Cut-through Optical Path Control Technology for a Terabit-class Super-network

Kenichi Matsui†, Takeshi Yagi, Yuuichi Naruse, and Junichi Murayama

Abstract

Based on traffic monitoring data, our cut-through optical path control technology allocates cut-through optical paths between provider edge routers on routes with a high traffic demand. To optimize the optical path arrangement, a network control server monitors traffic demand, calculates the optimal path allocation, and continuously configures the optical paths. Since this server ensures highly efficient network resource usage, high-performance packet forwarding is achieved economically.

1. Introduction

As described in the first paper [1] in this issue, we are designing a terabit-class super-network (TSN) architecture that applies IPv6-based connectionless forwarding to a generalized multiprotocol label switching (GMPLS) [2] network. In the TSN architecture, traffic is concentrated at electrical provider (P) routers and this could decrease network capacity. To ease this traffic concentration and keep the forwarding performance high, we apply cut-through optical paths. A cut-through optical path is established between provider edge (PE) routers that are exchanging heavy traffic. This path increases the bandwidth capacity between PE routers because the capacity of a cut-through optical path is larger than that of IPv6 routes passing through electrical P routers. Moreover, the use of cut-through optical paths avoids traffic congestion at the electrical P routers. As the traffic demand pattern changes, ineffectual, obsolete cut-through optical paths are deleted and new ones are established, improving the efficiency of packet forwarding. Therefore, a TSN network can achieve the same forwarding performance as a GMPLS network with fewer optical paths.

To achieve the maximum efficiency, cut-through optical paths should be recalculated and reassigned dynamically according to changes in traffic demand. Additionally, the reassignment should be optimized according to traffic demand taking account of the network resources available. There are two issues in implementing such cut-through optical path control technology.

The first issue is resource management. At present, multiprotocol label switching (MPLS) routers monitor traffic demand, exchange the monitored traffic data, and calculate path allocations and set up the paths accordingly [3]. If network resources are managed by each PE router in such a distributed manner, it is difficult to achieve uniform resource allocation. For example, there is a risk of deadlock in resource allocation. Thus, this type of network resource management could not provide optimal resource sharing.

The second issue is calculation time. To achieve highly efficient cut-through optical path, we should solve the combinatorial problem concerning the mapping of IPv6 flows to an optical path [4]. Since the path calculation time increases in proportion to the number of network constraints, it is difficult to finish a calculation during a dynamic path reassignment.

To solve these problems and provide dynamic and optimum cut-through optical path control, we propose two methods: a server-based resource management method and a method of grouping IPv6 flows according to their efficiency. The first method enables network resources to be managed under a unified pol-

† NTT Information Sharing Platform Laboratories
Musashino-shi, 180-8585 Japan
E-mail: matsui.kenichi@lab.ntt.co.jp

icy because a network control server (NCS) can manage resources centrally. The second method reduces the path calculation time because only the effective IPv6 flows are selected for calculation from among the massive number of IP flows.

This paper is organized as follows. Section 2 briefly reviews cut-through optical paths. Section 3 describes their control issues and section 4 proposes methods that solve these issues. Section 5 describes the steps in path assignment. Section 6 evaluates the effectiveness of cut-through optical paths.

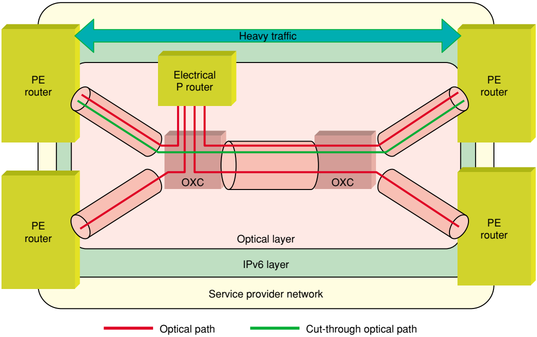
2. Cut-through optical paths

2.1 Optical path layouts in a GMPLS network and TSN

In a GMPLS network, to ensure full reachability between PE routers by optical paths alone, the paths would need to be arranged in a mesh topology. Since the number of optical paths is the same as the square of the number of PE routers, this full-mesh arrangement is not scalable. In the TSN architecture, IP-in-IPv6 overlay networking technology [5] arranges optical paths in a hub-and-spoke topology with electrical P routers forming the hub nodes and PE routers forming the leaf nodes. Since the number of optical paths is the same as the number of PE routers, this hub-and-spoke arrangement is scalable.

