Resilient Routing Layers for Network Disaster Planning

Audun Fosselie Hansen[1 2], Amund Kvalbein[1], Tarik Čičić[1], and Stein Gjessing[1]

[1] Networks and Distributed Systems Group, Simula Research Laboratory, Oslo, Norway
{audunh, amundk, tarikc, steing}@simula.no
[2] Telenor R&D, Oslo, Norway

Abstract. Most research on network recovery has been centered around two common assumptions regarding failure characteristics: Failures do not occur simultaneously and failures do mostly strike links. Even this may be the characteristics of everyday failures, we argue that disasters like earthquakes, power outages and terrorist attacks impose other failure characteristics. In this paper we demonstrate how our method, called 'Resilient Routing Layers', can be used as a tool for recovery from failures adhering to such disaster characteristics.

Keywords: Network resilience, Recovery in various networks

1 Introduction

Resilience against physical attacks was one of the primary design goals of the Internet from the outset [1]. The distributed nature of control information and routing algorithms allows the Internet to recover from link or node failures by calculating new valid paths in the remaining network. However, special challenges arise for the Internet when faced with disastrous events like earthquakes, floods, hurricanes, large-scale accidents, power outages or terrorist attacks.

From a networking point of view, disasters like the ones just mentioned have two key characteristics. First, a *large number of nodes* can be taken out at the same time. This makes many traditional protection mechanisms unsuitable, since they are often designed to protect against single failures, and most methods focus on link failures only. Second, the failing nodes are geographically near each other, giving poor connectivity in the disaster area. After the power outage on the US east coast on September 28 2003, as much as 1% of the Internet was disconnected for several minutes, and several thousand networks were still unconnected after eight hours [2].

At the same time, the need to communicate often increases dramatically in the disaster area. After the 9-11 terrorist attack in New York, there was a sharp increase in the traffic load in the mobile phone networks, which suffered from damaged infrastructure and experienced regional congestion [3]. For the Internet, the overall traffic load did not increase, although the pattern of use

diverged from the normal. However, with the convergence between the data and the telecommunications infrastructures, we believe that the increased demand experienced in the telephone network, will be increasingly relevant also for the Internet in a disaster situation.

Given the reduced availability and the increased need for communications near and in the disaster area, we believe that a disaster recovery scheme should aim at treating the affected area isolated from the rest of the network. Since the affected area of the network must be considered unreliable, other parts of the network should not be dependent on this area for routing or traffic forwarding. At the same time, the remaining resources in the affected area and between the affected area and the rest of the network, is likely to be scarce, and put under heavy pressure. These communication resources should therefore be available for intra-area traffic and traffic originating or terminating in the area, not transit traffic.

In this paper, we propose using Resilient Routing Layers (RRL) [4] for protecting networks against large-scale disasters. RRL complies well with the goals stated above. Surrounding networks do not rely on the affected area for routing updates, since all forwarding is done using pre-calculated routing information. Traffic that did not pass through the affected area before the failures, is not affected. Our scheme removes transit traffic from the disaster are, while still allowing traffic originating or terminating in the area.

2 Background and Related Work

Planning for communication abnormalities in the event of disasters should be considered in all levels of the communication hierarchy. Research reported in the literature is scarce and what can be found has mainly focused on the application and transport layers.

Because communication during a disaster is mainly a government responsibility, some government funded research has been performed on how network applications and supporting infrastructure should be designed in order to be available in a disaster situation [5][6]. A catastrophe will cause stress in the Internet, and access and admission control methods have been developed that can be beneficial after a disaster [7]. At the transport level special socket communication that is reliable also in the case of catastrophic failures is reported in [8].

The lack of research on network disaster planning may be caused by the fact that the IP protocol and the Internet itself is designed for failures [1]. Many of the common recovery and rerouting techniques developed for more regular errors and failure situation may, however, be used as a starting point when more specific research into network disaster recovery is performed.

