Extending path computation element for lightpath restoration in wavelength-switched optical networks

Lianxing Hou (侯连兴)*, Yueming Lu (陆月明), Junhua Wang (王俊华), Yuefeng Ji (纪越峰), and Yiqiang Hua (华一强)

Key Laboratory of Information Photonics and Optical Communications, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China

**E-mail: lancyhou@gmail.com*

Received April 24, 2009

Current generalized multi-protocol label switching (GMPLS) standards do not include adequate models for wavelength-switched optical networks (WSON) in recovery mechanisms. In this letter, GMPLS/path computation element (PCE) extensions are applied for the restoration of the lightpaths disrupted by collision or optical impairment. A reserved deflection routing scheme is proposed to achieve fast restoration. It uses the expanded PCE component to compute and assign the backup paths for lightpath recovery. Numerical results demonstrate that this scheme is effective and low cost.

OCIS codes: 060.4261, 060.4510.

doi: 10.3788/COL20100802.0142.

In transparent wavelength-switched optical networks (WSON)[1], dynamic lightpath provisioning requires effective routing and wavelength assignment (RWA) schemes, taking into account of both bandwidth availability and wavelength continuity constraints. However, lightpath restoration requires a faster restoration speed and a simpler deployment system. Toward these goals, the centralized path computation element (PCE) has been developed[2]. It is able to potentially simplify the implementation of network nodes that may avoid complex routing modules and to provide effective network resource assignment. Thus, the path computation clients (PCCs) can request the computation of an explicitly routed path given a set of constraints, which are motivated by mesh restoration algorithms or heuristics. Some researchers are opposed to use PCE for restoration, since it takes time. In our experimental setup, star topology is applied to PCE and PCCs, and the recovery time could be reduced to less than 50 ms.

However, the major challenge in the distributed all-optical networks is that the current global information on resource availability cannot be guaranteed at any particular place and time, which can lead to contentions. We may soon have to support an increasing number of burst traffic loads in the future such as in wavelength-routed optical burst switched (WROBS) networks or all-optical packed-switched (OPS) networks[3−11]. It is expected that the connection requests will arrive at a very high speed, while the average duration of each connection is only several dozens or hundreds of milliseconds. One possible routing mechanism that can be used to reduce loss or congestion due to sub-optimal path selection is deflection routing[8]. In this mechanism, each node maintains several paths to a destination, with one path designated as the primary (default). When the primary path of an incoming connection is not available, the node deflects the connection to any of the secondary paths. Although deflection routing is inexpensive and simple with the capability of high resource utilization, it may result in optical packets looping in a multi-hop network for a long time. Moreover, if a deflected service takes a longer path to reach its destination, it would lead to overtime and cause network congestion, especially at high traffic loads[9]. To solve this problem, a deflection routing protocol for optical burst switched (OBS) networks has been proposed in Ref. [10], and the optical bursts might have likely arrived out of sequence at the destination in Ref. [11]. In addition, deflection routing is, by nature, suboptimal because it only considers the congestion of the current switch, not the state of the links further along the path, and may cause undesirable vibration effects. To solve these problems, Teng *et al.* presented an approximate integer linear optimization scheme to path selection with the objective of balancing the traffic across the network links to reduce congestion and improve overall performance[12]. However, on account of the complex computation and huge backup route data for each node, this scheme has become more inefficient, resulting in an inability to handle larger non-Poisson traffic such as link failure. While the integer linear programming (ILP) formulation can avoid the defects mentioned above in WSON, and achieve high computation efficiency and high capacity efficiency. The PCE charges the centralized network resource assignment, while the pre-configured cycle fast restoration scheme provides self-adapting protection not only for on-cycle links but also for straddling links[13].

We propose the enhancement of the PCE/generalized multi-protocol label switching (GMPLS) network architecture for the reserved deflection routing scheme to support the fast restoration of lightpaths. Unless otherwise specified, we focus on the GMPLS-based multi-fiber networks that use the resource reservation protocol-traffic engineering (RSVP-TE) signaling protocol. Wavelength reservation is bidirectional. To ease our analysis, the constraint called as "wavelength continuity constraint" is considered; whereas the limited range wavelength conversion or sparse converters in WSONs is left for further evaluation in the future research.

