Performance Evaluation of the Spatial Reuse Protocol fairness algorithm (SRP-fa) used in DPT Networks

Julio Moyano (1), Salvatore Spadaro (1), Bruno Bostica (2), Josep Solé Pareta (1)

(1) Universitat Politècnica de Catalunya - UPC,
Barcelona, Catalunya (Spain)
Email: jmoyano@retemail.es, {sspadaro, pareta}@ac.upc.es
(2) Telecom Italia Lab - TILAB,
Torino, Italy
Email: bruno.bostica@tilab.com

ABSTRACT

The Dynamic Packet Transport (DPT) is introduced by CISCO as a metropolitan area network solution providing an innovative and optimized network architecture for ring-based delivery of IP services. The DPT functionality is enabled by a new Media Access Control (MAC) layer protocol called Spatial Reuse Protocol (SRP). This paper describes the simulation model and presents essential results from the performance evaluation of the Spatial Reuse Protocol fairness algorithm (SRP-fa). The simulations were carried out with Opnet (Optimized Network Engineering Tool) and the main objective was to verify and validate the features and benefits inherited from the SRP-fa, particularly concerning to *the efficient use of bandwidth*, *the fairness among nodes accessing to the ring,* and *the support for priority traffic*. This study has been developed within LION Project (IST-1999-11387).

I. INTRODUCTION

In an IP over DWDM scenario, one of the main topics to be solved is how the IP layer can be best transported directly over a DWDM network. As it is not possible to send IP datagrams directly over a physical medium because IP does not provide neither bit synchronization nor packet delineation, different mapping/framing solutions have been proposed in order to encapsulate and adapt IP packets to the DWDM layer in a cost-efficient manner, i.e. trying to bypass intermediate layers such as ATM and SDH [1].

The Dynamic Packet Transport (DPT) is introduced by Cisco as a metropolitan area network solution providing an innovative and optimized network architecture for ring-based delivery of IP services.

The DPT functionality is enabled by a new Media Access Control (MAC) layer protocol called Spatial

Reuse Protocol (SRP) operating over a dual-ring network topology.

DPT ring is composed of nodes that are interconnected by two counter-rotating rings that are referred to as "outer" ring and "inner" ring. They both transport concurrently data and control packets and while data information is sent in one direction, called 'downstream', the corresponding control messages are sent in the opposite direction, called 'upstream' (see Figure 1) [1], [2].

The ***spatial reuse*** concept refers to the fact that packets (unicast) only circulate along spans between the source and the destination node rather than the whole ring as in other protocols such as FDDI and Token Ring. This operation, in which destination nodes remove the packets from the ring, is known as ***destination stripping***. Thanks to the spatial reuse (or destination stripping), the bandwidth of the ring is used in a more efficient manner since bandwidth is only used on traverse segments allowing nodes to transmit concurrently and, thus, to improve the network rate. Network throughput can exceed 1, although utilization of any part of the ring channel itself is, of course, less than 1. In addition, the use of the bandwidth is dynamic; there are no bandwidth reservation or provisioned connections.

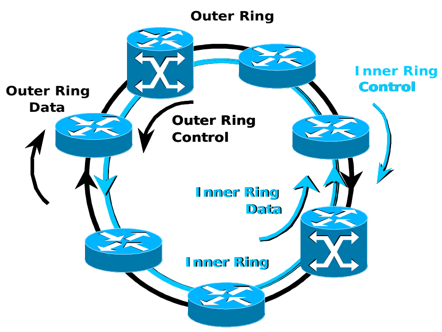


Fig. 1: Dynamic Packet Transport ring

The Spatial Reuse Protocol is a new MAC layer protocol for ring configurations and takes its name from the ***spatial reuse*** concept [3].

SRP is layer 1 (media) independent and can be used over a variety of underlying technologies such as SONET/SDH, WDM, dark fiber, or mixed environments. SRP

provides redundancy and protection in the event of a failed node or fiber cut (Intelligent Protection Switching) and

offers an efficient use of bandwidth (Spatial Reuse – Destination Stripping).

