

One-to-Many Multicast Restoration Based on Dynamic Core-Based Selection Algorithm in WDM Mesh Networks

I-Shyan Hwang, *San-Nan Lee, Zen-Der Shyu and Kang-Peng Chen

Abstract-This paper proposes a novel dynamic core-based selection (DCS) algorithm for the multicast restoration in WDM mesh networks. The core-based fault tolerance scheme provides a flexible way to control a number of core nodes with less control overheads for searching the routing path, wavelength assignment (RWA) and restoration paths when fault occurs in the one-to-many multicast domain. Compared with the source-based scheme, core-based schemes are easier to maintain, and specifically scalable in large-scale topologies. In the core-based fault tolerance scheme, k -tuple domination nodes are selected to form a minimum sized vertex subset such that each vertex in the graph is dominated by at least k vertices, where the k is defined as two in this paper. The proposed DCS algorithm is defined as each node in multicast tree session must be directly connected to at least one core node in multicast tree session and also has to be directly connected to at least one core node out of multicast tree session. The primary aim of this work is to provide the scalable and fast local survivability based on the information from core nodes. Simulation results show that the proposed algorithm outperforms the Dual Tree and MRLR algorithms in terms of total hop counts needed for all recovery paths and blocking probability for different network topologies.

Keywords: DCS, Multicast restoration, Local survivability, Dual Tree, MRLR.

I. Introduction

Optical networks with Dense Wavelength Division Multiplex (DWDM) can provide multiple data channels to supply high speed, high capacity to perform bandwidth-intensive multicast transmission service [1] technology, which uses wavelength division, markedly increases the bandwidth of existing optical fiber networks. Multicasting is the simultaneous delivery of data stream from one or more sources to multiple receivers by using one or more tree-based structures in the network. The benefit of multicast is that the source sends a single copy to the entire destinations, such that multicast can utilize bandwidth efficiency. In general, the one-to-many multicast routing algorithm can be classified as *source-based* mode and *core-based* mode.

In source-based mode, multicast tree rooted at the source node and connected to each member in the multicast session [2]. Then, data packets originating from

the source node are sent to all the destination nodes via the links of multicast tree. The Distance Vector Multicast Routing Protocol (DVMRP) [3] is a *dense mode* routing protocol which has good performance in network environments with high bandwidth and densely distributed multicast session members. In heterogeneous Internet environment, it potentially has to support many thousands of multicast groups, each of which may be sparsely distributed, so this technique does not scale well. The Protocol Independent Multicast-Dense Mode (PIM-DM) [4] is an efficient protocol when most receivers are interested in the multicast data, but does not scale well across larger domains in which most receivers are not interested in the data. The main difference between DVMRP and PIM-DM is that the PIM-DM introduces the concept of protocol independence such that any underlying unicast routing protocol to perform reverse path forwarding checks. The Multicast Open Shortest Path First (MOSPF) [5] is an extension to the Open Shortest Path First (OSPF) protocol to support multicast routing and allows routers to share information about group memberships.

In core-based mode, some nodes for each group are selected as the core nodes and multicast tree rooted at core node and constructed to span all the group members. Data packets flow from any source to its parent and children, such as Core-based Tree (CBT) [6] and Protocol Independent Multicast-Sparse Mode (PIM-SM) [7]. The CBT may build one or multiple core-based bidirectional trees which are shared by all of the group's senders and receivers with core node selection algorithm. However, the traffic may be concentrated at core nodes, the scalability and more management cost should be considered. The PIM-SM is an efficient routing protocol to multicast groups that span all destinations distributed sparsely in inter-domain network environment. The core-based scheme is preferable to source-based for multiple sources in the multicast group. The advantage of core-based scheme has less control overhead of a single shared tree rather than multiple trees. Recent researches in core-based scheme focus on the multi-core selection by k -center or r -dominating in multicast trees [8]. Core node selection algorithms can be classified as (1) one core node, and (2) multiple core nodes selection [9].

