

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
1 September 2005 (01.09.2005)

PCT

(10) International Publication Number  
WO 2005/081474 A1

(51) International Patent Classification<sup>7</sup>: H04L 12/56,  
12/26, H04B 10/08

(21) International Application Number:  
PCT/SG2005/000056

(22) International Filing Date: 25 February 2005 (25.02.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/547,729 25 February 2004 (25.02.2004) US

(71) Applicant (for all designated States except US): NATIONAL UNIVERSITY OF SINGAPORE [SG/SG];  
10 Kent Ridge Crescent, Singapore 119260 (SG).

(72) Inventors; and

(75) Inventors/Applicants (for US only): PHUNG, Minh, Hoang [VN/SG]; Optical Networks Lab, Blk E4 #06-18, 4 Engineering Drive 3, Dept. of Electrical and Computer Engineering., National University of Singapore, Singapore 117576 (SG). CHUA, Kee, Chaing [MY/SG]; Dept. of Electrical and Computer Engineering, National University

of Singapore, 10 Kent Ridge Crescent, Singapore 119260 (SG). MOHAN, Gurusamy [IN/SG]; Dept. of Electrical and Computer Engineering, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260 (SG). MOTANI, Mehul [US/SG]; Dept. of Electrical and Computer Engineering, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260 (SG).

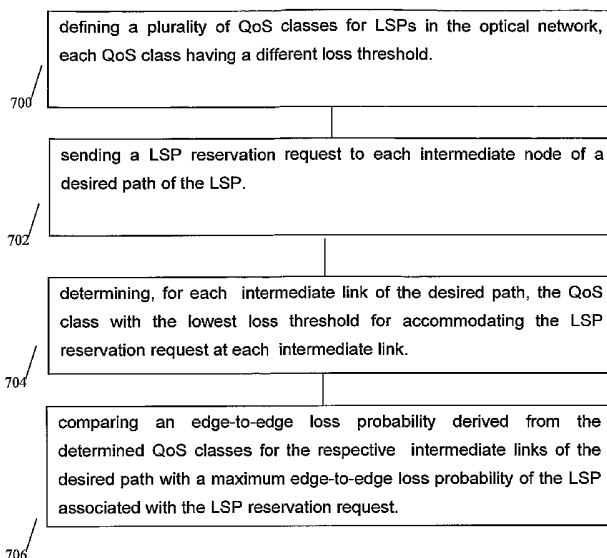
(74) Agent: ELLA CHEONG SPRUSON & FERGUSON (SINGAPORE) PTE LTD; P O Box 1531, Robinson Road Post Office, Singapore 903031 (SG).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, 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, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH,

[Continued on next page]

(54) Title: QOS IN OPTICAL NETWORKS



(57) Abstract: A method of determining whether a quality of service (QoS) request for a label switched path (LSP) in an optical network with label switching capability can be accommodated, the method comprising defining a plurality of QoS classes for LSPs in the optical network, each QoS class having a different loss threshold; sending a LSP reservation request to each intermediate node of a desired path of the LSP; determining, for each intermediate link of the desired path, the QoS class with the lowest loss threshold for accommodating the LSP reservation request at each intermediate link; and comparing an edge-to-edge loss probability derived from the determined QoS classes for the respective intermediate links of the desired path with a maximum edge-to-edge loss probability of the LSP associated with the LSP reservation request.

WO 2005/081474 A1



GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— *with international search 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.*

## QoS in Optical Networks

### FIELD OF INVENTION

The present invention relates broadly to a method of determining whether a quality of service (QoS) request for a label switched path (LSP) in an optical network with label switching capability can be accommodated, and to an optical network with label switching capability.

### BACKGROUND

In optical networks in which each wavelength is shared by different data streams, Optical Burst Switching (OBS) or Optical Packet Switching (OPS) is typically applied. Unlike for circuit switching, loss mechanisms must be considered in both OBS and OPS.

OBS has been proposed for the efficient transport of data network traffic over all-optical Wavelength Division Multiplexing (WDM) networks. OBS seeks to combine the best virtues of the circuit switching approach and the packet switching approach. It achieves high utilization through statistical multiplexing associated with the latter approach and the relative simplicity of the former approach. OBS is generally regarded as a promising solution for WDM networks in the medium-to-long term.

Another relevant trend is the appearance of more real-time multimedia applications connected to the Internet. Due to their stringent delay requirements, the impact of packet loss cannot be mitigated by retransmission. Therefore, these applications require certain quantitative guarantees on loss probability to operate properly. It is desirable that the next-generation Internet offers those guarantees.

The basic OPS and OBS paradigms offer only a best effort service with unbounded loss probability when the offered load is high. Since the OPS/OBS layers have better knowledge about the optical backbone network, supporting QoS at the

OPS/OBS layers would lead to improved network utilization, scalability and response time. It would be desirable if the OPS/OBS layers offer certain quantitative guarantees on edge-to-edge loss probabilities for bursts traversing the optical network.

To support QoS in e.g. OBS networks, various schemes have been developed. In general, they classify burst traffic into a limited number of classes and provide QoS differentiation among the traffic classes on a per-hop basis, but do not offer quantitative guarantees on any QoS metric. The differentiation method can be either isolation-based or proportional. The isolation-based differentiation assigns each traffic class a priority and tries to ensure that in a contention between two bursts of different priorities, the higher priority burst always wins. The proportional differentiation assigns a QoS factor to each traffic class. Each core node is required to keep a predefined QoS metric of each class proportional to its QoS factor. The QoS metric can be either loss probability or allocated bandwidth.

Such schemes do not provide quantitative QoS guarantees. Furthermore, e.g. in OBS networks the loss probability of different traffic flows depends heavily on their burst characteristics such as offset time and burst length distribution. Thus, although the overall loss probability of a class may be below a specified limit, some traffic flows of that class may experience loss probability that is well above the limit.

## **SUMMARY**

In accordance with a first aspect of the present invention there is provided a method of determining whether a quality of service (QoS) request for a label switched path (LSP) in an optical network with label switching capability can be accommodated, the method comprising defining a plurality of QoS classes for LSPs in the optical network, each QoS class having a different loss threshold; sending a LSP reservation request to each intermediate node of a desired path of the LSP; determining, for each intermediate link of the desired path, the QoS class with the lowest loss threshold for accommodating the LSP reservation request at each intermediate link; and comparing an edge-to-edge loss probability derived from the determined QoS classes for the respective intermediate links of the desired path

with a maximum edge-to-edge loss probability of the LSP associated with the LSP reservation request.

The method may further comprise keeping, for each intermediate link of the desired path, distances between the thresholds of all classes and respective measured loss probabilities of all classes at an output link of each intermediate node of the desired path non-negative.

The method may further comprise keeping, for each intermediate link of the desired path, respective measured loss probabilities of all LSPs within one QoS class similar.

A lowest loss threshold of the loss thresholds of the QoS classes may be based on a maximum number of intermediate nodes along possible paths in the network and a lowest required loss probability over the maximum number of intermediate nodes.

A highest loss threshold of the loss thresholds of the QoS classes may be based on a highest required loss probability over a one-hop possible path in the network.

Remaining ones of the loss thresholds of the QoS classes may be based on an even distribution between the lowest and the highest loss thresholds.

The method may further comprise dropping, when there is a contention between an incoming data packet/burst and existing data packets/bursts at an output link of an intermediate node in the optical network, one or more of the data packets/bursts belonging to the QoS class having the largest distance between said QoS class's threshold and measured loss probability at the output link, to resolve the contention.

Only data packets/bursts the single dropping of which would resolve the contention may be considered as candidates for dropping.

If more than one data packets/bursts belonging to the QoS class having the largest distance between said QoS class's threshold and measured loss probability at the output link, the dropping may be based on dropping probabilities calculated based on respective lengths of said data packets/bursts.

The dropping probability of a data packet/burst may be inversely proportional to the length of the data packet/burst.

The LSP reservation request comprises a requested amount of bandwidth at each intermediate link.

The method may further comprise calculating, at a link, a new weighted average threshold and a new overall loss probability of all classes if the requested bandwidth is to be accommodated and giving a rejection response if the new weighted average threshold is smaller than the new overall loss probability and giving a acceptance response otherwise.

The method may further comprise, if the LSP reservation request is to decrease the bandwidth of an existing LSP or to terminate the existing LSP, decreasing the data rate of said existing LSP or terminating the LSP before the LSP reservation request is sent to each intermediate node of the desired path.

The method may further comprise, upon receiving the LSP reservation request at an intermediate node to decrease the bandwidth of an existing LSP or to terminate the existing LSP, the intermediate node updates a label information database at the intermediate node based on information contained in the request.

The method may further comprise, if the LSP reservation request received at an intermediate node is to increase the bandwidth of an existing LSP, the intermediate node updates a label information database at the intermediate node and gives an acceptance response if the increased bandwidth can be accommodated in the QoS class registered for the existing LSP, or gives a rejection response otherwise.

The method may further comprise, if the LSP reservation request is to increase the bandwidth of the existing LSP, the data rate of the existing LSP is increased only after acceptance responses are received from all intermediate nodes of the desired path.

The LSP reservation request may be rejected if no QoS class for accommodating the LSP reservation request is determined at one of the intermediate links of the desired path.

Determining of the QoS class may comprise starting with the QoS class having the lowest threshold and continuing with other classes in increasing order of loss thresholds.

