Scalable Hybrid Path Computation Procedure for PCE-Based Multi-Domain WSON Networks

G. Hernández-Sola, J. Perelló, F. Agraz, S. Spadaro, J. Comellas, G. Junyent
Advanced Broadband Communications Center (CCABA), Universitat Politècnica de Catalunya (UPC)
Jordi Girona, 1-3, 08034, Barcelona (Spain)
Tel: (+34) 93 4010966, Fax: (+34) 93 4017200, e-mail: guillem.hernandez@tsc.upc.edu

ABSTRACT
In the Hierarchical Path Computation Element (H-PCE) architecture, an optimum end-to-end path is computed using a hierarchical relationship among PCEs. An H-PCE-based path computation procedure provides low network blocking probability meanwhile it results on a great amount of control overhead messages. In this paper we propose a hybrid path computation procedure based on the H-PCE architecture and the Backward Recursive PCE-Based Computation (BRPC). Extensive simulation results show that the proposed approach performs better than H-PCE in terms of network control overhead.
Keywords: Multi-domain, path computation element, hierarchical path computation.

1. INTRODUCTION

Lately research efforts related to optical transport network infrastructures have been mainly focused on single-domain scenarios, where scalability and confidentiality do not represent a problem. Nonetheless, the next generation Wavelength Switched Optical Networks (WSONs) will comprise several domains, each one managed by a different service provider. From the routing point of view, a domain is a collection of network elements within a common address management or path computational responsibility. Compute a path between nodes within a single domain is reasonably simple. However, computing an end-to-end path where the source and destination nodes belong to different domains requires cooperation between the computational entities, being each one responsible for its correspondent domain.

The PCE has been proposed as a network entity able to calculate end-to-end routes with computational constraints in both single and multi-domain networks [1]. In a single domain scenario, the PCE computes optimal paths using its Traffic Engineering Database (TED), which is generally constructed by means of the network routing protocol (i.e., OSPF-TE [2] in GMPLS [3]). Nonetheless, in multi-domain network scenarios, the information exchange between PCEs is normally reduced, only including the shared links and border nodes information. Moreover, the lack of knowledge of the general domain mesh connectivity bounds the PCE capabilities in order to calculate efficient inter-domain end-to-end paths. The IETF has proposed two general approaches for the inter-domain path computation (i.e., per-domain [4] and Backward Recursive PCE-based Computation (BRPC) [5]). In the former, the entry border node of each domain requests to its responsible PCE the computation of the correspondent piece of Label Switched Path (LSP) and the final end-to-end LSP is a concatenation of local intra-domain route segments. In the latter, the PCEs collaborate to compute the end-to-end LSPs through a domain sequence previously predetermined. Both procedures drive the route computation to be sub-optimal mainly because only one sequence of domains is used to collect routing information to compute the end-to-end LSP. Notice that when it is possible to reach a destination node using different domain sequences from a source node, the assumption of a predetermined domain path drastically hinders the network performance of the end-to-end inter-domain computation procedure.

The IETF has also introduced the concept of the Hierarchical PCE (H-PCE) architecture [6], showing how to coordinate several PCEs in order to collect information from the whole set of domains in a network and then derives an optimal end-to-end path without assuming a predetermined domain sequence. The idea of hierarchy is defined to control and manage a group of network entities called child PCEs interconnected to a higher hierarchically level PCE, called the Parent PCE. Specifically, this architecture defines the interfaces, functionalities and the end-to-end path procedures at each hierarchical level using a client-server architecture. This hierarchical coordination is based on the Path Computation Element Protocol (PCEP) [7] which basically uses two different types of control messages, namely, Path Computation Request (*PCReq*) and Path Computation Reply (*PCRep*). H-PCE provides low connection blocking probability at expenses of a huge amount of control overhead messages. However, in H-PCE, for each end-to-end LSP computation, updated route information to the complete set of PCEs (domains) is requested, thus arising scalability problems when increasing the number of domains.

In this paper, a hybrid path computation procedure is proposed. The aim is to maintain the low blocking probability provided by the H-PCE approach but drastically reducing the generated network overhead (control messages required for the LSPs establishment). Additionally, a trade-off between shared routing information, scalability and security across domains is also obtained.

978-1-4577-0882-4/11/$26.00

The rest of this paper is organized as follows. Section 2 describes the H-PCE architecture and its limitations. Our proposed mechanism; the hybrid path computation procedure is described in Section 3. The performance studies are highlighted in Section 4. Finally, some conclusions are discussed in Section 5.

