Survivability Performance Analysis of Rerouting Strategies in an ATM/VP DCS Survivable Mesh Network

Kyamakya Kyandoghere
University of Hagen, Electrical Engineering, FTK
58084 Hagen, Germany

Abstract — *Several self-healing protocols utilizing virtual paths have been proposed in the relevant literature. Those which work in a mesh topology function according to three main rerouting strategies (though specific flooding administrations differ): local rerouting, source-destination rerouting, and local-destination rerouting. Most performance studies of self-healing protocols have considered restoration time as sole performance metric. This would have to be within the 2s threshold in order to guarantee service continuity. This one-sided metric needs to be completed. In this paper, I propose an extended performance(or goodness) metric framework in order to catch more performance aspects. In addition, this analysis uses survivability functions to measure the performance of rerouting strategies.*

Key words— ATM/VP transport network performance metrics, rerouting strategies, survivability functions.

I. INTRODUCTION

Today, an ever-increasing number of end users totally depend on telecommunications. They require fully-reliable services, whether the services involve voice or data traffic, and whether switched or private lines. The demand for high-quality performance, assured service continuity, and transparency to failures has never been greater. Users requiring enhanced levels of network survivability include not only the government, military, and emergency organizations, but also the general public.

Network affecting/degrading disasters come in many varieties. They range from natural disasters such as hurricanes, earthquakes, floods, and fire, to inadvertent errors such as software problems, operator error, or even to deliberately planted viruses and worms. Survivability strategies in the backbone mesh network are critical, considering the high concentration of traffic, and telecommunications operators' limited experience in provisioning radically new kinds of network services. These strategies include adequate rerouting and efficient control mechanisms.

A major benefit of setting survivability performance objectives is to ensure that, under given failure scenarios, network performance will not degrade below predetermined levels. To be useful for planning, such a set of performance objectives should realize network design and management goals. Since network survivability deals with network integrity in the wake of disaster, the specific event being addressed is an issue. It is important to specify the disaster type clearly (i.e. potential failure scenarios it may cause).

This paper clarifies several important issues and proposes a probabilistic framework for the study of disaster-based survivability performance of rerouting strategies, for an ATM/VP transport network. The objectives are namely: to investigate rerouting strategies and restoration capacity allocation; to propose a unified performance metrics framework for ATM-VP networks; and to use survivability functions in order to evaluate the performance of rerouting strategies. Section II discusses rerouting strategies in detail. In section III, I explain the proposed network performance metrics. The general procedure for finding survivability functions, the simulation model, and simulation results, are presented in sections IV, V and VI. Finally, conclusions of the study and future research avenues are outlined in section VII.

II. REROUTING STRATEGIES IN AN ATM/VP DCS (Digital Crossconnect Systems) MESH NETWORK

For ATM/VP-based networks there are several intrinsic features that could potentially be exploited to provide improved restoration techniques, beyond those established for STM-based transport networks [1-4]. The most important of these features are: ATM cell-level error detection, inherent rate adaptation, and non-hierarchical multiplexing. The ATM cell-level error detection mechanism increases the overall error check sampling rate per transmission interface, and thus provides a means for enhancing failure detection, and implementing alarm threshold policies. Inherent rate adaptation and non-hierarchical multiplexing allow flexible interface structures and eliminate multiplex stages within the network. This, in turn, allows increased link capacity utilization, flexible network reconfiguration, and dynamic bandwidth control — all of which can potentially be combined to yield a fast network reconfiguration with significantly lower spare capacity requirements for ATM-based restoration, when compared with STM-based network restoration [5].

The VP concept for ATM networks has already been studied in much detail [6]. A VP route is established by setting routing tables at VP connection points between VP connection end-points. Because of non-hierarchical multiplexing, path capacity does not need to be explicitly assigned at VP connection points during VP establishment time, but is handled by separate management procedures, including call admission control and usage monitoring. These are carried out at the ingress VP connection end-point. Thus connection points on the VP route perform no processing for VP capacity management and are therefore not affected by changes in VP capacity management procedures. This independence of route and capacity management leads to two important features of ATM networks: adaptive network reconfiguration and dynamic bandwidth allocation, both of which can be exploited to provide restoration strategies.

There are two approaches to provide network restoration under failure. In one approach the affected traffic is rerouted to the extend possible of facilities with spare capacity. Some form of capacity-search algorithm is used to hunt for this spare capacity, and a contention-resolution protocol is employed to assign routes to VPs in order to prevent a dead-lock (in the event that more than one VP is trying to claim usage of the same spare capacity). This approach is called dynamic restoration. Its shortcomings include: (a) considerable time is expended in discovering routes with spare capacity; and (b) restoration is possible only to the extent that spare capacity is available. A second approach overcomes the above shortcomings, by providing redundant capacity and precomputing alternate paths for all VPs under all failure scenarios. The philosophy of this approach is to plan redundancy for the most frequently occurring failure conditions. The precomputed alternate (or backup) routes are stored in the network nodes. Upon receiving failure information, every node carrying VPs affected by the failure will activate stored alternate paths. The time required to perform restoration consists, then, of the time taken to propagate the failure information to all nodes, and the time taken to activate the alternate paths at all nodes.

The merits and demerits of the two approaches are not one-sided (see Table I). The first approach requires less capital investment in the network and fundamentally aims at developing efficient strategies which optimize the use of available network resources after a failure (link or node), rather than plan redundancy into the network. After all, it may be fairly expensive to plan for all possible failure scenarios and precompute alternate paths for large scale networks. Hence, tradeoffs for both approaches exist. In consideration of this fact, two solutions are conceivable: (a) a combination of the two approaches, i.e., to plan redundancy for a set of the most probable failure scenarios, and to use dynamic rerouting for all other possible failure scenarios; (b) several survivability/availability grades may be defined so that not all VPs have 100%-restoration guarantee in all failure scenarios. This solution considers the fact that there are just some cases in which the communication system availability is critical (on-line banking, military communication, etc.), and that the communication system is generally used in parallel with other non-critical services. It is obvious that offering different levels of network availability has potential, not only to save network resources, but also to allow the operator to provide customers with services at their most appropriate cost. I have studied this problematic in two earlier papers [7,8].

Table I: comparison of preplanned and dynamic restoration

attribute	**preplanned**	**dynamic**
system complexity	lower	higher
network adaptability	difficult	easy
restoration speed	faster	slower
memory requirement	higher	lower

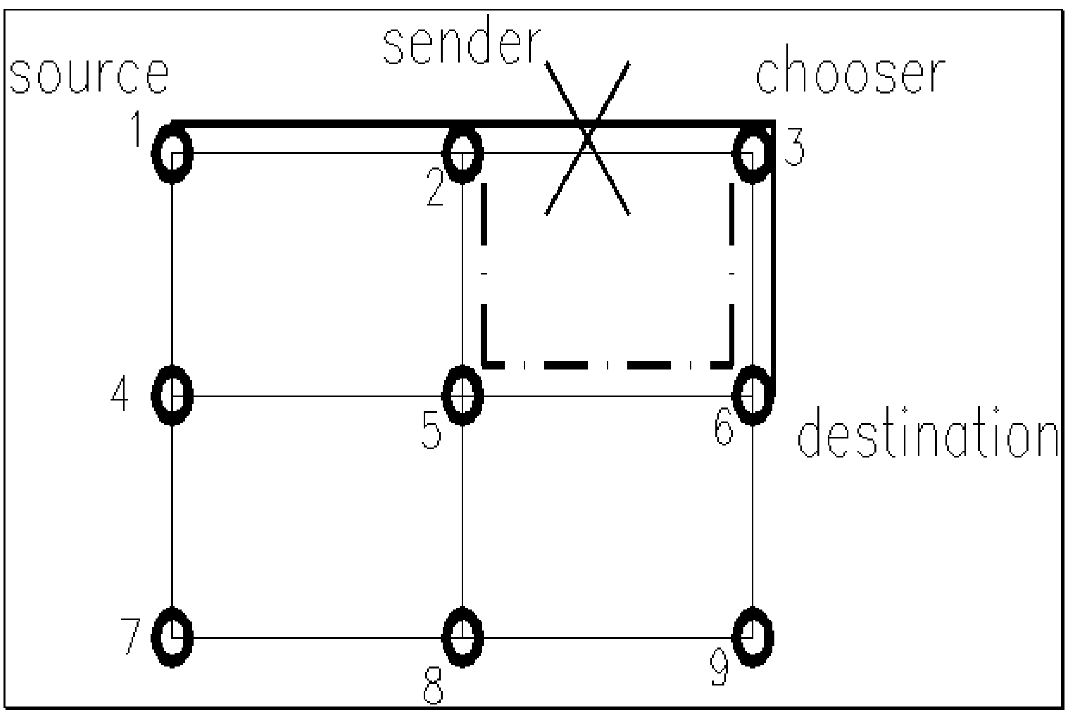

Fig.1: Local Rerouting

I now turn our attention to rerouting algorithms used for computing alternate routes. Generally speaking, there are tradeoffs between using restoration speed, processing/memory burden and restoration capacity. This study examines three rerouting strategies which handle these trade-off in different ways and degrees, as will be also outlined in the discussions summarized in sections VI and VII.

A. LOCAL REROUTING (LR)

When a link fails, in local rerouting (also called link restoration), all VPs carried on that link are rerouted locally around the failed link, without regard to their point of origin or their destination. This is illustrated in Fig.1. This rerouting algorithm is very simple to implement, but is achieved at the expense of higher restoration capacity. Since all VPs are processed locally, this scheme involves the fewest number of nodes in the rerouting process and, hence, can potentially be the fastest, in terms of restoration speed. Further, the memory requirement (in case of preplanned restoration) in this scheme is relatively small, since a node is affected (i.e. involved in the restoration process) only by failures of either incident links or links in the vicinity.

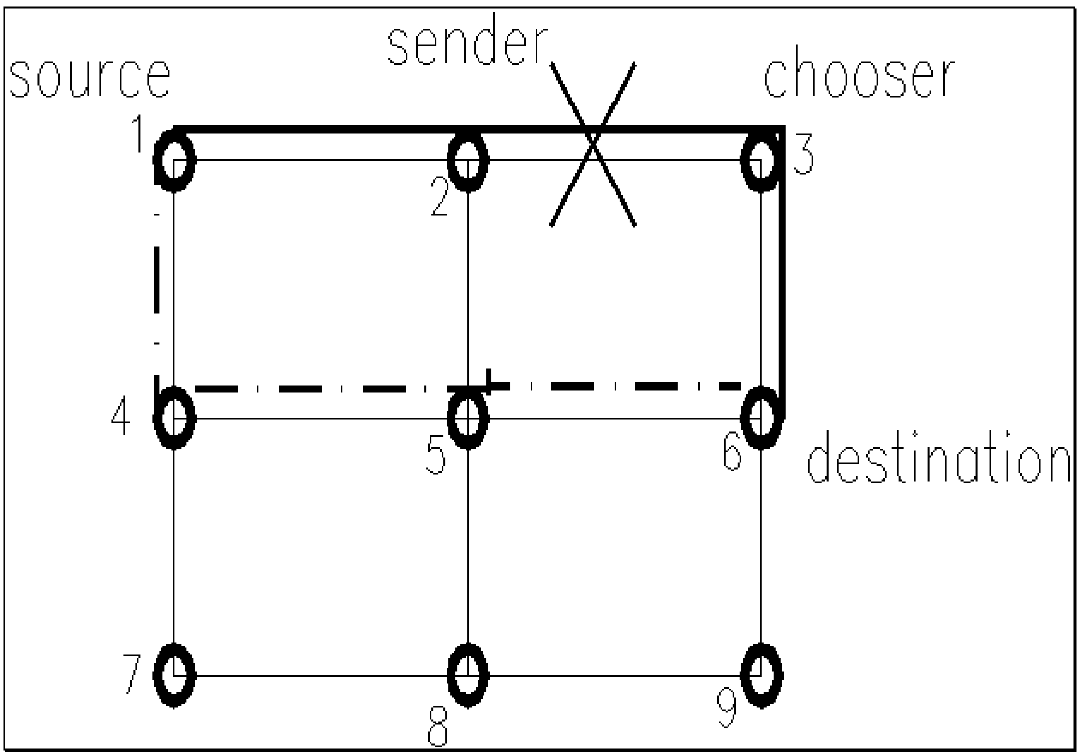

Fig.2: Source-Destination Rerouting

B. SOURCE-DESTINATION REROUTING (SDR)

In contrast to local rerouting, in source-based rerouting (also called path restoration), each VP affected by a link failure is processed and rerouted individually. Each affected VP is traced back to its source node, which reroutes the VP on an alternate path. This is illustrated in Fig.2. While the source-based scheme is optimal from the point of view of restoration capacity, the memory burden placed on the network nodes (in case of preplanned restoration), is much larger than in the case of local rerouting — as any given node could be affected by many more link failures. Further, the restoration time may be longer than in the case of local rerouting as the failure information has to propagate back to the end-nodes (source and destination) before restoration can be performed.

