

Management-Augmented Stateless PCE for Wavelength Routed Optical Networks

Piero Castoldi, *Member, IEEE*, Francesco Paolucci, Alessio Giorgetti, *Filippo Cugini
Luca Valcarengi *Member, IEEE*, °Shaowei Huang, °Ken-ichi Kitayama, *Fellow, IEEE*
Scuola Superiore Sant'Anna, Via Moruzzi 1, 56124 Pisa, Italy, e-mail: piero.castoldi@sssup.it
* *CNIT, Via Moruzzi 1, 56124 Pisa, Italy*
° *Osaka University, Osaka, Japan*

ABSTRACT

Two stateless Path Computation Element (PCE) implementations to be used in GMPLS-based Wavelength Routed Optical Networks (WRONs) are discussed (Fully Stateless PCE, F-SL-PCE, and Contention Avoidance Stateless PCE, CA-SL-PCE). The proposed CA-SL-PCE implementation exploits static management information for avoiding resource contentions among subsequent lightpath requests, while preserving network scalability. Simulation results show the capability of CA-SL-PCE implementation in strongly reducing resource contentions, thus resulting in a significant reduction of the overall blocking probability.

Keywords: PCE, stateless, contention avoidance, management plane.

1. INTRODUCTION

Wavelength Routed Optical Networks (WRONs) consist of optical nodes (e.g., optical cross-connects) interconnected by wavelength-division multiplexed (WDM) links. In such networks, the path computation has to take into account not only bandwidth availability but also other WRON-specific constraints (e.g., wavelength continuity).

The Generalized Multi-Protocol Label Switching (GMPLS) control plane [1] can be applied to WRONs for dynamically establishing lightpath requests in the optical domain. In particular, in the current implementation of GMPLS-based WRONs the Routing and Wavelength Assignment (RWA) is supported in a fully distributed way by exploiting a two-step procedure. Routing (i.e., route computation) is performed at source node by resorting to the locally available network status information. While wavelength assignment is performed during the signaling phase of the lightpath set up by resorting to specific RSVP-TE objects (e.g., Label Set) [2] and mechanisms (e.g., Crankback) [3]. However, this fully distributed (FD) architecture has some important drawbacks since, in dynamic traffic conditions, network status information are difficult to be kept updated at every network node. As a consequence subsequent lightpath requests may contend for the same network resources (i.e., resource contentions may arise) resulting in an degraded network resource utilization [4].

To overcome the above problem, a centralized network element, i.e., the Path Computation Element (PCE) [5], has been recently proposed to be used in GMPLS-based WRONs for both performing an effective centralized routing and avoiding the utilization of complex routing modules in network nodes. Two types of PCE architectures are currently under investigation in the standardization bodies: Stateful PCE and Stateless PCE. In the Stateful PCE (SF-PCE) architecture, the PCE is responsible for handling detailed wavelength availability information and the exact route of each lightpath established in the network. The PCE could resort to both extended OSPF-TE implementations [6] and the network management plane information [7] for collecting a vision of the current network status. Although this architecture may potentially provide optimal TE solutions, it may suffer from scalability issues overloading the control plane with a huge amount of messages necessary for collecting detailed network status information. On the contrary, the alternative Stateless PCE architecture (SL-PCE) represents a promising tradeoff between effectiveness and scalability. On the one hand, in this architecture the PCE simply relies on the aggregated bandwidth availability information advertised by the standard OSPF-TE, thus preserving network scalability. On the other hand, the centralized routing approach can be exploited for providing effective TE solutions.

This paper compares two feasible implementations of the SL-PCE architecture. In the first one, named Fully Stateless PCE (F-SL-PCE), the PCE simply applies a least congested routing scheme exploiting the aggregated bandwidth availability information provided by the standard OSPF-TE protocol. The second implementation, named Contention-Avoidance Stateless PCE (CA-SL-PCE), extends F-SL-PCE by resorting to additional static information provided by the network management plane and utilized by the PCE for estimating the lightpath setup time between every source-destination pair. The estimated lightpath setup times are then utilized for avoiding resource contentions among subsequent lightpath requests.

