Benefits of restoration signaling message aggregation

Mukul Goyal, Jennifer Yates*, Guangzhi Li*, Wu-chi Feng§

CIS Department, The Ohio State University, DL395, 2015 Neil Ave, Columbus OH 43210, mukul@cis.ohio-state.edu

**AT&T Labs-Research, 180 Park Ave, Florham Park NJ 07932, {jyates,gli}@research.att.com*

§Dept. Comp. Sci. and Engg., Oregon Graduate Institute, 20000 NW Walker Rd, Beaverton OR 97006, wuchi@cse.ogi.edu

Abstract: Signaling aggregation provides a practical approach to fast failure recovery in path-based restoration schemes. This paper proposes several signaling aggregation schemes and demonstrates their potential for improving restoration performance compared with schemes without aggregation.

1. Introduction

Path based restoration schemes [1] have emerged as a cost effective approach to achieve fast recovery from failures in optical networks. These schemes use signaling along restoration (backup) paths to re-establish connectivity after a failure disrupts the primary paths of the connections. Current restoration signaling proposals [1,2] can be characterized as "per-connection" in nature since each failed connection is restored using a separate set of signaling messages. As the number of connections affected by a failure increases, there is a corresponding linear increase in the number of signaling messages generated to restore these connections and hence in the queuing delays suffered by these messages at the optical cross-connects (OXCs) [2]. These queuing delays impact failure recovery times which become unacceptably high for moderately large number of connections in the network. Clearly, if a single signaling message could restore multiple connections, the number of signaling messages can be reduced, presumably with a corresponding reduction in the queuing delays and recovery times. In this paper, we propose several restoration signaling aggregation schemes that combine individual signaling messages into an aggregate message. We demonstrate significant improvements in recovery times achieved with these schemes in comparison to per-connection signaling.

2. Signaling Aggregation Schemes

While we experimented with several proposals for the path-based restoration signaling [1,2], in this paper we focus on the procedure described in [1]. In this scheme, the OXC detecting a failure sends a *failure indication* (or *alarm*) message to the source OXC of each of the failed connections. Upon receiving the *alarm*, the source OXC initiates failure recovery by sending a *switchover request* message (or simply a *request*) towards the destination OXC along the restoration path. As the *request* travels through the intermediate OXCs, they select channels for the failed connection. Upon receiving the *request*, the destination OXC generates a *switchover response* message (or a *response*), which travels back towards the source OXC along the restoration path. As the intermediate OXCs receive the *response*, they initiate channel cross-connection for the failed connection and forward the *response* further upstream without waiting for the cross-connection to complete. The connection recovery is completed when all OXCs along the restoration path have finished channel cross-connections for the connection.

The first step in signaling aggregation is aggregating the *alarms* generated by the OXC detecting a failure. Rather than sending individual *alarm*s for each failed connection, the *alarm*s going to the same OXC can be combined into a single *aggregated alarm*. We refer to this as *alarm aggregation*, which, besides reducing the number of *alarms*, allows a source OXC to simultaneously learn about multiple failed connections.

Among the failed connections identified by an *aggregated alarm*, the connections with the same restoration path can be restored using common *aggregated request* and *response* messages. We refer to this scheme as *aggregation over common path* signaling. Further reduction in the number of signaling messages can be obtained if, rather than aggregating signaling messages going over same end-to-end path, the OXCs aggregate signaling messages going to the same neighboring (next-hop) OXC. We refer to such a scheme as *aggregation over next hop* signaling. In this scheme, the source OXC groups the failed connections according to the next OXC along their restoration paths and sends a single *aggregated request* for each group. If the OXC receiving an *aggregated request* is the destination for a sub-set of the connections identified therein, it generates an *aggregated response* for these connections and sends it back towards the source OXC along their common restoration path. The remaining connections are re-grouped according to the next OXC on their restoration paths and the OXC transmits a single *aggregated request* message for each group. This procedure is repeated at all OXCs receiving the *aggregated request* message. Thus, an *aggregated request* splits as the restoration paths of the failed connections diverge. Note that *aggregation over common path* is actually a special case of *aggregation over next hop*.