The rest of this paper is organized as follows. Section II describes related work of survivability scheme in multicast. Next, Section III proposes a novel dynamic core-based selection (DCS) algorithm for one-to-many multicast restoration in WDM mesh networks, which is capable of providing overall protection for optical nodes and fibers. Section IV compares the system performance of the proposed algorithm with the Dual Tree [10] and Multiple Ring-based Local Restoration (MRLR) [11] algorithms in terms of total hop counts needed for all

I-Shyan Hwang is a professor in the Department of Computer Science & Engineering at the Yuan-Ze University, Chung-Li, Taiwan (ishwang@saturn.yzu.edu.tw).

San-Nan Lee is a assistant professor in the Department of Computer Science and Information Engineering at Vanung University, Chung-Li, Taiwan (Tel:886-3-4515811~72607, Fax:886-3-4515811~72000, grand@vnu.edu.tw).

Zen-Der Shyu is currently pursuing the Ph. D degree at the Yuan-Ze University. Since 1995, he served as an instructor in the Army Academy R.O.C.

Kang-Peng Chen received M.S. degree in Department of Computer Science and Engineering from the Yuan-Ze University, Chung-Li, Taiwan in 2008.

recovery paths and blocking probability for different network topologies. Finally, conclusions are drawn in Section V.

II. Multicast Fault Tolerance

With the intensive bandwidth in DWDM optical fibers, a network function failure, such as network equipment corrupted or fiber cut, will cause serious data loss. For protecting multicast session against failure, network survivability must be considered in designing DWDM networks. The light-tree protection for multicast has more challenging than the unicast in optical networks, since a link or a node failure will affect several destinations at the same time. So keeping multicast session functionally needs a heuristic procedure and the objective is to achieve high reliability and fast recoveries with minimum backup resource and cost in the multicasting scheme. The preplanned fault tolerance with backup rerouting algorithm can provide fast recovery time, and it is suitable for one-to-many multicast with different backup structures belonging to the reduced topology which including a set of backup paths to minimize the multicast tree cost after recovery. There are three categories in multicast fault tolerance schemes as shown in Fig. 1.

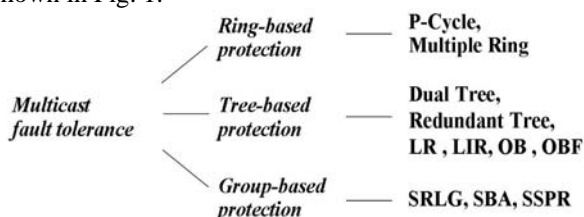


Fig. 1 Three categories of multicast fault tolerance algorithms.

- **Ring-based protection:** Ring-based protection has been part of the transport network landscape for some time due to the wide deployment of ring network topologies. The first mesh-ring decomposition is known as the node cover [12], in which the set of rings is selected such that each node belongs to one or more rings and each link belongs to no more than one ring. A node cover does not necessarily cover all the links and the uncovered links remain unprotected. This problem can be eliminated using the ring cover [13]. Each link must belong to at least one ring and all links are protected, however, spare resources may be redundant. Unfortunately, deciding the minimal ring cover in non-planar networks is an NP-hard problem and is not scalable. This type of algorithm can, therefore, recover quickly from failures and determine how to allocate recovered traffic loads based on current traffic load and the network bandwidth along the restoration paths.
- **Tree-based protection:** Link protection and path protection [14] are parts of the tree-based protection. The link protection is to find a backup link between nodes, but it causes a huge cost and wasted bandwidth. The path protection establishes a disjoint path between each source-destination pair. The *redundant tree protection* is proposed in [15] to establish a new

multicast tree from original source to the destination nodes bypassing the existing tree nodes. In *dual tree protection* [10], affected leaf nodes are connected to an unaffected leaf node. It requires that the underlying network topology is a bi-connected graph, in which there are at least two vertex-disjoint or link-disjoint paths between any two nodes. Four multicast tree fault tolerance schemes are proposed in [16], namely,

- **Group-based protection:** The concept of Shared Risk Link Group (SRLG) [17] is a group of network links that share a common physical resource. After an initial failure and before reprovisioning, it cannot support restoration for additional failures to other light-paths in the same restoration group. Therefore, two or more working paths under the same failure risk cannot share the same protection resource. The Shared Bandwidth Assignment (SBA) [18] provides the ability using integer linear programming to allocate bandwidth and find backup segments by dynamic programming when several multicast working paths sharing the common backup path. The Shared Segment Protection with Reprovisioning (SSPR) [19] is similar to the SRLG that the SSPR can provide dual links fault tolerance and increase the bandwidth utilization.

