Handling Outdated Information in Optical PCE Environments


*Advanced Network Architectures Lab (CRAAX), Technical University of Catalonia (UPC), Spain
†Universidade Federal de Víncosa, vbarbosa@ac.upc.edu

Abstract—The rapid emergence of new network scenarios and architectures, such as Data Centers Networks (DCNs), Path Computation Element (PCE), and Software Defined Networking (SDN), has refreshed some on-line routing-related problems, objective of many research efforts in the past. As a result, new scalable and efficient path computation algorithms are required to address particular characteristics and demands of on-line scenarios, such as those brought by inaccurate Network State Information (NSI), strongly affecting the overall blocking probability. In this paper, we propose a prediction-based PCE scheme, referred to as PPCE. PPCE is devised for highly dynamic network scenarios, aiming at reducing the amount of signaling messages as well as the blocking probability.

I. INTRODUCTION AND MOTIVATION

At present, the design of carrier-grade networks follows a multi-layer model based on the convergence of IP/MPLS and Optical technologies, such as Wavelength Division Multiplexing (WDM). In multi-layer networks, an IP/MPLS (Virtual) layer is used for providing a variety of services, whereas WDM (transport) layer is used as a transport medium. It is common that the transport layer follows an all-optical communication model, where Wavelength Routed (WR) nodes are capable of switching data in the optical domain without Optical-Electrical-Optical (OEO) conversion at intermediate nodes. This feature is highly demanded in today’s optical networks in order to reduce both power consumption and communication latency. As a result, all-optical WDM networks are a widespread practice conversely to electrical and opaque networks. This is noticeable by the adoption of WRs in novel network scenarios, such as Data Center Network (DCNs) [1].

Novel network scenarios demand several features including fast connection-provisioning, recovery actions, and Traffic engineering (TE). These features are commonly achieved by source-based routing strategies using a distributed control plane scheme (e.g., GMPLS orASON). However, due to the ever-increase of traffic demand, sourced-based routing strategies present significant weaknesses when facing path computation actions. Specifically in large and highly dynamic network scenarios with stringent constraints. As a result, new network paradigms such as, the Path Computation Element (PCE) [2], and Software Defined Networking [3], have emerged with the aim of replacing the conventional distributed source-based path computation strategies.

A PCE scheme is a centralized routing architecture that decouples path computation actions from WRs with the aim of improving the overall routing performance and scalability, i.e., blocking probability and path computation time. In a similar manner to PCE schemes, the rationale behind SDN is to decouple control plane features, such as the set-up of lightpaths, resulting in lower provisioning times and augmented TE features.

Clearly the overall routing performance strongly depends on the accuracy of the Network State Information (NSI) used for lightpath computation. For example, the accuracy of the NSI is significantly important in order to meet the so-called Wavelength Continuity Constraint (WCC), which states that a lightpath can be solely established, if the same wavelength is available on the path selected from the source to the destination WR.

The negative effects of having inaccurate NSI is a well-known problem referred to as the Routing Inaccuracy (RI) problem. Many references may be found in traditional backbone networks based on distributed control planes (see for example [4]), largely assessing that an incorrect selection of wavelengths in optical scenarios, may undoubtedly increase the amount of blocked connections, which can be catastrophic in highly dynamic scenarios. In highly dynamic scenarios such as DCNs, where traffic demands are above 10 Gbps and lightpaths are set-up and torn-down in a short-term basis, it is a must to maintain low blocking probability with the minimum amount of signaling. The simplest, from a CAPEX and OPEX point of view and theoretically most efficient vehicle to offset the negative effects caused by an inaccurate NSI, is to decrease the time interval (updating time) to disseminate NSI. However, this solution yields a signaling overhead increment, which may lead to scalability issues, especially on highly dynamic network scenarios such as DCNs.

Moreover, there are two main facts motivating the need of novel research efforts in the domain of RI: 1) the RI effects in DCN scenarios are highly aggravating the routing performance, hence deserving special attention, and; 2) despite that a wealth of research efforts have been devoted to RI in distributed source-based routing scenarios, there are few studies addressing the RI problem considering both a centralized routing and control approach. It is important to remark that the huge amount of traffic in DCNs has a negative impact on the accuracy of the NSI managed by path computation algorithms. This argument converges on the increase of the signaling overhead.

