Design of contention resolution over backup path in path computation element based wavelength-switched optical networks

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

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 August 10, 2009

A contention resolution scheme, P-cycle auto-switching (PCAS), is proposed in wavelength-switched optical networks (WSONs) based on generalized multi-protocol label switching (GMPLS) to solve backward blocking. This system permits the preferable request and switches the other to a backup path generated by the P-cycle. Simulation results show that the PCAS scheme is effective in mitigating backward blocking, especially under concentrated load.

OCIS codes: 060.2330, 060.4251.
doi: 10.3788/COL20100804.0372.

In transparent wavelength-switched optical networks (WSONs)[1], dynamic light path provisioning requires effective routing and wavelength assignment (RWA) schemes that consider both bandwidth availability and wavelength continuity constraints. To adapt to the growth of optical networks, distributed schemes have been proposed and are being standardized in the framework of generalized multi-protocol label switching (GMPLS)[2]. However, the major challenge is that the current global information on wavelength availability cannot be guaranteed at any particular place and time in the distributed system, which undoubtedly leads to contention.

The simplest and most basic wavelength assignment schemes are the source-initiated reservation (SIR) method and the destination-initiated reservation (DIR) method[3,4]. In the SIR method, a reservation request control message is sent from the source to the destination, reserving one or more wavelengths along the way as it proceeds towards the destination. The destination node will select one of the reserved wavelength channels (if available) and send a confirmation request back to the source, informing it of the selected wavelength and releasing the other reserved wavelengths. In the DIR method, a control message is forwarded from the source to the destination, on the way collecting the wavelength availability information along the path. Based on this information, the destination node will select an available wavelength (if one is available along the path) and send a reservation request back to the source node to reserve the selected wavelength. However, several studies have shown that neither method efficiently utilizes network resources. Therefore, the more centralized path computation element (PCE) method, which can provide a strict route, presents an appealing solution[5]. The PCE method is able to simplify the implementation of network nodes potentially, thereby avoiding complex routing modules and providing effective network resource assignment.

In the future, we may have to deal with the contention resolution and avoidance problems in order to support an increasing number of burst traffic loads[6−8]. This study focuses on the utilization of backup paths in WSONs for contention resolution. A contention resolution method called P-cycle auto-switching (PCAS) is proposed in a signaling-based approach with DIR method. To simplify our analysis, we make the following assumptions. Unless otherwise specified, the multi-fiber GMPLS-enabled WSONs use resource reservation protocol-traffic engineering (RSVP-TE) signaling protocol. Wavelength reservation is bidirectional, and there is no wavelength converter or opto-electronic conversion.

The basic assumption is that new arrival requests for the network and blocking events are mutually independent. The number of fibers between nodes is defined by F, the wavelength in the link l is w, and the newly launched requests dynamically arriving according to a Poisson process are distributed with an average of $\lambda_{w,l}$. The holding time of wavelength resource for requests compliance has a general distribution average of $1/\mu_{w,l}$. The route of path message requests is K, and K_i refers to a link to a number of i hops. The model can therefore be considered an M/G/F/F[9] queuing system.

The key performance metric in the dynamic lightpath establishment schemes is the connection blocking probability. A light path connection request will be blocked when a route with sufficient free capacity cannot be found on the path from the source to the destination. In the case of wavelength-continuous light paths, if a wavelength cannot be found between the source and the destination, the connection request will be blocked even if there is free capacity on each other's hop of the path.

There are different definitions of the concepts "forward blocking" and "backward blocking". As signaling is based on the DIR method, we refer to these two types of blocking as blocking due to insufficient network capacity (called forward blocking) and blocking due to having outdated global information (called backward blocking)[4]. As explained earlier, delays are caused by the need to collect and transmit the link state information or propagation. When a control message reaches a link to reserve wavelength channels, it is possible that the available capacity when the state information of the link is collected has, in the meantime, been reserved by another connection request. There are three types of blocking in the above model.