Moreover SRP is fairness among nodes accessing to the ring (SRP fairness algorithm) and is scalable across a large number of nodes. Finally, it offers support for priority traffic and multicasting.

This paper focuses on simulation results obtained in order to verify and validate the features and benefits inherited from the SRP-fa that the system DPT announces. The evaluation was done in terms of *Network Throughput* and *Delay End-to-End*. The simulation environment used for this work was based on OPNET (Optimized Network Engineering Tool).

II. SPATIAL REUSE PROTOCOL

Two protocols or algorithms lay the foundations of the SRP protocol: the ***SRP fairness algorithm*** (SRP-fa) and the ***Intelligent Protection Switching protocol*** (IPS). The former controls the access to the shared media ensuring fairness, bounding latency and avoiding privileged nodes or conditions while undertaking to prevent congestion, and, the latter consists of a protection scheme of the dual-ring.

Figure 2 illustrates the SRP packet processing (reception and transmission) that is operated at each DPT interface. The two levels of priority are managed both in the *transit buffer* (insertion buffer) and in the *transmit buffer* by means of two separate FIFO queues (high and low priority). The packet processing consists of three functions named *Reception*, *Insertion Register* and

Transmission.

Reception: When a packet arrives to the node, it is checked its source and destination address (address lookup) and its mode (type of packet) since packets can either be data or control packets. Control packets are always stripped whereas data packets are stripped if the destination MAC address corresponds to the node MAC address and the packet is unicast.

Insertion Register (Transit Buffer): the transit buffer together with the SRP fairness algorithm control the access to the ring. This technique makes use of an insertion or transit buffer, placed in each interface unit in which the nodes can store packets sent by other nodes while transiting through the ring [4].

Transmission: A transmission algorithm is needed since the packets can be sent from either the transit buffer (data generated by other nodes) or the transmit buffer (data generated by the own node) and both handling different priorities (high or low).

•

The transmission algorithm is the following:

1. High priority packets from the transit buffer are always sent first.
2. High priority packets from the transmit buffer are sent as long as the low priority transit buffer is not full.
3. Low priority packets from the transmit buffer are sent as long as the low priority transit buffer has not crossed a threshold indicating this situation and the SRP-fa rules allow it.
4. If nothing else can be sent, low priority packets from the low priority transit buffer are sent.

It is worth noting that the SRP-fa contributes to the transmission algorithm controlling the access to the ring (step 3 of the transmission algorithm).

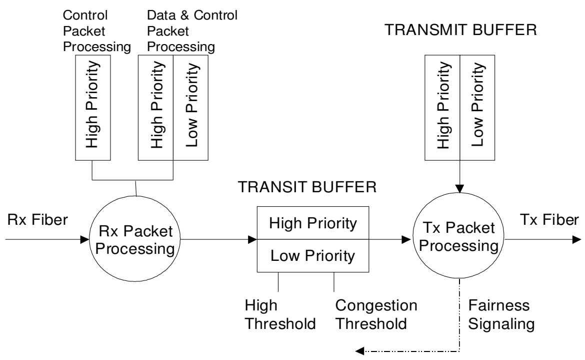


Fig. 2: DPT Interface

A. SRP-fa algorithm

The SRP-fa is a distributed algorithm that takes charge of ensuring *Global fairness* (each node gets a fair share of the ring bandwidth) and *Local optimization* (each node maximally leverages the spatial reuse properties of the ring).

The SRP-fa, together with the *Insertion Register* mechanism, controls access to the ring. The algorithm only applies to low priority traffic. High priority traffic does not follow any SRP-fa rules. In addition, the SRP-fa is a preventive control of congestion since nodes, periodically, generate *Usage* messages to propagate bandwidth information to upstream nodes in order to adapt their rates. These *Usage* messages also perform a keepalive function. The *Insertion Register* mechanism offers access that is totally distributed throughout the ring because each node is responsible for access control.

