

BROADBAND SWITCHING SYSTEM

This invention relates to a broadband switching system for the switching of asynchronously transferred data cells, and
5 to a method of switching asynchronously transmitted data cells.

Broadband switching networks for switching asynchronously transferred cells are known, in which a predetermined level
10 of bandwidth is allocated to a transmission channel connecting a first customer to a second customer. In some of these known systems, a communications channel is provided over a significant period of time, effectively of the leased-line type, and manual measures are implemented in
15 order to set up such a connection or to modify a connection according to the particular terminations and the level of traffic being conveyed. Consequently, it is usual for customers to incur a fixed rate charge as part of the overall charge for the connection, resulting in payment
20 being made irrespective as to whether the connection is being used or not.

Alternative systems have been proposed or are available. In particular, it is possible for connections to be established
25 on a dial-up basis, requiring termination equipment to be provided with facilities for establishing connections by issuing signalling commands and responding to similar commands issued by the network.

30 The use of permanent circuits to support a private communications network is widespread. The demand for such circuits is expected to grow to include broadband rates above 2 Mbit/s, the circuits carrying traffic multiplexed from sources which are inherently bursty, possibly together
35 with traffic which is transmitted at constant bit rates and is delay sensitive, such as voice transmission and constant bit-rate video.

REPLACED SHEETS SG

- | | | | |
|-------------------------------------|-------------|------------------|-------------------|
| ☐ | Description | : paged | to paged |
| ☒ | Claim(s) | : paged ...29... | to paged ...33... |
| ☐ | Drawing(s) | : paged | to paged |
| ☐ | Abstract | : paged | to paged |

Asynchronous transfer mode (ATM) cells all have a fixed information field of forty eight octets which can carry customer traffic or customer-originating control information (signalling). These two types of data transmission are distinguished by setting virtual path (VP) and vertical circuit (VC) values in the cell headers. Another field provided in the ATM header is known as cell loss priority, which enables low priority cells to be distinguished from high priority cells. In the event of congestion, the low priority cells may be discarded first.

For private circuits within an ATM based network, the desired route, the required bandwidth, and the quality of service (QOS) are set up using network management procedures. The private circuits are known as permanent virtual circuits (PVCs) because there is no actual physical circuit, only a VP/VC value or "label" which is associated with information stored in the switches to determine the route and preserve the bandwidth and QOS requirements.

20

A disadvantage of all known permanent circuits is that the bandwidth remains assigned to the circuit, even when the customer has nothing to transmit. This means that the customer may have to pay higher charges than would be obtained if the bandwidth was only made available when needed. The assumption being made here is that charging is related to reserved bandwidth, and this is not necessarily correct in terms of the way public network operators may choose to charge for virtual circuits. However, it is expected that charging based on reserved bandwidth will become a significant factor in the future.

A common practice is to set up a permanent virtual circuit so that it is only available during certain hours of the day, or during certain days of the week. A difficulty with this approach is that it does not allow the customer to

change the pattern of usage quickly, and it may only crudely reflect the usage required by the customer.

5 A second proposal has been to provide the customer with a separate communications channel to the network management plane, thereby allowing a permanent virtual circuit to be reconfigured. A difficulty with this approach is that some time delay will be incurred before the customer can start to use the virtual circuit.

10

A third proposal is to introduce equipment at every switching point in the network that recognises a fast resource management cell, indicating that bandwidth should now be assigned to the circuit. A difficulty with this
15 approach is that there is no internationally agreed standard for a bandwidth-requesting cell that would be recognised by the switching equipment produced by the various manufacturers.

20 According to a first aspect of the invention, a broadband switching system having at least one ingress for connection to a respective signal source and at least one egress for connection to a receiving system, the switching system having at least one switch for transmitting information-
25 carrying asynchronously transferred data cells from the ingress to the egress, system control means for accepting and establishing a connection between the ingress and egress via the said switch, and bandwidth control means arranged to feed a indicator signal back to the ingress for transmission
30 to the signal source, to detect incoming cells supplied to the ingress from the signal source, and, automatically in response to such cell detection, to cause the system control means to allocate a predetermined bandwidth for transmission of the cells to the egress.

35

It will be appreciated that in practice, the automatic allocation of the predetermined bandwidth may occur some

time after incoming cells are detected. This is because the required bandwidth may not immediately be available for allocation.

- 5 A connection established in this way is particularly suitable for bandwidth sensitive transmissions such as data communications between networked personal computers. This is because if bandwidth is allocated, it will always be at the predetermined level. The above types of transmission do
10 not operate satisfactorily when operating below the predetermined level of bandwidth (which is a pre-registered level that has been calculated as being sufficient for satisfactory operation) and thus if that level is not available, no bandwidth is allocated, i.e. lower bandwidth
15 amounts are not allocated.

However, in some circumstances the bandwidth control means may be arranged to cause allocation of bandwidth to a signal source at a level less than the predetermined bandwidth
20 associated with that signal source and in conjunction with downgrading the priority of cells received at the ingress such that the bandwidth of non-downgraded cells supplied to the switch does not exceed the allocated bandwidth. This will typically occur when the predetermined level of
25 bandwidth is no longer available on the system.

Once the cells have been downgraded, the system may delete the cells if the system becomes overloaded i.e. has more cells for transmission than can be transmitted in the
30 available bandwidth on the system. Thus the user is taking a conscious risk that some of a message may not be transmitted when transmitting in this optional priority-downgrading mode. Conveniently, the cell deletion technique used is an "intelligent" technique which does not randomly
35 delete cells (thereby corrupting an unknown number of messages) but deletes cells wherever possible, only from a single message.