2.2 Cut-through optical path control

In a hub-and-spoke topology, an electrical P router can become a forwarding bottleneck depending on the traffic concentration. To ease this traffic concentration and keep the forwarding performance high, we apply cut-through optical paths (**Fig. 1**). Since a cut-through optical path is established only between PE routers that are exchanging heavy traffic, the total number of optical paths is smaller than in a full-mesh topology. Therefore, when cut-through optical paths are combined with the hub-and-spoke topology, the arrangement is still scalable. To achieve maximum efficiency, cut-through optical paths should be reassigned dynamically according to changes in traffic demand. The reassignment should be optimized according to traffic demand taking into account the network resources available. To operate cut-through optical paths, we require a control technology that implements these functions.



PE router: provider edge router, P router: provider router, OXC: optical cross connect

Fig. 1. Cut-through optical path.

3. Issues in cut-through optical path control

3.1 Problems in network resource management

For the present, some router vendors are implementing the dynamic label switched path (LSP) reassignment mechanism on MPLS routers. They are doing this by the distributed approach, appending a traffic engineering function to routers. In a service provider (SP) network, the MPLS routers are arranged as PE routers and P routers. PE routers are interconnected by LSPs in a full-mesh configuration. Each PE router monitors traffic on the terminated LSPs and exchanges the monitored traffic information by extended routing protocols. To ease the traffic concentration on P routers, a PE router calculates and sets up an LSP arrangement according to traffic demand.

Since each router calculates paths separately, it is difficult to manage all of the network resources under a unified policy. Thus, quick and optimal resource sharing is impossible. For example, if many routers try to reserve the same link resource or the same PE router interfaces, deadlock will result. **Figure 2** shows PE router 1 and PE router 2 trying to reserve

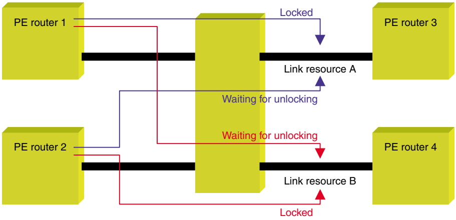


Fig. 2. Deadlock in a distributed resource management.

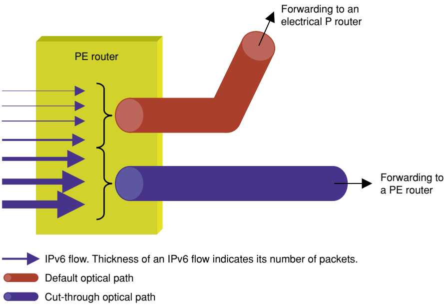


Fig. 3. Mapping IPv6 flows to optical paths.

link resources A and B simultaneously and ending up deadlocked. This deadlock will cause a delay in path setup and degrade the forwarding performance.

3.2 Problems in path calculation

The second problem is in path calculation. Since optical paths are expensive to operate, the required number of cut-through optical paths should be reduced. To achieve highly efficient cut-through optical paths, as many as possible of the IPv6 flows should be mapped to a cut-through optical path. To optimize this problem, we should solve the combinatorial problem concerning the mapping of IPv6 flows to an optical path (**Fig. 3**). The objective of the combinatorial problem of packing IPv6 packets is to maximize the number of IPv6 packets in an optical path. The main constraint in the problem is the bandwidth of IPv6 flows [4].