The proposed dynamic core-based selection (DCS) algorithm is defined as each node in multicast tree session must be directly connected to at least one core node in multicast tree session and also has to be directly connected to at least one core node out of multicast tree session. The DCS algorithm is extended to core nodes selection which produces the core-based tree to enhance restoration ability.

III. Proposed Algorithm

In general, finding the k -tuple dominating set (k -DS) of a graph $G = \{V, E\}$, where nodes set $V = \{v_1, \dots, v_m\}$ with m nodes and links set $E = \{e_1, \dots, e_n\}$ with n links, that dominates the multicast tree nodes in a subset of core nodes S , where $S \subseteq V$, such that each node in the subset $V \setminus S$ is adjacent to at least k core nodes in S . The value of k is given by two in this paper. The proposed dynamic core-based selection (DCS) algorithm is defined as each node in multicast tree session must be connected to at least one core node in multicast tree session and also the node has to be connected to one core node out of multicast tree session.

Each core node has a core routing table (CRT) which records the information between current core node and other core nodes, such as hop counts, available wavelengths, and the dominating nodes of the current core node. The available wavelength is used to gather wavelengths usage statistic for each link of recovery path. In addition, each non-core node in multicast session has a non-core routing table (NCRT) which records the information of its dominated core nodes in terms of available wavelengths. Both tables are broadcasted through control channels periodically to other nodes for updating the information based on the network status.

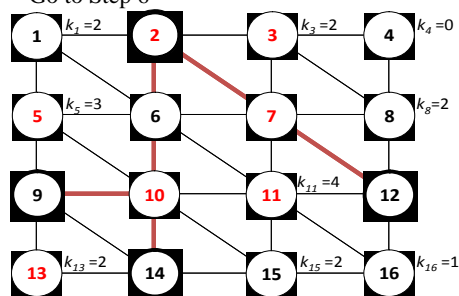
The proposed DCS algorithm is a distributed core selection algorithm which has two subsets: (1) multicast session member dominating set (MDS), (2) non-multicast session member dominating set (NDS). In such a way, each node in one-to-many multicast session will be dominated at least twice to enhance the restoration ability. Both MDS and NDS are established by the DCS algorithm and shown as follows:

Dynamic core-based selection (DCS) algorithm

Input: Multicast session tree T , network topology

Output: MDS and NDS

1. Initialize the MDS= $\{s\}$, where s is root
2. All parent nodes of leaf nodes in T are added to MDS
3. Delete leaf nodes and parent nodes of leaf nodes, and then form new sub tree T'
4. If T' is not null
5. Go to Step 2
6. Calculate $K=\{k_1, \dots, k_i\}$, where k_i is the number of connections node i in non-multicast session to the members in multicast session with one hop
7. The node which has the maximal degree in K is added to NDS
8. If all multicast session members are dominated by core nodes in NDS or no one can be dominated
9. Done
10. else
11. Go to Step 6



Source = {2} Destinations = {9,12,14}
 MDS = {2,7,10} NDS = {3,5,11,13}

Fig. 2 Example of the MDS and NDS

The MDS is defined as each node in the multicast session is dominated at least once by core nodes resided in multicast session. Initially, the root node of multicast session as the source node is assumed to be a core node, and then finds the core nodes from multicast leaf nodes to the source node, recursively. First, the parent nodes of leaf nodes are chosen as core nodes and the leaf nodes and parent nodes of the chosen core nodes are deleted to form a new sub tree. This process is to make sure that each node can be dominated once by core nodes. Based on the new sub tree, the parent nodes of the new leaf nodes are chosen as core nodes and the new leaf nodes and parent nodes of the chosen core nodes are deleted to form another new sub tree. Repeating the process until the source node or null node is left. For example, shown in Fig. 2.

