Optical quorum cycles for efficient communication

The final publication is available at Springer via `http://dx.doi.org/10.1007/s11107-015-0561-8`

Cory J. Kleinheksel · Arun K. Somani

arXiv:1608.05168v1 [cs.NI] 18 Aug 2016

Abstract Many optical networks face heterogeneous communication requests requiring topologies to be efficient and fault tolerant. For efficiency and distributed control, it is common in distributed systems and algorithms to group nodes into intersecting sets referred to as quorum sets. We show efficiency and distributed control can also be accomplished in optical network routing by applying the same established quorum set theory.

Cycle-based optical network routing, whether using SONET rings or p-cycles, provides the sufficient reliability in the network. Light-trails forming a cycle allow broadcasts within a cycle to be used for efficient multicasts. Cyclic quorum sets also have all pairs of nodes occurring in one or more quorums, so efficient, arbitrary unicast communication can occur between any two nodes. Efficient broadcasts to all network nodes are possible by a node broadcasting to all quorum cycles to which it belongs ($O(\sqrt{N})$).

In this paper, we propose applying the distributed efficiency of the quorum sets to routing optical cycles based on light-trails. With this new method of topology construction, unicast and multicast communication

Cory J. Kleinheksel · Arun K. Somani
Iowa State University, Ames, IA 50011
Tel.: 515-294-0442, Fax: 515-294-9273
E-mail: {cklein,arun}@iastate.edu

requests do not need to be known or even modeled a priori. Additionally, in the presence of network link faults, greater than 99% average coverage enables the continued operation of nearly all arbitrary unicast and multicast requests in the network. Finally, to further improve the fault coverage, an augmentation to the ECBRA cycle finding algorithm is proposed.

1 Introduction

We developed a novel method to enhance fault tolerance capabilities of cycles in a network. The cycle-based routing algorithm is to be used in the optical networks to achieve routing for both point-to-point and multi-point communication. The actual cycles are created using quorums. Within a cycle, multicasts to all nodes in that cycle is possible. The quorum intersection property and the use of cyclic quorums sets provide all of the unicast capabilities. Exploiting the same properties, we can achieve efficient broadcasts with $O(\sqrt{N})$ multicasts. These are significant results for modern and future optical networks facing dynamic heterogeneous traffic requests.

Fiber-optic lines make up the foundation of many networks across the globe. Some networks stretch hundreds of kilometers, while others are contained within buildings or rooms. These optical circuits are depended upon for high-speed communications in distributed algorithms, as much as they are needed for the arbitrary point-to-point communications.

Failures within a network are to be expected and can happen as much as every couple days [1]. Protecting

against these optical circuit faults is critical and there are many different approaches depending on the network needs and individual circumstances. SONET rings can be used to protect point-to-point and shared paths while enabling failure location. Using a pre-configured p-cycle backup [2], all node pair connections can be protected also.

Knowing the unicast or multicast requests a priori is often not possible. This constraint makes protection against faults in those arbitrary communication paths a challenge. An efficient all-nodes-to-all-nodes protection scheme supporting both unicast and multicast communication is necessary.

For efficiency and distributed control, it is common in distributed systems and algorithms to group nodes into intersecting sets referred to as quorum sets. In this paper (and the prior, shorten conference version [3]), it is shown that efficiency and distributed control can also be accomplished in optical network routing by applying the same established quorum set theory.

The rest of the paper is organized as follows. Sections 2, 3, and 4 establish the network model, node communication, and path routing / fault tolerance. In Section. 5, we apply the distributed efficiency of the quorum sets to routing optical cycles. This method of route construction allows for the efficient handling of dynamic unicast and multicast communication requests. Sections 6 and 7 address routing and solution efficiencies. Lastly, Section 8 analyzes the performance of our quorums set cycle routing techniques in the presence of network link faults.

2 Network model

No two fiber-optic networks are the same. Some stretch hundreds of kilometers, while other networks are contained within buildings or rooms. Regardless of the physical environment, these optical circuits are depended upon for high-speed communications. Thus, it is important to extract the network's critical components that affect its ability to deliver reliable, arbitrary point-to-point and multi-point communications.

These fiber-optic networks consist of several transmitters and receivers interconnected by fiber-optic cables. As you might expect, transmitters and receivers are typically found together and generically called an optical node. The cables form the links (i.e., edges) between those nodes, which leads to a convenient model of a network in terms of a graph $G = (V, E)$. V are the set of nodes in the network and E are the set of edges.

Edge (a_i, a_j) is a fiber-optic link connecting nodes a_i and a_j in the network, where $a_i, a_j \in V$ and $(a_i, a_j) \in$
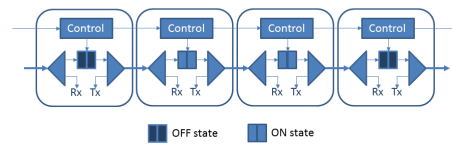

Fig. 1: Four nodes in a light-trail architecture

E. It is a general assumption that the same set of optical wavelengths are available on all edges in E. The number of wavelengths available per optical fiber is dependent on the fiber-optic cables and the transmitter/receiver pairs.

3 Light-trails

Lightpaths were a critical building block in the first optical communications, but required significant traffic engineering and aggregation to support point-to-point communication, or pay the penalty of low resource utilization on the fiber-optic link. Lightpaths cannot support multicast traffic. Light-trails were proposed in [4,5] as a solution to the challenges facing lightpaths and could be built using commercial off-the-shelf technology. In the years since the introduction of light-trails, significant contributions have been made to enable adoption and advance the architecture [1,6,7,8, 9].