The algorithm works as follows:

a set of usage counters monitors the rate at which low priority transmit data and forwarded data are sent. These counters are MY_USAGE (tracks the amount of traffic sourced on the ring), FWD_USAGE (tracks the amount of traffic forwarded on the ring), and ALLOWED_USAGE (indicates the current maximum transmit usage for that node). The ALLOWED_USAGE is built up periodically by all nodes during no congestion situation. Each node has a parameter, MAX_USAGE, which limits the maximum amount of low priority traffic that the node can send. The ALLOWED_USAGE value must not exceed this maximum.

As exposed when defining the transmission algorithm, SRP-fa rules (step 3 of the algorithm) also control the access to the ring. This rule allows the low priority transmit buffer to transmit if MY_USAGE is less than ALLOWED_USAGE. When this condition fails, the low priority transit buffer takes control of transmission instead of the low priority transmit buffer even if the high threshold has not been crossed yet.

There is a congestion threshold in the low priority transit buffer used to detect congestion. If the congestion threshold is crossed (congestion detected), the node begins to tell to upstream nodes the value of its transmit usage counter (MY_USAGE). In order to send this bandwidth information, *Usage* control packets are used. These usage messages are generated periodically even if there is no new bandwidth information to send (a null value of all ones is sent) since these packets also inform to the destination node that a valid data link exists (keepalive function).

Nodes receiving usage messages adjust their transmit rates not to exceed the told value, and propagate the usage messages received whenever the node is not congested. If the node receiving the usage information is also congested, it propagates the minimum value of its transmit usage and the usage message received (see Figure 3).

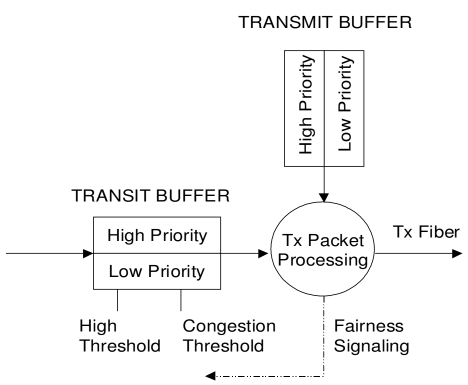


Fig. 3: SRP-fa procedure

III. SIMULATION ENVIRONMENT

The objective of these simulation cases studies is to verify and validate the features and benefits inherited from the SRP-fa that the system DPT announces. The DPT capabilities given by the fairness algorithm are *Efficient use of bandwidth*, *Support for priority traffic*, *Fairness among nodes using the ring*, *Scalability across a large number of nodes attached to a ring*. The simulation cases studies particularly focuses on them.

For this purpose, the OPNET network simulation tool has been used, which offers a powerful graphical environment for creating, executing and analyzing event-driven simulations of communication networks. This software package has many applications, typically communication protocol development and testing, analysis of protocol interactions and network optimization.

OPNET is based on a series of hierarchically related editors that directly parallel the structure of actual networks. Models specifications performed in the Project Editor rely on elements specified in the Node Editor and, when working with the Node Editor, the developer makes use of models defined in the Process Editor. The remaining editors are used to define various data models such as links and packet formats.

The three main OPNET's editors are *Project Editor, Node Editor*, *Process Editor. Project Editor* develops network models. Network models are made up of subnets and node models. This editor also includes basic simulation and analysis capabilities. The *Node Editor* develops node models. Node models are objects in a network model. Node models are made up of modules with process models. Modules may also include external system models and parameters. The *Process Editor* develops process models. Process models control module behavior and may reference parameter models.

A. Simulation parameters

Three main parameters have been taken into account when defining the simulations to be carried out in order to evaluate the SRP-fa in different situations or conditions: the traffic source, the traffic load distribution (traffic matrix) and the number of nodes which conforms the ring (topology).

As the DPT system is a data-oriented solution, a traffic source modeling burstiness has been created (On-Off traffic source). A Poisson traffic source has also been used to compare results.

