Chapter 5

Multi-ring architecture

5.1 State-of-the-art

The multi-ring architecture consists of a number of unidirectional slotted WDM rings of metropolitan dimensions, which collect traffic from several ring nodes. The WDM rings are interconnected to other rings via the Hub, and to a packet switch in the core. The rings can be either physically disjoint, or be obtained by partitioning the optical bandwidth into disjoint portions. The use of a Hub node that is in control of the resources makes the multi-ring different from other optical ring networks like e.g., the HORNET (and without any limiting relation between node counts and the number of wavelength). The hub node is used to forward optical packets between ring networks, as well as to interconnect the metro area to the backbone through an electronic Gateway. The Hub is an SOA-based optical packet switch capable to cope with a very high level of traffic (Terabit/s). The lack of real optical memories is compensated through the use of an extended MAC protocol. The optical Hub is configured by a controller which exploits the control channels of each connected ring network, in order to calculate the switching permutation. The details of the technological implementation of the Hub architecture are described in [43].

Nodes are composed of an electronic part and an optical part. The electronic part realizes the adaptation with client layers, and packet buffering (electronically). At the optical level, two node architectures are proposed in the DAVID project to propose a progressive introduction of optical packet technologies.

Targeting a short/medium term approach, a first proposal was made to limit the use of advanced optical technologies and use commercial and mature ones instead. Based on passive structures as described in [73], the architecture uses optical couplers and off-line optical filters to minimize physical issues when cascading the nodes (Figure 5.1a). This passive structure is lacking the packet drop stage, so that nodes keep all packets on the ring, simply copying packets addressed to them. We call this solution *passive multi-ring* (PMR). To allow simultaneous add and drop operation within the same slot, up- and downstream channels are spectrally separated. Nodes access the ring using separate sets of wavelengths: W for transmission, and W for reception. Data are sent by nodes on transmission wavelengths and switched from

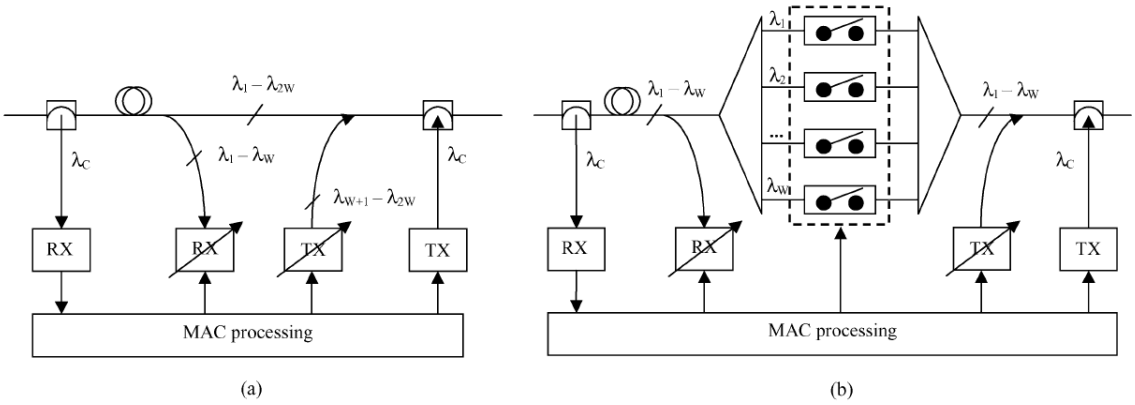

Figure 5.1: Architectures of a) PMR node with transmission and reception decoupling and b) MR node with erasure capability

transmission wavelengths to reception wavelengths at the Hub, which must provide wavelength conversion. Packets are received by nodes from reception wavelengths and are dropped when they reach the Hub, which, in turn, generates empty transmission channels for downstream nodes.

A second node structure is also considered as a longer-term approach allowing this erasing capability (Figure 5.1b) which allows packet removal at the destination, hence wavelength reuse (i.e., packets only circulate along ring spans between source and destination nodes). We call this solution *active multi-ring* or simply multi-ring (MR). In this case, nodes use the same set of W wavelengths for both transmission and reception.

In the following sections we only focus on the active multi-ring architecture. We use the passive multi-ring architecture as a reference in Section 5.5.4 and in the benchmarking analysis of Chapter 6.

5.1.1 MAC protocol

The rings are shared media, requiring a MAC protocol to arbitrate access to its slots, in order to regulate both time and wavelength dimensions. The overall system works as a combined wavelength/time/space distributed multiplexer. Contention and collision is avoided by an allocation algorithm and intelligent operation in the ring nodes using the control channel. Only the control channel is converted to the electrical domain for processing at each ring node, while the bulk of user information remains in the optical domain until its final destination in the end ring. The control slot information includes the state (empty or used) of the data slots, and the destination address of the corresponding data packets.

The inlet-outlet Hub allocation algorithm works as follows: a measurement cycle is defined (the length of which is denoted by F), during which the Hub monitors the use of the slots allocated to any ring pair. The monitoring of ring-to-ring traffic can be based either upon measurements of the load on the different rings at the Hub,

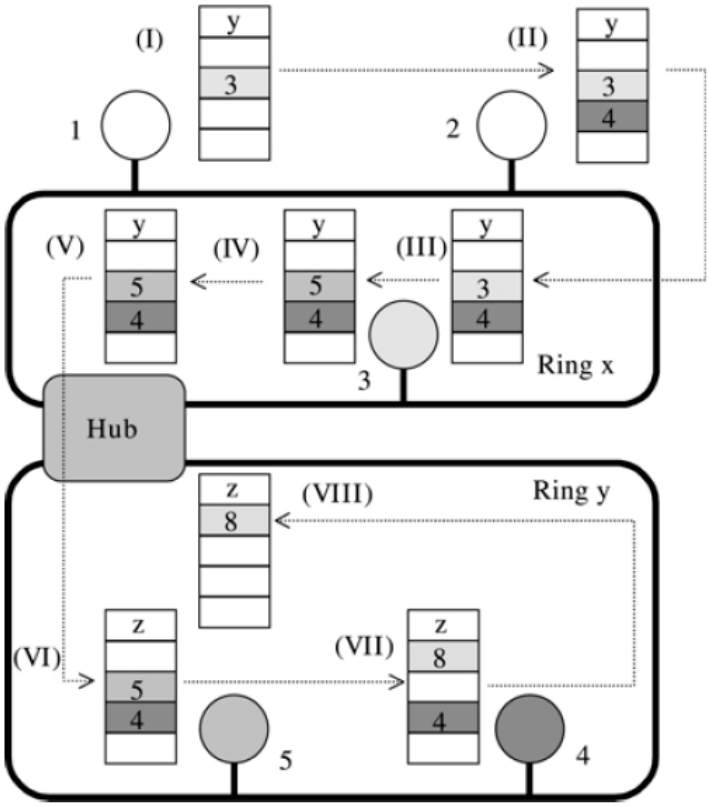

Figure 5.2: Multi-slot forwarding in the multi-ring. Colors in slot represent packet destinations

or upon explicit reservations issued by ring nodes. At the end of the measurement cycle, the Hub issues a new set of switching permutations, to be used for the coming measurement cycle. The Hub acts as a nonblocking switch that is reconfigured in every time slot and can exploit wavelength conversion to solve contention. In every time slot, the Hub operates a permutations from input rings to output rings. This permutation is the same for all wavelength of each ring and is known for each time slot in each ring: each multi-slot is labeled by the Hub with the identity of the ring which packets transmitted in the multi-slot will be forwarded by the Hub.

Given this behavior, each multi-slot traverses a sequence of ring, e.g., as illustrated in Figure 5.2. Nodes of ring x transmit data to be received by nodes of ring y (steps 2 to 4). Ring x can be viewed as the "upstream" ring, where transmission occur, while ring y can be viewed as the "downstream" ring, where receptions occur.

5.1.2 Scheduling algorithm

The computation of the sequence of permutations operated by the Hub is a scheduling problem as shown in Figure 5.3. The Hub scheduler is driven by a ring-to-ring request matrix $\mathbf{R}$, each element $\mathbf{R}_{i,j}$ contains the number of multi-slots that must be transmitted from input ring i to output ring j. The request matrix $\mathbf{R}$ is decomposed

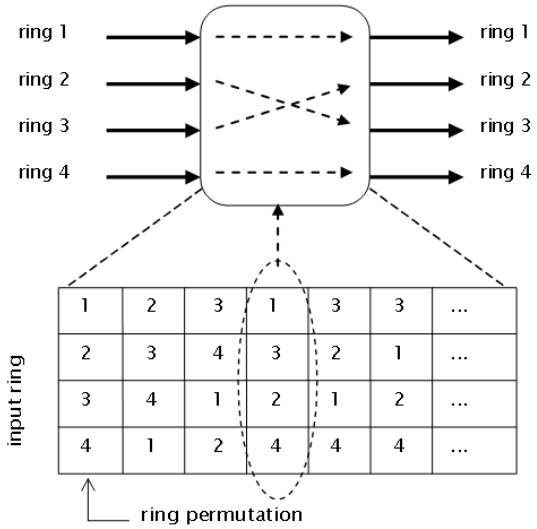

Figure 5.3: Scheduling at the Hub

into F switching matrices $\mathbf{P}$ through iterated application of the critical maximum matching algorithm [63].

