Engineering Robust Next-Generation Networks

G. Panza Senior Member IEEE[1], A. Capone Senior Member IEEE[2], D. Pinarello[3] and P. Belotti[4]

1, 2, 3 – CEFRIEL/Politecnico di Milano, via Fucini 2, 20133 Milano, Italy (gianmarco.panza@cefriel.it, capone@elet.polimi.it)

4 - Tepper School of Business, Carnegie Mellon University, 5000 Forbes Ave, 15213 Pittsburgh PA (belotti@lehigh.edu)

Abstract — **Next-Generation Networks (NGNs) employ the Internet Protocol (IP) over a wide variety of packet-switching technologies, which often lack in fault resilience enabling features. An overlay MPLS infrastructure with its fast-reroute mechanisms can be deployed to overcome such an issue. Addressing NGNs robust to single link and node failures, an off-line method to effectively calculate working and recovery paths for highly demanding services, is proposed and analyzed.**

The strength of our work is the ability to address two recovery techniques in a very simple manner, by formulating an Integer Linear Programming (ILP) problem, optimizing either the overall switching delay experienced by the user in case of failure or the bandwidth allocation thanks to a shared protection, while limiting the recovery time to some tens of ms as in SONET/SDH networks.

Index Terms — **Fault resilience, FRR, ILP, MPLS, NGN, Shared Protection, Traffic Engineering.**

I. INTRODUCTION

NOWADAYS, the Internet scenario is more and more characterized by value added services based on applications (e.g. VoIP, videoconferencing, video streaming) with stringent Quality of Service (QoS) requirements.

A critical issue to cope with is the robustness to network faults: link or node failures should be recovered quickly (as fast as some tens of ms to be comparable with current SONET/SDH networks), transparently to users.

Multi-Protocol Label Switching (MPLS) [1] is an advanced forwarding technology which uses the control plane of the IP routing protocol. MPLS can react rapidly to faults by switching failed connections to secondary paths. General specifications and bandwidth reservation for traffic engineering and path recovery are discussed in [2], [3] and [4]. Providing reliable services in MPLS is studied and fast rerouting techniques are proposed in [5][6][14][15].

This work[1] aims at addressing in a straightforward manner the traffic engineering problem of a MPLS network fully recoverable against single link or node failure, relying on fast-reroute mechanisms. We assume that the MPLS network has a two-connected topology (i.e. any two nodes admit at least two disjoint paths between them) with given link capacities and traffic demand set. For such a demand set, we calculate the working and recovery paths for all the requested connections, subject to capacity and recovery time constraints.

The algorithm is based on a simple and computationally efficient Integer Linear Programming (ILP) model, which optimizes either the bandwidth allocation by applying the Shared Protection (SP) or the overall switching delay, as seen by the user in case of failure, while constraining the recovery time to a given bound (i.e. some tens of ms).

The remainder of this paper is organized as follows. The next section overviews and analyzes the issued path restoration schemes. Then a description of our method and the discussion of some practical cases are provided. Finally, the last section summarizes the work and outlines the main conclusions.

II. ROBUSTNESS WITH MPLS

Path recovery consists in rerouting traffic around a failure, i.e. all packets routed through a link or node that have failed are rerouted along an alternative path (called recovery path) [5].

There are two basic models for path recovery: rerouting and protection switching. Rerouting is a model that establishes a recovery path (RP) after a failure on its working path (WP). Protection switching is a model that establishes a RP prior to any failure on the WP, therefore it can support fast restoration schemes. Depending on how the repairs are carried out upon the occurrence of a failure on the WP, we deal with global repair or local repair. In global repair, an alternative path is established from the source to the end of the WP, protection is always activated on an end-to-end basis, irrespective of where a failure occurs. In local repair, a portion of WP is protected by a dedicated alternative path and protection is activated by each LSR that has detected a failure. Local and global repair have different advantages [7]. The former can handle concurrent link failures more easily, but it needs more network resources as an end-to-end path has to be established. Both local and global repair techniques can employ a shared restoration [9][10] for an efficient resource utilization. A scheme using protection switching has been designed by Haskin and Krishnan [8].

The merit of such a scheme is that almost no packet loss occurs during link/node failure and a buffering technique [14] can be employed to avoid re-ordering of packets.

