Heuristic Strategies for Routing 100-GbE over OTN in Distributed Differential Delay Compensation Architectures

João Santos[1,2], **João Pedro**[1,2], **Paulo Monteiro**[1,3], **and João Pires**[2]
1 - Nokia Siemens Networks Portugal S. A., R. Irmãos Siemens, 1, 2720-093 Amadora, Portugal, joao.santos@nsn.com
2 - Instituto de Telecomunicações, Instituto Superior Técnico, Av. Rovisco Pais, 1049-001 Lisboa, Portugal
3 - Instituto de Telecomunicações, Universidade de Aveiro, Campus de Santiago, 3810-193 Aveiro, Portugal

Abstract: A heuristic optimization framework is proposed for routing virtually-concatenated 100Gb/s Ethernet over optical transport networks with distributed differential delay compensation. Under short computing times, reduced buffer sizes and limited link capacity requirements are obtained.

OCIS codes: (060.4250) Networks; (060.4510) Optical communications

1. Introduction

The increasing adoption of IP-based applications (e.g. HDTV, file sharing) is leading to an impressive traffic demand growth. With the goal of enlarging the aggregation network capacity, IEEE is currently developing the novel 100 Gb/s Ethernet (100-GbE) interface. For the core segment, ITU-T is also proposing a new digital container for encapsulation of 100-GbE signals into a single wavelength over optical transport network (OTN) systems [1]. This new optical channel data unit (ODU), identified as ODU-4, will complement the existing specifications of 2.5 Gb/s (ODU-1), 10 Gb/s (ODU-2) and 40 Gb/s (ODU-3) [2]. From a technical perspective, the serial 100 Gb/s transmission requires the use of advanced modulation formats to fit into the existing optical fiber infrastructure, thus resulting in complex and expensive transceiver equipment. Therefore, given the immaturity of the serial technology, network operators may favor, at least in the near term, the deployment of parallel transmission approaches.

In the OTN, inverse-multiplexing is enabled by the virtual concatenation (VCAT) protocol. With VCAT, OTN is able to convey a 100-GbE signal over a group of eleven ODU-2 or three ODU-3 containers, denoted as ODU-2-11v and ODU-3-3v, respectively [2]. Furthermore, VCAT permits to route each group member independently, using either a co-routing (CR) or a diverse-routing (DR) approach. In opposition to CR, the DR strategy is able to explore more efficiently the network connectivity by distributing the ODUs through the network in a balanced way and, as a consequence, reduce the maximum link capacity [3]. However, the fact that DR can make VCAT group members traverse spatially diverse paths introduces differential delay at the reception, requiring the incorporation of a large amount of high-speed memory to temporarily store the incoming ODUs when reassembling the client signal, as illustrated in Fig. 1(a). Since building large-sized memories to operate at these speeds is very costly, a distributed differential delay compensation (DDC) scheme with buffering in the intermediate path nodes (see Fig. 1(b)) was recently proposed [4].

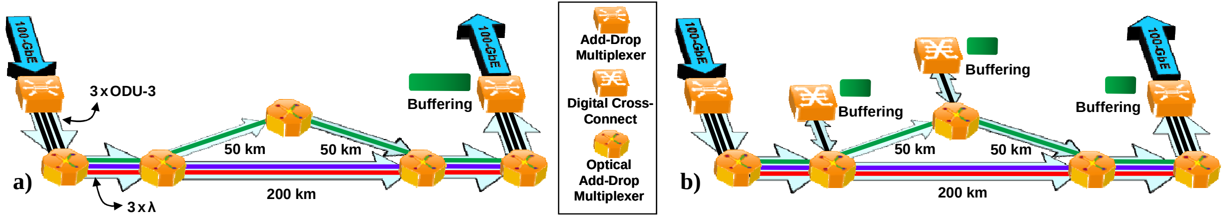

Fig. 1. (a) Centralized and (b) distributed differential delay compensation architectures.

