Evaluation of the APS Protocol for SDH Rings Reconfiguration

Pascal ANELLI, Michel SOTO

Abstract— **The SDH (Synchronous Digital Hierarchy) architecture is one of the underlying technologies used by ATM networks. The SDH includes various protection mechanisms. One main design issue is probably the reconfiguration process in case of failure. In case of SDH works on optical fibres with a ring network topology, the Automatic Protection Switching (APS) protocol can be used.**

This paper addresses the problem of maximum allowed recovery time in four fibers ring architecture. We analyze the APS protocol and derive upper bounds for the processing time in each node of the network in order to cope with the maximum reconfiguration time of 50 ms, as specified in the standard.

We finally analyze the behavior of the system in case of two interleaved failures. A worst case analysis is carried out, showing that a 100 ms reconfiguration time can be guaranteed.

Keywords— **Synchronous Digital Hierarchy, B-ISDN, Protocols, Communication System Performance**

I. Introduction

IN the area of B-ISDN, the ATM architecture of protocols is widely studied. This architecture needs a physical layer able to cope with the high throughput involved in the application developments they are supposed to support. The Synchronous Digital Hierarchy (SDH) technology offers technical possibilities to build an infrastructure of a high speed transport network which conveys the broadband services.

Broadband network services rely on high throughput and high reliability. SDH is an ITU-T digital transmission standard [1] that defines common interfaces for vendor compatibilities, digital hierarchy for fibre optic transmission and a frame structure for multiplexing. This standard was initially proposed by ANSI under the name Synchronous Optical Network (SONET)[2]. ITU-T has further refined and generalized the concept to produce the SDH (Synchronous Digital Hierarchy) of which SONET is a subset. An introduction in [3] presents the basic concepts about SDH and its frame formats.

SDH based networks will have many advantages over the digital networks currently in use. They well meet the requirements of the new broadband network services: the operation and maintenance functions provide automatic restoration of services when equipments or links fail like cable cuts for instance. Each level of SDH network disposes of overhead and elements of quality monitoring needed to exchange informations about its operation and maintenance [4].

To achieve the highest network reliability and survivability such networks must be designed according to a ring architecture [5] and uses physical redundancy. The ring topology allows to protect against link or node failures. whereas linear architecture is useful only to protect against link failures. The redundancy in the ring deals with bandwidth as well as network components. Thank to that redundancy, the ring has the ability to be self-healing. Moreover, the ring topology benefits include simplicity, flexibility and potential fast restoration time. The restoration times must be less than 50 ms after detection of the failure according to current network protection switching time requirements at the SDH level [6].

These architectures of self-healing ring (SHR) are divided into two categories: the bidirectional SHR's (B-SHR's) and unidirectional SHR's (U-SHR's)[7]. The type of the ring is defined according to the direction of the traffic flow under normal working conditions. In a B-SHR, the duplex traffic is on the same path and traverses the same set of nodes for both directions of transmission. Conversely, in U-SHR, the duplex traffic travel over opposite path and all the nodes of the ring are involved i.e. transmitting and receiving in normal condition is done on the same fibre. The SHR can be further categorized into path or multiplex section protection depending on which layer the protection switch is made [5]. The path protection switching uses path layer indications and restore individual end-to-end service channel. While multiplex section protection switching uses indications located in the section overhead (SOH) to restore multiplex demand from a failed equipment. The various protection schemes for fibre networks are presented in [8]. In [6], ITU-T defines a protection scheme for the multiplexage section level. The restoration service at this level is achieved by using Automatic Protection Switching (APS) system to perform a loop-back function when a failure occurs. Each node of the network participates in the restoration process.

For the moment, the APS system is well-defined for the Multiplex Section Shared Protection Rings (MS-SPRing) [6]. A MS-SPRing is a bidirectional ring that uses the multiplex section level status and performance parameters to initiate APS. The APS system is based on three different elements: first, the presence of a protection channel which is able to substitute to the channel carrying the traffic under working conditions; second, the capability in the nodes to switch and to select the traffic between protection channel and main channel; third, a protocol that provides self-healing capability to mitigate network component failure and uses the SDH multiplex section layer indications for

This work was supported by Alcatel CIT, Lannion center, France.

Université Pierre et Marie Curie, LIP6, Paris, France. E-mail: `Pascal.Anelli@lip6.fr`

Université René Descartes, LIP6, Paris, France. E-mail: `Michel.Soto@lip6.fr`

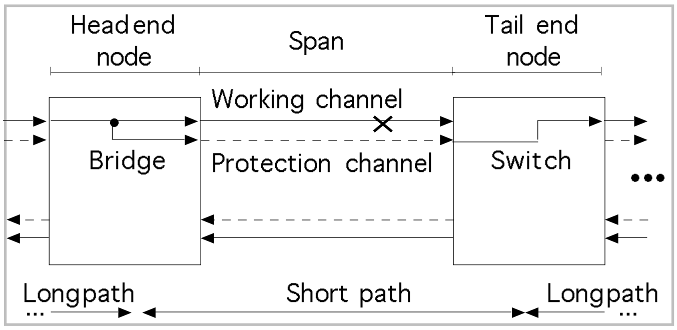

Fig. 1. APS vocabulary

protection switching.