5.1.3 Traffic measurements

The request matrix $\mathbf{R}$ is computed at the end of each measurement cycle F, as the sum of 3 contributions:

$$\mathbf{R} = \lceil \mathbf{SM} + \beta\mathbf{IC} + \gamma\mathbf{EC} \rceil$$

with β and γ positive constants, where:

- **SM** is a measure of the average number of multi-slots transmitted among ring pairs during each OW. To smooth out measurement errors, SM is passed through an exponential filter. This is an absolute throughput measure.
- **IC** is the percentage of filled slots which indicates potential congestion situation. This is a relative throughput measure.
- **EC** takes into account explicit congestion signals sent by nodes. This congestion signals are triggered at nodes by SAT: if a node receive the SAT and the length of its queue exceeds the quota Q, the node sends a congestion signal to the Hub.

In order to perform traffic measurement at the Hub, the slot reuse capability (spatial reuse) is not exploited; all traffic is forced to pass through the Hub before being removed from the ring.

5.1.4 Fairness control

The empty-slot operation can exhibit fairness problem under unbalanced traffic. This is particulary true in the ring topology, in which, as already mentioned, upstream nodes have generally better access chances than downstream nodes.

Credit-based schemes is used to enforce throughput fairness. A control signal called SAT is circulated in store-and-forward mode from node to node along ring; a node forwarding the SAT is granted a transmission quota Q and can transmit up to Q packets before the next SAT reception. When a node receives the SAT, it immediately forward the SAT to the next node on the ring if it is satisfied (hence the name SAT), i.e., if no packets are waiting for transmission or if Q packets were transmitted since the previous SAT reception. If the node is not satisfied, the SAT is kept at the node until one of the two conditions above is met.

5.1.5 QoS provisioning

A QoS strategy providing 2 priority classes (i.e., 2 asynchronous services) is proposed in [2]. This is a connection-less approach, therefore it does not guaranteed neither the delay nor the bandwidth. The nodes requiring to send a first priority class packet but no free slot is found, mark a busy slot using an additional flag available in the control channel. This reservation is not meant for the use of the specific node which did the marking but it is at the disposal of any node with first priority packet to be sent. Since the marking is done on already busy slot it is likely that the slot will travel around the rings. In order to not proliferate the marks inefficiently, every node keeps track of its action so that it also removes a reservation mark for every first priority packet sent in an unmarked empty slot. The second priority packets can only use unmarked empty slots.

5.2 Contributions

Our contributions include: (i) the performance evaluation of these architectures and the identification of the drawbacks and of the open issues, (ii) the optimization of the proposed architectures and MAC protocols, and finally (iii) the proposal of different QoS mechanisms to support guaranteed and best-effort services.

5.2.1 Simulation scenario

The performances of the proposed mechanisms are evaluated in order to assess their merits. The simulation results presented in the following sections have been obtained by means of an ad-hoc event-driven simulator reproducing a real scale configuration of the multi-ring network. The parameters of the network are:

- R indicates the number of rings;
- n indicates the number of nodes per ring;

- W indicates the number of wavelengths per fiber;
- B_w indicates the bit-rate. It is set to 10 Gbit/s in every simulation scenario;
- P_s indicates the duration of the time-slot. It is set to 1 μs;
- F indicates the number of time-slots per frame;
- RTT indicates the ring round trip time;
- L indicates the length of the ring;
- ρ indicates the offered load;
- $\mathbf{M}$ indicates the ring-to-ring traffic matrix, whose generic element $\mathbf{M}_{i,j}$ is a real number ranging between 0 and 1 representing the percentage of traffic coming from input ring i and going to output ring j with respect to ρ. Four different traffic matrix are defined named: *uniform* $\mathbf{M}^U$, *diagonal-x* $\mathbf{M}^{Dx}$, *power-of-ten* $\mathbf{M}^P$, and *very unbalanced* $\mathbf{M}^V$. For the case of $R = 4$ and $x = 7$, the matrices are as follows:

$$M^U = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \qquad M^{D7} = \frac{1}{10} \begin{bmatrix} 7 & 1 & 1 & 1 \\ 1 & 7 & 1 & 1 \\ 1 & 1 & 7 & 1 \\ 1 & 1 & 1 & 7 \end{bmatrix}$$

$$M^P = \frac{1}{1111} \begin{bmatrix} 1 & 10 & 10^2 & 10^3 \\ 10 & 10^2 & 10^3 & 1 \\ 10^2 & 10^3 & 1 & 10 \\ 10^3 & 1 & 10 & 10^2 \end{bmatrix} \quad M^V = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{10} & \frac{1}{3} & \frac{1}{15} \\ 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} & 0 \end{bmatrix}$$

  Other traffic matrices used to evaluate specific behavior will be directly introduced in the evaluation description.

The electrical queues at nodes are considered infinite.

Two types of input traffic were considered for the best-effort traffic: the classical model with Poisson distribution of the interarrival times (*Poisson model*) and a self-similar process (*Self-similar model*) implemented as a superposition of 16 strictly alternating independent and identically distributed ON/OFF sources. The duration of each ON/OFF period was assumed to be a random variable with a Pareto distribution with shape $\alpha = 1.2$, which leads to a Hurst parameter of $H = 0.9$ [103]. All packets have the same size and fit in one slot.

Instead, guaranteed traffic is connection-oriented; each node generates connection requests occupying one slot per frame. Both the connection interarrival time and the connection duration are geometrically distributed.

The mean value of the interarrival times for best-effort and guaranteed traffic is selected accordingly to generate the required offered load ρ.

The number of simulated packets is chosen big enough to reach steady-state results and a 95% confidence interval is calculated.

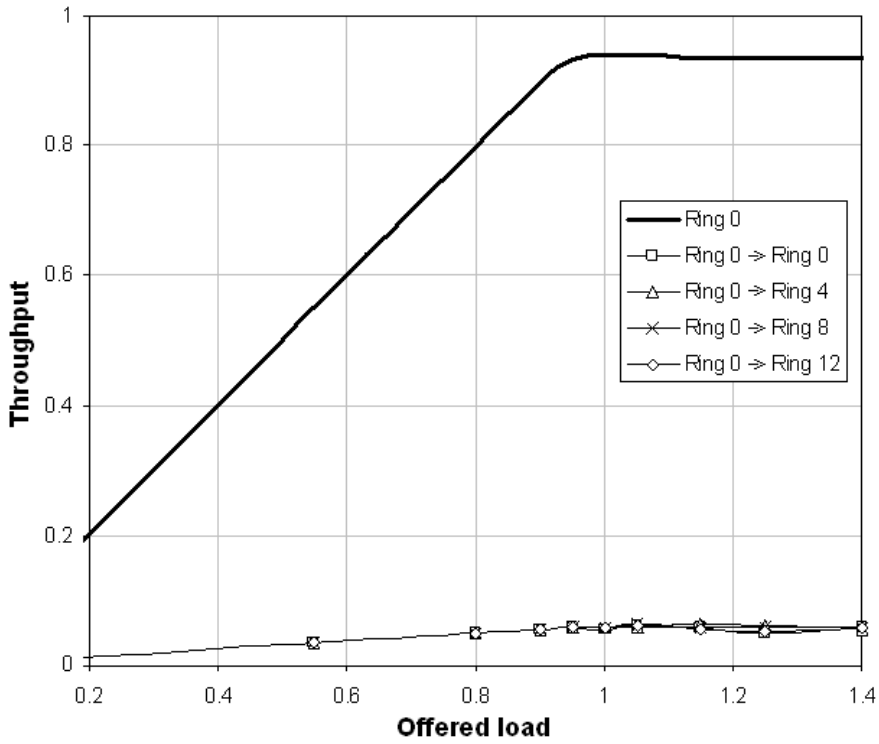

Figure 5.4: Throughput as a function of the offered load under uniform traffic matrix

To evaluate the performance of the multi-ring network, we use both absolute and relative throughput measures. The former calculated as the ratio between used and available slots, the latter as the ratio between used and required slots.

5.3 Performance evaluation

The main issues to address regarding the performance of the multi-ring metro network are the achievable throughput and the fairness. The more the throughput is close to one (the ideal value) the better the MAC protocol. This is not enough because we also want the MAC protocol to share as evenly as possible the bandwidth between the nodes.

In Figure 5.4, Figure 5.5, and Figure 5.6 we set up a network with $R = 16$ rings, $n = 10$ nodes, $W = 4$ wavelengths, $Q = 500$ packets, $L = 100$ km which means that a slots needs $RTT = 0.5$ ms (500 slots) to circulate around a ring, and $F = 10000$ slots per frame. The traffic model is the Poisson one.

Figure 5.4 shows the throughput per destination ring 0, 4, 8 and 12 on ring 0, and the overall throughput of ring 0 (solid line) as a function of the offered load under uniform traffic matrix. Although we report the throughput for some rings, the same behavior holds for all other rings due to the traffic symmetries. We can observe that the throughput increases with the offered load until it reaches the saturation. The total network utilization is close to 0.935 and each destination ring is treated fairly.