1671-7694/2010/020142-04

The PCE is a multi-threaded, asynchronous process (Fig. 1[14]). A dedicated thread in PCE is responsible for updating the traffic engineering database and policies through the traffic engineering database (TED) from a control plane. Meanwhile, another thread is responsible for the actual path computation through a connection manager. The open shortest path first-traffic engineering (OSPF-TE) protocol is used for the synchronization mechanism to update the TED. In addition, the TE Link sub-TLVs (type/length/value) contained within the OSPF-TE link-state advertisement (LSA) is defined as the carrier of the TED information. The TED constructs double databases: the working and backup resource databases. The PCE has a local policy that affects path computation and selection in response to a path computation request. Such a policy may act on the information provided by the requesting PCC. The result of applying the policy includes, for example, rejection of the path computation request or provision of a path that does not meet all of the requested constraints. Furthermore, the policy may support administratively configured paths or a selection of transit providers. The inclusion of the policy within the PCE may simplify the applications of the policy within the path computation/selection process. However, multiple paths need to be computed to support a single service (e.g., for protection or load sharing). A PCC that requires more than one path to be computed may send a series of individual requests to the PCE. In this case of non-synchronized path computation requests, the input/output (I/O) system may assign multiple individual path computations to generate the paths.

According to different requirements (lightpath establishment or restoration), different algorithms are implemented in dynamic shared libraries, following the application programming interface (API) algorithm that abstracts the underlying optical network resources as a directed graph. If the request to PCE is a new lightpath request, the path computation engine will apply the specified RWA algorithm within the path computation and selection process, and the ILP algorithm will assume responsibility for the backup path requests from the lightpath restoration.

Three types of protocols need to be deployed for the PCE/WSON requirements. As recommended by the internet engineering task force (IETF), common control and measurement plane (CCAMP), and PCE workgroups, the OSPF-TE works as the interior gatway protocol (IGP) TE routing protocol; the RSVP-TE does the signaling; and the PCE communication protocol (PCEP) deals with the path computation[1]. The PCEP is used to support PCC-PCE and PCE-PCE communication. The OSPF-TE is expanded to advertise the TED information (e.g., the wavelength available status on a per link basis including working and backup link) by carrying newly defined sub-TLVs. Afterwards, the RSVP message adopts a flag to mark the path identification (ID) of deflection routing while the lightpath is restored in the backup path. The requests and responses for deflection routing interact in both PCE and PCCs through the extended PCEP.

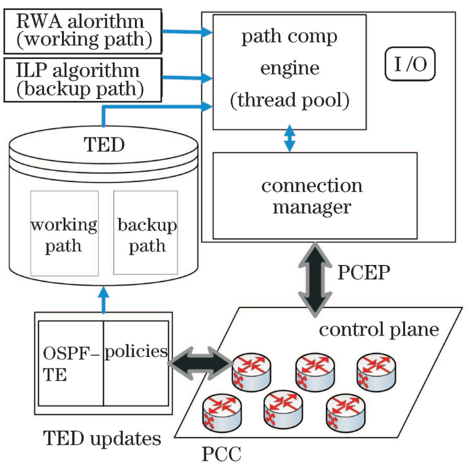

Fig. 1. PCE extension.

To solve the congestion problems, the reserved deflection routing (Rsv-DR) is proposed. In this scheme, the backup routing for the lightpath in the working path is not random. This is because it has already been generated in the PCE based on the ILP. The ILP algorithm is used for the pre-configured cycle (p-cycle) design[13,15], which has a higher capacity efficiency because it can provide protection not only for on-cycle links but also for straddling links.

We model the wavelength division multiplexing (WDM) optical network as an undirected graph $G=(V, E)$, where each node in V represents an optical switch, and each edge in E represents a network link. The conventional reserved resource capacity efficiency (RE) can be defined as[16]

$$\mathrm{RE}(p)=\frac{\sum\limits_{\forall e\in E}X_{p,e}\delta_p(e)}{\sum\limits_{\forall e\in E|X_{p,e}=1}c_e},$$
$$X_{\mathrm{p,e}}=\begin{cases}2 & e \text{ is straddling link in p-cycle } p\\ 1 & e \text{ is on-cycle link in p-cycle } p\\ 0 & \text{otherwise}\end{cases}, \quad (1)$$

where p is a variable defined as a p-cycle; c_e is the weight of link e, which denotes the link capacity compared with the standard band width with a default value of 1, and $\delta_p(e)$ is the binary variable, which is 1 if the p-cycle p occupies the wavelengths in link e, and 0 otherwise.