C. Local-Destination Rerouting (LDR)

This scheme is a combination of link and path restoration. VPs are rerouted locally and individually based on their destination. This allows the local node to determine the best (according to some metric such as distance, number of hops, etc.) alternate route for each VP, from itself to the VP node destination. This scheme is illustrated in Fig.3. The restoration speed, as well as the memory requirement of the local-destination scheme is intermediate between link and path restoration.

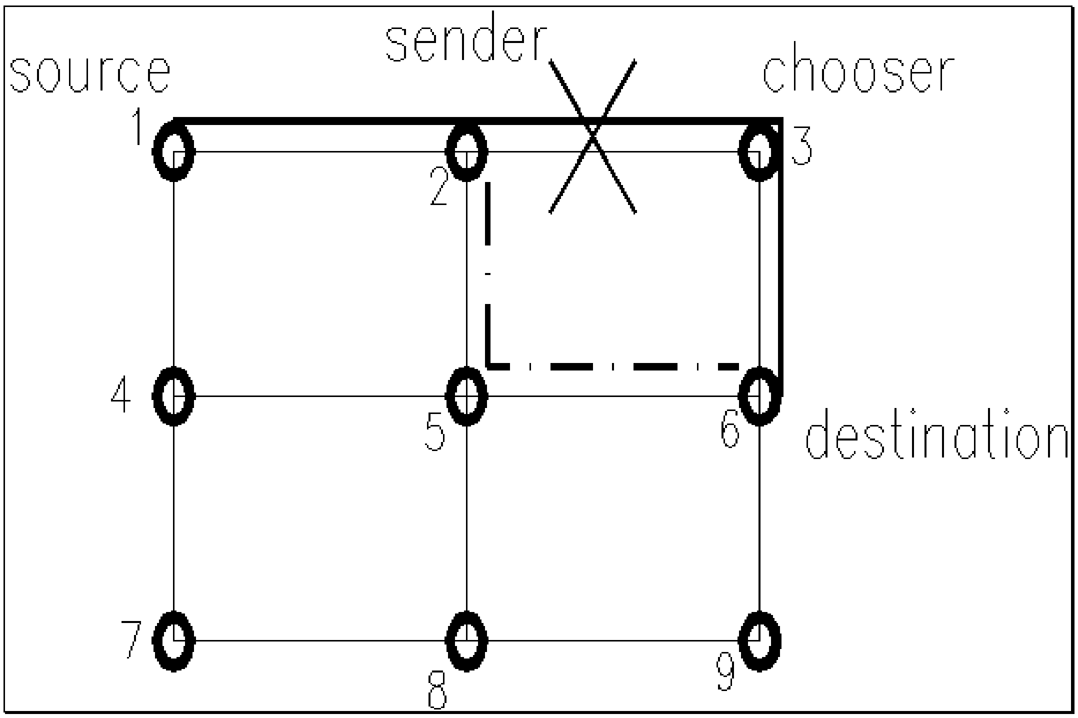

Fig.3. Local-Destination Rerouting

III. ATM/VP TRANSPORT NETWORK PERFORMANCE METRICS

A set of performance metrics that captures all critical aspects of the ATM/VP transport network integrity is needed. Considering restoration needs, I propose the following unified performance metrics framework: residual traffic volume; network protection ratio; network VPI capacity ratio; connectability; and restoration time.

A. RESIDUAL TRAFFIC VOLUME (RTV)

It is the traffic volume ratio that survives the disaster after restoration,

$$RTV = tv/tn \tag{1}$$

where tn is the traffic volume before failure, and tv the traffic volume after failure and rerouting.

B. NETWORK VPI CAPACITY RATIO (NVCR)

The VPI numbers are limited resources. For ATM/VP networks, this is the expression of the more general problem concerning all path layer technologies (PDH, SDH, ATM and even Optical), of limited maximum number of paths per one transmission link (see Fig.4).

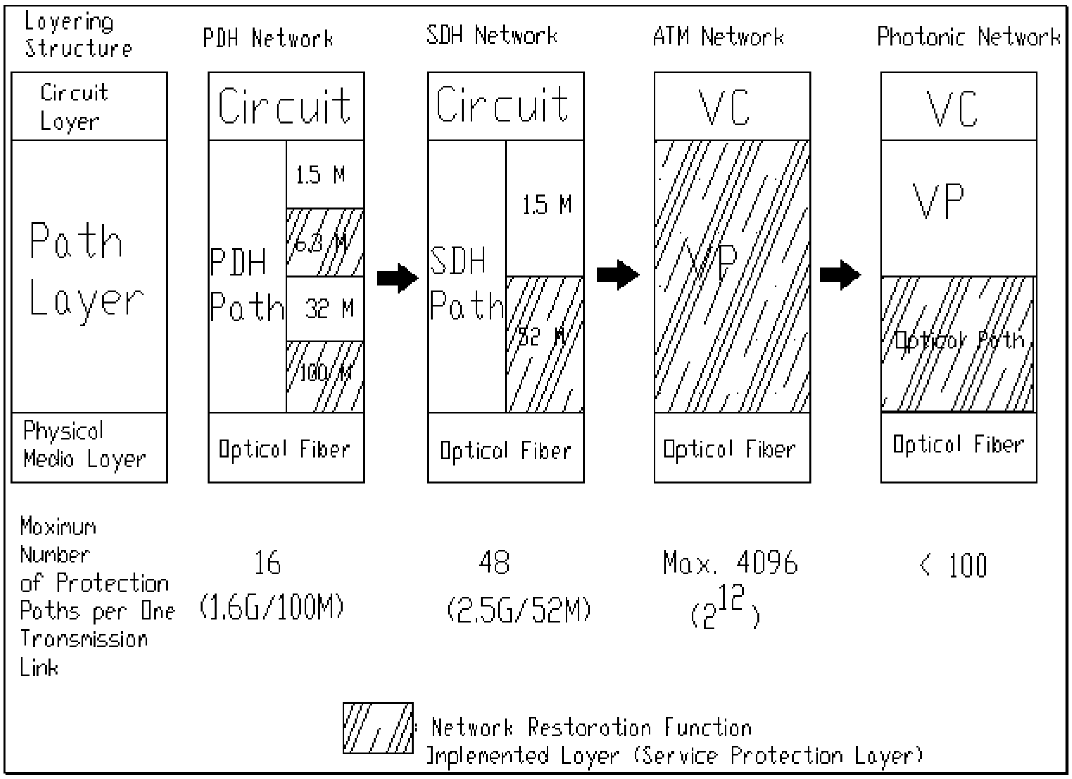

Fig. 4: Path layer evolution and allocation of network restoration function.

While pre-assigning alternate VPs (but also in case of dynamic alternate path computing), the VPI numbers resource can become a serious problem [3]. Zero bandwidth backup VPs do not consume the spare capacity of a link, however, as many VPI numbers as target VPs must be reserved. Thus, if a large number of backup VPs are accommodated on the same link, link capacity becomes unusable because of a shortage of VPI numbers. This problem can be assessed by calculating the VPI capacity ratio (NVCR). The NVCR is one index that indicates the possibility of insufficient VPI numbers on one link or in a network. The larger NVCR is, the worse.

$$\textit{Network VPI Capacity Ratio}:\ NVCR = \frac{\sum_i C_{1,i} / (C_{1,i} + C_{2,i})}{\sum_i N_{1,i} / (N_{1,i} + N_{2,i})} \tag{2}$$

where (see Fig.5): $C_{1,i}$ and $C_{2,i}$ are, respectively, occupied and spare capacity on link i; $N_{1,i}$ and $N_{2,i}$ are quantities, respectively, of occupied and for backup reserved VPI numbers on link i.

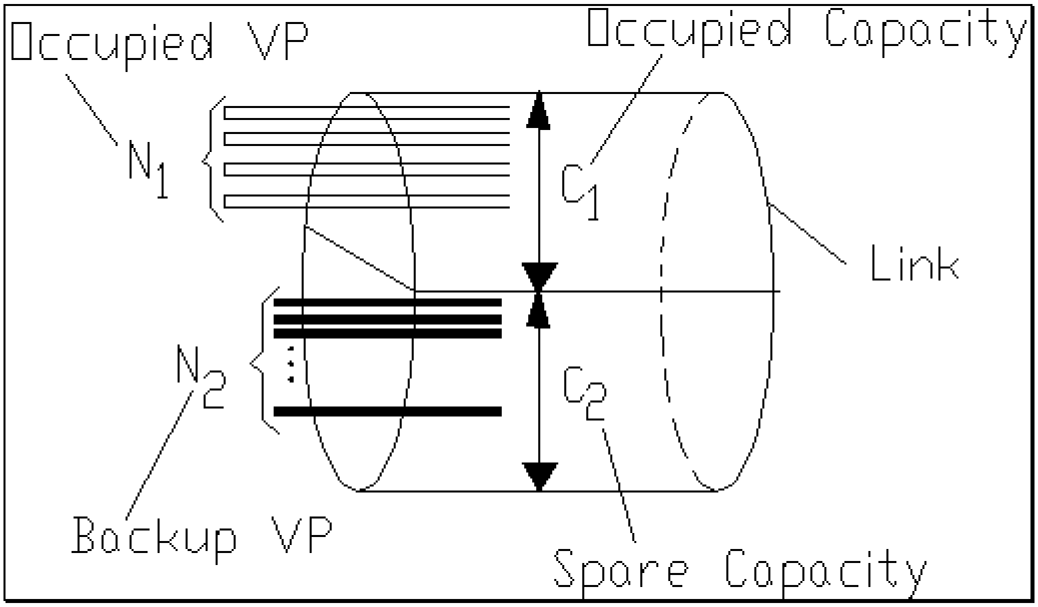

Fig.5: Distribution of VPI Numbers and Capacity on a link.

C. NETWORK PATH PROTECTION RATIO (NPR)

The path protection ratio expresses how much a path is protected. I will use an example to explain it. If the possible alternate paths (according to the rerouting algorithm) provide more capacity than the original path, the protection ratio is then 100%; however, if the alternate paths provide only 30% of the original path capacity, for example, the protection ratio is then 30%.

$$\textit{Path Protection Ratio}:\ PR_i = \frac{\min(w_i, k_i)}{w_i} \tag{3}$$

where: w_i, the capacity for VP i; k_i, the capacity of possible alternate (backup) VPs.

$$\textit{Network Protection Ratio}:\ NPR = \frac{\sum_i \min(w_i, k_i)}{\sum_i w_i} = \frac{\sum_i (PR_i . w_i)}{\sum_i w_i} \tag{4}$$

D. PATH LENGTH EFFICIENCY (PLE)

The path length efficiency shows the transport delay increase, consecutive to rerouting (restoration). I express here the path length in number of hops, but one may also use the physical length of the path (for example in km).

$$P_{LE} = \frac{\sum_i L(P_{ref}, i)}{\sum_i L(P_{back}, i)} \tag{5}$$

where: $L(P_{ref}, i)$, length of the original VP i; $L(P_{back}, i)$, length of the (shortest) alternate VP (assigned to the original VP i after restoration).