Information identifying the determined QoS classes at the respective intermediate links of the desired path may be sent to a common point of the desired path.

The method may further comprise determining whether the LSP reservation request can be accommodated over the entire desired path based on calculating a lowest edge-to-edge loss probability for the LSP reservation request based on the QoS classes received at the common point and giving a rejection response if the calculated loss probability is greater than the maximum edge-to-edge loss probability of the LSP associated with the LSP reservation request or giving an acceptance response otherwise.

The rejection or acceptance responses may be sent to an ingress node of the optical network.

If the request is to establish a new LSP, the new LSP may start only after the acceptance response is received.

The method may further comprise determining a remaining capacity of each link of the desired path after determining of the QoS classes; sending remaining capacity information to the common point together with the determined QoS classes

and utilizing the remaining capacity information in mapping the LSP's label to one of the QoS classes for each intermediate node of the desired path to facilitate maximizing network resource utilization.

The common point of the desired path may be the egress node.

The LSP reservation request may be sent to each intermediate node of the desired path along the desired path and information to be forwarded to the egress node is recorded in the LSP reservation request at each node of the desired path.

The QoS class mapping may be recorded in an acknowledgement message and the acknowledgement message may be sent back along the desired path in the opposite direction.

The method may further comprise regulating data rates of LSPs by marking data packets/bursts of an LSP that exceeds a submitted traffic profile of said LSP as out-of-profile and giving best effort service to out-of-profile packets/bursts inside the optical network.

In accordance with a second aspect of the present invention there is provided an optical network with label switching capability comprising means for defining a plurality of QoS classes for LSPs in the optical network, each QoS class having a different loss threshold; means for sending a LSP reservation request to each intermediate node of a desired path of the LSP; means for determining, for each intermediate link of the desired path, the QoS class with the lowest loss threshold for accommodating the LSP reservation request at each intermediate link; and means for comparing an edge-to-edge loss probability derived from the determined QoS classes for the respective intermediate links of the desired path with a maximum edge-to-edge loss probability of the LSP associated with the LSP reservation request.

In accordance with a third aspect of the present invention there is provided an optical network with label switching capability comprising a label information data base defining a plurality of QoS classes for LSPs in the optical network, each QoS



class having a different loss threshold; a transmitter sending a LSP reservation request to each intermediate node of a desired path of the LSP; and an admission control system determining, for each intermediate link of the desired path, the QoS class with the lowest loss threshold for accommodating the LSP reservation request at each intermediate link and comparing an edge-to-edge loss probability derived from the determined QoS classes for the respective intermediate links of the desired path with a maximum edge-to-edge loss probability of the LSP associated with the LSP reservation request.

## **BRIEF DESCRIPTION OF THE DRAWINGS**

Embodiments of the invention will be better understood and readily apparent to one of ordinary skill in the art from the following written description, by way of example only, and in conjunction with the drawings, in which:

Figure 1 is a schematic drawing illustrating an OPS/OBS network in an example embodiment.

Figure 2 is a schematic drawing illustrating an architecture of an OBS ingress node in an example embodiment.

Figure 3 is a schematic drawing illustrating an architecture of an OBS core node in an example embodiment.

Figure 4 is a schematic drawing illustrating an architecture of an OBS egress node in an example embodiment.

Figure 5 is a schematic drawing illustrating the construction of a contention list in an example embodiment.

Figure 6 is a schematic drawing illustrating a preemption scenario in an example embodiment.

Figure 7 is a graph showing overall loss probabilities of various traffic scenarios versus overall loading in an example embodiment.

Figure 8 shows a flowchart illustrating a method of determining whether a quality of service (QoS) request for a label switched path (LSP) in an optical network with label switching capability can be accommodated, in an example embodiment.

Figure 9 is a schematic drawing illustrating an architecture of an OPS ingress node in another example embodiment.

Figure 10 is a schematic drawing illustrating an architecture of an OPS core node in another example embodiment.

Figure 11 is a schematic drawing illustrating an architecture of an OPS egress node in another example embodiment.

## DETAILED DESCRIPTION

Figure 1 shows a general block diagram of an OPS/OBS network 1. The OPS/OBS network includes multiple ingress nodes 2 and multiple egress nodes 3 connected to a meshed network of multiple core nodes 4. The connections 5 between ingress nodes 2, egress nodes 3 and core nodes 4 are made using optical fibre links. Each optical fibre can carry multiple channels at different wavelength, which are used as either control channels or data channels. Each ingress node 2 is connected to multiple Internet Protocol (IP) channels 6 at its input port and each egress node 3 is connected to multiple IP channels 6 at its output port. The IP channels can be either electrical or optical.

Figure 2 shows a block diagram of an OBS ingress node 2 in an example embodiment. The ingress node 2 receives IP packet flows 200 via the multiple IP channels (6 in Figure 1). The IP packets carry routing and QoS information 202 in their packet headers.

The incoming IP packets 200 are directed by the packet classifier 7 to appropriate assembly queues in the burst assembler 8, which creates data bursts using one of a known burst assembly algorithms, e.g. one described in [Y. Xiong et al, "Control Architecture in Optical Burst-Switched WDM Networks," IEEE Journal on Selected Areas in Communications, vol. 18, no. 10, Oct. 2000, pages 1838-1851]. The burst assembler 8 outputs assembled data bursts to data channels 10. The burst assembler 8 also sends burst information to the header packet generator 9 for header packet generation. The corresponding header packets of the data bursts are sent to the control channels 11.

The ingress node 2 also receives QoS reservation information 204 for IP flows. The ingress node 2 includes a QoS reservation unit 12, which receives the reservation information 204, and a policing unit 13 for QoS management purposes. The QoS reservation unit 12 is responsible for reserving enough bandwidth at intermediate links to meet the QoS requirement of burst flows, based on the received reservation information 204. The QoS reservation unit 12 is coupled to the policing unit 13 and keeps the policing unit 13 updated on QoS reservation information 204. The policing unit 13 is responsible for keeping burst flows from exceeding their QoS reservation. For this purpose, the policing unit 13 is coupled to the header packet generator 9. The policing unit 13 will check each header packet generated by the header packet generator 9 and mark them as out-of-profile if they exceed the QoS reservation of their respective burst flow. Out-of-profile bursts receive best-effort treatment inside the network. QoS signalling messages, Resv and Ack, are sent by the QoS reservation unit 12 coupled to the control channels 11 respectively.

Figure 3 shows a block diagram of a OBS core node 4 in an example embodiment. Header packets and control packets from control channels 11 go through an optical/electrical (O/E) converter 14. After the O/E converter 14, QoS signalling messages, Resv and Ack, go through an admission control unit 15, which determines whether a QoS reservation request is admissible. The QoS signalling messages then go through an electrical/optical (E/O) converter 16 coupled to the control channels 11. Header packets 300, on the other hand, go from the O/E 14 to the router 17, which directs the header packets 300 to one of the schedulers 18 associated with each output port 302 of the core node 4 coupled to the data channels 10. The scheduler 18 on each output port 302 assigns a free channel to each incoming header packet. If a free channel cannot be found, a contention resolution unit 19 coupled to the scheduler 18 is called up to resolve the channel contention. Burst loss information is measured by the loss monitoring unit 20 coupled to the scheduler 18.

The admission control unit 15, the router 17, the contention resolution unit 19 and the loss monitoring unit 20 are coupled to and in close contact with a Label Information Base (LIB) 21, which acts as an information repository of the core node 4. The LIB 21 and the admission control unit 15 are bi-directionally coupled. The LIB 21 and the router 17 are bi-directionally coupled. The LIB 21 receives input from the lost monitoring unit

20, and provides output to the contention resolution unit 19. On the data plane (data channels 10), the optical switching matrix 22 coupled between the input and output data channels 10 is continuously configured by the scheduler 18 to switch bursts from input channels to appropriate output channels. The scheduler 18 is bi-directionally coupled to the contention resolution unit 19, and provides output to the lost monitoring unit 20 and the E/O converter 16. The scheduler 18 further provides output to the optical switching matrix 22.

Figure 4 shows the block diagram of an OBS egress node 3 in an example embodiment. The admission control and class mapping unit 24 receives Resv messages from the control channels 11, and processes the messages as described later and returns the result recorded on Ack messages back to the ingress node. The data bursts from the data channels 10 are converted into electronic form by the O/E converter 14, put through the burst disassembler 25. The resulting IP packets are then sent out to appropriate IP flows 200.

The QoS model in the example embodiment, assumes a label switching architecture such as Multi-Protocol Label Switching (MPLS) to be installed in the OBS network. Other label switching architectures that may be used in different embodiments include Generalized Multi-Protocol Label Switching (GMPLS) and Labeled Optical Burst Switching (LOBS). The architecture defines a number of Forward Equivalence Classes (FECs) that are unique across the whole network. At ingress nodes 2, bursts are classified into one of the FECs. Each burst header 300 carries a label to identify the FEC the header 300 belongs to. When a core node 4 receives a header packet 300, the label in the header is used to look up routing and QoS information associated with the label from the Label Information Base (LIB) 21, which has previously been downloaded to the node 2. A burst stream carrying the same label and following a particular path between an ingress node 2 and an egress node 3 constitutes a burst flow or a Label Switched Path (LSP). The required bandwidth of an LSP is the sum of the required bandwidths in the OBS of all the component IP flows and the required edge-to-edge loss probability of the LSP is equal to the most stringent required edge-to-edge loss probability of the IP flows.