Simulation results show that the proposed CA-SL-PCE implementation is able to considerably reduce the overall lightpath setup blocking in several traffic scenarios.

2. FULLY STATELESS PCE (F-SL-PCE) IMPLEMENTATION

The F-SL-PCE implementation has been proposed with the main objective of preserving network scalability [5]. The introduction of the F-SL-PCE allows to move the route computation from source nodes (as in the FD architecture) to the PCE. In particular, upon a lightpath request the source node, also referred to as Path Computation Client (PCC), resorts to the PCE Protocol (PCEP) for sending a PCReq message to the PCE specifying the destination node of the requested lightpath. After path computation the PCE replies with a PCRep message which includes in the Explicit Routing Object (ERO) the strict route towards destination. However, since the PCE only relies on the same information utilized in the FD architecture, it is not able to improve the network utilization.

The path computation in the F-SL-PCE implementation relies on the aggregated bandwidth availability information advertised by the OSPF-TE routing protocol. In particular, besides the network topology, the PCE Traffic Engineering Database (TED) contains the number of available wavelengths on each WDM link. Upon reception of a PCReq message requiring a lightpath from s to d , the PCE selects the path with the largest number of available wavelengths on its most congested link, within the set $P_{s,d}$ of candidate paths from s to d .

The signaling protocol (RSVP-TE) is responsible for reserving the wavelengths along the links of the computed path. In particular, upon reception of the PCRep message, the source node sends a Path message towards the destination. Path messages include the ERO and the Label Set object that gathers information about the current status (i.e., available or reserved) of the wavelengths along the traversed links. At destination, the Label Set contains the available wavelengths along the end-to-end path. The destination node selects one of the wavelengths in the received Label Set using a specific wavelength assignment strategy (e.g., first-fit). Thus routing is performed at PCE, while wavelength-assignment is performed at destination by utilizing the wavelength availability information gathered by the Path message. After wavelength selection, the destination node sends a Resv message back to the source for reserving the selected wavelength on each traversed link.

An RSVP-TE instance may be blocked in the forward direction i.e., during the Path message propagation from source to destination node, or in the backward direction i.e., during the Resv message propagation from destination to source node. *Forward blocking* is due to insufficient network capacity, i.e., when in an intermediate node the intersection operation between the received Label Set and the set of available wavelengths on the outgoing link results in an empty set. *Backward blocking* is due to outdated information, i.e., when the selected wavelength, that the Path message detected as available, has been reserved on a link by some other ongoing RSVP-TE instances and the wavelength availability state changed by the time the Resv message arrived to that link. This specific occurrence is defined as resource contention.

3. CONTENTION AVOIDANCE STATELESS PCE (CA-SL-PCE) IMPLEMENTATION

In the proposed CA-SL-PCE implementation the OSPF-TE routing protocol operates, as in F-SL-PCE, announcing only the aggregated bandwidth availability information. However, the PCE also resorts to some static information obtained from the management plane.

Besides the presence of the standard TED (containing network topology and the number of available wavelengths along each WDM link), an additional Management-TED (M-TED) is introduced in the PCE to store the information provided by the management plane. In particular, the proposed implementation requires to know the WDM links physical length, the bit-rate of the control plane interfaces and an estimation of the processing time of RSVP-TE and PCEP messages. This information is directly retrieved from the management plane during network bootstrap or when new links, nodes and interfaces are installed in the network. The information collected from the management plane is static and therefore do not impact the network scalability. The PCE uses the M-TED for estimating the setup time of each requested lightpath. The estimated values are then utilized for reducing backward blocking implementing a dynamic path computation in order to prevent resource contentions.

The lightpath setup procedure consists of the two steps detailed in Fig.1. The first step comprises the path computation request phase: the source node asks the PCE for a path towards a destination node by resorting to the PCReq and PCRep message exchange. The time needed for this operation is the round trip time between the source node and the PCE, i.e., $t_2 - t_0 = RTT_{PCE}$. Thus, the RSVP signaling is triggered towards the destination node and the time needed is the round trip time between the source node and the destination node, i.e., $t_4 - t_2 = RTT_{SD}$.