In Figure 5.5, the throughput per node is plotted against the number of nodes

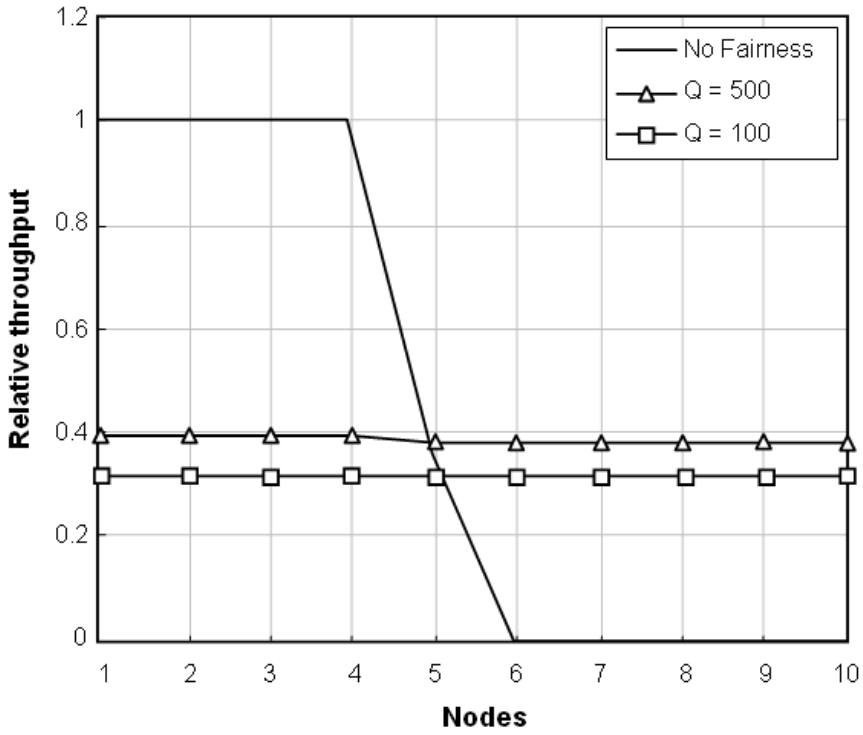

Figure 5.5: Relative throughput per node for total load on the ring of 0.7, without fairness control (solid line) and with SAT for two values of Q

active on the ring. Two cases have been simulated with the SAT quota set at $Q = 500$ and $Q = 100$. The traffic is uniform with a load of 0.7 per ring, and node unbalanced: all nodes send packets only to node 0 of ring 0, except node 0 of ring 0 that sends packets uniformly to the rest of nodes of ring 0. As expected without fairness control (solid line), the bandwidth utilization is unfair: upstream nodes (node 1, 2, 3, 4, and partially the node 5) use all empty slots (the relative throughput is one) and the downstream nodes (node 6, 7, 8, 9, and 10) cannot transmit (the used bandwidth is zero). By introducing the fairness control the bandwidth utilization becomes fair both for the case $Q = 500$ (triangle markers) and $Q = 100$ (square markers).

The fairness problem for inter-ring communications is addressed in Figure 5.6. Here, the diagonal-3 traffic matrix is considered, where about 30% of the traffic is intra-ring while the remaining 70% is evenly spread among the remaining rings, and per-ring uniformly distributed among the nodes. We can see that for low values of the offered load (not congesting the network), the throughput is proportional to the traffic matrix weights. For higher values of the offered load, the rings are treated according to a max-min fairness criteria: the intra-ring throughput decreases to give a fair bandwidth portion to the inter-ring connections.

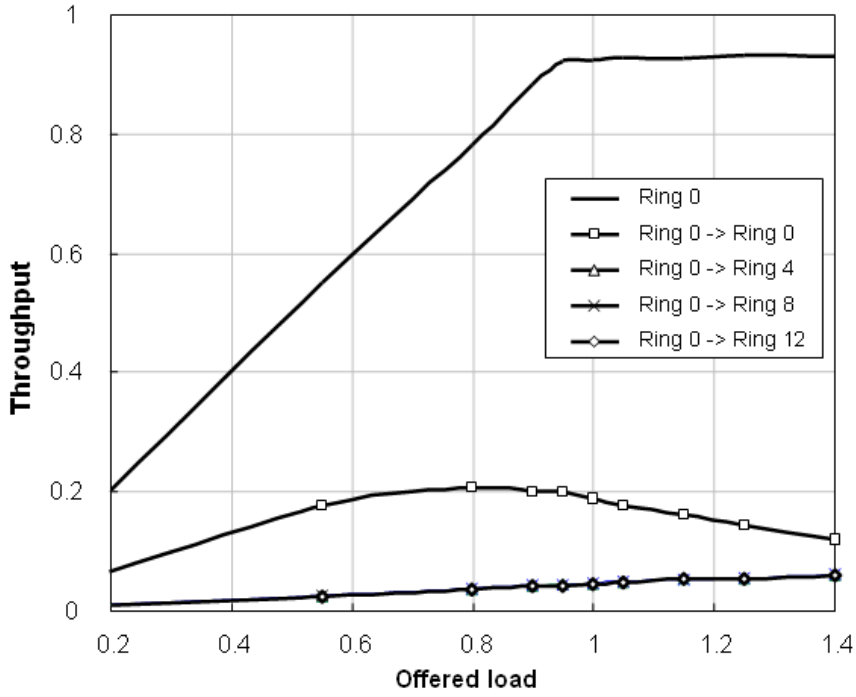

Figure 5.6: Throughput as a function of the offered load under diagonal traffic matrix

5.4 Optimization

From the performance evaluation section, we identify two main weaknesses.

The first one regards the impossibility of exploiting the spatial reuse capability which is one of the main advantageous of the ring topology. Indeed, in the original proposal, the Hub needs to measure every transmitted packet to estimate the traffic request matrix and generate the ring-to-ring permutations.

To perform traffic measurement in presence of spatial reuse, we introduce an additional field in the control channel which we call SR. If any node receives a slot previously transmitted by a node on the same ring before it passes through the Hub, the node marks the SR field. So then, the Hub has to react according to the following three possible actions: 1) when the Hub receives empty both the slot and SR, the Hub count 0 slots; 2) when the Hub receives an empty slot but SR is marked, the Hub measures one slot; and 3) when the Hub receives a full slot, the Hub measures one slot independently of the status of SR.

The benefits of exploiting the spatial reuse in the MAC protocol is studied in Figure 5.7 which shows the throughput with (dashed line) and without (solid line) exploiting the spatial reuse for a metro with $R = 16$ rings and $W = 4$ wavelengths (Figure 5.7a), and with $R = 4$ and $W = 16$ (Figure 5.7b); the diagonal-3 and diagonal-7 traffic matrices are used respectively. For both figures, the interarrival traffic follows the Poisson model. The rest of parameters is set as in the previous study.

As expected, the higher the percentage of intra-ring traffic the more the gain of performance due to the spatial reuse. This is more evident in small networks (low

number of rings). This result is quite obvious if we consider that the spatial reuse is exploited only in presence of intra-ring traffic.

The second weakness regards the notification of the explicit congestion signals. In the original MAC protocol, the SAT token has been used for providing two functions: 1) controls the fair access by limiting the number of transmitted slots, and 2) triggers the notification of the congestion situation from nodes to the Hub. Under particular traffic conditions, the Hub may get the urgencies provided by SATs with considerable delay.

In order to overcome this problem, a new mechanism is introduced in such a way that multi-slots operate as triggers for notifying the explicit congestion signals from nodes to the Hub independently of the SATs. Hence, the Hub will take into account the signals to generate the new set of ring permutations.

As an illustrative example, in Figure 5.8 we study the capacity of the network to overtake congestion situation comparing the original and the optimized solution. The parameter of the network are $R = 4$, $W = 16$, and $n = 10$. The rest of parameters remains unchanged with respect to the previous study. The Self-similar traffic model and diagonal-7 traffic matrix are used. To enforce the congestion, at time 700000 slots, the first and second columns of the traffic matrix are interchanged which simulates a drastic traffic fluctuation. The offered load is set to $\rho = 0.85$.

Both figures show a transient behavior due to the time required for adapting the network to the new traffic situation. The original proposal lasts around 200000 time-slots to return to a stable condition, while it is only 100000 time-slots for the optimized solution. This result indicates that the modification really improve the capacity of the network to solve congestion situations.

5.5 QoS provisioning

5.5.1 Problem formulation

In Chapter 3 we discuss that we are interested to support 3 different services: guaranteed, priority and best-effort. Since work in [2] proposed a method to support priority and best-effort (GS) services, we concentrate our contribution in proposing a method to support the guaranteed service.

We hence consider a guaranteed service (GS) traffic class with guaranteed bandwidth, and a best effort (BE) traffic class. The management of BE traffic is a relative easy task as has been studied in [9] and briefly summarized in Section 5.1. The Hub can schedule the permutations according to both traffic measurements and congestion signals issued by nodes allowing the nodes to decide whether to transmit and/or to receive by checking on the control slot of the multi-slot what has been already transmitted by upstream nodes. On the other hand, a connection-oriented approach is necessary to guarantee the GS traffic requirements, where the Hub establishes connections between the nodes reserving the required resources along the rings. Therefore, the scheduling algorithm performs two tasks: (1) maintaining the virtual, time-slotted connections between the nodes, and (2) fairly sharing the unreserved