It is crucial to know exactly the processing time allowed for a node to meet the current network protection switching time requirements (i.e. 50 ms). This consideration is taken into account in the design of the node architecture when implementing the APS system in a node.

The paper is organized as follows: section 2 describes the APS protocol in the SDH environment; section 3 presents the modeling approach and the chosen performance criteria; section 4 gives the modelization results; finally, section 5 presents some design issues for the SDH nodes.

II. The APS protocol

We will now present the APS principles and the vocabulary used in this paper [6]. Figure 1 shows the basic terms of the system. The traffic travels in the working channel bidirectionally on separate fibres through the same ring nodes. The working channel is the channel used for traffic transport under normal conditions, i.e. without any failure. A switch event occurs when a network component failure occurs. The working channel is protected by an alternate channel (the protection channel) which is used to transport the traffic during a switch event. Otherwise, protection channel is used to carry extra traffic which is not protected and could be preempted in case of a switch event. MS-SPRing can be implemented on 2 or 4 fibres. With a two-fibre MS-SPRing, protection channel is provided by reserving a part of the bandwidth on each fibre. In this case, no fibre is dedicated for protection, so the protection channel and the working channel share the same fibre. With the four-fibre MS-SPRing, the working and protection channels are carried over separate fibres. In 2 or 4 fibres case, two adjacent nodes are connected by a set of four channels which is called a span.

A failure on a span is detected and corrected by its adjacent nodes. These nodes are called switching nodes and they use the bridge and switch actions for the protection of the working channel. The bridge action consists in transmitting identical traffic on both the working and protection channels. While the switch action consists in selecting traffic from the protection channel rather than the working channel. When a node detects a failure, it becomes a switching node and more precisely a tail end i.e. it requests that the previous node behaves like a bridge. Conversely, a node that is notified for a failure is the second switching node called head end i.e. it executes a bridge. The requests for switching are transmitted on both the short and long path. The short path is defined as "the path segment on the span for which the request is initiated". This span is always the one to which both the head end and the tail end are connected. The long path is the other path segment which connects the head end and the tail end but including all the other spans. Therefore, other nodes belong to this path segment.

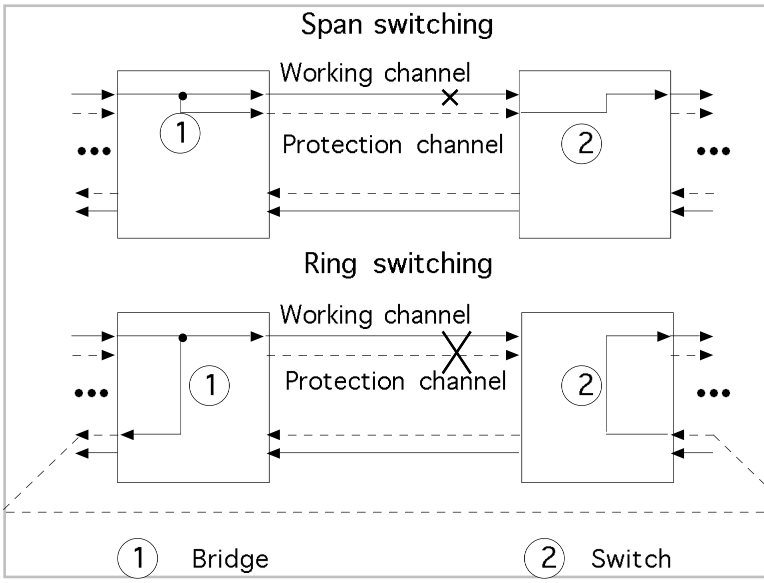

Fig. 2. Protection mechanisms

There are two protection mechanisms as shown on figure 2: ring switching and span switching. The latter is only used on a four-fibre MS-SPRing in order to provide an additional degree of protection. A span switching consists in the transmission of the working traffic on the protection channel of the span where the failure occurs. In this case, the protection channel is substituted to the working channel in fault. A restoration using the ring switch is only needed if both the protection and the working channels on the same span are affected by failures. Here, the working traffic is transmitted to the other switching node, through the protection channel of the long path.