The bandwidth control means may include feedback means arranged to transmit a cell rate value derived from a bandwidth-indicating signal provided by a stored table associating fixed bandwidth levels with signal sources. The
5 stored table be located in the system control means and/or the bandwidth control means. The cell rate value is fed back to the ingress for transmission to a recognised signal source to indicate to the signal source a permitted rate of supply of cells to the ingress.

10

The feedback means may also be arranged to transmit a "predetermined bandwidth not available" signal to the ingress when the system control means determines that system needs to operate in the priority-downgrading mode for a
15 signal source. Having received such a signal, a signal source may choose to take advantage of the opportunity of transmitting at a lower priority (with the risk of lost cells) but at the same rate as before or to cease transmitting until a subsequent opportunity exists to
20 transmit at the predetermined level (without cells being downgraded in priority).

Preferably, the bandwidth control means is arranged automatically to cause the system control means to
25 deallocate bandwidth for cells from a particular signal source when that signal source is determined to be inactive. This may be when the rate of supply of cells received at the ingress is zero.

30 The system control means may be arranged periodically to determine whether the predetermined bandwidth is available on the system. The bandwidth controller may then "offer" the predetermined bandwidth so determined, to a signal source preferably using the indicator signal. The offer may
35 be made to a signal source which has been halted because it has had its bandwidth deallocated or it may be offered to a signal source which is inactive. Preferably, the system

control means determines that the offer has not been accepted by recognising that the bandwidth has not be caused to been allocated within a predetermined acceptance time period.

5

Conveniently, the bandwidth control means are arranged to detect the rate at which cells are supplied to the input port of the network, primarily to determine whether or not cells are being supplied.

10

The bandwidth control means may include a buffer for delaying transmission of the cells to the switch until bandwidth has been allocated for the cells.

- 15 Preferably, the feedback means is arranged to transmit a reduce-traffic-level instruction to the ingress for reception by a signal source, when receipt of cells from a signal source has been detected and no bandwidth has been allocated for cells from that signal source. The reduce-
20 traffic-level instruction may be a halt instruction for instructing a signal source to cease supplying cells to the ingress. The bandwidth control means may also include timing means for measuring the period of time for which the cells are delayed and may include cell deletion means for
25 deleting cells from the buffer after they have been delayed for a predetermined period of time. A signal source may be arranged to allow for this to occur if it has not received a cell rate value informing it that the predetermined bandwidth has been allocated.

30

- Buffering may be used on other occasions when an signal source is transmitting at a higher rate than the rate capable of being accepted by the network at a given time. Indeed, it is preferable for the buffer to have means for
35 detecting when it is filled to a predetermined threshold level, the feedback means being responsive to the detection to cause a cell rate value (usually requesting the signal

source to halt transmission) to be transmitted to the signal source, the signal source having the ability then to stop its output to avoid buffer overflow and a consequent loss of data.

5

An activity detector forming part of the bandwidth control means may include a cell counter for counting cells received from respective signal sources coupled to the bandwidth control means. The cell count so obtained, may be used to
10 generate charging signals for customer billing and for other purposes.

The invention also includes, according to a second aspect thereof, a method as claimed in claim 13 of the accompanying
15 claims. Third and fourth aspects of the invention are set out in accompanying claims 16 and 17.

The invention will now be described in greater detail, by way of example, with reference to the accompanying drawings,
20 in which:-

Figure 1 is a diagram of a broadband switching system in accordance with the invention;

25 Figure 2 is a diagram of another broadband switching system in accordance with the invention;

Figure 3 is a diagram of part of a broadband switching system showing how a single bandwidth controller can be
30 shared by several end-systems;

Figure 4 is a block diagram of a bandwidth controller for use in the systems of Figures 1 and 2;

35 Figure 5 is a specification description language diagram (SDL) for the activity detector module shown in Figure 4;

Figures 6-1 and 6-2 are SDLs for the controller module of Figure 4;

Figure 7 is a diagram of a resource management (RM) data cell;

Figure 8 is an SDL for the feedback module of Figure 4;

Figure 9 is a diagram of a buffer for the bandwidth controller of Figure 4;

Figure 10 is an SDL for the buffer;

Figure 11 is a block diagram of a shaper/multiplexer module and its connection to the buffer of Figure 6;

Figures 12-1, 12-2, and 12-3 are SDLs for the shaper/multiplexer module;

In its preferred form, the invention is concerned with a broadband switching network which may form part of or may constitute a public switching network for the transmission of asynchronously transferred data cells between end-systems.

25

Referring to Figure 1, the public network 10 has a plurality of switches operable in asynchronous transfer mode (ATM). In this simple example, the switches include two local switches 12 each having a port for connection to a respective end-system 14, and a transit switch 16 interconnecting the local switches 12. Associated with the switches is a connection admission control function (CAC) 18 and a dynamic bandwidth controller (DBC) 20 for controlling traffic entering the network through one of the local switches 12. This switch 12 also includes a usage parameter control device 22 for dynamically altering the priority of

data cells received at the input port 24 of the network from the end-system 14.

It will be understood that, in practice, the network 10 will include large numbers of local and transit switches 12, 16 and several DBCs 20 all interconnected to form a network having a plurality of ports such as port 24 for connecting several end-systems such as end-system 14. Using the DBC 20, the public network 10 is able to provide an available bit rate (ABR) service, the DBC acting to detect incoming cells supplied to the input port 24 and, automatically in response to this detection, to cause the CAC 18 to allocate bandwidth for the transmission of the cells to the destination end-system. Generally, end-systems 14 requiring the ABR service are allocated to a fixed DBC 20. There may be more than one DBC 20 for each local ATM switch 12. In the case of a fault, end-systems can be rerouted to a standby DBC (not shown).