Light-trails enable fast, dynamic creation of an unidirectional optical communication channel. This communication channel, unlike prior lightpaths, allows for channel receive and transmit access to all connected nodes, making them more suitable for IP-centric traffic [6]. Point-to-point communications from an upstream node to a downstream node can be scheduled on the shared light-trail. Similarly, an upstream node can multicast to any number of downstream nodes.

A scheduling protocol is in place to avoid collisions within a light-trail and controls when nodes are able to transmit to downstream nodes. The scheduling is generally assumed to occur over a control channel, which may or may not be separate from the shared optical fiber that is being used for the light-trail.

An example four-node light-trail can be seen in Figure 1. Optical shutters allow for wavelength reuse within the network. Start and end nodes have their optical shutters in the *off* state, while intermediate nodes have their optical shutters in the *on* state. This effectively isolates an optical signal to a specific light-trail and allows for reuse of optical wavelength(s) elsewhere in the network.
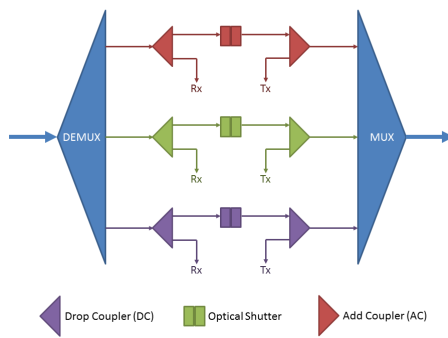
Fig. 2: Example light-trail node structure

Fig. 3: Cycle formed using the light-trail architecture

Nodes can receive from the incoming signal while the signal is simultaneously continuing to downstream nodes, sometimes referred to as a drop and continue function. The node structure can be seen in Figure 2. DWDM fiber-optic networks use reconfigurable add/drop multiplexers (ROADMs) to demux incoming signals into separate wavelengths, then to mux the wavelengths before being output once again. Each wavelength can separately support the light-trail architecture, allowing multiple light-trails to share the same edge in the network. Next-generation ROADMs further increase add/drop and switching flexibility while reducing costs [10]. Early technology supported only a few wavelengths; however, the latest devices may support over 100 channels for over 1 Terabits/s [11].

Light-trail communication is all optical and uses the same wavelength(s) from start to end node. Being all optical avoids any energy inefficiencies and time delays associated with unnecessary Optical-to-Electrical-to-Optical (O/E/O) conversions at intermediate hops. Transmissions within long haul networks, potentially passing through one or more nodes, used to be limited by the optical signal to noise ratio (OSNR). In recent years, several advancements and mitigating techniques have allowed for this limitation to be reduced and in some cases completely removed. One such advancement was the erbium- and ytterbium-doped optical fiber amplifier. These amplifiers compensate for signal losses and allow for signals to travel thousands of kilometers [11].

4 Light-trail, cycle routing, and fault-tolerance

Point-to-point and multi-point traffic requests have a set of nodes $C = \{a_i, ..., a_j\}$ that wish to communicate and need to be protected against network faults. Establishing a primary and backup multicast path fr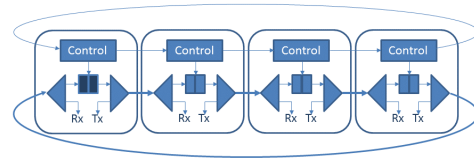om every node to every other node in C can be a waste of resources. Several methods protect the path or links along the route through an independently found tree or cycle. In this work, we utilize the light-trail architecture in the form of a cycle (Fig. 3). The bidirectional cycle will both route the multi-point request and protect it at the same time using fewer resources.

Figure 3 is simply a light-trail where the start and end node is the same node, referred to as the *hub node*. The hub node has its optical shutters in the *off* state, while intermediate nodes have their optical shutters in the *on* state. The resources at each hub node can be utilized to allow all-to-all communication on the cycle using only one light-trail. Traffic from a node to nodes downstream requires a single transmission. Traffic from a node to an upstream node must undergo Optical-to-Electrical-to-Optical (O/E/O) conversion at the hub node and be transmitted on the light-trail a second time.

Alternately, we choose to set up two light-trails, one in each direction. This enables upstream communications without the energy inefficiencies and time delays associated with O/E/O conversions. It also has fault tolerance properties.

Additionally, if traffic in the network is expected to be significantly different than equal traffic between all node pairs, multiple light-trails can be used to support those specific cycles expecting heavier traffic.

Failures within an optical network are to be expected. The generalization of p-cycle protection to allow for path and link protection was proposed by [2]. P-cycle protection of unicast and multicast traffic networks requires preconfiguration, and the offline nature allows for the efficient cycles to be selected [12,13]. [14] examined both path-based and link-based protection schemes in WDM networks. The use of path-pair protection, link-based shared protection, spanning paths, and p-cycles to protect multicast sessions have all been proposed for WDM networks as well [15,16,17,18,19]. The Optimized Collapsed Rings (OCR) single link protection heuristic was developed to address the heterogeneous, part multicast / part unicast, nature of WDM traffic [20].

The multi-point cycle routing algorithm (MCRA) uses bidirectional cycles for fault tolerance and is capable of supporting SONET rings and p-cycles [1]. Al-

though finding the smallest cycle supporting the multi-point communication is NP-Complete, the authors were able to show that their heuristic performed within 1.2 times of the optimal cycle size. ECBRA is a significant improvement of MCRA and outperforms the OCR heuristic [9].

