

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
14 November 2002 (14.11.2002)

PCT

(10) International Publication Number  
**WO 02/091672 A2**

- (51) International Patent Classification<sup>7</sup>: **H04L 12/00**
- (21) International Application Number: PCT/US02/14264
- (22) International Filing Date: 7 May 2002 (07.05.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/288,802 7 May 2001 (07.05.2001) US
- (71) Applicant (for all designated States except US): **VITESSE SEMICONDUCTOR CORPORATION** [US/US]; 741 Calle Plano, Camarillo, CA 93012 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **TAGORE-BRAGE, Jens P.** [DK/DK]; Edvard Glaeselsvej 3, DK-2000 Frederiksberg (DK). **ANDERSEN, Martin E.** [DK/SE]; Sundspromenaden 25, S-211 16 Malmo (SE). **ANDERSEN, Ole C.** [DK/DK]; Oktobervej 119, DK-2730 Herlev (DK).
- (74) Agent: **CASTELLANO, John A.**; Harness, Dickey & Pierce, P.L.C., P.O. Box 8910, Reston, VA 20195 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT (utility model), AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ (utility model), CZ, DE (utility model), DE, DK (utility model), DK, DM, DZ, EC, EE (utility model), EE, ES, FI (utility model), FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK (utility model), SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**  
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*



**WO 02/091672 A2**

(54) Title: A SYSTEM AND A METHOD FOR PROCESSING DATA PACKETS OR FRAMES

(57) Abstract: In flow control, such as in Ethernet communication, the data is divided into queues of different priorities so that, when transmission is enabled, quality of service may be provided by first transmitting data of high priority. A networking element, such as a switch, a router, a bridge or a MAC, is adapted to transfer flow control information without awaiting its own buffer to fill. This reduces latency and facilitates the providing of networking elements without buffers.

## A SYSTEM AND A METHOD FOR PROCESSING DATA PACKETS OR FRAMES

The present method relates to a system and a method for processing data packets, and in particular to a new manner of handling flow control in networking elements.

5

A problem encountered in networking elements when the transmission rate is increased is seen in e.g. an Ethernet where, at 10 Gbit/s, each bit is 2 cm long on the transmission fibre. Thus, a 40 km fibre actually holds 2Mbit at any point in time. According to the standard of flow control, this will mean that when the receiver's buffer gets too full, the receiver transmits a flow control stop signal to the transmitter. In a worst-case scenario, the transmitter has just started transmitting a maximum frame (1522 bytes) when receiving the FC stop signal. However, before the transmitter stops transmitting, the receiver still has to have a buffer for holding a total of  $2 \times 2\text{Mbit} + 2$  maximum frames – that is, a total of approx. 1 Mbyte. This is a fact due to the chosen implementation of the flow control.

When the receiver reopens transmission, the transmitter will have buffered an amount of data received during the transmission pause. In prior art transmitters, this data will be stored in order and output in order. This, naturally, means that Quality of Service is not provided and that there is a risk of losing high priority packets due to the buffers being full or due to ageing.

This problem is caused by the prior art transmitters – or transmitting parts of standard network equipment – having the buffers placed before the analysers.

25

The present invention relates to a solution to that problem, and in a first aspect, the invention relates to a system for processing data packets or frames, the system comprising:

- 30 - means for providing the data packets or frames,  
- analysing means for providing, for each data packet or frame provided, one priority of a plurality of priorities,  
- queue means for storing the data packets or frames in each of a number of queues each relating to one or more of the plurality of priorities,

- one or more outputs for outputting the provided, analysed, and stored data packets or frames,
- means for transferring each data packet or frame from the queue means to an output,

5

wherein the transferring means are adapted to receive stopping information and to, upon receipt thereof, stop transfer of data packets or frames from at least one of the queues to at least one output.

10 Thus, the data to be buffered is analysed and put into prioritised buffer queues. Each queue may relate to more than a single priority, but the queues all relate to different (sets of) priorities. In this manner, upon reopening of transmission to a receiver, the data may be output as desired from these queues.

15 In this context, the means for providing the data may be an input port or a means actually generating the data. An input port may comprise e.g. a MAC for having the communication on the port conform to e.g. the Ethernet standard. Naturally, the same is the situation for the output(s).

20 A priority may be defined in a standard, such as the Differentiated Service in Ethernet packets. Alternatively or additionally, one or more parts of the packets or frames may be analysed in order to identify one or more features or combination of features therein. Each such feature or combination of features may have a corresponding priority, and the packet or frame will be assigned the priority of one of the features identified.

25

Normally, a priority will be an integer between 0 and 63, such as between 0 and 8. However, a priority may be represented by any sign or character as long as it is determinable (and this may simply be defined) which of two different priorities is the highest. Finally, a priority may also be an interval of e.g. integers.