One of the most studied topics within the field of recovery research has been efficient algorithms for finding alternative paths between sources and destinations in a network [9] [10]. Network recovery management is, however, difficult if the only view of the network is a large set of unstructured backup paths. The

literature provides some alternatives for more structured recovery. Such schemes are based on building a set of subtopologies of the network, serving as a more intuitive abstraction of the alternative paths. These schemes can serve as input to restoration and protection, both global and local. Examples of such approaches are Redundant Trees [11] [12] and Protection Cycles [13] [14].

On the IP level, recovery relies on the IP routing protocols to complete a global rerouting process. In case of normal failure situations this rerouting can take several seconds to complete. Although fast recovery is always welcome, we may tolerate some seconds of disrupted communications during a disaster. However, experiments have demonstrated that IP routing protocols cause very unstable routing for long periods when recovering from failures [15] [16]. During serious outages caused by disasters we expect this instability to be even worse. Another potential obstacle derived from IP rerouting may occur in situations where most network components in an area have failed, but some components have survived and are offering viable routes through that area. This may cause transit traffic to congest the limited amount of resources available in the struck area. This is resources that should be exclusively available for traffic originated and terminated in that area.

The authors have previously described Resilient Routing Layers (RRL) as a method for recovery from node and link failures [4]. RRL can be put in the category of schemes that is based on building subtopologies of the network. RRL differs from other schemes in that it is not bound to a particular kind of subtopologies like trees or cycles. As opposed to most other schemes that focus on link failures and often only single failures, RRL is designed to also isolate many nodes simultaneously [17]. Most schemes presented above have their applicability in connection-oriented networks, while RRL can be applied to both connection-oriented and connectionless networks.

In this paper we will demonstrate how RRL can serve as a tool for pre-planning fast isolation of multiple nodes or even whole areas. RRL only isolates the nodes or areas from carrying transit traffic. Traffic that is originated in or destined for these nodes or areas get exclusively access to all available local resources.

3 Resilient Routing Layers (RRL)

RRL organizes the network topology in subtopologies that we call *routing layers*. In each layer there are some nodes or areas that do not carry transit traffic, and we call these nodes or areas the *safe nodes* of this layer.

Layers are constructed so that all nodes are present in each layer, and there exist a path between all node pairs in each layer. Each node should be safe in at least one layer to guarantee single node fault tolerance. There are numerous ways to construct the layers so that different protection properties are optimized [4] [17] [18].

Figure 1 shows an example of how a network can be covered by two layers. The 8 nodes of figure 1 can also represent 8 areas or subnetworks, as is more

thoroughly described later in the article, and depicted in the topology in figure 3. Figure 1a) shows the original full topology. In layer 1 (figure 1b)), nodes 1, 2, 3 and 5 (dashed) are safe. The links connecting these nodes to the rest of the network are also dashed, indicating that the routing on these links must adhere to some specific rules. The rest of the nodes (4, 6, 7 and 8), can be made safe in layer 2 as shown in figure 1c). When a node fails, the traffic affected by the failure will be routed according to the safe layer of the failed node. Traffic not originally routed through the failed node will still be routed according to the full topology.

The following guidelines will arrange that all pairs of nodes can communicate with each other in all layers, and also that safe nodes will not carry any transit traffic, only traffic originating and terminating in the safe nodes:
1) non-dashed links can carry all kinds of traffic originated and terminated anywhere.
2) dashed links can only be used as the first hop or last hop of the communication, meaning that traffic originated in a safe node can use a dashed link as first hop, and that traffic terminated in a safe node can use a dashed link as last hop towards the safe node.

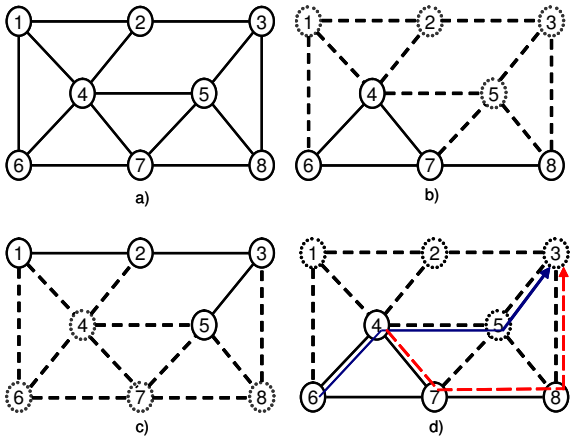

Fig. 1. a): An example network with 8 nodes and 14 links. b): layer 1 (L_1) generated based on a). c): layer 2 (L_2) generated based on a). d): An example on how traffic is routed in a failure situation