ECBRA heuristic balances optimality and speed, taking $O(|E| |C|^3)$ steps to find a close to optimal cycle. First, a modified breadth first search is performed on each node in set C of required communication nodes. The goal is to find a shortest path in G that also has the best ratio of nodes from set C versus total nodes on the path. The heuristic gives preference to paths with 2-degree C nodes as these 2-degree nodes are required to be a part of the cycle.

To complete the cycle, a path from the sink node returning to the source node must be found. No links may be used twice. If all nodes in C are in the cycle, then the cycle search is complete. Otherwise, a third step is required to add any missing nodes.

If needed, the final step iteratively removes edges from the cycle and inserts paths through missing nodes in C. Because insertion of the node can be cheaper by removing some links from the cycle rather than others, the optimal edge removal from the cycle and path replacement is computed for each missing node insertion.

5 Quorums

In distributed communication and algorithms, coordination, mutual exclusion, and consensus implementations have grouped N nodes into small sets called quorums. This organization of nodes can minimize communications in operations like negotiating access to a global resource.

A quorums set minimally has the property that all quorums in the set must intersect. Specifically for distributed implementations, it is also desirable that each node has equal work and equal responsibility within the quorums set [21].

Not every grouping of nodes into sets (quorums) will result in having these three properties, nor will the quorum sizes be minimal. [21] proved the lower bound on the size of quorums set having these three properties. Cyclic quorums sets with these properties, like other quorums sets in general, are difficult to find and require an exhaustive search [21,22]. Quorums sets are proposed in this paper (and a conference paper [3]) for cycle-based routing to efficiently support arbitrary point-to-point and multi-point communication.

5.1 Defining quorums set

A is a set of $N = |V|$ nodes. A set S_i is a subset of A. When set Q of subsets (Eq. 3) covers all nodes in A (Eq. 4) and all subsets also have non-empty intersections (Eq. 5), then set Q is called a quorums set.

$$A = \{a_1, ..., a_N\} \tag{1}$$

$$S_i = \{a_j, ...\}, \quad j \in 1, 2, ..., N \tag{2}$$

$$Q = \{S_1, ..., S_N\} \tag{3}$$

$$\bigcup_{i=1}^{N} S_i = \{a_1, ..., a_N\} = A \tag{4}$$

$$S_i \cap S_j \neq \emptyset, \quad \forall \ i, j \in 1, 2, ..., N \tag{5}$$

The lower bounds for the maximum individual quorum size (i.e., $|S_i|$) in a minimum set is K, where Equation 6 holds and $(K - 1)$ is a power of a prime, proved through equivalence to finding a finite projective plane [21]. Additionally it is desirable that each quorum S_i in the quorum set be of equal size (Eq. 7), such that there is equal work and it is desirable that each node be contained in the same number of quorums (Eq. 8), such that there is equal responsibility.

$$N \leq K(K - 1) + 1 \tag{6}$$

$$|S_i| = K, \quad \forall \ i \in 1, 2, ..., N \tag{7}$$

$$a_i \text{ is contained in } K \, S_j's, \quad \forall \ i \in 1, 2, ..., N \tag{8}$$

To the best of our knowledge, no efficient algorithm is known to find quorums of minimum size that maintain the properties used by [21]. Cyclic quorums adhere to these properties and are based on cyclic block design and cyclic difference sets; however, searching for optimal cyclic quorum requires an exhaustive search [22]. Cyclic quorums are unique in that once the first quorum (Eq. 9) is defined the remaining quorums can be generated via incrementing the entity ids (modulus to keep entity ids within bounds is not shown in Eq. 10 for conciseness). For simplicity assume $a_1 \in S_1$ without loss of generality (any one-to-one re-mapping of entity ids can result in this assumption).

$$S_1 = \{a_1, ..., a_j\} \tag{9}$$

$$S_i = \{a_{1+(i-1)}, ..., a_{j+(i-1)}\} \tag{10}$$

For our work, we used the $N = 4, ..., 111$ optimal cyclic quorums from a paper by Luk and Wong [22].

5.2 Quorums sets for routing

It is important to establish quorums relationship to the network model $G = (V, E)$. Quorums sets cover N entities, in this case $N = |V|$ optical network nodes. The

number of entities, N, also defines the number of small subsets, i.e., quorums, that will be used in our solution.

In this paper (and a conference paper[3]), the use of quorums for efficient point-to-point, multi-point, and all-to-all traffic requests in optical networks are proposed. This is important because traffic in many optical networks is heterogeneous meaning the routing framework must be able to handle all types.

Point-to-point, multi-point, and all-to-all traffic can be routed through an optical network with N cycles based on cyclic quorums. The following breaks down the handling of different traffic request types on the quorum supporting rings. Sections 5.2.1 and 5.2.2 expand upon a conference paper presented by Lastine and Somani [3].

5.2.1 Point-to-point traffic

Using cyclic quorums sets (see additional definitions in Sect. 5.1) all possible node pairs occurred in at least one cycle. This ensures for any dynamic point-to-point request there is at least one cycle that can support it.

For many networks, node pairs will appear in multiple cycles. When new requests arrive, they can select a route on the least loaded light-trail available. Taking this one step further, if cycles with cyclical quorums are intentionally made larger than the minimum required, then it can be ensured that all networks have at least some opportunity for load balancing point-to-point connections.

