PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H04L 12/26 — A2

(11) International Publication Number: **WO 96/38955**

(43) International Publication Date: 5 December 1996 (05.12.96)

(21) International Application Number: PCT/EP96/02335

(22) International Filing Date: 30 May 1996 (30.05.96)

(30) Priority Data:
9510931.0 31 May 1995 (31.05.95) GB

(71) Applicant *(for all designated States except US)*: 3COM IRELAND [-/-]; Upland House, P.O. Box 309, Georgetown, Grand Cayman (KY).

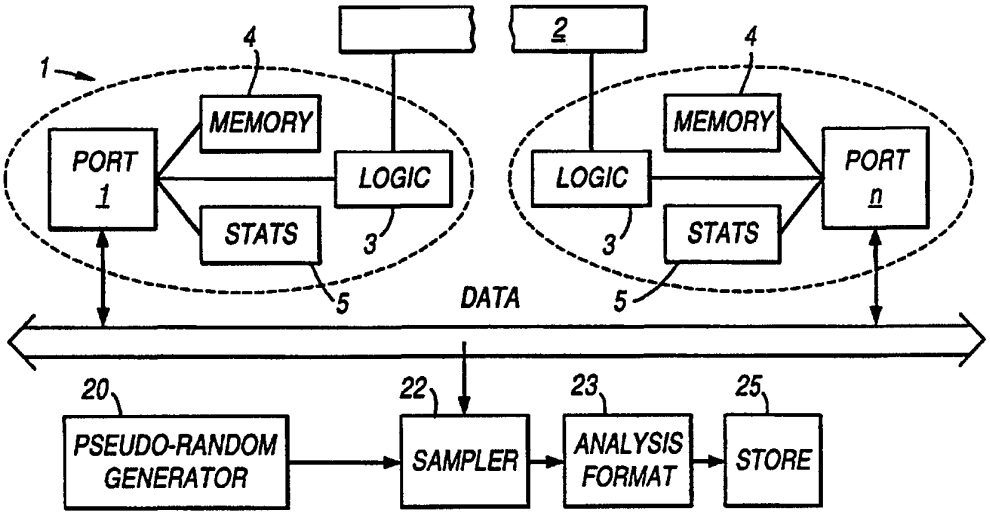
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: CREEDON, Tadhg [IE/IE]; Colsmeagmore, Furbo, Co. Galway (IE). GAVIN, Vincent [IE/IE]; 18 The Old Rectory, Chapel Hill, Lucan, Co. Dublin (IE).

(74) Agent: CRAWFORD, Andrew, Birkby; A.A. Thornton & Co., Northumberland House, 303-306 High Holborn, London WC1V 7LE (GB).

(81) Designated States: AU, CA, GB, JP, KR, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: TRAFFIC MONITORING AND CONTROL IN A SWITCH



(57) Abstract

In order to monitor the traffic in a part of a network a switch or bridge (1) includes a processor and a store. The processor is arranged to sample the traffic on a random or pseudo-random basis in real time and store the results for analysis by a management entity. The whole of a packet (10) may be captured and stored by the sampling device or, for more specific monitoring, only a predetermined portion of the packet may be stored.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

TRAFFIC MONITORING AND CONTROL IN A SWITCH

The present invention relates to computer networks and more particularly to ethernet, token ring and ATM networks.

In ethernet network systems it is known to provide management facilities which can accumulate traffic statistics from stores located at the individual ports. These statistics are required in order to efficiently manage the network. In many cases, the production of these statistics is deemed to be a long trend analysis lasting many hours or days in order for a network manager to take decisions regarding an appropriate layout of the network.

While some current systems are exact they are unweildy and costly since the level of detail is not necessary in order to give an accurate profile of traffic. Others gather all data in-real time and process it later to compress it for reduced storage. We propose a scheme to gather only extracts from the data and process it in real time.

It is known to sub-divide a network into discrete zones by switches or bridges which already have a microprocessor and an amount of memory in order to handle traffic passing through the bridge.

The present invention provides a switch or bridge including a processor and a store which will be used to store data relating to traffic on a part of the network. We propose a scheme to gather only extracts from the data and process it in real time which is more cost effective. Traffic will be sampled on a random or pseudo random basis and details of the traffic at that time stored in the store for analysis by a management entity.

The advantage of this approach is that it has a low memory overhead and requires little increase in the number of circuits involved and consequently has little effect on the overall cost of the network. Further, the sampling and storing can be done by the processor of a switch or bridge in the background e.g. by an interrupt routine. It is, however, useful for logging one or more of a

number of parameters in relation to one or more ports in the bridge or switch. For example, it may be used to provide the profile of the most frequent users on that port identified by MAC addresses or the profile of generators of errors or of small or larger packet sizes and the profile of generators of broadcast/multicast transmissions. For more flexible capabilities, the facility can be programmed to capture data from any specified port or indeed any specified packet.