Data is transmitted in the form of asynchronous transfer mode (ATM) cells, each having an information field of forty-eight octets, in addition to a header of five octets, which includes information facilitating transmission through the network itself. Thus, routing is controlled on a cell-by-cell basis and a plurality of transmission paths and time multiplexed slots may be employed for any particular link. ATM cells are, therefore, transmitted via virtual paths and virtual circuits, as defined by the header information.

The virtual paths and virtual circuits are identified by a virtual path identifier (VPI) and a virtual channel identifier (VCI) in the five octet header which effectively defines the connection between the end-systems so that cells forming part of a common message will be transmitted over the same connection. ABR traffic enters the public network 10 by routing cells according to their VPIs and VCIs through the DBC 20 and then out to external routes, as shown in

Figure 1. From the DBC 20, the traffic on each virtual path and virtual channel is restricted to a cell rate (which will be referred to hereinafter as "CR") determined by the CAC 18.

5

An alternative illustrative arrangement is shown in Figure 2. In this case, end-system 14A is subject to the control of more than one DBC. In fact, the connection between two end-systems 14A, 14B is routed through two public networks
10 10-1 and 10-2. Each network 10-1, 10-2 has its own DBC 20-1, 20-2 responsible for restricting traffic entering the network according to the bandwidth allocated by its own connection admission control function (CAC) 18-1, 18-2. Each DBC 20-1, 20-2 is also responsible for advising the
15 end-system 14A of the current applicable CR.

In the systems of both Figure 1 and Figure 2, the DBCs 20, 20-1, 20-2 request bandwidth from the respective CAC 18, 18-1, 18-2 whilst buffering any incoming data cells which
20 cannot immediately be transmitted to the respective switches 12, 16. The CAC 18, 18-1, 18-2 then allocates a bandwidth. This allocation is then indicated to the DBC 20, 20-1, 20-2 which communicates a maximum CR to the transmitting end-system 14. The allocation only occurs if sufficient
25 bandwidth is available on the system to allocate a predetermined bandwidth (pre-registered by the customer) to the end-system.

It is possible for a single dynamic bandwidth controller
30 (DBC) to be shared by several end-systems or signal sources. For example, referring to Figure 3, a DBC 20-3 is shown connected to a broadband ATM switch 12-3 forming part of the network 10, the traffic of three sources 14C being handled using an output buffer 28. The number of sources which can
35 be handled by the DBC 20-3 is determined by the link rate L (i.e. there must be not so many sources that it is always the link rate L which is the limiting factor determining the

available rate.) The aggregate cell rate of the ABR traffic from the sources 14C must not exceed L. This implies that if the traffic from each source is bursty, there may be times when the output buffer 28 is congested. This can be
5 avoided by supplementing the sustained cell rate (CR) feedback to the end-systems 14C with generic flow control (GFC) signals which operate to stop all transmissions from each source immediately.

- 10 Whenever the dynamic bandwidth controller (DBC) is incorporated in the arrangements of Figures 1, 2, or 3, its main functions are as follows.

Firstly, it provides buffering of incoming data cells, the
15 degree of buffering at any given time being determined according to the transmission containing the cells, the transmission being identified by the VPI and VCI information referred to above. The DBC further controls or "shapes" the traffic fed to the network 10 so as to be equal to the
20 current CR applicable to that particular transmission, the CR depending on the allocated bandwidth.

The allocated bandwidth, and hence the CR, for any given transmission is determined by the CAC 18 (see Figure 1) on
25 the basis of determining the route to be followed by the transmission and assessing a fair share of the available capacity on the route based upon the known number of active transmissions and on the predetermined bandwidth which needs to be allocated.

30

When a transmission begins it is detected in the DBC, which immediately transmits a halt signal to the relevant end-system 14 (see Figure 1). The halting of the end-system ensures that a newly active transmitting source does not
35 cause overload in the system 10 before the CAC 18 has been able to allocate bandwidth and derive an CR for that transmission. Such overload would typically cause cell loss

for that transmission. This is part of a second main function of the DBC, i.e. to send a feedback signal to the end-system for the purpose of controlling its transmitted cell rate. Indeed, each time the CAC 18 derives a new CR
5 for a transmission, a CR advice signal is fed back to the end-system. In this case, the CR will be either zero (i.e. halt) or a CR corresponding to the predetermined bandwidth for the end-system 14.

10 The pre-transmission buffering of the DBC is used to allow a cooperating end-system sufficient time to adjust its output to the latest CR feedback advice. This implies that there is sufficient buffering in the DBC to allow excess cells to enter for a period at least equal to the round trip
15 delay between the DBC and the end-system. If cells continue to arrive from the end-system 14 at a rate greater than the advised feedback CR (for instance, because the CR was lost en route, or because of a faulty end-system) the excess cells will be dropped in the DBC by overflow of the buffer.

20

In the preferred DBC, it is also possible to include fault tolerance by making use of a buffer threshold. When the stored cells relating to a given transmission reach the threshold, retransmission of the CR advice feedback to the
25 end-system is triggered. This feature is useful also as a mechanism for policing end-systems to prevent inefficient use of bandwidth, whether due to a faulty terminal or due to deliberate non-compliance with contracted transmission rules. In this way, interference with the quality of
30 service provided for other, compliant end-systems is prevented. In effect, the DBC defines the ABR traffic contract with the network 10.

The modules of the DBC 20 will now be described in more
35 detail with reference to Figure 4.

The DBC 20 is shown in Figure 4 is a discrete unit having an input port 30 for receiving asynchronously transmitted data cells, an output for feeding data cells onto a switch 12 or 16 (see Figures 1 and 2) forming part of the switching network 10. The unit also has another input 34 for receiving messages back from the switch 12 or 16 and a feedback output 35 for transmitting feedback messages to the end-system 14 (shown in Figure 1). Although the DBC 20 is shown as a discrete unit, it will be appreciated that Figure 4 can be regarded as a functional diagram representing a subset system of a larger data processing unit, much of which may be embodied as software functions.