5.2.2 Multi-point traffic

Optimally, if all multi-point participants belong to the same cycle, then one cycle can be used. Realistically though, dynamic requests often will not be of this nature. Requests will span multiple quorums and/or be larger than a single quorum cycle. Hence in the worst case, no more than K cycles are required to efficiently route and protect traffic for each multicast traffic request (more discussion on this bound in the broadcast traffic discussion, Sect. 5.2.3).

Additionally, in [7] the problem of multicast in the general case for light-trails and for light-trail WDM networks is studied. A way of transforming the general case to a minimum Steiner tree problem is given. A polynomial time algorithm is given for the special case of light-trail WDM ring networks. Work from this paper can be used to plan multicast requests that become known after the cycles based on quorums have been selected.

If we know of multicast requests the network will be serving a priori, we can number nodes to reduce the number of light-trails required to support the multicast.

The simplest case is when there is only one multicast that needs to be supported and the nodes involved in the multicast is equal or smaller than the number of nodes in a quorum. In this case, we can renumber the nodes such that they are all in one cyclic quorum.

If there are two known multicasts, each not larger than a quorum, mapping may still be fairly simple. If they both use some of the same nodes, there may be a pair of quorums that also has at least the same number of overlaps. So two multicast could be accommodated by numbering nodes common to multicasts with numbers common to quorums, and then naming the rest of the nodes in the multicasts after nodes from one or the other quorum.

With each additional known multicast to support, the mapping to quorums can become more of a challenge. We leave this as an open question on how to handle the general case of numbering nodes to handle multiple multicast requests known a priori.

5.2.3 Broadcast traffic

Broadcast traffic is simply the worst case of a multicast traffic request. The upper bound of requiring no more than K cycles to route and protect up to broadcast traffic can described as follows. In Section 5.1, any entity will occur in at most K quorums. In those K quorums, all other entities must be present in order to form necessary point-to-point pairs as described before (Sect. 5.2.1). Hence, any optical node can communicate on all K quorum cycles that it is a member and reach all other optical nodes, thereby efficiently serving any dynamic broadcast request.

6 Efficiency analysis

The results presented in this section were adapted from results in a conference paper presented be Lastine and Somani [3].

NSFNET is a 14 node 22 link network (Fig. 5a). At a size of 14 nodes, it has cyclic quorums of size five. An improved version of ECBRA [23] was used to find cycles that included the specified set of nodes in each quorum.

Routing efficiency Quorums for NSFNET are given in Table 1. The first column numbers these quorums from 1 to 14. The second column lists the actual quorums. The third column lists cycles in the network that contain the quorums. The fourth column provides the length (size) of each cycle as a convenience to the reader.

Table 1: Cyclic quorums and cycles found in NSFNET

#	Quorum	Cycle	Size
1	1 2 3 4 8	1 4 12 13 8 3 2	7
2	2 3 4 5 9	2 1 4 5 6 9 10 7 3	9
3	3 4 5 6 10	3 7 10 9 6 14 4 5	8
4	4 5 6 7 11	4 12 11 10 7 3 5 6 14	9
5	5 6 7 8 12	5 6 9 10 7 8 13 12 4	9
6	6 7 8 9 13	6 9 10 13 8 7 3 5	8
7	7 8 9 10 14	7 8 3 2 14 6 9 10	8
8	8 9 10 11 1	8 13 12 11 10 9 6 14 1 2 3	11
9	9 10 11 12 2	9 6 14 2 1 4 12 11 10	9
10	10 11 12 13 3	10 11 12 13 8 3 7	7
11	11 12 13 14 4	11 10 13 8 3 2 14 4 12	9
12	12 13 14 1 5	12 13 8 3 5 4 14 1 4	9
13	13 14 1 2 6	13 10 9 6 14 2 1 4 12	9
14	14 1 2 3 7	14 6 9 10 7 3 2 1	8
		Average:	**8.57**
		Total:	**120**

Cycle lengths fall in the range 7-11 links. Their average length being 8.57 links and combined length is 120 links.

Supporting all-to-all traffic using traditional point-to-point connections would be more expensive. There are 91 possible node pairs. With only 22 of the pairs connected via a direct link, it would take over 200 links to support all point-to-point communication paths. The quorums set cycle routing required only 120 links.

Resource efficiency Less hardware is required by the quorum solution as well. The minimum number of transmitters and receivers at a node depends on the number of times it appears in the set of quorums. From Section 5.1, this lower bound is K for optimal cyclic quorums. In the example of the 14 node NSFNET, where $K = 5$, a node minimally needs only 5 transmitters and 5 receivers. This is a significant saving over establishing lightpaths between all pairs of nodes which would require 13 transmitters and 13 receivers at each node.

7 Cycle length

The results presented in this section were adapted from results in a conference paper presented be Lastine and Somani [3].

7.1 Random graph solution cost analysis

Nodes are randomly placed inside a grid and links are randomly inserted with probability given by the Waxman model. In general, the probability of a link existing depends on the distance between a node pair and the desired density of the network. For context, graphs with a mean of approximately 2 edges per node are sparse in optical networking. Whereas a mean of $log_2 n$ edges per node would be dense in this context.

Graphs with various parameters had to be considered, in order to empirically approximate performance of our quorums set cycle routing on a general network. Parameters considered were number of nodes (20, 35, 50, and 65), edge density (sparse, medium, and dense), and edge length (short and long).