In our previous work [3] we addressed the problem of the absence of a suitable network design method for routing virtually-concatenated ODU containers over OTN with optimal DDC distribution and minimum capacity requirements, by proposing an integer linear programming model. Although the ILP-based optimization guarantees the return of an optimal solution, and since the optimal DDC distribution is a NP-complete problem, it also suffers from extended running times due to the exponential dependency with the number of constraints and variables of the ILP. Therefore, for OTN networks of practical size, only time-constrained executions of the model are possible, leading to sub-optimal solutions. In order to be able to solve large instances of this optimization problem and to improve the quality of the solutions computed in a short time interval, we propose a heuristic optimization solution based in a multi-stage framework for solving the multi-constrained ODU routing problem [3].

2. Heuristic Framework Description

In order to properly tackle the multi-constrained problem, our routing and delay compensation (RDC) framework is implemented with a multi-stage optimization process, as presented in Fig. 2. The idea behind this strategy is to divide a highly complex problem into several sub-problems individually solved via simple heuristic methodologies.

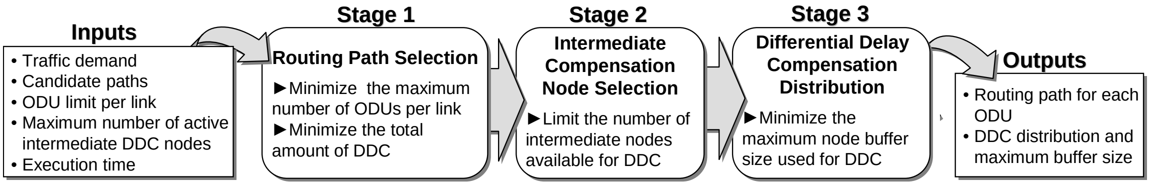

Fig. 2. Building blocks of the multi-stage RDC framework.

<u>Stage 1:</u> Initially, the forwarding path for each ODU is selected in such way that the maximum link capacity and the total amount of differential delay in the network are minimized. Existing solutions [5] exhibit several drawbacks. Firstly, shortest path routing can lead to faster exhaustion of the maximum link capacity by creating bottlenecks in the more frequently used network links. Another aspect that impacts the performance of the static routing algorithm is the order by which the connections are provisioned. Since each node pair is processed sequentially in accordance with a given pre-established (and fixed) order, these local choices have an impact on the final outcome. In view of these limitations, we have developed an iterated search algorithm that rearranges the ordered sets belonging to the node pairs, X, and the candidate paths, Y. Prior to each iteration of the algorithm, the ordered sets X and Y are randomly reordered to enable the exploration of the large solution space. Inside each iteration, the traffic demand of each active node pair is satisfied by individually assigning the ODUs to the first path of the ordered path set X that does not surpass the ODU limit per link received as input (see Fig. 2). If no path in this condition is available, the ODUs are routed over the least loaded path. By repeating this process for all node pairs in each iteration, a different routing solution is obtained. The solution requiring the minimum capacity per link and maximum DDC per node is saved and, after stage 1 is concluded, is delivered to stage 2. The total number of iterations executed by the iterated search algorithm depends on the execution time bound received as an input by the RDC.

<u>Stage 2:</u> After receiving, from stage 1, the ODU paths that require DDC, this stage orders the preferred nodes for intermediate path DDC. The maximum number of intermediate DDC nodes is defined beforehand as an input parameter. Here, each intermediate node traversed by a routing path is added with the differential delay affecting that path. When all paths are processed, the intermediate nodes with the highest quantity of differential delay passing through are placed on top of the selection list. With this method, the preferred intermediate DDC nodes are the ones capable of sustaining a higher amount of DDC and alleviate other DDC nodes, if required, to reduce the buffer size.