To take advantage of the resilient routing layers a packet network implementation must fulfill certain requirements. Each packet must be marked so that each node on the path will know what layer the packet should be routed on. If n is the maximum number of layers, $log_2(n)$ bits in the packet header should identify the currently valid layer. The node that moves a packet to another layer, marks

the packet header with the global identification of the new layer. In the case of failures, only traffic affected by the failed node should be moved to another layer. All packets not affected by the fault will still be routed based on the full topology.

Fig. 1d gives an example of how traffic is switched between layers when node 5 fails. The dotted links may not be used for transit traffic in layer 1, i.e. the safe layer of node 5. Before node 5 fails, all traffic may use the full topology, e.g. traffic from node 6 to node 3 will follow the path 6-4-5-3. When node 5 fails, traffic transiting node 5 must be routed according to layer 1 (using dotted links for first hop and last hop only), while all other traffic (traffic not originally transiting node 5) can still be routed according to the full topology. In the case of local rerouting, traffic is routed from node 6 to 4 according to the full topology. Node 4 detects the failure, and switches traffic to layer 1. The path for traffic between node 6 and node 3 will then be 6-4-7-8-3. If node 6 is notified about the failure (global rerouting) of node 5, the transition to layer 1 could be done by node 6. The path would then be 6-7-8-3.

3.1 RRL Evaluations

Scalability: As demonstrated in [4], RRL seems to scale well, and requires few layers even for very large networks. Table 1 presents a collection of results regarding the number of layers for different real network topologies and some very large synthetic topologies. The real world topologies have been collected from Oliver Heckmann [19] and Rocketfuel [20]. In addition we have generated 100 synthetic topologies using the brite topology generator with 1024 nodes and an average node degree of 4 [21].

Table 1. Percentage of topologies requiring from two till six layers

Topology type	**%2**	**%3**	**%4**	**%5**	**%6**
Rocketfuel	17	50	33	0	0
Heckmann	33	17	33	0	17
Brite	0	0	58	42	0

Backup path lengths: Since RRL restricts the number of links that can carry transit traffic in the case of a failure, backup routes will be longer compared to having all links available. Fig. 2 shows that the increased backup path length is relatively modest, i.e., about 0.8 hops in average compared to the most optimal backup path. These figures stem from [4] where we have used 100 brite topologies with 32 nodes and an average node degree of 4.

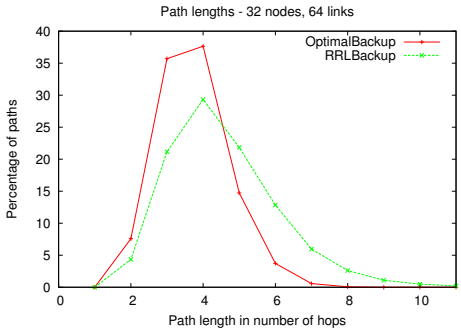

Fig. 2. Distributions of path lengths when introducing one node failure in the original path. The figure shows optimal backup paths and RRL backup paths for 100 brite topologies with 32 nodes and 64 links. Optimal backup path refers to finding a new path in a topology where only the failed node is removed

Resisting multiple node failures: As reviewed in Sec. 2 most contributions on resilience have focused on failures of one component only and particularly link failures. The failure characteristics during disastrous events differ from what have been focus for previous research on recovery. Such events have a tendency to strike numerous nodes instead of a single fiber conduit.

With respect to this context, RRL offers a flexible way of isolating nodes, including many nodes at the same time. RRL resists simultaneous failures of all nodes that are safe within the same layer. What nodes that should be safe in the same layers can be decided based on knowledge of what nodes have a greater risk of failing simultaneously. Some considerations regarding RRL's flexibility and ability to resist multiple failures can be found in [17].