In order that the present invention be more readily understood, an embodiment thereof will now be described by way of example only with reference to the accompanying drawings in which:-

Fig. 1 shows the form of a typical packet of information; and

Fig. 2 shows diagrammatically a circuit layout for achieving the monitoring.

As indicated in Fig. 1, a typical packet 10 of information on an ethernet network consists of a block of information 11 indicating the destination address of the packet of information, a source address 12 indicating the device from which the information is derived, other control information 13, actual data 14 to be transmitted, and finally a block 15 indicating any errors which have occurred.

The present invention provides a system for storing data relating to one or more of the destination address, source address, some of the control information and the errors in stores to allow analysis of traffic on the system. In order to reduce costs, instead of storing every single transmission, it is intended simply to sample the traffic and store only predetermined parts of the packet. We consider that this is best done on a random or pseudo random basis in order to avoid any possibility of missing cyclically repeating information. One form of apparatus according to the present invention is shown in more detail in Figure 2. A switch or bridge 1 is represented as comprising a number of ports 1...n, only two of which are shown. Data flow through the ports is controlled by a management entity 2. Each port is identical and includes a number of circuits including a port logic circuit 3, a data FIFO store 4 for data received by and transmitted by the port and

a statistics memory section 5 which stores details of all packets generated by the device(s) (not shown) connected to the port.

In order to sample data flowing through the bridge or switch a pseudo random generator 20 is provided which generates a sampling pulse. This sampling pulse operates a gating circuit 21 which feeds information from the next complete packet 22 of a stream of data on the data bus of the network after the timing pulse through an analysis and formatting circuit 23 into the additional memory 25 and then closes the gate at the end of that packet until such time as the next timing pulse is generated by the pseudo random timer.

For the sake of completeness, the way in which the packet is captured will be described. When the pseudo-random generator 20 generates a pulse, the data is examined for a data sequence indicating the start of a packet. Any data existing prior to such an indication is simply passed through. Once the start of a packet is detected, an extract of the packet is stored in a formatted fashion

The sampler 22 is arranged to identify the start of the next complete packet after the generation of the psuedo-random timing pulse by the generator 20 and then only extract those relevant portions of the packet. This is possible because the structure of an ethernet packet is such that the addressing and control blocks are of known size and in a known position in the packet. Once the start of a packet is detected it is simply a matter of control in order to capture the desired part.

Using this technique it is possible to capture and store information from within the data section 14 of the packet simply by appropriate programming of the sampler 22.

This low cost technique can be readily implemented and can be used, if necessary, to provide details of traffic which could be further utilized in order to adaptively control various parameters of the network and assist in debugging the network.

One use of the above described arrangements in the area of so-called

storm protection and this will now be described in detail.

It is customary in networks for every device to have a unique address associated with it. Usually networks are designed to allow a one-to-one communication between ports of the network to which the devices are attached, but it is often the case that one port wishes to broadcast the same message to all other ports within the network. While this is acceptable, one has to guard against the possibility of a so-called storm of such broadcast data occurring in view of the fact that this would normally jam buffer stores associated with each port and also with the fact that the whole network would be slowed down.

When the switch or bridge monitors traffic through it, certain parameters of the statistics related to the traffic through the bridge or switch can be utilised in order to detect the onset of a storm condition and takes action to control the level of such traffic within limits.

Preferably, each port has associated with it one bit which can be controlled in order to prevent a storm of multicast/broadcast data being forwarded to all the ports of the switch and jamming the system.

The switch or bridge can simply identify from previous receptions of requests for multicasts or broadcasts that it will become overloaded and thus refuse to handle any further requests for multicasts or broadcasts until a suitable time.

In more detail, it is assumed that the network is an ethernet network comprising one or more switches or bridges each having a number of ports as well as a management entity for the switch or bridge. Each of the ports of a switch is known to the management entity of the switch and each port has storage and logic circuits.

The logic circuits may be arranged to add to the port number a digital bit whose value can be altered in response to instructions from the management entity. The storage circuits may include stores for logging all traffic through the port including accurate statistics of all broadcast and/or multicast messages created

by the device attached to the port.

The management entity monitors traffic through all ports of the switch or bridge and compiles statistics relating to the traffic. Included in the statistics are the number of occasions multicast/broadcast data is being supplied. On the basis of the statistics, the management entity makes a decision as to whether to permit a further multicast/broadcast data message to pass through the switch or bridge. If the management entity determines that a storm would result from a further multicast/ broadcast data message which would normally jam the FIFO's related with each port and also slow down the network, the management entity then changes the additional bit per port which directs any subsequent multicast/broadcast message to a non-existent port of the switch. In our system this is port 31.

It is considered that this system of diversion to a non-existent port is superior to that of simply turning off a port which is attempting to transmit a multicast/broadcast message or of detecting all broadcast or multicast traffic input through the port since the statistics relating to usage of the network are being maintained and consequently the management entity can detect when traffic on the network reduces to such a level that further multicast/broadcast message can be accommodated. At that point, the additional bit per port can be changed back to permit multicast/broadcast messages.

This particular method permits unicast messages to be handled by the network as usual whether or not the storm protection mechanism is in operation.

