Safety and Reliability for Managing Risk – Guedes Soares & Zio (eds)
© 2006 Taylor & Francis Group, London, ISBN 0-415-41620-5

Multi-scheme recovery in MPLS networks

L. Jorge
Polytechnic Institute of Bragança, INESC Coimbra, Portugal

T. Gomes
Dept. of Electrical and Computer Engineering, Coimbra University, INESC Coimbra, Portugal

ABSTRACT: *MultiProtocol Label Switching* (MPLS) networks have been proposed as a solution to offer reliable, efficient and differentiated telecommunication services. Connection oriented technologies allow the use of traffic engineering approaches to select routes and are also expected to enhance the reliability of IP networks. Routing protocols can be robust and survivable but take a long time to recover from faults, which will not be acceptable for many applications. Therefore several schemes and frameworks for MPLS recovery have been proposed, some of which allow network recovery in the tens of milliseconds.

In this work is proposed a multi-scheme recovery methodology, which intends to increase overall network resilience, while using network resources efficiently by taking into account resilience requirements of different class types. The objective of the methodology will be to offer protection to a set of services by choosing the most appropriate recovery scheme, taking into account the service class, the network state, and the characteristics of available recovery schemes. The appropriate recovery scheme will therefore be chosen based on a combination of quantitative measures and qualitative classification.

1 INTRODUCTION

MPLS networks have been proposed as a solution to offer reliable, efficient and differentiated telecommunication services. Connection oriented technologies as MPLS allow the use of traffic engineering approaches to select routes, *Label Switched Paths* (LSPs), and are also expected to enhance the reliability of *Internet Protocol* (IP) networks. Routing protocols (OSI layer 3) are robust and survivable but they take a long time to recover from faults, which will not be acceptable for many applications. Therefore several schemes and frameworks for MPLS recovery have been proposed, some of which allow network recovery in the tens of milliseconds. Compared to the (usually faster) lower layer recovery, MPLS recovery can make use of additional information to provide differentiated recovery, only to those traffic flows that require it, therefore saving resources.

Service providers must satisfy agreed throughput, maximum delay, and maximum down times, among other performance measures, as stated in contractual *Service Level Agreements* (SLAs). MPLS has been chosen by network service providers because it allows explicit control of traffic flows in the network and also quick recovery in the presence of link and/or node failures, and therefore reduces the risk of disrespecting SLAs.

The paper is organised in the following manner. First, MPLS networks and the concept of MPLS-based recovery procedures will be introduced. Then, a short overview of recovery approaches, in the context of MPLS networks, will be presented. Finally a multi-scheme recovery methodology, which intends to increase network resilience, while using network resources efficiently, will be proposed.

1.1 *MPLS networks*

MPLS is a network technology that offers *Quality of Service* (QoS), *Traffic Engineering* (TE) and many new applications, to overcome some the IP based network limitations. An MPLS network is made of routers known as *Label Switching Routers* (LSRs). At the ingress points of an MPLS network (the ingress LSR), labels are added to incoming packets, and all forwarding in the MPLS domain, along a *Label Switched Path* (LSP), is made based on those labels.

A packet label is removed at the egress LSR, before the packet is sent to a non MPLS-router. Figure 1 presents an active LSP, as well as the corresponding Ingress and Egress LSRs. Because packet forwarding in MPLS is based only in label switching, it tends to be faster than IP. MPLS supports two kinds of routing: implicit (hop-by-hop) routing, and explicit routing. In hop-by-hop routing for a LSP, each node chooses the

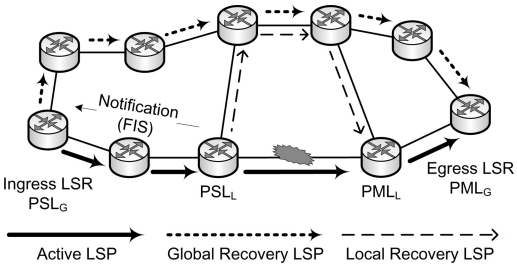

Figure 1. Local and global recovery.

following one independently. On explicit routing however, a single LSR specifies several (or even all) the LSRs that define the path of the LSP. Usage of explicit routes by MPLS traffic engineering allows improved network resource usage while ensuring the desired characteristics along a LSP.