This paper presents a centralized, high performance path
computation algorithm leveraging prediction concepts, as a solution to severely minimize the effects produced by the RI problem. We focus in dynamic optical network scenarios where connection requests arrive in a random manner and centralized control strategies such as the PCE and SDN based on OpenFlow are used for lightpath computation and provisioning.

With aim of providing realistic results, the proposed path computation algorithm is evaluated in two realistic network models: 1) a traditional optical backbone network design, and; 2) a DCN-like topology. The rationale driving this strategy is that the proposed solution is envisioned open enough to be deployed in different network scenarios, handling typical routing parameters but also the specific requirements, characteristics and functionalities coming from new network scenarios and control paradigms.

The rest of this paper is organized as follows. Section II presents in a nutshell several works related to the RI problem in the optical domain. Section III introduces the proposed scheme. Section IV presents the evaluation strategy and provides extensive simulation results of the proposed algorithm versus other path computation algorithms considering different updating times and traffic loads. Finally, Section V concludes the paper and suggests avenues for future work.

II. RELATED WORK

The design of new routing strategies, properly addressing the RI problem in new network scenarios, must leverage past contributions, what requires a comprehensive understanding of past proposals that may feed and fuel novel initiatives.

This section positions the RI problem in the literature and also introduces some seed studies related to the RI problem in optical networks. Even though there is a wealth of studies providing analytical and numerical results of distinct path computation algorithms, most of these studies consider accurate NSI, hence completely ignoring the RI problem -studies already available in the literature also refer to inaccurate NSI as imprecise, stale, delayed or outdated NSI.

A pioneering work related to the study of the RI problem in optical networks can be found in [5], where authors analyze the performance of conventional path computation algorithms under inaccurate NSI. This study successfully positions the RI problem in optical networks by effectively demonstrating that in highly dynamic network scenarios under inaccurate NSI, the performance of conventional path computation algorithms significantly decreases. This study also showed that path computation algorithms that are considered optimal under accurate NSI, conduct suboptimal performance in comparison with other schemes. Recent contributions dealing with the RI problem in optical networks can be found in [6], [7] and [8]. However, none of these studies consider to position the RI problem in a DCN scenario, taking into consideration topologies and traffic models common of these type of scenarios.

It is important to remark that the RI problem is not exclusive to distributed source-based routing architectures. It also affects centralized routing architectures such as the PCE. Notice that a PCE scheme follows two main approaches: 1) stateful PCE, and; 2) stateless PCE [2]. A stateful PCE remembers every path that is successfully set-up or tear-down in order to keep NSI related to wavelengths availability per optical link. This approach has several handicaps such as: 1) a high level of synchronization which might be highly complex to achieve with multiple PCEs, and; 2) latencies related to the control plane, e.g., setup delay as well as race condition might lead to inaccurate NSI. On the other hand, a stateless PCE does not keep state of the lightpaths that are set-up or torn-down; hence, NSI is kept on the so-called Traffic Engineering Database (TED). The TED is built by NSI that is periodically disseminated by routing mechanisms. In comparison with a stateful PCE scheme a stateless PCE has less complexity. However, the TED may contain outdated NSI in highly dynamic scenarios. Nevertheless, it is worth mentioning that the RI problem might still be present in stateful PCE scenarios, motivated by control plane delays and race-conditions caused by collaborative PCE schemes.

Authors in [9] propose a pre-reservation mechanisms to cope the RI problem in stateless and statefull PCE scenarios. However, this approach requires enhancements in both PCE and Path Computation Element Protocol (PCEP). Moreover, there are other available contributions that target the integration of SDN and PCE schemes such as [10] to improve the performance of stateless PCE schemes. Nevertheless, authors do not consider the RI problem. In light of this, we consider that there is a need for path computation mechanisms, devised for PCE schemes, that are no disruptive with current control architectures that can deal with negative effects of RI and adopting new control architectures. To fill this gap, this papers introduces a novel stateless PCE scheme called Predictive Path Computation Element (PPCE), which adopts Fine-Granularity Predictive (FGP) counters for modeling lightpaths availability that, unquestionably yields low blocking probability while keeping low signaling overhead. We do demonstrate that fine-granularity counters are more suitable to model lightpaths availability in comparison with Coarse-Granularity Predictive (CGP) counters, and that this strategy is more suitable to face the RI problem.

