A NEW FAULT-TOLERANT CONFIGURATION FOR THE CAMBRIDGE RING: THE HIERARCHICAL RING-STAR

by

Thet Ngian Chen, B.Sc.

A Doctoral Thesis

Submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy of the University of Technology, Loughborough.

December, 1985

Supervisor : Professor J.W.R. Griffiths

Department of Electronic and Electrical Engineering
Loughborough University
England

To my parents

&

KENLEN

ABSTRACT

The primary objective of this research is to look at ways of resolving the reliability problems of the Cambridge Ring local area network system. The result is a novel design to enhance the Cambridge Ring with fault tolerance by introducing redundant communication paths with dynamic reconfiguration. The proposed Ring-Star system combines the advantages of ring and star networks to create a network which is topologically resilient while retaining the efficient communication advantage of rings.

Although as a local area network the ring has much in its favour, it inherently suffers a major drawback - reliability. Ring networks have by virtue of their design a point-to-point communication medium, and hence a single node failure can bring down the entire network. Likewise, if the transmission cable is cut, the entire network will fail completely. Over the years several techniques have been put forward to mitigate this problem but they have their limitations. A literature review was carried out to survey and evaluate existing fault tolerant techniques, and consequently to identify key design issues. As a result of the analysis, and taking into account the particular properties of the Cambridge Ring network, the author was able to propose a new design.

The design itself evolved in two stages. Phase 1 produced an ideal design while Phase 2 improved on the basic prototype by converting the configuration into one more suitable for practical use. Both prototypes were built and tested.

The Ring-Star concept offers a highly resilient system. It has a star-like ring structure, being a ring network encapsulating within its periphery a star structure. It is this star component which provides the high degree of fault tolerance to the ring - it can tolerate multiple faults. The design employs off-line redundancy techniques to provide automatic fault isolation and recovery, and it requires no human intervention once it is initiated. In essence the Ring-Star provides a non-stop communication facility. However as with any star structured network, installation may be a problem. The basic design was evolved into a multiple Ring-Star architecture, the "Hierarchical Ring-Star" to ease this problem. It also offers an installation more flexibility in layout and growth. An experimental system has been built and successfully tested. From this, an estimate of cost for adding fault tolerance to the Cambridge Ring has been made. The approximate cost per node came to £60. The overall theoretical reliability of the system has also been evaluated. When compared to a basic ring without fault tolerance an improvement of about 50% was obtained while the overall reliability approaches 90%.

Finally, it should be noted that since this design is concerned only with the physical layer configuration: it may be applied to any ring network. Thus it can be adopted to enhance reliability for a ring designed according to the IEEE802.5 standard, IBM token ring or any other proprietary ring.

ACKNOWLEDGEMENTS

The author would like to acknowledge the financial assistance received from Loughborough University and the Rutherford and Appleton Laboratory which made this research possible.

The author would also like to thank Professor J.W.R. Griffiths for his guidance throughout the research, for his help in many aspects and in particular for securing the financial support.

Special thanks must also extent to several friends, Mrs. K. Carey-Smith, Mrs. S. Ip and Dr. D. Parish for reading and commenting on the final draft of this thesis, and to all my colleagues in the Signal Processing Laboratory of Loughborough University for their help and with whom I have enjoyed every moment of my research period.

LIST OF SYMBOLS AND ACRONYMS

CB	CSU Block
CSU	Centre Switching Unit
EOP	End Of Packet
K	Kilo or one thousand
lsi	large scale integration circuit
led	light emitting diode
L	Links or section of ring cables
m-p	minipacket
msi	medium scale integrated circuit
M	Mega or one million
MCSU	Master Centre Switching Unit
m	reliability of MCSU
n	reliability of node
N	Nodes
p	reliability of link
pr	probability
PSU	Power Supply Unit
q	unreliability of link
RP	Relay Port
s	reliability of SCSU
ssi	small scale integrated circuit
SCSU	Slave Centre Switching Unit
SOP	Start Of Packet

U	Union (set theory)
vlsi	very large scale integrated circuit

DEFINITIONS AND TERMINOLOGY

Availability

Is the probability that a system is available to perform its functions at an instant of time.