When a node decides that a switch is required, it sends the appropriate request in both directions, i.e. the short and long path. Consequently, a node can receive the same request from different directions. Span switching fully relies on the request transmitted on the short path. The long path signaling informs the other nodes that a span switch exists elsewhere in the ring. At the opposite, the ring switching fully relies on the requests transmitted on the long path. The failure of a channel can be any one of two: a hard failure mainly due to a loss of signal, or a soft failure asserted when the observed bit error ratio (BER) exceeds a preselected threshold. In the first case, the failure is called

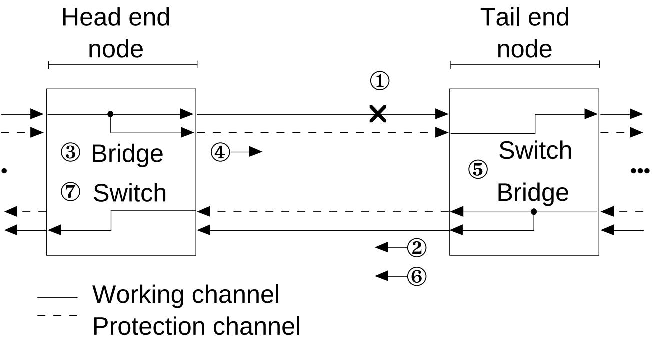

Fig. 3. SDH automatic protection switching principle

a Signal Fail (SF) and in the later case, it is called a Signal Degrade (SD).

A. APS algorithm

The APS protocol uses in-band signaling for protection switching through K_1 and K_2 bytes within the SDH multiplex section overhead. K_1 identifies the switch priority and the destination node while K_2 designates the source node and indicates the path and the status of the switching operation. The basic idea of the protocol is to ensure the working traffic protection by local actions. The two neighbors of the failure are involved in the protocol exchange. The protocol operation follows this general pattern as shown on figure 3:
1. The node which detects a failure (1) (tail end) sends a request (2) to the other adjacent node (head end) of the failed span. The request depends on the type of the failure.
2. The head end, after receiving the request, executes a bridge or a bridge and switch (3). Then it sends a acknowledgment to the tail end (4).
3. On receiving the confirmation, the tail end executes a bridge and switch in one or two steps (5). It terminates by sending its status (6).
4. Finally, the head end finishes by a switch if necessary (7).

The requests to perform protection switching can be initiated either externally or automatically. The external commands are initiated by the operating system or by a craft person. Afterwards, only the automatically initiated protection switching commands are studied because they are based on multiplex section and equipment performance criteria. So, the lockout of channel, the forced switch, the manual switch and the exercise are excluded. Likewise, we do not take into account failures of the protection channel. Only the traffic carried over the working channel is protected and consequently, no mechanism is provided to protect information possibly transferred over the protection channel when no switch even took place. There are four automatically initiated commands. When a hard failure affects the working channel only, the restoration is done in using the span switching method. In this case, the command is Signal-Fail-Span (SF_S). The Signal-Fail-Ring (SF_R) command is used for the same type of failure affecting both the working and protection channels over the same span. In case of soft failure affecting the working channel, the command initiated is the Signal-Degrade-Span (SD_S) and a Signal-Degrade-Ring (SD_R) when the failure affects both channels. When the failure has been cleared, the ring is restored in its initial state (if no other failure condition exists).

Since the protection channels are shared among all spans, contention among the nodes for the protection facility may arise when multiple failures occur. Consequently, the APS protocol defines a priority between the commands. The basic principle is that the priority of signal degrade is less than signal fail, and that span switching has priority over ring switching when facing the same type of failure (e.g. SF or SD). So the priority for the automatically initiated commands is in increasing order : SD_R, SD_S, SF_R, SF_S.

The priority is examined by each node before performing any protection switching activity. The APS protocol contains a preemption mechanism between ring and span switches. According the priorities determined by the standard, five couples of preemption schemes may exist. These couples are: (SD_R - SD_S), (SD_R - SF_R), (SD_R - SF_S), (SD_S - SF_R) and (SF_R - SF_S).

The standard allows the co-existence of switches under some special circumstances. Any span switch can co-exist with any other span switch; so SF_S and SD_S co-exist because these commands are processed locally and independently from each others. The SF_R command co-exist with SF_R command, the ring will recover by segmenting into multiple sub-rings. In this case, only the connections limited at one segment will be recovered. Conversely, when multiple SD_R commands exist over different spans, no action is executed and existing bridges and switches are dropped. If the failures can co-exist, each one is repaired but in case of preemption, only the last failure will be repaired at the end of the switch completion time. If the protection can be considered good for one failure, with two or more failures ring functionnality will be degraded.

III. The model

ITU-T in [6] defines the general network objectives. The maximum number of nodes on a MS-SPRing is sixteen. The end-to-end switch completion time should be within 50 ms after detection of a fault condition when no previous switch request exists for a ring of less than 1 200 km of fibre. The switch completion time is defined as the *interval from the decision to switch to the completion of the bridge and switch operation at a switching node initiating the bridge request.* This means that the detection time is not included in the switch completion time. This latter is completed as soon as the switching nodes complete their operation. So this time represents the amount of time following the failure detection until the effective recovery of the service (i.e., Bridge and Switch executed in each switching node).