Simulation results show that regardless of edge density or length, networks with more nodes intuitively required larger cycles in order to connect all nodes to each other. Denser graphs allow for more route options for the cycle, which allows for smaller cycles to be formed including fewer non-quorum nodes in the cycle. Lastly, graphs with higher probabilities for long edges on average had smaller cycles than graphs of the same size and density, but shorter edges. The intuition for this comes from the need to connect all nodes and longer edges would allow "skipping" a few middle nodes when connecting to a node across the network.

A rough measurement of routing efficiency is produced by computing the ratio of average total nodes in a cycle to the required quorum nodes. The quorums for graphs of size (20, 35, 50, and 65) nodes, respectively, contain (6, 7, 8, and 9) nodes. As expected, simulation results show that the denser graphs had more efficient routes. Also graphs with longer edges tended to have more efficient routes. Interestingly though, increases in network size did not always result in decreases in efficiency as one might expect. This could be partly due to the nature of which nodes are included in which quorum, which motivates the next set of results on whether altering the node number scheme in a graph could impact solution quality.

7.2 Impact of node numbering analysis

Choosing of a node id for a particular node is likely a result of convenience rather than science. Examining several random node numberings for graphs found in literature (Fig. 5), the impact of node numbering on routing efficiency was investigated.

Total links used for routing (routing efficiency) did vary depending on the node numbering. All networks experienced greater than 4% difference between their observed minimum and maximum total links used in their randomly numbered graphs. That is like having an extra cycle routed in each optical network, but not needing it.

The biggest result from this investigation was that all random node numberings for all networks had worse
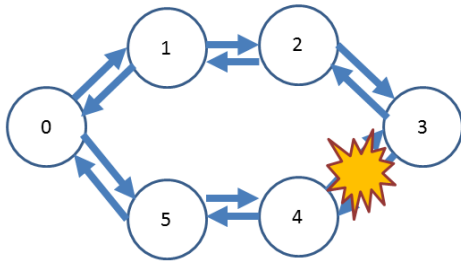
Fig. 4: Example route fault tolerance using light-trails

routing efficiency than the original node numberings in Figure 5. There is some intuitiveness to this result. The cyclic quorums in [22] tend to include several nodes with closely spaced id numbers and include fewer id numbers spaced further away. When the original node numberings of the networks had nodes close to one another with closely spaced id numbers, this worked to the advantage of forming smaller cycles based on quorums.

8 Fault tolerance analysis

Much of the discussion so far has addressed non-fault scenarios, specifically showing efficiencies in performing unicast, multicast, and broadcast operations using cycles based on quorums sets. Optical networks are highly depended upon. The fault tolerance aspect of this route design is important and a paired light-trail implementation is used to address this. When a link fails, both cycles break. In the Figure 4 example, hub node 3 would no longer have a downstream edge to node 4; however, upstream communication can be used to still reach node 4.

8.1 Fault model

The fault model assumed for our work is the single edge failure. While a simple model, it does cover most real single fault scenarios.

The most direct fault to consider is the optical link fault. This occurs when a link is broken, like planned maintenance or the accidental severing during land excavation. Modeling link faults as a single edge failure is straightforward.

Each modeled node needs a pair of transmitters and receivers for each occurrence in a cycle. These pairs of devices can fail too. Short of a natural disaster, pairs will likely fail independently of one another. When a transmitter/receiver pair fails within a modeled node, the affect on the global network is similar to that link failing. Modeling as a single edge failure, while not an exact fault mapping, is an appropriate abstraction.

We do not examine multiple simultaneous faults in our simulation models.

8.2 Paired cycle fault simulation

Maintaining the ability to serve all dynamic point-to-point traffic requests despite fault is important. We examined the fault tolerance of the NSFNET, ARPANET, American backbone, Chinese backbone networks (Fig. 5). For each of these four networks, we created 1000 random numbering schemes for the nodes and found cycles to support cyclical quorums given in [22].

To model the fault, we simulate the failure of each edge, $e \in E$, in the network model, $G = (V, E)$. We then examine the network's ability to serve all potential point-to-point requests by counting pairs of nodes that would be able to communicate and conversely those pairs that are unable to communicate. In the non-fault case, all nodes can unidirectionally communicate with all other nodes for a total of $|V| * (|V| - 1)$ pairs.

The results in Table 2 show acceptable fault tolerance performance. The mean number of missing communication pairs is less than 3 (95% CI). This means that any edge fault in the network will typically be recoverable and few point-to-point connections will experience connection loss when using quorums set-based cycle routing.

To calculate fault coverage in this scenario, we calculated the mean connected pairs divided by the total pairs.

$$1 - \frac{Mean\ Missing\ Pairs}{Total\ Pairs}$$

The results of our simulation showed that greater than 99% fault coverage can typically be expected.

8.3 Improving fault-tolerance

Many of the simulated networks typically performed well, but ideally we would like to see the fewer missing pairs per fault case. A reduction in the highest missing pairs observed is also important.

8.3.1 Additional cycle fault protection

Recall from Figure 3 in Section 4 that hub nodes in cycles have their optical shutters in the *off* state. When a fault occurs like in Figure 4, it is possible that the hub node is 0 and the optical shutters could prevent a necessary communication path between 3 and 4.