Critical Faults

Are defined as faults which disable ring operation completely. Thus node failures and broken links are classified as critical faults.

Fault Tolerance

Is the ability of a system to continue to perform its specified tasks after the occurrence of faults.

Link break

Is a condition when the ring cable is severed. This is a critical fault.

Node Failure

Is the condition when a ring node fails in such a way as to disrupt ring operation by corrupting passing packets.

Off-Line Redundancy

Redundancy wherein the alternative means of performing the function is inoperative until needed and is switched on upon failure of the primary means of performing the function.

Reliability

Is the probability that an item will perform a required function under stated condition for a stated period of time.

Resilience

May be generally defined as the ability of a system to survive failures in an organised manner.

Redundancy

Is defined as the addition of information, resources, or time beyond what is needed for normal system operation.

The following definitions are specific for the Cambridge Ring:

Boot Server

A device which provides bootstrapping service for various computers throughout the Cambridge Ring system.

Name Server

The Name Server plays a fundamental role in the Cambridge Ring system. When presented with the text name of a service, a process or a computer, anywhere in the system, the name server returns the appropriate ring address. It is also capable of performing the converse translation.

Repeater

An electronic circuit that receives clock and data, demodulates them and presents them to the Station - Repeater interface; modulates either the received data or data gated to it from the Station, and then transmits regenerated signals on around the ring. Synchronism is achieved using a local oscillator phase locked to the incoming clock frequency.

Station

An electronic circuit that interfaces both to the Repeater and the Interface Unit of an attached device. It peforms serial to parallel and parallel to serial data conversion, controls communications across the ring by synchronising to the minipacket structure. It detects and reports certain error types.

Node

An electronic circuit that combines Repeater and Station functionality.

Interface Unit

Logic that interfaces a Node to a particular type of attached device.

Minipacket

The unit of transmission between Stations controlled by the ring access mechanism.

Slot Structure

A regular framing structure imposed upon the circulating bit stream to carry minipackets.

Monitor

A unique node on the ring with responsibility for initialising and maintaining ring operation and slot structure.

Ring Connectors

Fixed and free connectors for the physical connection of Nodes into the Ring.

TABLE OF CONTENTS

CHAPTER 3 - THE CAMBRIDGE RING NETWORK, A DESCRIPTION

CHAPTER 4 - PRELIMINARY INVESTIGATIVE STUDIES

CHAPTER 5 : DEVELOPMENT OF THE HIERARCHICAL RING-STAR

CHAPTER 6 : IMPLEMENTATION OF THE HIERARCHICAL RING-STAR

CHAPTER ONE

INTRODUCTION

1.1 Background

Ring networks have a number of desirable properties for applications in local area networks (LAN). The protocols can be arranged to allow guaranteed bandwidth to all nodes on the ring and hence provide a deterministic response to user services. For real-time applications such as voice or image, this property is vital. Their point-to-point medium simplifies hardware, allowing rings to operate on a wide range of media from twisted pairs to optical fibres. Also in contrast to bus networks, the transmission speed of ring networks is not limited by propagation time and they can therefore operate at much higher speeds efficiently.

However, rings have one major disadvantage - reliability. A single failure on its transmission path will disable the entire network. Thus damage to the cable at any point could disrupt ring operation and similarly nodes must operate reliably at all times.

There have been several schemes proposed to improve reliability. An early and obvious way is simply to run a duplicated ring in parallel. In the event of any failure on the primary ring, a switch is made to the duplicated ring. This technique is however far from ideal, it is too simplistic and can tolerate only one fault. Moreover it can be very expensive, approximately doubling the cost. Similarly, the other techniques have their advantages and their drawbacks. By analysing them, a better approach is sought.

1.2 Objectives

The primary objective of this research project was to devise a scheme to improve the reliability of the Cambridge Ring. This objective was subject to a number of initial constraints.