A local repair mechanism, called One-to-One backup [6], can be used to reduce the recovery time. The idea is to deploy an alternative path, called detour path, between the ends of each protected portion of the working path. In this way the length of the recovery path is decreased and consequently, the recovery time shortened. For fast reroute, the detour path needs to be pre-setup for each portion of the working path. Therefore, to protect a working path composed of N nodes, N-1 detour paths are required.

III. WORK DESCRIPTION

Our work aims at designing a method to compute a set of WP-RP pairs for global repair, and a set of detour paths (DPs) for the local protection of portions of a working path in a network, such that the whole working path is protected against single faults. The computed WPs, RPs and DPs must minimize the overall resource allocation (and thus maximize the residual network capacity).

As a starting point, we consider a generic network topology with a set of ingress, transit and egress routers and a set of traffic demands. The network links may have different capacities and costs. The problem we tackle is to select, for every traffic demand, working and recovery paths that satisfy the following requirements.

- Capacity constraint: for every link, the overall required capacity on the link must not exceed its capacity.
- Protection constraint: each WP must always be protected by one recovery path (or set of DPs).
- Recovery time constraint: the recovery time should be as in a SONET/SDH network (i.e. some tens of ms).

An ILP formulation for a problem similar to ours has been introduced in [18], but it can only be applied to small networks. A Dynamic Programming scheme is then used to obtain a solution in more realistic cases.

A novelty of our approach is that we use ILP model with an exact ILP solver, to obtain an optimal solution, even for large real networks.

A. Recovery path decomposition – the U-shaped paths

In order to search for recovery paths, we point out that for a given WP we can find a U-shaped detour path between the ingress and egress nodes of the WP portion to be protected. Let us call "U-path" this basic detour path, as depicted in Fig. 3. We can decompose the U path into four portions between the four delimiting routers A, B, C, and D, as shown in Fig. 4. From this basic decomposition we see that every U has three portions, with the following meaning in a network topology.

- First portion: reverse WP segment (A → B);
- Second portion: a shortcut from WP to RP (B → C);
- Third portion: a RP segment (C → D).

An extra portion, consisting of a return path from the RP to the WP (D → A), is used only in the case of One-to-One backup (i.e. for the DP determination). It is worthwhile to highlight that every U implements both the previously referenced target Fast Re-Routing (FRR) techniques at the same time.

In order to form a complete WP-RP path pair where every portion of a given WP is protected, we only need to chain several U by collapsing the "fourth portion" of a detour path with the "second portion" of the following one, as shown in Fig. 5. With this scheme, every link belonging to the WP has a protection DP (essentially, the U) and the "third portions" of the chained U form a complete recovery path. Our model implements the shortcut mechanism as defined by Haskins, since the "second portion" of the U path can be considered as

the shortcuts to the RP (see Fig.5). also, every node belonging to the WP has a protection DP, as natural with Haskin's method and by activating also the following chained U, when a fault happens in the node A of a given U with One-to-One Backup (in this case the restored traffic is to be routed from the "third portion" of the given U, directly to the "third portion" of the following U). Considering a set of directed arcs identified by points A, B, C, and D of Fig. 4. We associate a portion of the working and restoration resources to a single variable and a solution will be defined by a set of binary variables u_{AB}^{CD} that represent the concatenated *U-paths* and that are equal to one if the corresponding U-path is used in the solution.

In practice, we introduce special graph structures (U-paths) and use them in an ILP model instead of a flow ILP model, where one variable is defined for each arc in the graph and the protection mechanisms are described by constraints. Indeed, we embed these mechanisms in the model variables, obtaining a model with fewer complicating constraints and hence easier to manage.

B. Least cost path determination

The first step of our method consists in calculating a set of minimum cost paths between each pair of ingress and egress routers. Then, the Us to protect the portions of all possible WPs are determined.