III. THE PREDICTIVE PATH COMPUTATION MECHANISM

This section is devoted to: i) justify the utilization of FGP counters (link/wavelength) instead of a CGP approach (route/wavelength) by highlighting its strengths and benefits; ii) present a comprehensive description of the proposed PCE scheme, namely PPCE, and; iii) provide a scalability analysis of the proposed PCE scheme that clearly validates its utilization in scenarios with outdated TED. For the sake of understanding, Table I lists the set of symbols and terminology used in this paper.

A. FGP Counters

The PPCE mechanism, as a concept, is proposed as a predictive path computation strategy that computes lightpaths based on historic information about lightpaths availability. PPCE is technically supported by the utilization of predictive counters. Indeed, PPCE uses two-bit FGP counters
Table I

<table>
<thead>
<tr>
<th>Symbols and Terminology</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>⌈ ⌉</td>
<td>The ceiling operator.</td>
</tr>
<tr>
<td>λ_n</td>
<td>An optical wavelength where n ∈ {1, ...,</td>
</tr>
<tr>
<td>W</td>
<td>The set of wavelengths available in any WR, where W ∈ {λ_1, λ_2, ..., λ_{</td>
</tr>
<tr>
<td>M_s,d</td>
<td>The set of candidates paths for a source-destination pair, where s, d ∈ V and s ≠ d, and V is the set of WRs.</td>
</tr>
<tr>
<td>C_{i,λ}</td>
<td>FGP counter where i ∈ E, λ ∈ W, and E is the set of optical links.</td>
</tr>
<tr>
<td>l_{j,λ}</td>
<td>Availability of a lightpath computed using route j and wavelength λ using FPG counters.</td>
</tr>
<tr>
<td>N_j</td>
<td>Is the number of links of route j.</td>
</tr>
<tr>
<td>W_i</td>
<td>Available wavelengths on link i.</td>
</tr>
</tbody>
</table>

(C_{i,λ}), whose values from 0 to 1 and values from 2 to 3, model the lightpath availability and unavailability respectively, on link i using wavelength λ, i.e., if a lightpath might be successfully established or not. Notice that predictive counters values are locally computed by the PPCE. By means of predictive counters, a PCE scheme computes the availability of a lightpath (l_{j,λ}) as shown in Equation (1). As a consequence, local NSI enables lightpath selection without any need of global NSI.

\[
l_{j,λ} = \left( \frac{\sum_{i \in j} (C_{i,λ})}{N_j} \right)
\]  

The procedure to update the predictive counters is the following. In case that a lightpath selected by the predictive PCE cannot be setup on link i using wavelength λ, the counter C_{i,λ} is increased (if and only if the counter value is lower than 3), otherwise it is decreased (if and only if the counter value is greater than 0). We assume that the OpenFlow controller informs the PCE about if the lightpath was successfully established or not. Moreover, we use two-bit counters because, as stated by [7], higher or lower counter values add a high or low hysteresis level respectively, both leading to the improper modeling of lightpaths availability. Unlike other contributions in this area using two-bit CGP counters (a counter per route/wavelength), PPCE implements two-bit FGP counters (a counter per link/wavelength), what undoubtedly suits requirements for highly dynamic scenarios, such those previously introduced in this paper.

The rationale driving the fine-granularity approach for predictive counters is illustrated in Fig. 1. The posed scenario analyzes the behavior of: i) the First-Fit (FF) wavelength assignment algorithm combined with Shortest-Path routing algorithm using global periodically updated NSI stored in the TED; and ii) two predictive algorithms solely using local NSI, which is stored in the so-called predictive TED, one of them using CGP counters, and the other one FGP counters. In this scenario, accurate NSI does mention to real (accurate) wavelengths availability on an optical link.

Let’s assume that a Connection Request (CR) arrives to the OpenFlow controller demanding a lightpath between WR nodes S_1 and D. For this purpose, the OpenFlow controller—acting as a Path Computation Client (PCC)—requests a lightpath computation to the available PCE, see step. 1 in Fig. 1a. The PCE scheme by means of FF (using the conventional TED), or Predictive path computation algorithms using either FGP or CGP counters (available in the predictive TED) will select lightpath S_1, 1, 2, D using wavelength λ_1 (step 2). Unfortunately, once the OpenFlow controller attempts to setup the selected lightpath it will be blocked, because the NSI available in the TED related to the available wavelengths (specifically for link 1 – 2) is outdated (inaccurate) (step 3).