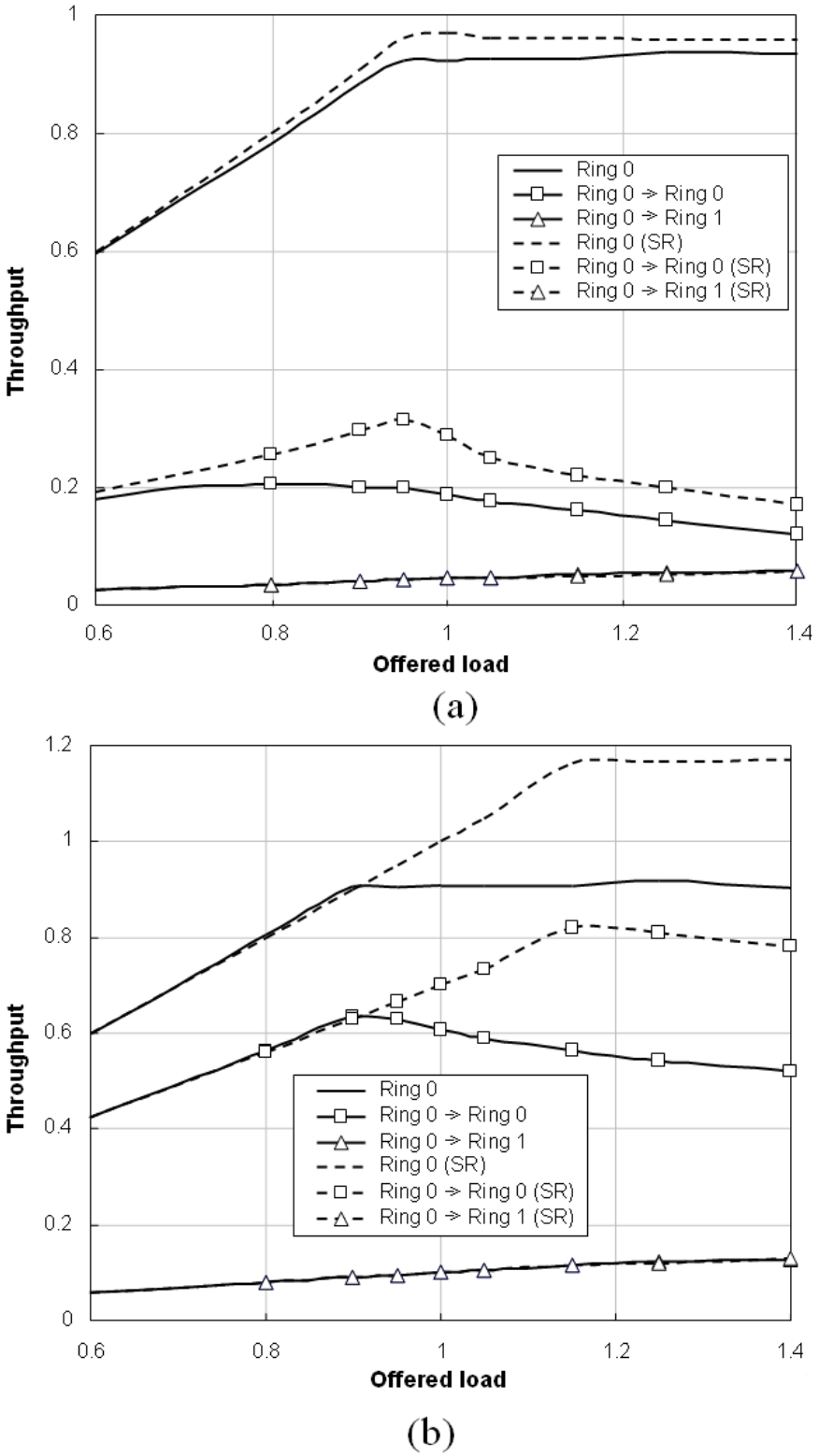

Figure 5.7: Throughput as a function of the offered load under diagonal traffic matrix with spatial reuse (dashed line) and without spatial reuse (solid line). a) Network with 16 rings and 4 wavelengths per ring, b) Network with 4 rings and 16 wavelengths per ring

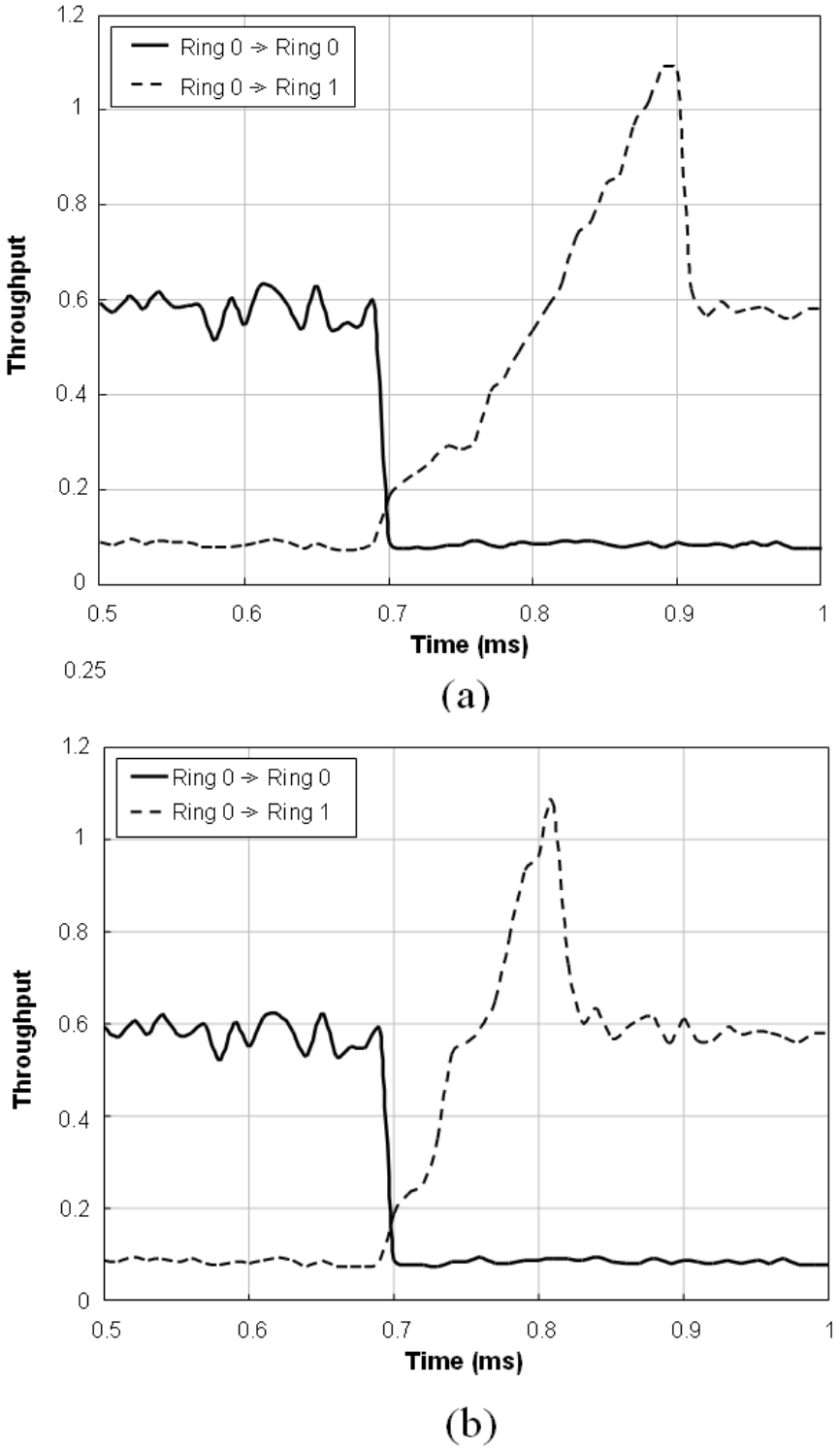

Figure 5.8: Throughput as a function of the offered load under diagonal-7 traffic matrix with traffic fluctuation comparing a) the original and b) the optimized solution

bandwidth to the BE traffic (traffic subject to fairness control).

To set up the virtual connection, the nodes send explicit GS reservation requests to the Hub which collects them in a node-to-node connection request matrix **A**. At the same time, the Hub estimates the BE traffic requirements by measuring the BE traffic load and calculates a ring-to-ring matrix **B***.

The scheduling is based on a fixed-size frame of length F slots, which is considered the most suited solution for a system with guaranteed bandwidth allocation. Indeed, with a fixed-size frame, a reservation issued in terms of bit rate can easily be translated into an equivalent number of slots per frame. The frame length F must be chosen trading the allocation granularity (asking for longer frames) for the access delay and scheduling complexity (asking for shorter frames). The selection of optimal values for F is outside the scope of this thesis, but we typically envision several thousand slots in the frame.

Therefore, the problem of scheduling multi-class traffic can be stated as follows:

- GIVEN
  the node-to-node GS connection request matrix **A**
  the ring-to-ring BE request matrix **B***
- FIND
  a conflict-free slot allocation within a frame of length F
- SUCH THAT
  the number of allocated GS connection is maximized
  then, the number of transmitted BE packets is maximized

Satisfying the following constraints:

1. *Priority of GS traffic over BE traffic.* GS requests must be satisfied before serving BE traffic.
2. *Persistent allocation of GS connections.* The allocation of new GS and BE traffic must not affect currently established GS connections.
3. *Atomic allocation of GS requests.* GS requests are accepted only when they can be fully satisfied; otherwise, they must be refused. Atomic allocation typically makes sense when requests correspond to single real-time user connections, whereas it does not apply to elastic BE traffic, nor to sources that multiplex several data flows.
4. *No contentions.* Since each node is equipped with only one tunable data transceiver, it can transmit and receive at most one packet in each multi-slot.
5. *Avoid in-transit collision at the Hub.* A packet collision may occur at the Hub between the packets injected in the downstream ring (after having been switched at the Hub) and the slots reserved by the Hub in the upstream ring for the transmission of GS traffic.

The multi-class scheduling problem is NP-hard because it is generalization of the well-known knapsack optimization problem, which is NP-hard [83]. In fact, GS traffic requests can be considered as a set of objects of different sizes that must be fit in a knapsack of capacity F. The problem can be solved in polynomial time if we adopt heuristic solutions, which decrease the complexity of the scheduling accepting some degree of worse performance.