2. HIERARCHICAL PCE PATH COMPUTATION

The selection of the domain sequence is essential to determine the optimal end-to-end path in multi-domain networks. Authors in [8] have proposed an enhancement of the BRPC procedure based on using different end-to-end domain-disjoint sequences. They concluded that it becomes essential to gather routing information from different domain sequences in order to compute more accurate end-to-end inter-domain LSPs, resulting in reduced overall blocking probability. The Hierarchical PCE (H-PCE) architecture [6] assumes that the selection of a sequence of domains for an end-to-end path is basically a hierarchical path computation problem. Namely, one mechanism is used to determine a path across a domain (an intra-domain path computation), and a separate mechanism is used to determine the sequence of domains to be traversed by the end-to-end connections. In H-PCE, a Parent PCE maintains a domain topology map that contains the child domains (seen as vertexes in the topology) and their interconnections (edges in the topology). Each domain has at least one PCE (child PCE) capable of computing paths across the domain, and at the same time, it is managed by the Parent PCE. Fig. 1 (left) shows the physical topology of a multi-domain network and Fig. 1 (right) shows the domain topology map which contains the inter-domain links and the border nodes from each domain. In such architecture, the Parent PCE knows the identity and location of the child PCEs responsible for the child domains. To maintain domain confidentiality, the Parent PCE is aware of the topology and connections between domains, but is not aware of the topology of the child domains. On the other hand, each child PCE does not know the topology of the other domains.

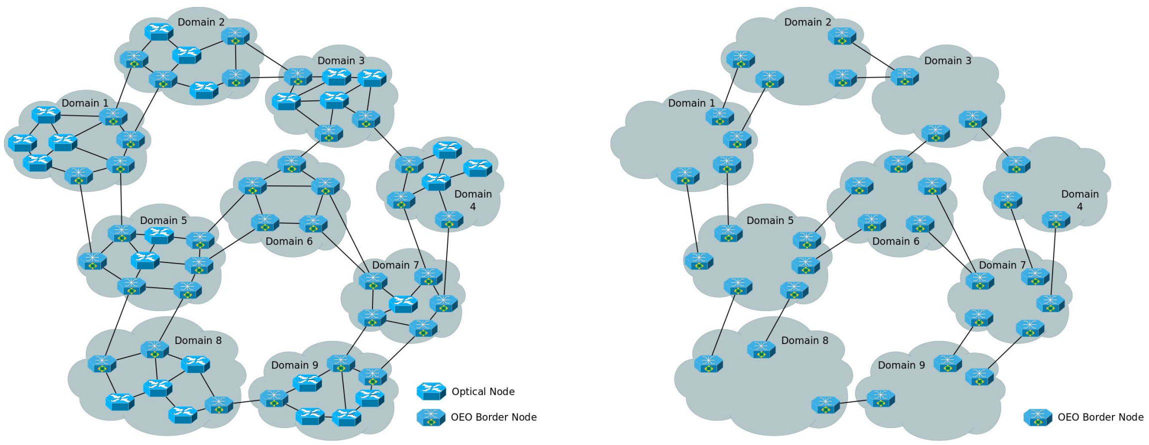

Figure 1. An example of a physical transport topology (left); the domain map managed by the Parent PCE used to calculate the domain sequences (right).

2.1 Scalability and path computation accuracy limitations

H-PCE assumes that if no restricting policies are applied, upon a LSP request, the Parent PCE requests route information to its entire group of child PCEs in order to compute an optimal path for the end-to-end LSP. Note that, for example, when a child PCE is requested for a path where the destination node is allocated in its adjacent domain, it seems unnecessary to request the domain sequence to its Parent PCE or gather routing information from the most geographically separated child PCEs in the network. Additionally, as a multi-domain network increases in size and complexity, it becomes necessary to introduce more accurate policies to reduce the amount of network control overhead messages. Several approaches in order to enhance the H-PCE architecture have been carried out. An experimental H-PCE test-bed using only one domain sequence selection policy is analyzed in [9]. The authors conclude that its solution is sub-optimal due to not having complex TE aggregation schemes. Another approach aiming to compute optimal routes in H-PCE architecture is presented in [10]. This work describes two lightweight path computation schemes based on H-PCE [6]. In the first scheme, the Parent PCE only provides a unique sequence of domains. In the second scheme, the Parent PCE provides a sequence of the domains and the list of border nodes that should be traversed by the inter-domain LSPs. As expected, the authors in [10] conclude that the second scheme, the one with more multi-domain routing information to compute the path provides better connection blocking probability. However, both proposals only collect network state information from the child domains of a single domain sequence. This assumption drives the proposed approaches to discard valuable routing information in order to calculate an accurate end-to-end path.