The restoration probability (RP) can be defined as

$$\mathrm{RP}=\frac{\sum\limits_{\forall p\in P}\sum\limits_{\forall i\in p}\Delta s_e}{\sum\limits_{l\in L}c_l\delta_l(x)}, \quad (2)$$

where Δs_e is the working capacity increment in link e; i is the link on cycle: L is the set of label switched paths (LSPs); p is the set of p-cycles; x is the broken link; l is an LSP; and $\delta_l(x)$ is the binary variable, which is 1 if the LSP l occupies the wavelengths in link x, and 0 otherwise.

In this letter, p-cycle generates the efficient set of cycles via the PCE based only on the topology of network and static spare capacity, while the local node selects the Rsv-DR in the p-cycle.

The objective of the deployed ILP algorithm is

$$\mathrm{Min}(\sum_{\forall e\in E}c_e s_e). \quad (3)$$

The constraints are

$$s_e = \sum_{p \in P} \delta_p(e), \quad (4)$$

$$\sum PL_p = \sum S_l, \quad (5)$$

$$s_e \leq w_{\max}. \quad (6)$$

The objective of Eq. (3) is to minimize the total capacity for restoration. In this scheme, c_e is the weight of link e, and s_e is the reserved wavelength in link e. When c_e is the real distance between nodes, the objective is to minimize the total length. When e is 1, the objective is to minimize the hop-count. Equation 4 specifies that the reserved wavelengths s_e in link e amount to the sum of the protected wavelengths of each p-cycle in link e. Equation (5) ensures 100% protection against a single link failure. Here, PL_p refers to the total wavelengths that the p-cycle p can protect, and S_l is the weight of the LSP l that specifies the hop-count of this service. Equation (6) confines the reserved wavelengths s_e in link e less than the maximum available wavelengths $w_{\max}$. According to the wavelength continuity constraint, one wavelength can be used in a single fiber once. Thus, if $w_{\max} = 1$, each link only has one fiber of spare capacity for the p-cycle.

The algorithm is as follows: firstly, select all the simple cycles from the specified topology using the depth first search (DFS) algorithm; secondly, form the p-cycles set and select M cycles that have the maximum reserved resource capacity efficiency with 100% protection; finally, complete the switch of the cross connection matrix for the on-cycle link (backup link).

After the lightpath is disrupted, the responsible node starts to compute the rerouting algorithm for the optimal Rsv-DR in the selected p-cycle which matches the rest path the most. During the Rsv-DR trip, the intermediate nodes catch the service from the backup link and forward it as a special route. The destination node then switches the Rsv-DR to a normal working link.

The example shown in Fig. 2 illustrates the restoration process that occurs in a failure. As it can be seen that (ABCDE) is one p-cycle for the reserved deflection routing. When the primary path (EC) is not available, the node E requests the PCE to compute the path and select the preferred p-cycle as the deflection routing path for restoration. It then deflects the connection to the other secondary backup path. If one path is not enough for burst data, two routes are enabled. As all the restored connections transfer through the backup path, the intermediate nodes such as nodes A, B, and D process these lightpaths according to the deflection routing mark and then switch them to the next related node. In the deflection routing's destination, node C reads the original path message saved in the path state block/reservation state block (PSB/RSB) database[17] and recovers the original path.

In this section, we evaluate this scheme using the COST239 topology (11 nodes). In order to achieve a realistic network load, lightpath provisioning requests are dynamically generated according to a Poisson process uniformly distributed among the source-destination pairs. The network is provisioned with lightpaths using the RSVP-TE. The results are obtained with a network testbed (Fig. 3) and collected after single link failure under various network loads. Nodes are then connected by 16 fibers in one domain with single PCE server, with each fiber containing 20 wavelengths.

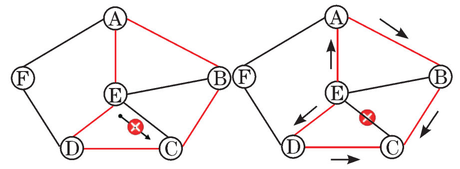

Fig. 2. An example of the restoration process.

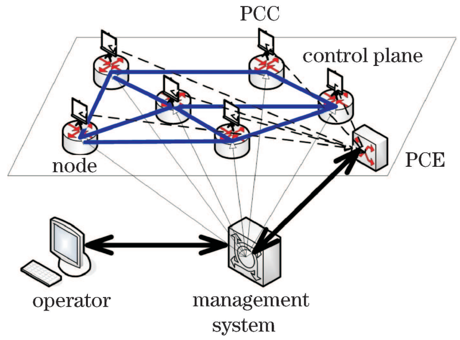

Fig. 3. Schematic of experimental setup.