Type I blocking: the Resv message arrives at one node and finds that the wavelength on the required port has been reserved by another connection. As shown in Fig. 1, when the Resv message of request R0 reaches node F, the wavelength w that is available when the state information of the link is collected has, in the meantime, been reserved by request R1. Thus, collision occurs.

The probability (occupied m channels) can be calculated by the Erlang B formula when wavelength w in link l only has m channels to be occupied[9]:

$$P_{w,l}(m) = \frac{\left(\lambda_{w,l}/\mu_{w,l}\right)^m/m!}{\sum\limits_{i=0}^{F}\left[\left(\lambda_{w,l}/\mu_{w,l}\right)^i/i!\right]}, m = 0, \cdots, F. \qquad (1)$$

The transition matrix for M/G/F/F model is $Q_w^{K_i}$. The average of time interval for Resv message on collision, Tx_{K_i}, can be approximated into the maximum time for light path establishment. During this period, the transition matrix for w is

$$Hx_{w,K_i}^K = e^{Q_w^{K_i} \times Tx_{K_i}}, \qquad (2)$$

and the blocking probability is

$$X_{w,K_i}^K = 1 - \sum_{i=0}^{F-1}\left[P_{w,K_i}(i) \times Hx_{w,K_i}^K(i+1, F+1)\right]. \qquad (3)$$

Type II blocking: if pair nodes of the link reserve the same wavelength at the same time, each Resv message will find that the assigned wavelength has been reserved after arriving at the opposite node of the link. As shown in Fig. 2, there are two Resv messages from the connection requests R0 and R1. Collision occurs in the link between E and F when the pair nodes reserve the same wavelength and send the Resv message at the same time. The contention should then be resolved by other means[2]. This kind of contention resolution is used to assign the wavelength to one Resv message so the other will be dropped. (If there is no alternative mechanism, both will be dropped.) Node E sends a ResvErr message to node F indicating such an error.

The average of time interval for Resv message on collision is Ty_{K_i}, and it can be approximated into a transmission delay in each link for the RSVP package.

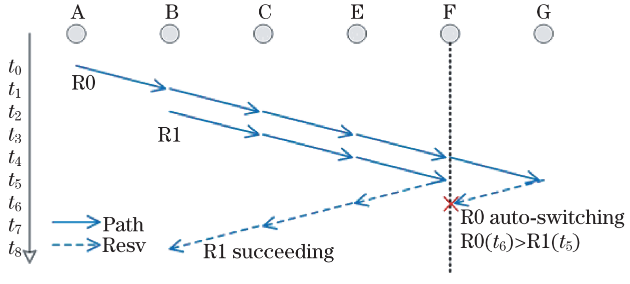

Fig. 1. Type I blocking.

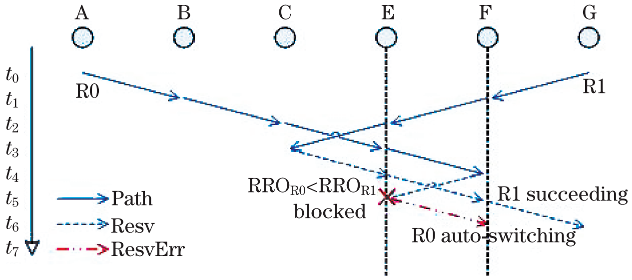

Fig. 2. Type II blocking.

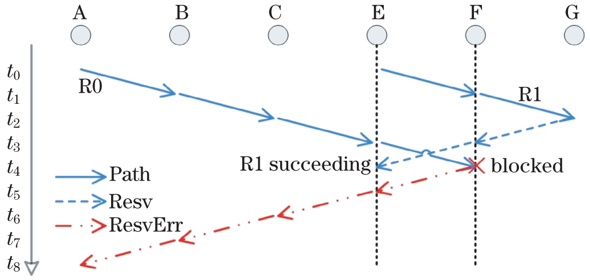

Fig. 3. Type III blocking.