Therefore, when the PCE receives the PCReq message at time t_1 , it is able to estimate the time in which the lightpath establishment will be completed, i.e., $t_4 = t_1 + \frac{1}{2} RTT_{PCE} + RTT_{SD}$. In order to avoid resource contentions the PCE stores and keeps updated the *Contention Table (CT)* containing an entry for each lightpath currently under establishment.

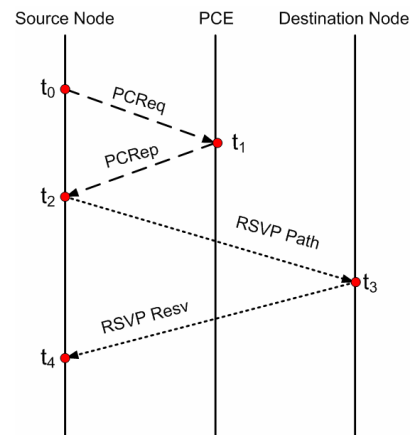


Fig. 1. Path Computation and Signaling Timeline.

The i -th entry of CT contains the PCReq reception timestamp t_1^i , the ERO included in the PCRep message, and the estimated time t_4^i .

When a new lightpath request j arrives at the PCE at time t_1^j each entry i of the CT is checked:

- If $t_4^i \geq t_1^j + \frac{1}{2} RTT_{PCE}^j$ then the RSVP-TE signaling phase of lightpath j will start before the end of the signaling phase of lightpath i . Lightpaths i and j are therefore potentially contending and all the links used by lightpath i are marked as critical.
- If $t_4^i < t_1^j + \frac{1}{2} RTT_{PCE}^j$ then the RSVP-TE signaling phase of lightpath j will start after the end of the signaling phase of lightpath i . Lightpaths i and j are therefore not potentially contending.
- If t_4^i is elapsed then the entry i is removed from the CT.

The path computation is then performed taking into account the list of critical links. In particular, upon reception of a PCReq message, the network topology contained in the TED is modified by increasing the routing cost of the identified critical links. With these modified costs the set of candidate paths $P_{s,d}$ from s to d is dynamically computed. Then the path with the largest number of available wavelengths on its most congested link is selected within $P_{s,d}$. After path computation, the CT is updated by inserting a new entry for the lightpath j . The average dimension of the CT depends on the network traffic load (i.e., the rate of incoming connections requests) and is limited by the number of incoming lightpath requests within few hundreds of milliseconds, resulting in an acceptable computational load at the PCE.

The resource reservation utilized for the CA-SL-PCE implementation is the same RSVP-TE based reservation adopted for the F-SL-PCE.

4. SIMULATION RESULTS

The F-SL-PCE and the CA-SL-PCE implementations are compared by means of a custom built C++ event driven simulator. The simulation is carried out on a Pan European network topology (Fig. 2) with 27 optical nodes and 55 WDM bidirectional links. Each link has 32 wavelengths per direction. The traffic is uniformly distributed among all the node pairs, lightpath requests arrive following a Poisson process with an exponentially distributed inter-arrival time (with average $1/\lambda$ seconds) and holding time (with average $1/\mu$ seconds). The load offered to the network is therefore expressed in Erlang as λ/μ . In the F-SL-PCE implementation, the set of candidate paths $P_{s,d}$ includes all the paths from node s to node d whose hop length is within one hop from the shortest path. For uniformity in the CA-SL-PCE implementation $P_{s,d}$ includes all the paths from node s to node d whose cost is within one unit from the least cost path. In this case the link cost is equal to 1 for not critical links while it is set to 10 for critical links. First-fit wavelength assignment is used in both implementations at the destination node within the received Label Set. In all the figures the simulation points are plotted with the confidence interval at 90% of confidence level.