5.5.2 Heuristic solution

Given the node-to-node matrix **A** which contains the GS connection requests and the ring-to-ring matrix **B*** which contains the BE traffic load measurements, the heuristic approach is an incremental algorithm, which consists of three steps:

1. At first, all the slots that were allocated in the previous frame to BE traffic, as well as those corresponding to ended GS connections are released, so that only persistent GS connections remain allocated.

2. New GS requests are scheduled scanning the resources in a round-robin way. This steps ends when either all requests have been satisfied or all slots in the frame have been considered.

3. The remaining slots are used to allocate BE traffic contained in the matrix **B*** which is scheduled independently of GS traffic through an iterated critical maximum size matching [83]. The set of ring-to-ring permutations obtained must be fit in the unused part of the frame by selecting the permutations that allocate the maximum number of slots.

The complexity of this algorithm is $O(N^2F)$.

This algorithm satisfies all the constraints described above except the in-transit collision. Different methods can be used to fulfill this requirement.

For example, the Hub can take into account nodes' positions along the rings, while computing the scheduling and can inform all the nodes about the GS reservation before issues the new set of permutation matrices. In such a solution, which we call *Full Shared* (FS), the in-transit collision constraint can be satisfied since the nodes know the slots reserved for the other nodes.

The problem can be made easier by separating transmissions and receptions either in frequency or in time. The first approach leads to the *Frequency Decoupling* (FD) solution where the nodes use $\frac{W}{2}$ wavelengths to transmit, and the rest to receive packets. This method precludes the spatial reuse capability since all packets must be switched at the Hub before being received. In the second approach, dubbed *Time Decoupling* (TD), the same set of wavelengths is shared in the time domain and used only one RTT (Round-Trip Time) out of two to transmit the packets. In this case, the spatial reuse capability is available since the node can receive in every moment. For both solutions, the Hub can schedule the matrices without knowing the nodes' position along the rings.

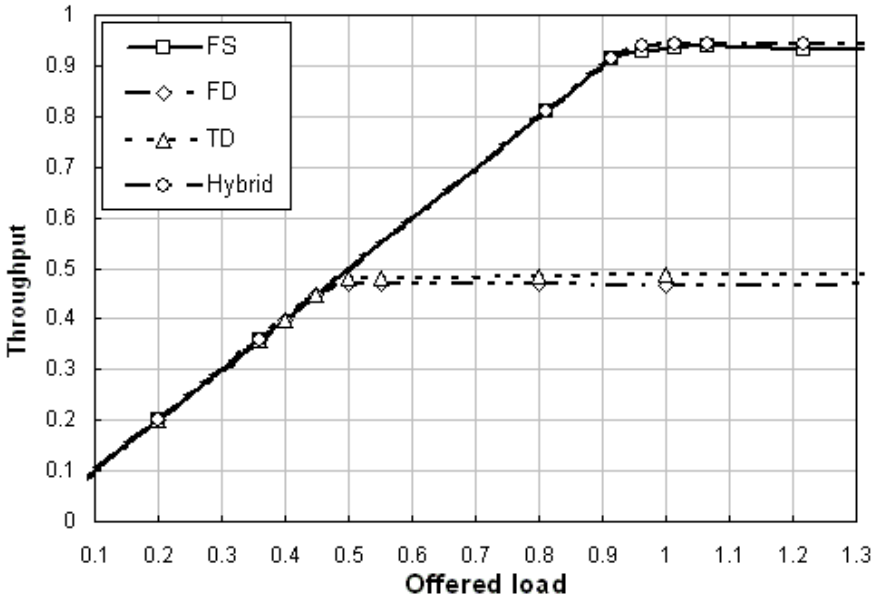

Figure 5.9: Throughput as a function of the offered load under uniform traffic matrix. GS load is fixed to 30%

Nonetheless, both FD and TD have the drawback that the available bandwidth is halved by the fixed resource partitioning. This problem can be overtaken if the decoupling is only applied to the GS traffic, and not to the BE traffic. In this case, the Hub can schedule the matrices without knowing the nodes' position along the rings but must inform the nodes about the GS reservation as in the FS approach.

5.5.3 Performance evaluation

The simulated network consists of $R = 4$ rings and $n = 10$ nodes per ring, each ring with $W = 4$ wavelengths (plus 1 for the control channel). The length of the ring is $L = 100$ km and the quota for the BE traffic is $Q = 500$. The GS traffic is not subject of fairness control. The frame is $F = 10000$ slots. The traffic is the self-similar model.

Figure 5.9 shows the total (BE+GS) throughput of the four solutions as a function of offered load (with GS = 30% of total load) considering uniform traffic distribution. The Hybrid solution obtains the best performance.

Figure 5.10 shows the throughput as a function of the HP relative load percentage assuming 100% total load and unbalanced traffic distribution; 70% of traffic is intra-ring (where the space reuse is possible), the rest is uniformly distributed among the inter-rings. The figure depicts the BE and total (BE+GS) throughput. The space reuse exploitation makes TD better than FD, while Hybrid still achieves the best results as far as GS traffic requires less than 50% of total resource.

Table 5.1 compares the complexity of the four solutions showing the average running time to obtain a point of the previous figures.

The simulation results show that the Hybrid solution is preferable when the GS traffic is less than 50% of total resources. TD and FD solutions can be adopted when

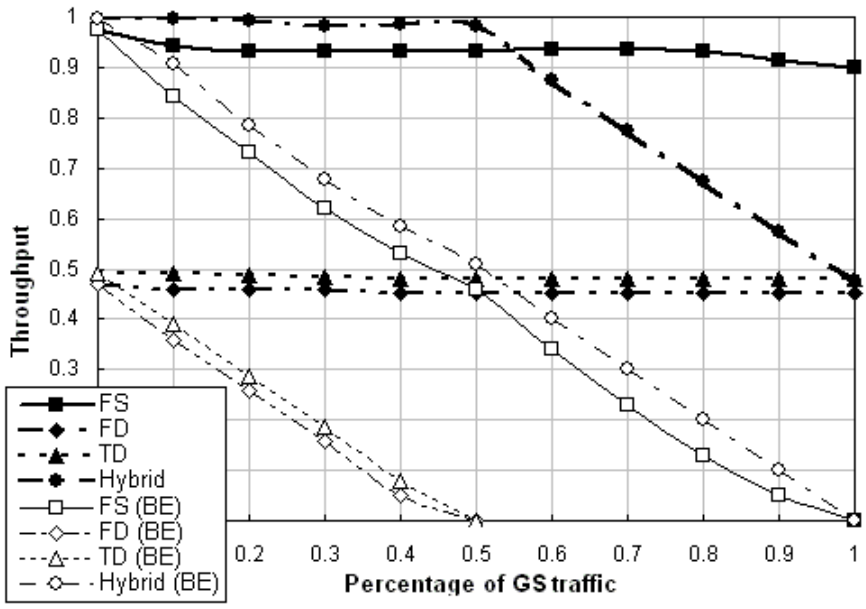

Figure 5.10: Throughput as a function of GS traffic relative load assuming 100% total load under diagonal traffic matrix

Table 5.1: Average running times for the four solutions

	FS	**FD**	**TD**	**Hybrid**
Time	2298 s	1354 s	1298 s	1711 s

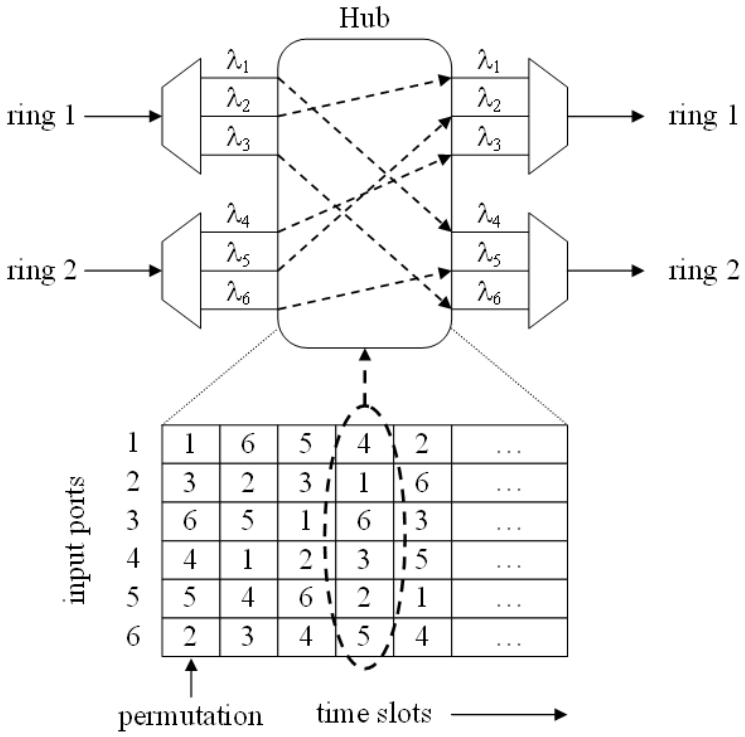

Figure 5.11: Scheduling wavelength-to-wavelength permutations at the Hub

the bandwidth on network links is not a bottleneck.

5.5.4 Optimization of the QoS mechanism

Scheduling can aim at different levels of performance guarantees. In general, an allocation of slots in the frame should provide average rates to node pairs in accordance with the traffic matrix. The scheme proposed in [9] for BE traffic allocates ring-to-ring rates at the Hub, and access decisions are decentralized at nodes, which, however, do not have guaranteed access. Advantages of that approach are the very small amount of information in the control channel and the good scalability properties.