3. PROPOSED MECHANISM

This section describes the Hybrid Path Computation (HPC), a proposal focused on the reduction of the number of control overhead messages for each computed end-to-end inter-domain LSP while maintaining low connection blocking probability. Our proposed procedure assumes that a child PCE executes BRPC as the default path computation procedure; the pre-determined domain sequence is statically provided by the Border Gateway Protocol (BGP) [11]. Nonetheless, once a child PCE is not able to provide a proper end-to-end inter-domain path using BRPC or the number of domains to be crossed surpasses a pre-defined threshold *k*, the child PCE delegates to the Parent PCE the path computation, and the standard H-PCE procedure is thus applied. In standard H-PCE, a child PCE is requested to compute two different intra-domain computations, namely, a node-to-edge or an edge-to-edge path segment computation. In the former, a child PCE provides a set of path segments from the requested node to each border node in the domain. In the latter, a child PCE builds a complete graph of virtual links between all its border nodes (i.e., a full-mesh abstraction). Both routing information sets are forwarded to the Parent PCE. It is assumed here that each child PCE is managed by a Parent PCE with the capacity to compute an inter-domain end-to-end LSP concatenating the collected routing information from its child domains and the global domain topology map. Note that HPC supposes that a shortest domain path in the topology map is more probable to be the optimal one to allocate the requested end-to-end path. Furthermore, it is also assumed that a longer domain sequence provides a sub-optimal end-to-end route essentially given that it is more likely to dismiss appropriate routing information. As the number of domain hops increases, it becomes more difficult to select the most appropriate sequence to allocate an end-to-end path and for this reason it is more suitable to execute H-PCE. Our hybrid path computation approach enhances the child PCEs with the capacity to choose the computation procedure between BRPC and H-PCE depending on a formerly established number of domains in the sequence to be traversed by an end-to-end path.

4. PERFORMANCE EVALUATION

This section evaluates the performance of the hybrid path computation scheme by running OMNeT++ discrete event simulations in a 9-domain optical network. Figure 1(left) depicts the network topology which is composed of 61 nodes and 95 links (19 inter-domain) carrying each one 8 wavelengths per link. Figure 1 (right) shows the domain topology map managed by the Parent PCE. This abstracted topology map is composed by 19 inter-domain links and 36 border nodes. In all simulations, 10^5 LSP/connection requests are generated following 70/30% intra/inter-domain ratio. For inter-domain connections, source and destination domains are uniformly selected and source/destination nodes are uniformly chosen in their respective domains. Source and destination nodes are randomly selected for intra-domain connections. All generated requests demand a whole wavelength capacity and mean holding time (HT) is set to 600 seconds following an exponential distribution. Request inter-arrival time is also exponentially distributed and varies with the network offered load.

Figure 2 depicts the overall network connection blocking probability (B_p) achieved by the H-PCE, the HPC (with $k = 2, 3$) and BRPC respectively, as a function of the global offered load. For the sake of clarity, by using the HPC procedure with $k = 2$, BRPC is applied when the destination node is allocated in an adjacent domain while the HPC procedure with $k = 3$ uses BRPC whether the domains sequence includes 2 transit domains. BRPC presents the worst B_p because only collects routing information a pre-configured domain sequence meanwhile H-PCE presents the best B_p because it computes an end-to-end path using all the gathered routing information from the child PCEs. For low offered loads, HPC generates a similar B_p in comparison to H-PCE. HPC uses the enhanced child PCE capacity to switch between BRPC and H-PCE, providing a not significant increase of the B_p for higher offered loads. In particular, for 200 Erlang, HPC with $k = 2$ performs a B_p around 1%, being slightly higher than H-PCE.

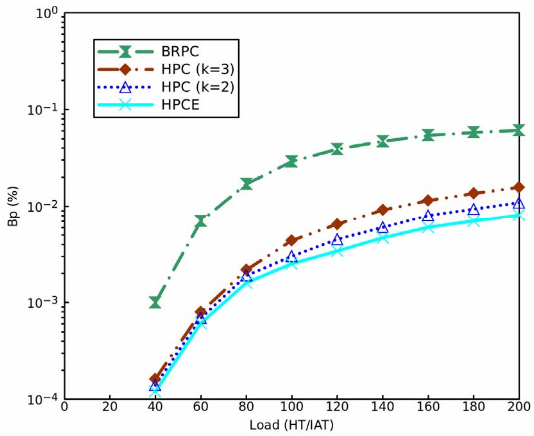

Figure 2. Connection blocking probability in a BRPC, HPC (k = 2, 3) and HPCE.