E. CONNECTABILTY

The dynamic reconfiguration capability of transport networks is a very important aspect of network integrity. It enhances the network's ability to react to unpredicted high load situations. I consider the notion of adapting the underlying transport network (ATM/VP) to the time-varying behavior of one or more logical networks (ATM/VC) which it supports. The transport network is subject to time-varying demands, which are a function of the offered load, and of the management techniques used in the logical network. When a node in the ATM/VC network cannot satisfy the aggregate level of offered load over a significant time period, it requests additional transport capacity from the transport network. The transport network reacts by attempting to synthesize a new transport path using a shortest path algorithm or providing more bandwidth to existing paths. It will then be possible to carry more traffic.

Connectability is a metric for characterizing the ability of the transport network to satisfy such requests. It estimates the probability that a path will exist between two nodes in a reconfigurable transport network, at some arbitrary point in the future. Connectability is used to measure the number and distribution of spare links in the transport network, and to constrain transport network management techniques which utilize these spares in order to enhance the service offered by the network by reconfiguring spare capacity to meet changing network traffic conditions. Theoretical studies show that an exact calculation of connectability is exponentially complex [9][10]. However, one may calculate an upper bound of this. This upper bound was found to be more responsive to changes in network configuration than the lower bound [11].

The following formula has been derived for the upper bound of connectability:

$$C(s,t,p) \leq (1-(1-p)^k)^{\min(L_i)} \tag{6}$$

where k shortest paths are found between the nodes s and t; p is the probability that an edge is free for use[1]; L_i is the length of each path (found), then the maximum number of edge-disjoint cut sets equals the length of the shortest path, min(L_i). The number of edges in each cut set is the number of paths found k.

F. RESTORATION TIME

This is the conventional metric used in many studies [3,12,13]. Nevertheless, let us review the essentials. Restoration time refers to the time required by a restoration algorithm to complete execution. Since it is desirable to accomplish restoration as quickly as possible in order to avoid call dropping, this is

---

[1] p is the probability of (link) utilisation rather than failure. For simplicity, we assume it here to be the same for all links of the network and equals 1/2. See [11] for a deeper discussion.

an important metric. Ideally, an algorithm must achieve full possible restoration in less than two seconds [17].

The restoration time depends on the following: (a) the overall organization of restoration messages' flooding; (b) the restoration messages' format and structure; (c) the message propagation time; (d) the node processing time; and (e) the path activation time.

In studies evaluating restoration time, the modeling of network elements has been brought down to three parameters: (a) cross-connect delay, the time needed to set up a cross-connect point in a node; (b) internal communication delay, the time needed to pass alarms and restoration messages through the control architecture of a node; and (c) processing delay, the actual CPU time needed to process an algorithm code.

Thus, the restoration time depends on both restoration protocol details, and node processing power. Because of this strong implementation dependence, and because several studies have already been conducted (on it), restoration time will not be considered in this study.

IV. GENERAL PROCEDURE FOR FINDING SURVIVABILITY FUNCTIONS

A. GENERALITIES

The concept of survivability is similar to the concept of reliability, only more general. Several studies to evaluate reliability, use routing models [18-20]. In Addition, recent works on network survivability [21-22], provide us with a good background to formulate a general procedure for analysis of ATM transport network survivability performance based on rerouting models. My formulation both includes and extends the existing definitions for network survivability measurement. I propose a conditional approach to survivability analysis, which measures the network performance after failure has occurred. This approach may either use probabilistic weighting of the resulting network restoral and/or repair after the failure, or use a deterministic analysis of these states. This method can be used to evaluate different restoral, repair, or preventive methods, depending on which types of comparison characteristics are critical. In this study, the performance of rerouting strategies is evaluated. Because of the computational complexity, automated computation of survivability functions is necessary.

The reason for concentrating on survivability function rather than on a single survivability value (as in many previous studies) is that a number of different, interesting quantities can be derived from the function, each capturing a particular characteristic. These particular characteristics will henceforth be called survivability parameters (i.e. expected survivability; worst-case survivability; r-percentile survivability; probability of zero survivability). While considering a single survivability value, one usually takes the 'worst-case' single or isolated failure, as has been exemplary done in references [4] and [23]. The network performance metric (generally the percental residual traffic volume) is then evaluated in this most critical failure scenario. Its drawbacks are, however, the following: (a) nothing is said about the probability occurrence of this worst-case failure; (b) for certain metrics, and in any given topology, the worst-case failure may not necessarily be evident; and (c) nothing is said about the distribution of the survivability values. The survivability function, which is in essence a distribution of the survivability, overcomes the above shortcomings, by providing all information regarding survivability; even the worst-case survivability may be extracted from it.

B. METHODOLOGY

1) Specify disaster type to be studied: this indicates which types of failures may occur (link failure, node failure, etc.).
2) Define network performance metric Z of interest. For an ATM transport network, a performance metrics framework has been discussed in section III.
3) Consider the kind of failures which can occur in consideration of the disaster type being considered (e.g. link failure, node failure, etc.); then partially enumerate the most probable events. This is an easy task for network carriers, since they have failure statistics for many years of experience at their disposal [14,15]. An event is a disaster. It may cause one or several link failures, and/or one or several node failures.
4) Determine or assign a probability Pm of each event m.
5) For each event m, carry out the rerouting, and calculate the survivability value Zm for the performance metric Z. From this we derive the survivability function:

$$P[Z = z] = \sum_{m:Zm=z} Pm \tag{7}$$

The basis for a linear combination of the effects of different events in equation (7) comes from the assumption that the events are independent, since there is any evident reason to assume a priori the contrary.

6) Derive the following survivability parameters from the survivability function:

a) expected survivability:

$$E\{Z\} = \sum_{z} z.P[Z = z] \tag{8}$$

b) worst-case survivability, that I also call 'absolute worst-case' survivability:

$$z^{*} = \min_{P[Z=z]>0} z \tag{9}$$

c) r-percentile survivability, that I also call 'relative or probabilistic worst-case' survivability:

$$z_r = \max_{P[Z\le z]\le r/100} z \tag{10}$$

d) probability of zero survivability Po:

$$Po = P[Z = 0] \tag{11}$$

NOTE

The survivability function is a probabilistic performance measure (of the survivability) for a single metric. However, it is conceivable to calculate a performance vector $G = \begin{bmatrix} G_1 \\ G_2 \\ G_3 \end{bmatrix}$ which integrates all performance metrics.

The components of G could be calculated as follows:

$$\begin{aligned} G_1 &= \alpha_1 E\{Z_1\} + \bullet\bullet\bullet + \gamma_1 E\{Z_n\} \\ G_2 &= \alpha_2 z_1^* \quad + \bullet\bullet\bullet + \gamma_2 z_n^* \\ G_3 &= \alpha_3 z_{r_1} \quad + \bullet\bullet\bullet + \gamma_3 z_{r_n} \end{aligned} \tag{12}$$

where Z_i $(i = 1,\ldots,n)$ are the different performance metrics; $\alpha_i,\ldots,\gamma_i$ are weighting coefficients which express the importance of each metric. A similar way of thinking has been presented in reference [25]. The weighting coefficients should take possible correlations between different metrics into account.

I do not judge it necessary to follow this aspect in more depth here, since a separate analysis of the different metrics already provides us with the essential background on rerouting strategies' performance.

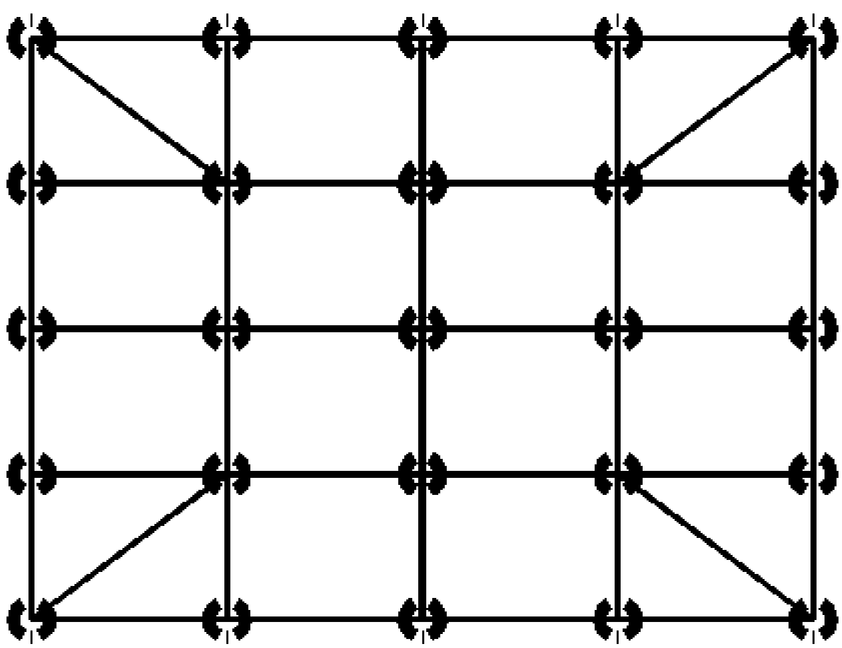

Fig.6: 25 Node Mesh Network

V. SIMULATION MODEL

The focus here is on restoration within the transport network. Statistical multiplexing is done inside the access network, the VPs in the (backbone) infrastructure typically serve as digital pipes with little or no bandwidth fluctuation.

A. MESH TOPOLOGY JUSTIFICATION

As explained in references [16], the following line of reasoning pleads for a mesh topology for the backbone network. A basic motivation for mesh restoration is to exploit the physical route diversity of real networks for capacity efficient survivability. For full restoration with APS (Automatic Protection Systems) 100% redundancy is required. Rings require 100% or more redundancy. Mesh restorable networks, on the other hand, are based on generalized rerouting over all diverse routes of the network. Spare capacity on each span/link contributes to the survivability of many other spans/links. Such networks are called mesh-restorable, not to imply that the network is a full mesh, but to reflect the routing mechanism's ability to exploit a mesh-like topology.

Mesh networks not only achieve lower redundancy levels when compared with rings, but they enable a working base that is already an efficient, shortest path placement for working capacity. The total capacity (working plus spare) for mesh networks is, therefore, generally several times less than that of ring-based networks for the total service base [17]. This does not necessarily mean that mesh networks are more cost-efficient — particularly considering current electrical interfaced crossconnect machines (for example, DS3 interface costs are high). The capacity difference, however, is of such magnitude (three-to six-fold in typical networks [17]), that it signals the opportunity for developers to convert the intrinsic efficiency of mesh networking into cost savings for telecommunication companies. The development of integrated optical terminations on the (digital) crossconnect systems will be key for this conversion.

B. REDUNDANT CAPACITY ASSIGNMENT

The different rerouting schemes have different redundant capacity requirements. In order to quantify the tradeoffs for the different rerouting schemes, a simulation study on a 25 node mesh network has been conducted. The mesh network studied is that of Fig.6.

Normally, it is easier to implement effective failure prevention measures for nodes (and other (geographically) localized network elements, in general) than for cables. This would include fire control systems, and fault-tolerant architectures, utilizing high-reliability components that ensure that internal failures are circumvented in milliseconds, with minimal loss of service. Therefore, only link failures are considered here.

Supposing that the disaster can cause at most three link failures, backup paths for all failure scenarios are computed, and VPI numbers are reserved for each. Then, redundant capacity is only placed to ensure 100% restoration in any one link-failure case. The dimensioning basis is so 'to place enough spare capacity that 100%-restoration for any one link-failure scenario will be guaranteed.