The QoS model in the example embodiment defines a limited number of QoS classes and assigns each class a loss threshold. The threshold of a class is the upper bound on the loss probability that bursts belonging to that class will experience at a particular link 5. One or more labels are mapped by the Network to a QoS class for a particular link 5. The class mapped to a particular label may be different at different links 5. A core node 4 is only required to keep per-class information, which includes the predefined loss threshold, and the amount of admitted traffic measured based on a suitable traffic profile, e.g. average data rate and the current average loss probability. A quantity called the distance to threshold is defined as the difference between the predefined loss threshold and the measured loss probability.

The core node 4 in the example embodiment ensures that the loss probabilities of all QoS classes are below their respective thresholds at all time. In other words, the distances to thresholds of all QoS classes are required to be non-negative. The core node 4 does this using a differentiation mechanism and an admission control mechanism in the example embodiment.

For the example embodiment, a differentiation scheme assists the admission control routine by shifting burst loss from classes that are in danger of breaching their thresholds to classes that are not. This makes the task of the admission control routine easier because the routine now only has to keep the total offered load and therefore the total burst loss below a certain level.

Another requirement of the differentiation scheme in the example embodiment is to ensure traffic flows with different burst characteristics in the same class have the same loss probability. This requirement is particularly relevant to OBS networks. Burst characteristics such as offsets and burst length distribution have significant impacts on burst loss probability. Hence, without intervention from the differentiation scheme, some flows with unfavourable traffic characteristics may experience loss probabilities above the threshold although the overall loss probability of the class is still below the threshold.

The above requirements for the differentiation scheme in the example embodiment are summarised below.

1) Ensure that as the offered load to the core node increases, the distances to thresholds of all classes present at the node converge to zero, and that if the offered load increases further, all thresholds will be exceeded simultaneously. The meaning of “simultaneously” is related to the time scale at which loss probabilities are measured.

2) Ensure that bursts belonging to the same class experience the same loss probability at a particular core node regardless of their offsets and burst lengths.

For scalability reasons, the example embodiment uses class-level differentiation and only requires a core node to keep per-class information, which includes the predefined loss threshold, the amount of admitted traffic and the current average loss probability. The average loss probability is continuously updated using an exponentially weighted averaging algorithm.

In the example embodiment, when a burst header arrives at a node and fails to reserve an output channel/wavelength, the scheduler 18 (Figure 3) constructs a *contention list* that contains the incoming reservation and scheduled reservations that overlap (or contend) with the incoming one. A scheduled reservation on each wavelength is included if preemption of that reservation helps to schedule the new reservation. This is illustrated in Fig. 4 where only the ticked reservations 400 and 402 among the ones overlapping with the incoming reservation 404 on wavelength  $W_1$ ,  $W_2$  and  $W_3$  are included in the contention list. The contention resolution unit 19 (Figure 3) then selects one reservation from the list to drop according to some criteria described later. If the dropped reservation is a scheduled one then the incoming reservation will be scheduled in its place. In that case, the incoming reservation *preempts* the scheduled reservation.

When preemption happens, a special NOTIFY header packet 304 (Figure 3) will be immediately generated and sent on the control channel 11 (Figure 3) to the downstream nodes to inform them of the preemption. The downstream nodes then remove the reservation corresponding to the preempted burst. Although one NOTIFY packet is required for every preemption in the example embodiment, the rate of preemption is bounded by the loss rate, which is usually kept very small. Therefore, the additional overhead by the transmission of NOTIFY packets will not be significant.

There are two criteria for selecting a reservation from the contention list to drop in the example embodiment. The first criterion is that the selected reservation belongs to the class with the largest distance to threshold in the contention list. This criterion ensures that all the distances to thresholds of the classes present at the node are kept similar, thereby facilitating satisfying the first requirement above. The second criterion is applied when there are more than one reservation belonging to the class with the largest distance to threshold. In that case, only one of them is selected for dropping. Let the burst length of the  $i$ th reservation be  $l_i$ , ( $1 \leq i \leq N$ ), where  $N$  is the number of reservations belonging to the class with the largest distance to threshold in the contention list. The probability of it being dropped is

$$p_i^d = \frac{1/l_i}{\sum_{j=1}^N 1/l_j} \quad (1)$$

This is because the probability that a reservation is involved in a contention is roughly proportional to its length, assuming Poisson burst arrivals. So  $p_i^d$  is explicitly formulated to compensate for that burst length selection effect. In addition, the selection is independent of burst offsets. Therefore, it achieves the second requirement.

The following assumptions and simplifications are used in the analysis of the overall loss probability at an output link of a node in an example embodiment. Firstly, for the sake of tractability, only one QoS class is assumed to be present at a node. This assumption has negligible effect on the result. Secondly, burst arrivals follow a Poisson process with total rate  $\lambda$ . This is justified by the fact that a core node usually has a large number of traffic flows and the aggregation of a large number of independent and identically distributed point processes results in a Poisson point process. Thirdly, the incoming traffic consists of a number of traffic components with the  $i$ th component having constant burst length  $1/\mu_i$  and arrival rate  $\lambda_i$ . This assumption results from the fact that size-triggered burst assembly is a popular method to assemble bursts. This method produces burst lengths with very narrow dynamic range, which can be considered constant. Finally, we assume that no Fibre Delay Line (FDL) buffer is present and the offset difference among incoming bursts is minimal.

The lower bound on loss probability  $P_l$  is derived in the example embodiment by observing that preemption itself does not change the total number of lost bursts in the system. Thus, it is determined using the following Erlang's loss formula for an  $M|G|k|k$  queueing model:

$$P_l = B(k, \rho) = \frac{r^k / k!}{\sum_{i=0}^k r^i / i!} \quad (2)$$

where  $k$  is the number of wavelengths per output link;  $\rho$  is the total offered load and  $r = k\rho$ .

Although preemption does not directly affect the number of lost bursts, preemption affects the probability that a burst whose header arrives later is successfully scheduled. Depending on the reservation intervals of later bursts, the preemption may have detrimental or beneficial effects. Considering a preemption scenario as illustrated in Figure 6 where burst 501 is preempted by burst 502. Let bursts 503 and 504 be two bursts whose headers arrive after the preemption. For burst 503, the preemption is detrimental because had there been no preemption, burst 503 would be successfully scheduled. On the other hand, the preemption is beneficial to burst 504. However, for that to happen, burst 504 has to have a considerably shorter offset than other bursts, which is unlikely due to our assumption that the offset difference among bursts is minimal. For other preemption scenarios, it can also be demonstrated that a considerable offset difference is required for a preemption to have beneficial effects. Therefore, it can be argued that preemption generally worsens the schedulability of later bursts.

To quantify that effect, it is observed that from the perspective of burst 503, the preemption is equivalent to dropping burst 502 and extending the effective length of burst 501 as illustrated in Figure 6 at numeral 505. Therefore, the preemption increases the time that the system spends with all  $k$  wavelengths occupied. The upper bound on burst loss probability is derived by assuming that the loss probability is also increased by the same proportion. Let  $\delta = (l' - l)/l$  where  $l'$  is the new effective length and  $l$  is the actual length of the preempted burst. The upper bound on loss probability is then given as



$$P_u = \begin{cases} (1 + \delta\rho)B(k, \rho) & \text{if } B(k, \rho) < \frac{1}{1 + \delta\rho} \\ 1 & \text{otherwise} \end{cases} \quad (3)$$

An approximate formula for the loss probability can be derived based on equation (3) by observing that the increase in effective length of a preempted burst increases the overall loss probability only if another incoming burst contends with it again during the extended duration. The probability that this does not happen is

$$p = \sum_{i=0}^{\infty} \frac{e^{-\delta r} (\delta r)^i}{i!} \left( \frac{k}{k+1} \right)^i = e^{-\frac{\delta r}{k+1}} \quad (4)$$

From equations (3) and (4), the loss probability  $P$  is given as

$$P = P_u - e^{-\frac{\delta r}{k+1}} \delta\rho B(k, \rho) \quad (5)$$

We will now derive  $\delta$ . Suppose the incoming traffic has  $N_c$  traffic components with  $N_c$  different burst lengths. Let  $a$  and  $b$  denote the component indices of the incoming burst and the preempted burst, respectively. The probability of a particular combination  $(a, b)$  is given by the formula

$$P(a, b) = \frac{\lambda_a}{\sum_{i=1}^{N_c} \lambda_i} \times \frac{\rho_b}{\sum_{j=1}^{N_c} \rho_j} \times \frac{\mu_b}{\sum_{k=1}^{N_c} \mu_k} \quad (1 \leq a, b \leq N_c) \quad (6)$$

The first and second factors in equation (6) are the probabilities that an incoming burst and a scheduled burst belong to components  $a$  and  $b$ , respectively. The third factor accounts for the length selective mechanism of the preemption scheme. For a preemption situation  $(a, b)$ , the effective length is increased by  $\frac{1}{\mu_a} - \frac{1}{2\mu_b}$ . Therefore, it

follows that

$$\delta = \sum_{a=1}^{N_c} \sum_{b=1}^{N_c} P(a, b) \left( \frac{\mu_b}{\mu_a} - \frac{1}{2} \right) \quad (7)$$

While the above description assumes that no FDL buffer is present, the example embodiment can be readily extended to work with FDL buffers by repeating the preemption procedure for each FDL and the new reservation interval.