For subsequent CRs (connections arriving before the next
updating time.) FF and predictive algorithms will work differently. FF will continue selecting lightpath $S_1$, 1, 2, $D$ using wavelength $\lambda_1$, as long as the NSI available in the TED is kept outdated, see Fig. 1b. Conversely, Predictive path computation algorithms can capture the unavailability of lightpath $S_1$, 1, 2, $D$ using wavelength $\lambda_1$, thus selecting wavelength $\lambda_2$ for the same path as shown in Fig. 1b.

It is also important to illustrate the different behavior shown by FGP and CGP counters. Fig. 1c shows how CGP counters cannot model the unavailability of wavelength $\lambda_1$ on the path $S_1$, 1, 2, since CGP counters can model the unavailability of a route, but cannot model the unavailability of the links within this route. This is a strong handicap for CGP counters that may be overcome by using FGP counters. Therefore, by means of FGP counters it is possible to select lightpath $S_1$, 1, 2 using wavelength $\lambda_3$, reducing in this way the amount of blocked connections.

B. The Predictive Path Computation Element scheme

The PPCE scheme computes lightpaths as follows. The first step (step. 1) consists in randomly selecting –assuming uniform probability distribution for all wavelengths– a wavelength for each path to a given destination $d$. We consider that under inaccurate NSI a random selection of wavelengths is more suitable in order to offset the impact of inaccurate NSI. Then, by means of FGP counters, PPCE computes the availability of the selected lightpath $l_{s,d}$ (cf. Equation 1). If the lightpath availability is less than 2 and the selected wavelength is available on the output link towards the destination, the lightpath is then selected. It is worth mentioning that in the case that all lightpaths are predicted to be unavailable, a PCE scheme selects a lightpath based solely on its output links availability (cf. step. 2 in PPCE). This policy is also useful to unblock the lightpaths that are in an unavailable state ($l_{s,d} \geq 1$). Finally, if there are no lightpaths available the incoming connection is blocked. The pseudo-code for the PPCE is shown in Fig. 2.

C. Scalability and Computational Complexity

The PPCE scheme works in two phases: an offline route generation phase, and an online lightpath selection phase. In the route generation phase, $|M_{s,d}|$ pre-computed (candidate) routes are (offline) generated for each source-destination pair by means of Dijkstra’s algorithm with a complexity of $O(|E| + |V| \times log(|V|))$, where $E$ and $V$ are the set of links and WRs in an optical network respectively. On the other hand, assuming the worst case scenario the complexity of the online phase is $O(2(|M_{s,d}| \times |W|))$.

We adopt a route pre-computation strategy in order to both minimize the complexity of the lightpath selection phase and ease the implementation in comparison with an adaptive routing strategy.

On the other hand, the proposed path computation algorithm is intended to be deployed as the lightpath computation mechanism of a PCE scheme. Moreover, to orchestrate control operation such as, the set-up and tear-down of lightpaths, PPCE relies on a centralized control architecture, i.e., SDN based on OpenFlow. Indeed, OpenFlow is used as Southbound interface for the PCE in the domain of an optical network.

From a deployment perspective, PPCE requires minor modifications in current control planes schemes since NSI dissemination is not required. This means that the TED is populated with local NSI (predictive counters) that is collected by the PCE scheme every time a lightpath computation is made. A local TED based on predictive counters avoids complex operation such as the synchronization required to build a global NSI.

IV. EVALUATION STRATEGY AND RESULTS

This section introduces an analysis of the experimental results obtained by the PPCE scheme, shown as a proof-of-concept for the proposed routing mechanism. Extensive simulation trials were conducted to obtain the blocking probability of distinct path computation algorithms. We consider two optical topology configurations: 1) the well-known NSFNET topology (see Fig. 3.a) similar to authors in [11], and; 2) a fragment of a full-mesh Clos fabric network topology often used within DCNs (see Fig. 3.b), hereinafter referred to as DCN-like topology [12].

The simulation results presented in this section were obtained using the well-known network simulation framework Omnetpp [13], and all the plotted values have a 95% confidence interval not larger than 0.5% of the plotted values.
Figure 3. Evaluated network topologies: a) NSFNET topology (14 nodes, 21 links); b) DCN-like topology.