This problem is similar to a knapsack problem[*1] in terms of finding the most valuable set of IPv6 flows that fit in an optical path of fixed capacity. The calculation time of the knapsack problem increases with the number of constraints. To ensure the optical path calculation method can efficiently provide dynamic cut-through optical path assignment, the calculation time of the problem should be shortened.

4. Proposed methods for cut-through optical path control

4.1 Server-based path control architecture

To solve the resource management problem, we use a server-based optical path control architecture. As shown in **Fig. 4**, the network control server (NCS) calculates a path centrally and assigns an optical path.

*1 The knapsack problem is a combinatorial problem. Given items of different values and volumes, the problem is to select the most valuable set of items to fit in a container of fixed volume.

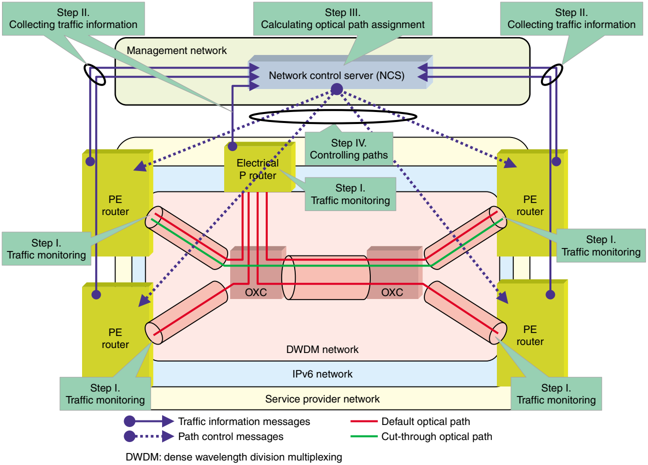


Fig. 4. Operations of cut-through optical path assignment.

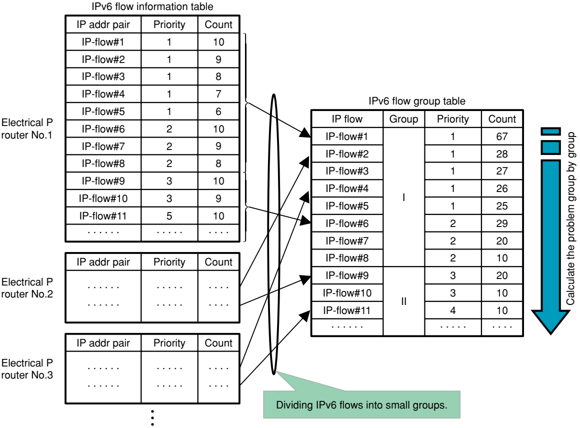


Fig. 5. Grouping IPv6 flows.

Since the server approach makes it possible to manage all of the network resource information and traffic information, it can consider the priority of all paths and allocate the network resources effectively (Details of Fig. 4 are described in section 5.).

4.2 Grouping IPv6 flows according to their efficiency

To shorten the calculation time of the combinatorial problem of packing IP flows, we propose a method that reduces the number of IP-flows to be calculated. In this method, we focus on the difference in influence of IP-flows. IP-flows with a high traffic volume and high priority will improve the efficiency of cut-through optical paths more than ones with a low traffic volume and priority. Therefore, the efficiency of cut-through optical paths that are calculated using only a subset of IP-flows that have high traffic volume and high priority may differ little from the efficiency of those calculated using the whole set of IP-flows.

We divide the IPv6 flows into small groups as shown in **Fig. 5**. First, we make an IPv6 flow list of each electrical P router and sort it according to priority and number of IPv6 packets. Second, we merge these tables and sort the flows and divide them into small groups. Each group consists of IPv6 flows having the same priority and number of packets. Third, we sort the groups according to total priority and number of IPv6 packets. Then, we start to calculate the combinatorial problem of packing IP flows from the top of the group list. When we finish the calculation for one group, we go on to the next group in the list. If there are few optical path resources, the calculation is terminated. Since each group contains a small number of IPv6 flows, the calculation time of this step should be shorter.