3.2 Isolating whole areas

We have previously argued that disastrous events often strike one or more particular areas. These areas can be states, cities, campuses or buildings. An area that has been struck by a disaster will experience a decreased amount of available communication resources. On the contrary the need for communication to and from that area may increase. Therefore, communications that are transiting this area should find other routes, thus not using precious resources if not absolutely necessary. Also, if the area does not manage to handle any traffic, the traffic only transiting this area must be rerouted.

To accomplish the isolation of whole areas and not only single nodes, we suggest to associate localized nodes as one area. Then we can use RRL on the area-level, meaning that a node represents an area with respect to the RRL overview in Sec. 3.

Fig. 3 shows the hypothetical pan-European cost239 network [22], where the areas correspond to a node in the cost239 network. Each area consists of several

intra-area nodes which offer different connections to intra-area nodes in other areas. To arrange that each area is safe at least once, we need two layers.

Upon a disastrous event striking for instance area 6, some intra-area nodes may survive and hence still offer connectivity to other areas, however the capacity has probably decreased. The idea behind RRL is that traffic originally originated and terminated in area 6 will still be routed from or to that area, while traffic originally only transiting area 6 will now be routed around. This is accomplished by letting the affected traffic be routed according to the safe layer of area 6 (Fig. 4). Traffic not originally passing area 6 will still be routed according to the full topology.

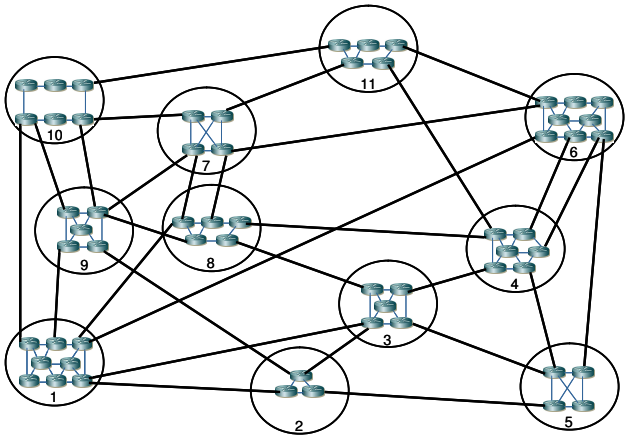

Fig. 3. The original fully connected Cost239 network

Practical support for areas: The most common Internet routing schemes support grouping of nodes and subnetworks into areas. Inter-domain routing, e.g., BGP, is based on modeling Autonomous Systems (ASs) that represent localized nodes and subnetworks [23]. Also Intra-domain routing, e.g., OSPF, support such grouping of localized nodes into areas [24].

Hierarchical RRL: As demonstrated above, RRL can be used to isolate many nodes at the same time, resisting multiple simultaneous failures. In addition, it can be used to isolate whole areas. These two ways of application could be combined to form a hierarchical solution. RRL could be used for recovery of multiple node failures within each single area, and at the same time for preventing transit traffic through the failed area.

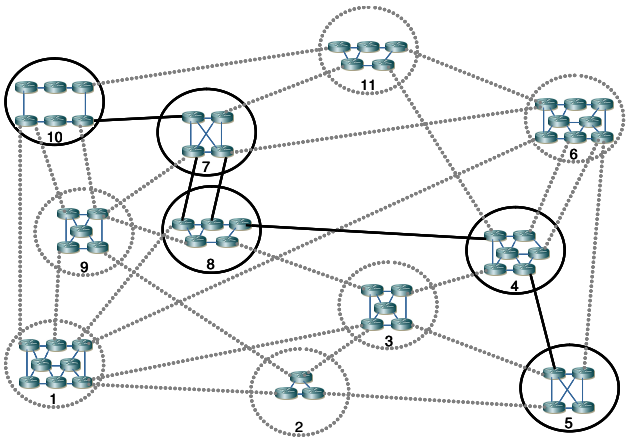

Fig. 4. The layer where area 6 is isolated and safe

4 Conclusion

In this paper we have argued that network recovery schemes, as specified today, are not designed to handle the types of failures that are induced by typical disastrous events. Such events are likely to strike many nodes localized in the same area. We have demonstrated how 'Resilient Routing Layers' (RRL) better adhere to the failure characteristics imposed by disasters. RRL can be used to isolate nodes and areas in such a way that no transit traffic will be routed through a struck node or area, while traffic originated or terminated in such nodes or areas will still be transmitted if surviving connections allows it.