Incoming cells on input 30 arrive as a user cell stream which is fed firstly to an activity detector 36. The purpose of the activity detector is to provide state information to a controller module 38 about each received transmission, each transmission being identified by its VPI and VCI contained in the cell headers. A transmission is labelled active by the activity detector 36 if it was previously quiet and a cell having the appropriate VPI and VCI values is observed to be transmitted from an end-system to input 30. Synchronisation of the activity detector 36 with the start of a cell header may be carried out using an error check field contained in the cell header. A transmission is considered to be in an inactive state if it was previously active and no cell having the appropriate VPI and VCI values has been detected for a period of time t . The error check field provides a degree of redundancy, by which error checking may be performed on the header information. Thus, the principal reason for providing the header error check field is to ensure that the header information is correct, thereby ensuring that cells are not transmitted to erroneous addresses.

35

Activity detector 36 maintains a timer and state table for each VPI/VCI value pair. Preferably, t is set to be several

seconds so that active-inactive-active transitions relating to any given VPI/VCI value pair which are of the order of several milliseconds remain undetected so that the transmission is indicated as remaining in the active state under these conditions. This has the effect of reducing the frequency of messages sent by the DBC 20 to the CAC 18 at some expense to lowering utilisation of the network.

Another function of the activity detector 36 is that of counting the cells for transmission during an interval after receiving a "start cell count" signal from the controller 38. This information can be used, for example, for charging purposes and also by the controller 38 for assessing the actual cell rate of received transmissions.

15

Pseudocode for the activity detector is listed below and the corresponding SDL appears in Figure 5.

```

20      BEGIN {cell arrival}
        cell arrival from end-system
        read VC
        reset VC inactivity timer
        IF VC is newly active THEN
          update state table
25      advise CONTROLLER of newly active VC
        ELSE IF counting.cells(VPI/VCI) THEN
          increment cell.count(VPI/VCI)
        ELSE
          do nothing
30      END

      BEGIN {VC timer expires}
        VC inactivity timer expires indicating quiet VC
        update state table
35      advise CONTROLLER of quiet VC
        counting.cells := FALSE
      END

      BEGIN {receive a start.cell.count signal}
40      receive a start.cell.count(VPI/VCI) signal from CONTROLLER
        cell.count(VPI/VCI) := 0
        counting.cells := TRUE
      END

45      BEGIN {cell.count timer expires}
        cell.count timer expires
        send cell.count(VPI/VCI) to CONTROLLER

```

restart cell.count timer
END

It will be seen that, as far as the incoming user cell
5 stream is concerned, the activity detector 36 reads the
VPI/VCI values in each cell header of the arriving cell
stream, and this information is used to update the state
table which it maintains for each VPI/VCI value pair. As
far as communication with the controller 38 is concerned,
10 the detector 36 informs the controller of a change of state
of any VPI/VCI value pair. The controller can inform the
activity detector of the timer value *t* to be used.
Preferably, the same value of *t* is used for all VPI/VCI
value pairs. Cell count information is sent to the
15 controller 38 by the activity detector 36 at the end of each
timer expiry.

The cells of the user cell stream arriving on input 30 are
transmitted without delay to a buffer module 40 where they
20 are stored in first-in, first-out (FIFO) buffer queues, each
queue comprising cells having a given VPI/VCI value pair.
The detector 36 is non-specific to cell type. Thus, the
arrival of any data cells will be detected and can
potentially affect the activity state associated with a
25 VPI/VCI value pair, independently of the existence or
absence of control or management cells. Buffered cells are
fed from the buffer 40 to a shaper multiplexer module 42
prior to being fed to an ATM switch via output 32. Operation
of the buffer and shaper/multiplexer modules 40, 42 will be
30 described in more detail below. For the time being, it is
sufficient simply to say that the buffer module is capable
of signalling to the controller 38 when any buffer queue has
reached a predetermined buffer fill threshold. The shaper
multiplexer module 42 is responsible for removing cells from
35 the buffer module 40 and transmitting them onwards towards
their destination. It includes a multiplexer function and
the shaper stores an CR value for each VPI/VCI value pair so
that the cell stream fed from the output is shaped to ensure

that the capacity of the respective path through the network for each transmission, as determined by the allocated bandwidth, is not exceeded. The controller 38 also controls a feedback module 44 for receiving feedback messages from the network on input 34 and from the controller 38 itself, for onward transmission to the end-system 14 via output 35. The functions of the buffer, shaper/multiplexer, and feedback modules 40, 42 and 44 will be described in more detail below. The controller 38 will be considered first.

10

The purpose of the controller 38 is to signal to the CAC that an ABR type transmission identified by any given VPI/VCI value pair should have bandwidth in the system allocated or re-negotiated. In this embodiment, the controller 38 sends a halt signal to the end-system via the feedback module 44 as soon as the activity detector 36 detects that the end-system has become active.

The CAC 18 is then sent a request for bandwidth. This is interpreted by the CAC 18 as a request for the predetermined bandwidth associated with the end-system. If this bandwidth cannot be granted, the end-system is kept in a halt state, the cells already received by the DBC are buffered (by setting the shaper module 42 rate to zero) and a timer is started for monitoring for how long the cells are buffered.

The CAC 18 periodically (preferably at periods just less than the maximum time for which cells are buffered by the shaper) attempts to find the requested bandwidth and offers it to the DBC, which in turn offers it to the end-system in the form of a signal via the feedback module 44. When the end-system 14 takes up the offer of bandwidth, the bandwidth is allocated, and the shaper is notified of the CR corresponding to the allocated bandwidth. If the CAC 18 cannot find the bandwidth, it may remove bandwidth allocation from other end-systems to obtain sufficient

bandwidth to be allocated. Suitable bandwidth balancing techniques for use in the CAC, are described below.