The blocking probability is

$$Y_{w,K_i}^K = \frac{1}{2}\left\{1 - \sum_{i=0}^{F-1}\left[P_{w,K_i}(i) \times Hy_{w,K_i}^K(i+1, F+1)\right]\right\}. \qquad (4)$$

Type III blocking is similar to Type II. There are also two collision requests, R0 and R1, on both ends of the link (Fig. 3). However, one of them is a path message used to detect capacity for request. If all available wavelengths have been already reserved, the path message will be blocked. As in the DIR method, this is a forward blocking (blocking due to insufficient network capacity) type, and this request has not actually occupied any resource (the path message carries the labels set, which identifies a set of available wavelengths), so it is dropped without handling. This type of blocking will be ignored in this contention resolution scheme.

In the normal method, abandoning the loser in the contention will lead to wasted resources as the loser may have already reserved several links before the collision occurs. To make matters worse, the request will be repeated. The repeated requests will greatly burden the network load, especially in high load situations. Now we introduce a new contention resolution system based on protection switching: PCAS.

The P-cycle or pre-configured restoration cycle is one of the researching protection methods for link failure[10]. It offers advantages of both ring and mesh restoration schemes: fast restoration time as in ring protection and high capacity efficiency as in mesh restoration. Moreover, the capacity efficiency is higher because the P-cycle can provide protection not only for on-cycle links but also for straddling links. In this study, the P-cycle method is improved for the application of contention resolution. To provide a backup path for a work link (fiber), the reserved resources are used for protection switching in the PCAS configuration. The generating algorithm for the P-cycle is in charge of centralization. This can be simplified as the traffic of protection switching for collision is much less than the traffic for link failure in a steady-state network. At the same time, it facilitates dealing with protection switching as the switch of the cross connection matrix has been already done before collision occurs, guaranteeing to avoid request overtime.

Before collision occurs, the P-cycle is designed from reserved fibers to carry out the protection of work fibers because the reserved path made by the P-cycles not only protects the wavelength in the other on-cycle work links but also those in the straddling links. When collision occurs, the contention algorithm[2] is improved by the record route object (RRO) sum contention priority. The RRO in both collision Resv messages are compared. The request with the bigger sum of reserved links wins the contention, whereas the lesser one is rejected. If the requests are equal, the node with the higher node ID wins. As for the failed connection request, the responsible node checks the "RRO" used to record the route in the Path message phase and is saved in the path/reservation state block (PSB/RSB) database, and then it selects the best P-cycle by matching the rest path of the Resv message. It then completes the reserved path by switching the connection of both end nodes, and when the Resv messages reach the reserved path end node, it continues its original path.

Three types of protocols need to be deployed for the PCE/WSON requirements. As recommended by the Internet Engineering Task Force (IETF), the open shortest path first-traffic engineering (OSPF-TE) works as the interior gateway protocol-traffic engineering (IGP-TE) routing protocol, the resource reservation protocol-traffic engineering (RSVP-TE) does the signaling, while the PCE communication protocol (PCEP) deals with the path computation[1]. The PCEP is used to support path computation client (PCC)-PCE and PCE-PCE communication. A dedicated thread in PCE is responsible for updating the traffic engineering database (TED) from a control plane. TED is extended to construct double databases: the working and backup resource databases. Each node has a database that includes all the available P-cycle information related to it, and PCC is used for the synchronization mechanism to update TED. Sender template object extends the function of Resv message. When collision occurs, the Resv message adopts a flag to mark the path ID of the P-cycle, which is selected from a local node database, and it then notifies PCE through PCC to update TED. When the downstream node receives the Resv message in the backup link, it deals with the request according to the path ID of the P-cycle until it reaches the destination node in the reserved path.

After the contention occurs, the responsible node (node F) begins to compute the rerouting algorithm for the optimal reserved path in the selected P-cycle, as shown in Fig. 4. Two requests, route DCF and route ZACFG, are shown in Fig. 4. route DCF is completed, whereas the Resv message of route ZACFG is blocked in node F because the wavelength has been reserved by route DCF. The responsible node thus switches route ZACFG to route ZABFG, and it then notifies PCE server to update TED.

The algorithmic details are presented below.