The NDS is defined as each node in multicast session has at least once dominated by core nodes from non-multicast session. The selection of core node from non-multicast session is based on the connectivity for the

nodes in non-multicast session to the multicast session. Given $K = \{k_1, \dots, k_i\}$, where k_i is the number of connections from node i in non-multicast session to the members in multicast session with one hop. After calculating the $K = \{k_1, \dots, k_i\}$ of each node in non-multicast session to the multicast session, the node with maximum k_i is chosen as a core node. Repeating the process until all multicast session members are dominated or no other core nodes can be dominated. For example, shown in Fig. 2, calculating the k_i of all nodes in non-multicast session to the multicast session, e.g., node 11 have the k_{11} of 4, node 5 has the k_5 of 3, node 16 has k_{16} of 1 and node 4 has k_4 of 0, the other nodes have the k_i , where $i = \{1,3,8,13,15\}$ of 2. Then, the node 11 with k_{11} of 4 is chosen as a core node which can dominate the nodes 6, 7, 10 and 12; the node 5 with k_5 of 3 is chosen as another core node which can dominate the nodes 6, 9 and 10; the node 3 with k_3 of 2 is chosen as a core node which can dominate the nodes 2 and 7, and node 13 with k_{13} of 2 is chosen as a core node to dominate nodes 9 and 14. Then, the NDS includes nodes 3, 5, 11 and 13.

A restoration scheme for one-to-many multicast communication (RSOMMC) algorithm to resolve the node or link faults is proposed in this session. When fault occurs, the upstream node u , where $u \in T$, adjacent to the failed link or the failed node will broadcast the failure type and location, and determine which destination node(s) is affected based on the CRT from S_{core} or itself if the node u is a core node. Each affected destination node needs to find the recovery paths P with minimum hop counts based on the following equation:

$$P = P_v + P_{core} + P_d$$

, where P_v is the path from upstream node of the failed link or the parent node of the failed node to the dominating core node, S_{core} , P_{core} is the path between the dominating core node, S_{core} , near the fault and core node dominates the affected destination nodes, D_{core} , and P_d is the path from the dominating core node, D_{core} to the affected destination nodes. The P_v is null when the upstream node of the failed link or the parent node of the failed node v is a core node. Node fault can be seen as multiple links fault and tries to find the path from the upstream node of the failed link or the parent node of the failed node to each destination node individually. The proposed algorithm can maintain multicast session and choose the minimum hop count to be P_{core} . Once the

P_{core} is determined, P_v and P_d are also known then the recovery path is established. The pseudo code of restoration scheme for one-to-many multicast communication (RSOMMC) for link failure or node failure and the domination fault recovery scheme is described as follows:

Restoration scheme for one-to-many multicast

communication (RSOMMC)

Input: Core routing table (CRT), Non-core node routing table (NCRT)

Output: Recovery path

1. If a link fault occurs
2. The upstream node of the failed link, u , finds the core node S_{core} that dominated the node u , and core node D_{core_i} that dominates the affected destination nodes where i is the sequences of destination nodes
3. According the CRT and NCRT, the path P_v is selected from the source node to S_{core} based on the minimum hop counts, P_{core} is the path between S_{core} and D_{core_i} , and the P_d is the path between the D_{core_i} and destination node.
4. If there is another destination node
5. Go to step 2
6. else
7. Return $P = P_v + P_{core} + P_d$
8. If a node fault occurs
9. Sets all links connecting to the node are failed
10. If all links are recovered, then done
11. else
12. Go to step 2
15. Done

Domination fault recovery algorithm

Input: MDS, NDS, CRT and NCRT

Output: CRT and NCRT

1. If fault occurs
2. Executes the RSOMMC algorithm
3. Execute the dynamic core-based selection (DCS) algorithm to update the MDS and NDS
4. Each core node updates the CRT
5. Each non-core node updates the NCRT
6. Done

● **Link failure**

Based on the Fig. 2, a minimum spanning tree with minimum hop counts for multicast session tree is established, node 2 is the source node and nodes 9, 12, and 14 are destination nodes. If the link 6–10 is cut and the affected destination nodes are nodes 9 and 14, the source node of the failed link is node 6 and the recovery path of source node to affected destination nodes can be obtained based on $P = P_v + P_{core} + P_d$, which is shown Fig. 3.