Regarding to the traffic load distribution, two situations have been considered. The first case proposes a traffic matrix in which the traffic generated by each node is uniformly distributed to all the remaining nodes. The second one propounds an unbalanced traffic destination distribution where all nodes but one communicate with a node selected as a sink. In both cases, the main objective is to evaluate how the nodes access to the ring and how the bandwidth is used and distributed.

The payload of the SRP packets is supposed to be an IP datagram and, therefore, a real statistic of the packet size distribution of Internet has been chosen as reference for the size of the packets generated during the simulation [5]. Figure 4 illustrates the probability distribution made up with the OPNET's PDF Editor to apply as packet size distribution in the model.

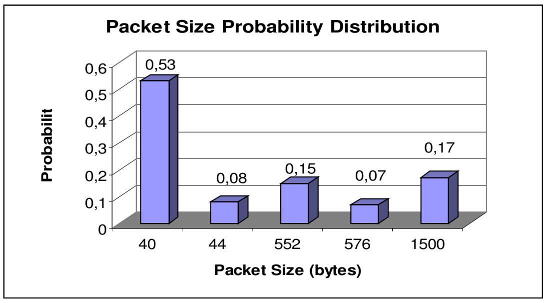


Fig. 4: Packet Size Distribution

IV. EXPERIMENTS AND RESULTS

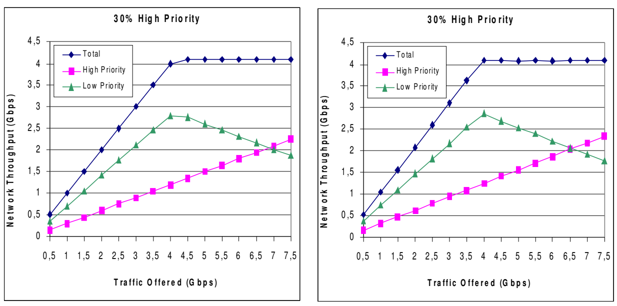
The dual counter-rotating ring of the DPT system poses two different situations depending on the presence of either an odd or an even number of nodes. In fact, in the first case the network is symmetrical in terms of network load (independent of the traffic load distribution used) due to the balanced position of the nodes with regard to the one that generates the traffic. In the second case (even number of nodes) the network is *unbalanced* since each node can choose either the inner or the outer ring to communicate with the farthest one. This election may cause either the outer or the inner ring supporting more traffic than the other does leading to an *unbalanced* situation.

Fig. 5: DPT ring topologies

A. Simulation Environment

A set of simulations with both DPT ring topologies have been carried out with the following conditions:

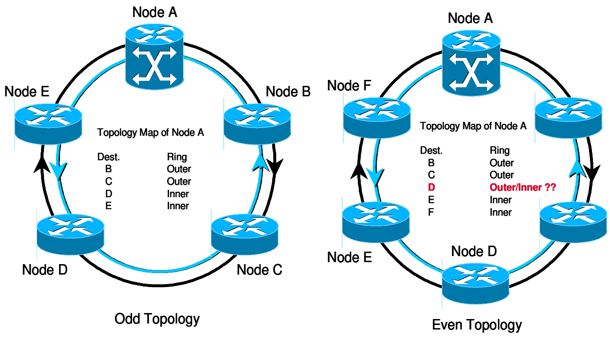
- Traffic Load Distributions: two different traffic distributions have been used. A uniform traffic



  load distribution and the unbalanced traffic load distribution, where all nodes but one communicate with a selected node.
- *On-Off traffic parameters:* The burstiness (peak

rate/average rate) has been set to b = 10, and the mean burst length has been set to BL = 10 packets.

- *Packet size:* A real statistic of the packet size distribution of Internet has been chosen as



reference for the size of the IP packets (payload of the SRP packets) generated during the simulation [5].