30

Naturally, the queue means may be implemented in any manner – such as as a number of separate memories each holding one or more queues or a single memory holding all queues. When more queues are present in a memory, each queue may be implemented as a linked list. Also, the full data packet or frame may be stored in one piece or a number  
35 of separate pieces. In addition, it may be desired to actually only store part of the packet

or frame, if the remaining part(s) thereof may be regenerated when the packet or frame is to be retrieved and further processed.

In most instances, the transferring means will transfer a data packet or frame to a single  
5 output. However, if the packet or frame is a multicast frame or a broadcast frame, the packet or frame may be transmitted to a plurality of outputs.

The stopping information may be received from a networking element receiving data packets or frames output from one or more of the outputs, such as an input connected to  
10 the same networking element. In normal flow control operation, the present system will also be adapted to receive data packets or frames from the receiving networking element – and part of those packets or frames will be flow control packets informing the present system of the status of the receiving networking element.

15 Another manner of identifying that the receiving networking element is not able to receive more data will be when it does not acknowledge receipt thereof. In this manner, the stopping information may be generated internally in the system by a monitoring means monitoring the transmission and acknowledgements and which signals when an expected acknowledgement is not received.

20

Naturally, under all circumstances, the data packets or frames may conform to any known standard – or to none at all – and may be of a fixed or variable size.

The stopping information may inform the system to only stop certain (normally the lower)  
25 priorities from being transmitted – while higher priorities may still be transmitted.

Alternatively, the transferring means could be adapted to, upon receipt of the stopping information, stop transfer of data packets or frames from all queues to at least one output – normally in an embodiment where the system comprises a plurality of outputs. In this manner, the other outputs may be able to transmit even though one receiver is congested.

30

Preferably, the transferring means are adapted to receive the stopping information via the one or more outputs – or inputs connected to the same networking element(s). Normally, the transferring means will receive transmission information at regular intervals. Also, a congested receiver may transmit stopping information at regular intervals, and the

transferring means may stop transfer until no stopping information has been received in a period of time exceeding the regular interval.

In one embodiment, the system will comprise a plurality of outputs and wherein the  
5 transferring means comprises output queue means for holding data to be output from each of the outputs, wherein the transferring means are adapted to transfer the data packets from the queue means to the relevant output queue means and from the output queue means to the respective outputs.

10 In this situation, it is preferred that the output queue means actually implements not only a single queue for each output but a number of queues so that, for each output, prioritised queues (as the queue means) receive data packets or frames with different priorities. An easy implementation would be one where the number of output queue means for each output corresponds to the total number of priorities or queues in the queue means.

15

Especially when the transferring means are adapted to transfer data packets or frames from a higher priority queue means before a lower priority queue means to the respective output queue means, any stopping and reopening of transmission to a receiver will implement Quality of Service in that the buffered higher priority packets are transmitted  
20 first. In the situation where the output queue means also have prioritised queues, an even better performance is achieved.

In a preferred embodiment, in fact, the functionality of the above queue means and output queue means are combined and the queue means are provided with one queue for each  
25 priority for each of a plurality of outputs. In this manner, each data packet or frame may be transferred to the correct output/priority at once.

When the providing means and each output is adapted to provide and output data packets or frames at a predetermined rate (such as measured in frames /sec or bits/bytes /sec),  
30 preferably the analysing means are adapted to analyse the data frames or packets with at least the predetermined rate. It is even more preferred that the analysing means are adapted to analyse the data frames or packets at least a rate corresponding to the total predetermined outputting rate of all outputs.

Preferably, the queue means are adapted to, in each queue, hold data packets or frames in the order in which they were received. Such queues may be implemented as a FIFO.

As described, the queue means may comprise a single memory and each queue may be implemented as a linked list of data packets or frames. When the total amount of data to be buffered is defined by latency in a flow control system – the amount of data to buffer is defined by the transmission rate and the time between a receiver receiving data from the system is congested until a transmitter transmitting data to the present system is stopped. In this manner, a single memory with the queues as linked lists may be defined with no prior knowledge of the composition of the data (as to priority). Otherwise, the individual prioritised queues might be required to be dimensioned to handle the full amount of data. This amount of data will depend both on the transmission rate and distance. Preferably, the queue means are adapted to hold at least 1Mbit, such as at least 4Mbit, preferably at least 6Mbit, such as at least 1Mbyte of data.

15

In the same manner, the output queue means may comprise a single memory and each output queue may be implemented as a linked list of data packets or frames. Also, the output queue means may be adapted to hold at least 1Mbit, such as at least 4Mbit, preferably at least 6Mbit, such as at least 1Mbyte of data.

20

It is especially desired that the prioritised queues are adapted to buffer a large amount, if not all, of the data to be buffered.