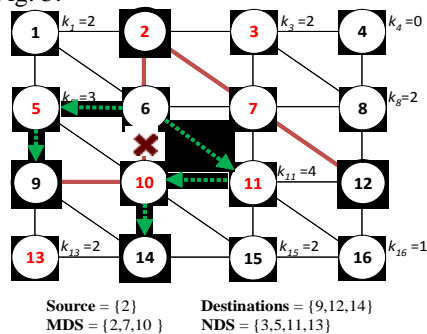


Fig. 3 Example of link failure

● **Node failure**

Based on the Fig. 2, if node 6 is corrupted, the links 2–6, 1–6, 5–6, 6–10, 6–11, 6–7 are disconnected and the affected destination nodes are 9 and 14. The parent node of the failed node is node 2 and the recovery path of parent node to affected destination nodes can be obtained based on $P = P_v + P_{core} + P_d$,

which is shown in Fig. 4.

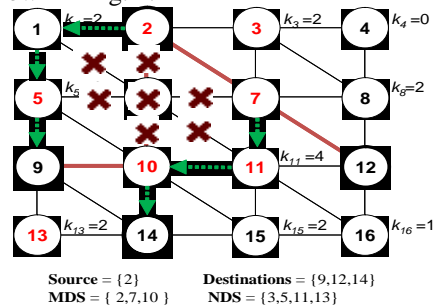


Fig. 4 Example of node failure

IV. Performance Evaluation

The performance of the proposed algorithm is analyzed by simulating the USANet (24 nodes and 43 links), ChinaNet (39 nodes and 72 links) and NTTNet (57 nodes and 81 links) networks. In the experiments, each link includes only one fiber on which 32 and 64 wavelength data streams can be transmitted, including two wavelengths used as bidirectional control channels and the conversion time of each converter is 1μs. Source node and destination nodes of multicast session are randomly selected, and the multicast group size (session size) is also randomly assigned from the interval {6,7,8}. Then, the minimum spanning tree (MST) approach is applied to construct the multicast tree for each multicast session. Multicast connection requests arrive according to a Poisson distribution $\lambda = \{1,2,\dots,10\}$ per minute, and their holding time is exponentially distributed (approximately a few minutes to hours). After the working tree is constructed for the multicast session, a single failure point (link failure or node failure) for each multicast session is randomly set in the multicast tree. Simulation programs are developed using the OPNET and the performance of the proposed algorithm herein is compared with those of the Dual Tree and MRLR algorithms in terms of the average hop count and blocking probability. The average hop count is defined as average hop count of all recovery paths per session and the blocking probability is defined as one or more restoration paths can not be established when fault occurs.

a. Link failure

● **Average hop count**

Figure 5 shows the average hop count of link failure versus request rate with the different numbers of channels for three different algorithms in the USANet, ChinaNet, and NTTNet. The proposed algorithm has the least average hop count needed than the other two algorithms for different network topologies. Moreover, when the request rate is higher and the channel number is smaller, the average hop count of recovery paths of three algorithms are all increased. It is interesting to notice that the average hop count of recovery paths of the three algorithms are increased when the network topology is more complicated and has more protection paths can be chosen in the USANet. Dual tree protection needs to discover fully disjoint paths; hence the hop count of each recovery path is larger than other algorithms. It can be also observed that the average hop count of the MRLR is

higher than the proposed algorithm because the ring-based protection in some situations gets more hop count route even shorter paths exist.