The LSPs are set up using signalling protocols, such as *Resource Reservation Protocol with TE* (RSVP-TE), *Constraint-based Routing – LDP* (CR-LDP) or just *Label Distribution Protocol* (LDP).

In MPLS it is possible to route traffic satisfying specified QoS constraints, using teletraffic engineering approaches resulting in a more efficient use of network resources. MPLS networks were first deployed in core networks, have moved successfully to metropolitan area networks and are starting to be found in access networks. All these networks must offer reliable services with guaranteed QoS, and because protection mechanisms at the network layer take too long to recover from fault, in (Sharma et al. 2003) a standard framework for MPLS-based recovery was proposed.

1.2 *Recovery in MPLS networks*

To provide dependable services MPLS networks make use of a set of procedures (detection, notification and fault recovery) which seek to ensure appropriate protection for the traffic carried in the LSPs. When a fault happens in the active LSP, the recovery mechanism must re-direct the traffic to a recovery LSP which bypasses the fault.

The two basic recovery models used to redirect traffic, defined on (Sharma et al. 2003), are rerouting and protection switching. These approaches differ mainly on when the recovery path is established. When rerouting is used the *Recovery Path* (RP) (the path by which the traffic is restored after the occurrence of a fault) is signalled only upon fault detection in the *Active Path* (AP) (the path which carries traffic before the occurrence of a fault). Protection switching pre-establishes a RP before any failure detection in the protected AP. So, protection switching is faster than rerouting but cannot handle simultaneous faults in the active and the recovery paths. On the other hand, rerouting is generally slow, and cannot offer QoS guaranties upon failure, but can use resources in a more efficient way.

The recovery scope can be global or local. When a failure occurs the LSR responsible for switching or replicating the traffic between the AP and RP is the *Path Switch LSR* (PSL). The *Path Merge LSR* (PML) is the LSR responsible for receiving the RP traffic. If the PML is not the egress LSR it will merge the RP traffic back onto the AP.

Figure 1 presents an example of a global recovery path, for the depicted active path, and the corresponding PSL$_G$ and PML$_G$. For the same active path, its also presented an example of a local recovery path, for recovering a fault in the marked link, and the corresponding PSL$_L$ and PML$_L$. Global recovery intents to protect against any link or node fault in a path, whereas local recovery intents to protect against a link or node fault and to minimise the time required for fault propagation. Local recovery is attempted by the node immediately upstream of the fault according to (Sharma et al. 2003). Global recovery is usually slower than local recovery because the failure notification message, the *Fault Indication Signal* (FIS), has to travel to the *Point of Repair* (POR), a LSR that is setup for performing MPLS recovery. The POR can be a PSL or PML depending on the type of recovery scheme employed (in Figure 1, the node PSL$_G$ is the POR for global recovery).

When the node immediately upstream of a fault is unable to recover a failed path (either because that node is not a POR or because the POR was unsuccessful in its attempt), sends a FIS to the node immediately upstream, where possibly a new attempt of local recovery takes place. This type of recovery appears in (Hong et al. 2004; Yoon et al. 2001), and even if the node that successfully recovers the AP coincides with the head-end node, we consider it a mechanism of local recovery.

When a recovery path is pre-established, its resources can be pre-reserved. To provide more efficient resource usage, pre-reserved resources can be shared by multiple primary resources that are not expected to fail at the same time – like in (Kodialam & Lakshman 2002). Interdemand sharing takes place when backup bandwidth of different active LSP is shared. This type of approach is used for 100% protection in single failure scenarios. When an active LSP is protected using several backup LSPs and they share bandwidth, we say intrademand sharing has taken place.

A network service provider should apply different recovery options according to the QoS characteristics of the carried traffic flow (service class), while satisfying service-level agreements and simultaneously consuming minimum resources. This has spurred the

proposal of a huge number of recovery schemes that offer dependability by fault tolerance. However, in most of these studies, the subject of QoS is either not addressed or is reduced to bandwidth guarantees (Kodialam & Lakshman 2002). In (Jorge & Gomes 2006) a revision of MPLS recovery schemes can be found.

