

[14] S. OHTA, K. SATO, and I. TOKIZAWA, "A Dynamically Controllable ATM Transport Network Based on the Virtual Path Concept," *Proc. IEEE Globecom'88*, pp. 1272-1276, 1988.

[15] K. R. Krishnan and R. H. Cardwell, "Routing and Virtual-Path Design in ATM Networks," *Proc. IEEE Globecom'94*, Vol. 2, pp. 765-769, 1994.

[16] F. Y-S. Lin and K-T. Cheng, "Virtual Path Assignment and Virtual Circuit Routing in ATM Networks," *Proc. IEEE Globecom'93*, Vol. 1, pp 436-441, 1993

[17] K. Mezger, D. W. Petr and T. Kelley, "Weighted Fair Queueing vs. Weighted Round Robin: A comparative Analysis", *IEEE Wichita Conference on Communications, Networking and Signal Processing*, April 1994.

[18] Lixia Zhang, "Virtual Clock: A New Traffic Control Algorithm for Packet-Switched Networks," *ACM Transactions on Computer Systems*, Vol. 9, No. 2, pp. 101-124, May 1991.

[19] K. Sriram, "Dynamic Bandwidth Allocation and Congestion Control Schemes for Voice and Data Multiplexing in Wideband Packet Technology," *Proc. of IEEE ICC'90*, pp. 1003-1009, April 1990.

[20] M. Katevenis, S. Sidiropoulos and C. Courcoubetis, "Weighted Round-Robin Cell Multiplexing in a General-Purpose ATM Switch Chip," *IEEE JSAC*, Vol. 9, No. 8, pp. 1265-1279, October 1991.

[21] Y. Wang, T. Lin and K. Gan, "An Improved Scheduling Algorithm for Weighted Round-Robin Cell Multiplexing in an ATM Switch," *Proc. IEEE ICC'94*, pp. 1032-1037, 1994.

[22] H. Kroner, "Comparative Performance Study of Space Priority Mechanisms for ATM Networks," *Proc. IEEE Infocom'90*, pp. 1136-1143, 1990.

[23] D. W. Petr, V. S. Frost, "Nested Threshold Cell Discarding for ATM Overload Control Optimization Under Cell Loss Constraints," *Proc. IEEE Infocom'91*, pp. 1403-1412, 1991.

[24] G. Hebuterne and A. Gravey, “A Space Priority Queueing Mechanism for Multiplexing ATM Channels,” *Computer Networks and ISDN Systems*, pp. 37-43, December 1990.

[25] Q. Hu, D. W. Petr, and C. Braun, “Self-tuning Fuzzy Traffic Rate Control for ATM Networks,” *Proc. IEEE ICC'96*, pp.424-427, 1996.

[26] A. Kolarov, G. Ramamurthy, “ A Control Theoretic Approach to the Design of Closed Loop Rate Based Flow Control for High Speed ATM Networks,” *Proc. IEEE Infocom'97*, 1997.

[27] K. Liu, D. W. Petr, and C. Braun, “A Measurement-Based CAC Strategy for ATM Networks,” *Proc. IEEE ICC'97*, 1997

Biographies

Kunyan Liu (S'95) received his MSEE degree in 1993 from Beijing University of Posts and Telecommunications and his BSEE degree in 1990 from Peking University. Since March 1995 he has been a research assistant in Telecommunications and Information Sciences Laboratory of the University of Kansas, working on funded research project in ATM networking. His current research interests include bandwidth management and CAC in ATM networks, video traffic characterization, network performance analysis and broadband signaling.

Hongbo Zhu (S'94-M'97) received the BSEE and MSEE degrees from the University of Kansas (KU) in December 1993 and May 1995, respectively. From May 1994 to January 1996, he was a Research Assistant in the Telecommunications and Information Sciences Laboratory of KU, working on funded research in ATM based B-ISDN. He is now a Member of Technical Staff at Telogy Networks Inc., Germantown, Maryland, working on Control and Management Plane protocol software development for ATM based B-ISDN. His current research interests are in protocol software development techniques, real-time systems, and network performance analysis.