Using the above procedures, an OXC may receive multiple *aggregated request* and *response* messages during a short period of time. If we allow a signaling message to wait for a short while at an OXC after it is processed, then it is feasible for the OXC to re-aggregate multiple aggregated messages based on common next hops. We refer to this approach as *aggregation over next hop with delay*. Note that *aggregation over next hop* is a special case of *aggregation over next hop with delay*, when the delay is set to 0.

3. Benefits of Restoration Signaling Aggregation

To evaluate the performance of restoration signaling aggregation, we implemented the signaling schemes described above in the NS2 simulator [3]. The signaling message processing times were obtained from measurements in the AT&T prototype testbed [4] and are listed in Table 1. We simulated several topologies that demonstrated similar conclusions; we present results here for the 21-node, 26-link ARPA2 network [5] with connections established between randomly selected source and destination nodes. Each simulation involved sequentially failing every link in the network and observing the recovery times for the affected connections; the reported results are the 90th percentile of the observed recovery times. It is assumed that the OXCs are capable of executing cross-connections in parallel – this was shown in [2] to be a prime requirement for fast restoration.

We begin the performance evaluation by comparing the recovery times obtained with per-connection signaling with those obtained using the signaling aggregation schemes for different numbers of connections, as reported in Figure 1. As discussed earlier, the recovery times with per-connection signaling increase linearly with the number of connections. The *aggregation over common path* scheme dramatically improves recovery times by using common signaling messages to restore all connections sharing a common restoration path. As the number of connections increase from 2000 to 8000 in the ARPA2 topology, the average number of connections sharing a common restoration path increases from 4.8 to 19. Assuming that message processing times are unaffected by the increase in message processing load, the increase in number of connections has virtually no impact on the recovery times. The more sophisticated *aggregation over next hop* scheme provides further improvements in recovery times, as illustrated in Figure 2 which also shows marginally improved recovery times achieved with the *aggregation over next hop* when *response* messages are delayed by 1ms at each OXC to allow for greater aggregation. In general, the benefits of delayed signaling are marginal and it is difficult to determine the correct delay duration. Hence *aggregation over next hop* signaling appears to be the most appropriate practical choice.

The performance of the signaling aggregation schemes reported in Figures 1 and 2 is based on the assumptions that 1) the processing times for aggregated messages is the same as for non-aggregated messages; 2) there is no limit on the number of connections that can be signaled within a single message; and 3) the restoration path is calculated using shortest path routing. The first assumption is based on the observation made in other control plane protocols that the per-packet processing overhead dominates the total message processing [6]. The validity of the second assumption depends on how much connection-related information is carried in the signaling message. If it is only the connection ID, which would be adequate for shared mesh restoration [1], we can assume that fairly large numbers of connections can be signaled together in a message. Finally, the third assumption will not be valid if the restoration paths are selected so as to increase the resource sharing among connections [7]. In this case, connections with a common source and destination may not share the same restoration path which will presumably have a significant effect on the performance of the *aggregation over common path* scheme. The performance of other signaling aggregation schemes will also be affected since the restoration path will no longer be the "shortest" available path. Figure 3 illustrates the combined effect on the performance of signaling aggregation schemes in comparison to per-connection signaling when the above assumptions do not hold and instead 1) the message processing time increases with the number (n) of signaled connections - specifically the message processing time is ($A + B \times n$) where A is the fixed per-packet overhead and B (= $A/10$) is the time required to process each connection signaled by the message; 2) one message can signal at most 10 connections and 3) the restoration paths are calculated so as to increase restoration resource sharing using the scheme described in [7]. Note that the recovery times obtained with the signaling aggregation schemes under these conditions remain less than one-third of the recovery times obtained with per-connection signaling. Additionally, there appears to be little difference in the recovery times achieved with different signaling aggregation schemes.

4. Conclusions

In this paper, we proposed and evaluated several signaling aggregation schemes that reduce the number of signaling messages, thus avoiding long queuing delays during restoration signaling for large numbers of connections. By incorporating the proposed aggregation mechanisms, restoration signaling can continue to provide fast recovery from network failures even for very large number of connections in the network.