- First, the ensuing design should require minimal modifications to the current range of Cambridge Ring equipment. Although restrictive, this helps to protect the present investments made in Cambridge Ring hardware.

- Second, it is also important to take into consideration likely future developments. Operational speed will almost certainly improve from the present 10MHz to 100MHz or more. To facilitate such speeds it is likely that optical fibres will be used as the transmission medium.

- If the design conceived is applicable to ring networks other than the Cambridge Ring, it would probably gain wider acceptance. The ideal is a general design which can be used with the emerging ring standards, or any other new ring innovations.

- Cost is an obvious constraint and especially so in an academic environment. Moreover the less expensive it is the more likely it will find applications.

1.3 The Need for a Reliable Cambridge Ring System

The Cambridge Ring is a type of ring network. Therefore it exhibits most of the properties of a ring network, notably susceptibility to faults in the transmission path. However the current Cambridge Ring standard (CR 82) does not directly address the reliability problems of ring networks. Instead, the design philosophy attempts to reduce the chances of failures, and if a fault was to develop, to pinpoint the source quickly.

Although the quality of Cambridge Ring equipment is very high (Spratt 80, Binn 82), continual operation of the ring cannot be guaranteed. There will always be an element of doubt. For example, a technician may cut a ring cable by accident. This is totally unpredictable.

The reliability of the Cambridge Ring is probably sufficient for University or research environments. But it is hardly suitable for commercial or industrial appplications and certainly not for military use.

Many schemes have been implemented to improve the reliability of rings, for example the IBM token ring has some fault tolerance built in, but with the Cambridge Ring, less has been done. Although Racal's Planet ring network is fault tolerant, its implementation of the Cambridge Ring do not conform with the CR82 standard. The other major suppliers of CR82 rings, Logica, SEEL and CAMTEC do not incorporate fault tolerance in their designs. This may have to change. Among several surveys, a recent study by the market research organisation, Frost & Sullivan (Financial Times, 30 October, 1985) found that the demand for fault tolerant computers is soaring. The reason cited is that businesses as diverse as banking and chemical manufacturers are looking to safeguard themselves against disaster.

Although Frost & Sullivan surveyed computers, their findings should apply to local networks too when they are more widely used. Thus, if the Cambridge Ring is to be more widely accepted in industry, it should also be enhanced with fault tolerance to reassure potential users.

1.4 Research Plan

At the inception of the project, it was decided that the work would be conducted in several phases.

Phase 1 : To examine the literature in order to survey developments in fault tolerant rings, to evaluate them and to establish design guidelines.

Phase 2 : To propose a suitable fault tolerant design for the Cambridge Ring based on the guidelines set out in phase 1, and to develop the necessary concepts and techniques.

Phase 3 : To implement in hardware the fault tolerant design for experimentation and evaluation.

1.5 Summary : Chapter by chapter

Chapter 2 is a preliminary phase providing the necessary background information. The historical development of fault tolerant computing is included to provide the reader with an insight into its significance. The theory and concepts relevant to reliability studies have also been presented.

Chapter 3 describes the Cambridge Ring, paying particular attention to the reliability and maintainability aspects.

Chapter 4 reviews the relevant literature to provide an overview in the field of fault tolerant ring networks.

Existing fault tolerant techniques are compared and evaluated in order to identify key design issues. The result is a set of objectives drawn up to guide development towards a suitable fault tolerant Cambridge Ring system.

Chapter 5 forms the first stage in the development of a new fault tolerant ring configuration. It describes the proposed Ring-Star concept, the techniques adopted and explains the operation. However there is a limitation with the basic system and an improved design is suggested. An implementation of the Hierarchical Ring-Star system is proposed.

Chapter 6 describes the detailed design work carried out to implement the Hierarchical Ring-Star system. The prototype hardware is described generally at system level although the unique aspects are explained in detail. Operational software and algorithms are covered, the latter including flow charts to simplify the explanation.

Chapter 7 demonstrates the operation of the Hierarchical Ring-Star in an experimental study. The system was shown to be resilient to node failures and breaks in the ring cable. Furthermore it tolerates several of these faults without any partition problems. Manual control of the system was also attempted. This proved to be easily configurable and in all cases, normal operation of the ring resumed after a short delay.