We experimented with adding an additional pair of cycles to form quad cycles. The pair has its hub node
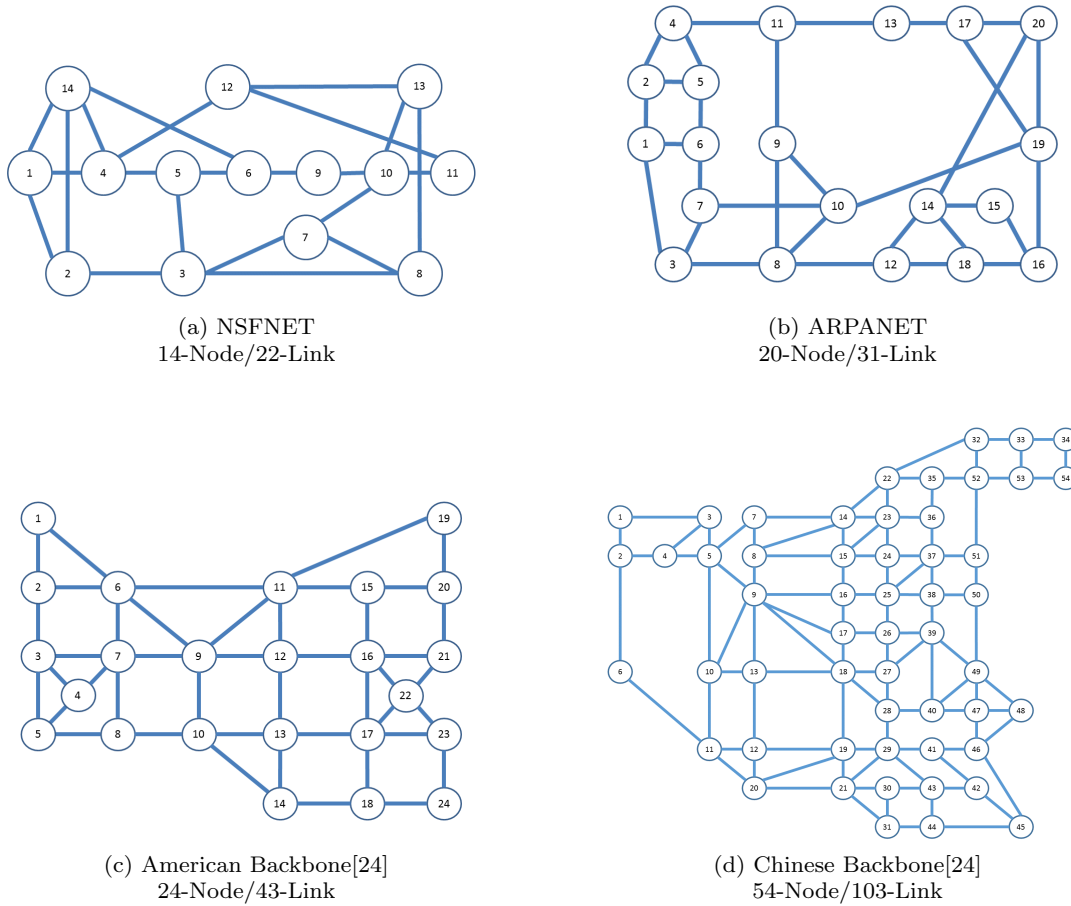
(a) NSFNET
14-Node/22-Link

(b) ARPANET
20-Node/31-Link

(c) American Backbone[24]
24-Node/43-Link

(d) Chinese Backbone[24]
54-Node/103-Link

Fig. 5: Networks used for simulations

Table 2: Paired quorum cycle fault simulation results

Network	Nodes	Total Pairs	Missing Pairs			Fault Coverage (%)
			High	Mean (95% CI)	Low	
NSFNET	14	182	12	0.93644 ± 0.02070	0	99.485
ARPANET	20	380	16	0.76051 ± 0.01715	0	99.800
American	24	552	26	2.05273 ± 0.02812	0	99.628
Chinese	54	2862	56	2.77809 ± 0.02400	0	99.903

directly across from the original pair's hub node, i.e., at position $\left\lfloor \frac{CycleLength}{2} \right\rfloor$. In the Figure 6 example, hub nodes 0 (inner blue light-trail) and 3 (outer red light-trail) are across from one another. Node 3 still does not have a downstream edge to node 4 on either the inner or outer cycle, but there does exist an upstream path on the outer cycle to node 4.

The results in Table 3 show that most of the networks had the mean number of missing pairs improved by an order of magnitude (95% CI). Any edge fault in the quad cycle network will typically be recoverable and fewer point-to-point connections will experience
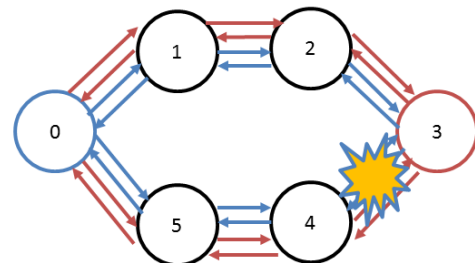

Fig. 6: Quad light-trails to provide additional cycle fault protection

Table 3: Quad quorum cycle fault simulation results

Network	Nodes	Total Pairs	Missing Pairs			Fault Coverage (%)
			High	Mean (95% CI)	Low	
NSFNET	14	182	6	0.09130 ± 0.00574	0	99.950
ARPANET	20	380	6	0.08710 ± 0.00496	0	99.977
American	24	552	12	0.28661 ± 0.00833	0	99.948
Chinese	54	2862	20	0.52939 ± 0.00768	0	99.982

connection loss than with only the paired cycles. This translates into higher fault coverage values as well. Similarly the edge faults that generated the highest missing pairs observed in the simulated networks also decreased by 50% or more.