The minimum cost paths are generated through Recursive Enumeration Algorithm (REA) described by Jiménez and Marzal [11].
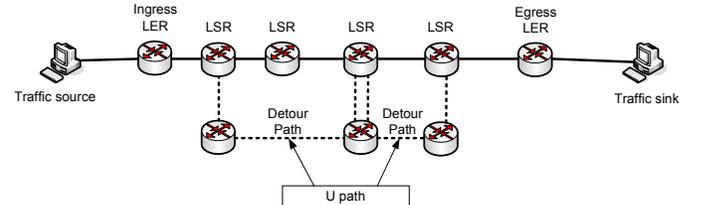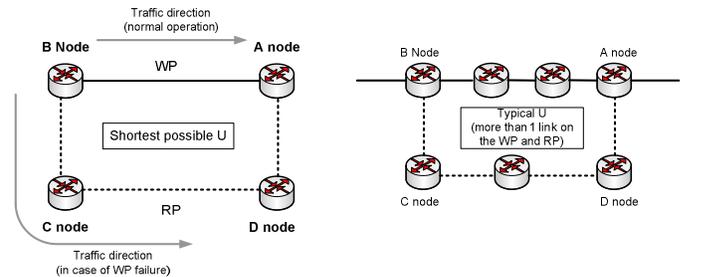
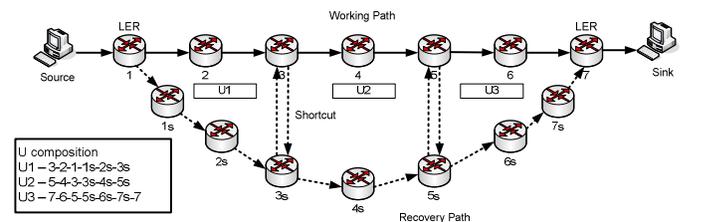

Fig.1. U path examples

Fig. 2. U-path decomposition

Fig. 3. Example of resulting U chain.

The set of paths is created as follows.

- First, calculating a set of K minimum cost paths between each pair of border routers.
- Second, from the K paths finding all the completely disjoint pair of paths that can form the WPs and RPs between every pair of border routers.
- Third, for every WP-RP pair, obtaining all the minimum cost shortcuts that connect the routers of the WP to the routers of the RP, thus identifying all the Us.

The recovery time constraint implies some restrictions on the length of the U. In particular, the portion of the WP in the U (named "first portion") has at most 3 links.

The result of the last step is a set of U paths to which our formulation is restricted. This is equivalent to establishing an initial set of variables in a Column Generation approach.

C. Selecting the best Us

The next step of our algorithm is to calculate the optimal WP-RP pairs, or equivalently, the optimal chains of Us. An Integer Linear Programming (ILP) problem has been formulated for the purpose.

Let us define $L_u \subset L$ the set of links contained in a specific u, for all $u \in U$. As we have specified before, this set has three portions if we are running our algorithm for engineering the network with Haskin's method and 4 portions in the case of the One-to-One backup method.

We also point out that we allow demands not to be accepted, but for each unsatisfied demand we introduce a high penalty in the objective function that is given to the ILP solver.

This avoids unfeasibility of the routing problem and obtains a solution with a minimum number of unsatisfied traffic demands. Recovery time is instead guaranteed by the constraint on the length of the "first portion" of the U and the a priori generation of U variables. The optimal solution of the ILP problem leads to a network engineering with working and recovery paths for every accepted traffic demand, which maximizes the overall residual capacity, allowing the maximum possible number of traffic demands to be satisfied.

An efficient resource allocation is addressed by applying the shared protection [9][10]. On each link the maximum bandwidth needed for all possible working conditions (i.e. also in case of single link/node failure) is allocated (hence, considered in the related constraint), rather than the sum of the bandwidth to be allocated for each satisfied traffic demand both in the WP and RP. The SP is fully integrated in the formulated ILP model in order to minimize the resulting resource allocation.

IV. ANALYSIS OF PRACTICAL CASES

As an evaluation process, we have fed the developed engineering tool with several networks and traffic demand sets. All tests were performed using a laptop PC equipped with a 1.73 GHz Intel Core Duo T2250 of 2 GB RAM, and the GLPSOL ver. 4.9 (the solver LP/MIP GPLK standalone [13],

based on GNUMathProg, which is a subset of the well-known AMPL modelling language [12].

Hereafter, we provide results for topologies that resemble some realistic cases: partially mesh and resulting from either tangent or intersected rings (with a higher capacity of the core links for the latter ones), more often referring to metropolitan networks, where recovery time within some tens of ms was first supported (by technologies as SONET/SDH).

The aim is to show the simplicity, flexibility and correctness, as well as the good performance of our proposal, while reporting numerical results. The set of traffic demands for the network Mesh (49 nodes, 83 links and link capacity of 100 Mbit/s) are shown in TABLE I. Every triplet of numbers in the table indicates ingress node identifier, egress node identifier and amount of requested bandwidth for a given demand.