David W. Petr (BSEE '76 SMU, MSEE '78 Stanford, PhD '90 Kansas) is an Associate Professor in the Electrical and Engineering and Computer Science Department and the Telecommunications and Information Sciences Laboratory at the University of Kansas. His current research interests are focused on the design and analysis of network resource management and congestion control mechanisms, particularly for integrated traffic Asynchronous Transfer Mode (ATM) networks. From 1977 to 1986 he was employed by AT&T Bell Laboratories, working in the areas of digital PBXs, digital transmission systems, packet communication networks and protocols, and speech coding. He was awarded three patents related to this work and was a Bell Laboratories

nominee for the Eta Kappa Nu Outstanding Young Electrical Engineer Award in 1984. While pursuing his doctoral degree, he received the 1987 IEEE Frank A. Cowan Scholarship for graduate studies in communications. Dr. Petr is a Senior Member of the IEEE and a member of Eta Kappa Nu, Tau Beta Pi, Sigma Xi, and the American Society for Engineering Education (ASEE).

Victor S. Frost is currently the Executive Director for Research for the Information and Telecommunications Technology Center at the University of Kansas. Professor Frost has been involved in research for numerous corporations, including Sprint, NCR, BNR, NEC, Telesat Canada, AT&T, McDonnell Douglas, DEC, and COMDISCO Systems. His research has also been sponsored by government agencies, including, NSF, DARPA, Rome Labs, and NASA. From 1987 to 1996 Dr. Frost was Director of the Telecommunications and Information Sciences Laboratory. He has published over 35 journal articles. He has served as a Guest Editor for the IEEE Communications Magazine (March 1994) and the IEEE Journal on Selected Areas in Communications (May 1995) and is currently an Associate Editor for the IEEE Communications Letters and the ACM Transactions on Simulation and Modeling of Computer Systems. His current research interest is in the areas of integrated communication networks, high-speed networks, communications system analysis, and simulation. He is currently involved in research on the MAGIC and AAI high speed wide area testbeds. Dr. Frost received a Presidential Young Investigator Award from the National Science Foundation in 1984, an Air Force Summer Faculty Fellowship, a Ralph R. Teetor Educational Award from the Society of Automotive Engineers, and the Miller Professional Development Awards for Engineering Research and Service in 1986 and 1991 respectively. He is a member of Eta Kappa Nu and Tau Beta Pi and a senior member of the IEEE. He served as Chairman of the Kansas City section of the IEEE Communications Society from June 1991 to Dec. 1992. He

has also served on State of Kansas NSF EPSCoR and DoD DEPSCoR committees as well as the Kansas Inc. Telecommunications Task Force. He is a member of the Board of Trustees for the University of Kansas Center for Research Inc. and the Self Fellowship program. Dr. Frost received the BS., MS.. and Ph.D. degrees from the University of Kansas, Lawrence in 1977, 1978, and 1982, respectively. In 1982 he joined the faculty of the University of Kansas, where he is currently the Dan F. Servey Distinguished Professor of Electrical Engineering and Computer Science.

Cameron Braun received his B.S. and M.S. degrees in electrical engineering from the University of Kansas in 1992 and 1994, respectively. Since 1994 he has been a member of technical staff in the Technology Planning & Integration group at Sprint, Overland Park, KS. His research interests are in the areas of traffic management, call admission control, dynamic resource allocation, and routing for ATM networks. He represents Sprint at the Traffic Management and PNNI SWGs of the ATM Forum. He has been an IEEE member since 1990 and an ACM member since 1995. His email address is cbraun@sprintcorp.com.

William L. Edwards is the Chief Scientist at Sprint Corp. in Overland Park, Kansas. As Chief Scientist Dr. Edwards is responsible for Broadband Integrated Services Digital Network/Asynchronous Transfer Mode (ATM) technology research for Sprint Corp.. He has been an active member of the ATM Forum, the industry group responsible for developing standards for the future global high speed telecommunications infrastructure. Prior to joining Sprint in 1991, he was Chief Engineer for Independent Telecommunications Network, a consultant to the Arizona state government and First Pacific Networks. He was the Director of Broadband Network Research and Director, Quantum Physics and Electronics Research at Pacific Bell and he participated in the systems engineering and planning of the DATAKIT local area network while at AT&T Bell Laboratories. He has been

the Visiting McKay Professor and Visiting Industrial Fellow University of Cal, Berkeley, EE/CS Associate Adjunct Professor University of Colorado- Boulder, Visiting Professor University of Texas- Dallas, Visiting Assistant Professor Texas A&M -College Station.