Adding the additional pairs of cycles and the still having missing communication pairs may have been a bit surprising. While the additional cycles significantly helped in fault reductions, the underlying missing pairs come from the intermediate nodes common to both cycles. Depending on the failing edge, these nodes may not have a downstream or upstream path on any of the cycles to nodes on the opposite half of the cycle.

8.3.2 Modifying the cycle routing algorithm

Rather than addressing the missing point-to-point communication pairs with additional cycles, we could try to address the underlying cause of the missing pairs by changing the cycles themselves. Here, we briefly outline additional algorithm steps that could be added to ECBRA.

In Section 5.2.1, every point-to-point request was supported by at least one quorum cycle. Being opportunistic, even if one cycle is experiencing a fault, there may be another quorum cycle that also is supporting that point-to-point request. Hence, quantifying missing pairs and attempting to compensate should only occur once all cycles for a graph's quorums set is found.

To enumerate the missing pairs, simulate the failure of each edge and examine the network's ability to serve all potential point-to-point requests by counting pairs of nodes that would be able to communicate and conversely those pairs that are unable to communicate. Given results in Section 8.2 the mean number of missing pairs is expected to be less than 3 per fault edge (95% CI). Form a *Missing-Pair* tuple $t = <\{s, d\}, e>$, where s is the source and d is the destination of the missing communications pair and e is the edge whose failure was being simulated.

Every pair missing can be isolated back to a cycle responsible for the pair (Sect. 5.2.1). For each tuple t, remove the faulty edge e in the responsible cycle for pair $\{s, d\}$. To complete the cycle, find the shortest path between the two disconnected nodes, such that cycle edges are not reused and avoiding the use of e. Once all Missing-Pair tuples have been processed, repeat the enumeration and removal of missing pairs once again. Repetition is required to confirm that in the process of addressing one Missing-Pair tuple that another new missing pair was not added.

There is a possibility that a cycle may have to use edge e. An example would be a node with only two edges and the node being a member of the quorum for that cycle. This is a known limitation and is a challenge in general for establishing appropriate fault tolerance for communications to/from this node, as well as, through this node serving others. This limitation is discussed further in [9].

9 Conclusion

Many optical networks face heterogeneous communication requests requiring topologies to be efficient and fault tolerant. We show efficiency and distributed control can also be accomplished in optical network routing by applying the quorums set theory.

Specifically we show cyclic quorums set routing have all pairs of nodes occurring in one or more quorums. This allows for efficient, arbitrary unicast and multicast communications requests that do not need to be known or even modeled a priori. Even efficient broadcasts to all network nodes are possible by a node broadcasting to all quorum cycles to which it belongs.

We analyzed the fault tolerance of our quorums set routing approach. In the presence of network link faults, greater than 99% average coverage enables the continued operation of nearly all arbitrary unicast and multicast requests in the network. Adding additional cycles showed an order of magnitude reduction in mean missing point-to-point pairs for most networks (95% CI). Finally, to address faults at their source, an augmentation to the ECBRA cycle finding algorithm was proposed.

10 Future work

In this paper, ECBRA was used to route each of the quorums-based cycles. It was shown that the quorums set approach required far fewer links to accomplish the routing of all-to-all traffic, when compared to using point-to-point connections.

As an unintended benefit, some quorums sets result in node pairs occurring in more than one quorums-based cycle. It was these occurrences of node pairs multiple times that motivates us to examine whether redundant node pairs could be generated intentionally. Predictably redundant pairs can improve the dependability of the optical network, by guaranteeing even if a cycle failed, all node pairs for point-to-point and multi-point communications could still be present in the network.

References

1. D. Lastine, S. Sankaran, and A. K. Somani, "A fault-tolerant multipoint cycle routing algorithm (mcra)," in *Broadband Communications, Networks, and Systems.* Springer, 2012, pp. 341–360.

2. W. D. Grover and G. Shen, "Extending the p-cycle concept to path-segment protection," in *Communications, 2003. ICC'03. IEEE International Conference on*, vol. 2. IEEE, 2003, pp. 1314–1319.

3. A. K. Somani and D. Lastine, "Optical paths supporting quorums for efficient communication," in *Optical Communications and Networks (ICOCN), 2014 13th International Conference on.* IEEE, 2014.

4. A. Gumaste and I. Chlamtac, "Light-trails: a novel conceptual framework for conducting optical communications," in *High Performance Switching and Routing, 2003, HPSR. Workshop on.* IEEE, 2003, pp. 251–256.

5. I. Chlamtac and A. Gumaste, "Light-trails: A solution to ip centric communication in the optical domain," in *Quality of Service in Multiservice IP Networks.* Springer, 2003, pp. 634–644.

6. J. Fang, W. He, and A. K. Somani, "Optimal light trail design in wdm optical networks," in *Communications, 2004 IEEE International Conference on*, vol. 3. IEEE, 2004, pp. 1699–1703.

7. Y. Li, J. Wang, A. Gumaste, Y. Xu, and Y. Xu, "Multicast routing in light-trail wdm networks," in *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE.* IEEE, 2008, pp. 1–5.

8. W. Zhang, F. Kandah, C. Wang, and H. Li, "Dynamic light trail routing in wdm optical networks," *Photonic Network Communications*, vol. 21, no. 1, pp. 78–89, 2011.