Two different point-to-point traffic demands are considered here: in one case the traffic demand of 10 units (or bandwidth units) exists between all node pairs, while in the other the traffic demand is asymmetric, in that only 60 % of all node pairs have traffic between them. Each unit demand can be thought of as 1 DS3 for convenience. Details on algorithms used for designing the alternate (backup) routes and spare resource distribution can be found in works from Kawamura et al., Anderson et al., and Chujo et al. [3,5,13]. Table II summarizes the redundant capacities needed for the 25-node mesh network.

Table II: Redundant capacity required. Dimensioning basis: 'to ensure 100%-restoration for any one link-failure scenario'.

rerouting strategy	asymmetric traffic demand	symmetric traffic demand
local rerouting	73%	75%
local-destination rerouting	69%	59%
source-destination rerouting	54%	41%

One can see also that the simplicity of local rerouting is achieved at the expense of higher restoration capacity. One of the main reasons for this is the so-called (and well known) 'backhauling effect'. In fact, since redundant capacity has to be assigned on each new hop, this results in unnecessary assignment of redundant capacity. The backhauling experienced in local rerouting is avoided by local-destination rerouting. Thus this scheme requires lower restoration capacity than the local rerouting, while retaining the flavor of local rerouting. Except in the case of local rerouting, a symmetric traffic demand conducts to a lower percental redundant capacity requirement.

C. FINDING SURVIVABILITY FUNCTIONS

The network considered here consists of 25 nodes and 44 links (Fig.6). Further, it is supposed that the disaster will cause only link failures. As already explained, the reason for this is that it is easier and more efficient to protect and take preventive measures for nodes. Nevertheless, preventive measures by cable construction and placement are also in rapid progress [23,24]. Considering this, it was assumed that a maximal of three (simultaneous) link failures may be caused from the disaster.

To determine Pm, the probability of each event m, one assumes that link failures are independent and that the probability that a link i fails is only proportional to its length l_i, that is

$$P[\text{link} \;\; \text{i fails}] = \varepsilon . l_i = \rho_i \tag{12}$$

whereby $o < \rho_i < 1$, and $o < \varepsilon < 1$. It can also be assumed that all links have almost the same length[2]. So, ρ_i is the same for all links. The link failure considered here is a 'physical link failure' (or span failure). A physical link connects two nodes in the topology of Fig.6, and its failure involves multiple ATM VP paths.

The probability of no link failure is, then

$$\text{P[no link failure]} = \prod_i (1-\rho_i) \tag{13}$$

With ρ_i determined, the probabilities of single, double, and triple link failures are as follows:

$$\text{P[only link i fails]} = \rho_i \prod_{n \neq i} (1-\rho_n) \tag{14}$$

[2] As may be seen on Fig.6, all physical links have the same length, except four of them.

$$P[\text{only link i and j fail}] = \rho_i \rho_j \prod_{n \neq i,j} (1 - \rho_n) \qquad (15)$$

$$P[\text{only link i, j and k fail}] = \rho_i \rho_j \rho_k \prod_{n \neq i,j,k} (1 - \rho_n) \qquad (16)$$

Concerning rerouting, the network is formalized by a graph. The physical network with several VP's (paths) is considered as a multicommodity network (where the VPs are the commodities). How the graph is manipulated is, next, briefly explained. The fundamental tasks performed are the following:

- Finding alternate paths: the possible restoration paths are identified according to the rerouting mechanism chosen. All paths which cannot be used in restoration are identified.
- Resolving spare capacity contention: since not all of the candidate restoration paths are disjoint, there is some competition among them for available spare capacity. Such contention is resolved by an algorithm. The method used to resolve spare capacity contention can have critical impact on the level of restoration achieved.
- Selecting restoration paths: this refers to the way to select which candidate paths are to be used for restoration. One basic selection criterion used here is the „shortest path“ (relative to hops number). However, it is possible to use other criteria, such as the bandwidth of the path. But, since all paths in the model have been assumed to have the same bandwidth, this is then irrelevant. Besides, since the selection process involves contention for spare capacity, there is often (particularly in multiple link failure case) a strong relationship between selection of restoration paths and resolution of spare channel contention. A very simple contention resolution mechanism is used here, since one takes one of the contenting paths randomly.

Based on the values of Zm and Pm (derived as explained in section IV.B) above, survivability functions are calculated for all proposed performance metrics. By reading and examining the following results, one must have in mind the required redundant capacities (summarized in Table II) for the different rerouting schemes.

VI SIMULATION RESULTS

A. NOTATIONS & ACRONYM

The following abbreviations are used in this section are several times:

lpf	link failure probability
Sym	symmetric traffic demand
Asym	asymmetric traffic demand
LR	local rerouting
SDR	source-destination rerouting
LDR	local-destination rerouting

In addition, when not explicitly indicated, a link failure probability (lfp) of 0.025 is used for the results presented below.

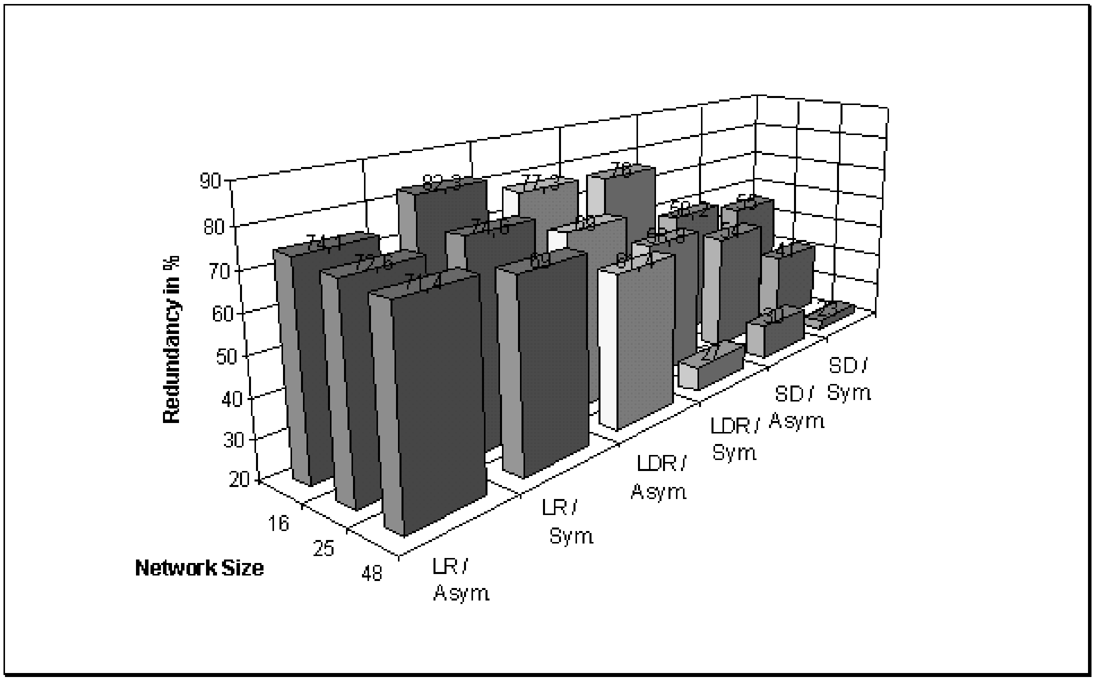

Fig.7: Network Size Influence on the Percental Redundancy

B. NETWORK SIZE INFLUENCE ON THE PERCENTAL REDUNDANCY

The percental redundancy decreases with the network size for all rerouting strategies, and for both traffic demands (see Fig.7). The decrease is more pronounced by SDR, and in case of symmetric traffic demand for all rerouting schemes. The topologies used are similar to the one of Fig.6 (of the Manhattan-city type).

Noteworthy is the fact that, with LR in small size networks, the symmetric traffic demand requires more percental redundancy than the asymmetric one. This is not the case for the other rerouting strategies, where we observe, rather, the contrary. In terms of percental redundancy requirements in small-scale networks, LDR is not at all better than LR.

A large scale network requires indeed less percental redundancy, but it is difficult to manage. So, a tradeoff must be found. This problem is generally solved by building clusters, i.e. dividing the large-scale network into smaller subnetworks which are easier to manage, with an acceptable complexity and processing power. It is a kind of decentralization of control, since each subnetwork may be assigned a unique controller.

Larger network sizes may be tolerable by LR as this reduces (high) values of the percental redundancy required, and because easy network handling may still be possible because of the (very) low complexity of this rerouting scheme(*). In contrast, a very large network size may be prohibitive for SDR since the redundancy savings do not necessarily compensate the negative effects and problems resulting from the combination of a very high (management and processing) complexity and higher restoration times.

(*) $O(n^2)$ is the complexity of finding a single shortest replacement route in a simple graph. It can be solved with Dijkstra's shortest path algorithm. Note that path restoration (SDR) has against it a complexity of $O(n^4)$ [17].

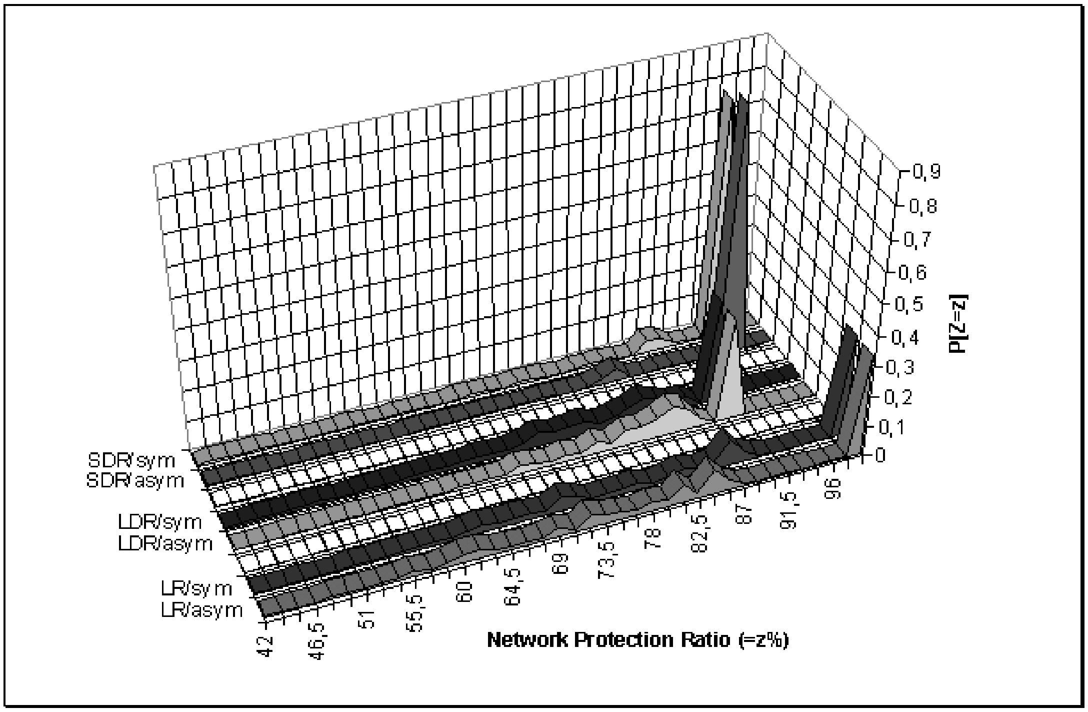

Fig.8: Network Protection Ratio Survivability Functions.

B. NETWORK PROTECTION RATIO (NPR) SURVIVABILITY

When one thoroughly inspects the NPR survivability functions presented in Fig. 8, one realizes that they seem to consist in two parts: (a) an important component, constituted by the highest impulse on the right-hand side. This is certainly the contribution of the most probable failure events. It could constitute a first approximation of the NPR survivability. (b) a second part, constituted by a group of several short impulses, on the left-hand side of the ordinate axis. These are likely to be the contributions of the less probable failure events.