Chapter 8 evaluates the Hierarchical Ring-Star system. Its topology is evaluated in terms of cabling requirements, and compared with other fault tolerant approaches. The overall reliability of the experimental Hierarchical Ring-Star is computed to gauge its improvement over a ring without fault tolerant enhancements. Finally an estimate of the cost for adding fault tolerance is also provided.

Chapter 9 discusses the merits and limitations of the Ring-Star concept and concludes the thesis. The research project is reviewed in the context of the objectives set out earlier in the thesis.

CHAPTER TWO

FAULT TOLERANCE - HISTORICAL DEVELOPMENTS AND CONCEPTS

2.1 Historical Developments and the Need for Fault Tolerance

Early computers were extremely unreliable, requiring maintenance several times a day just to keep them operational. But since then, users have been demanding more reliable computers. This trend is reflected in the development of computing.

In the early days, computers were based on vacuum tubes, which like lightbulbs tend to burn-out relatively quickly. One of these machines was built in 1946 at the Moore School of Engineering in Philadelphia, U.S.A. The computer known as ENIAC had 18000 electronic tubes, and it did operate over short periods of time. Maintenance engineers had to be present to keep replacing burnt vacuum tubes to keep it working. The invention of semiconductor devices changed all this. It was the first major step in the process of improving reliability. Today computers built with vlsi components must be thousands of times more reliable than the old vacuum tube machines.

Yet, an increasing demand is evident for even higher levels of reliability (Strube 85). Computers have become so integrated into the working environment that in many applications downtime is not acceptable. The call is for "full availability from assembly to obsolescence." A typical example is banks. Banks were among the first to widely adopt fault tolerant computers since vast sums of money may be lost if their computers are down even for a few hours. Another example is industrial control. In this application, unreliable computer operations may have disastrous consequences which could endanger human life. For these reasons the use of fault tolerant computers have been forecast to grow even more significantly in the future.

The invention of computer networks was the next logical step in this process (though it must be stressed that it was only one of several factors leading to this development). It was realised that an installation could be vulnerable if it depended on a single computer. The numerous cases of major disasters involving computers and the loss of revenue reported by companies serves to illustrate this point. By distributing computing power to several computers situated physically apart this problem is reduced. In the earlier days of networking, computers tended to operate more or less in isolation with the occasional exchange of data. Now, distributed processing is favoured where one could view the entire network as one huge computer. Thus the focus now is on the reliability of the communication channels.

There are many different interconnection schemes to connect together the processors of a distributed system: Star, Mesh, Bus and Rings. When LANs were first developed, the Star and Mesh topologies were rejected in favour of Bus and Ring topologies. The primary objectives then were performance and costs. For example, a fully connected mesh-like network would not be very flexible and is difficult to install. They are also more complex in routing and greatly increase delay in message passing.

The development of networks appears to be mirroring that of computers. Initially, performance was the prime criteria and as technology matured, attention was switched to other issues - resilience being one significant factor. With the continuing decrease in computing cost, the extra expense of more reliable communication is becoming acceptable. But is this argument acceptable? If one is to look back at the development of computing, the question of reliability came as a consequence rather than as a result of planned goals. The transistor was not invented because it was thought the vacuum tube was unreliable. The integrated circuit did not appear because solder connections are less dependable. These improvements came about by normal development efforts in solid-state physics and were not driven by a reliability goal. The problem is, are we going to be that lucky again? This is unlikely. Reliability must be set as a goal, not just a hope that it will appear as a side effect.

The bulk of the potential market for LANs will be in the office and factory environments. Ideally all communication within an organisation, including telephony, data, text and image transfer should be carried out on a single LAN medium. Currently very few suppliers market such an advance product but with the present pace of research, there is little doubt that it will be achieved. Thus the whole activity of a business may depend on the correct functioning of the LAN. A catastropic failure of the LAN may be tolerated only extremely rarely and then only for short intervals. Failure of a single LAN attachment may be tolerated as long as the rest of the LAN continues to operate with no appreciable break. The system must be very resilient.