5. Cut-through optical path assignment

With the methods mentioned in section 3, cut-through optical path will be reassigned by the following steps: monitoring IPv6 traffic demands, collecting the traffic information, calculating optical path assignment, and setting up optical paths (Fig. 4). In this section, we briefly explain what the NCS, PE routers, and electrical P routers do during each step.

5.1 Traffic monitoring

Traffic monitoring in TSN is economically achieved by use of IP-in-IPv6 connectionless forwarding technology. To suppress the amount of information collected, each PE router counts only the number of IPv6 packets sent over cut-through optical paths; the total number of IPv6 packets from each PE router sent to electrical P routers is counted at the electrical P routers themselves. This monitoring method is discussed in more detail in section 3.1 of [5].

5.2 Collecting traffic information

The NCS periodically accesses electrical P routers and PE routers and collects the monitored information from them. Then it combines these two sets of information to provide an overall IPv6 traffic picture. Additionally, to detect a significant upsurge in the number of IPv6 packets in IP flows between polling intervals, electrical P routers and PE routers send a trap message to the NCS if the number exceeds a threshold value. When the NCS receives the trap, it polls the routers immediately to update the present distribution of IP flows.

5.3 Calculating optical path assignment

Based on the collected traffic information, the NCS periodically optimizes the cut-through optical path arrangement that maximizes the efficiency of optical paths. To do this, first, the NCS calculates the combinatorial problem of packing IP flows with the IPv6 flow grouping method and allocates network resources optimally. Second, the NCS compares the total number of packets transmitted over the currently assigned cut-through optical paths with that in the newly calculated ones. It subtracts the former from the latter. Then, if the difference exceeds a threshold value, the NCS decides to reassign the cut-through optical paths.

5.4 Setting up cut-through optical paths

When the NCS decides to reassign the cut-though optical path, it removes one cut-through optical path and sets up another. To set up a cut-through optical path, it orders the ingress PE router to set up an optical path to the egress PE router. The ingress PE router sets up a new one by generating a path setup signaling message. The signaling message will be forwarded by the optical P routers on the optical path route, and each optical P router sets up the optical path. In this way the cut-through optical path is set up. The details of setting up cut-through optical paths, such as the cooperative reassignment of IPv6 routes and optical paths, are described in [4].

5.5 Feedback

The above four steps in the operating process are repeated continually. This iterative operation keeps high efficiency of cut-through optical paths in spite of changes in traffic demands.

6. Evaluation of the effect of cut-through optical paths

As described in section 2, the cut-through optical path control technology can achieve the most efficient optical paths by matching the forwarding capacity of optical paths to traffic demand changes. In this section, for the same traffic demand situation, we compare the number of optical paths required for a conventional full-mesh optical path arrangement and the cut-through optical path arrangement. Through this comparison, we evaluate the efficiency of cut-through optical paths.

6.1 Calculation of the number of optical paths

- *Notation*
  - N_{op} total number of optical paths
  - N_{pe} total number of PE routers
  - T_{ij} number of IPv6 packets between the nth PE router (PE_i) and PE_j
  - T_i total number of IPv6 packets from PE_i
  - N_{dop} total number of default optical paths
  - N_{copi} number of cut-through optical paths from PE_i
  - N_{cop} total number of cut-through optical paths

- *N_{op} in a full-mesh optical path arrangement*

  When optical paths are arranged between PE routers in a full-mesh topology, the number of optical paths (N_{op}) is given by

$$N_{op} = N_{pe}\,(N_{pe} - 1). \tag{1}$$

- *N_{op} in a hub-and-spoke optical path arrangement with cut-through optical paths*

  The number of optical paths (N_{op}) in a hub-and-spoke topology is given by the sum of the number of default optical paths (N_{dop}) and the total number of cut-through optical paths (N_{cop}).