During this period, the head and tail end nodes must execute the bridge and switch and thus must exchange the appropriate information. The request sent on the long path takes a time due to the propagation delay plus the processing time spent in each intermediate node. An intermediate node is located between 2 switching nodes on the long path. The propagation delay depends on the length of the ring and on the signal speed through the fibre. For a network of maximum size (i.e. sixteen nodes and 1 200 km of fibre), the propagation delay is constant and depends on the physical properties of the fibre.

The question that arises now is to know the time that the switching and intermediate nodes can spend to process the current command without exceeding the maximum switch completion time, that is 50 ms. Unfortunately, the APS protocol has different behaviors for each command. For example, a node executing a span switch sends the request on the short path, while the long path is used for a ring switch. In general, the request sent on the long path takes more time to be received. Facing this problem, we will try to determine command which is critical with respect to the time it consumes, so that the reconfiguration can take place within the maximum allowed duration.

The processing time in the switching node includes the time for the K_1 and K_2 bytes generation, the K_1 and K_2 bytes software processing and loop back switch control. This paper evaluates the acceptable maximum processing time when a failure occurs, the ring being initially in normal state. When this maximum processing time is determined, we study the switch completion time when a second failure occurs during the processing of the first one. The second failure occurs at $t + \Delta t$, if t is the date of the first failure. The protection switching step is time critical. Consequently, the ring restoration in its initial state when the failure(s) has cleared is not studied. The study relies on real situations and is mainly focused on the worst case recovery time of the SDH ring. The results are obtained by simulation using the OPNET tool [9]. Simulation is the only way we can use, because of the complexity of the protocol, in particular its inherent parallelism.

According to [6], there are three node states: the idle state, the switching state, and the pass-through state. Conceptually, these states apply to a single APS controller which is the part of the node that is responsible of performing protection switching operations. The idle state means that the ring is running under normal conditions without any fault condition detected. The switching state is devoted to a switching node (tail or head end) near the outage span. The pass-through state is reserved to the intermediate nodes; these nodes are located between the switching nodes along the long path.

In the pass-through state, a node transmits on one side exactly what is received from the opposite side. The tail and head nodes go into the switch state and then begin the exchange of control informations using bytes K_1 and K_2 in both directions on the ring. The intermediate nodes only propagate K_1 and K_2 bytes and go into pass-through state as soon as they know the occurrence of the failure. In case of contention between commands, the priorities of these commands can modify the node state, in particular between switch and pass-through states.

For modeling purposes, we choose to divide the APS controller of the node in 2 parts, one for each span. Thus, a node is made of 2 state machines in our simulation model. The general behavior of an APS controller is to look at all incoming informations, then to choose the highest priority input, and to take action based on that choice. In order to model this behavior, each state machine communicates with the other, so that only the one in charge of the request of higher priority is active.

The switch completion time depends on the values of the APS protocol parameters we will now present. The processing time T_{proc}, as already explained, is one of the main parameters. It is the time for a node to move from the idle state to the switch state. The determination of a bound for this parameter is the main goal of this study. The rule on the K_1 and K_2 bytes validation also influences the switch completion time. This rule applies to nodes which are in the idle state and in the switch state. The K-byte validation rule is the following: *before accepting the K-bytes as valid by the node, the value must be received identically in three successive frames.* The time to detect the change of K_1 and K_2 bytes and the time to move from the idle state to the pass through state is represented in an intermediate node by T_s. The time T_s depends on the implementation of the change detection from idle to pass-through in an intermediate node. It is 0 ms in a full hardware processing implementation and 1 ms in a software processing implementation. Once an intermediate node is in the pass-through state, it is supposed to transmit transparently K_1 and K_2 bytes from one span to the other with a pass through delay D_p. This maximum delay is equal to 125 μs. We assume that the average size of a span is 80 km (obtained by dividing the maximum total size of 1 200 km by the maximum number of nodes over the ring, i.e. 16). The propagation delay (T_p) over a span is supposed to be constant and equal to 380 μs. Finally T_{base} represents the transmission time for a SDH frame i.e. one frame (including K_1 and K_2 bytes) every 125 μs.

The span switch commands are initiated on a MS-SPRing with four fibres. Accordingly the network model uses four fibres. The processing of a request received on the short path is not interrupted by the reception of a request with the same priority on the long path. The last request received will be processed by the switching node when the processing of the current request is completed. Circuits are ignored because they do not matter in re-establishing traffic continuity.