In (Ricciato et al. 2003) a scheme is proposed that provides some sort of differentiated handling of protection, but only as far as it refers to bandwidth guarantees. (Autenrieth & Kirstadter 2002) shows that differentiating recovery options for traffic flows can provide reduced resource usage. Several combinations of recovery mechanisms (e.g. protection versus rerouting, local scope versus global scope) were analysed based on bandwidth usage. Other measures were proposed in a similar study (Dong et al. 2003) for optical and GMPLS networks. In (Calle 2004), is proposed a scheme which provides an explicit choice of protection mechanisms to be used by different traffic classes. His work also puts forward alternative protection mechanisms for the same traffic class, and describes the expected and observed performance of those mechanisms. However, the operational indication on how to choose among the protection mechanisms (presented by an analytical algorithm based on packet loss, failure recovery time and resource consumption) is limited to protection within three different recovery scopes. According to (Calle 2004) including other protection models in this approach may lead to scalability problems.

The proposal presented here considers service differentiation in a way similar to the previously introduced, but intends to provide a scheme that adapts its behaviour according to a set of variables, including network state, and to provide operational guidance on which recovery scheme (through the aggregation of several protection mechanisms) to use at each moment according to the service class characteristics.

2 REVIEW OF SOME MPLS RECOVERY SCHEMES

We will now present a description of some fault recovery schemes proposed in the context of MPLS networks. Recovery schemes similar to these (i.e. which share *some* of their characteristics) can be the templates for the schemes in Table 1 which will be used in Section 3.1. Table 1 presents 9 recovery schemes that are characterized by recovery model, scope and other criteria. Their names describe whether rerouting (R) or protection (P) is used, and whether the scope is local (L) or global (G). Therefore, a PG scheme provides global protection, a RL scheme local rerouting, and so on.

Table 1. Recovery schemes.

Name	Description
PG1	With reservation sharing
PG2	No reservation
PL1	Reservation, no interdemand sharing
PL2	Reservation, with interdemand sharing
PL3	No reservation
RG	No reservation
RL1	With pre-calculation
RL2	No pre-calculation
RL3	Like RL2, different path choice algorithm

We will start by reviewing some approaches for protection switching and later discuss the characteristics of some rerouting approaches.

2.1 *Protection switching*

Recently, two *Fast Reroute* (FRR) schemes were standardised. According to (Pan et al. 2005), such schemes allow redirection times of tenths of milliseconds, which is required by real-time applications. One-to-one backup is one of these schemes. In this recovery scheme, for each protected LSP, a protection path, named "Detour LSP", is required in each *Point of Local Repari* (PLR), the head-end of a local RP. This means that each LSR on a protected LSP must compute a protection path from itself to the egress LSR. Each detour LSP may be recomputed and updated by the PLR at any time.

This scheme (one-to-one backup) needs a large number of recovery paths. This number is a function not only of the number of protected paths, but also of the length (hop count) of those paths. To minimise this number, path merging, among detours of a protected LSP and that LSP, is implemented whenever possible. This recovery scheme will be available only for protected LSPs (LSPs which requested protection when they were first signalled) possibly with guaranteed bandwidth. In Table 1 these schemes are denoted by PL1 and PL3, respectively with and without reservation.

In (Kodialam & Lakshman 2002), when a request for a new protected LSP is issued, the working and the recovery paths are simultaneously calculated on-line. If the requested bandwidth cannot be guaranteed for the AP and the RP, the request is rejected.

This scheme requires knowledge of the total bandwidth used by the APs, of the total bandwidth used by the RPs in each link (just the aggregated information, not broke down for each LSP), and also of the free bandwidth in each link. The authors call this the *Partial Information* (PI) model. Using this model, a dynamic routing heuristic algorithm to achieve local recovery is presented, computing both the active path

and the recovery paths. The algorithm presented considers intrademand sharing (although not explicitly between the AP and corresponding RPs) and also some interdemand sharing. This scheme will be selected as PL2 in Table 1, and will be used in the proposed recovery system in section 3.

In (Kodialam & Lakshman 2002 algorithms for global recovery are also presented. Global recovery under the PI model allows just some interdemand sharing. It selects the active and recovery paths via linear programming models (or heuristics, mostly when facing NP-hard problems). The AP and RP are computed at the ingress node without full information, but when actual reservation occurs, it is possible to reserve the exact amount of protection bandwidth in each link, provided that the RP is signalled carrying the complete path of the corresponding protected AP. This scheme handles single link failures, but the authors showed it can easily be extended to handle also single node failures. The previous scheme and also a possible variant, resulting from replacing the sub-algorithm used for obtaining a pair of disjoint paths by the approach in (Gomes et al. 2005) are two possible choices for PG1. For illustrating the approach in section 3 we will select the PI global protection scheme in (Kodialam & 2002) for PG1 in Table 1.