In Fig. 4, we compare the performances of 1:1 protection, source node rerouting, local node original deflection routing (Lcl O-DR), and Rsv-DR. The figure shows that 1:1 protection performs best under low traffic because it has the fastest and the most effective restoration capability for its specified protection path. However, the capacity efficiency is so low that the performance deteriorates sharply under higher load. When the load is more than 1,000 Erlang, the available resource is not enough for all the new restoration requests; thus, the restoration probability of the source node rerouting performs less than the other method under high load. In contrast, the source node rerouting performs the worst under a lower load because it has to return the source node error message and reroute, thereby increasing the probability of failure for the lightpath. When the load is 1,200 Erlang, its restoration probability slips to almost 70%. The local reserved deflection routing performs better than the original deflection routing. In the original deflection routing, each node maintains several paths to a destination that computed itself. Moreover, the paths designated as slaves are naturally suboptimal, and the congestion probability of the current switch brought about by outdated

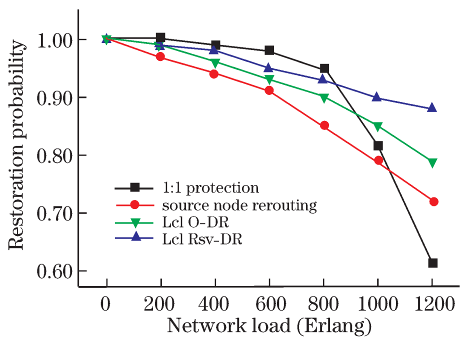

Fig. 4. Restoration probability versus load.

global information rises with the increased traffic load. In terms of the local Rsv-DR, the responsible node is the local node, and the protection resource is already reserved for each working path in the PCE, which has real-time TED. Thus, under a higher load, the local Rsv-DR can achieve better performance than the local O-DR.

In conclusion, we have shown that the PCE of GMPLS-based WSONs could be extended for lightpath restoration. The local Rsv-DR scheme uses the PCE to assign backup links for each working link. It has ring-like speed with mesh-like capacity for lightpath restoration. Simulation shows that it achieves better performance under various network loads. Therefore, the local node Rsv-DR scheme is a suitable lightpath restoration scheme in WSONs.

This work was supported in part by the National "863" Program (No. 2007AA01Z252), the National Natural Science Foundation of China (Nos. 60772024 and 60711140087), the National "973" Program of China (No. 2007CB310705), the Specialized Research Fund for the Doctoral Program of Higher Education of MOE (No. 200800130001), the Program for Changjiang Scholars and Innovative Research Team (No. IRT0609), the International S&T Cooperation Program of China (No. 2006DFA11040), and the New Century Excellent Talents in University (No. 06-0090).

References

1. Y. Lee and G. Bernstein, Internet Draft, Feb. 2008.
2. A. Farrel, J.-P. Vasseur, and J. Ash, RFC4655, IETF, Aug. 2006.
3. M. Düer and P. Bayvel, J. Lightwave Technol. **20,** 574 (2002).
4. G. Wang, H. Wang, and Y. Ji, Chin. Opt. Lett. **6,** 628 (2008).
5. A. K. Garg and R. S. Kaler, Chin. Opt. Lett. **6,** 244 (2008).
6. S. Yao, B. Mukherjee, and S. Dixit, IEEE Commun. Mag. **38,** 84 (2000).
7. C. Wu and S. Xiao, Chin. Opt. Lett. **7,** 384 (2009).
8. J. Bannister, F. Borgonovo, L. Fratta, and M. Gerla, IEEE/ACM Trans. Net. **3,** 509 (1995).
9. T. S. El-Bawab and J.-D. Shin, IEEE Commun. Mag. **40,** 60 (2002).
10. X. Wang, H. Morikawa, and T. Aoyama, Proc. SPIE **4213,** 242 (2000).
11. L. Xu, H. G. Perros, and G. Rouskas, IEEE Commun. Mag. **39,** 136 (2001).
12. J. Teng and G. N. Rouskas, in *Proceedings of ONDM 2005* 1 (2005).
13. W. D. Grover and D. Stamatelakis, in *Proceedings of IEEE ICC* **1,** 537 (1998).
14. R. Casellas, R. Muñoz, and R. Martínez, in *Proceedings of ECOC 2008*, 1 (2008).
15. D. A. Schupke, C. G. Gruber, and A. Autenrieth, in *Proceedings of IEEE GLOBECOM* **3,** 2761 (2004).
16. J. Doucette, D. He, W. D. Grover, and O. Yang, in *Proceedings of DRCN 2003* 212 (2003).
17. R. Braden and L. Zhang, RFC2209, IETF, Sept. 1997.