- *Percentage of priority traffic:* This parameter indicates the percentage of high priority traffic offered onto the ring whereas the rest is low priority traffic. The simulations have been carried out with percentages of 30% for the high priority traffic.
- *Transmit buffer configuration:* The transmit buffer is composed by two infinite FIFO queues (corresponding to high and low priorities).
- *Transit buffer configuration:* Like the transmit buffer, the transit buffer is split into two FIFO queues but, in this case, with a finite size of 512 KB. The thresholds are placed at 192 KB for the low threshold (also called congestion threshold), and 384 KB for the high threshold. Both thresholds are only placed at the low priority queue.
- *Ring speed:* The channel bit rate implemented in the DPT ring is an OC-12 (622 Mbps). The DPT ring is assumed error-free.

B. Experiment I: Odd Topology

In the first experiment, a DPT network composed by 5 nodes (odd number of nodes) has been created.

The focus of the experiment has been to determine how the nodes access to the ring, whether this access is fair among nodes or not, and to check the performance of the SRP-fa in terms of packet delay since it has been assumed an infinite size for the transmit buffers and there is no packet loss.

SRP uses the SRP-fa together with the *Register Insertion* (Insertion Buffer) mechanism to control access to the ring opening the possibility of using parts of the ring for two or more simultaneous communication paths (spatial reuse concept). When routing allows this type of use, network throughput can exceed 1, although utilization of any part of the ring channel itself is strictly less than 1.

Fig. 6: Bandwidth Efficiency, Poisson vs On-Off

The results in Figure 6 show the network throughput versus the traffic load offered to the ring. Distinction between high and low priorities has also been made. Before reaching high load traffic conditions, as offered load increases, so does network throughput. Since the network is assumed error-free and no packets are lost (no retransmissions), network throughput and offered traffic are equal until saturation occurs (4.15 Gbps). These results show that, unlike other different shared media access algorithms, the SRP-fa due to the spatial reuse concept (destination stripping) allows a very efficient control access and bandwidth use. In extreme load traffic conditions (traffic offered > 4.15 Gbps), the network is not able to support all the traffic offered and the low priority queues of the transmit buffers begin growing indefinitely. Low priority traffic throughput decreases permitting high priority traffic to be transparent to network saturation as long as, obviously, low priority traffic does not arrive to zero. Consequently, a good service of high priority traffic is assured.

The performance of the DPT ring is also evaluated in terms of packet delay in order to compare the SRP ring (DPT ring) with other ring solutions. It can see that the DPT ring is able to support priority traffic and, thus, to offer services with soft QoS requirements. The average delay value obtained for the high priority traffic, even in heaviest load conditions is acceptable.

The Figure 7 shows a comparison between Poisson and On-Off traffic sources in terms of average packet delay. In both pictures, the On-Off curves show the worst performance. Low priority traffic degrades in presence of high load conditions; nevertheless, the values obtained for high priority traffic are acceptable even for the On-Off curve.

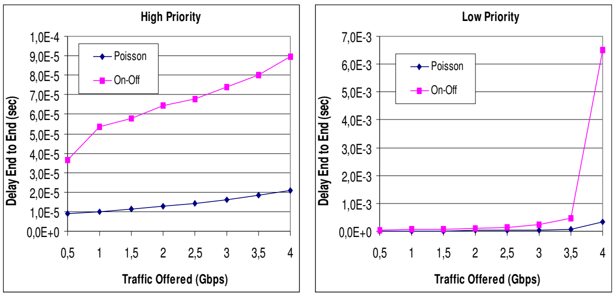
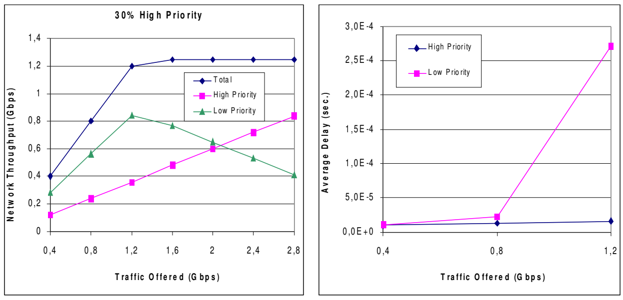