Scheme PG2 could be implemented using any algorithm for global recovery without pre-reservation.

2.2 *Rerouting*

In Yoon et al. 2001, Ahn et al. 2002, Hong et al. 2004) the recovery path is established only after the fault occurs. These schemes implement local recovery, establishing a recovery path for each failed active path. The recovery path is only used for the duration of the fault. While (Yoon et al. 2001) pre-calculates the recovery paths, (Ahn et al. 2002, Hong et al. 2004) only make this calculation upon detection of the fault.

In (Yoon et al. 2001), for each working path, the ingress LSR and each intermediate LSRs with recovery capability pre-calculate a recovery path to the nearest downstream LSR. These paths are updated whenever network state changes (link usage information) arrive at the path computing LSRs. This scheme assumes that all the possible faults are either failure or degraded usage of a single link in a single MPLS protection domain. Facing such a fault, the PSL will be the nearest upstream LSR with recovery capability. If the direct upstream LSR does not have such capabilities, a FIS will be sent through a *Reverse Notification Tree* (RNT). Since the intermediate LSRs can only handle single link failures, a FIS (if sent) will always be propagated to the ingress LSR. Scheme RL1 in Table 1 is modeled on this scheme.

In (Ahn et al. 2002), the PSL is also the nearest LSR upstream from the fault. It chooses the recovery path as the least cost path among the alternative paths, from the PSL up to a LSR downstream on the working path that may become a PML. On Table 1 this scheme will be used for RL2.

In (Hong et al. 2004) an iterative process for building the recovery path is proposed. In this rerouting scheme, each iteration searches an optimal alternative path from the PSL to the working path egress LSR. In order to minimise the recovery time, the scheme begins to look for such a path on a subset of the LSRs, where the tentative PSL is the LSR closest to the failure and the PML is always the egress LSR. This subset (delimiting the range of recovery) is successively enlarged whenever a recovery path is not found. Every time the subset is extended, in a process called hierarchical restoration, a new PSL is used, the next upstream LSRs in the working path. This procedure, if unsuccessful, is repeated until the PSL is the ingress node. In (Hong et al. 2004), the number of iterations required defines the speed of recovery. As the total number of operations increases, each successive iteration is harder to solve. However, as the authors (Hong et al. 2004) pointed out, using too small recovery ranges may result in ignoring bandwidth available elsewhere in the network. Scheme RL3 of Table 1 corresponds to this scheme.

Finally, scheme RG of Table 1 is a generic global rerouting scheme, without any kind of reservation. All these schemes seem to be very efficient on resource usage, which is partially explained by the rerouting model.

3 A RECOVERY SYSTEM PROPOSAL

We will now present a recovery system proposal which will choose among a set of schemes the ones which best suit each particular situation. It is proposed that this system will be implemented in a set of LSRs in the network, and as such it should have either modest requirements or be able to be scale according to the power available on those LSRs. Therefore, the system will not be monolithic, but rather made of a series of components that can be expanded, for instance incorporating additional schemes or characteristics to be considered.

3.1 *System structure*

The objective of the methodology will be to offer protection to a set of services by choosing the most appropriate recovery scheme. This choice will be made taking into account the service classes (which usually may require a desired level of protection), the network state and the characteristics of the available or considered recovery schemes. This will result in a preplaned recovery approach for each LSP, which can

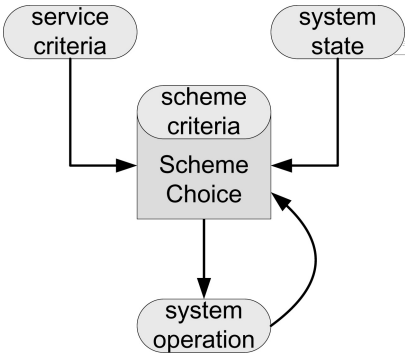

Figure 2. Main diagram.

be reviewed periodically, either because of network behaviour changes or to perform optimisation.