Since the distances to thresholds of the classes at a node are kept similar by the differentiation scheme in the example embodiment, the admission control routine only needs to keep the average of the distances to threshold greater than zero. In other words, the overall loss probability is kept smaller than the weighted average threshold. If there are  $M$  QoS classes at the node and  $T_i$  and  $B_i$  are the predefined threshold and the total reserved bandwidth of the  $i$ th class, respectively, the weighted average threshold is calculated as:

$$T = \frac{\sum_{i=1}^M T_i B_i}{\sum_{j=1}^M B_j} \quad (8)$$

The overall loss probability  $P$  can be calculated using equations (3) or (5). In a different embodiment, an empirical graph may be used.

Returning to Figure 1, edge-to-edge signalling and reservation mechanisms are responsible for coordinating the reservation setup and teardown for LSPs over the edge-to-edge paths in the example embodiment. During the reservation process of an LSP, the signalling mechanism polls all the intermediate core nodes 4 about remaining capacity on the intermediate links 5 and conveys the information to the egress node 3. Using that information as the input, the egress node produces a class allocation that maps the LSP to an appropriate class for each link 5 on the path. The signalling mechanism then distributes the class allocation to the core nodes 4. As a simple illustration, it is considered that an LSP with an edge-to-edge loss requirement of 5% needs to be established over a 4-hop path and the second hop is near congestion. The network may decide to allocate the LSP a threshold of 3.2% for the second hop and 0.4% for the other hops to reflect the fact that the second node is congested. The resulting guaranteed upper bound on edge-to-edge threshold would roughly be 4.4%, satisfying the LSP's requirement.

The QoS requirements of an LSP consists of its minimum required bandwidth and its maximum edge-to-edge loss probability. The loss probability requirement is usually fixed for a particular LSP. On the other hand, the required bandwidth usually fluctuates as new IP flows join the LSP or existing IP flows within the LSP terminate. The reservation scenarios for an LSP in the example embodiment can be categorised as follows:

1) A new LSP is to be established with a specified minimum bandwidth requirement and a maximum edge-to-edge loss probability. This happens when some IP flow requests arrive at the ingress node and cannot be fitted into any of the existing LSPs.

2) An existing LSP needs to increase its reserved bandwidth by a specified amount. This happens when some incoming IP flow requests have edge-to-edge loss requirements compatible with that of the LSP.

3) An existing LSP needs to decrease its reserved bandwidth by a specified amount. This happens when some existing IP flows within the LSP terminate.

4) An existing LSP terminates because all of its existing IP flows terminate.

The detailed reservation process for the first scenario is as follows in the example embodiment. The ingress node 2 sends a reservation message towards the egress node 3 over the path that the LSP will take. The message contains a requested bandwidth  $b_o$ , which is equal to the sum of the requested bandwidth of all the component IP flows, and a required edge-to-edge loss probability  $P_o$  equal to the most stringent loss probability of the IP flows. When a core node 4 receives the message, the admission control routine at that node 4 checks each class using methods described above to determine if the requested bandwidth can be accommodated in that class. The check starts from the lowest index class (corresponding to the lowest threshold) and moves up, and stops at the first satisfactory class. The node 4 then records the class index  $k$  in the message before passing the message on. It also locks in the requested bandwidth by setting the total reserved bandwidth  $B_k$  of class  $k$  as  $B_k(new) = B_k(old) + b_o$  so that the

LSP will not be affected by later reservation messages. On the other hand, if all the classes have been checked unsuccessfully, the request is rejected and an error message is sent back to the ingress node 2. Upon receiving the error message, the upstream nodes release the bandwidth locked up earlier.

The final admission control decision for the LSP is made at the egress node 3. The reservation message that reaches the egress 3 contains the class indices  $k$ 's for all the intermediate links. The egress node 3 calculates the lowest possible edge-to-edge loss probability  $P_{e2e}^0$  as follows.

$$P_{e2e}^0 = 1 - \prod_{i=1}^n (1 - p_i^0) \quad (9)$$

where  $p_i^0$  is the lowest threshold offered at the  $i$ th link on a  $n$ -link path.  $p_i^0$  is derived from the class index  $k(i)$  in the reservation message. That is,  $p_i^0$  is the threshold of the class with index  $k(i)$

If  $P_{e2e}^0 \leq P_0$ , the request is admitted. The egress node 3 then allocates each core node one of the predefined classes in which to support the LSP such that

$$\begin{cases} p_i \geq p_i^0 \\ P_{e2e} = 1 - \prod_{i=1}^n (1 - p_i) \leq P_0 \end{cases} \quad (11)$$

where  $p_i$  is the threshold of the class allocated to the LSP at the  $i$ th link.

This class allocation is signalled back to the core nodes 4 via the control channel 11 (Figure 2) using a returned acknowledgement message that contains the old indices  $k$ 's and another array of allocated class indices  $k_a$ 's. In the example embodiment,  $k_a(i) \geq k(i)$  since  $p_i \geq p_i^0$ . Upon receiving the acknowledgement message, a core node 4 moves the reserved bandwidth of the LSP from class  $k$  to class  $k_a$ . The new LSP is allowed to start only after the ingress node 2 has received the successful acknowledgement message.

If  $P_{e2e}^0 > P_0$ , the request is rejected and an error message is sent back to the ingress node 2. The intermediate core nodes will release the locked bandwidth upon receiving the error message.

The reservation process for the second scenario is relatively simpler in the example embodiment. In this case, the ingress node 2 sends out a reservation message containing the requested bandwidth  $b_0$  and the LSP's label. Since there is already a QoS class associated with the LSP at each of the intermediate link 5, a core node 4 on the path only needs to check if  $b_0$  can be supported in the registered class. When the request arrives, the admission control routine substitutes  $B_k$  with  $B'_k = B_k + b_0$  and recalculates the weighted average threshold  $T'$  and the overall loss probability  $P'$  as described above. If  $P' \leq T'$ , the request is admitted and the node locks in  $b_0$  and passes the reservation message on. Otherwise, an error message is sent back and the upstream nodes release the bandwidth locked previously. If the reservation message reaches the egress node 3, a successful acknowledgement message is returned to the ingress node 2 and the LSP is allowed to increase its operating bandwidth.

In the third and fourth scenarios, the reservation processes are similar in the example embodiment. The ingress node 2 sends out a message carrying the amount of bandwidth with a flag to indicate that it is to be released and the LSP's label. The released bandwidth is equal to the reserved bandwidth of the LSP if the LSP is to terminate. At intermediate core nodes 4, the total reserved bandwidth of the class associated with the LSP is decreased by that amount. No admission control check is necessary. Since the core nodes 4 do not keep track of bandwidth reservation by individual LSPs, the processing at core nodes 4 is identical for both the third and fourth scenarios. It should be noted that when an LSP terminates, there is a separate signalling process to remove the LSP's information from core nodes' LIBs.

QoS class definition is an important part of configuring the framework in the example embodiment. Usually, the number of classes  $M$ , which is directly related to the complexity of a core node's QoS differentiation block, is fixed. Hence, in this process, one decides on where to place the available thresholds, namely the lowest and highest loss thresholds  $T_l$  and  $T_h$  and those between them.

In an OBS network with a maximum network diameter of  $H$  hops, an LSP in the network can have a required edge-to-edge loss guarantee anywhere between  $P_l$  and  $P_h$  (not counting best-effort and out-of-profile traffic). One can see that the case requiring the lowest loss threshold  $T_l$  occurs when an LSP over the longest  $H$ -hop path requires the lowest edge-to-edge loss probability of  $P_l$ . Thus, one may set  $T_l \leq P_l / H$ . Similarly, the highest threshold is  $T_h = P_h$  for the case when a one-hop LSP requires the highest edge-to-edge loss probability of  $P_h$ .

Before considering how to place the remaining thresholds between  $T_l$  and  $T_h$  in an example embodiment, it should be noted that since the possible edge-to-edge loss probability  $P_o$  is continuous and the threshold values are discrete, the edge-to-edge loss bound  $P_{e2e}$  offered by the network will typically be more stringent than  $P_o$ . This "discretization error" reduces the maximum amount of traffic that can be admitted. The thresholds needs are preferably spaced so that this discretization error is minimised. Under the traffic model assumptions in the example embodiment, Figure 7 shows an approximately linear relationship (curve 600) between the overall loss probability in logarithmic scale and the total admitted traffic in linear scale. Therefore, in order to increase the maximum amount of admitted traffic, a simple solution can be to minimise the gaps between the thresholds in logarithmic scale. Assuming  $P_o$  is uniformly distributed on the logarithmic scale, the thresholds may be distributed evenly. That is, the thresholds are assigned the values  $T_l, \gamma T_l, \dots, \gamma^{M-1} T_l$  where  $\gamma = \sqrt[M]{T_h / T_l}$

Returning to Figure 1, when a new LSP is being established, the network allocates a QoS class to the LSP at each of the links 5 on the LSP's path. In the example embodiment, the LSP is allocated a class with a high threshold at a heavily loaded link and with a low threshold at a lightly loaded link. With bottleneck links required to support a less demanding QoS level, more traffic can be admitted onto the path. This class allocation is carried out at the egress node 3 based on the information recorded in the reservation message by the core nodes.