Preferably, the analysing means are also adapted to determine, on the basis of one or more predetermined parts of each data packet or frame, from which output the data packet or frame should be output. In normal systems, this is obtained by learning and creation of a lookup table or function where networking information (such as MAC addresses or IP addresses) is combined with output port numbers.

30 One type of networking element where the present system is useful is one where the providing means comprise a number of inputs adapted to receive the data packets or frames from an external network. An element of this type may be a switch or a hub.

In a second aspect, the invention relates to a method for processing data packets or frames, the method comprising:

35

- providing the data packets or frames,
  - providing, for each data packet or frame provided, one priority of a plurality of priorities,
- 5 - storing the data packets or frames in one of a number of queues each relating to one or more of the plurality of priorities,
- receiving each data packet or frame from the queues and outputting the data packet or frame from one or more outputs,
- 10 wherein the receiving and outputting step comprises receiving stopping information and stopping transfer of data packets or frames from at least one of the queues to at least one of the outputs.

In one embodiment, the receiving and outputting step comprises stopping transfer of data  
15 packets or frames from all queues to at least one of the outputs.

Preferably, the receiving and outputting step comprises receiving the stopping information via the one or more outputs.

- 20 In one embodiment, the receiving and outputting step comprises the steps of:
- transferring the data packets or frames from the queues to an output queue means for holding data to be output from the respective one of a plurality of outputs,
  - transferring the data packets from the queues to the relevant output queue means and from the output queue means to the respective outputs.

25

Then, the receiving and outputting step may comprise transferring data packets or frames from a higher priority queue before a lower priority queue to the respective output queue means.

- 30 When the step of providing the data packets or frames and each output outputs data packets or frames with a predetermined rate (such as measured in frames /sec or bits/bytes /sec), the step of providing a priority is preferably performed with at least the predetermined rate. Even more preferably, the providing of the priority is performed with at least a rate corresponding to the total predetermined outputting rate of all outputs.

35

In the preferred embodiment, the storing step comprises storing each data packet or frame in a queue corresponding to an output from which it should be output and to the priority assigned. In this manner, a queue is provided for each priority and for each output.

- 5 In another embodiment, the receiving and outputting step comprises receiving data packets from each of the queues and outputting the data packets or frames from one of a plurality of outputs.

Preferably, the storing step comprises storing the data packets or frames in each queue in  
10 the order in which they were received.

Also, the storing step may comprise storing the data packets or frames in a single memory, each queue being implemented as a linked list of data packets or frames.

- 15 Preferably, the storing step comprises storing the data packets or frames in queues implemented in a memory adapted to hold at least 1Mbit, such as at least 4Mbit, preferably at least 6Mbit, such as at least 1Mbyte of data.

In addition, the storing step may comprise storing the data packets or frames in a single  
20 memory, each output queue being implemented as a linked list of data packets or frames.

Further, the transferring step may comprise transferring the data packets or frames to output queue means implemented in a memory adapted to hold at least 1Mbit, such as at least 4Mbit, preferably at least 6Mbit, such as at least 1Mbyte of data.

25

In general, the system preferably comprises the step of analysing one or more predetermined parts of each data packet or frame in order to determine from which output the data packet or frame should be output.

- 30 As mentioned above, an especially interesting embodiment is one where the step of providing the data packets or frames comprises providing the data packets or frames from a number of inputs adapted to receive the data packets or frames from an external network. A method of this type may be used in a switch or a hub.



A very interesting networking element may be seen in a third aspect relating to a networking element having:

- 5 - a first input/output connected to a first link communicating with a second networking element, the first I/O being adapted to receive data from the second networking element and to transmit stopping/congestion information to the second networking element,
- 10 - a second input/output connected to a second link communicating with a third networking element, the second I/O being adapted to transmit data to the third networking element and to receive stopping/congestion information from the third networking element,
- first means for receiving data from the first I/O and transmitting data to the second I/O,
- second means for receiving stopping/congestion information from the second I/O and transmitting stopping/congestion information to the first I/O,

15

the second means being adapted to, upon receipt of the stopping/congestion information from the second I/O, transmit stopping/congestion information to the first I/O without delay.

20 In this context, the first networking element may be a switch, a router or just a Media Access Controller. This element is adapted to transfer data (maybe with some intermediate analysis or alteration) from one I/O to another. In the opposite direction, stopping/congestion information may be transferred (maybe also with some intermediate analysis or alteration) in order for the third networking element to signal that it is no longer  
25 interested in – or able to – receive additional data.

Normally this stopping/congestion information may be pause frames, flow control information or other congestion signalling adapted to provide e.g. back pressure or the like and to ensure that buffers do not overflow in the system.

30

In normal networking elements, all networking elements would have buffers and merely stop transmission of data to the congested receiver. The networking element would then itself keep receiving data and simply store this. When the buffer of the networking element is also congesting, the networking element would then assert back pressure on the  
35 networking element from which it receives data in order for that to then stop transmitting.