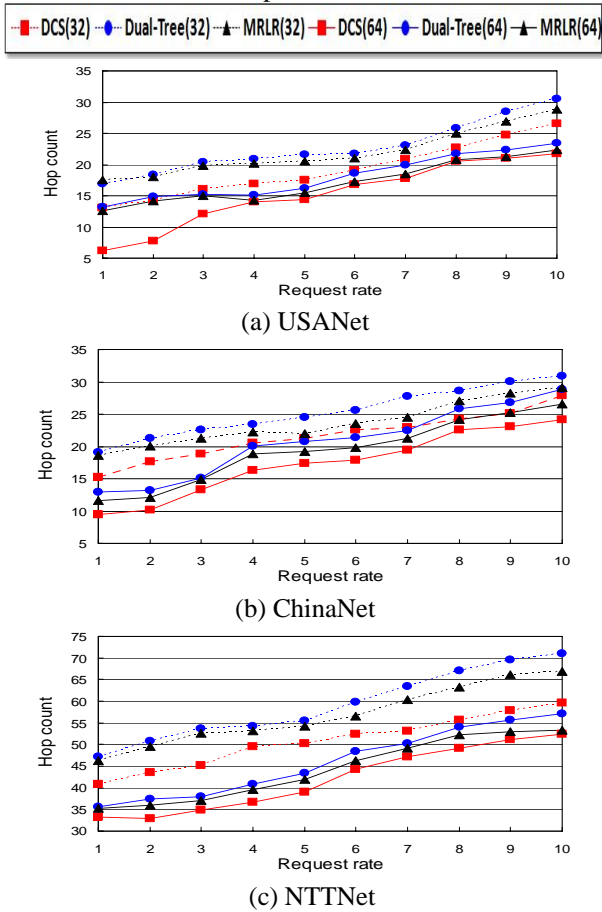


Fig. 5 Average hop count versus request rate with different numbers of channels.

● **Blocking probability**

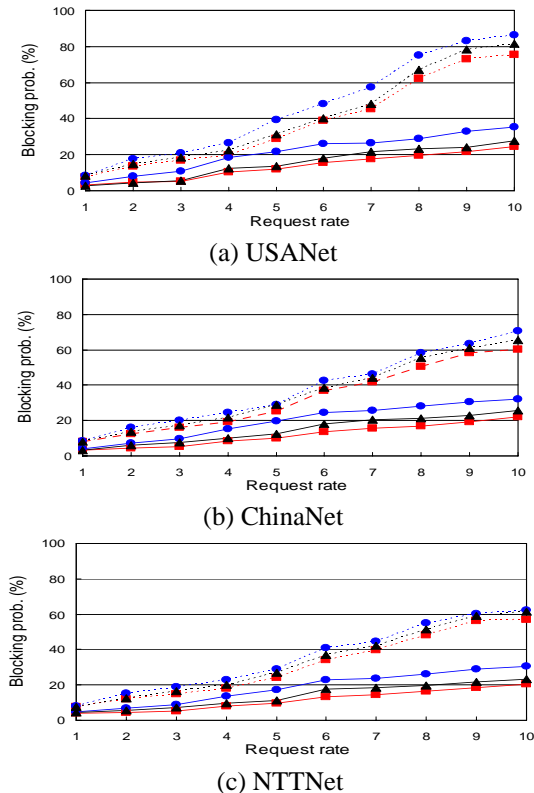


Fig. 6 Blocking probability versus request rate with different numbers of channels.

Figure 6 compares the blocking probability of link failure versus request rate with different numbers of channels for three different algorithms in the USANet, ChinaNet, and NTTNet. When the request rate is higher and the channel number is smaller, all the blocking probabilities of the three algorithms are increased. However, the blocking probabilities in all algorithms are improved when the network topology becomes complicated and more recovery paths can be chosen. Our proposed algorithm has the lowest blocking probability compared to the other two algorithms and the blocking probability increases smoothly as the request rate is getting higher.

b. Node failure

● **Average hop count**

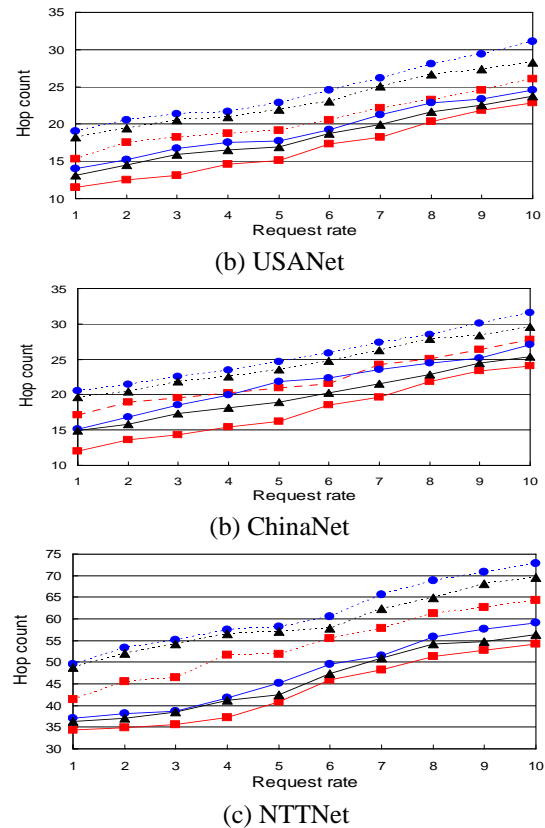


Fig. 7 Average hop count versus request rate with different numbers of channels.