A recovery scheme will be chosen based on a combination of quantitative measures and qualitative classification. Examples of quantitative measures are link failure probability and actual resource utilisation, including bandwidth sharing. Some other quantitative measures, like recovery time, could also be used. Qualitative classification used may be, for example, whether a scheme uses protection switching or rerouting, and the kind of faults it addresses.

The network state will be characterised by the description of the traffic flows carried locally, available bandwidth in every link (and possibly bandwidth reserved for active and protection paths), link/node average reliability, and the present network topology (nodes may fail and link capacities may be reduced, also due to failure) including the description of *Shared Risk Link Groups* (SRLGs).

The structure of the proposed system is presented in Figure 2. In it, the final chosen scheme is function of a group of criteria:

service criteria – the *Diff-Serv-aware Traffic Engineering* (DS-TE) class types (CT), as defined in (Faucheur & Lai 2003), can form the base of differentiated handling (service class handling) of the offered traffic.

system state – a set of parameters describing the system state at *choice time*. Some criteria based on network state may demand the calculation of performance measures which could be hard to compute and/or require bulky or difficult to obtain data, about the network state. In this case the system may be built to operate using only available information, and a sub-set of possible criteria based on the obtainable performance measures, which may depend on the LSR processing capabilities/load at the evaluation time.

scheme criteria – the available schemes are described in terms of their global characteristics – like those presented in (Foo 2003; Jorge & Gomes 2006). Some preliminary choices may focus only on those general characteristics, with the actual scheme choice being made by later stages, respecting (or not) previous choices.

system operation – after a scheme has been chosen, changes in the network state (e.g. subsequent failures) or performance targets may require its change, or at least, its review.

All the previous elements influence the choice of the actual scheme to be used to protect a class type on a particular LSP. The details of the choice can be operationalised by a set of decision tables, encoding the decision preferences according to the previous criteria.

It is assumed that the set of recovery schemes that can be used is known beforehand, although there is a measure of freedom in choosing the actual recovery schemes to be used. We presented a set of 9 recovery schemes (in Table 1) that *may* be used in such a system.

While most of the schemes presented differ in externally observable characteristics, in Table 1 we see also a pair of methods (RL2,RL3) with similar characteristics, differing only on internal mechanisms. In this situation, to choose among them will require a keen knowledge of how the internal characteristics translate into actual performance. In fact, it may be easier to just choose one of them for default usage.

Notice also that the reservation described in the method characteristics refer to bandwidth reservation for backup paths, upon their establishment (it could also be denoted pre-reservation). Some presented schemes may be ill-suited to a DS-TE usage, since, e.g. using protection without reservation (that is, just performing path signaling) may be inadequate if the class properties require performing actual bandwidth reservation before usage.

The scheme choice mechanism will consist on several stages. These stages will be applied sequentially, although some stages may be skipped (if so desired), according to the recovery scheme selected in previous stages. Every stage will contain a set of decision tables. The number of stages is predefined, but the content of the tables is not, and neither is the number of tables to be used, although some care must be taken not to create an incoherent table set.

In each table, a set of conditions is described, and an action is selected whenever the criteria (on service, scheme or system state) match those conditions, although for some matches no action will be provided – null matches. The action will be usually the selection of a scheme to be used, although sometimes it will just be the selection of some additional conditions or parameters which will eventually lead to the final scheme. Any match (including null matches) ends the scanning of the current table, and by default allows for the scanning of the next one in the same stage. However, an action may also describe which stage should follow it,

thereby allowing skipping to the next stage, or to any other following stage. In the end of each stage there is a "default" table, which will be used in the following situations: a) whenever previous tables of the same stage failed the match; b) whenever previous choices are allowed to be amended (this situation occcurs when no jumps to a subsequent stage were made, after the action in previous tables of this stage); c) and finally to provide default values. It is desired, that in simple installations, just using the default tables will provide an acceptable service.

The overriding of previous choices can emerge from conflicting conditions. In such cases, the final choice is the last one, which requires care in the ordering of the tables (an example will be presented later). Other more elaborate methods for dealing with conflicting conditions may be used, provided they can resolve the conflict in the adequate time.