VC type	PCR (cells/sec)	SCR (cells/sec)	τ_s (ms)
CBR (64kbps voice)	167	N/A	0
VBR (MPEG)	1.06×10^5	9434	49.8
VBR (Non-MPEG)	1.06×10^5	34433	49.8

Tab. 1: Traffic parameters

	CBR	MPEG VBR	Non-MPEG VBR
Max queueing delay (ms)	12.14	50.0	50.0
MQL at ingress edge (cells)	1	471	1716
MQL at ingress VP MUX (cells)	1	1	1
MQL at 1st core switch (cells)	2	2	2
MQL at 2nd core switch (cells)	3	3	3
MQL at egress edge (cells)	2	472	1717

Tab. 2: Maximum delay and queue length

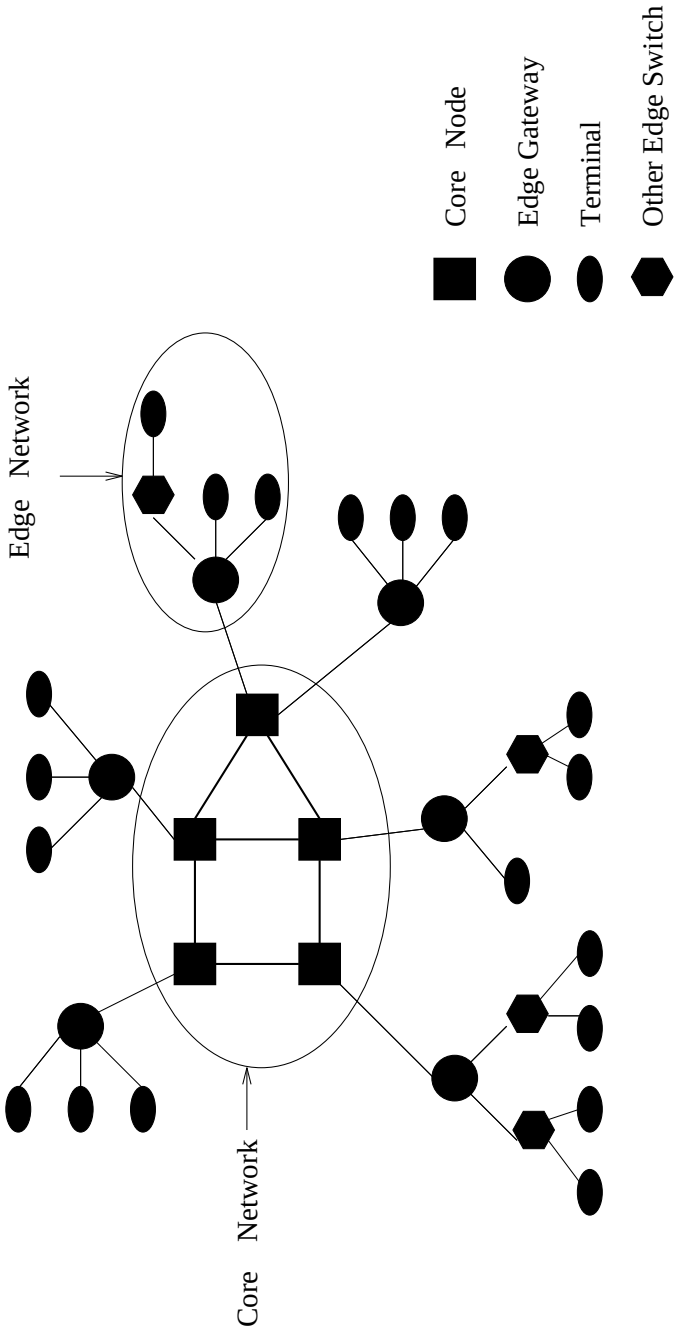

Fig. 1: Network Model: Core and Edge Concept

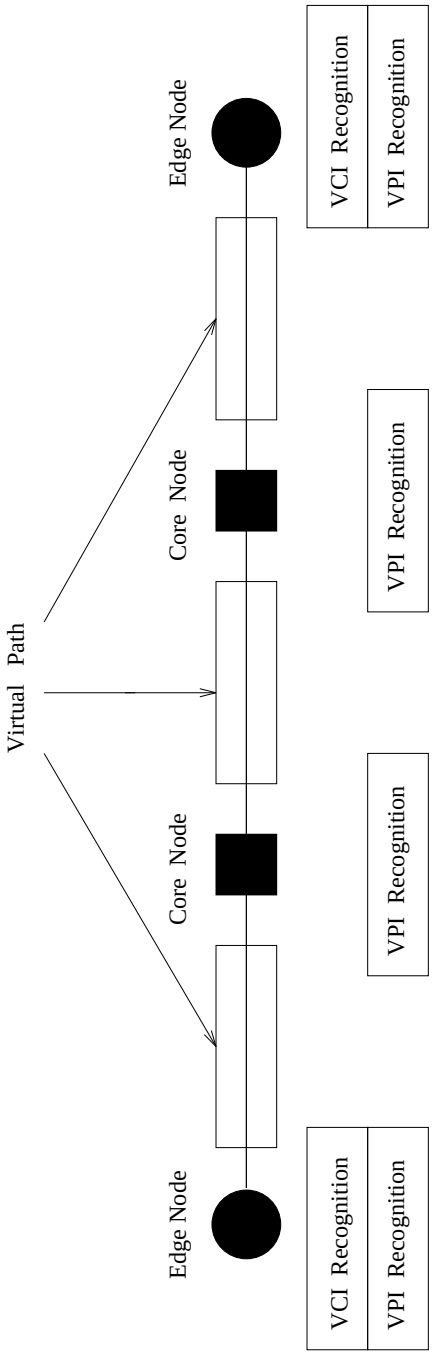

Fig. 2: Cell Switching