<u>Stage 3:</u> In order to distribute the DDC through the nodes, the order by which the ODU paths are processed has to be defined. Here, the routing paths with the highest differential delay are placed first in the processing list. Afterwards, and for each of the listed paths, the correspondent differential delay is distributed over the available DDC nodes. Note that, by definition of the distributed compensation architecture [4], only the termination node and the intermediate nodes (specified in Stage 2) are eligible to perform DDC in any given routing path. To distribute the delay, an equalization algorithm is used to level the amount of differential delay compensated in the path nodes. With this strategy, we aim at balancing the overall DDC distribution and reduce the maximum node buffer size.

3. Performance Evaluation

In this section, the proposed RDC heuristic is assessed against our previous ILP solution [3] in NSFNET and EON network topologies (depicted in Fig. 3). We assume that each 100-GbE stream is splitted and routed over three ODU-3 (40 Gb/s) containers and consider different static traffic demand matrixes. For each traffic matrix only a fraction of the total node pairs, varying from 20% to 100%, exchanges 100-GbE flows. The running time of the RDC framework is limited to 1 hour whereas the ILP results correspond to an execution time of one week (168 hours). For each node pair, 10 candidate paths are pre-computed with a k-shortest paths algorithm.

The first set of results reflect the performance of the ILP and the iterated search algorithm applied in stage 1 of RDC regarding the minimization of the maximum link capacity. Fig. 3(c) shows that the performance of the heuristic approach is the same as that with the ILP for EON but not for NSFNET. Although RDC is still able to attain gains over CR, in less connected topologies such as NSFNET it is frequently surpassed by the ILP. These gains may vary from 13% to 3%, in EON, and between 11% and 3% for NSFNET. The total accumulated differential delay, i.e. the sum of the differential delay affecting all virtually concatenated ODUs, resulting from the path selection in stage 1 is also compared in Fig. 4. In EON, RDC substantially outperforms the ILP, even with the inclusion of additional intermediate compensation nodes. However, in NSFNET the difference between both methods is less significant and, for the fully loaded network scenario, the ILP can outmatch RDC.

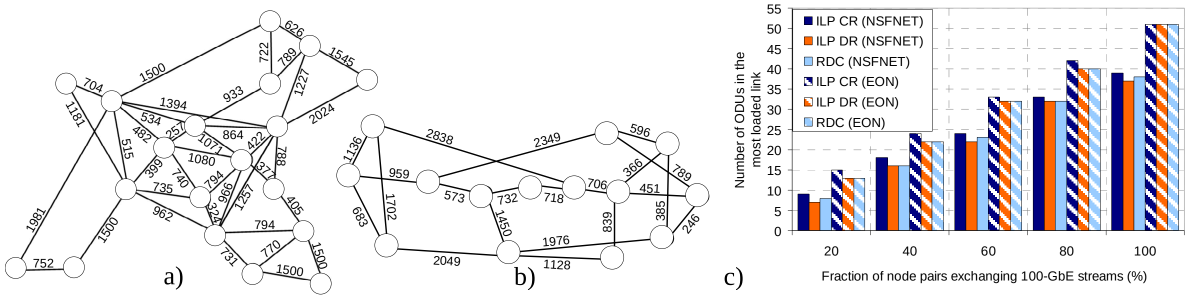

Fig. 3. (a) EON and (b) NSFNET topologies (distances in km); (c) Maximum link capacity requirements for NSFNET and EON.

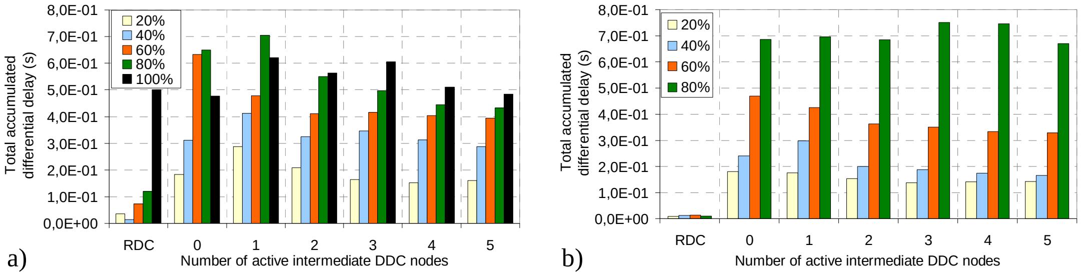

Fig. 4. Accumulated differential delay obtained with ILP and RDC in (a) NSFNET and (b) EON for different network loads.