Fig. 7: Delay End-to-End, Poisson vs On-Off

Since now it has been shown how the SRP-fa provides access to the ring in a very efficient manner, but it is necessary to verify this controls access mechanism fair. Figure 8 attempts to verify the fairness use of the ring. In this case, both outer and inner output ports of node A have been used to confirm the fairness examining the source address of each outgoing packet. For instance, all packets traveling through the segment from node A to node E (i.e. sent by the inner output port of node A) correspond to communications from node A to E, from A to D and from B to E. A fair use of this segment (A to E path) should reveal an equal use of the bandwidth among the three communications mentioned.

The results show the percentage of packets circulating through the segment connecting node A with E (called Bandwidth Usage Inner) according to the originator node (source address). Figure 8 illustrates a fairness use of bandwidth even in extreme load conditions (traffic offered equal to 4.15 Gbps), since the percentages obtained are approximately 67% of packets originated by node A and 33% created by node E. Thus, the three different communications sharing this ring segment (communications from node A to E, from A to D and from B to E) have a fairly percentage of the bandwidth.



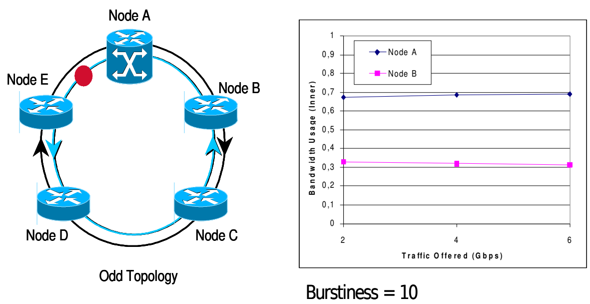


Fig. 8: Fairness (On-Off)

Subsequently we considered the unbalanced traffic load distribution (all nodes against one). Figure 9 shows the virtual network created by this traffic load distribution. Node A has been chosen as a sink without generating any kind of traffic whereas the rest only send packets to node A.

In this situation, it was interesting to test that nodes farthest to node A (nodes C and D) have the same percentage of bandwidth use than the nearest ones (nodes B and E). Figure 9 shows the percentage of packets received in node A through its outer segment classified in line with their source address. Fairness is assured, even in highest load conditions (offered traffic greater than 1.2 Gbps), since node A receives the same amount of traffic from both node D and node E. The worst case shows approximately a percentage of 54% (node E) and 46% (node D). Regarding to the fairness there is no difference in using Poisson or On-Off traffic sources.

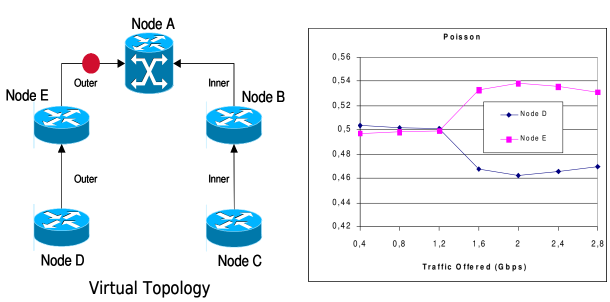


Fig. 9: Fairness – All vs One (Poisson)

Unlike the uniform traffic distribution, the network saturation only occurs when node A cannot accept more traffic. Because both outer and inner interfaces are OC-12 (622 Mbps), the congestion condition takes place at 1.2 Gbps (2*OC-12). The figure 10 illustrates this circumstance and how the low priority throughput decreases when the saturation condition is surpassed. Like the preceding studied cases, high priority traffic seems to be transparent to this circumstance.

The average delay curves show how the low priority delay suddenly increases close to the network saturation point. However the high priority delay is not affected by network congestion assuring, again, a good treatment of high priority even under severe conditions.

Fig. 10: Bandwidth Efficiency

C. Experiment II: Even Topology

In this case an even topology situation has been created using a DPT network constituted by 6 nodes (Figure 5). The DPT ring has been simulated under a uniform traffic distribution illustrating the *balanced* condition mentioned before. The simulation conditions used were the same described in subsection IV.A.