Now that the general structure of the survivability functions is understood, let us comment and compare them. The SDR survivability functions distinguish themselves in that their first components are very important, when compared to their second ones. Overall, this conducts to a very good network protection ratio. The NPR survivability parameters presented in Fig.9 confirm this fact.

Although the first component impulses of the LR survivability functions are well positioned on the left-side border of the ordinate axis, their weights are unfortunately low. This explains the modest survivability performances, as depicted on Fig.9. In addition, LDR reaches rather moderate performance values. The power of SDR is also noted by the fact that it has the lowest redundancy at its disposal.

Further, one does not notice any substantial influence of the traffic demand, except a relatively small percental redundancy requirement difference (this difference is but more sensible by SDR, where it reaches 13 %, see Table II).

According to the definition of NPR, a good NPR should conduct to a better residual traffic volume (RTV) survivability. The following section will show whether these expectations are fulfilled.

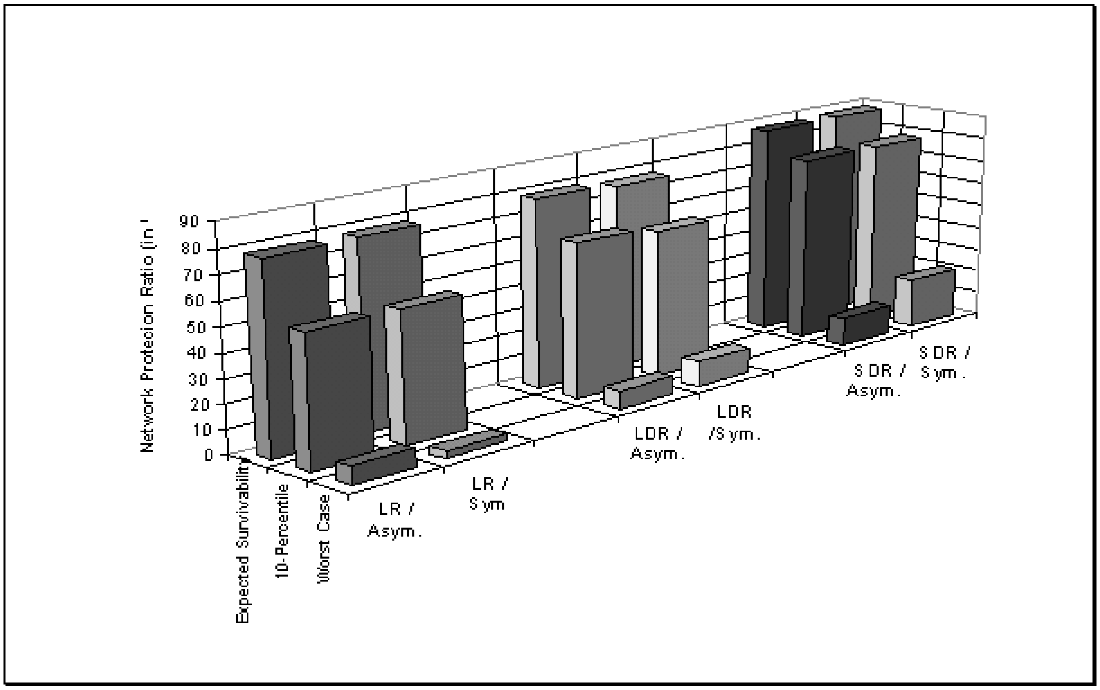

Fig.9: Network Protection Ratio Survivability Parameters.

C. RESIDUAL TRAFFIC (RTV) VOLUME

As seen in Fig.10 the survivability values are well-concentrated near the right-side extremity of the ordinate axis. One cannot notice any substantial difference between the different rerouting schemes, referring to both survivability functions (Fig.10) and survivability parameters (Fig.11). A foundation (or explanation) of this behavior is the fact that a 'common dimensioning basis' was taken for all rerouting schemes in both traffic demand scenarios.

In fact, the redundancy dimensioning was stated as: „providing enough redundancy in order to guarantee 100%-restoration for any one link-failure scenario". An important realization to take from Fig.10 and Fig.11 is that the three rerouting schemes have the same RTV survivability performance (expected survivability, 10-percentile survivability), provided that the dimensioning basis is the same.

The only obvious difference is observed by the „RTV worst-case survivability". The values here are still (almost) the same in case of symmetric traffic demand. One observes, however, relatively lower values for asymmetric traffic demand, whereby the highest difference occurs by LR (see Fig.11). However, this difference may be neglected in a first approximation (regarding also the very low probability occurrence), so that (it can be stated that), "*the three rerouting schemes conduct to the same RTV survivability performance, provided that the dimensioning basis is the same*".

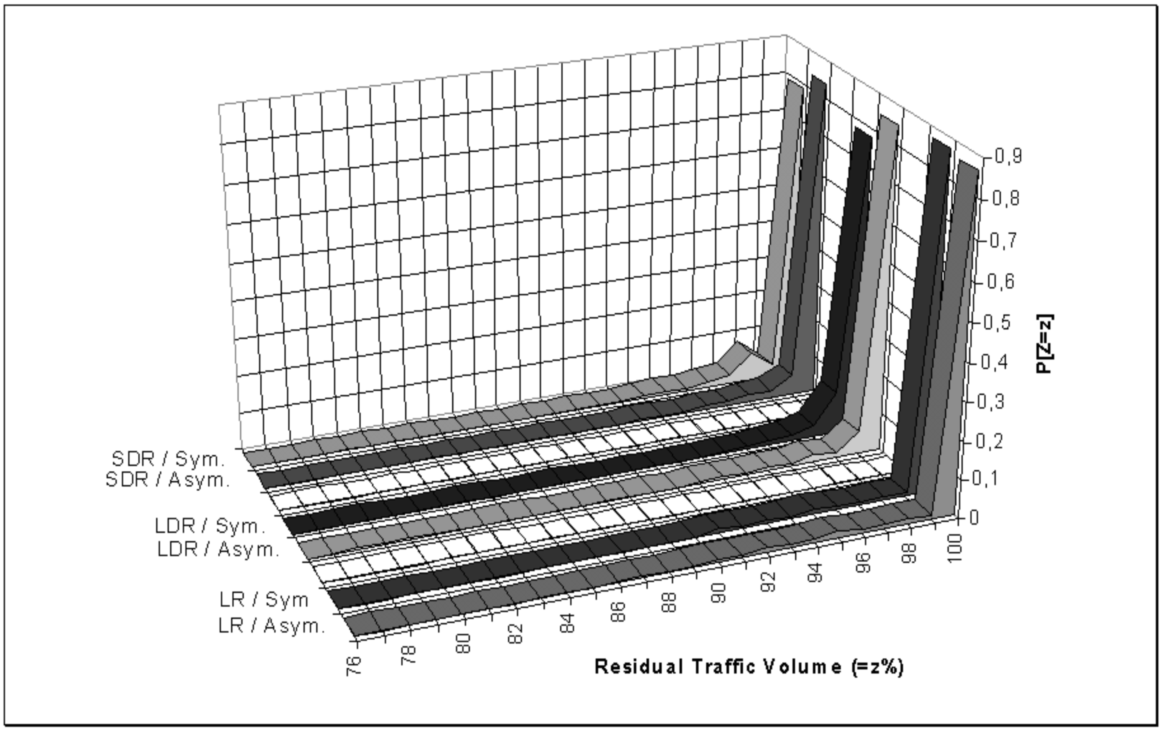

Fig.10: Residual Traffic Volume Survivability Functions.

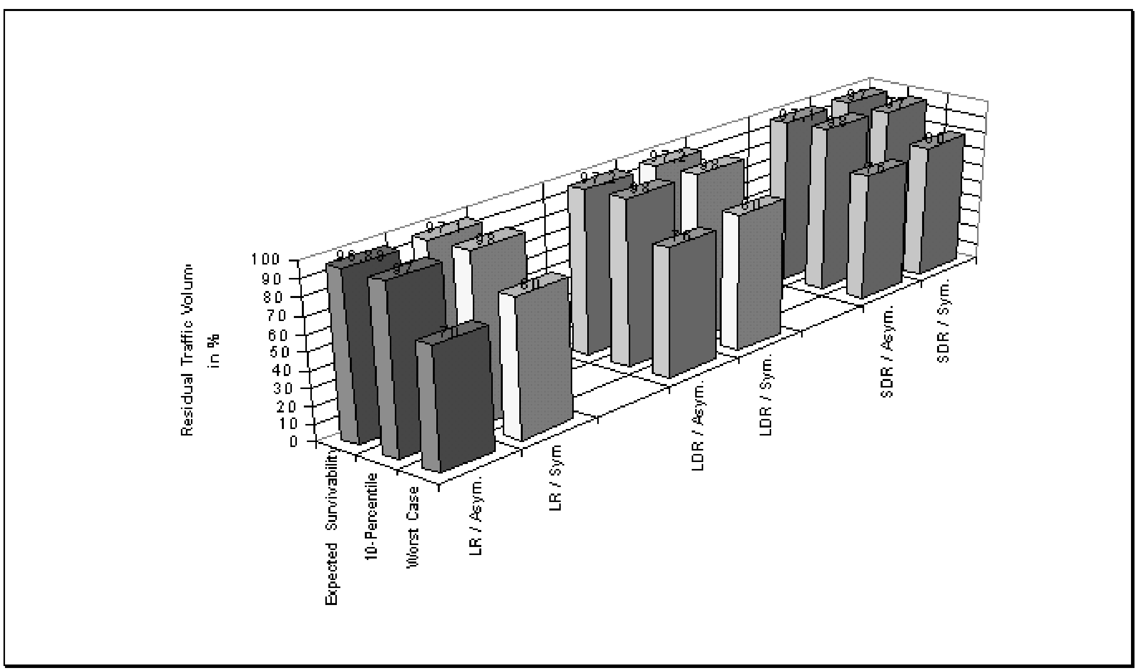

Fig. 11: Residual Traffic Volume Survivability Parameters.

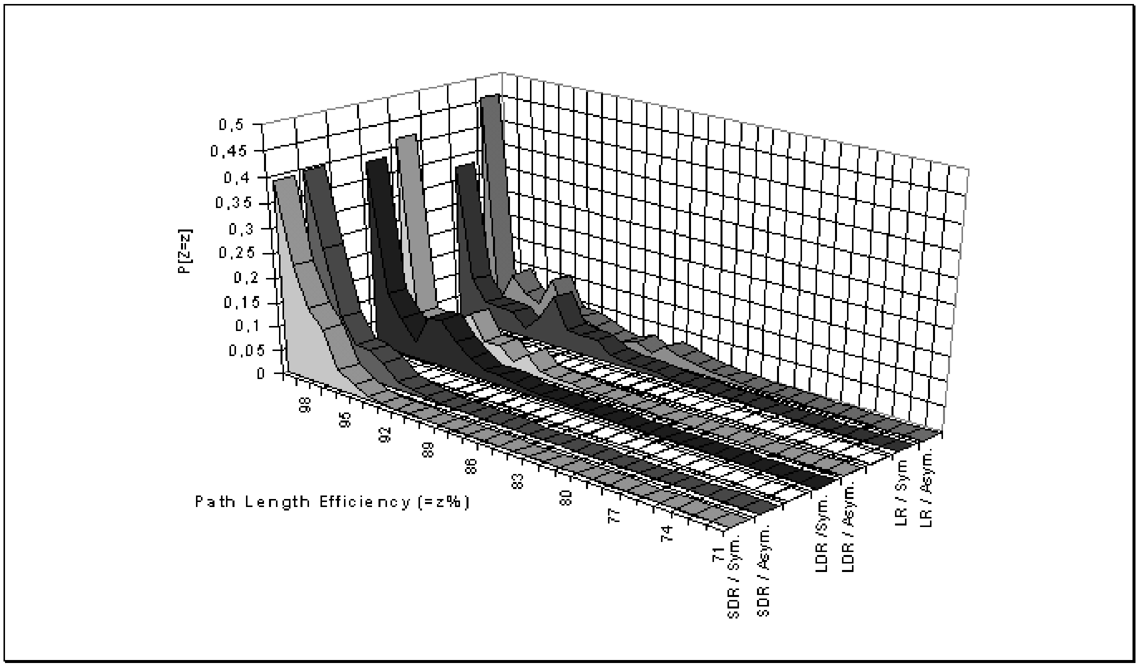

Fig.12: Path Length Efficiency Survivability Functions.