A simple class allocation scheme in an example embodiment is to use the class index array  $k$  recorded in the reservation message as indicators of the utilisation levels at the links 5. In this scheme, the egress node 3 repeatedly increments all the elements of  $k$  until the guaranteed edge-to-edge loss probability  $P_{e2e}$  is greater than  $P_o$ . At that

point, the scheme backs off and decrements the indices one at a time starting from the lowest one until  $P_{e2e} \leq P_0$ .

To illustrate the above scheme, an OBS network with 64 wavelengths per link and 8 predefined QoS classes with indices  $\{0, 1, \dots, 7\}$  is considered. The lowest threshold is  $T_l = 0.05\%$  and the ratio between two adjacent threshold is  $\gamma = 2$ . An LSP with an edge-to-edge loss requirement of 1% is to be set up over a three-hop path. The utilisation levels at the intermediate links on the path are  $\{0.3, 0.6, 0.35\}$ . The required bandwidth of the LSP is assumed to be very small compared to the link capacity. From Figure 7, one can see that all the links can accommodate the LSP in class 0 or  $k = \{0,0,0\}$ . By following the above scheme, the allocated classes will be  $k_a = \{2,3,3\}$  corresponding to thresholds of  $\{0.2\%, 0.4\%, 0.4\%\}$ . However, this allocation does not reflect the utilisation levels at the links 5 since the bottleneck link is assigned the same class index as other links. The reason is because the minimum index that can be recorded in  $k$  is zero. Therefore, utilisation levels lower than what corresponds to  $T_l$  may not be distinguished.

To reduce the above effect, in another example embodiment a core node 4 may record another parameter  $\kappa$  to the reservation message in addition to  $k$ . Upon receiving a reservation message, as described above, the node 4 pretends to put the request in a class, starting from class 0 and estimates the new overall loss probability  $P'$  and weighted average threshold  $T'$ . If  $P' \leq T'$  at class 0,  $\kappa$  is calculated as

$$\kappa = \left\lceil \log_{\gamma} \left( \frac{T'}{P'} \right) \right\rceil$$

$\kappa$  is recorded in the reservation message along with  $k = 0$ . Otherwise, the node 4 keeps on checking higher indexed classes until just before  $P' > T'$  or all the classes have been checked. If the request is admitted, the selected class index  $k$  is recorded along with  $\kappa = 0$ . The parameter  $\kappa$  acts as the distance between  $T'$  and  $P'$  in logarithmic scale.

In a pseudo-code in Algorithm 1 given below,  $k$ ,  $\kappa$ ,  $k_a$  represent whole arrays. The egress node 3 uses  $k - \kappa$  as indicators of the utilisation levels at links 5 in the class allocation process. In the first two lines, the algorithm sets  $k_a$  such that the maximum element is  $M - 1$  and the differences among the elements are the same as in the array  $k$

-  $\kappa$ . Next, it repeatedly decrements all the elements of  $k_a$  until  $P_{e2e} \leq P_0$ . Finally, the elements of  $k_a$  are incremented one by one until just before  $P_{e2e} > P_0$  in order to push  $P_{e2e}$  as close as possible to  $P_0$  without exceeding it.

**Algorithm 1:** Class allocation algorithm

$$d_{max} = \text{MAX}(k - \kappa)$$

$$k_a = (k - \kappa) - d_{max} + M - 1$$

**while**  $P_{e2e} > P_0$

$$k_a = k_a - 1$$

Increment elements of  $k_a$  one by one until just before  $P_{e2e} > P_0$

Applying the scheme of algorithm 1 to the previous example,  $k = \{0,0,0\}$ . If  $\kappa = \{50,2,35\}$ , going through algorithm 1,  $k_a = \{-41,7,-26\}$  on line (3) and  $k_a = \{-44,4,-29\}$  on line (5). The final result is  $k_a = \{1,4,1\}$  corresponding to thresholds of  $\{0.1\%, 0.8\%, 0.1\%\}$ . This shows that the scheme in this example embodiment successfully allocates the maximum possible class index to the bottleneck link.

When there is only one bottleneck link, the class allocation scheme in this example embodiment treats the non-bottleneck links essentially the same regardless of their utilisation levels. This is because even a small change in the total utilisation level causes a large change in the overall loss probability at a link as can be observed from Figure 7. Therefore, provided that the utilisation levels at non-bottleneck links are not very close to that at the bottleneck link, loss probabilities at non-bottleneck links may be so small compared to that of the bottleneck link that any differentiated treatment is inconsequential.

Figure 8 shows a flowchart illustrating a method of determining whether a quality of service (QoS) request for a label switched path (LSP) in an optical network with label switching capability can be accommodated, in an example embodiment. At step 700, a plurality of QoS classes are defined for LSPs in the optical network, each QoS class having a different loss threshold. At step 702, a LSP reservation request is sent to each intermediate node of a desired path of the LSP. At step 704, for each intermediate link of the desired path, the QoS class with the lowest loss threshold for accommodating the LSP reservation request at each intermediate link



is determined. At step 706, an edge-to-edge loss probability derived from the determined QoS classes for the respective intermediate links of the desired path is compared with a maximum edge-to-edge loss probability of the LSP associated with the LSP reservation request.

It will be appreciated by a person skilled in the art that numerous variations and/or modifications may be made to the present invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects to be illustrative and not restrictive.

For example, while the embodiment described with reference to Figures 2 to 8 has been with reference to an OBS network, it will be appreciated that the present invention is applicable to other optical networks with label switching capabilities. For example, Figure 9 shows a schematic drawing of an OPS network ingress node 900, in which corresponding components when compared with the ingress node 2 described above with reference to Figure 2 have been indicated with the same reference numeral. The ingress node 900 does not include a burst assembler (compare 8 in Figure 2). Rather, the header package generator 9 generates the header for each IP packet of an IP packet flow 200, which is added to the individual IP packets as indicated at 902 before transmission of the optical packets on the data channels 10.

Similarly, Figure 10 shows an OPS network core node 1000 in which the same components when compared with the core node 4 described above with reference to Figure 3 have been indicated with the same reference numeral. At the core node 1000, headers of the IP packets received on the data channels 10 are extracted as indicated at numeral 1002, and are provided to the router 17 via the O/E converter 14. After the assigning of free channels including contention resolution, if required, as described above with reference to Figure 3, the headers are re-combined with the respective IP packets as indicated at 1004, via E/O converter 16. Suitable fiber delay lines (FDLs) 23 are provided to accommodate the required processing time.

Finally, Figure 11 shows a schematic drawing of an OPS network egress node 1100, in which the same components when compared with the egress node 3 described above with reference to Figure 5 have been indicated with the same reference numeral. In this embodiment, the egress node 1100 comprises an O/E converter 14 for sending out the IP packets received from the data channels 10 to appropriate IP flows 200.

**CLAIMS**

1. A method of determining whether a quality of service (QoS) request for a label switched path (LSP) in an optical network with label switching capability can be accommodated, the method comprising:

defining a plurality of QoS classes for LSPs in the optical network, each QoS class having a different loss threshold;

sending a LSP reservation request to each intermediate node of a desired path of the LSP;

determining, for each intermediate link of the desired path, the QoS class with the lowest loss threshold for accommodating the LSP reservation request at each intermediate link; and

comparing an edge-to-edge loss probability derived from the determined QoS classes for the respective intermediate links of the desired path with a maximum edge-to-edge loss probability of the LSP associated with the LSP reservation request.

2. The method as claimed in claim 1, further comprising keeping, for each intermediate link of the desired path, distances between the thresholds of all classes and respective measured loss probabilities of all classes at an output link of each intermediate node of the desired path non-negative.

3. The method as claimed in claim 2, further comprising keeping, for each intermediate link of the desired path, respective measured loss probabilities of all LSPs within one QoS class similar.

4. The method as claimed in any one of the preceding claims, wherein a lowest loss threshold of the loss thresholds of the QoS classes is based on a maximum number of intermediate nodes along possible paths in the network and a lowest required loss probability over the maximum number of intermediate nodes.

5. The method as claimed in any one of the preceding claims, wherein a highest loss threshold of the loss thresholds of the QoS classes is based on a highest required loss probability over a one-hop possible path in the network.

6. The method as claimed in claims 4 and 5, wherein remaining ones of the loss thresholds of the QoS classes are based on an even distribution between the lowest and the highest loss thresholds.

7. The method as claimed in any of the preceding claims, further comprising dropping, when there is a contention between an incoming data packet/burst and existing data packets/bursts at an output link of an intermediate node in the optical network, one or more of the data packets/bursts belonging to the QoS class having the largest distance between said QoS class's threshold and measured loss probability at the output link, to resolve the contention.

8. The method as claimed in claim 7, wherein only data packets/bursts the single dropping of which would resolve the contention are considered as candidates for dropping.

9. The method as claimed in claims 7 or 8, wherein, if more than one data packets/bursts belonging to the QoS class having the largest distance between said QoS class's threshold and measured loss probability at the output link, the dropping is based on dropping probabilities calculated based on respective lengths of said data packets/bursts.

10. The method as claimed in claim 9, wherein the dropping probability of a data packet/burst is inversely proportional to the length of the data packet/burst.

11. The method as claimed in any of the preceding claims, wherein the LSP reservation request comprises a requested amount of bandwidth at each intermediate link.

12. The method as claimed in claim 11, further comprising calculating, at a link, a new weighted average threshold and a new overall loss probability of all classes if the requested bandwidth is to be accommodated and giving a rejection

response if the new weighted average threshold is smaller than the new overall loss probability and giving a acceptance response otherwise.