Figure 7 shows the average hop count of node failure versus request rate with the different numbers of channels with three different algorithms in the USANet, ChinaNet, and NTTNet. The simulation results of single node fault are similar to link fault for different network topologies and it have more hop counts than the link fault scenario. When the request rate is higher and the channel number is smaller, the average hop count of recovery paths of three algorithms are all increased. It is noticed that the average hop count of recovery paths of the three algorithms increases when the network topology is more complicated and has more recovery paths can be chosen.

● **Blocking probability**

Figure 8 compares the blocking probability of node

failure versus request rate with different numbers of channels with three different algorithms in the USANet, ChinaNet, and NTTNet. It is similar to the link fault scenario that when the request rate is high and the channel number is small, all the blocking probabilities of the three algorithms are increased. The proposed algorithm has the lowest blocking probability compared to the other two algorithms and the blocking probability is increased smoothly as the request rate is getting higher. It also shows that the proposed algorithm still has stronger availability to handle node failure situations and higher scalability than other two algorithms.

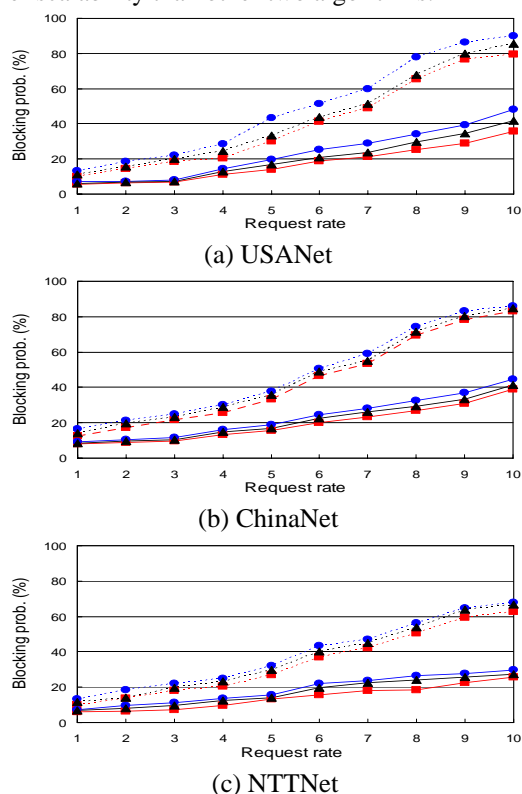


Fig. 8 Blocking probability versus request rate with different numbers of channels.

V. Conclusion

In this paper, we have proposed a dynamic core-based selection (DCS) algorithm for one-to-many multicast traffic with minimum spanning tree in WDM mesh networks. The proposed algorithm can provide the local recovery for link failure or node failure based on the information from core nodes. The simulation results show that the proposed algorithm outperforms the dual tree and MRLR algorithm in terms of the average hop count of recovery path and blocking probability, especially in the complicated network topology. The proposed algorithm is scalable to large-size mesh networks and can achieve high network survivability and reliability. Furthermore, when multiple faults occur, the proposed algorithm also can provide fast recovery ability and further research is the extension of more efficient core nodes selection scheme to satisfy the many-to-many multicast services. We believe that the proposed work

will meet the requirements of future networks for high performance, scalability and reliability.