If the timer expires before bandwidth can be allocated, the
5 cells held in the DBC (in the buffer) are deleted. In this
case, the end-system knows that the cells have been deleted
since it also has a timer and unless a CR is fed-back within
a predetermined time, it assumes that bandwidth cannot be
allocated and that the few cells which were sent before the
10 halt signal was received, have been deleted.

If the CAC 18 needs to remove the bandwidth, the whole
bandwidth is removed and the end-system is halted as
described above.

15

As a strategy which is an alternative to the detection of
cells before bandwidth is allocated, the CAC 18 may
continuously poll end-systems to offer bandwidth (to the
predetermined level required by an end-system) to the end-
20 system via the DBC. If the end-system starts transmitting,
the bandwidth is allocated.

This polling, offer and acceptance procedure may also be
used to cause transmission to re-start after an end-system
25 has been halted as described above.

As mentioned above, the controller 38 is also arranged to
receive a signal from the buffer module 40 when the buffer
fill for a given VPI/VCI value pair has reached a given
30 threshold. This signal causes the controller 38 to command
the feedback module 44 to issue a so-called resource
management (RM) cell, which will be described in more detail
hereinafter. The controller 38 may also receive a DBC
identity value for each new transmission (identified by a
35 new VPI/VCI value pair) which is established, this DBC
identity value being received from the CAC. Alternatively,
the DBC may use a default identity if none is supplied.

Pseudocode for the controller appears below:-

```

5      BEGIN {receive an active VPI/VCI from ACTIVITY DETECTOR}
        receive active (VPI/VCI) from ACTIVITY DETECTOR
        send halt (VPI/VCI) to feedback
        send bandwidth request (VPI/VCI) to CAC
        send active (VPI/VCI) and CR=0 to shaper
        start timer for buffered cells
10     END
        BEGIN {timer expires}
        timer for buffered cells expires
        send delete buffered cells (VPI/VCI) to shaper
15     END
        BEGIN {receive a CR from CAC}
        receive a CR from CAC
        send CR to shaper
        send CR to feedback
20     END
        BEGIN {receive an inactive signal}
        receive an inactive signal
        advise CAC of inactive VPI/VCI
25     END
        BEGIN {receive a VPI/VCI buffer threshold signal}
        receive a VPI/VCI buffer threshold signal
        signal feedback to retransmit CR to end-system
30     END

```

SDLs for the controller are shown in Figures 6-1 and 6-2.

The controller 38 is arranged to write a DBC, VPI/VCI
 35 identity value pair into the feedback module 44. It is also
 arranged to instruct the feedback module 44 to issue a
 resource management command for a specific VPI/VCI value
 pair. This instruction may also contain an appropriate CR
 pair T, τ , (T is an average cell inter-arrival time and τ is
 40 a burst tolerance). It should be noted that only one change
 in the values specified in an RM cell is sent for each new
 VPI/VCI value pair whenever the CAC updates the CR value.
 Typically, this may be once every 30 seconds or more in a
 public network, and depends upon the sensitivity setting of
 45 the activity detector in the DBC 20. It follows that the
 required feedback control bandwidth can be relatively small.

As will be seen from the pseudocode, the controller 38 receives signals from the buffer module 40 whenever a buffer fill threshold is reached by cells having a specific VPI/VCI value pair.

5

The interface with the activity detector 36 has already been described.

The purpose of the feedback module 44 will now be described
10 briefly.

As mentioned above, the feedback module 44 transmits current CR values (as signalled by the controller 38) to the end-system via output 35. The CR is transmitted using a
15 resource management cell as shown in Figure 7. Optionally, one field of this cell is the DBC identity value which is used to enable an end-system 14 (see Figure 1) to distinguish between CR advices from different DBCs (e.g. DBCs 20-1 and 20-2 as shown in Figure 2) in the end-system
20 to end-system path. This DBC identity field is indicated as field 50 in Figure 7. The CR is placed in field 52. This RM cell, like other cells, has a five octet header which contains a PT field 54 indicating that the cell is a resource management (RM) cell.

25

It is proposed that, if used, DBC identity values are not fixed but are chosen at the time of setting up the transmission path through the network for a given a VPI/VCI value pair. This implies that the CAC 18 assigns a value
30 for the DBC identity for each VPI/VCI value pair, and the feedback module 44 maintains a table of (DBC, VPI/VCI) identity pairs. For example, in Figure 2, public network 10-1 is arranged to choose a DBC identity for a given VPI/VCI pair and signals this information forwards so that
35 public network 10-2 does not chose the same value (e.g. public network 10-1 assigns identity 1, public network 10-2

assigns identity 2, etc). The DBC identity value is stored in a table maintained by the feedback module 44.

The CR field 52 in the RM cell (see Figure 7) contains the CR advice from the CAC which is provided as the average cell inter-arrival time T , plus a burst tolerance τ .

Operation of the feedback module 44 is triggered by the controller 38 (a) when a new CR is advised by the CAC 18, and (b) when the buffer fill level in buffer module 40 corresponding to any VPI/VCI value pair rises above the buffer fill threshold. A resource management (RM) cell is then sent to the end-system.

The pseudocode for the feedback module 44 is as follows and the corresponding SDL is shown in Figure 8.

```

BEGIN {Receive a CR}
  receive a CR for a VPI/VCI from Controller
  default_CR := CR
END {Receive an CR}

BEGIN {RM.cell timer expires}
  RM.cell timer expires
  create RM.cell
  write default_CR into RM.cell
  send RM.cell to end-system
  restart RM.cell timer
END {cell arrival from network}

```

Next the buffer module 40 will be considered.