This type of communication has the disadvantage that all networking elements need buffers and need to assert back pressure or the like upon a certain degree of filling or congestion of the buffer. This provides some latency in the stopping of the actual data transmitter – and requires all networking elements to have buffers.

5

The present networking element transmits the stopping/congestion information as soon as it receives it. In this manner, there is no waiting for any buffer of the networking element to fill. Also, the actual transmitter of the data is stopped sooner. In this context, “without delay” will mean that the stopping/congestion information is transmitted as soon as possible. Inevitable delays may occur due to analysis/alteration of the stopping/congestion information or waiting for the link to the second networking element to become available.

Thus, in one embodiment, the first means and the first and second I/O’s each comprise no more data storage than that required for holding a single data packet or frame. This is not required in that the second networking element is stopped as soon as possible.

In another embodiment, at least one of the first means and the first and second I/O’s comprises a data storage, but then the second means is adapted to transmit the stopping/congestion information independently of a degree of filling/congestion of the data storage. Naturally, a storage may be provided if desired.

The first means could be adapted to stop transmission of data to the second I/O when the first means has received stopping/congestion information. In this situation, a storage would be desired in order to then store any data received until the second networking element stops transmitting data.

Alternatively, the first means could be adapted to maintain transmission of any data received from the first I/O to the second I/O when the first means has received stopping/congestion information. In this situation, which may be termed a cut-through operation, no buffer is required.

Another aspect relates to communication system comprising:

- a first networking element according to the third aspect,
- a second networking element communicating with the first networking element via the first I/O of the first networking element and

35

- a third networking element communicating with the first networking element via the second I/O of the first networking element,

the second networking element being adapted to transmit data to the first I/O, to receive  
5 stopping/congestion data from the first I/O, and to stop transmission of data upon receipt of stopping/congestion data,

the third networking element being adapted to receive data from the second I/O and to transmit stopping/congestion data to the second I/O.

10

A fifth aspect relates to a method of receiving and transmitting data in and from a networking element, the method comprising:

- receiving, on a first input/output of the networking element connected to a first  
15 link communicating with a second networking element, data from the second networking element and transmitting stopping/congestion information to the second networking element,
- transmitting, on a second input/output of the networking element connected to a second link communicating with a third networking element, data to the third  
20 networking element and receiving stopping/congestion information from the third networking element,
- in a first means, receiving data from the first I/O and transmitting data to the second I/O,
- in a second means, receiving stopping/congestion information from the second  
25 I/O and transmitting stopping/congestion information to the first I/O,

the second means, upon receipt of the stopping/congestion information from the second I/O, transmitting stopping/congestion information to the first I/O without delay.

30 Again, the first means and the first and second I/O's could each store, at each point in time, no more data storage than that required for holding a single data packet or frame.

Alternatively, at least one of the first means and the first and second I/O's could store data, and wherein the second means could transmit the stopping/congestion information  
35 independently of an amount of data stored.

Also, the first means could stop transmission of data to the second I/O when the first means has received stopping/congestion information. Alternatively, the first means could maintain transmission of any data received from the first I/O to the second I/O when the  
5 first means has received stopping/congestion information.

A final aspect relates to a method of operating the system according to the third aspect, the method comprising:

- 10 - during transmission of data from the second networking element to the first I/O of the first networking element and from the second I/O of the first networking element to the third networking element, transmitting stopping/congestion information from the third networking element to the second I/O of the first networking element,
- 15 - with no delay transmitting stopping/congestion information from the first I/O to the second networking element, and  
stopping, in the third networking element, transmission of data to the first I/O.

In the following, three embodiments of the invention will be described with reference to  
20 the drawing wherein:

Fig. 1 illustrates a first embodiment of a device for receiving, processing and outputting data packets,

25 Fig. 2 illustrates a second embodiment of a device for receiving, processing and outputting data packets, and

Fig. 3 illustrates a third embodiment of handling stopping/congestion information.

30 In Fig. 1, the device has an input 10 receiving data packets or frames from e.g. a data network or data provider. The device is able to categorize data packets or frames, in categorizer 12, into one of four priorities.

These packets or frames are analysed or categorised in order to provide a priority for  
35 each. This categorisation is performed on the basis of the contents of certain fields of the

packet or frame. A suitable categorizer may be seen from the US application No. 09/965,078 titled "A method and system for analysing a data packet or frame" and filed on 28 September, 2001 – and which is hereby incorporated by reference.

- 5 The input 10 may be one or a number of separate inputs receiving data packets or frames from a number of different networks or parts of a network (such as via a number of different networking elements).

The device further has a number of outputs 18, and four queues 14 (one for each priority)  
10 for each output 18.