References

1. Clark, D.D.: The design philosophy of the DARPA internet protocols. SIGCOMM, Computer Communications Review **18** (1988) 106–114
2. McGrath, D.: Measuring the 4:11 effect: The power failure and the internet. IEEE Security and Privacy **1** (2003) 16–18
3. Partridge, C., Barford, P.: The Internet Under Crisis Conditions: Learning from September 11. The National Academic press, Washington, D.C. (2003)
4. Hansen, A.F., Cicic, T., Gjessing, S., Lysne, O.: Resilient routing layers: A simple and flexible approach for resilience in packet networks. Technical Report 13, Simula Research Laboratory (2004)
5. Smith, D.R., et al.: Contingence/disaster recovery planning for transmission systems of the defense information system network. IEEE Journal on Selected Areas in Communications **12** (1994) 13–22
6. Luka, G., Fergus, P.: AIN applications to support NS/EP disaster response and recovery. In: Conference Record Military Communications Conference, MILCOM '1995. (1995) 843–847

7. Beard, C.C., Frost, V.S.: Prioritized resource allocation for stressed networks. IEEE/ACM Transactions on Networking **9** (2001) 618–633
8. Haungs, M., Pandy, R., Barr, E.: Handling catastrophic failures in scalable internet applications. In: Proceedings of the 2004 Interantional Symposium on Applications and the Internet (SAINT'04), Tokyo, Japan (2004)
9. Suurballe, J.W.: Disjoint paths in a network. Networks (1974) 125–145
10. Macgregor, M.H., Groover, W.: Optimized k-shortest-paths algorithm for facility restoration. Software-practice and experience **24** (1994) 823–834
11. Medard, M., Finn, S.G., Barry, R.A.: Redundant trees for preplanned recovery in arbitrary vertex-redundant or edge-redundant graphs. IEEE/ACM Transactions on Networking **7** (1999) 641–652
12. Bartos, R., Raman, M.: A heuristic approach to service restoration in MPLS networks. In: Proc. ICC. (2001) 117–121
13. Grover, W.D., Stamatelakis, D.: Cycle-oriented distributed preconfiguration: Ring-like speed with mesh-like capacity for self-planning network restoration. In: Proc. ICC. Volume 1. (1998) 537–543
14. Stamatelakis, D., Grover, W.D.: IP layer restoration and network planning based on virtual protection cycles. IEEE Journal on selected areas in communications **18** (2000)
15. Labovitz, C., et al.: Origins of Internet routing instability. In: Proceedings of IEEE/INFOCOM. (1999)
16. Labovitz, C., Ahuja, A., Bose, A., Jahanian, F.: Delayed Internet routing convergence. IEEE/ACM Transactions on Networking **9** (2001) 293–306
17. Cicic, T., Hansen, A.F., Gjessing, S., Lysne, O.: Applicability of resilient routing layers for k-fault network recovery. In: Proc. ICN'05. (2005)
18. Kvalbein, A., Hansen, A.F., Cicic, T., Gjessing, S., Lysne, O.: Fast recovery from link failures using resilient routing layers. In: submitted to the 10th IEEE Symposium on Computers and Communications (ISCC 2005), La Manga, Spain (2005)
19. See http://dmz02.kom.e-technik.tu-darmstadt.de/∼heckmann/.
20. See http://www.cs.washington.edu/research/networking/rocketfuel/.
21. Medina, A., Lakhina, A., Matta, I., Byers, J.: BRITE: An approach to universal topology generation. In: IEEE MASCOTS. (2001) 346–353
22. O'Mahony, M.J.: Results from the COST 239 project. ultra-high capacity optical transmission networks. In: Proceedings of the 22nd European Conference on Optical Communication (ECOC'96), Oslo, Norway (1996) 11–14
23. Rekhter, Y., et al.: A border gateway protocol 4 (BGP-4). IETF, RFC 1771 (1995)
24. Moy, J.: OSPF version 2. In: IETF, RFC 2328 (1998)