The buffer module is shown in more detail in Figure 9. Its purpose is to store incoming data cells on the basis of the VPI/VCI value pairs contained in the cells. Buffering the cells allows an end-system 14 (Figure 1) time to respond to a feedback signal from module 44. Another function of the buffer module 40 is to send a signal to the controller 38 when the buffer fill threshold is reached, indicating that an end-system is not responding to a feedback signal (this in turn causes the controller 38 to re-send a CR to the end-system, as mentioned above). The buffer module 40 also

drops received cells when the maximum buffer allocation for a given VPI/VCI value pair is exceeded. This is done by buffer overflow.

- 5 The size of buffer required for a DBC 20 controlling access to the switching system 10 could be relatively small. For example, if the DBC 20 has a combined input rate from all sources of 150 Mbit/s, then, if the round trip delay to the end-system is 100 μ s, there will be less than 35 cells in
10 flight whenever the CR values are changed.

The size of the shared memory area 56 is mainly to cater for changes in the burst tolerance, because a change in this rate leads to only a small number of excess cell arrivals
15 (e.g. around 35 cells). The fixed cell positions assigned to respective VPI/VCI value pairs are designated by the reference numeral 58 in Figure 9. The cells in these positions represent the front cells of a plurality of queues, each queue having its own VPI/VCI value pair. In
20 other words, the queues can be visualised as running horizontally in Figure 9 with the front cells at the right hand side. Cells arriving in the buffer module 40 are placed in the queues in a first-in, first-out (FIFO) order.

- 25 Cells are removed from the buffer module 40 when an appropriate signal is received from the shaper section of the shaper/multiplexer module 42, as defined by the buffer module pseudocode which follows:-

```

30      BEGIN {Receive a cell}
          receive a cell
          IF there is room in the buffer THEN
              put cell in buffer
              increment buffer-fill level
35      IF buffer-fill level = Threshold THEN
          transmit buffer-full signal to CONTROLLER
          ELSE
              do nothing
          END {Receive a cell}
40      BEGIN {Receive a fetch}
  
```

```

        receive a VPI/VCI fetch signal from the SHAPER/MUX
        pass cell from buffer to the SHAPER/MUX
        decrement buffer-fill level
    END {Receive a fetch}

```

5

The corresponding SDL appears in Figure 10.

Referring now to Figure 4 in combination with Figure 11, the shaper/multiplexer module 42 operates to remove cells from the buffer module 40 and to transmit them onwards towards their destination via the network switches. Module 42 has two parts which are a multiplexer 60 and a shaper 62. For each VPI/VCI value pair, the shaper 62 maintains a sustained cell rate (CR) value and a timer.

15

The cell stream fed to output 32 is shaped by the shaper so that bursts which are not greater than the burst tolerance τ pass without being delayed by the shaper 62. However, the multiplex function may delay a cell if several transmissions represented by different VPI/VCI value pairs are bursting simultaneously. In this case, the multiplexer 60 assigns each active VPI/VCI value pair a fair share of the DBC output bandwidth. It does this by polling active VPI/VCI value pairs in a round-robin fashion. Cells which are waiting for a period equal to or greater than the rate interval T are flagged with a higher priority "cell must go" value. The multiplexer picks up these cells first (see Figure 11). Cells will be forced to wait by the shaper function if bursts arrive which are longer than the burst tolerance credit value. The detailed operation of the shaper/multiplexer module 42 will become apparent from the following pseudocode:-

```

1.  BEGIN {STATE = ACTIVE}
35  receive a cell.waiting[VPI/VCI] signal from buffer
    IF-burst credit ok THEN
        cell.can.go:= TRUE
        STATE:= WAIT for multiplexer
    ELSE {burst credit not ok}
40  STATE := WAIT for credit timer to expire
    END

```

```

2.   BEGIN {STATE = WAIT for credit timer to expire}
      credit timer expires
      increment burst tolerance credit counter
      cell.can.go := TRUE
5    cell.must.go := TRUE
      STATE := WAIT for multiplexer
      END

10   3.   BEGIN {STATE = WAIT for multiplexer}
      receive a fetch.cell[VPI/VCI] from multiplexer
      decrement credit counter
      cell.can.go := FALSE
      cell.must.go := FALSE
      STATE := ACTIVE
15   END

20   4.   BEGIN {STATE = WAIT for multiplexer}
      credit timer expires
      IF credit counter <  $\tau$  THEN
        increment credit counter
      ELSE
        do nothing
      cell.must.go := TRUE
25   END

5.   BEGIN {STATE = SHAPER ACTIVE}
      credit timer expires
      IF credit counter <  $\tau$  THEN
        increment credit counter
30   ELSE
        do nothing
      cell.must.go := TRUE
      END

35   6.   BEGIN {STATE = SHAPER.CR ACTIVE}
      new CR advised (T,  $\tau$ )
      nextT := T
      nextcredit :=  $\tau$ 
      END

40   7.   BEGIN {STATE = SHAPER TIMER ACTIVE}
      timer expires
      reset timer (nextT)
      END

45   8.   BEGIN {STATE = MULTIPLEXER ACTIVE}
      output cell timer expires
      index := pointer
      REPEAT {1st loop of searching for cell.must.go}
50   increment index
      IF cell.must.go[index] THEN
        pointer := index
        fetch cell[index] from buffer
        send fetch cell signal to SHAPER
55   STATE := MUX.ACTIVE
      ELSE

```

```

        IF index = max.buffer.size THEN
            index := 0
        UNTIL index = pointer

5      REPEAT {2nd loop of searching for cell.can.go}
        increment index
        IF cell.can.go[index] THEN
            pointer := index
            fetch cell[index] from buffer
10      send fetch cell signal to SHAPER
            STATE := MUX.ACTIVE
        ELSE
            IF index = max.buffer.size THEN
                index := 0
15      UNTIL index = pointer
            send no.cell.waiting.signal
        END

```

When the CAC 18 receives a bandwidth request from the
20 controller 38, it must first determined whether sufficient
bandwidth is available to meet the predetermined minimum
bandwidth. If sufficient bandwidth is not available by the
time the timer for monitoring how long calls have been
buffered, is about to expire, bandwidth is "robbed" from
25 other users as described below. Confirmation of allocated
bandwidth is then sent to the controller 38 which forwards
the bandwidth allocation to the end-system via the feedback
module 44.