CLAIMS:

1. A method of monitoring traffic on a network comprising,

sampling packets on the network at intervals to generate a statistical representation of traffic flow.

2. The method of claim 1, wherein the sampling is carried out at random or pseudo random intervals.

3. The method of any preceding claim wherein whole packets are sampled.

4. The method of any of claims 1 to 3, wherein pre-determined parts of packets are sampled.

5. The method of claims 1 to 4, wherein analysis of the network traffic is carried out in real time.

6. The method of any of claims 1 to 5, wherein analysis of the network traffic is carried out not in real time.

7. The method of any of claims 1 to 6, further comprising preventing, on the basis of the representation of traffic flow, devices found to be producing large quantities of traffic from transmitting messages when such transmission would be likely to cause substantial slowing of the network.

8. The method of claim 7, wherein devices are prevented from transmitting multicasts or broadcasts in particular when a storm is deemed likely.

9. The method of either of claims 7 or 8 wherein

devices which are found to be generating large numbers of multicasts or broadcasts are prevented from transmitting multicasts or broadcasts.

1/1

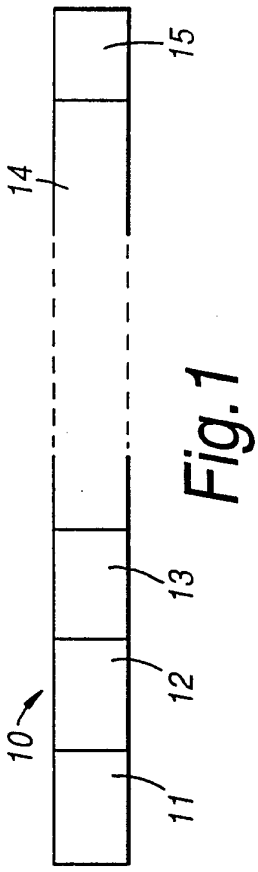

10
11
12
13
14
15

Fig. 1

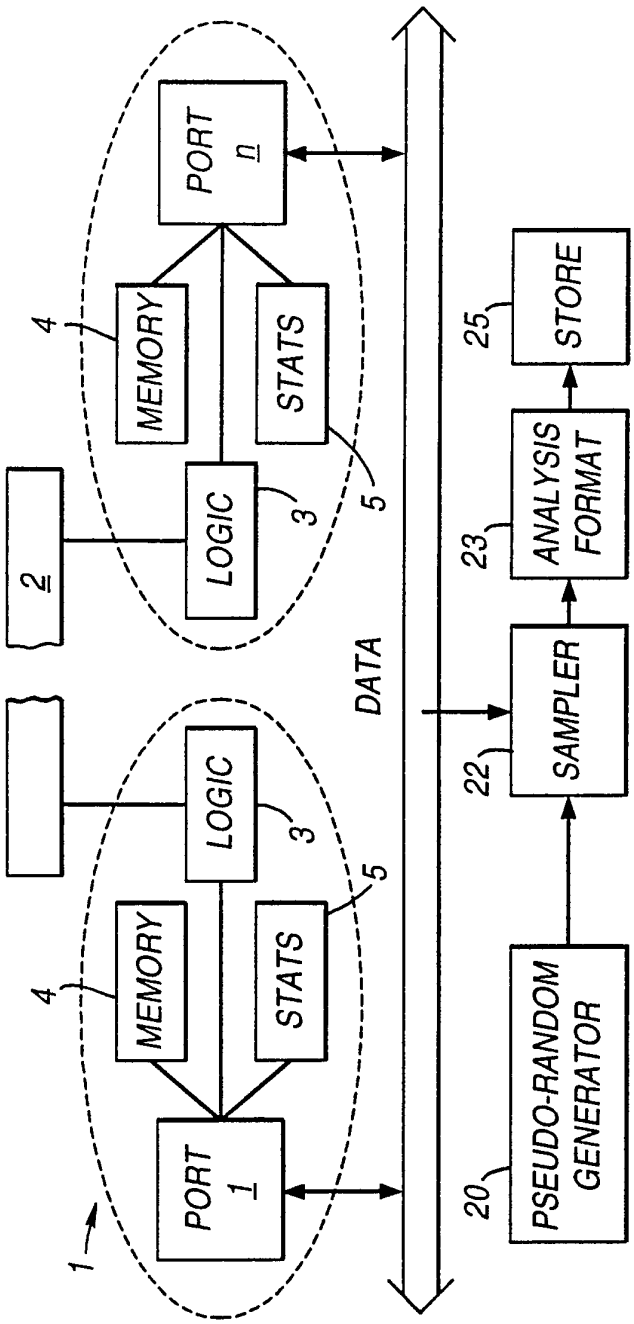

PORT 1
PORT n
MEMORY
STATS
LOGIC
DATA
SAMPLER
ANALYSIS FORMAT
STORE
PSEUDO-RANDOM GENERATOR
1
2
3
4
5
20
22
23
25

Fig. 2