13. The method as claimed in claims 11 or 12, further comprising, if the LSP reservation request is to decrease the bandwidth of an existing LSP or to terminate the existing LSP, decreasing the data rate of said existing LSP or terminating the LSP before the LSP reservation request is sent to each intermediate node of the desired path.

14. The method as claimed in claim 13, further comprising, upon receiving the LSP reservation request at an intermediate node to decrease the bandwidth of an existing LSP or to terminate the existing LSP, the intermediate node updates a label information database at the intermediate node based on information contained in the request.

15. The method as claimed in claims 11 or 12, further comprising, if the LSP reservation request received at an intermediate node is to increase the bandwidth of an existing LSP, the intermediate node updates a label information database at the intermediate node and gives an acceptance response if the increased bandwidth can be accommodated in the QoS class registered for the existing LSP, or gives a rejection response otherwise.

16. The method as claimed in claim 15, wherein, if the LSP reservation request is to increase the bandwidth of the existing LSP, the data rate of the existing LSP is increases only after acceptance responses are received from all intermediate nodes of the desired path.

17. The method as claimed in any one of the preceding claims, wherein the LSP reservation request is rejected if no QoS class for accommodating the LSP reservation request is determined at one of the intermediate links of the desired path.

18. The method as claimed in any one of the preceding claims, wherein determining of the QoS class comprises starting with the QoS class having the

lowest threshold and continuing with other classes in increasing order of loss thresholds.

19. The method as claimed in any one of the preceding claims, wherein information identifying the determined QoS classes at the respective intermediate links of the desired path is sent to a common point of the desired path.

20. The method as claimed in claim 19, further comprising determining whether the LSP reservation request can be accommodated over the entire desired path based on calculating a lowest edge-to-edge loss probability for the LSP reservation request based on the QoS classes received at the common point and giving a rejection response if the calculated loss probability is greater than the maximum edge-to-edge loss probability of the LSP associated with the LSP reservation request or giving an acceptance response otherwise.

21. The method as claimed in claim 20, wherein the rejection or acceptance responses are sent to an ingress node of the optical network.

22. The method as claimed in claim 21, wherein, if the request is to establish a new LSP, the new LSP starts only after the acceptance response is received.

23. The method as claimed in any of claims 20 to 22, further comprising determining a remaining capacity of each link of the desired path after determining of the QoS classes; sending remaining capacity information to the common point together with the determined QoS classes and utilizing the remaining capacity information in mapping the LSP's label to one of the QoS classes for each intermediate node of the desired path to facilitate maximizing network resource utilization.

24. The method as claimed in any of the claims from 20 to 23, wherein the common point of the desired path is the egress node.

25. The method as claimed in claim 24, wherein the LSP reservation request is sent to each intermediate node of the desired path along the desired path and information to be forwarded to the egress node is recorded in the LSP reservation request at each node of the desired path.

26. The method as claimed in claim 25, wherein the QoS class mapping is recorded in an acknowledgement message and the acknowledgement message is sent back along the desired path in the opposite direction.

27. The method as claimed in any one of the preceding claims, further comprising regulating data rates of LSPs by marking data packets/bursts of an LSP that exceeds a submitted traffic profile of said LSP as out-of-profile and giving best effort service to out-of-profile packets/bursts inside the optical network.

28. An optical network with label switching capability comprising:  
means for defining a plurality of QoS classes for LSPs in the optical network, each QoS class having a different loss threshold;  
means for sending a LSP reservation request to each intermediate node of a desired path of the LSP;  
means for determining, for each intermediate link of the desired path, the QoS class with the lowest loss threshold for accommodating the LSP reservation request at each intermediate link; and  
means for comparing an edge-to-edge loss probability derived from the determined QoS classes for the respective intermediate links of the desired path with a maximum edge-to-edge loss probability of the LSP associated with the LSP reservation request.

29. An optical network with label switching capability comprising:  
a label information data base defining a plurality of QoS classes for LSPs in the optical network, each QoS class having a different loss threshold;  
a transmitter sending a LSP reservation request to each intermediate node of a desired path of the LSP; and  
an admission control system determining, for each intermediate link of the desired path, the QoS class with the lowest loss threshold for accommodating the

LSP reservation request at each intermediate link and comparing an edge-to-edge loss probability derived from the determined QoS classes for the respective intermediate links of the desired path with a maximum edge-to-edge loss probability of the LSP associated with the LSP reservation request.