The value of T_{proc} is supposed to be the same for each switching nodes. When a node moves from idle state to switch state, T_{proc} is divided in 2 components: T_{p_1} represents K_1 and K_2 bytes software processing time, that is the reaction time when the switching node has detected a failure or has been notified that a failure has occurred, plus the time for K_1 and K_2 bytes generation; T_{p_2} is the bridge and switch time. The bridge and switch time is assumed to be equal to the bridge time $T_{p_{2_a}}$ plus the switch time $T_{p_{2_b}}$. So, we have:

$$T_{proc} = T_{p_1} + T_{p_2}$$

and

$$T_{p_2} = T_{p_{2_a}} + T_{p_{2_b}}$$

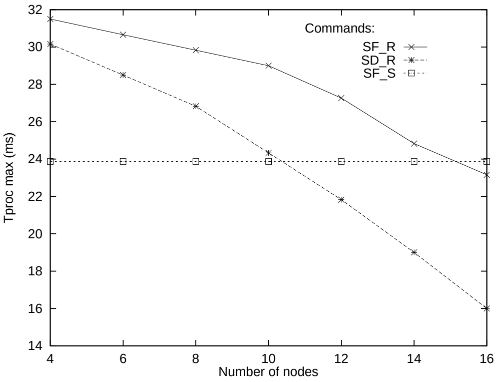

Fig. 4. Maximum admissible processing time for each command depending on network configuration

When a ring request is preempted, the time to undo a bridge and switch is equal to the time to execute the bridge and switch, T_{p_2}.

We begin by the calculation of the maximum processing time, $T_{proc_{max}}$, that each switching node can use without violating the maximum completion time of 50 ms from the failure detection to the recovery. The first approach uses the maximum value of T_s, that is 1 ms (software processing implementation case). Under that hypothesis, the time for a request to go from one switching node to the other by the long path is the bigger possible.

Unidirectional or bidirectional failures may occur on a ring. A unidirectional failure means that only one way of the span is damaged. In this case, the failure is detected by the tail end node and notified to head end node on the short path. The notification of the failure on the short path by the tail end node is a time consuming operation. When a bidirectional failure occurs, the span is damaged on both ways. As a consequence, a bidirectional failure is simultaneously detected by both of the switching nodes and both of them become a tail end node i.e. request a bridge. Thus, a bidirectional failure needs less time to restore traffic than a unidirectional failure since it is not necessary to notify the failure on the short path. Following these remarks, we focus the study only on unidirectional failures, involving the following commands: SF_S as a result of signal failure on the working channel only, SF_R when the signal failure is also on protection channel and SD_R as a result of signal degradation on both working and protection channels. The SD_S request is not studied because it is equivalent to SF_S request sequence. For convenience, in the following, the name of the command is also used as a synonym of the corresponding failure.

IV. Results

The value of $T_{proc_{max}}$ is evaluated in the worst cases of failure that is unidirectional failures and for 4-fibre rings ranging from 4 to 16 nodes. Assuming $T_{p_1} = T_{p_2}$, the values of $T_{proc_{max}}$ are shown on figure 4.

The maximum processing time, $T_{proc_{max}}$ is expressed by:

$$T_{proc_{max}}(i,j) = T^{-1}_{reconfig}(50ms, T_s)(i,j)$$

where $i \in \{SF\_S,\ SF\_R,\ SD\_R\}$, $j \in \{4,\ 5,\ \ldots,\ 16\}$ and $T^{-1}_{reconfig}$ is the inverse function of $T_{reconfig} = f(T_{proc}, T_s)$.

The analysis of figure 4 exhibits a constant $T_{proc_{max}}$ for span switching command, whatever is the number of nodes of the ring. For this command, the traffic protection switching procedure mainly uses the short path. The long path is not used and thus, the number of nodes has no effect.

It is interesting to notice that there is no command giving always the lowest value for $T_{proc_{max}}$, whatever is the number of nodes. When the ring has a small number of node, span switch commands give $T_{proc_{max}}$ values less than those obtained for ring switch commands in the same configuration. This means that a ring switching is faster than a span switching on small rings. This could be surprising because span switch only uses the short path and, thus, the time consumed to exchange commands is very low. The explanation of this phenomenon is given by the span switch procedure itself: the operations involved in a span switch for the switching nodes are purely sequential so that the needed processing steps alternate between the two switching nodes. Conversly, in the ring switch procedure, propagation of the commands and processing steps are done simultaneously. In this case, a command can arrive during the processing of previous one and can be processed immediately after. The parallelism of the the ring switch procedure is favored while the propagation time of the command through the long path is less than the processing time at the switching nodes; this situation occurs with a small number of nodes. When the number of node is greater than 10, SD_R becomes the worst failure case.

Figure 4 also shows that the value of $T_{proc_{max}}$ for the SD_R procedure is always less than the $T_{proc_{max}}$ for the SF_R procedure. The reason is also given by the SD_R procedure which needs one more exchange of command on the long path than the SF_R procedure.

Finally, the SF_R command allows the highest value for $T_{proc_{max}}$ for a number of nodes less than 15. Thus, the worst situation for T_{proc} is not related to the type of command executed. For commands using the long path (typically ring switch), the value of $T_{proc_{max}}$ is related to the number of node in the ring and the ring length. As the number of nodes and the ring length increase, the propagation of K_1 and K_2 bytes through the ring consumes time and thus, the acceptable value of $T_{proc_{max}}$ decreases. This fact is amplified by the software implementation detection of the change from idle to pass-through in an intermediate node. Consequently, $T_{proc_{max}}$ is different for every ring configuration.