5. References

1. J.P. Lang *et al.*, "Generalized MPLS recovery functional specification," draft-bala-gmpls-recovery-functional-00.txt, Internet Draft – Work in Progress, August 2002.
2. M. Goyal, G. Li, and J. Yates, "Shared mesh restoration: a simulation study," *Optical Fiber Commun. Conf.*, March 2002.
3. The Network Simulator – ns-2 , http://www.isi.edu/nsnam/ns/.
4. D. Wang, G. Li, J. Yates, P. Sebos, C. Kalmanek, and A. Greenberg, "Optical NNI inter-working demonstration," *IEEE/LEOS Summer Topical Meeting,* July 2002.
5. K.C. Lee and V. Li, "A Wavelength-Convertible Optical Network," *IEEE/OSA Jour. Lightwave Tech.,* Vol. 11, No. 5/6, pp. 962-970, May/June 1993.
6. A. Shaikh, and A. Greenberg, "Experience in black-box OSPF measurement," *ACM SIGCOMM Internet Measurement Workshop,* November 2001.
7. G. Li, D. Wang, C. Kalmanek, and R. Doverspike, "Efficient distributed path selection for shared restoration connections," *IEEE Infocom 2002,* June 2002.

Table 1 Simulation Parameters

Processing delay for Request and ALARM messages.	0.418ms
Processing delay for Response messages.	0.326ms
Forwarding delay for a message in transit.	0.1ms
Channel cross-connection delay	2-3ms

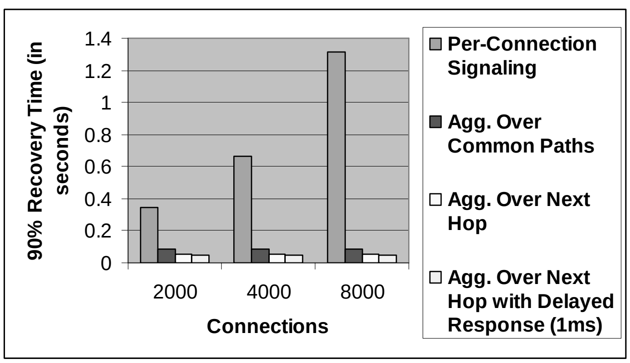

Figure 1 Comparing 90% Recovery Times obtained with Per-Connection Signaling and different Signaling Aggregation schemes.

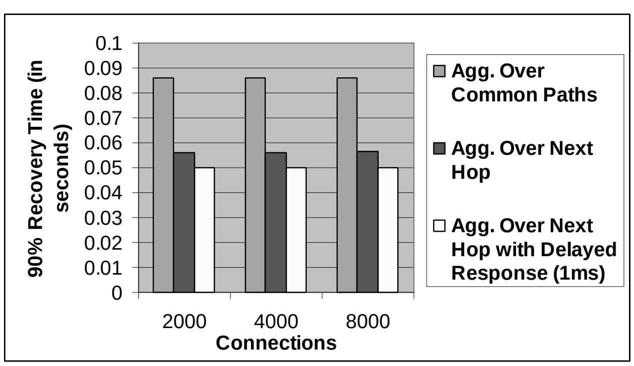

Figure 2 Performance of different Signaling Aggregation schemes.

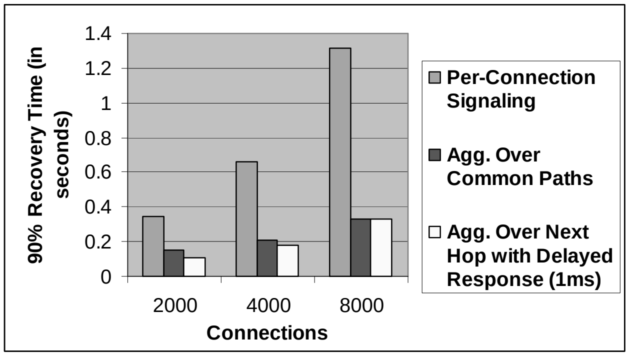

Figure 3 Performance of Signaling Aggregation with message processing times dependent on number of connections being signaled, limited numbers of connections signaled per message and non-shortest path routing.