D. PATH LENGTH EFFICIENCY (PLE) SURVIVABILITY

The survivability functions presented in Fig.12 behave approximately as exponentially decreasing ones, whereby SDR experiences the highest decrease in speed. This results in a smallest transport delay increase for this rerouting scheme, as confirmed in Fig.13. The high path length increase caused by LR results in an inefficient use of the available redundant capacity.

One of the main teachings of these results is that local rerouting is not the appropriate rerouting scheme for networks (e.g. satellite networks) and applications (real-time applications), where delay and/or capacity efficiency are critical.

E. NETWORK VPI CAPACITY RTIO (NVCR) SURVIVABILITY

As explained earlier, for this goodness metric, the higher the values are, the worst. Fig.14 and Fig.15 reveal to us that SDR distinguishes itself through its more efficient use of VPI resources. LR is worst and LDR performs rather moderately. A symmetric demand results in a somehow more efficient VPI use.

In general, VPI numbers become critical when the path (VPs) numbers (in the network) are high. And since a large scale network results in a larger number of paths (because of the larger number of point-to-point connections), the obtained results (Fig.14 and Fig.15) signal that LR implementation limits the network size (or better, the subnetwork size) respectively path numbers. This is because of the danger of a shortage of VPI numbers.

Although for LR a large (sub)network results in smaller redundancy requirements (see Section VI.B), the NVCR survivability values indicate that the (sub)network size cannot be large at will (while LR implemented) because of the VPI numbers bottleneck.

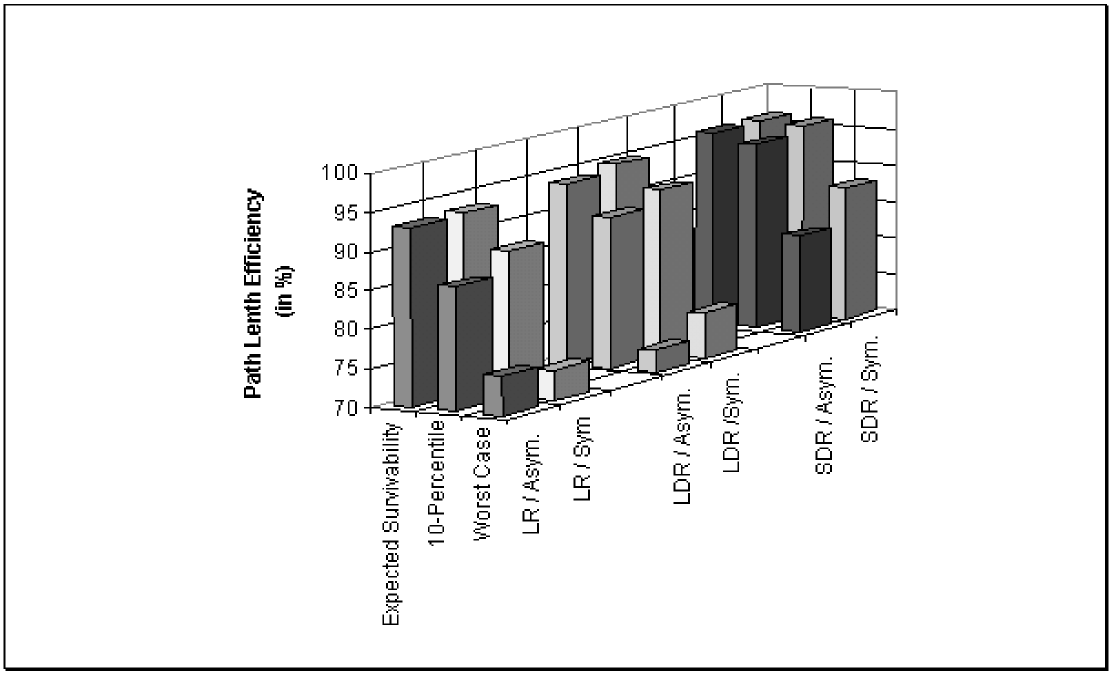

Fig.13: Path Length Efficiency Survivability Parameters.

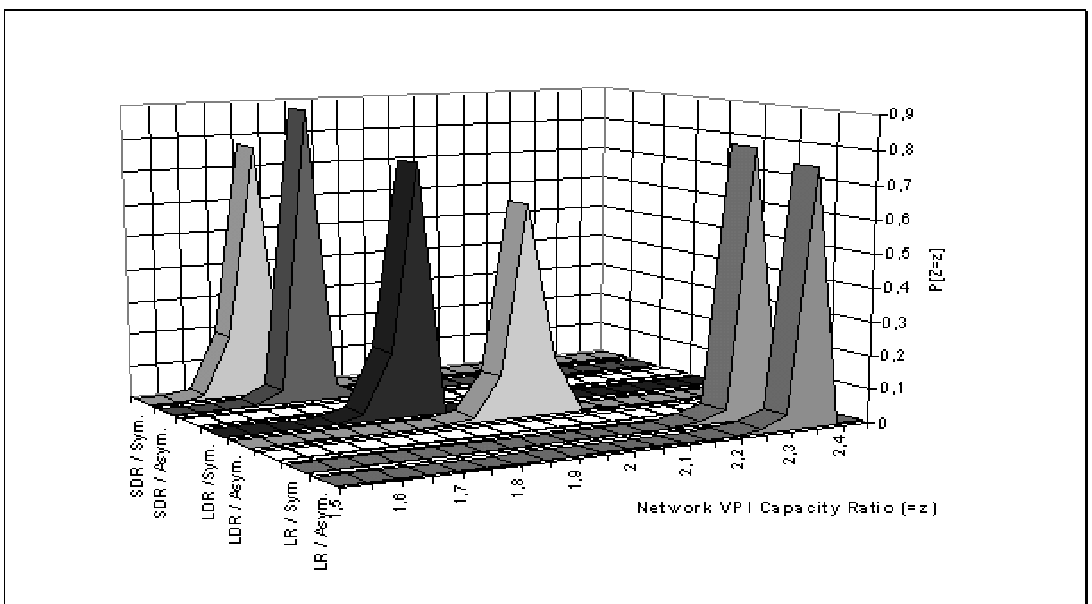

Fig.14: Network VPI Capacity Ratio Survivability Functions.

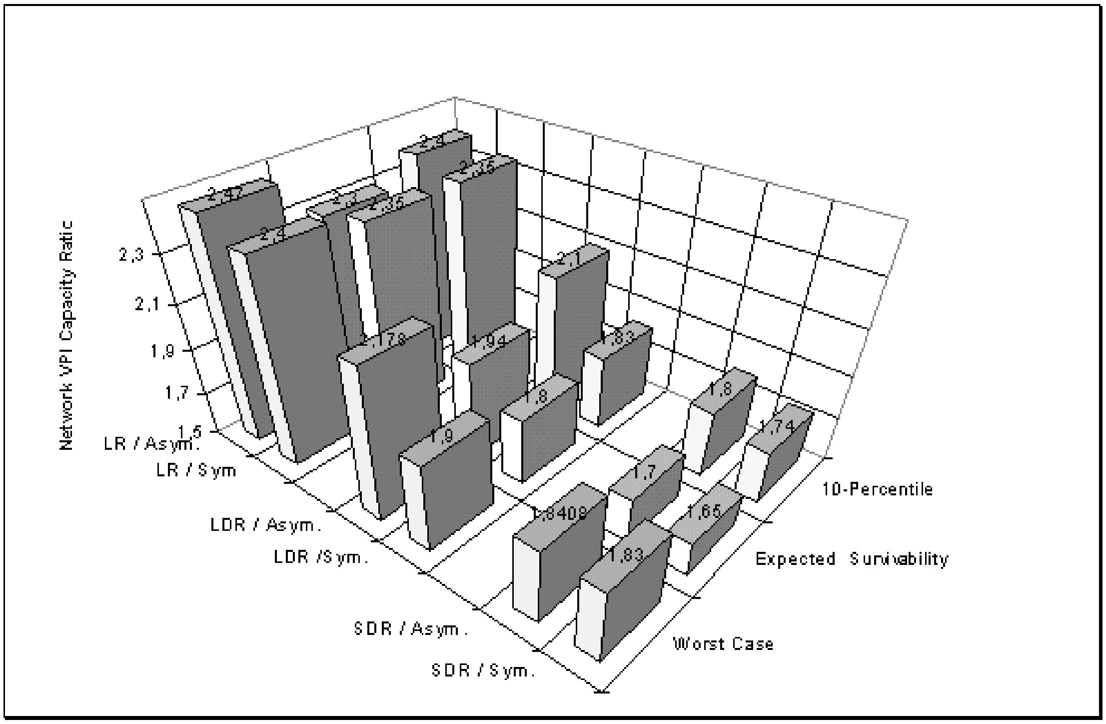

Fig.15: Network VPI Capacity Ratio Survivability Parameters.

F. CONNECTABILITY SURVIVABILITY

For clarity (of the figure), I have chosen to plot the survivability functions for only one traffic demand scenario (Fig.16). The first statement about Fig.16 and Fig.17 is that LDR performs the better connectability. This means that the way LDR places the different backup paths results in the best dynamic reconfiguration capability of the transport network.

The expected survivability and the worst-case survivability are quite insensitive to traffic demand type. The 10-percentile survivability against it is very sensitive. An asymmetric traffic demand is obviously untoward relative to connectability.

The survivability function for LR (has a particular form) is characterized by quantum leaps between the different impulses. This strange behavior results from the fact that local rerouted paths (constitute or) place an irregularity in the path system; the connectability degradation is not more smooth because of this. The SDR connectability performance is the worst, but one cannot complain, considering its modest redundant capacity requirements.

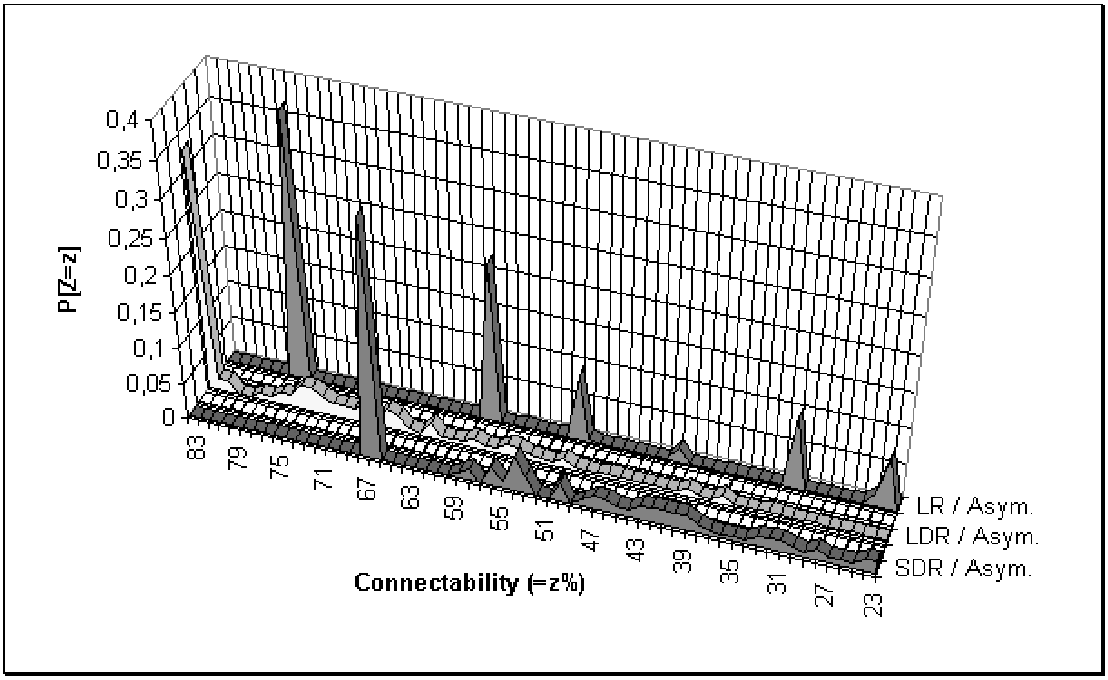

Fig. 16: Connectability Survivability Functions.