From the set of values obtained for $T_{proc_{max}}$, we select the value $T_{proc_{limit}}$ defined as follow:

$$T_{proc_{limit}} = \min_{i,j} T_{proc_{max}}(i,j)$$

where

$$i \in \{SF\_S, SF\_R, SD\_R\} \text{ and } j \in \{4, 5, \ldots, 16\}$$

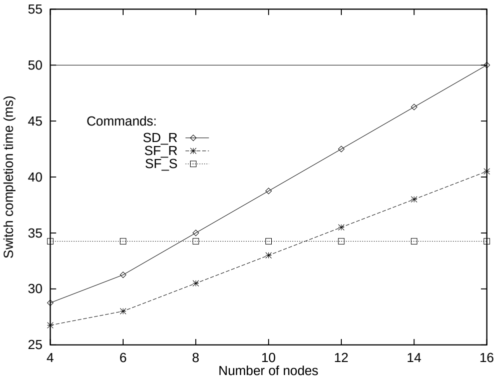

Fig. 5. Switch completion time for different network configurations

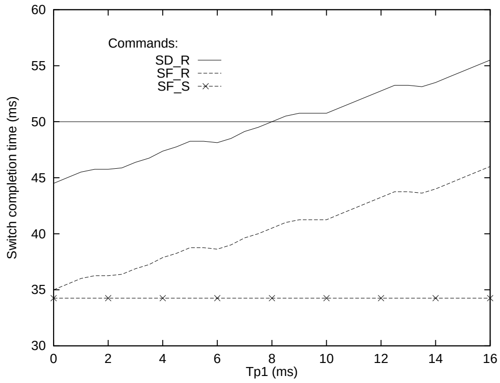

Fig. 6. Time distribution inside T_{proc}

The value of $T_{proc_{limit}}$ is the processing time not to be exceeded by switching nodes to guarantee that the switch completion time is less or equal to 50 ms for any ring ranging from 4 up to 16 nodes and for any command. As shown on figure 6, this value is met with a 16 nodes ring on unidirectional SD_R failure. In this situation, $T_{proc_{limit}}$ is equal to 16 ms. Figure 5 shows the switching completion times for the different failures, when $T_{proc_{limit}}$ is used and T_s is equal to 1 ms.

Figures 4 and 5 show that for span switch commands, the only important parameter affecting the completion time is T_{proc}. In this case, the long path is not used. Further, when requests propagation times are small compared to processing times, ring switch procedures result on smaller reconfiguration times than span switch procedures. This is due, as we mentioned before, to the intrinsic parallelism of ring switch procedure. Finally, the protection procedure following a SF_R is faster than that following a SD_R, whatever is the configuration of the ring. The reason is the extra request through the long path needed by the later one.

For convenience, $T_{p_1} = T_{p_2}$ was assumed. Nevertheless, actual distribution of T_{p_1} and T_{p_2} are strongly implementation dependent. In most of cases, T_{p_1} will be different from T_{p_2}. So, in the following, this assumption is relaxed and the completion time is studied when T_{p_1} ranges from 0 to $T_{proc_{limit}}$. The value of T_{p_2} is given by:

$$T_{p_2} = T_{proc_{limit}} - T_{p_1}$$

The switch completion time for a 16 nodes ring is shown on figure 6. As expected, the completion time of the SF_S command is not sensitive to the respective values of T_{p_1} and T_{p_2}. At the opposite, the completion times of the SD_R and SF_R commands are influenced: when $T_{p_1} < T_{p_2}$, the completion time is reduced. As T_{p_1} is the part of T_{proc} devoted to K_1 and K_2 processing, a small value for T_{p_1} favors the parallelism of the ring switch procedures. In consequence, implementors should focus on optimization of K_1 and K_2 bytes processing time to get efficient implementations of the APS protocol.

APS standard only defines a switching completion time requirement of 50 ms for a single failure on a clean ring (i.e, no extra traffic and no previous bridge request). Under all other conditions, the required completion time may be exceeded but must remain less than 100 ms. The priority scheme allows APS to deal with more than one failure. It is interesting to evaluate how APS behaves in worst case regarding the switch completion time, that is when a second failure occurs during the traffic restoration process due to a first failure.

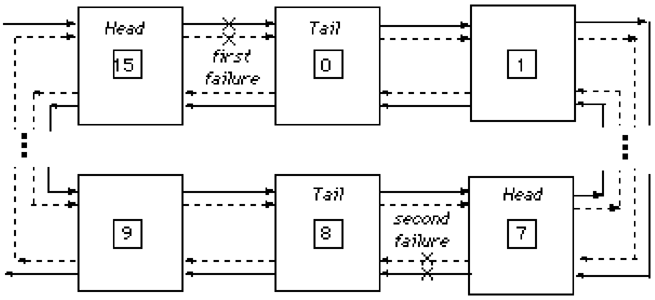

Fig. 7. Network configuration for 2 interleaved failures

In this section, the switch completion time is studied when two interleaved failures occur on four-fiber ring. We selected 3 scenarios of paired failures:

1. SD_R then SF_S, preemption of ring request by a span request,
2. SD_R then SF_R, preemption of a ring request by another ring request,
3. SF_R then SF_R, coexisting ring requests. At the end of the switch completion time, the ring is segmenting into two sub-rings.