The categorizer 12 is further adapted to determine which output a data packet or frame is to be output from and to transfer each categorized packet or frame into the queue 14 corresponding to the priority assigned and to the output determined.

15

The queues 14 may be implemented as a number of separate FIFO's.

Finally, a shaper 16 is provided for each output 18. Each shaper 16 is adapted to transfer data packets from the queues 14 to the corresponding output 18. The shaper 16 is  
20 adapted to transfer data packets or frames from that non-empty queue 14 having the highest priority to each output.

The arrows of the outputs 18 point in two directions in that the device is flow control enabled and therefore able to receive flow control information from networking units  
25 receiving the data packets or frames output from the device on outputs 18.

This networking information is used in the receiving shaper 16 to stop sending data packets or frames. Upon "reopening" of transmission, any later received high priority data packets or frames will be sent before lower priority packets or frames. In this manner,  
30 preference is made toward the higher priority packets, thus implementing Quality of Service and preventing Head of Line blocking.

It should be noted that the stopping of sending data may be only from one or more of the queues of an output. It may be desired that only lower priority data is stopped in order to

give the receiver a chance of reducing its work load or the amount of data in its buffers – but still maintain data flow at the higher priorities.

In Fig. 2, the queues 14 are implemented in a single memory block, such as a RAM, as  
5 linked lists. In this manner, the overall storage capacity required may be reduced.

Fig. 3 illustrates a new manner of handling stopping/congestion information such as  
pause frames, flow control information or other congestion information transmitted  
between networking elements (such as switches, routers, bridges, MAC's or the like) in  
10 order to control data flow and to ensure that data is not lost.

The system of Fig. 3 comprises a networking element 20 according to the third aspect.  
This networking element communicates, via two-way links illustrated by arrows, with a  
second networking element 22 and a third networking element 24. The networking  
15 element 22 transmits data to the element 20, which transmits data to the element 24.  
Also, the element 24 is able to transmit stopping/congestion information to the element 22  
via the element 20 in order to reduce or stop the amount of data transmitted.

Naturally, the data on the links may have different formats (such as Ethernet, SPI4v2,  
20 XAUI, XGMII, GMII, MII – or any other format – packet based, cell based, frame based)  
and the element 20 may perform any analysis or alteration thereof (in the means 30 or 26)  
before sending the data to the element 24.

Also, the stopping/congestion information transmitted from the element 24 may have any  
25 form – and a form different from that on which stopping/congestion information is  
transmitted between the elements 20 and 22.

The element 20 has, in this embodiment, a buffer 26 for buffering data received from the  
element 22 before transmission to the element 24. This buffer may have any size. In one  
30 embodiment where the link between the elements 22 and 20 is cell-based, and the  
communication between the elements 20 and 24 is frame based, the buffer has a size of  
at least one frame/packet of the largest possible size in that the frame needs to be  
assembled from the cells before transmission to the element 24.

When the element 24 transmits stopping/congestion information, the element 20 immediately transmits stopping/congestion information to the element 22. It should be noted that different types of stopping/congestion information may be used (such as for stopping all - or only part of the data traffic).

5

The element 22 will then stop transmission of data either for a predetermined period of time or until instructed to start transmission.

The element 20 may also stop (using means 30) transmitting data to the element 24 and 10 instead buffer any data received in the buffer 26. Alternatively, the buffer 26 may be provided only for e.g. reassembly purposes and the element 20 will keep transmitting any data received to the element 24 which will then receive and buffer the data.

Thus, in the last situation, the element 20 may be a cut-through element not delaying the 15 data transmission more than what is required for the analysis/alteration/reassembly necessarily handled by that element.

## CLAIMS

1. A system for processing data packets or frames, the system comprising:
- 5 - means for providing the data packets or frames,
  - analysing means for providing, for each data packet or frame provided, one priority of a plurality of priorities,
  - queue means for storing the data packets or frames in each of a number of queues each relating to one or more of the plurality of priorities,
  - 10 - one or more outputs for outputting the provided, analysed, and stored data packets or frames,
  - means for transferring each data packet or frame from the queue means to an output,
- 15 wherein the transferring means are adapted to receive stopping information and to, upon receipt thereof, stop transfer of data packets or frames from at least one of the queues to at least one output.
2. A system according to claim 1, wherein the transferring means are adapted to, upon  
20 receipt of the stopping information, stop transfer of data packets or frames from all queues to at least one output.
3. A system according to claim 1 or 2, wherein the transferring means is adapted to  
25 receive the stopping information via the one or more outputs.
4. A system according to any of claims 1-3, the system comprising a plurality of outputs  
and wherein the transferring means comprises output queue means for holding data to be  
output from each of the outputs, wherein the transferring means are adapted to transfer  
the data packets from the queue means to the relevant output queue means and from the  
30 output queue means to the respective outputs.
5. A system according to claim 4, wherein the transferring means are adapted to transfer  
data packets or frames from a higher priority queue means before a lower priority queue  
means to the respective output queue means.