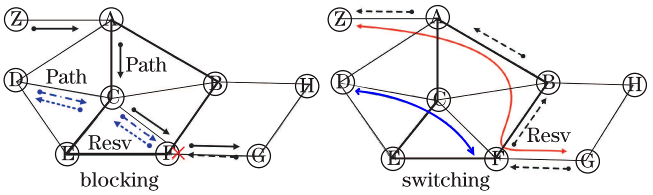

Fig. 4. Blocking and switching using P-cycle.

1) Checking the RRO from the Path message saved in the request phase and obtaining the subsequent route (FCAZ) of the Resv message.

2) Confirming a P-cycle (ABFEC) that matches the rest path of the Resv message the most and calculating the reserved path (FBA).

3) Pushing the sender template object in the Resv message, including "the source node (Z) ID" (the destination node (Z) of the Resv message) and "layered service provider (LSP) ID", and updating for the destination node (A) ID and the P-cycle ID of the backup reserved path.

4) Pushing the RSVP hop object in the Resv message, and updating for the local node (F) IP.

5) Switching and completing the reserved path as a new Resv message (from F to A).

6) Modifying the RSVP hop in order in each node in the reserved path and recording the RRO information (GFBAZ).

In the reserved path's destination (A), the sender template and RSVP hop are popped, the original Path message saved in the PSB/RSB database is read, and the original path (from A to Z) is recovered.

The P-cycle generation algorithm is based on the integer linear programming (ILP)[11]. It generates an efficient set of cycles based only on the topology of the network and the static spare capacity to reduce its complexity dramatically.

The objective of the deployed ILP algorithm is

$$\mathrm{Min}\Big(\sum_{\forall e\in E} c_e s_e\Big). \tag{5}$$

The objective of expression (5) 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.

Referring to Ref. [12], the conventional reserved resource capacity efficiency (R_E) can be defined as

$$R_E(p) = \sum_{\forall e\in E} X_{p,e}\delta_e \Bigg/ \sum_{\forall e\in E|X_{p,e}=1} c_e, \tag{6}$$

where δ_e is a binary variable; it is 1 if cycle p occupies the wavelengths in link e, and 0 otherwise. And

$$X_{p,e} = \begin{cases} 2 & e \text{ is straddling link in } p \\ 2 & e \text{ is on-cycle link in } p \\ 0 & \text{otherwise} \end{cases}$$

The algorithm is as follows.

1) Searching all the simple cycles from the specified topology using the depth first search (DFS) algorithm.

2) Ensuring that all the on-cycle links come from the

backup link and the straddling links from the work link.

3) Forming the P-cycles set, and selecting several cycles with the maximum reserved resource capacity efficiency to protect more work links.

4) Completing the switch of the cross connection matrix for the on-cycle link (backup link).

With the PCAS, the blocking probability is modified as

$$X'^{K}_{w,K_i} = 1 - \sum_{i=0}^{F-1-f+R_E f} \left[P_{w,K_i}(i) \times Hx^{K}_{w,K_i}(i+1, F+1-f)\right], \quad (7)$$

$$Y'^{K}_{w,K_i} = \frac{1}{2}\left\{1 - \sum_{i=0}^{F-1-f+R_E f} \left[P_{w,K_i}(i) \times Hy^{K}_{w,K_i}(i+1, F+1-f)\right]\right\}, \quad (8)$$

where f is the serial number of the backup fiber for protection switching. If only the F fibers are entirely occupied, the collision will occur. The total probability is

$$Z'^{K}_{w,K_i} = \sum_{n=0}^{F} X'(n) Y'(F-n) . \quad (9)$$

We evaluate the PCAS in terms of the aforementioned performance parameters. Simulation experiments are carried out using a widely used network topology: the NSFNET topology. Parameters of the network topology is given in Table 1. To achieve a realistic network load, the light path provisioning requests are dynamically generated according to a Poisson process. Both inter-arrival time of the requests and the light path holding time are exponentially distributed with an average of $1/\lambda = 1000$ s and $\frac{1}{\mu}$ s, respectively. For the given load (i.e., λ/μ = Erlangs), the network is provisioned with light paths using RSVP-TE. The simulation is triggered through a custom-built JAVA event-driven simulator (GLASS), and the results are collected. We use GLPK-4.31 to implement the ILP formulation[13].