However, if higher node-to-node guarantees are required, as it is the case for GS traffic, the scheduler must allocate slots to single node-to-node requests and perform wavelength-to-wavelength permutations among rings. The resource allocation problem becomes mostly centralized, and the amount of information in the control channel increases. It must be noted that a centralized scheduling is often desired by network operators, who may want to apply different control policies for different nodes (e.g., introducing an admission control policy for GS traffic requests).

Figure 5.11 shows a metro network comprising two rings conveying three wavelengths each, where the Hub operates wavelength-to-wavelength permutations.

Figure 5.12 illustrates the operation of a network when the Hub performs wavelength-to-wavelength permutations. For simplicity, we consider a network with $R = 2$ rings, $W = 3$ wavelengths per ring, and $n = 3$ nodes per ring. Control slots are on the

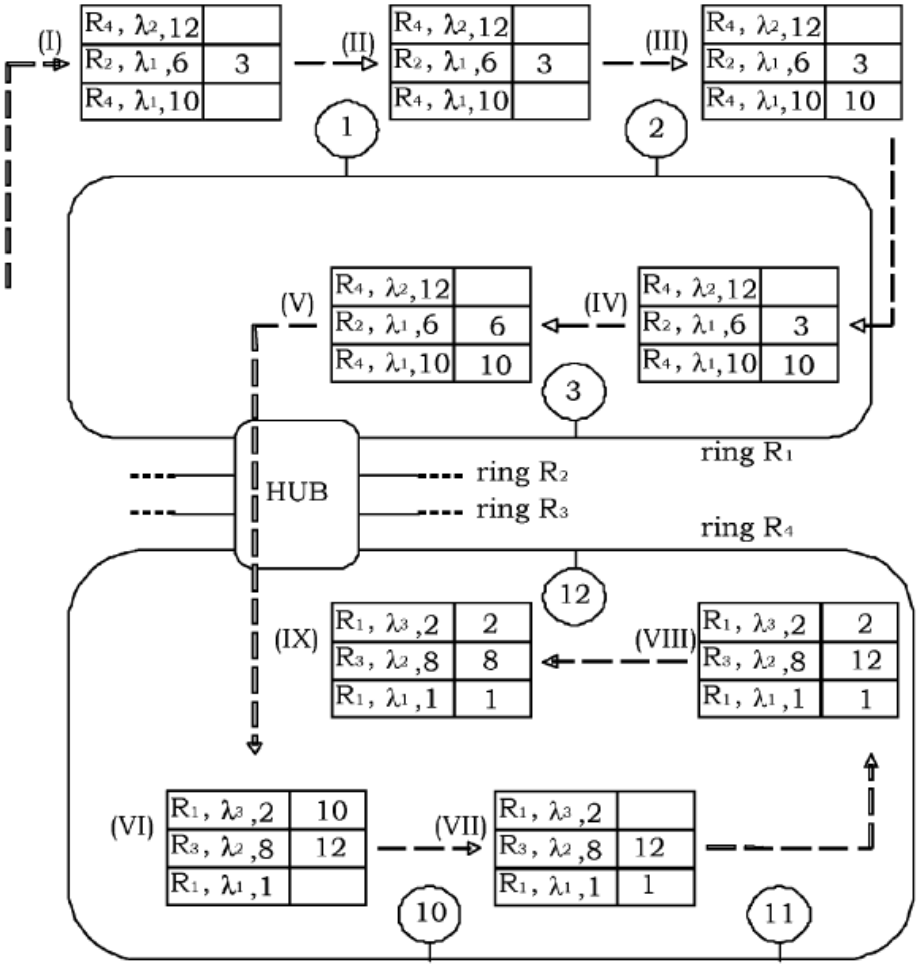

Figure 5.12: Example of slot forwarding in the multi-ring network with wavelength-to-wavelength permutations

left-hand side of multi-slots, whereas data slots are on the right-hand side. Data slots are reserved for packets addressed to nodes identified in the triple (ring, wavelength, node) inserted by the scheduling algorithm running at the Hub in the corresponding control slot. Numbers in data slots denote the destination node to which the packets they convey are addressed. We describe the slot forwarding process in successive steps showing the path followed by a multi-slot as it travels from ring R_1 to ring R_4.

1. Control slots indicate that the first data slot is reserved for sending data to node 12 on ring R_4 using wavelength λ_2, the second data slot is reserved for a packet to be delivered on λ_1 to node 6 on R_2, and the third slot is booked for transporting on λ_1 a packet addressed to node 10 on R_4. The second data slot is carrying a packet headed to node 3.

2. Node 1 leaves the multi-slot untouched.

3. Node 2 has a packet for node 10 on ring R_4 and inserts it in the corresponding data slot.

4. Node 3 drops the packet transported in the second data slot and reuses it to send a packet to node 6 on ring R_2.

5. The multi-slot reaches the Hub.

6. The Hub switches data slots to their intended output ring/wavelength as indicated by the triple contained in the relevant control slot. Therefore, the third slot containing a packet headed to node 10 is switched to wavelength λ_1 (first slot) of ring R_4, and a packet coming from another ring and with destination node 12 is inserted in the second slot. Finally, the scheduling algorithm inserts new destination triples in the control slots.

7. Node 10 drops the packet sent from node 2 and transmits a packet to node 1 using the third data slot.

8. Node 11 transmits a packet to node 2 on ring R_1 using the first data slot.

9. Node 12 drops the packet in the second slot, and transmits a new packet to node 8 on ring R_3.

In the presence of multi-class traffic, the Hub collects bandwidth requests issued by nodes for both GS and BE services, builds two node-to-node traffic request matrices (**A** and **B**), and schedules them in a sequence of wavelength-to-wavelength permutations. Both request matrices store the number of slots that must be transmitted from a node to any other node within a frame of F slots.

Now, the Hub knows all the packets that will be transmitted during the next frame, therefore it is easy to satisfied all the constraints described above using the following heuristic scheduling algorithm which is an adaptation of the previous one:

1. All the slots assigned to BE traffic and all the slots reserved for ended GS connections are released.

2. New GS connection requests are scheduled scanning the resources in a round-robin way. This step ends when either all requests have been satisfied or all slots in the frame have been considered.

3. BE requests are scheduled using wavelength-to-wavelength permutations scanning the resources in a round robin way taking into account the in-transit collisions.

The complexity of this algorithm is $O(N^2FW)$.

In order to evaluate the performance of the optimized solution, we also consider two frequency decoupling solutions proposed in [10] for the passive multi-ring node architecture. Indeed, the passive multi-ring well adopts the FD method since no erasure stage is required. These solutions comprise an optimum polynomial algorithm with no atomicity constraint of the GS connection requests and a faster greedy algorithm based on the same round-robin method adopted in our heuristic algorithm.

We study by simulation a network configuration comprising $R = 4$ rings and $n = 16$ nodes per ring, each node sharing $W = 4$ wavelengths. The latter means that for the multi-ring architecture each ring conveys 5 wavelengths (4 for data and 1 for control), whereas, for the passive multi-ring architecture, the wavelengths carried on each ring are 9 (4 for upstream traffic, 4 for downstream traffic, and 1 for control). The ring round-trip time to $RTT = 512$ μs, which means that the propagation delay on each ring is 512 times the slot duration, and the frame duration is $F = 10240$ slots (20 RTTs).

All figures consist of three plots showing the performance of the optimum algorithm for the passive configuration and the heuristics for both the passive and the active multi-ring architecture.

All plots show the throughput as a function of the amount of GS traffic present in the network, when the total BE offered load is exactly 1. In other words, when the GS traffic load on the horizontal axis of the figures is 0.2, the total network load is 1.2. Note, however, that both BE and GS traffic are distributed among different rings according to the chosen ring-to-ring matrix $\mathbf{M}^x$.

The plots in Figures 5.13-5.15 show the throughput for each destination ring on source ring 1 for GS traffic (white markers), the total GS throughput (black square markers), the total BE throughput (dashed line without markers), and the total throughput on ring 1 (solid line without markers). Although we plot the throughput for a single ring, the same behavior holds for all the other rings due to traffic symmetries.

Figure 5.13 compares the three solutions under uniform traffic. Figure 5.13(a) shows that GS throughput increases with the offered load until it reaches the value of 1 in overload. The overall throughput is constantly equal to 1, as it is always possible to fill with BE traffic the slots left free by GS connections. Figure 5.13(b) presents the same behavior as Figure 5.13(a) except when the offered load exceeds 1. In this case, the heuristics is not capable of maintaining the overall network throughput to 1.

In the active multi-ring configuration shown in Figure 5.13(c), the network behaves differently. Indeed, in this case, we have half wavelengths and more contention probability due to shared transmission and reception channels; therefore, when only BE traffic is present in the network the overall network throughput is 0.97. As the GS traffic increases, the network throughput increases as well, reaching values even higher than 1. Note that in this case, the amount of BE traffic never drops to 0. Indeed, since we measure throughput as the ratio between the number of transmitted packets and the number of available slots, it may happen that a slot can be used more than once to transport different packets during one single round trip. For instance, a node can transmit a packet to a neighboring node on the same ring; the destination node can reuse the same slot to transmit another packet to a different node on the same ring, and so on. This is an implicit consequence of the wavelength reuse capability, and increases network throughput. Nevertheless, this gain can be exploited only for intra-ring traffic, i.e., when the transmitter and the receiver belong to the same ring. In any other case (i.e., for inter-ring traffic), packets must be switched at the Hub from their source ring to their destination ring and, therefore, slots containing such

packets cannot be reused. For this reason, when traffic is mostly inter-ring, the gain obtained from wavelength reuse is lower: in Figure 5.13(c), it reaches 1.05, and in Figure 5.15(c), it is not even noticeable. In Figure 5.14(c), instead, it becomes more evident, reaching 1.17, because the diagonal traffic pattern has a higher percentage of intra-ring traffic than the other scenarios.