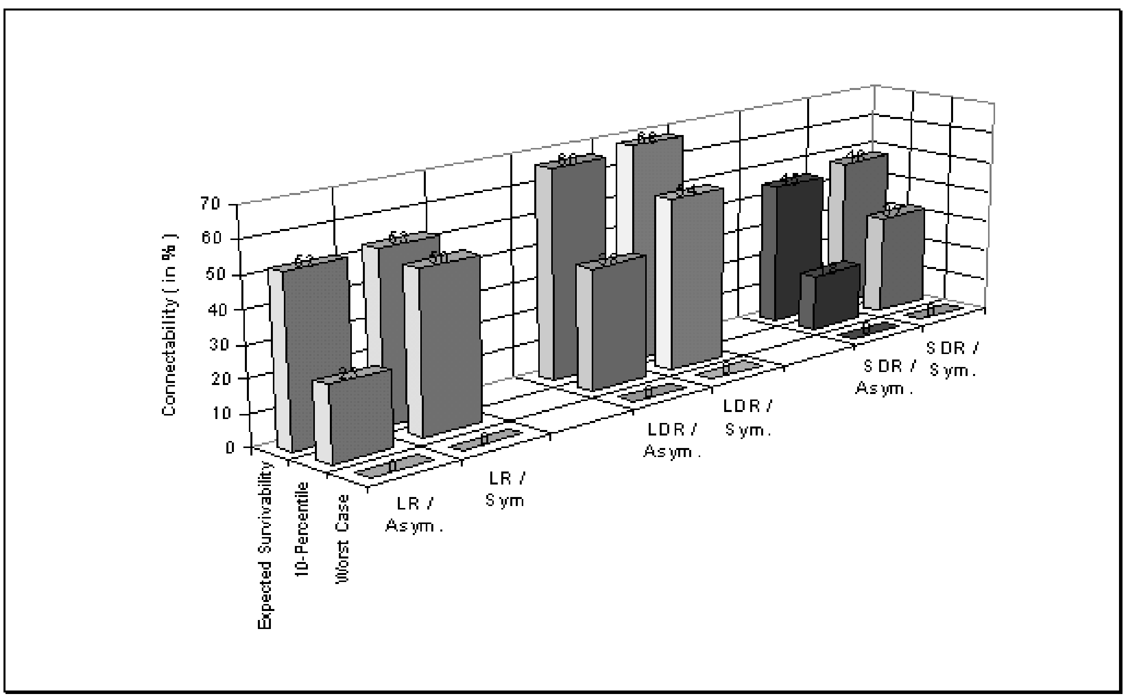

Fig.17: Connectability Survivability Parameters.

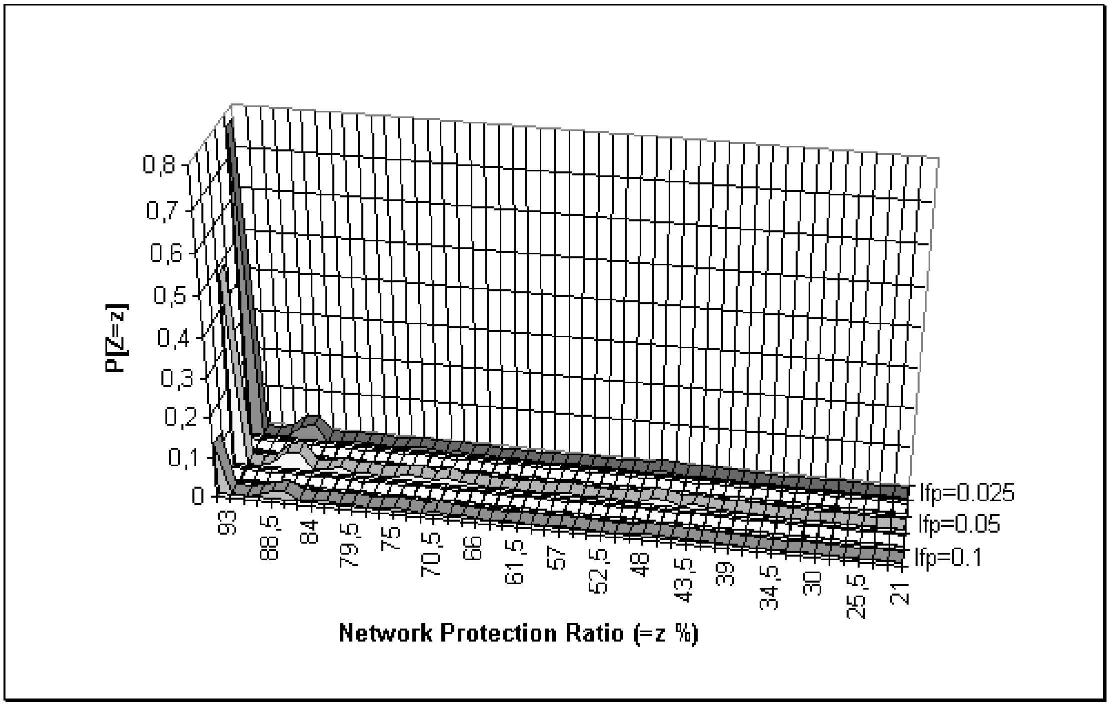

Fig. 18: Disaster Extent Influence on Network Protection Ratio Survivability functions; Source-Destination Rerouting; Symmetric Traffic Demand.

G. DISASTER EXTENT INFLUENCE ON SURVIVABILITY

Fig.18-21 are just a sample of survivability functions and parameters which aim to show disaster extent effect on network integrity. In practice, the link failure probability ρ is set to reflect the extent of damage expected from the disaster, i.e. the higher the extent of a disaster is, the higher ρ. Following values of ρ are considered: 0.025, 0.05 and 0.1. Besides, Fig.18-20 reveal that the effect of disaster extent on survivability functions is not at all linear. The throttling is very strong by higher values (of the survivability), while it is rather small and quite insignificant by lower ones. This behavior results in an important decrease of the survivability mean values (Fig.21).

The worst-case survivability, for its part, is absolute independent of the extent of disasters. Further, the effect on the 10-percentile survivability is rather small, despite a moderate or quite slight non linearity. One can assume then, as a first approximation, that the 10-percentile survivability also remains constant, independently of the extent of disaster. In this way, one obtains two intrinsic 'survivability quantities', which may be extracted from any survivability function (without worrying about the (precise) disaster extent considered): the worst-case survivability, and the 10-percentile survivability.

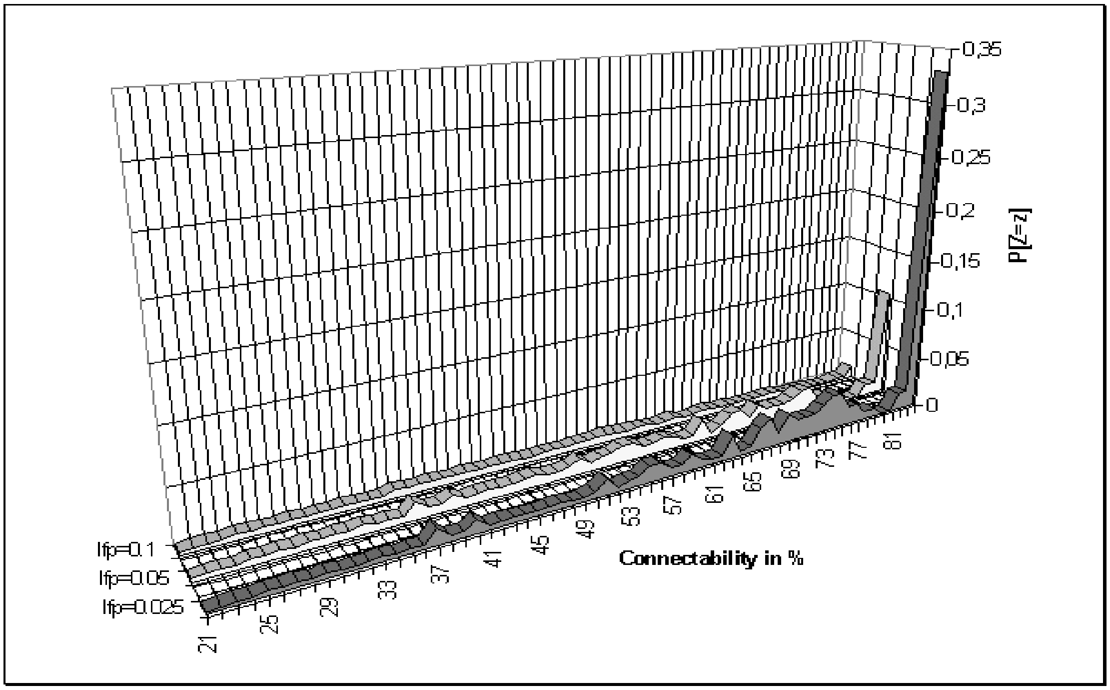

Fig. 19: Disaster Extent Influence on Connectability Survivability Functions; Local-Destination Rerouting; Asymmetric Traffic Demand.

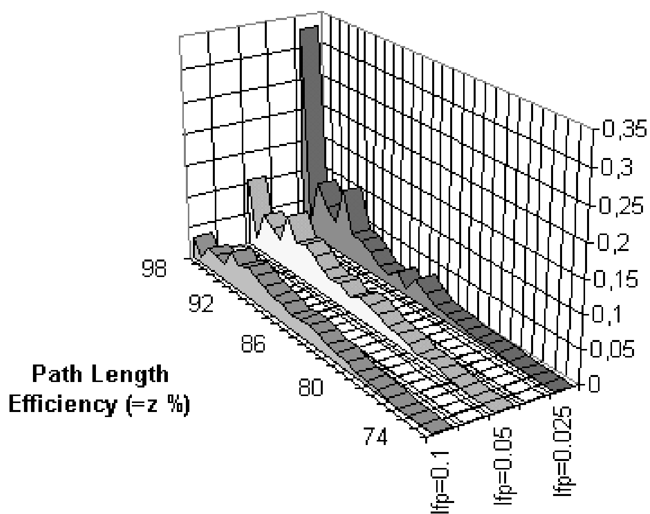

Fig. 20: Disaster Extent Influence on Path Length Efficiency Survivability Functions; Local Rerouting; Asymmetric Traffic Demand.