We assume the second failure is always located on the farest span from the one where the first failure occurs as shown on Figure 7. The first failure occurs on the span between nodes 15 and 0 while the second failure occurs on the span between nodes 7 and 8.

The delay between the first failure and the second failure ranges from 0 to the completion time of the first failure. Figure 8 shows the results obtained with $T_{proc_{limit}}$, and T_s = 1 ms. For any pair of interleaved failures the completion time stays under 50 ms while the second failure occurs within 16 ms after the first one. In any case, the completion

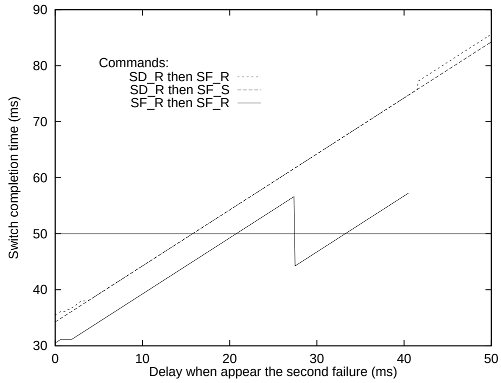

Fig. 8. 2 interleaved failures

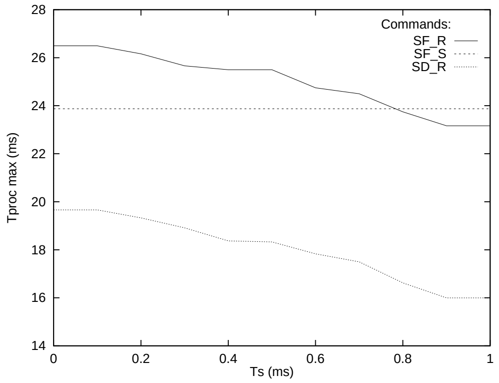

Fig. 9. $T_{proc_{max}}$ as a function of T_S

time is less than 100 ms.

The unusual shape of the curves is explained by the location of the first request when the second failure occurs. If we focus on the curve of the two interleaved SF_R failures, which is the most unusual, we can see the switch completion time growing lineary while the second failure happens in a delay less than 27.5 ms after the first failure.

Until 27.5 ms, the first request can not pass on the span between nodes 7 and 8 before the second failure occurs. The ring reconfiguration process for the first failure is always delayed by the starting of the ring reconfiguration process due to second failure. The thrust of the curve occurs at time 27.5 ms. From this time, the first request always passes the span between nodes 7 and 8 *before* the second failure occurs. In consequence, the ring reconfiguration process for the first failure is no more delayed by the starting of the ring reconfiguration process due to second failure and the two ring reconfiguration processes are performed with a lot of more parallelism. Such a parallelism immediatly decreases the switch completion time. Then, the switch completion time restart growing lineary. In general, the time of the thrust on the curve is directly linked to the location on the ring of the second failure. Actually, on the long path, the closer is the second failure to tail of the first failure, the sooner the thrust of the curve occurs.

As mentioned before, the T_s parameter is related to the implementation of the change detection from idle state to pass-through state in an intermediate node. The value of T_s has a direct impact on the propagation of K_1 and K_2 bytes on the long path and thus on the value of $T_{proc_{max}}$. We have previously set T_s to 1 ms. We now study $T_{proc_{max}}$ when T_s ranges from 0 to 1 ms on a 16 nodes ring.

As shown on figure 9, the admissible value of $T_{proc_{max}}$ decreases as the value of T_s increases. As we can have predicted, SF_S is insensitive to different values of T_s. The reason is that it does not use the long path. On the other hand, for ring switch type procedures, it is to be noticed that the greater is T_s, the smaller is $T_{proc_{max}}$. This is due to the requests exchanged that consume more time to the detriment of the processing time in the switching nodes. To verify this relation between T_s and $T_{proc_{max}}$ we studied the switch completion time as function of T_{proc} for T_s = 0 ms and T_s = 1 ms. The results on figure 10 confirm that, whatever is the value of T_{proc}, the switch completion time is always greater for T_s = 1 ms than for T_s = 0 ms.