In Figure 5.14(a), obtained using the optimum algorithm for the passive multi-ring architecture, the overall throughput is always 1, and GS throughput proportionally increases with the offered load. It is interesting to note how intra-ring traffic, after reaching a value of about 0.7, starts to decrease and leaves resources to inter-ring traffic when the total GS traffic equals to the network capacity. This is due to the fact that the scheduler in overload tends to equalize the load on the rings, according to maxmin throughput fairness. As before, the heuristic algorithm in Figure 5.14(b) behaves similarly to the optimum one, except when the offered load exceeds 1. Moreover, the heuristics does not equalize rings' load.

In Figure 5.15(a), we observe the optimum behavior of the algorithm with the power-of-ten traffic pattern. All GS connections are allocated optimally and total throughput remains equal to 1. This is not true for both heuristic algorithms. The heuristics for the passive configuration in Figure 5.15(b) is not able to allocate all GS requests, reaching a throughput around 0.98 independent of the amount of GS traffic injected. The heuristics for the active configuration in Figure 5.15(c) performs even worse and the total network throughput does not rise above 0.9.

Finally, in Figure 5.16, we analyze the case of the very unbalanced traffic pattern: Since it is not symmetric, in each subplot, we show the throughput of GS traffic on each ring, the total network throughput of GS and BE traffic, and the overall network throughput. The offered load is normalized with respect to ring 2, where the traffic is higher. This means that when GS traffic load in the horizontal axis of Figure 5.16 is equal to 1, ring 1 is 50% loaded, ring 2 is 100% loaded, and rings 3 and 4 are 33.3% loaded; therefore, the total network load is about 0.54. Figure 5.16(a) shows that while the allocation of GS traffic remains very close to the optimum, the very unbalanced traffic pattern causes total traffic allocation to be suboptimal. This is due to the non optimality of a double matching with two traffic classes. Indeed, when only one class is present in the network, the total throughput is always equal to 100%. Furthermore, heuristic algorithms in Figure 5.16(b) and (c) are not able to allocate all traffic.

5.6 Summary

In this part of the thesis, we focussed on the multi-ring network architecture. Its performance has been evaluated by simulation considering several scenarios. The performance results have been obtained using a real scale simulator including self-similar traffic model and different traffic patterns between interconnected rings.

Two main weaknesses have been identified and overtaken by optimized mechanisms. The spatial reuse capability is exploited adding a special field in the control slot. The results indicate that, as expected, the higher the percentage of intra-ring

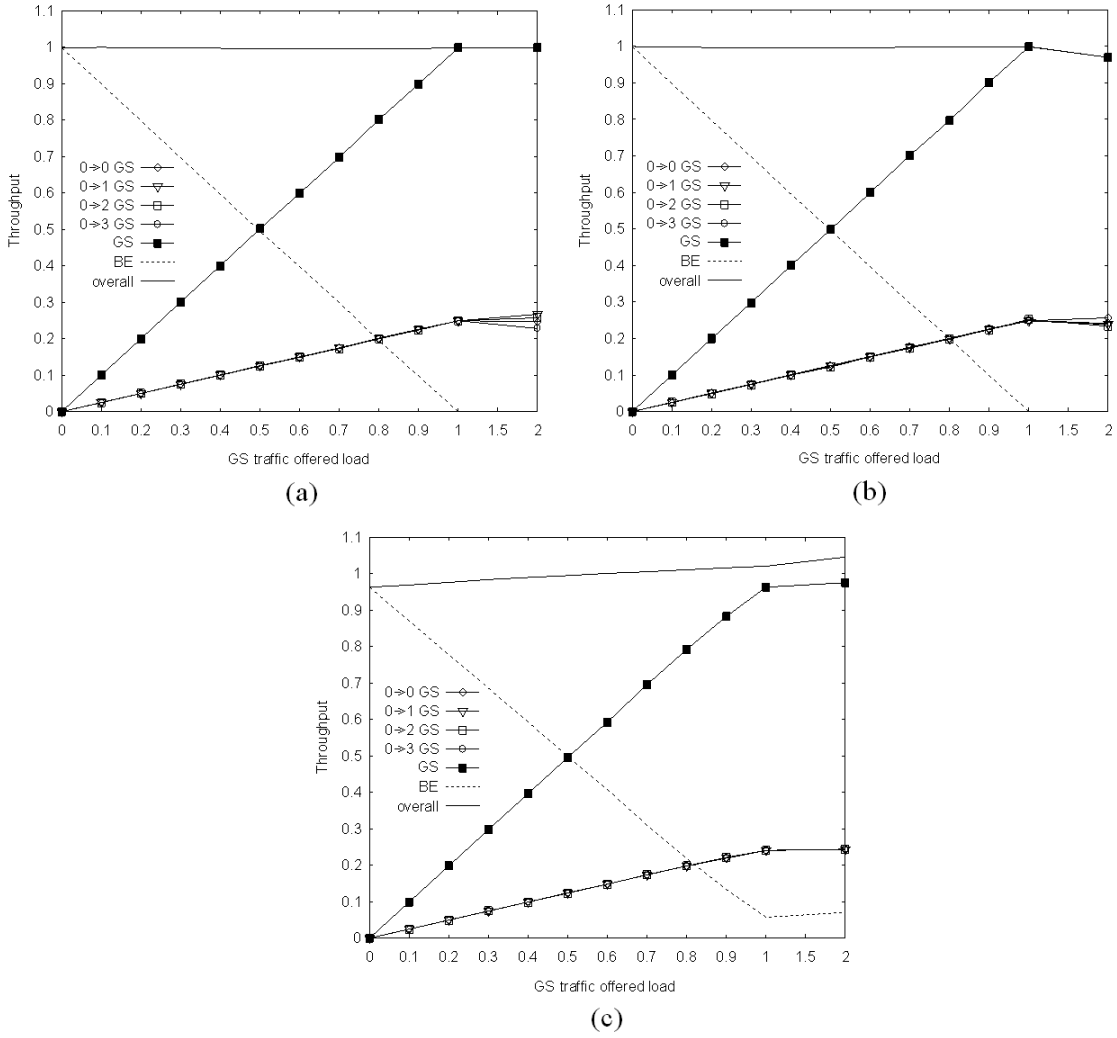

Figure 5.13: Throughput as a function of GS traffic relative load under the uniform traffic matrix: comparison of a) the optimal and b) the heuristic solution for the passive multi-ring configuration, and c) the heuristic solution for the active multi-ring configuration

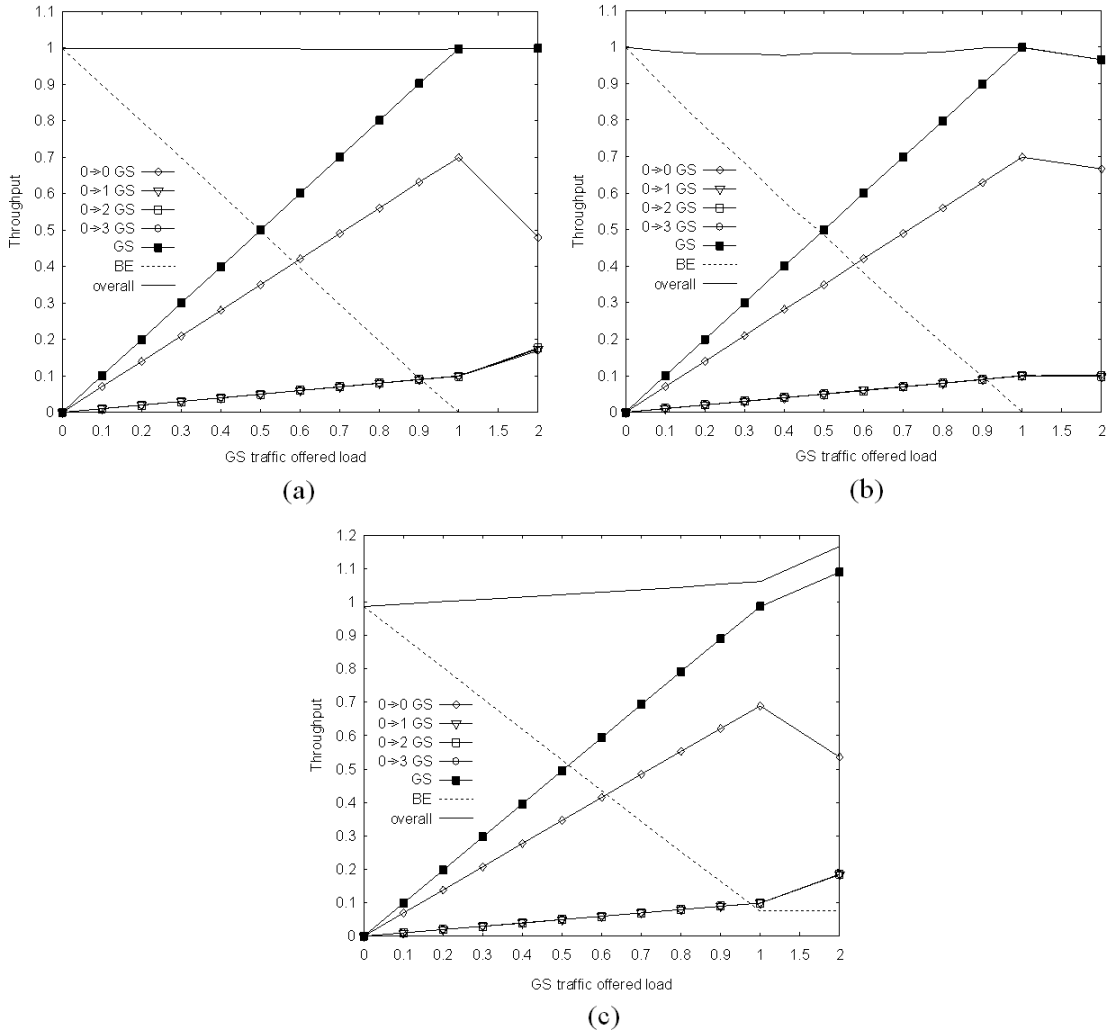

Figure 5.14: Throughput as a function of GS traffic relative load under the diagonal traffic matrix: comparison of a) the optimal and b) the heuristic solution for the passive multi-ring configuration, and c) the heuristic solution for the active multi-ring configuration