To improve reliability, each LAN component can be engineered to be highly reliable but this is no solution. However small the failure rate, it will be multiplied by hundreds or perhaps thousands of times in a large installation. Even on a passive bus LAN, some such failure will cause the network to fail completely. Repair can be time consuming and disruptive to the working environment. It might not be acceptable for the LAN to be out of operation for several hours while an engineer is called.

Therefore, if LANs are to be acceptable as the media for integrated communications in the factory for example, steps must be taken to enhance their reliability. In particular it must be ensured that no single failure will cause the

LAN to break down completely, and that all such failures are detected so that they can be repaired before they can be combined with later failures to cause a complete breakdown.

In other areas, the significance of a reliable LAN may be even more crucial, especially if it involves human life such as applications in process control or the nuclear industry. Other examples where reliable LANs are essential include applications in the military, and spacebound satellites. In the later case, the system also has to function where human intervention for maintenance or repair is impossible. Ring networks are especially vulnerable to reliability problems. In contrast to bus LANs, a single node failure could disrupt the entire network. Ring components may be designed to be extremely robust to reduce this problem but there is always an unpredictable element. Ring cables may be accidentally severed by human error. There is thus a need for reliable ring networks.

2.2 Reliability

Reliability is defined as the ability to perform a specified function under specific conditions for a specified time. Reliability can be viewed at several levels. In one, which is a basic network goal, is to provide high reliability by having alternative sources of supply. This by definition is the very reason why networks were invented in the first place. With unconnected computers, if a machine goes down due to hardware failure, the users are out of luck even though there may be substantial computing capacity available elsewhere. With a network, the temporary loss of a single computer is much less serious because users can often be accomodated elsewhere until the service is restored. This is the computing service level.

The level this thesis addresses is yet another level below the computing service level. This is the communication level. Here, if the node to which the computing device is attached to is down, the computer is unfortunately taken down as well. This is minor compared to the problem when the communication cable is severed. The whole network may be completely down. This is therefore a primary problem. For military, banking, industrial process control, and many other applications, a complete loss of computing power for even a few hours due to some catastrophe, natural or otherwise is completely intolerable.

The computer network can prevent such failures at the computing service level. But reliable networks adds yet another dimension to the security of computer users, the end result is ideally a system which does not stop working.

Reliability can be achieved in two ways. Fault avoidance requires the physical components and their assembly techniques to be as nearly perfect as possible. The drawback is that the cost of obtaining near-perfect components is excessive and that manual maintenance must be continously available because the system ceases to operate upon first failure. A better alternative is fault tolerance. In a fault tolerant system, redundant components allow the system to continue to operate when some components fail.

Fault tolerance involves five major issues: fault detection, fault location, fault diagnosis, fault isolation and fault recovery.

Fault detection is the ability of a system to recognise that a fault has occurred. Fault location is the system's ability to determine where the fault has occurred. Fault diagnosis should uncover the type of fault so that the appropriate recovery procedure can be implemented. Fault isolaton is the process of isolating the fault and preventing its effects from propagating throughout a system. Finally, fault recovery is the system's ability to regain operational status in the presence of faults.

2.3 Characterisation of Faults

Faults may be characterised into three classes: First, there is the permanent fault which remains in existence indefinitely if no corrective actions are taken. In the Cambridge Ring, there are two such causes: node failure and link break. Both these faults would disrupt ring operation completely. The first failure is relatively unlikely and can be controlled to a considerable extent by good engineering. In fact the range of Cambridge Ring equipment has proved to be extremely reliable (Spratt 80, Binn 82). The second failure is likely to occur in an uncontrollable manner, such as when a technician accidentally cuts the cable.

Second, there is the transient fault, which may appear and disappear within a very short period of time. Interference is one cause which results in one or more ring packets being destroyed. Transient faults are usually environmentally induced, and since they occur only occasionally, it is not economically worthwhile to try to prevent them.