Fig. 3 shows the overall blocking probability and its two contributions (i.e., forward and backward blocking) as a function of the network load. The mean duration time has been fixed to 60s while the mean inter-arrival time is varied for achieving the desired traffic loads. The plot shows that in the F-SL-PCE the backward blocking probability is almost a constant contribution around 10^{-4} . This contribution dominates the forward blocking for low traffic load. On the contrary, the proposed CA-SL-PCE implementation is able to completely avoid resource contentions achieving a null backward blocking. Fig. 4 and Fig. 5 show the overall blocking probability and its two contributions as a function of the mean inter-arrival time. Fig. 4 has been obtained with a constant traffic load of 250 Erlang, where forward blocking is null (see Fig. 3) and the only component of the overall blocking is the backward one.

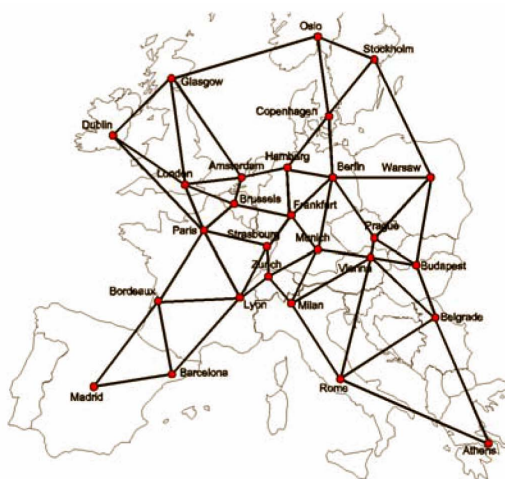


Fig. 2. Pan European Network.

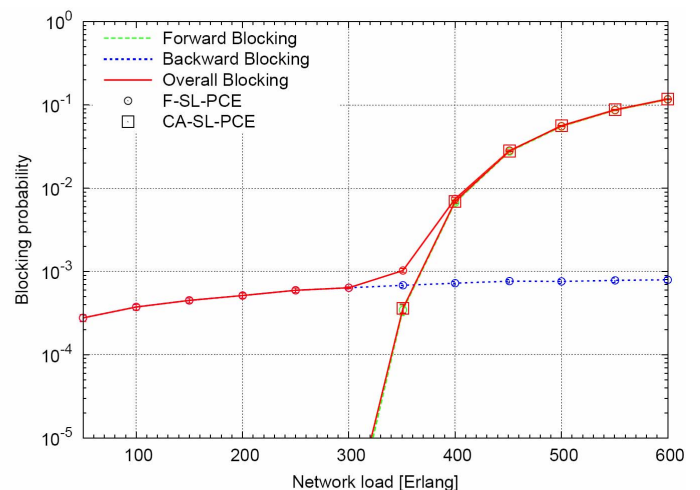


Fig. 3. Blocking probability in function of the network load.

In this case the CA-SL-PCE implementation is able to strongly reduce the overall blocking by avoiding resource contention also when arrivals are frequent. However, for very frequent arrivals (i.e., $1/\lambda < 10^{-3}$ s), the backward blocking obtained by CA-SL-PCE is similar to the one obtained by F-SL-PCE because in such dynamic condition numerous lightpath requests arrive in a very short period of time, thus a great number of links are identified as critical. In this situation it may be impossible to find a path toward destination not including critical links. Fig. 5 has been obtained with a constant traffic load of 350 Erlang, where forward blocking is not negligible. In particular, while the backward blocking is strongly reduced, the forward blocking is worsened by the proposed CA-SL-PCE in dynamic traffic conditions (i.e., $1/\lambda < 10^{-1}$ s). The reason is that, in such dynamic conditions, a wide set of links are identified as critical. Therefore, for avoiding critical links, the CA-SL-PCE implementation computes longer routes than the ones obtained with the F-SL-PCE implementation, resulting in inefficient network resource utilization. However, the overall blocking benefits from the CA-SL-PCE implementation since the backward blocking improvement is more relevant than the forward blocking degradation.