6. A system according to any of the preceding claims, wherein the providing means and each output are adapted to provide and output data packets or frames at a predetermined rate, and where the analysing means are adapted to analyse the data frames or packets with at least the predetermined rate.

5

7. A system according to any of the preceding claims, wherein the queue means are adapted to, in each queue, hold data packets or frames in an order in which they were received.

10 8. A system according to any of the preceding claims, wherein the queue means comprise a single memory and wherein each queue is implemented as a linked list of data packets or frames in the memory.

9. A system according to any of the preceding claims, wherein the queue means are  
15 adapted to hold at least 1Mbit, such as at least 4Mbit, preferably at least 6Mbit, such as at least 1Mbyte of data.

10. A system according to claim 3, wherein the output queue means comprise a single memory and wherein each output queue is implemented as a linked list of data packets or  
20 frames in the memory.

11. A system according to claim 3, wherein the output queue means are adapted to hold at least 1Mbit, such as at least 4Mbit, preferably at least 6Mbit, such as at least 1Mbyte of data.

25

12. A system according to any of the preceding claims, wherein the analysing means are adapted to determine, on the basis of one or more predetermined parts of each data packet or frame, from which output the data packet or frame should be output.

30 13. A system according to any of the preceding claims, wherein the providing means comprise a number of inputs adapted to receive the data packets or frames from an external network.

14. A method for processing data packets or frames, the method comprising:

35

- providing the data packets or frames,
  - providing, for each data packet or frame provided, one priority of a plurality of priorities,
  - storing the data packets or frames in one of a number of queues each relating to
- 5 one or more of the plurality of priorities,
- receiving each data packet or frame from the queues and outputting the data packet or frame from one or more outputs,

wherein the receiving and outputting step comprises receiving stopping information and

10 stopping transfer of data packets or frames from at least one of the queues to at least one of the outputs.

15 15. A method according to claim 14, wherein the receiving and outputting step comprises stopping transfer of data packets or frames from all queues to at least one of the outputs.

15

16. A method according to claim 14 or 15, wherein the receiving and outputting step comprises receiving the stopping information via the one or more outputs.

17. A method according to any of claims 14-16, where the receiving and outputting step

20 comprises the steps of:

- transferring the data packets or frames from the queues to an output queue means for holding data to be output from the respective output of a plurality of outputs, and
- transferring the data packets from the output queue means to the respective outputs.

25

18. A method according to claim 17, wherein the receiving and outputting step comprises transferring data packets or frames from a higher priority queue before a lower priority queue to the respective output queue means.

30 19. A method according to any of claims 14-18, wherein the step of providing the data packets or frames is provided with a predetermined rate and each output provides a predetermined rate, and where the step of providing a priority is performed with at least the predetermined rate.

20. A method according to any of claims 14-16, wherein the receiving and outputting step comprises receiving data packets from each of the queues and outputting the data packets or frames from one of a plurality of outputs.
- 5 21. A method according to any of claims 14-20, wherein the storing step comprises storing the data packets or frames in each queue in an order in which they were received.
22. A method according to any of claims 14-21, wherein the storing step comprises storing the data packets or frames in a single memory, each queue being implemented as  
10 a linked list of data packets or frames.
23. A method according to any of claims 14-22, wherein the storing step comprises storing the data packets or frames in queues implemented in a memory adapted to hold at least 1Mbit, such as at least 4Mbit, preferably at least 6Mbit, such as at least 1Mbyte of  
15 data.
24. A method according to claim 17, wherein the transferring step comprises transferring the data packets or frames to a single memory, each output queue being implemented as a linked list of data packets or frames.  
20
25. A method according to claims 17, wherein the transferring step comprises transferring the data packets or frames to output queue means implemented in a memory adapted to hold at least 1Mbit, such as at least 4Mbit, preferably at least 6Mbit, such as at least 1Mbyte of data.  
25
26. A method according to any of claims 14-25, further comprising the step of analysing one or more predetermined parts of each data packet or frame in order to determine from which output the data packet or frame should be output.
- 30 27. A method according to any of claims 14-26, wherein the step of providing the data packets or frames comprises providing the data packets or frames from a number of inputs adapted to receive the data packets or frames from an external network.
28. A networking element having:

- a first input/output connected to a first link communicating with a second networking element, the first I/O being adapted to receive data from the second networking element and to transmit stopping/congestion information to the second networking element,
- 5 - a second input/output connected to a second link communicating with a third networking element, the second I/O being adapted to transmit data to the third networking element and to receive stopping/congestion information from the third networking element,
- first means for receiving data from the first I/O and transmitting data to the  
10 second I/O,
- second means for receiving stopping/congestion information from the second I/O and transmitting stopping/congestion information to the first I/O,

the second means being adapted to, upon receipt of the stopping/congestion information  
15 from the second I/O, transmit stopping/congestion information to the first I/O without delay.