To simplify our calculation, we place 16 fibers (links) between each pair of nodes, and 4 (center) or 6 (edge) of them are reserved for protection. The P-cycles consist of 6 nodes on average. The transmission time in each link for the RSVP message is $T_{\text{delay}} = 5$ ms, and the maximum time of request is $T_{\max} = 100$ ms.

Figure 5 shows the effect of the PCAS on the average setup delay. The average time with PCAS is smoother than that without PCAS. Under a lower load, the average time without PCAS is shorter because it does not need extra computation, and it has more resources for work. However, under a higher load, the average time increases because too many fail requests need to be resumed. Actually, all schemes show good performance (near 50 ms) mostly because the limitation of the maximum time of request and all the fail requests are unrecorded.

Table 1. Network Topology Parameters

Topology	Number of Nodes	Number of Links	P-Cycles Number	Protection Ratio for Work Links
NSFNET	14	21	18	0.947

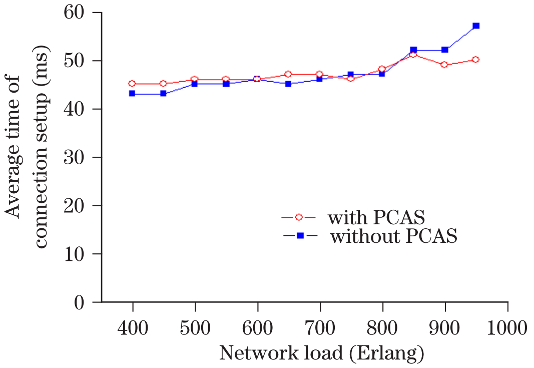

Fig. 5. Average setup delay.

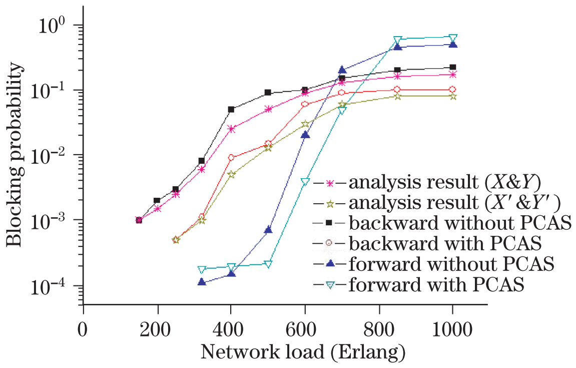

Fig. 6. Blocking probability under uniform load.

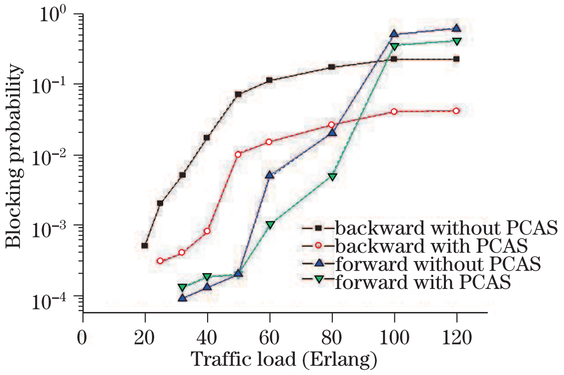

Fig. 7. Blocking probability under concentrated load/single link traffic.

The blocking probabilities with and without PCAS are compared. The light path provisioning requests are distributed both uniformly and concentrated among source-destination pairs. Figure 6 depicts the PCAS performance in the simulation and in the analytical model of uniformly distributed light path requests. As shown in the analysis results, both backward probabilities have similar treads. However, under the higher load, the simulation results are much greater than the analysis results. The repeated requests that fail to connect by contention will burden the network load significantly. The backward probability with PCAS is lower than that without PCAS. The forward probability with PCAS is lower than that without PCAS except at the highest and the lowest loads because the extra calculation increases the blocking probability in these situations. In a large network load, the forward probability becomes larger than the backward probability, and the four probability performances all level off in high load. This is because forward blocking (due to insufficient network capacity) becomes the main problem, and the forward probability with PCAS is actually slightly larger than that without it. This proves that the PCAS method partly sacrifices capacity for utilization rate. The final statistic figure is heavily dependent on the processing capacity of the simulation software.