The final outcome of the RDC framework is characterized by the maximum DDC buffer size per node. As shown in Fig. 5, the trends observed in Fig. 4 are approximately maintained, i.e., RDC is able to further reduce the buffer size obtained by the time-bounded ILP in the majority of the cases studied. In EON, this reduction varies between 40,4% and 87,7%. In NSFNET, RDC improves the ILP outcome by 9,8% to 88,6% for network loads up to 80%. For a 100% load, the ILP always offers smaller buffer dimensions. From Fig. 4 and 5, we can conclude that stage 1 is critical for the success of our heuristic approach, as the minimization of the maximum buffer size is closely dependent on the total accumulated differential delay. Additionally, we also demonstrate that even with long execution times and the incorporation of intermediate DDC nodes, the ILP often delivers sub-optimal solutions.

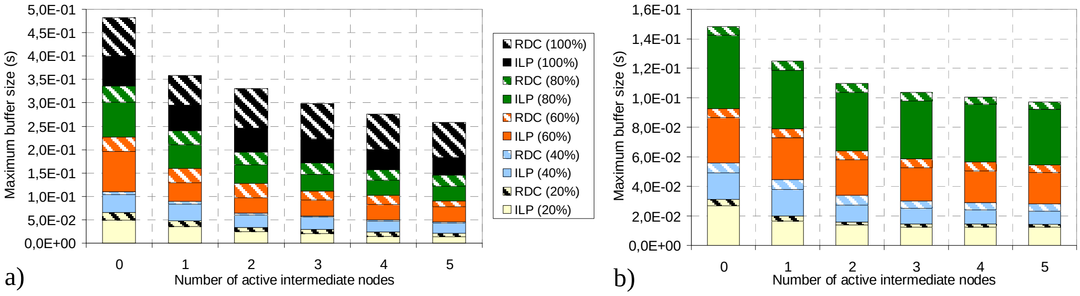

Fig. 5. Maximum DDC buffer size per node obtained with ILP and RDC in (a) NSFNET and (b) EON.

4. Conclusion

We have proposed and evaluated a multi-stage heuristic optimization framework applied to the routing of inverse-multiplexed 100-GbE flows over OTN with distributed differential delay architectures. Within a running time of only one hour, our approach is shown to outperform the equivalent time-constrained ILP model by further reducing the maximum buffer size and maintaining the link capacity gains, namely in highly connected topologies. In sparser networks, our algorithm slightly degrades the capacity requirements but still surpasses the co-routed scheme.

5. References

[1] S. Trowbridge, “High-Speed Ethernet Transport”, IEEE Communications Magazine, vol. 45, no. 12, pp. 120–125, 2007.
[2] ITU-T G.709, Interfaces for the Optical Transport Network (OTN), 2003.
[3] J. Santos, J. Pedro, P. Monteiro, and J. Pires, “Optimized ODU Routing for 100 Gb/s Ethernet over High-Speed Optical Networks with Distributed Differential Delay Compensation”, ECOC, 2009.
[4] M. Alicherry, C. Phadke, and V. Poosala, “Delay Distributed VCAT for Efficient Data-Optical Transport”, OFC, paper OThS2, 2005.
[5] A. Srivastana, *et al.*, “Differential Delay Aware Routing for Ethernet over SONET/SDH”, pp.1117-1127, INFOCOM, 2005.