30 It will be understood that when the DBC 20 requests a change
in the bandwidth allocated to a particular transmission, the
CAC must control other traffic in the network so that the
network capacity is used most effectively. The description
which follows deals with connection admission control
35 methods for overcoming the problem of traffic rebalancing.

Two connection admission control strategies will now be
described. Both tackle the problem of rebalancing traffic.
In other words, when a transmission becomes quiet or newly
40 active, it is necessary to determine how many other control
messages need to be generated for other transmissions. The
object is to make this number of control messages as small
as possible. The strategies described below apply generally

to other traffic on the system which does not require a predetermined bandwidth and which can accept varying levels of bandwidth.

5 The first strategy involves a relatively simple connection admission control method which involves no actual rebalancing. In this method, a newly active transmission (VPI/VCI value pair) is given a single sustained cell rate (CR) which is retained until the transmission goes quiet
10 again. Only when it is subsequently reactivated will the transmission get a different CR. This means that a quiet signal relating to one VPI/VCI value pair will cause no control signals to be generated for other VPI/VCI value pairs which were sharing capacity with it.

15

This is combined with a filling method which involves (i) giving a first newly active connection an effective capacity which is half of the total available capacity; (ii) giving the next newly active connection an effective capacity which
20 is half of the remaining capacity; (iii) giving the next newly active connection an effective capacity which is half of the still remaining capacity; and so on. This method is applied link-by-link over the entire route identified by the VPI/VCI value pair, and whichever yields the lowest
25 effective capacity is the determinant of the CR fed back to the DBC 20.

It follows that a newly active signal having one VPI/VCI value pair generates no control signals for the other
30 VPI/VCI value pairs which are sharing the capacity.

Since the DBC 20 is designed such that a user can only maintain a large effective capacity on the network so long as the VPI/VCI value pair remains in the active state in the
35 activity detector 36 (Figure 4), and the cell rate generated by the customer is close to the effective bandwidth value (refer to the cell-counting function of the activity

detector described above), it follows that users can only hold onto large effective bandwidths for as long as they are prepared to be charged for the proportionally larger loads which they are submitting.

5

This method is fair to users in the sense that, over a sufficiently long period, no user is systematically given a poorer capacity.

10 However, it is desirable in some circumstances to increase the number of users who are able to secure relatively large bandwidth allocations and this can be catered for by a second, modified method as follows.

15 In this case the underlying principle is that, if an active signal causes control signals for other VPI/VCI value pairs, let the signal be limited to only one per link, namely the richest (largest capacity) VPI/VCI value pair. This can be described as a limited rebalancing method or a "take-only-
20 from-the-richest" (Robin Hood) method.

This can best be illustrated with an example filling method:-

- 25 (i) the first newly active VPI/VCI value pair is assigned an effective capacity equal to half of the total available capacity;
- (ii) the next newly active connection is assigned half of
30 the remaining capacity plus a fifth of the effective capacity of the first VPI/VCI value pair (i.e. the current richest);
- (iii) the next newly active connection is assigned half of
35 the remaining capacity plus one fifth from the current richest; and so on.

To illustrate this process, it may be imagined that there is a single link with a capacity of 100 Mbit/s. The above steps then result in the following exemplary steps:-

- 5 (i) the first newly active VPI/VCI value pair gets 50 Mbit/s and there is 50 Mbit/s remaining;
- (ii) the next VPI/VCI value pair gets half of the remainder (which yields 25 Mbit/s) plus a fifth from
10 the first, which means that the first now has 40 Mbit/s, and the second has 35 Mbit/s;
- (iii) the next VPI/VCI value pair gets half of the remainder, which yields 12.5 Mbit/s plus a fifth from
15 the first, so that the first now has 32 Mbit/s, the second still has 35 Mbit/s, the third has 20.5 Mbit/s, and so on.

Note that more of the users are now getting large
20 capacities, but there is only one extra control message to send on the link. There is thus a limited rebalancing or "Robin Hood" strategy.

To extend the method to a route with many links, the above
25 process is repeated link-by-link. Whichever link yields the lowest effective capacity is the determinant of the CR value sent back to the DBC. Now, using this value of effective capacity, the CAC assigns it link-by-link by taking half of the remaining capacity on that link, and any extra which is
30 needed is taken from the richest VPI/VCI value pair on that link. Consequently, this generates at most one additional CR control message per link for each VPI/VCI active signal sent to the network. A quiet signal still generates no additional control messages.

35

This strategy also makes it impossible for a user to hold onto a very large capacity when others become active. In

addition, as many users as possible are given a reasonably large capacity while keeping the complexity of traffic rebalancing to a minimum.

- 5 In summary, there is provided a broadband switching system for the switching of asynchronously transferred cells of data, a dynamic bandwidth controller (DBC) controls the application of data cells to an input port of the system, the data cells being supplied by a number of transmitting
10 end-systems.

When an end-system begins transmitting data cells, the DBC detects the presence of incoming cells and requests bandwidth from a CAC forming part of the system.