Figure 7 depicts the PCAS performance under a concentrated load. To simplify our process for a typical environment, we select one single link to burst traffic, and the load of other links is fixed at 600 Erlangs. This model describes abrupt data within a small area. The simulation result shows that the backward probability with PCAS is much lower than that without it, and it levels off to a very low probability. This proves that backward probability can be restrained effectively. The forward probability with PCAS performs much better than that without PCAS even in high load. Therefore, the PCAS performs much better when dealing with abrupt data within a small area.

In conclusion, we propose a PCAS system for contention resolution in PCE-based WSONs with distributed light path establishment. This system increases the resource utilization rate and reduces the backward blocking due to outdated information. In this method, instead of a blocking and throwing system, we adopt a blocking and switching system. The reserved path design borrows the P-cycle scheme, which is used for failure restoration. Simulation results show that the PCAS scheme is effective in mitigating the backward blocking, especially under concentrated load. Therefore, the PCAS scheme over a backup path is a suitable contention resolution in WSONs.

This work was supported in part by the National “863” Program of China (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 (SRFDP) of Ministry of Education of China (No. 200800130001), the Program for Changjiang Scholars and Innovative Research Team (PCSIRT, No. IRT0609), the International S&T Cooperation Program of China (ISTCP, No. 2006DFA11040), the New Century Excellent Talents in University (NCET, No. 06-0090), and Huawei Technologies Co. Ltd.

References

1. Y. Lee and G. Bernstein, “Path computation element communication protocol (PCEP) requirements and extensions for the support of wavelength switched optical networks” (Oct. 29, 2007) http://tools.ietf.org/html/draft-lee-pce-wson-routing-wavelength-00 (June 26, 2009).
2. L. Berger, (ed.) “Generalized multi-protocol label switching (GMPLS): signaling functional description” (Jan. 2003) http://www.ietf.org/rfc/rfc3471.txt?number=3471 (Dec. 15, 2008).
3. X. Yuan, R. Melhem, R. Gupta, Y. Mei, and C. Qiao, Photonic Network Commun. **1,** 207 (1999).
4. K. Lu, G. Xiao, and I. Chalamtac, IEEE/ACM Trans. Networking **13,** 187 (2005).
5. A. Farrel and J.-P. Vasseur, “A path computation element (PCE)-based architecture” (Aug. 2006) http://www.rfc-editor.org/rfc/rfc4655.txt (Dec. 15, 2008).
6. L. Wang, J. Zhang, G. Gao, Y. Liu, X. Chen, Y. Zhao, Y. Yao, and W. Gu, IEEE Commun. Lett. **12,** 593 (2008).
7. G. Wang, H. Wang, and Y. Ji, Chin. Opt. Lett. **6,** 628 (2008).
8. A. K. Garg and R. S. Kaler, Chin. Opt. Lett. **6,** 244 (2008).
9. L. Kleinrock, *Queueing Systems Vol.1 Theory* (Wiley, New York, 1975).
10. W. D. Grover and D. Stamatelakis, in *Proceedings of IEEE International Conference on Communications* **1,** 537 (1998).
11. D. A. Schupke, C. G. Gruber, and A. Autenrieth, in *Proceedings of IEEE Global Telecommunications Conference 2004* **3,** 2761 (2004).
12. J. Doucette, D. He, W. D. Grover, and O. Yang, in *Proceedings of the Fourth International Workshop on Design of Reliable Communication Networks* 212 (2003).
13. H. Mittelmann, “Mixed integer linear programming benchmark” (Dec. 20, 2008) http://plato.asu.edu/ftp/milpf.html (Nov. 29, 2009).