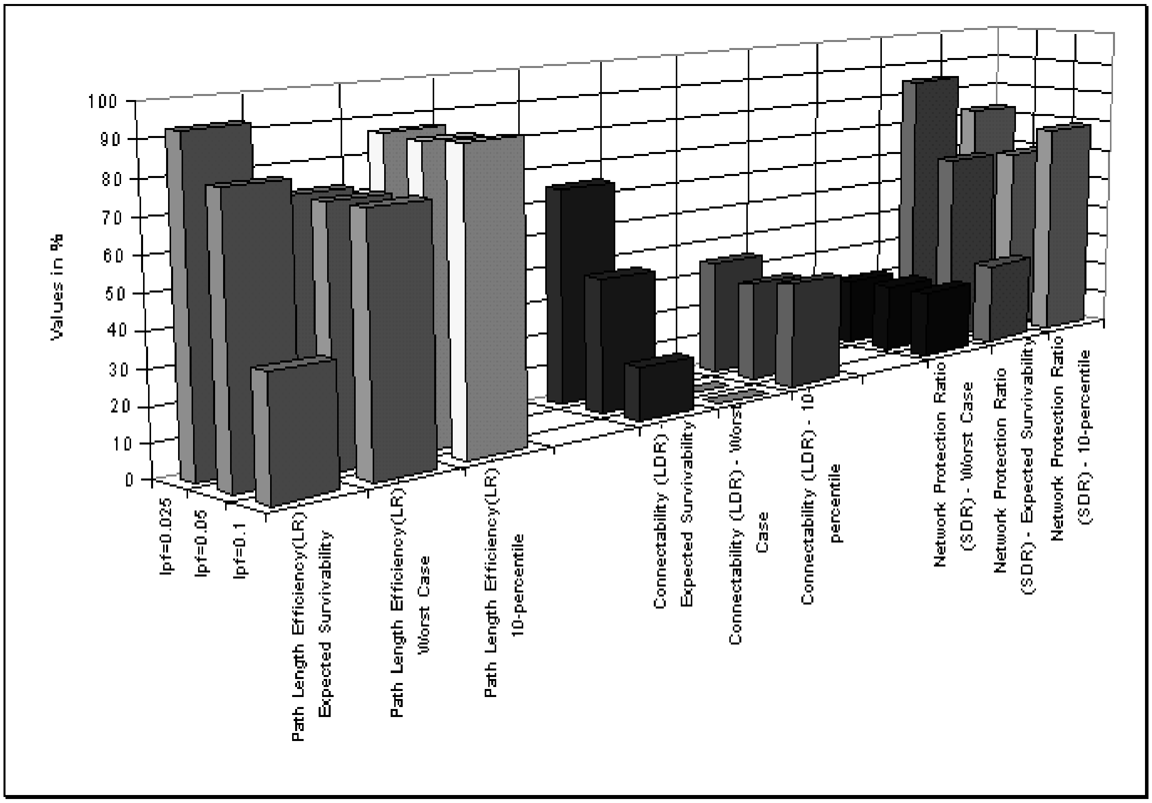

Fig.21: Disaster Extent Influence on Survivability Parameters: some examples.

H. SOME REMARKS ON THE MESH TOPOLOGY USED (Fig.6)

In this study, a physical network with a Manhattan-city type topology has been considered. An important question to ask is, to what extent can the results obtained be generalized for all realistic networks? In other words, since when can a topology be representative of real existing networks? This question was treated by Dunn et al. [26], who designed a study which randomly generated several transport networks models (topologies). The test networks varied systematically in average nodal degree (D) and average span locality (L). L is a measure specific to Dunn et al.'s study for reflecting the propensity of out-of-the-plane spans in a network (i.e. topological cross-overs). $L = \infty$ means any node is likely to have a direct span to any other node at any distance away; and L=1 implies that spans exist only between nodes that are geographically adjacent in the grid space in which the trial networks were synthesized. Joint parametric variations of L and D ensured that the sample space of test networks covered the range of L and D characteristics observed in real transport networks. In fact, published examples of Canadian, US, UK, German, Indian, and French national networks, plus several metro networks, are well imitated by random graphs synthesized with $2.5 < D < 5$ and $1 < L < 1.5$ [17]. Since the network used in this study has D=3.5 and $L \approx 1.1$, it can be stated that the results obtained are to some extent representative for real transport networks.

Table III: Simulation results summary

network goodness metric	local rerouting	local-destination rerouting	source-destination rerouting
redundancy(*)	75%	59%	41%
network protection ratio survivability			
• expected	≈75%	≈75%	≈75%
• 10-percentile	≈45%	≈50%	≈70%
• worst case	≈3%	≈10%	≈20%
residual traffic survivability			
• expected	≈97%	≈97%	≈97%
• 10-percentile	≈98%	≈98%	≈97%
• worst case	≈80%	≈80%	≈80%
path length efficiency survivability			
• expected	≈90%	≈94%	≈97%
• 10-percentile	≈89%	≈92%	≈96%
• worst case	≈74%	≈76%	≈88%
network VPI capacity ratio survivability			
• expected	≈2,35	≈1,8	≈1,65
• 10-percentile	≈2,35	≈1,83	≈1,74
• worst case	≈2,4	≈1,9	≈1,83
connectability survivability			
• expected	≈53%	≈68%	≈48%
• 10-percentile	≈50%	≈54%	≈32%
• worst case	≈0%	≈0%	≈0%

(*) One takes, exemplary, the symmetric traffic demand case

VII. CONCLUSIONS

The substance of the results presented in this paper are very meaningful and useful for the construction of future ATM networks. As illustrated in Tables III and IV, the merits and demerits of the three rerouting schemes are not one-sided. Local rerouting is simple and very fast, but characterized by a very inefficient use of network resources, namely of spare capacities and VPI numbers. It also causes the highest transport delay increase consecutive to rerouting. Source-destination restoration distinguishes itself by a very effective utilization of network resources, but it has the following drawbacks: a very complex rerouting decision, and potentially the highest restoration time. Complex rerouting decisions make the source-destination scheme unreliable, particularly in the case of dynamic & distributed alternate path calculation. The local-destination scheme realizes to some extent a compromise between the two preceding schemes. Its organization of the spare capacity results in a very good connectability performance. In addition, it performs moderate results for all other network goodness metrics.

It has also been shown that the 'worst-case survivability' and the '10-percentile survivability' are two intrinsic quantities, which can be extracted from any survivability function without worrying about the disaster extent (considered by the computation).

Out of consideration for the above conclusions, an idea could be, to superimpose, logically, several VP networks over the facility infrastructure. A VP network aimed for high priority services (with high availability requirements), such as tele-banking and military communications, could use the local rerouting restoration scheme. Since this scheme consumes network resources inefficiently, the billing functions could impose higher charges for this service class. Another VP network destined to low-priority services could use source-destination rerouting. This is quantitatively possibly the largest group. A very good network resources utilization would result in lower charges here. Finally, other various services with moderate priority could be transported on a VP network using local-destination scheme for restoration. The charges would then also be moderate.

A question for future study is, in what respect variations of the physical topology influence the performance of rerouting strategies (considering, of course, the performance metrics framework presented in this paper).

Table IV: simulation results commentary

goodness metric	local rerouting	local-destination rerouting	source-destination rerouting
rerouting decision complexity	<u>low(+)</u>	high	highest (-)
required redundancy	highest (-)	moderate	<u>lowest (+)</u>
network protection ratio	worst (-)	moderate	<u>best (+)</u>
residual traffic	good	good	<u>best (+)</u>
path length efficiency	worst (-)	moderate	<u>best (+)</u>
network VPI capacity ratio	worst (-)	moderate	<u>best (+)</u>
connectability	good	<u>best (+)</u>	worst (-)
restoration speed	<u>fastest (+)</u>	moderate	slowest (-)

VIII. ACKNOWLEDGMENTS

The research reported was carried out at the University of Hagen, Electrical Engineering, FTK, Hagen-Germany. I am very grateful to DAAD (Deutscher Akademischer AustauschDienst) for the scholarship program under which this work was conducted.

REFERENCES

[1] R. Kawamura, K. Sato, and I. Tokizawa, "Self-Healing Network Techniques Utilizing Virtual Paths," *Proc. 5th Int. Network Planning. Symp.*, Kobe, Japan, May 17-22, 1992, pp.129-134.

[2] K. Sato, S. Ohta, and I Tokizawa," Broadband ATM Network Architecture based on Virtual Paths," *IEEE Trans. Communications*, vol.38, no.8, August 1990.

[3] R. Kawamura, K. Sato, and I. Tokizawa, "Self-Healing ATM Networks based on Virtual Path Concept," *IEEE JSAC*, vol.12, no.1, January 1994.

[4] T. -H. Wu, *Fiber Network Service Survivability*, Boston, MA, Artech House, 1992.

[5] J. Anderson, B. T. Doshi, S. Dravida, and P. Harshavardhana, "Fast Restoration of ATM Networks," *IEEE JSAC*, vol.12, no.1, January 1994.

[6] H. Hadama, and S. Ohta, "Routing Control of Virtual Path in large-scale ATM-based Transport Networks", *IEICE Japan*, vol. J72-BI, no.11, pp.970-978, Nov. 1989.

[7] —, "Multiple-Availability-Level ATM Network Concept Integrating Multiple QOS, Dynamic Demand Admission and Logical Spare Capacity Assignment," *Proc. Fourth IFIP Workshop on Performance Modelling and Evaluation of ATM Networks*, 8th-10th July, 1996, Craigland Hotel, Ilkley, West Yorkshire. D.D. Kouvatsos (Ed.) Bradford, U.K.

[8] —, "A Customized Service Availability Concept based on Traffic Segregation, Path Diversity, and Hierarchical Restoration in Public Telecommunication Networks" *Proc. of the Third IEEE ICECS'96*, October 13-16, 1996, Rodos, Greece.

[9] M. H. Macgregor *et al.*, „The Self-Traffic Engineering Network*", Canadian Journal of Electr.&Comp. Eng.*, vol.18, no.2, 1993.

[10] C. J. Colbourn, *The Combinatorics of Network Reliability*, Oxford University Press, 1987.

[11] M. H. Macgregor, W.D. Grover, and U.M. Maydell, "Connectability: A performance Metric for Reconfigurable Transport Networks," *IEEE JSAC*, vol.11, December 1993.

[12] H. Fuji, "Restoration Message Transfer Mechanism and Restoration Characteristics of Double-Search Self-Healing ATM Network," *IEEE JSAC*, vol.12, January 1994.

[13] T. Chujo *et al.*, "The Design and Simulation of an Intelligent Transport Network with Distributed Control," *Proc. Network Operations and Management symposium*, San Diego, February 1990.

[14] A. L. Black *et al.*, "The Bellcore Reliability Prediction Procedure (RPP)," *Proc. IEEE ICC* 1988.

[15] D. S. Kobayashi and M. Tesfaye, "Availability of bi-directional line switched rings," *Rep. T1X1.5/91-070.*

[16] —, "The problematic of Communications Networks Integrity: an Overview," *Proc. joint IEEE ICCS'96/ISPACS'96*, 25-29 Nov. 96, Singapore.

[17] W. D. Grover, "Distributed Restoration of the Transport Network," in *Telecommunication Networks Management into the 21st Century: Techniques, Standards, Technologies, and Applications*, IEEE Press 1993.

[18] B. Sanso, *Reliability and Routing in a Telecommunication Network*, PhD thesis, publication #605, Oct. 1988, Centre de recherche sur les transports, Université de Montréal.

[19] K. K. Aggarwal, "Integration of Reliability and Capacity in Performance Measure of a Telecommunication Network," *IEEE Trans. Reliability*, vol.R-34, 1985, pp.184-186.

[20] B. Sansó and F. Soumis, "Communication & Transportation Network Reliability Using Routing Models," *IEEE Trans. Reliability*, vol.40, no.1, April 1991.

[21] A. Zolfaghari and F. J. Kaudel, "Framework for Network Survivability Performance," *IEEE JSAC*, vol.12, no.1, January 1994.

[22] *Technical Report on Network Survivability Performance*, Standards Committee T1A1.2, 1993.

[23] *Special Issue on Surviving Disaster, IEEE Communication Magazine*, vol. 28, June 1990.

[24] H. Koga *et al.*, "Future Maintenance Systems for Optical Fiber Cables," *Proc. IEEE ICC'91*, pp. 323-329.

[25] Y. Liu and W. Xiong, "Evaluating the Reliability of Telecommunication Networks with Multi-threshold Model," *Proc. IEEE ICC'92*, 1992.

[26] D. A. Dunn, W. D. Grover, and M. H. McGregor, "Comparison of the k-shortest paths and maximum-flow routing for network facility restoration," *IEEE JSAC*, vol.12, no.1., January 1994.