15

- The switching system stores a table associating a number of signal sources connected to the ingress with respective minimum transmission bandwidths and, preferably also, maximum delay times. When arrival of cells from one of the
20 sources at the input port is detected, the DBC sends a request signal for the relevant bandwidth to the CAC and delays transmission of the cells until at least the minimum bandwidth is allocated. This delay is typically effected by send a cell rate indicator signal back to the input port for
25 placing the source in a half mode. If no allocation of bandwidth has occurred before the respective maximum delay time, bandwidth is allocated by robbing bandwidth from other signal sources.

CLAIMS

1. A broadband switching system having at least one ingress for connection to a respective signal source and at least one egress for connection to a receiving system, the switching system having at least one switch for transmitting information-carrying asynchronously transferred data cells from the ingress to the egress, system control means for accepting and establishing a connection between the ingress and egress via the said switch, and bandwidth control means arranged to feed an indicator signal back to the ingress for transmission to the signal source, to detect incoming cells supplied to the ingress from the signal source, and, automatically in response to such cell detection, to cause the system control means to allocate a predetermined bandwidth for transmission of the cells to the egress the system including table storing means arranged for storing a table associating the predetermined bandwidth with the signal source, the bandwidth control means having means for recognising the source of the detected cells to identify the predetermined bandwidth associated with the detected cells, and the bandwidth control means being operable in response to the said cell detection to send a feedback signal to the ingress for receipt by the source to halt the supply of cells by the source until the said predetermined bandwidth is allocated.
2. A system according to claim 1, characterised in that the indicator signal is a cell rate indicator signal.
3. A system according to claim 2, characterised in that the bandwidth control means is arranged to delay cells initially received from the source at the start



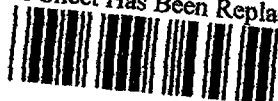
of the transmission and then subsequently to cause the said allocation of the predetermined bandwidth.

4. A system according to claim 3, characterised in that
5 table also associates a maximum transmission delay
A time with the signal source, and in that the
bandwidth control means is arranged to cause the
system control means to allocate the said
predetermined bandwidth at least before the end of
10 the said maximum delay time and, in conjunction
therewith, to feed the said cell rate indicator
signal with a cell rate communicate with the
allocated bandwidth to the ingress for transmission
to the signal source.
- 15 5. A system according to claim 1, wherein the bandwidth
control means is arranged to cause allocation of
bandwidth to a signal source at a level less than the
predetermined bandwidth associated with that signal
20 source and is further arranged to downgrade the
priority of cells received at the ingress such that
the bandwidth of non-downgraded cells supplied to the
switch does not exceed the allocated bandwidth.
- 25 6. A system according to claim 2, characterised in that
the bandwidth control means includes feedback means
arranged to transmit a maximum cell rate value to the
ingress for transmission to a recognised signal
source to indicate to the signal source a maximum
30 permitted rate of supply of cells to the ingress.
7. A system according to any preceding claim,
characterised in that the bandwidth control means is
arranged automatically to cause the system control
35 means to deallocate bandwidth for cells from a
particular signal source when that signal source is
determined to be inactive.





8. A system according to any preceding claim, characterised in that the bandwidth control means includes a buffer for delaying transmission of the cells to the switch until bandwidth has been allocated for the cells.
- 5
9. A system according to claim 6, characterised in that the feedback means is arranged to transmit a halt instruction to the ingress for reception by a signal source and for instructing a signal source to cease supplying cells to the ingress when receipt of cells from a signal source has been detected and no bandwidth has been allocated for cells from that signal source.
- 10
- 15
10. A system according to claim 4, characterised in that the bandwidth control means includes timing means for measuring the period of time for which the cells are delayed.
- 20
11. A system according to claim 10, characterised by including cell deletion means for deleting cells from the buffer after they have been delayed for a predetermined period of time.
- 25
12. A method of operating a broadband switching system for transmitting information carrying asynchronously transferred data cells from a ingress of the system to an egress of the system via at least one switch, the method comprising storing information associating a predetermined bandwidth with a signal source connectible to the ingress, detecting incoming cells from the said source at the ingress, automatically in response to such detection, providing a request signal in the system for allocating the said bandwidth, selectively causing a delay in the transmission of the cells to the egress if the said
- 30
- 35





predetermined bandwidth is not allocated in the system and subsequently allowing cell transmission when allocation occurs.

- 5 13. A method according to claim 12, characterised in that the delay is effected by transmitting a feedback signal to the ingress for placing the signal source in a halt mode.
- 10 14. A method according to claim 13, characterised in that the feedback signal is a cell rate indicator signal.
- 15 15. A broadband switching system having at least one ingress for connection to a respective signal source and at least one egress for connection to a receiving system, the switching system having at least one switch for transmitting information-carrying asynchronously transferred data cells from the ingress to the egress, system control means for accepting and establishing a connection between the ingress and egress via the said switch, and bandwidth control means arranged to feed a indicator signal back to the ingress for transmission to the signal source, to detect incoming cells supplied to the ingress from the signal source, and, the system being arranged automatically to determine that a predetermined bandwidth is available for transmission of cells from the source to the egress, and to offer this bandwidth to the source using the said indicator signal.
- 20
- 25
- 30
- 35 16. A broadband switching system having at least one ingress for connection to a respective signal source and at least one egress for connection to a receiving system, the switching system having at least one switch for transmitting information-carrying asynchronously transferred data cells from the





WO 96/27964

PCT/GB96/00534

33

5 ingress to the egress, system control means for
accepting and establishing a connection between the
ingress and the egress via the said switch, means for
storing a predetermined bandwidth value associated
with an identified receiving system, and transmission
control means arranged to send a signal-source-halt
feedback signal to the signal source when bandwidth
is not available to cells from the source at a
bandwidth level corresponding to the said
10 predetermined bandwidth value, whereby the signal
source is allowed to send cells via the ingress only
when the said bandwidth level is available.

This Sheet Has Been Replaced