The third class of fault, the latent fault may be data dependent. Its effect appears only at certain times and under certain conditions, and it is really an engineering problem. It is similar to and will be treated as a transient fault.

This thesis addresses the two permanent faults of link break and node failure, and attempts to prevent them from disrupting normal ring operation. Transient and latent faults are detected but nothing will be done about them except to report their occurrences.

2.4 Techniques of Fault Tolerance

Fault tolerance can be defined as the ability of a system to continue to perform its specified tasks after the occurrence of faults. The key ingredient in all fault tolerance techniques is redundancy. Redundancy is simply the addition of information, resources, or time beyond what is needed for normal system operation.

Redundancy may take several forms, including information, hardware, software and time redundancy. An example of information redundancy is the error-detecting code, formed by the addition of information to the basic data structure. The redundant information allows valid and invalid codes to be distinguished. Perhaps the simplest form of error detection coding is the single-bit parity check. The idea behind parity is to concatenate an additional bit to every binary data stream so that the resulting code is forced to have either an odd or even number of ones. If the parity bit achieves an odd number of ones, it is called "odd parity"; if the parity bit achieves an even number of ones, it is called "even parity." A relatively simple check of the number of ones in the data stream allows single-bit errors to be detected.

Another form of error detection coding is the checksum. Checksums are most applicable when blocks of data are to be transferred from one point to another. A simple technique is a "single-precision checksum," formed by adding all binary data that are to be transmitted and throwing away any overflow. An improved variation is the "double-precision checksum." Assuming the number of data bytes to be added is appropriately limited, overflow will not occur, and the information contained in the carry bits is not lost. Fault coverage can be substantially improved with the double precision approach. At the receiving point, the checksum is formed again and compared to the checksum that was generated at the transmitting point. Any discrepancies indicate an error during either transmission of the data or regeneration of the checksum.

Several other error-detecting coding techniques exist, but they will not be covered here. The interested reader should refer to the wide number of published texts on this subject. Parity bit checks and double precision checksum techniques are employed in the research work. They are simple, and they can be implemented in software at minimal cost but most of all they are good enough for the application.

Hardware redundancy is perhaps the most common technique used in fault tolerant systems. There are three forms. First, passive replication methods mask the occurrence of faults and do not offer detection, isolation or repair of a faulty module. Second, active replication methods do not mask faults, but detect and locate faults so that a spare component can be switched in to replace the faulty component. Third, hybrid methods combine the attractive features of both passive and active techniques. It uses fault masking to prevent the fault from affecting the system and fault detection to allow a spare module to be switched in to replace the faulty module.

A common passive method of redundancy is triple modular redundancy(TMR). The purpose of TMR, when used in a passive environment is to mask single faults by triplicating hardware and voting on the results. See Fig. 1. The voter examines all results and generates as the output, what it judges to be the correct result.

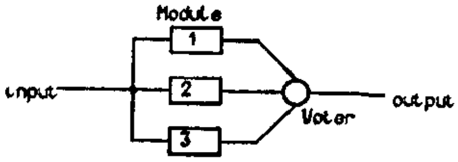


Fig. 1 Triple Modular Redundancy

The active redundancy concept in hardware replication techniques attempts to incorporate fault detection and fault recovery into the system at the expense of eliminating the fault-masking capability. One example is a simple duplication scheme that compares the results of two systems and generates an error message if a disagreement occurs. In the event of a disagreement, the system only reports the error and does not recover from it. See Fig. 2.

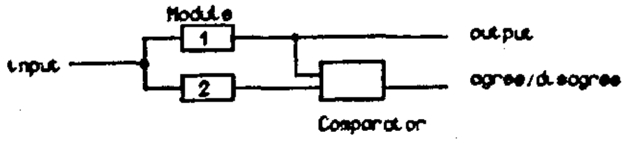


Fig. 2 Duplication of Processing Resources

A second technique of active replication is standby replacement. In this configuration, one unit is operational while one or more units serve as standbys. If a failure is detected with the on-line unit, it is removed from operation and a spare unit replaces it. This technique brings the system back to full operational capability after a single failure, but a disruption in processing is necessary while the spare is brought up. If this disruption in processing cannot be tolerated, "hot-sparing" may be used. In this technique, the spare operates in synchrony with the on-line unit and is prepared to take over at any time.