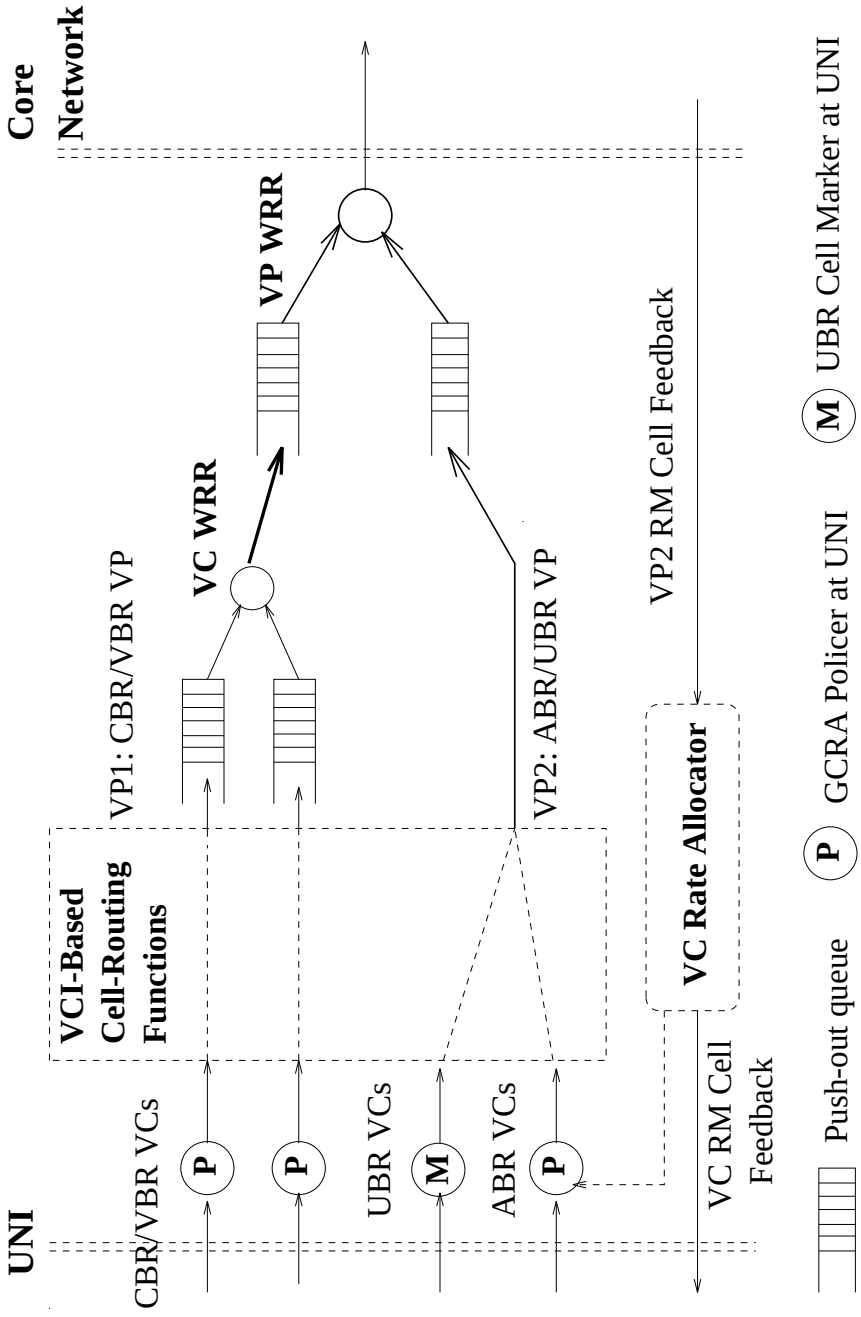

Fig. 3: Ingress function of the edge gateway

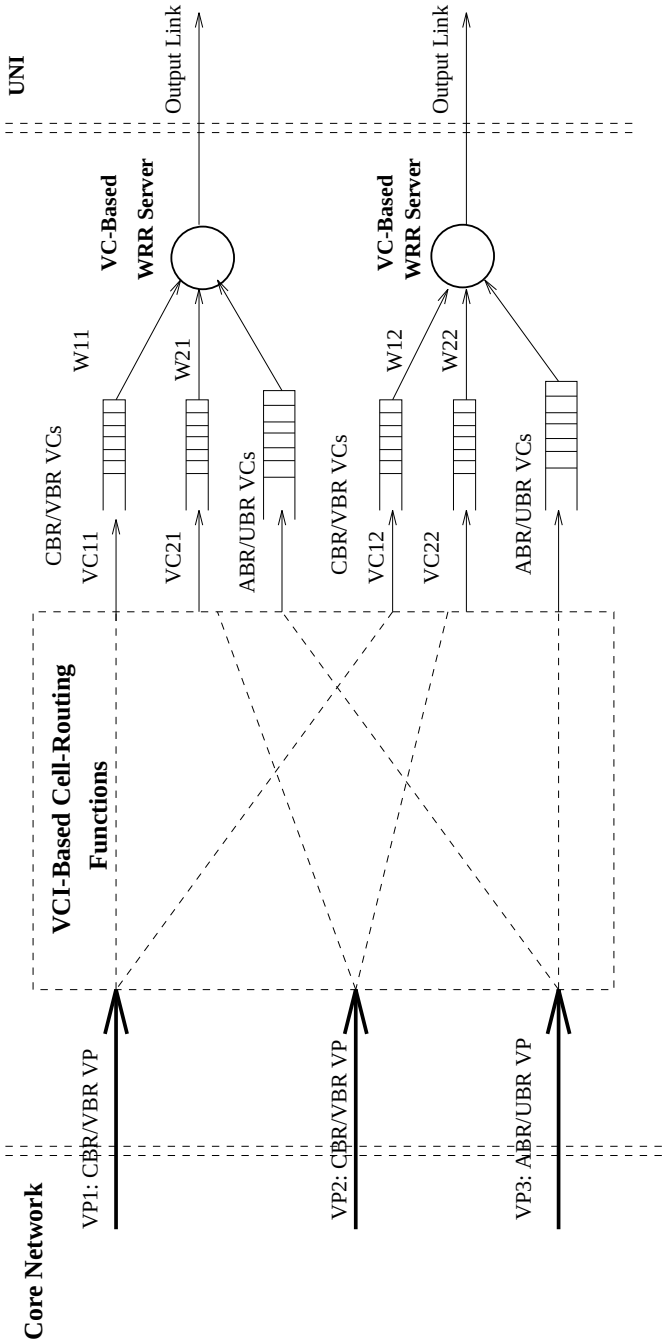

Fig. 4: Egress function of the edge gateway

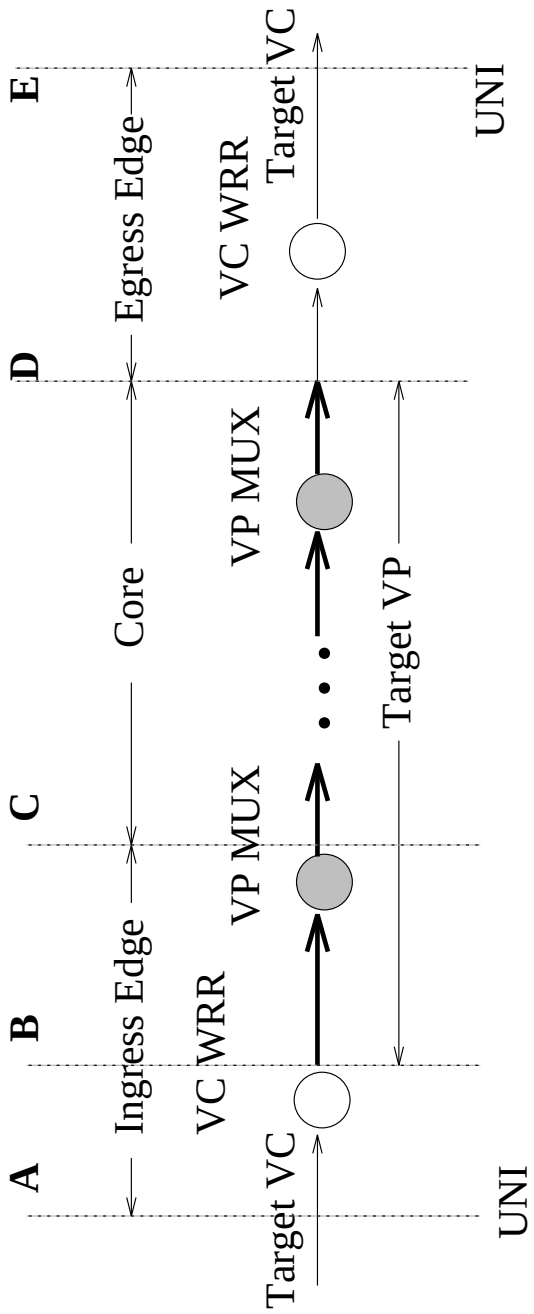