- *Calculation of N_{dop}*

  A default optical path is assigned from each PE router to an electrical P router. Thus, the total number of default optical paths between PE routers and electrical P routers is the same as the number of PE routers (N_{pe}). Therefore, N_{dop} is given by

$$N_{dop} = N_{pe}. \quad (2)$$

- *Calculation of N_{cop}*

A cut-through optical path is assigned according to the traffic distribution between PE routers. In this paper, we hypothesize that the traffic distribution between PE routers follows Zipf's law. The total number of IPv6 packets from PE_i (T_{ij}) is given by Eq. (3), where C is a constant that normalizes the value of the original Zipf's law.

$$T_{ij} = \frac{CT_i}{j}, \text{ where } C = \frac{1}{\sum_{i=1}^{N_{pe}} \frac{1}{i}} \quad (3)$$

A cut-through optical path is assigned to a pair of PE routers that are exchanging a high volume of traffic. We assume that a proportion α ($0 \le \alpha \le 1$) of the IP flows from PE_i to other PEs is accommodated by cut-through optical paths. Equation (4) expresses this situation.

$$\sum_{j=1}^{N_{COPi}} T_{ij} = \alpha T_i \quad (4)$$

Substituting Eq. (3) into Eq. (4) yields

$$\sum_{j=1}^{N_{COPi}} \frac{1}{j} = \alpha \sum_{i=1}^{N_{pe}} \frac{1}{i} \quad (5)$$

Equation (5) is a type of a harmonic progression[*2]. Since we hypothesize that N_{pe} is large, we can calculate approximate values of the total number of cut-through optical paths from PE_i (N_{copi}). N_{copi} is given by Eq. (6) where γ is Euler's number[*3].

$$N_{COPi} = e^{\{\alpha(\gamma + \ln N_{pe}) - \gamma\}} \quad (6)$$

- *N_{op} in a hub-and-spoke optical path arrangement with cut-through optical paths*

The number of optical paths (N_{op}) in a hub-and-spoke optical path arrangement with cut-through optical paths is given by the sum of Eqs. (2) and (6):

$$\begin{aligned} N_{op} &= N_{dop} + N_{cop} \\ &= N_{pe} + \sum_{i=1}^{N_{pe}} N_{COPi} = N_{pe}\left(e^{\{\alpha(\gamma + \ln N_{pe}) - \gamma\}} + 1\right). \quad (7) \end{aligned}$$

6.2 Evaluation

We compared the number of optical paths (N_{op}) of a full-mesh optical path arrangement (Eq. (1)) and that of a hub-and-spoke optical path arrangement with cut-through optical paths (Eq. (7)). As an example, consider the case where α =0.3. This means 70% of the traffic in an SP network is forwarded via electrical P routers and 30% via cut-through optical paths. **Figure 6** is a graph showing N_{op} in the situation. For 10,000 optical paths, the hub-and-spoke path arrangement can accommodate 1441 PE routers while the full-mesh path arrangement can accommodate only 100 PE routers. Thus, cut-through optical path technology increases the number of PE routers that can be accommodated by 1441%. This result indicates that cut-through optical paths improve the forwarding capacity of optical paths.

7. Conclusion

Cut-through optical path control technology provides highly effective dynamic cut-through optical path assignment. The network control server monitors all of the traffic information using a small amount of traffic monitored data, assigns network resources optimally, and calculates the cut-through optical path allocation in a short time. Applying this technology to a terabit-class super-network will achieve high-performance packet forwarding and effective network resource utilization.

8. Acknowledgment

This research was supported by a grant from the Telecommunications Advancement Organization of Japan (TAO).

---

*2 Harmonic progression (H_n) is a progression of terms whose reciprocals from an arithmetic progression.

H_n is defined as $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$, where n is an integer.

*3 Euler's number (γ) is defined as $\gamma = \lim_{n\to\infty}\left(1 + \frac{1}{2} + \cdots + \frac{1}{n} - \log n\right)$.

H_n can be approximated as $H_n = \gamma + \log n$ if n is large.