We applied our solution with the number K of initial minimum cost paths, obtained in the first phase with REA (as described in the previous section), equals to either 5 or 10. The designed tool computes the network configuration and the residual capacity for the two supported FRR methods (Haskin's and One-to-One backup), optimizing either the allocated capacity or the overall switching time (i.e. the extra-delay experienced by the user in case of link/node failure). In the process, the bandwidth is allocated for both the working and recovery paths, as for protection switching recovery techniques, with resource sharing (i.e. applying the SP).

Regarding the capacity allocation optimization, tables II report the outputs of our system for the different traffic demand sets A, B and A+B (in successive rows), with equal-cost links and related to the topology Mesh only for lack of room.

The first three rows refer to the results with the SP, while the last three rows refer to the results without it.

The four columns encompasses both the issued protection methods, Haskin (HK) and One-to-One Backup (DP) for either "fixed-length" Us (FIX), i.e. Us that have three links in the WP portion, or "variable-length" Us (VAR), i.e. Us that

TABLE I: TRAFFIC DEMAND SETS SPECIFICATION FOR NETWORK MESH

SET A			SET B		
1)	10	28	10	1)	45	10	10
2)	42	5	10	2)	25	6	10
3)	21	2	10	3)	14	32	10
4)	20	18	10	4)	1	11	10
5)	16	12	10	5)	21	18	10
6)	33	17	10	6)	7	12	10
7)	19	13	10	7)	12	5	10
8)	27	12	10	8)	41	24	10
9)	13	16	10	9)	28	5	10
10)	18	42	10	10)	16	12	10

TABLE II: RESULTS FOR NETWORK MESH, K=10

DP_Fix\Tc\M [Mbps/s/MB]	HK_Fix\Tc\M [Mbps/s/MB]	DP_Var\Tc\M [Mbps/s/MB]	HK_Var\Tc\M [Mbps/s/MB]
16470\2\7.1	16490\2.3\7.3	16470\5.2\9.4	16480\6.4\9.7
15950\3\8.1	16000\3.4\8.4	15970\7.2\10.6	16010\8.9\11
15790\4\9.2	15850\4.4\9.6	15810\12.3\12.1	15880\13.9\13.3
16410\1\4.1	16420\1\4.1	16410\2.1\6.3	16420\2.1\6.3
15870\1.3\5.2	15920\1.3\5.2	15870\2.7\7.8	15920\2.7\7.8
15690\1.6\6.6	15750\1.6\6.6	15690\4.7\8.8	15750\4.7\8.8

have three links in the WP portion at most. The residual capacity, computational time (T_c in seconds) and memory RAM usage (M, in Mb) are included.

The accepted traffic demands of the sets A and B are (2, 4) and (1, 2, 3, 4, 5, 6, 8) for K=10. With a higher value of K the number of admitted traffic demands increases, because REA finds (if existing) a bigger number of disjoint paths to form WP-RP pairs.

The gain in spare capacity by the SP grows with the increase of the number of accepted traffic requests. The peak is reached with Haskin FRR and variable-length Us, with 20 traffic demands and 10 initially calculated shortest paths (15.880 Mbit/s of residual capacity, 130 Mbit/s more than without the SP). Indeed, Haskin's doesn't involve the "fourth portion" of the Us and "variable-length" Us allow for a greater flexibility, being a superset of the "fixed-length" Us (i.e. Us having a number of links in the first segment lower than three can be also selected to optimise the objective function).

Data have been computed in very few seconds with a memory usage of just some Mbytes, for both "fixed-length" and "variable-length" Us. Whereas classical edge flow formulations usually fail to provide a provably good solution in short time. Such results demonstrate the scalability of our approach for large real networks.

As for path-based restoration, various solutions [14][16][17] have been proposed in the recent years. The advantage of such solutions over ours is a slightly higher network capacity efficiency actually due to the lack of shortcuts between the WP and RP. However, they are much slower in restoration, because in such proposals the source node is responsible for the traffic switching to the backup path and cannot perform the operation until it receives a fault notification from the node that has detected the failure. While, a key feature of our method is that it can address the same recovery time as in link-based solutions and having capacity efficiency close to that of path-based solutions.

All the investigated networks have been also tested with non-equal cost links. This has an impact on both the set of shortest path initially calculated by REA, and on the objective function to be optimized, in terms of weight assigned to the residual capacity of each link.