Table 1. Relative overhead reduction in comparison to H-PCE (%).

HPC (*k*=2)	72 %
HPC (*k*=3)	39 %
BRPC	20 %

Table 2. Parent PCE utilization (%).

	Parent PCE utilization (%)
HPC (*k*=2)	72 %
HPC (*k*=3)	31 %

Figure 2 also depicts that in some cases, executing BRPC using the domain sequence provided by the Parent PCE is enough to perfectly allocate an end-to-end inter-domain path. For the least loaded network scenarios, it seems the most suitable path computation method because the generated network overhead is lower and the performed B_p is acceptable. Table 1 analyzes the overhead reduction deploying the hybrid approach. The HPC procedure with $k = 2$ generates a reduction of the network control overhead around the 72%. Moreover, HPC with $k = 3$ improves around the 75% the B_p performance in comparison to BRPC only increasing by the 20% the network control overhead messages. Table 2 depicts the utilization of the Parent PCE by applying the HPC. Specifically, running the HPC procedure with $k = 3$ one third of the end-to-end inter-domain path computations requires the utilization of the Parent PCE. If HPC with $k = 2$ is run, the Parent PCE does not intervene in one fourth of the end-to-end path computations. Therefore, our proposed method reduces the Parent PCE computational load increasing the cooperation between the child PCEs.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed a hybrid path computation procedure based on the length of the domain sequence to be traversed by the end-to-end LSPs. To compute optimal paths in a multi-domain network, the exchange of some routing information of each domain is required, at expenses to increase the network control overhead among the PCEs. The proposed inter-domain path computation approach allows a drastically reduction of these control messages while keeping the connection blocking probability at the same level of H-PCE. From the simulation results, it can be concluded that HPC is the best solution to reach a trade-off between network control overhead (and thus routing scalability) and connection blocking probability. Simultaneously, it allows an efficient way to fulfill the confidentiality concerns among independent domains. Future efforts will be focused on the design of more advanced hybrid path computation policies, able to provide even more suitable solutions in terms of blocking probability and network scalability.

ACKNOWLEDGEMENTS

The work reported in this paper has been partially supported by the UPC through a FPI-UPC research scholarship grant and the Spanish Science Ministry through Project "Engineering Next Generation Optical Transport Networks (ENGINE)", (TEC2008-02634).

REFERENCES

[1] A. Farrel, JP. Vasseur, J. Ash, "A path computation element (PCE)-based architecture", RFC 4655, Aug. 2006.

[2] Katz, D., Kompella, K., and D. Yeung, "Traffic engineering (TE) extensions to OSPF version 2", RFC 3630, Sept. 2003.

[3] E. Mannie, "Generalized multi-protocol label switching (GMPLS) architecture", RFC 3945, Oct. 2004.

[4] JP. Vasseur, A. Ayyangar, R. Zhang,"A per-domain path computation method for establishing inter-domain traffic engineering (TE) label switched paths (LSPs)", RFC 5152, Feb. 2008.

[5] JP. Vasseur, R. Zhang, V. Bitar, JL. Le Roux, "A backward-recursive PCE-based computation (BRPC) procedure to compute shortest constrained inter-domain traffic engineering label switched paths", RFC 5441, Apr. 2009.

[6] D. King, A. Farrel, "The application of the PCE architecture to the determination of a sequence of domains in MPLS & GMPLS", IETF draft draft-king-pce-hierarchy-fwk-05.txt, Sept. 2010. Work in progress.

[7] JP. Vasseur, JL. Le Roux, "Path computation element (PCE) communication protocol (PCEP)", RFC 5440, Mar. 2009.

[8] G. Hernández-Sola, J. Perelló, F. Agraz, L. Velasco, S. Spadaro, G. Junyent, "Enhanced domain disjoint backward recursive TE path computation for PCE-based multi-domain networks", *Photonic Network Communications*, vol. 21, no. 2, 141-151, DOI: 10.1007/s11107-010-0288-5.

[9] R. Casellas, R. Muñoz, R. Martinez, "Lab trial of multi-domain path computation in GMPLS controlled WSON using a hierarchical PCE," in *Proc. OFC and NFOEC*, Mar. 6-10, 2011, Los Angeles, California, USA.

[10] A. Giorgetti, F. Paolucci, F. Cugini, and P. Castoldi, "Hierarchical PCE in GMPLS-based Multi-Domain Wavelength Switched Optical Networks," in *Proc. OFC and NFOEC*, Mar. 6-10, 2011, Los Angeles, California, USA.

[11] Rekhter, Y., Li, T.: A border gateway protocol 4 (BGP-4), RFC 4271 (2006)