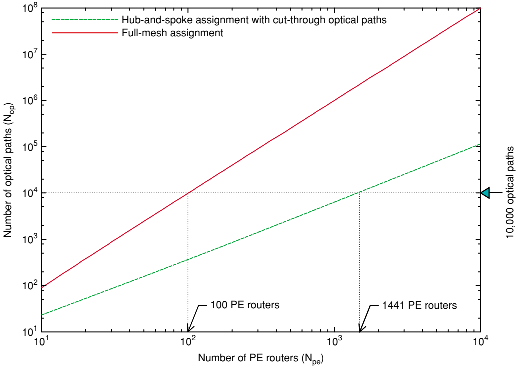


Fig. 6. Comparison of the number of optical paths.

References

[1] J. Murayama, K. Matsui, K. Matsuda, and M. Makino, "A Conceptual Design for a Terabit-class Super-networking Architecture," NTT Technical Review, Vol. 2, No. 3, pp. 12-20, 2004.

[2] A. Benerjee, J. Drake, J. Lang, B. Turner, D. Awduche, L. Berger, K. Kompella, and Y. Rekhter, "Generalized multiprotocol label switching: An overview of signaling enhancement and recovery techniques," IEEE Commun. Mag., pp. 144-151, Jul. 2001.

[3] X. Xipeng, A. Hannan, B. Bailey, and L. M.Ni, "Traffic engineering with MPLS in the Internet," Network, IEEE, Vol.14, issue 2, pp. 28-33, Mar./Apr. 2000.

[4] K. Matsui, T. Sakurai, M. Kaneda, and H. Ishii, "A Multi-Layered Traffic Engineering Architecture for the Electrical/Optical Hybrid Network," Proc. of 2003 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM'03), Victoria, B.C., Canada, Vol. II, pp. 293-296, Aug. 2003.

[5] Y. Naruse, T. Yagi, K. Matsui, and J. Murayama, "IP-in-IPv6 Overlay Networking Technology for a Terabit-class Super-network," NTT Technical Review, Vol. 2, No. 3, pp. 21-31, 2004.

Kenichi Matsui
Secure Communication Project, NTT Information Sharing Platform Laboratories.

He received the B.E. degree in information engineering and the M.S. degree in information sciences from Tohoku University, Sendai, Miyagi in 1995 and 1997, respectively. He joined NTT in 1997. His work focuses on IP networking and his research interests include traffic engineering for optical IP networks and MPLS, on-demand QoS management, and managed IP multicast platforms. He is a member of the Institute of Electronics, Information and Communication Engineers of Japan (IEICE), the Information Processing Society of Japan, and the IEEE Computer Society.

Yuuichi Naruse
Research Engineer, Secure Communication Project, NTT Information Sharing Platform Laboratories.

He received the B.S. and M.S. degrees in physics from Tohoku University, Sendai, Miyagi in 1990 and 1992, respectively. He joined NTT in 1992. He has been engaged in R&D of ATM-LAN systems, Fibre-Channel network systems, and IP-VPN service platforms.

Takeshi Yagi
Secure Communication Project, NTT Information Sharing Platform Laboratories.

He received the B.E. degree in electrical and electronic engineering and the M.S. degree in science and technology from Chiba University, Chiba in 2000 and 2002, respectively. He joined NTT in 2002. His studies focus on IP-VPN architecture and his current research interests include IP routing and forwarding technology, traffic monitoring technology, and layer cooperation technology. He is a member of the Institute of Electrical Engineers of Japan (IEEJ) and IEICE.

Junichi Murayama
Senior Research Engineer, Secure Communication Project, NTT Information Sharing Platform Laboratories.

He received the B.E. and M. E. degrees in electronics and communication engineering from Waseda University, Tokyo in 1989 and 1991, respectively. Since joining NTT in 1991, he has been engaged in R&D of ATM networks, large-scale IP networks, and IP-VPN service platforms. He is a member of IEICE.