A. Network Model

In the following we describe the network model used throughout this paper where the following settings apply:

- CRs arrive at a node according to a Poisson process with an inter-arrival mean time of 10 time units.
- The connection holding time is exponentially distributed with a mean of 50 time units.
- WRs without wavelength conversion capabilities. Each connection request requires a full wavelength on each link.
- NSI is periodically disseminated according to parameter updating time, and available at all WRs right after its dissemination.
- Blocked connection requests are not reattempted. The rationale behind this assumption is to avoid long setup times. Despite that connection reattempts are in the order of hundreds of milliseconds when control plane technologies such as SDN are used, the impact of connection reattempts cannot be neglected when the propagation delay is high or in highly dynamic scenarios where it is expected to provision lightpaths on a short-term basis.
- Once a lightpath is provisioned, it cannot be reconfigured. This assumption is made in order to avoid services disruption.
- 80 wavelengths per fiber. This assumption is based on the channel spacing standards defined by the International Telecommunication Union (ITU).

Figure 4 depicts the blocking probability versus a wide spectrum of updating time values for the two considered topologies for: 1) FF; 2) a random fit wavelength assignment strategy combined with Shortest-Path routing (RF); 3) a Predictive path computation algorithm using CGP counters hereinafter referred to as FRA [7], and; 4) the proposed PCE scheme. As shown in Fig. 4, both FF and RF algorithms yield the lowest blocking probability only when updating time is up to 15 time units for the NSFNET network and when updating time is up to 20 time units for the DCN-like network. However, when the updating time increases to higher (realistic values), the blocking probability of both FF and RR significantly increases and tends to settle when the holding time approximate to the updating time.

In order to put into context what we consider as realistic updating time values, consider (commonly assumed in the literature) that in dynamic optical scenarios the holding time of a connection is the time required for its data transmission. Therefore, assuming that an optical channel supports 100 Gbps bit rates (which is common in WDM networks), and the data to be transmitted is on average of 10 Gbits, having an update time of 15 time units leads to 3 update messages every 100 ms, or 30 update messages every second. Based on this reasoning, it is intuitive that relying on periodic NSI dissemination to build the TED has a significant impact on the scalability in dynamic scenarios.

Contrary to FF and RF, Predictive path computation algorithms such as FRA, and PPCE are not affected by the updating time. We note that Predictive path computation algorithms do not require periodical dissemination of NSI. Moreover, as it can be observed in Fig. 4, the PPCE outperforms the FRA in both virtual topology designs.

Based on the obtained results, it can be stated that 1) the performance of conventional path computation algorithms such as FF and RF, is strongly affected by the inaccuracy of NSI; 2) the performance of Predictive path computation algorithms such as PPCE and FRA schemes does not depend...
of Network State Information (NSI) dissemination strongly results, we conclude the following: i) the frequency Engineering Database (TED). Based on the obtained simu-
the blocking probability without any need for a global Traffic Element (PPCE) scheme with the aim of reducing under the presence of inaccuracy.

routing outperforms conventional path computation algorithms network topologies evaluated. Based on the results shown times.
conventional path computation algorithms with low updating times. Therefore, a tradeoff of predictive path computation algorithms is that they substantially reduce the signaling overhead, but they may slightly increase the blocking probability compared with conventional path computation algorithms with low updating times.

Figure 5 presents simulation results related to the blocking probability versus a wide spectrum inter-arrival values both network topologies evaluated. Based on the results shown in Fig. 5, we immediately conclude that indeed, predictive routing outperforms conventional path computation algorithms under the presence of inaccuracy.

V. CONCLUSIONS

In this paper, we present a so-called Predictive Path Computation Element (PPCE) scheme with the aim of reducing the blocking probability without any need for a global Traffic Engineering Database (TED). Based on the obtained simulation results, we conclude the following: i) the frequency of Network State Information (NSI) dissemination strongly impacts on the blocking probability of conventional PCE, and; ii) Predictive PCE schemes can outperform PCE schemes relying on a global TED. As a future line of work, we propose to consider other routing constraints such as flexible-spectrum grid and physical impairments.

ACKNOWLEDGMENTS

Authors acknowledge the support received by the Spanish Ministry of Economy under contract TEC2012-34682 (project partially funded by FEDER), and the Catalan Research Council under contract 2014 SGR 371.

REFERENCES