References

- [1] B. Mukherjee, "WDM optical communication networks: progress and challenges", *IEEE Journal on Selected Areas in Communications*, Vol. 18, Issue 10, pp. 1810-1824, Oct. 2000.
- [2] A. Karaman and H. Hassanein, "Core-selection algorithms in multicast routing - comparative and complexity analysis", *Computer Communications*, Vol. 29, Issue 8, pp. 998-1014, May 2006.
- [3] D. Waitzman, C. Partridge, and S. Deering, "Distance vector multicast routing protocol", *IETF RFC 1075*, Nov. 1988.
- [4] A. Adams, J. Nicholas, and W. Siadak, "Protocol independent multicast - dense mode (PIM-DM): Protocol specification (Revised)", *IETF RFC 3973*, Jan. 2005.
- [5] J. Moy, "Multicast extensions to OSPF", *IETF RFC 1584*, Mar. 1994.
- [6] A.M. Al-Sharafi, M. Othman, M.N. Sulaiman, and S. Subramaniam, "Modifying the multicast tree in AMTree protocol using random core selection", *International Conference on Computer and Communication Engineering 2008*, Kuala Lumpur, pp. 323-327, 13-15 May 2008.
- [7] B. Fenner, M. Handley, H. Holbrook, and I. Kouvelas, "Protocol independent multicast - sparse mode (PIM-SM): Protocol specification (Revised)", *IETF RFC 4601*, Aug. 2006.
- [8] A. Karaman and H. Hassanein, "QoS-constrained core selection for group communication", *Computer Communications*, Vol. 30, Issue 7, pp. 1600-1612, May 2007.
- [9] A. Chakrabarti and G. Manimaran, "A case for tree migration and integrated tree maintenance in QoS multicasting", *Computer Communications*, Vol. 26, Issue 9, pp. 1007-1017, June 2003.
- [10] A. Fei, J. Cui, M. Gerla, and D. Cavendish, "A dual-tree scheme for fault-tolerant multicast", *IEEE International Conference on Communications*, Helsinki, Finland, Vol. 3, pp. 690-694, June 2001.
- [11] I.S. Hwang, R.Y. Cheng, and W.D. Tseng, "A novel dynamic multiple ring-based local restoration for point-to-multipoint multicast traffic in WDM mesh networks", *Photonic Network Communications*, Vol. 14, No. 1, pp. 23-33, Aug. 2007.
- [12] G. Maier, A. Pattavina, S.D. Patre, and M. Martinelli, "Optical network survivability: protection techniques in the WDM layer", *Photonic Network Communications*, Vol. 4, No. 3, pp. 251-269, 2002.
- [13] A.F. Elrefaie, "Multiwavelength survivable ring network architectures", *IEEE International Conference on Communications*, Geneva, Vol. 3, pp. 1245-1251, May 23-26, 1993.
- [14] J. Tapolcai, P.H. Ho; D. Verchere, T. Cinkler, and A. Haque, "A new shared segment protection method for survivable networks with guaranteed recovery time", *IEEE Transactions on Reliability*, Vol. 57, No. 2, pp. 272-282, June 2008.
- [15] M. Medard, S.G. Finn, R.A. Barry, and R.G. Gallager, "Redundant trees for preplanned recovery in arbitrary vertex-redundant or edge-redundant graphs", *IEEE/ACM Transactions on Networking*, Vol. 7, Issue 5, pp. 641-652, Oct. 1999.
- [16] C. Boworntummarat, L. Wuttisittikulij, and S. Segkhoonthod, "Light-trees based protection strategies for multicast traffic in transport WDM mesh networks with multi-fiber systems", *IEEE International Conference on Communications*, Paris, France, Vol. 3, pp. 1791-1795, June 2004.
- [17] L. Guo, H. Yu, and L. Li, "A new shared-path protection algorithm under shared-risk link group constraints for survivable WDM mesh networks", *Optics Communications*, Vol. 246, Issue 4-6, pp. 285-295, Feb. 2005.
- [18] D. Xu, Y. Xiong, and C. Qiao, "Novel algorithms for shared segment protection", *IEEE Journal on Selected Areas in Communications*, Vol. 21, Issue 8, pp. 1320-1331, Aug. 2003.
- [19] L. Luhua, L. Lemin, and W. Sheng, "Multicast protection scheme in survivable WDM optical networks", *Journal of Network and Computer Applications*, Vol. 31, Issue 3, pp. 303-316, Dec. 2006.