The final active replication technique combines the duplication method and the standby replacement method. Here two units perform the same computations and a comparison of the results take place. In the event of a discrepancy between the two units, a spare is activated. This is also known as the "pair-and-a-spare" technique.

The basic concept of hybrid replication techniques is to combine the attractive features of passive and active methods to generate a system that has fault-masking, fault location, and fault detection as well as standby replacement capabilities. Several methods exist (Losq 76), but one of the most important is N-modular redundancy with spares (NMR). NMR hybrid redundancy uses N modules to create a voted output with a pool of spare resources. The system remains in the basic NMR configuration until the disagreement detector determines that a faulty unit exists. Once identified, the faulty unit is switched out and replaced with a spare, thus the reliability of the basic NMR system is maintained as long as the spare pool remains unexhausted. Voting always occurs among the active units, masking faults and ensuring continous error-free computations.

This brief survey covers the broad range of techniques available to implement hardware redundancy. In fact the techniques employed in the research uses the simpler methods. Standby replacement techniques form the basis of hardware redundancy to enhance reliability of the transmission medium. This is much less expensive when compared to the NMR concept but is certainly effective enough for the particular application. However the more complex techniques might be useful for the critical components of the Cambridge Ring system : the Monitor, Error Logger, power supplies and the Nameserver. Failure of any of these may be catastrophic to the operation of the Cambridge Ring system. Ring nodes are however not duplicated. The substantial cost involved is not worthwhile. They are simply isolated in the event of a failure to prevent them from affecting the operation of the rest of the ring.

Time redundancy can be used to detect transient error conditons in a system. When a fault is detected, two situations exist. First, a permanent fault may have occurred, and the correct course of action would be to isolate the faulty component. On the other hand, the fault may be transient in which case the hardware is healthy, and it would be a waste of resources to immediately shut down the processor. Time redundancy can be employed to distinguish between permanent and transient failure. For example, a timer may be set on detection of the first fault. If further faults are detected within a certain time period, then a permanent fault condition must have occurred, otherwise they are assumed to be transient.

Software redundancy is simply the addition of extra software to provide some fault tolerant features. This type of redundancy ranges from a complete duplication of software to the addition of small programs to perform validity checks. These techniques are often used to provide protection against software faults that may be present in a system.

One common technique is the validity or reasonableness check. Here additional software is added to verify that the results being produced are within certain ranges. Another type of software redundancy is the periodic self-test. Often, a large pecentage of faults can be detected by allowing software to periodically exercise the hardware and set a "watchdog" timer if the test is passed. The timer is designed to generate an error interrupt if it fails. A third type of software redundancy is the use of multiple copies of programs. To be effective, the programs are written by separate teams to protect against the occurrence of common problems. The basic idea is that multiple programs will perform the same tasks but might use different methods or at least different lines of codes. The multiple versions run simultaneously sequentially. The results are compared to provide a means of fault detection.

CHAPTER THREE

THE CAMBRIDGE RING NETWORK, A DESCRIPTION

3.1 Basic Operation

The Cambridge Ring system is based on the slotted ring principle so called because the bandwidth is divided into a number of fixed-sized packets or slots. Fig. 3 shows a conceptual model of a slotted ring.

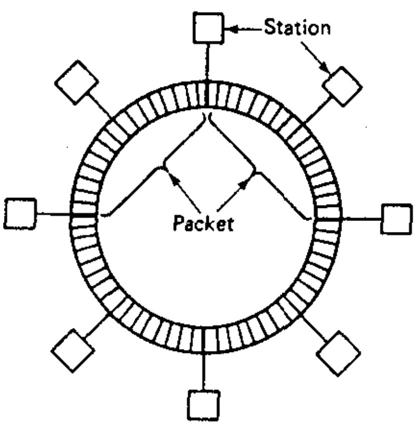


Fig. 3 A slotted ring

Unless the physical distance around the ring is very large or there are many nodes, it is unlikely that there will be enough delay to hold several packets, so artificial delays are needed. These can be obtained by putting shift registers into the ring interfaces. To send a message from one node to another, it is necessary to wait for an empty slot to come around, mark it as full and then load the destination address and data into the slot.