The recovery system has four stages:

- The first stage (A), to be performed when the service is required, provides the selection of the recovery model, the preliminary selection of the recovery scheme and sometimes the actual scheme selection;
- The second stage (B), to follow immediately after stage A, provides the actual scheme and/or which particular options of a scheme should be active;
- The third stage (F), which will be performed upon a fault, will try to use information regarding the fault characteristics to, if needed, improve or change the selected scheme;
- The fourth stage (O), will take place sometime after the fault occurrence (or possibly on a schedule) and allows optimization of recovery path, if desired.

3.2 *An application example*

As an example, we will now show how this multi-scheme system could be used in an application supporting four class types (Table 2). These classes are aggregates of traffic that should have common behaviour, as defined by DS-TE.

We will have a few tables for each stage, e.g. Tables 3 (#A.10), 4 (#A.20) and 5 (#A.99) for the first stage. The tables presented are illustrative and require validation by further study.

These tables will be, as stated, scanned sequentially. Notice that if pre-defined SLAs must be supported, other initial tables, regarding the conditions on such traffic, and desired behaviour, may be placed before #A.10 (Table 3), so such agreements may be honoured.

Whenever there is a match, besides taking the action defined by that match, a jump may be made to another stage (the jump destination is presented inside brackets, []). In Table 3, if the traffic load is light on CT2, a jump is made to stage B after selecting to use local protection, whereas if the class type is CT3, the scheme PL2 will be selected, and there will be a jump to stage F, which means additional configuration for this particular traffic will be made only upon a fault on this path. Frequently, no jump is selected upon a match (as it is shown on Table 3, for CT0 type traffic on heavily loaded networks), which means that the system should proceed to the following table, taking into account the selected action (in this example, Global scope recovery is selected for this traffic, though this choice may be reverted later on).

Whenever an action is selected by a match, this new action overrides previous choices. This override can be just partial, if the action itself just describes some characteristic of a scheme, rather than a full scheme – in this case, just the selected characteristic reflects the new action, and it is combined with the previous choices. In Table 4 such an override can occur for CT0, choosing Local recovery when the active path is long and bandwidth guarantees are required (assuming Global recovery was previously selected for this class).

On Table 4, notice also that two similar but different schemes are selected for CT1 and CT2 when the LSP size is short and bandwidth guarantees are required, since differences in the methods may make PL1 a little faster than PL2 in this particular situation.

Table 2. Traffic classes to support.

Class	Description
CT0	Best effort
CT1	High jitter tolerance; Medium/high recovery times
CT2	Medium jitter tolerance; Low recovery times
CT3	Low jitter tolerance; Very demanding recovery times

Table 3. #A.10 – on network traffic.

	Condition			
Load	CT0	CT1	CT2	CT3
Light	PG2[B]		PL[B]	PL2[F]
Heavy	G[ ]			

Table 4. #A.20 – on guaranteed bandwidth and active LSP path length.

		Conditions			
Bandwidth	*LSP length*	CT0	CT1	CT2	CT3
Guar.	*Short*		PG1[F]		
	Long	L[ ]	PL2[F]	PL1[F]	
No guar.	*Short*				
	Long				

Table 5. #A.99 – "default" rules for stage A.

	Condition			
	CT0	CT1	CT2	CT3
Always	R[ ]	R[ ]	P[ ]	P[ ]

Table 6. #B.10 – on the path failure probability.

		Conditions			
L/G	*failure prob.*	CT0	CT1	CT2	CT3
undef	*PFP1*	G[ ]			
	LFP1				G + L[ ]
	LFP2		G + L[ ]	G + L[ ]	G + L[ ]

Upon leaving the "default" table for a stage, which is the only one that *must* necessarily exist in each stage for the system to work, the system may proceed to the first table on the next stage if all the conditions are met. For example, after Table 5, the system may proceed immediately to Table 6. However, transitions to stage F require detection of an error on the path, and to stage O that some condition – e.g. an internal timer – signals to do so.

Some conditions, in the proposed tables, will only be true if the tested value is undefined – up to that point – e.g. Table 6 will only be required if there is no previous decision regarding using Local or Global recovery methods. Notice on Table 6 that the conditions may be quite complex. In this table PFP1 is true if the path failure probability is larger than some predetermined value, LFP1 can be activated whenever the link failure probability on the path is larger than some other significant constant on links near the end of the active path, and LFP2 may be triggered by having the link failure probability for some link on the path greater than twice the average link failure probability for links of the path. As can be expected, some of these tests can be expensive either computationally, on the data required, or on time needed to acquire it, and as such their usage must be balanced with the benefit expected from its usage.