Fig. 5: End-to-end Connection

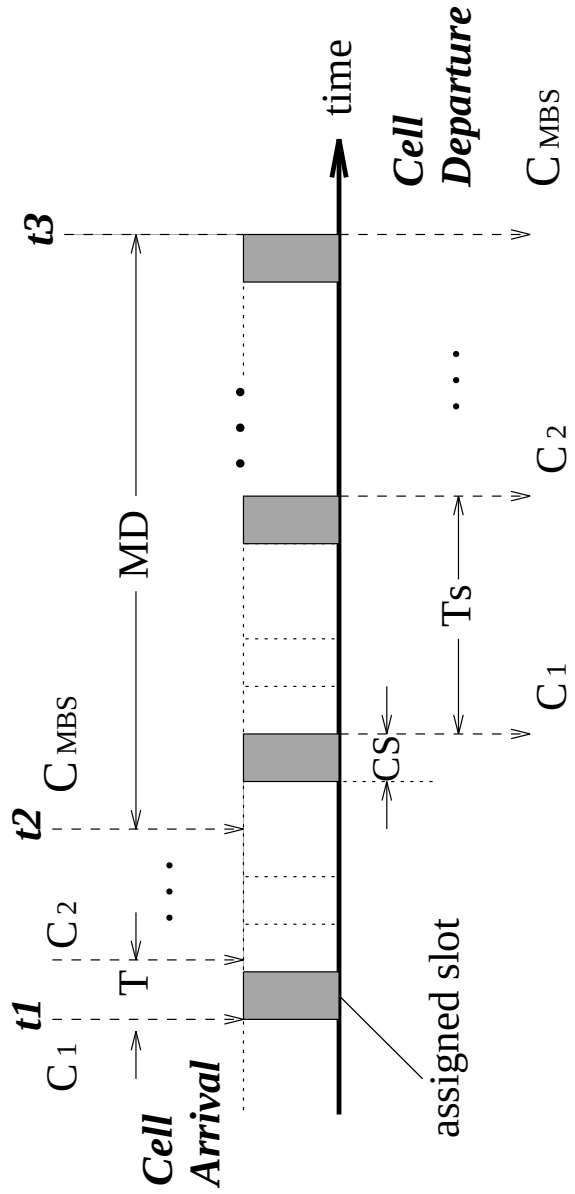

Fig. 6: Maximum queueing delay under distributed WRR

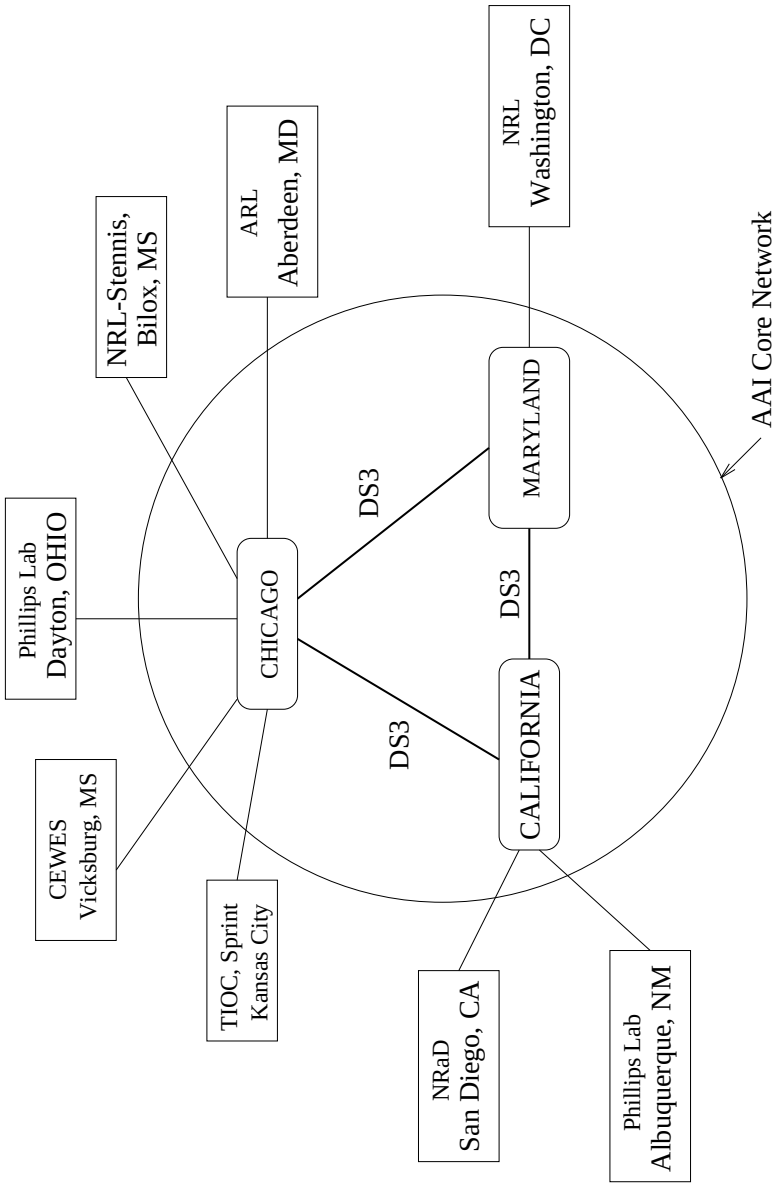

Fig. 7: AAI Network Topology