29. A networking element according to claim 28, wherein the first means and the first and second I/O's each comprise no more data storage than that required for holding a single  
20 data packet or frame.

30. A networking element according to claim 28, wherein at least one of the first means and the first and second I/O's comprises a data storage, and wherein the second means is adapted to transmit the stopping/congestion information independently of a degree of  
25 filling/congestion of the data storage.

31. A networking element according to any of claims 28-30, wherein the first means is adapted to stop transmission of data to the second I/O when the first means has received stopping/congestion information.  
30

32. A networking element according to any of claims 28-30, wherein the first means is adapted to maintain transmission of any data received from the first I/O to the second I/O when the first means has received stopping/congestion information.

35 33. A communication system comprising:

- a first networking element according to any of claims 28-32,
- a second networking element communicating with the first networking element via the first I/O of the first networking element and
- a third networking element communicating with the first networking element via  
5 the second I/O of the first networking element,

the second networking element being adapted to transmit data to the first I/O, to receive stopping/congestion data from the first I/O, and to stop transmission of data upon receipt of stopping/congestion data,

10

the third networking element being adapted to receive data from the second I/O and to transmit stopping/congestion data to the second I/O.

34. A method of receiving and transmitting data in and from a networking element, the  
15 method comprising:

- receiving, on a first input/output of the networking element connected to a first link communicating with a second networking element, data from the second networking element and transmitting stopping/congestion information to the  
20 second networking element,
- transmitting, on a second input/output of the networking element connected to a second link communicating with a third networking element, data to the third networking element and receiving stopping/congestion information from the third networking element,
- 25 - in a first means, receiving data from the first I/O and transmitting data to the second I/O,
- in a second means, receiving stopping/congestion information from the second I/O and transmitting stopping/congestion information to the first I/O,

30 the second means, upon receipt of the stopping/congestion information from the second I/O, transmitting stopping/congestion information to the first I/O without delay.

35 35. A method according to claim 34, wherein the first means and the first and second I/O's each store, at each point in time, no more data storage than that required for holding a single data packet or frame.

36. A method according to claim 34, wherein at least one of the first means and the first and second I/O's stores data, and wherein the second means transmits the stopping/congestion information independently of an amount of data stored.

5

37. A method according to any of claims 34-36, wherein the first means stops transmission of data to the second I/O when the first means has received stopping/congestion information.

10 38. A method according to any of claims 34-36, wherein the first means maintains transmission of any data received from the first I/O to the second I/O when the first means has received stopping/congestion information.

39. A method of operating the system according to claim 33, the method comprising:

15

- during transmission of data from the second networking element to the first I/O of the first networking element and from the second I/O of the first networking element to the third networking element, transmitting stopping/congestion information from the third networking element to the second I/O of the first networking element,
- 20 - with no delay transmitting stopping/congestion information from the first I/O to the second networking element, and
- stopping, in the third networking element, transmission of data to the first I/O.