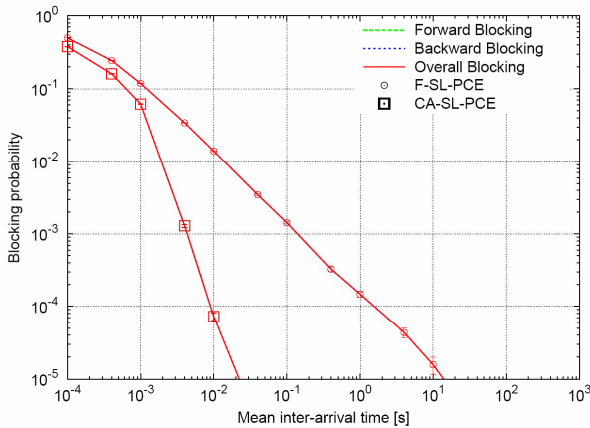


Fig. 4. Blocking probability in function of the inter-arrival time (constant load=250 erlang).

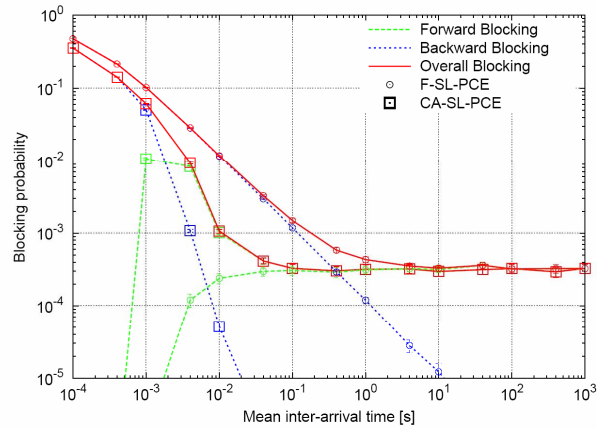


Fig. 5. Blocking probability in function of the inter-arrival time (constant load=350 erlang).

5. CONCLUSIONS

In this paper the Contention Avoidance Stateless PCE (CA-SL-PCE) implementation is proposed for improving the Fully Stateless PCE (F-SL-PCE) implementation. CA-SL-PCE resorts to additional static information provided by the management plane, and implements a contention avoidance mechanism aiming at preventing resource contentions, while using the same information utilized in F-SL-PCE. Extensive simulations showed the effectiveness of the proposed solution in several traffic scenarios. The CA-SL-PCE implementation demonstrates the possibility to improve the stateless PCE architecture without affecting the network scalability in terms of control plane load and information to be handled by the PCE.

ACKNOWLEDGEMENTS

The work described in this paper was carried out with the support of the BONE-project ("Building the Future Optical Network in Europe"), a Network of Excellence funded by the European Commission through the 7th ICT-Framework Programme and support of FIRB project "Software and Communication Platforms for High Performance Collaborative GRID (RBin043TKY)".

REFERENCES

- [1] E. Mannie, "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", *RFC 3945*, October 2004.
- [2] L. Berger, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling RSVP-TE Extensions", *RFC 3473*, January 2003.
- [3] A. Farrel et al., "Crankback signalling extensions for MPLS and GMPLS RSVP-TE", *RFC 4920*, July 2007.
- [4] A. Giorgetti, N. Sambo, I. Cerutti, N. Andriolli, P. Castoldi, "Label preference scheme for lightpath provisioning and restoration in distributed GMPLS networks", *J. Lightw. Technol.*, accepted for publication.
- [5] A. Farrel et al., "A Path Computation Element (PCE)-based architecture", *RFC 4655*, August 2006.
- [6] A. Giorgetti, N. Sambo, I. Cerutti, P. Castoldi, "Impact of Link-State Advertisement in GMPLS-based wavelength routed networks", *JWA98, OFC 2008*.
- [7] T. Tsuritani, M. Miyazawa, S. Kashihara, T. Otani, "Optical Path Computation Element interworking with Network Management System for Transparent Mesh Networks", *NWF5, OFC 2008*.