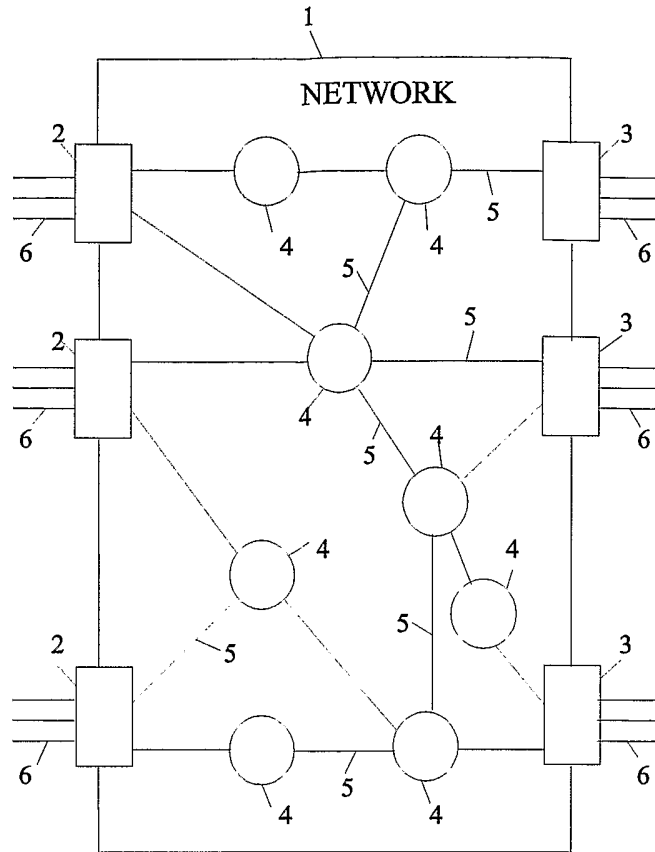


Figure 1

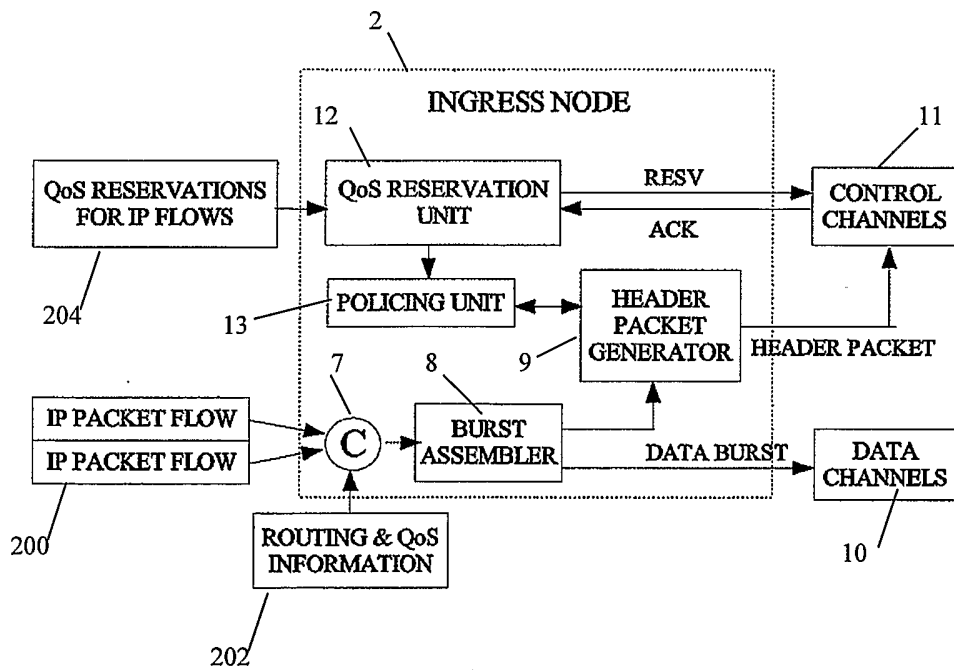


Figure 2

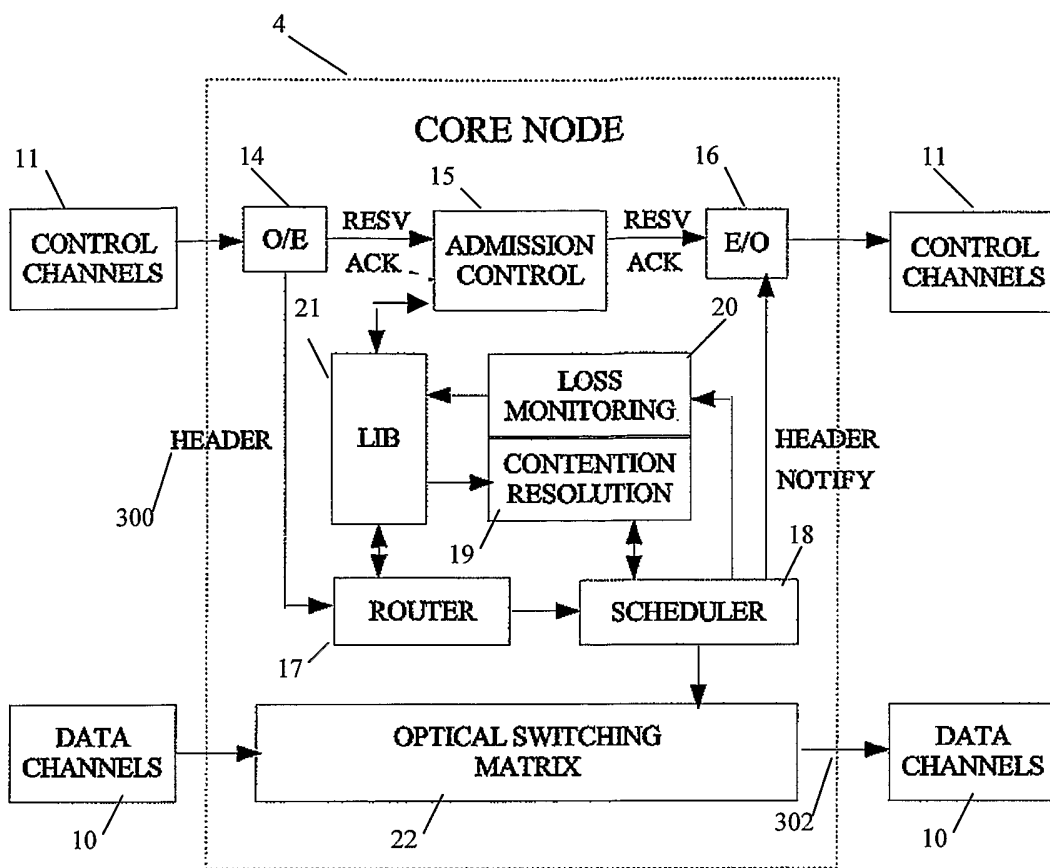


Figure 3

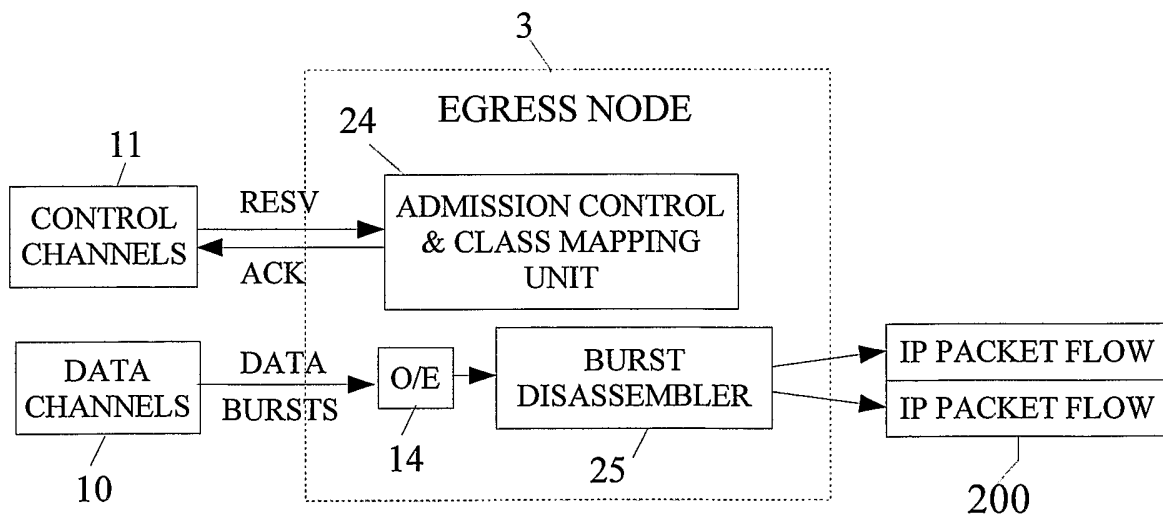


Figure 4

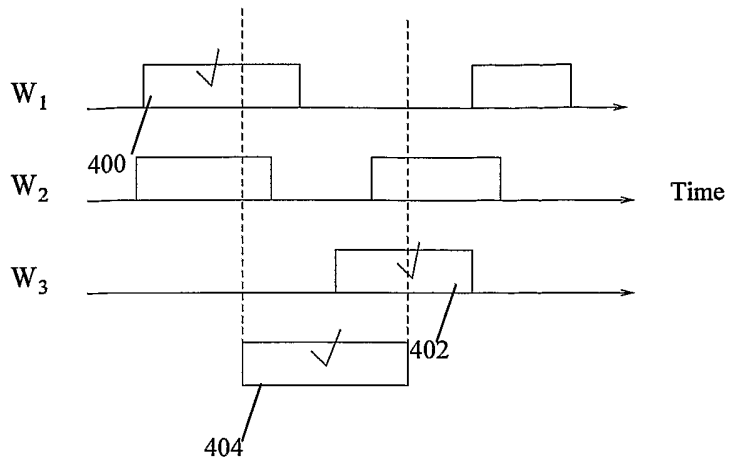


Figure 5

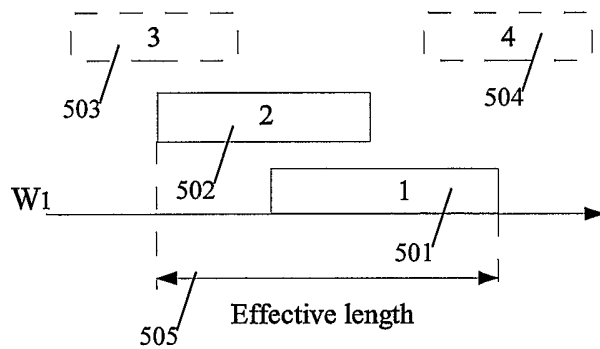


Figure 6

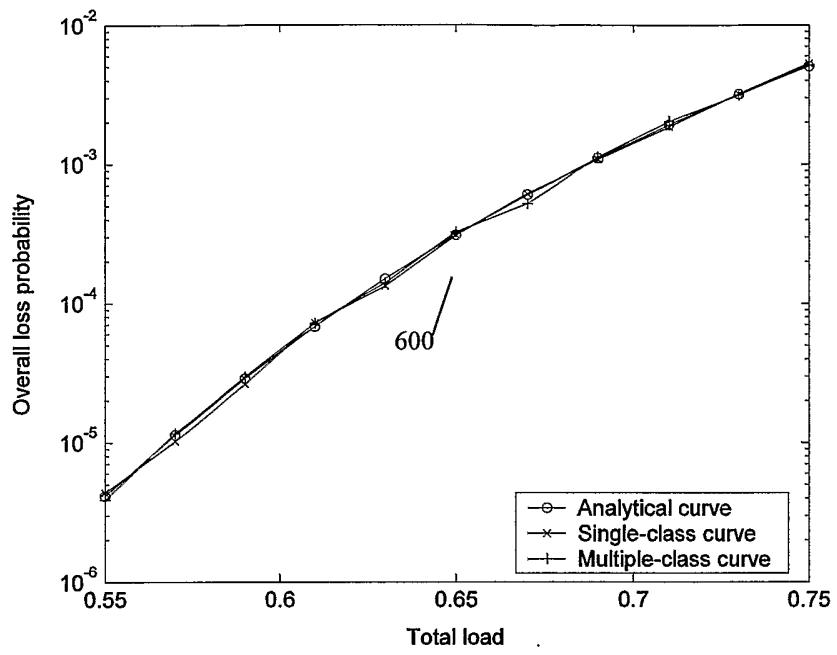


Figure 7

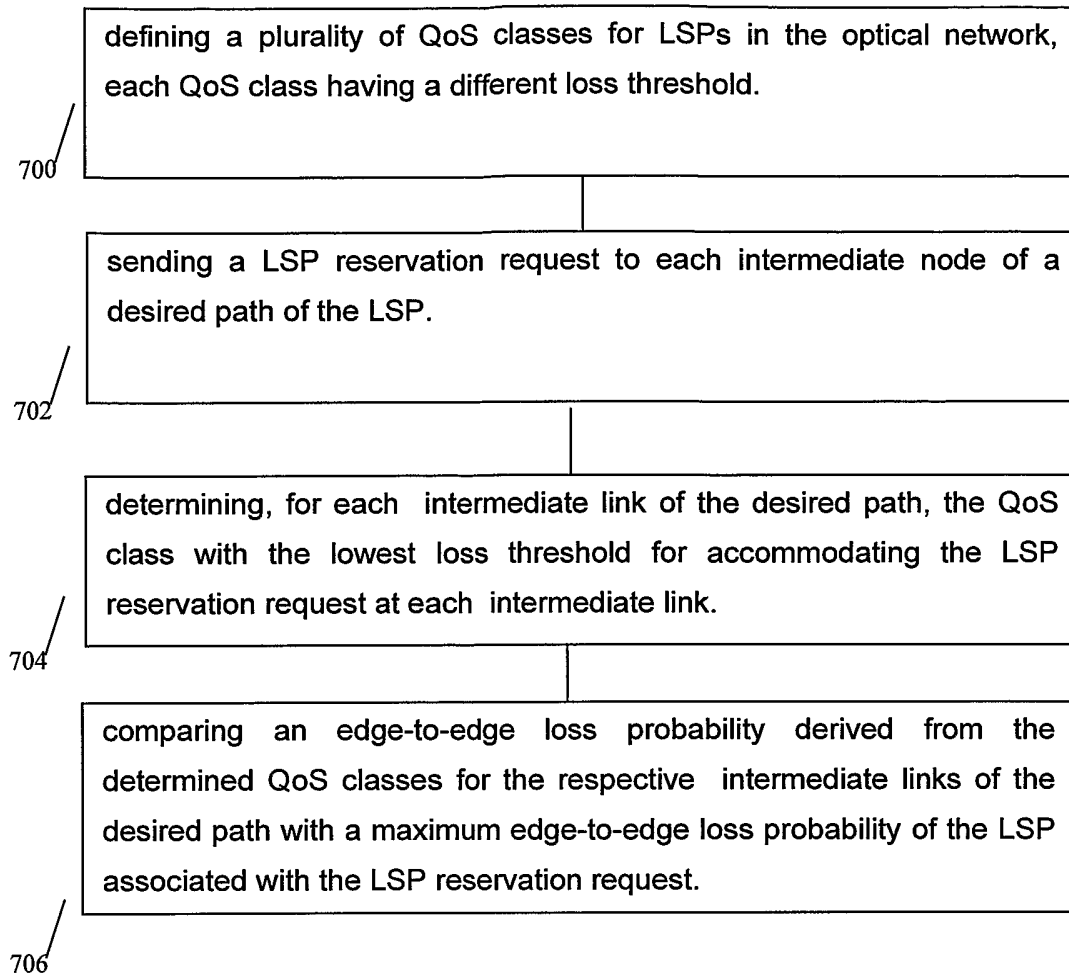


Figure 8

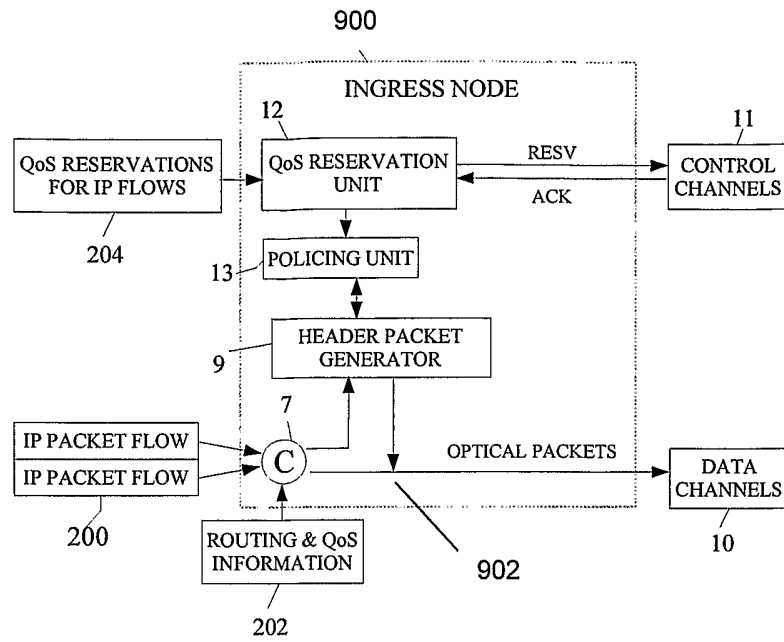


Figure 9

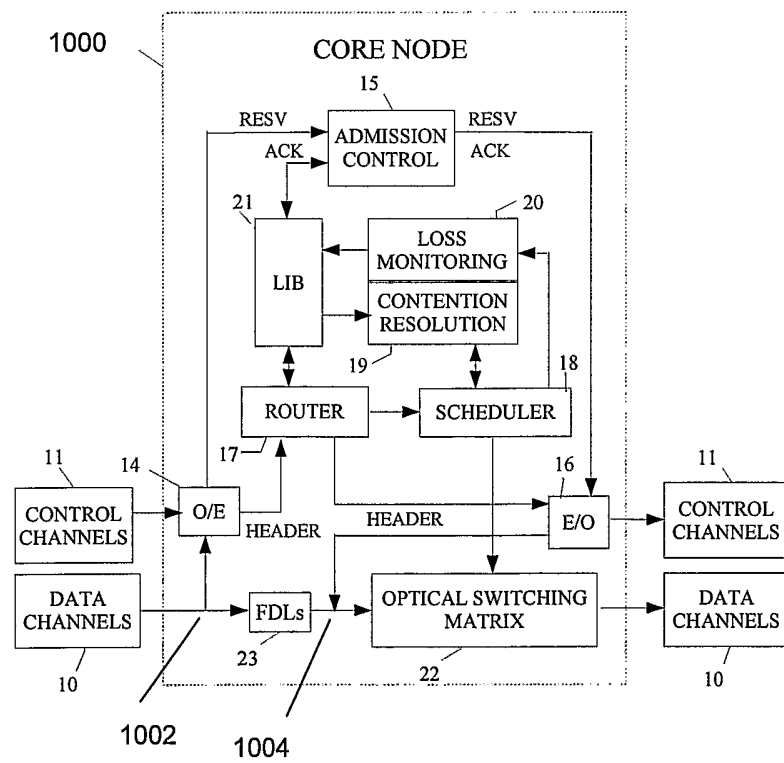


Figure 10

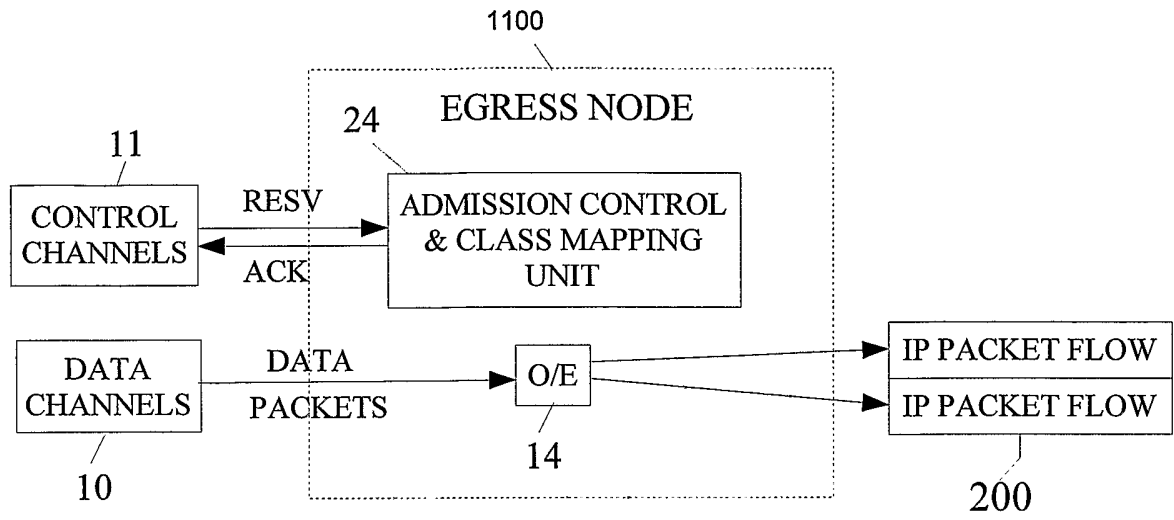


Figure 11

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/SG2005/000056

A. CLASSIFICATION OF SUBJECT MATTER		
Int. Cl. <sup>7</sup> : H04L 12/56, 12/26; H04B 10/08		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols)		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
DWPI & keywords: Optic, Label, Quality, Class and similar terms		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DEKERIS, B. and NARBUTAITE, L. Congestion Control Mechanism within MPLS Networks, Scientific Proceedings of RTU. Series 7. Telecommunications and Electronics, 2003, vol. 3 pages 43-46 (online), [retrieved on 5 May 2005]. Retrieved from the Internet <URL: <a href="http://www.rst.rtu.lv/Latvieshu%20lapa/publikacijas/sejums_3_pdf/09_Dekeris_Narbutaite_Congestion%20Control%20Mechanism.pdf">http://www.rst.rtu.lv/Latvieshu%20lapa/publikacijas/sejums_3_pdf/09_Dekeris_Narbutaite_Congestion%20Control%20Mechanism.pdf</a> >	1-29
P, A	US 6 744 767 B1 (CHIU <i>et al</i> ) 1 June 2004 See the whole document	1-29
A	US 2002/0097463 A1 (SAUNDERS <i>et al</i> ) 25 July 2002 See the whole document	1-29
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C <input checked="" type="checkbox"/> See patent family annex		
* Special categories of cited documents:		
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	
"E" earlier application or patent but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family	
"P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 6 May 2005	Date of mailing of the international search report 11 MAY 2005	
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaaustralia.gov.au Facsimile No. (02) 6285 3929	Authorized officer  <b>EVAN SULCS</b> Telephone No : (02) 6283 2442	



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/SG2005/000056

C (Continuation).		DOCUMENTS CONSIDERED TO BE RELEVANT
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2003/0206521 A1 (QIAO) 6 November 2003 See the whole document	1-29
A	CA 2 307 734 A1 (NEWBRIDGE NETWORKS CORPORATION) 27 November 2000 See the whole document	1-29

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

**PCT/SG2005/000056**

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report	Patent Family Member
US 6744767	
US 2002097463	CA 2363402
US 2003206521	
CA 2307734	US 6724881

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

END OF ANNEX