We considered the following approach. There is not only one specific ring selected to communicate with the farthest node. Half of the nodes use the outer ring and the other half the inner ring. Figure 11 shows how the network congestion coincides with the congestion, at the same time, of both outer and inner ring because of both rings sustain the same amount of traffic. The high and low priority curves behave as expected, assuring high performance for the high priority traffic to the detriment of low priority traffic.

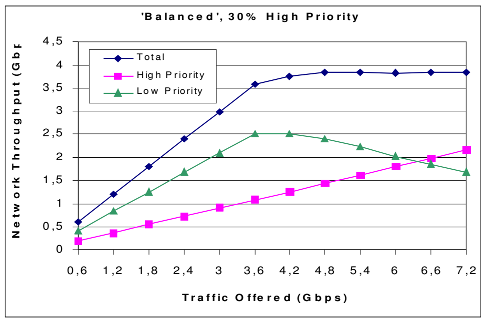


Fig. 11: Bandwidth Efficiency

Fairness is assured, even in highest load conditions, since both outer and inner part divide the bandwidth fairly (Figure 12).

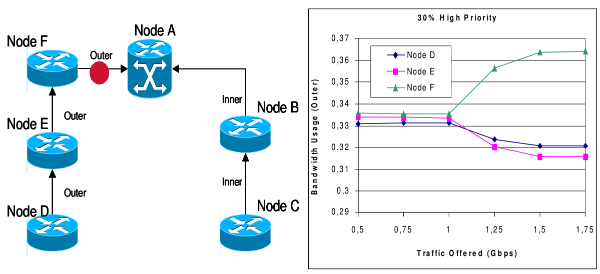


Fig. 12: Fairness (Poisson)

V. CONCLUSIONS

Dynamic Packet Transport (DPT) can be viewed as a ***data optimized*** solution with respect to SONET/SDH rings due to the limitations of SONET/SDH when carrying data traffic (IP traffic). DPT is cost efficient since eliminates the SONET/SDH equipment (ADMs) and uses the bandwidth efficiently. Finally, DPT introduces very interesting features for being used in LAN and MAN rings.

The results obtained have validated the following features related to the SRP-fa: *Efficient use of bandwidth*, *Fairness among nodes using the ring*, *Support for priority traffic*.

The SRP-fa together with the *Register Insertion* mechanism (spatial reuse concept) has appeared as a very efficient control access policy. All cases have always revealed a maximum network throughput greater than 1 (efficient use of bandwidth). Furthermore, this access technique has shown a fair share of ring bandwidth since neither starvation nor excessive delay conditions have been created within the DPT ring. Finally, it has been shown that a good service differentiation between low and high priority can be obtained. A good performance of high priority traffic is assured even under high traffic load conditions.

Currently the simulator is being extended in order to include the IPS (Intelligent Protection Switching) protocol in order to compare the DPT ring capabilities on resilience with the well-known capabilities of SDH.

VI. ACKNOWLEDGEMENTS

This work has been partially funded by the European Commission under the Information Society Technology (IST) program, (LION project, IST-1999-11387), and by CICYT (Spanish Ministry of Education) under contract TIC99-0572-C02-02.

REFERENCES

[1] Cisco Systems, "Dynamic Packet Transport Technology and Performance", White Paper.

[2] Cisco Systems, "Dynamic Packet Transport Technology and Applications Overview". http://www.cisco.com/warp/public/cc/cisco/mkt/servprod/opt/dpt/dpta_wp.htm

[3] D. Tsiang and G. Suwala, "The Cisco SRP MAC Layer Protocol", Request For Comments 2892, August 2000.

[4] Guy Pujolle et al., "Integrated Digital Communications Networks (VOL 2)", JOHN WILEY & SONS, 1988.

[5] Paul Lagner, "SDL Data Link Specification", Lucent Technologies – White Paper.