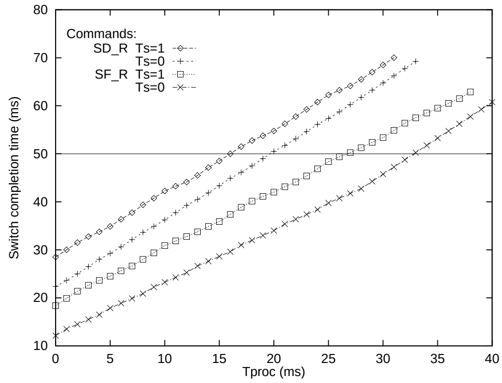

Fig. 10. Switch completion time for ring request with $T_S = 0$ and $T_S = 1$

Nevertheless, these results are not any longer valid with two interleaved failures. Figure 11 shows that the switch completion time is greater for T_s = 0 ms than for T_s = 1 ms. If we consider that T_{proc} is consumed twice (i.e., once for each interleaved failure) in the switch completion time we can make the following approximation to explain this contradictory result:

$$Switch\ completion\ time = 2 * T_{proc} + T_{exchange}$$

where $T_{exchange}$ represents the time consumed for exchanging the needed information between switching nodes. When T_s = 0 ms, the value of T_{proc} will be the highest possible. As T_{proc} is consumed twice, the switch completion time will be more important for T_s = 0 ms than for T_s = 1 ms. This result has consequence for design issues as explained in section 5.

V. Design issues

Implementors of the APS protocol should keep in mind there is no protection switching command that involves the lowest T_{proc} whatever the number of node of a SDH ring. At the opposite, the available T_{proc} is a function of both

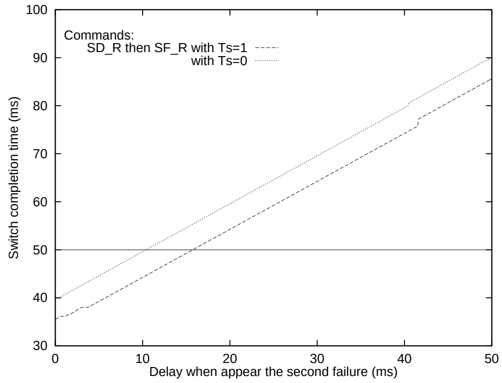

Fig. 11. 2 interleaved failures with 2 different T_S

the switching command and the number of node except for the SF_S command which exhibits a constant completion time in any case. An efficient implementation also involves an optimization of K_1 and K_2 processing (i.e., T_{p_1}) after a detection of an error condition in the switching nodes rather than an optimization of the bridge and switch operation (i.e., T_{p_2}). In this case, the switch completion time is reduced. For instance, if T_{p_1} is equal to T_{p_2}, T_{proc} must be less than 16 ms in order to keep the switch completion under 50 ms. In the intermediate nodes $T_s = 0$ allows the maximum value for T_{proc}. Nevertheless, with two interleaved failures the switch completion is greater when $T_s = 0$ ms. In others words, with two interleaved failures it is better to try to reduce T_{proc} rather than T_s. The switch completion time is more strongly reduced.

A good implementation of APS controller is one which minimizes the processing time in the switching node. So, in case of 2 failures interleaved, the switch completion is reduced when T_{proc} decrease and this whatever the value of T_s.

VI. Conclusions

This study has determined the maximum time processing available in a switching node to meet 50 ms completion time requirement whatever the ring configuration and the type of failure. An important result is that the completion time is not related to a single command. Particularly, when there are less than 8 nodes, ring switch commands are faster than span switch ones. This study has also shown with the maximum processing time available (with our hypothesis, T_{proc} = 16 ms) in a switching node, in any way, the completion time in the case of two interleaved failures meets the requirement of be under 100 ms. Finally the impact of the processing time in the intermediate node has been relativized from this of the switching node.

Acknowledgment

The authors would like to thank Pr. Éric Horlait for his careful reading and his helpful english writing of this paper.

References

[1] ITU-T Recommandation G.707, *Network node interface for the synchronous digital hierarchy (SDH)*, March 1996.

[2] American Standard for Telecommunications, *Telecommunications - Synchronous Optical Network (SONET) - Basic Description including Multiplex Structures, Rates and Formats*, 1995, T1.105-1995.

[3] C.G. Omidyar and A. Aldridge, "Introduction to SDH/SONET," *IEEE Communications Magazine*, vol. 31, no. 9, pp. 30–33, September 1993.

[4] M. Sexton and A. Reid, *Broadband Networking: ATM, SDH, and SONET*, Artech House, second edition, 1997.

[5] J. Hughes, D. Beckett, and J. Meloy, "Protection switching," *Telephone Engineer and Management*, vol. 97, no. 4, pp. 35–39, 1993.

[6] ITU-T Recommandation G.841, *Types and characteristics of SDH network protection architectures*, July 1995.

[7] T.H. Wu and R.C. Lau, "A class of self-healing ring architectures for SONET network applications," *IEEE Transactions on Communications*, vol. 40, no. 11, pp. 1746–1756, November 1992.

[8] T.H. Wu, "Emerging technologies for fiber network survivability," *IEEE Communications Magazine*, vol. 33, no. 2, pp. 58–74, February 1995.

[9] MIL 3 Inc, The INTELSAT Building, 3400 International drive, NW, Washington DC 20008, *OPNET: OPtimised Network Engineering Tool*, m version edition, 1991, M version.