The actions selected can also be complex. In Table 6 the selection is often G + L, which describes that two schemes are to be selected for this working path. The former (G) describes a property of the "regular" scheme to be applied throughout the working path, except on "exceptional" links that require different handling. In this case, a global recovery scheme is selected as the regular scheme, and local recovery is mandated for just some links in the working path, therefore creating a hybrid scheme. The actual links requiring "exceptional" handling are selected by conditions, either directly on this table or on subsequent tables (the details are omitted for clarity).

Table 7. #B.99 – "default" table for stage B.

		Conditions			
R/P	*L/G*	CT0	CT1	CT2	CT3
P	*L*	PL3[ ]	PL1[ ]	PL1[ ]	PL1[ ]
	G	PG2[ ]	PG2[ ]	PG1[ ]	PG1[ ]
	undef	PG2[ ]	PG2[ ]	PL1[ ]	PL1[ ]
R	*L*	RL2[ ]	RL2[ ]	RL1[ ]	RL1[ ]
	G	RG[ ]	RG[ ]	RG[ ]	RG[ ]
	undef	RG[ ]	RG[ ]	RL1[ ]	RL1[ ]

Table 8. #F.10 – changes on rerouting according to failure location.

			Conditions			
Loc	*R/P*	*L/G*	CT0	CT1	CT2	CT3
La1	*R*	*any*	RG[ ]	RG[ ]	RG[ ]	RG[ ]
La2	*R*	*any*	RG[ ]	RG[ ]	RG[ ]	RL3[ ]
La3	*R*	*any*	RG[ ]	RL2[ ]	RL3[ ]	RL3[ ]

Table 9. #F.99 – "default" table for stage F.

			Conditions			
Loc	*R/P*	*L/G*	CT0	CT1	CT2	CT3
Lr1	*R*	*L*	RG[O]	RL2[O]	RL3[O]	RL1[O]
	R	*G*	RG[O]	RG[O]	RL3[O]	RL1[O]
	P	*any*	RG[O]	RL2[O]	RL1[O]	RL1[O]

A few other tables could be placed before Table 7 which, in combination with Table 5, must provide a default scheme for every request that has not matched any specific table until then. After this stage, a protection scheme is supposedly selected for each request. Notice that such a scheme may be just a best effort to reroute the request upon failure. When a failure occurs, stage F is used and may lead to a modification of previously chosen recovery schemes for the affected LSPs.

In Table 8 the location of the failure on the *active* path is encoded. La1 refers to failure on a short active LSP, La2 refers to failure near the beginning of a long active path and La3 to a failure near the end long active path. As can be seen, the location of the fault may influence the kind of rerouting to be attempted.

In Table 9 (default table for F stage) a failure located on the *recovery* path is codified as Lr1. This table should only be triggered if simultaneous multiple

failures (or very quick sequential failures) cause the recovery path also to experience failure before this table is checked. This can thus be used to select an additional recovery scheme if the previous one is subject to failure. An example can be found in Table 9 where for (Lr1, R, G) and CT2 a local rerouting scheme is now the preferred solution.

While we present no table for stage O, that kind of table could be built like the previous ones. It should be used to provide an opportunity to decide when and how to optimize the "new" working path and select additional recovery paths after a failure, according to the class type. Also, sequential failures that were not handled by stage F could be addressed by this stage.

4 FINAL REMARKS AND FUTURE WORK

We proposed a multi-scheme recovery methodology, to increase overall network resilience, while using network resources efficiently by taking into account the different requirements of different class types. The methodology tries to offer protection to a set of services by choosing the most appropriate recovery scheme, taking into account the service class, the network state, and the characteristics of the available recovery schemes, as the depicted in a set of stages each composed of several decision tables. The methodology is embodied in a system, the details of which were described.

This is a preliminary study, and will be subject to further analysis and implementation, to check the validity of the approach. We hope however that the overhead this approach entails will be negligible and that it may improve on similar schemes regarding resource consumption and differentiated handling of services.

It should be noticed that fully defining the four stages may create a very complex system, with the possibility of poorly understood interactions among tables. Therefore the study of the behaviour of a subset of this system, containing only default tables and possibly reducing the number of stages can be a strong target for further research.