V. CONCLUSIONS AND FUTURE WORK

In this paper, a flexible and simple system to effectively calculate working and recovery paths for Haskin's method with shortcuts and One-to-One backup, has been proposed and analyzed. In order to take advantage of modern exact solvers for discrete optimization problems, we have formulated the proposal as an Integer Linear Programming (ILP) system, with either the bandwidth allocation or the overall switching delay experienced by the user in case of failure, as objective function. We have introduced a class of variables based on a decomposition of the working and recovery paths, which allows for a simple model also applicable to large networks.

The correctness and flexibility of our proposal, as well as its efficiency in the resource exploitation by SP, have been demonstrated. Haskin FRR scheme performs better with respects to the bandwidth usage , not involving the "fourth portion" of the Us. While, One-to-One Backup could improve

the overall switching time in some topologies, depending on the set of available initial shortest paths. A higher number of such minimum cost paths calculated by REA could lead to accept more traffic requests.

A driven distribution of the load within the network (also useful to cope with additional traffic demands) is possible but with a less efficient capacity allocation than in the case of equal cost links. Future work is about extending the designed system to accomplish also point-to-multipoint traffic requests.

REFERENCES

[1] E. Rosen et al.. Multi-Protocol Label Switching Architecture. IETF MPLS WG. RFC 3031, Jan. 2001.
[2] D. Awduche et al.. Overview and Principles of Internet Traffic Engineering. IETF TE WG. RFC 3272, May 2002.
[3] D. Awduche et al.. Requirements for Traffic Engineering Over MPLS. IETF TE WG. RFC 2702, Sept. 1999.
[4] Chen, T.M.; Oh, T.H. Reliable services in MPLS, Communications Magazine, IEEE, Volume 37, Issue 12, Dec. 1999 Page(s):58 - 62 .
[5] V. Sharma et al. Framework for Multi-Protocol Label Switching (MPLS)-based Recovery. IETF MPLS WG. RFC 3469, Feb. 2003.
[6] P. Pan et al. Fast Reroute Extensions to RSVP-TE for LSP Tunnels. IETF MPLS WG. RFC 4090, May 2005.
[7] L. Jorge, T. Gomes. Survey of Recovery Schemes in MPLS Networks. Dependability of Computer Systems DepCos-RELCOMEX '06, pages 110-118, 2006.
[8] D. Haskin and R. Krishnan. A method for setting an alternative label switched paths to handle fast reroute. IETF MPLS WG. Internet draft, May 2000.
[9] Guangzhi Li; Dongmei Wang; Kalmanek, C.; Doverspike, R.; Efficient distributed recovery path selection for shared mesh restoration; Networking, IEEE/ACM Transactions on, Volume 11, Issue 5, Oct. 2003 Page(s):761 – 771.
[10] F. Aslam, S. Raza, Z. Uzmi, Y. Kim.. Bandwidth Sharing with Primary Paths for Protection Routing in an MPLS Network, In Proc. of INFOCOM 2006, pages 1-6, 2006.
[11] VM Jiménez and A. Marzal. Computing the *K* shortest paths: A new algorithm. and an experimental comparison. WAE 1999, LNCS 1668: 15-29, 1999.
[12] R. Fourer, D.M. Gay and B.W. Kernighan. A Modeling Language for Mathematical Programming. Management Science 36 (1990), pp.519-554. http://www.ampl.com .
[13] The GLPK (GNU Linear Programming Kit) package for solving large-scale linear programming (LP), mixed integer programming (MIP), and other related problems. http://www.gnu.org/software/glpk/.
[14] Hundessa, L.; Domingo-Pascual, J.; Reliable and fast rerouting mechanism for a protected label switched path. Global Telecommunications Conference, 2002. GLOBECOM '02. IEEE Volume 2, 17-21 Nov. 2002 Page(s):1608 - 1612 vol.2.
[15] S.Petrov. An approach for MPLS Recovery. ACM International Conference Proceeding Series; Vol. 285 , Proceedings of the 2007 international conference on Computer systems and technologies, 2007.
[16] M. Kodialam and T. V. Lakshman. Dynamic routing of bandwidth guaranteed tunnels with restoration. In Proc. of IEEE INFOCOM'00, pages 902-911, 2000.
[17] G. Li, D. Wang, C. Kalmanek, and R. Doverspike. Efficient distributed path selection for shared restoration connections, In Proc. of IEEE INFOCOM'02, pages 140 - 149, June 2002.
[18] D. Xu, Y. Xiong, and C. Qiao,"Novel Algorithms for Shared Segment Protection," IEEE Journal on Selected Areas in Communications, Vol. 21, No. 8, pp. 1320-1331, Oct. 2003.