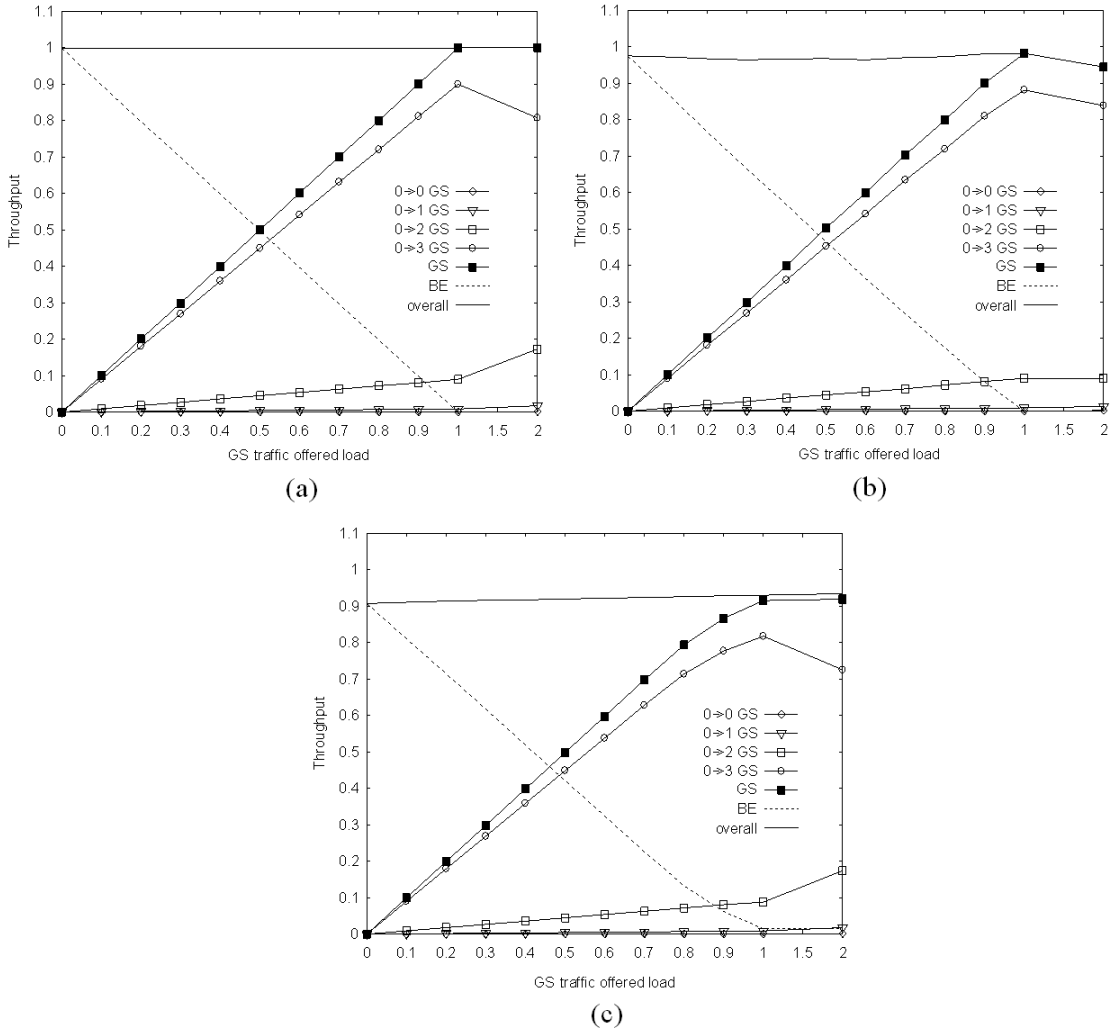

Figure 5.15: Throughput as a function of GS traffic relative load under the power-of-ten traffic matrix: comparison of a) the optimal and b) the heuristic solution for the passive multi-ring configuration, and c) the heuristic solution for the active multi-ring configuration

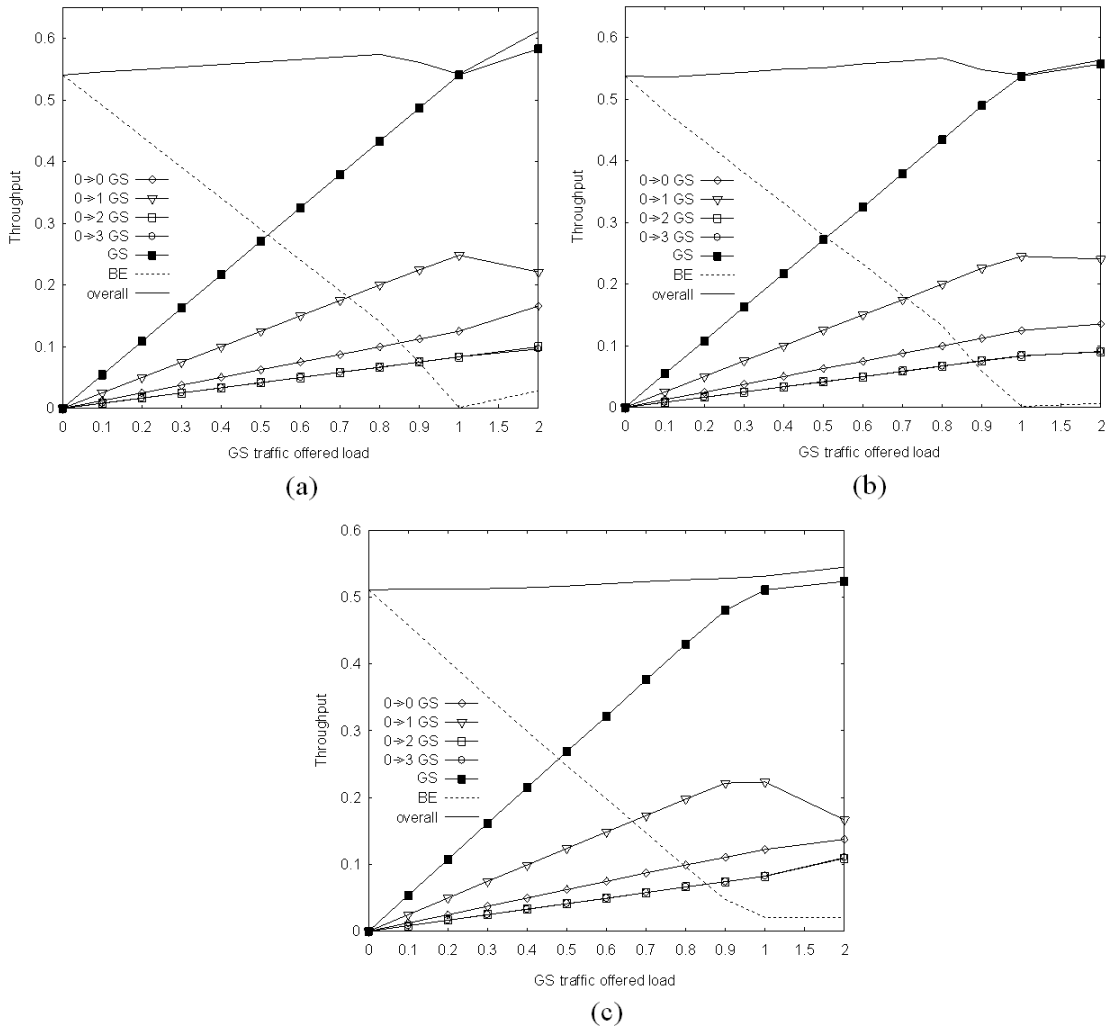

Figure 5.16: Throughput as a function of GS traffic relative load under the very unbalanced traffic matrix: comparison of a) the optimal and b) the heuristic solution for the passive multi-ring configuration, and c) the heuristic solution for the active multi-ring configuration

traffic the more the gain of performance due to the spatial reuse. This is more evident in small networks (low number of rings). This result is quite obvious if we consider that the spatial reuse is exploited only in presence of intra-ring traffic. The congestion notification mechanism has been improved using the multi-slots as triggers for notifying explicit congestion signals from nodes to the Hub. The validity of the proposals have been demonstrated by numerical results.

Finally we discussed the problem of allocating resources to provide guaranteed and best-effort services in different configurations of the DAVID metro network. We discussed architectural alternatives, and different formulations of the resource allocation problem. The latter is solved in a mostly centralized fashion at the Hub, but some access decisions may be de-centralized at network nodes, depending on the network configuration, and on the desired level of performance guarantees. Tradeoffs between optimality and complexity of the allocation schemes were observed by simulation. Our study shows the flexibility of the network architecture, and the effectiveness of the proposed strategies in accommodating very diverse traffic patterns.