As further work, we propose to tune the contents and number of tables, and to compare (by simulation) the performance of the proposed multi-scheme recovery system with the performance of individual schemes, and also with recovery approaches with different recovery solutions per service class.

ACRONYMS

AP *Active Path*
CR-LDP *Constraint-based Routing – LDP*
DS-TE *Diff-Serv-aware Traffic Engineering*
FRR *Fast Reroute*
FIS *Fault Indication Signal*
IP *Internet Protocol*
LDP *Label Distribution Protocol*
LSP *Label Switched Path*
LSR *Label Switching Router*
MPLS *MultiProtocol Label Switching*
PI *Partial Information*
PLR *Point of Local Repair*
PML *Path Merge LSR*
POR *Point of Repair*
PSL *Path Switch LSR*
QoS *Quality of Service*
RNT *Reverse Notification Tree*
RP *Recovery Path*
RSVP-TE *Resource Reservation Protocol with TE*
SLA *Service Level Agreement*
SRLG *Shared Risk Link Group*
TE *Traffic Engineering*

ACKNOWLEDGEMENTS

Work partially supported by programme POSI of the III EC programme cosponsored by FEDER and national funds.

REFERENCES

Ahn, G., Jang, J. & Chun, W., 2002. An efficient rerouting scheme for MPLS-based recovery and its performance evaluation. *Telecommunication Systems*, 19(3–4): 481–495.

Autenrieth, A. & Kirstadter, A., 2002. Engineering End-to-End IP resilience using resilience-differentiated QoS. *IEEE Communications Magazine*, 40(1):50–57.

Calle, E., 2004. *Enhanced fault recovery methods for protected traffic services in GMPLS networks*. Ph.D. thesis, Department of Electronics, Computer Science and Automatic Control – Universitat de Girona, Girona, Spain.

Dong, S., Phillips, C. & Friskney, R., 2003. Differentiated-resilience provisioning in wavelength-routed optical networks. In *Teletraffic Science and Engineering – ITC 18*, volume 5b, pages 921–930.

Faucheur, F.L. & Lai, W., 2003. Requirements for support of differentiated services-aware MPLS traffic engineering. IETF RFC 3564.

Foo, J., 2003. A survey of service restoration techniques in MPLS networks. In *Proceeding of the 2003 Australian Telecommunications, Networks and Applications Conference (ATNAC) – CD/ROM*.

Gomes, T., Craveirinha, J. & Jorge, L., 2005. An effective algorithm for obtaining the minimal cost pair of disjoint paths with dual arc costs. Submitted for publication.

Hong, D.-K., Hong, C.S. & Dongsik-Yun, 2004. A hierarchical restoration scheme with dynamic adjustment of restoration scope in an MPLS network. In *Network Operations and Management Symposium*, pages 191–204.

Jorge, L. & Gomes, T., 2006. Survey of recovery schemes in MPLS networks. In *International Conference on*

Dependability of Computer Systems (DepCos 2006). Wroclaw University of Technology, Poland. To be presented.

Kodialam, M. & Lakshman, T.V., 2002. Restorable dynamic quality of service routing. *IEEE Communications Magazine*, 40(6):72–81.

Pan, P., Swallow, G. & (Eds.), A.A., 2005. Fast reroute extensions to RSVP-TE for LSP tunnels. IETF RFC 4090.

Ricciato, F., Listasti, M., Belmonte, A. & Perla, D., 2003. Performance evaluation of a distributed scheme for protection against single and double faults for MPLS. In et al., M.A.M., ed., *QoS-IP 2003*, number 2601 in Lecture Notes on Computer Science, pages 218–232. Springer Verlag, Berlin.

Sharma, V., Hellstrand, F., Mack-Crane, B., Makam, S., Owens, K., Huang, C., Weil, J., Cain, B., Anderson, L., Jamoussi, B., Chiu, A. & Civanlar, S., 2003. Framework for multi-protocol label switching (MPLS)-based recovery. IETF RFC 3469.

Yoon, S., Lee, H., Choi, D., Kim, Y., Lee, G. & Lee, M., 2001. An efficient recovery mechanism for MPLS–based protection LSP. In *Joint 4th IEEE International Conference on ATM (ICATM 2001)*, pages 75–79. Seoul, Korea.