9. A. Somani, D. Lastine, and S. Sankaran, "Finding complex cycles through a set of nodes," in *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, Dec 2011, pp. 1–5.

10. P. N. Ji and Y. Aono, "Colorless and directionless multidegree reconfigurable optical add/drop multiplexers," in *Wireless and Optical Communications Conference (WOCC), 2010 19th Annual.* IEEE, 2010, pp. 1–5.

11. G. P. Agrawal, *Nonlinear fiber optics.* Academic press, 2007.

12. F. Zhang and W.-D. Zhong, "Performance evaluation of p-cycle based protection methods for provisioning of dynamic multicast sessions in mesh wdm networks," *Photonic Network Communications*, vol. 16, no. 2, pp. 127–138, 2008.

13. F. Zhang, W.-D. Zhong, and Y. Jin, "Optimizations of p-cycle-based protection of optical multicast sessions," *Lightwave Technology, Journal of*, vol. 26, no. 19, pp. 3298–3306, 2008.

14. S. Ramamurthy, L. Sahasrabuddhe, and B. Mukherjee, "Survivable wdm mesh networks," *Journal of Lightwave Technology*, vol. 21, no. 4, p. 870, 2003.

15. N. K. Singhal, L. H. Sahasrabuddhe, and B. Mukherjee, "Provisioning of survivable multicast sessions against single link failures in optical wdm mesh networks," *Journal of lightwave technology*, vol. 21, no. 11, p. 2587, 2003.

16. Y. Qing and G. Ning, "Protecting dynamic multicast sessions in optical wdm mesh networks," in *Information, Communications and Signal Processing, 2005 Fifth International Conference on.* IEEE, 2005, pp. 1187–1190.

17. H. Luo, H. Yu, L. Li, and S. Wang, "On protecting dynamic multicast sessions in survivable mesh wdm networks," in *Communications, 2006. ICC'06. IEEE International Conference on*, vol. 2. IEEE, 2006, pp. 835–840.

18. F. Zhang and W.-D. Zhong, "Applying p-cycles in dynamic provisioning of survivable multicast sessions in optical wdm networks," in *Optical Fiber Communication Conference.* Optical Society of America, 2007, p. JWA74.

19. T. Feng, L. Ruan, and W. Zhang, "Intelligent p-cycle protection for multicast sessions in wdm networks," in *Communications, 2008. ICC'08. IEEE International Conference on.* IEEE, 2008, pp. 5165–5169.

20. A. Khalil, A. Hadjiantonis, G. Ellinas, and M. Ali, "Preplanned multicast protection approaches in wdm mesh networks," in *Optical Communication, 2005. ECOC 2005. 31st European Conference on.* IET, 2005, pp. 25–26.

21. M. Maekawa, "An algorithm for mutual exclusion in decentralized systems," *ACM Transactions on Computer Systems (TOCS)*, vol. 3, no. 2, pp. 145–159, 1985.

22. W.-S. Luk and T.-T. Wong, "Two new quorum based algorithms for distributed mutual exclusion," in *Distributed Computing Systems, 1997., Proceedings of the 17th International Conference on.* IEEE, 1997, pp. 100–106.

23. D. Lastine, "Efficient communication using multiple cycles and multiple channels," Ph.D. dissertation, Iowa State University, 2014.

24. L. Tang, W. Huang, M. Razo, A. Sivasankaran, M. Tacca, and A. Fumagalli, "Multicast tree computation in networks with multicast incapable nodes," in *High Performance Switching and Routing (HPSR), 2011 IEEE 12th International Conference on.* IEEE, 2011, pp. 95–100.
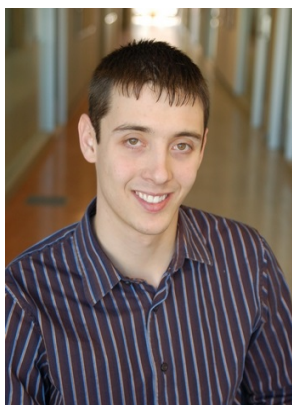
Cory J. Kleinheksel received the B.S. degree in Computer Engineering from Iowa State University in 2008. He is a Ph.D. candidate at Iowa State University pursuing a degree in Computer Engineering. He has received fellowships from the NSF Graduate Research Fellowship Program, IBM Ph.D. Fellowship Program, and Symbi GK-12 Fellowship Program at Iowa State University. His research interests are in the area of big

data communication and processing with emphasis on parallel and distributed systems.

Arun K. Somani is serving as Anson Marston Distinguished Professor and Philip and Virginia Sproul Professor of Electrical and Computer Engineering at Iowa State University. He also served as Scientific Officer for Govt. of India, New Delhi from 1974 to 1982 and as a faculty member at the University of Washington, Seattle, WA from 1985 to 1997. He earned his MSEE and PhD degrees in electrical engineering from the McGill University, Montreal, Canada, in 1983 and 1985, respectively.

Professor Somanis research interests are in the area of computer system design and architecture, fault tolerant computing, computer interconnection networks, WDM-based optical networking, and reconfigurable and parallel computer systems. He has also served as IEEE distinguished visitor, IEEE distinguished tutorial speaker, and IEEE Communication Society distinguished visitor and has delivered several key note speeches, tutorials and distinguished and invited talks all over the world. He was elected a Fellow of IEEE for his contributions to theory and applications of computer networks in 1999 and as a Distinguished Engineer of ACM in 2006. He also has been elected a fellow of AAAS in 2012.