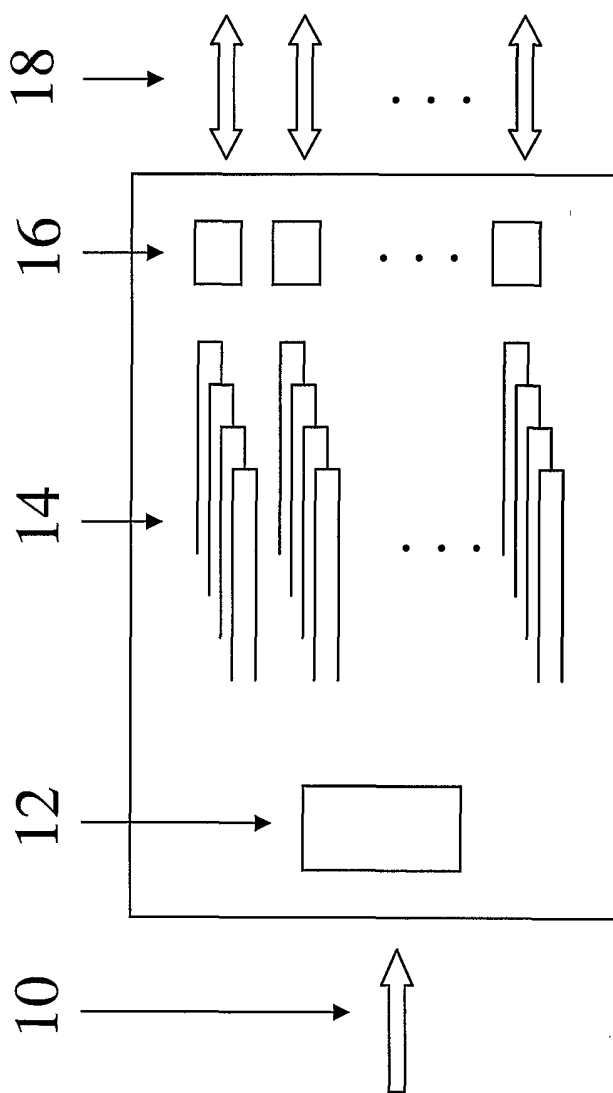


Fig. 1

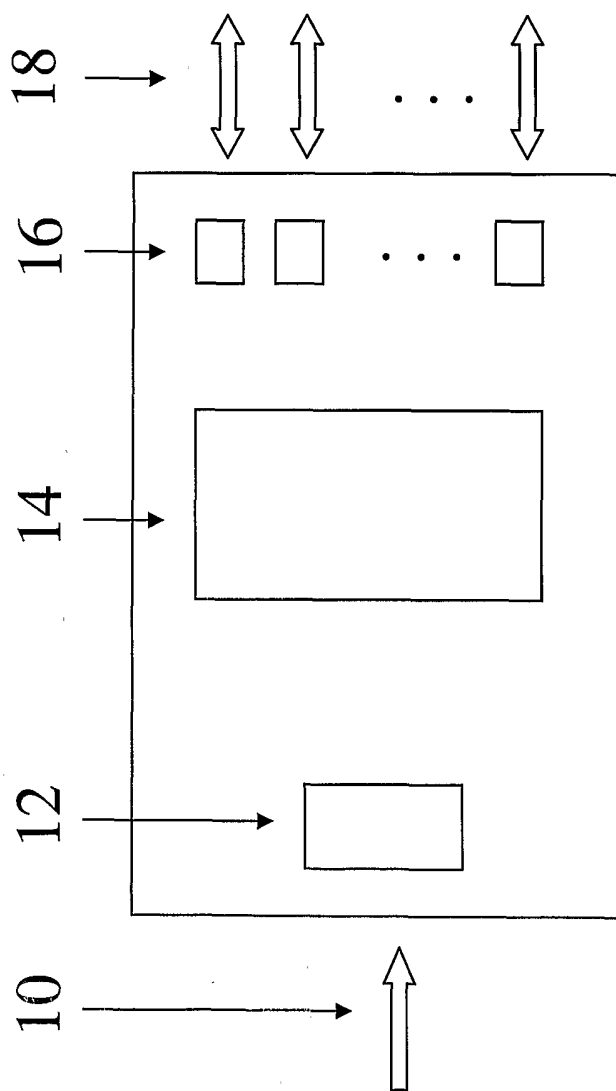


Fig. 2



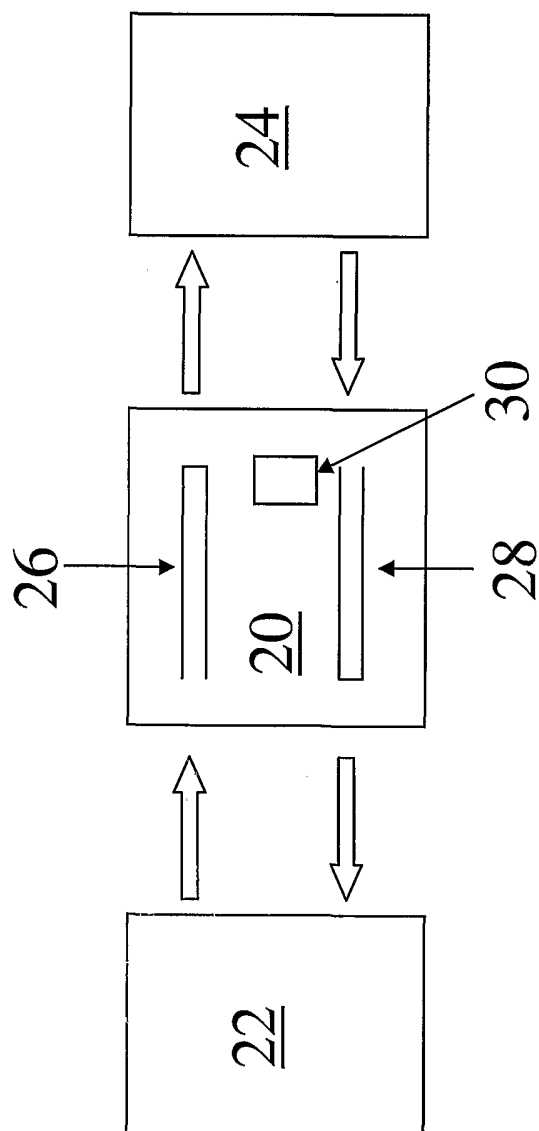


Fig. 3