The Cambridge Ring uses a more elaborate scheme to the above. The structure is shown in Fig. 4. At the lowest level the communication link comprises a closed ring of cable and active repeaters. The repeaters are used to regenerate the signals which transmit information round the ring. They may also be used to connect a device to the network.

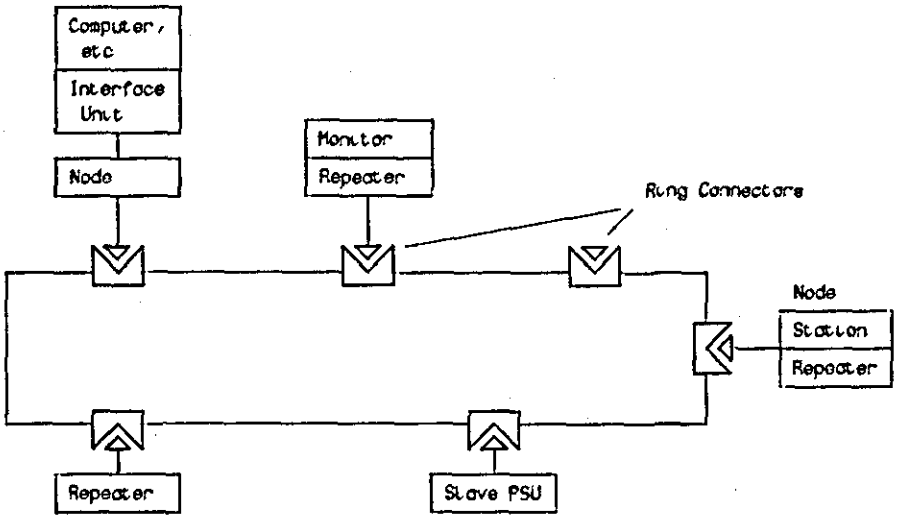


Fig. 4 Example Cambridge Ring System

The station interfaces the repeater to the interface unit of an attached device. This combination of repeater and station is known as a node. Communication takes place between nodes under the control of the host device. Basically, the host device calls upon the node to transmit and receive minipackets (m-p). The m-p is the basic unit of transmitted data between nodes and occupies exactly one slot. Each m-p (Fig. 5) is individually addressed, carrying eight bit source and destination addresses. Two bytes of data are carried together with several bits of control information. The first bit of every m-p is the leader bit indicating the start of the packet. The second bit is the full/empty marker, used to control access to the slot. This is followed by the destination address, source address, and two data bytes. The two control bits act as response bits. The last bit ensures the integrity of each m-p. The node checks and corrects the parity of all passing m-p, and any errors detected are reported.

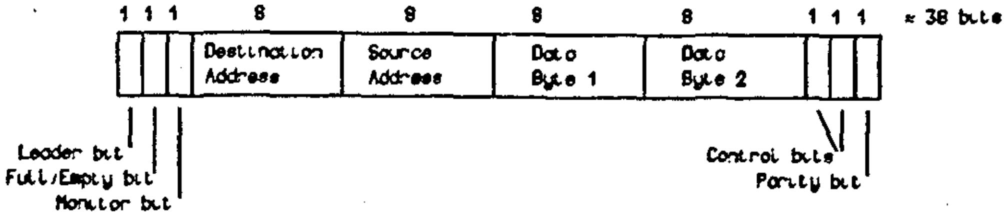


Fig. 5 Minipacket Format

The current implementation of the Cambridge Ring operates at 10MHz. The available bandwidth is subdivided into slots (ie. m-p) which continously circumnavigate the ring head to tail separated by one or more gap digits. These gap digits act as padding to permit an integral number of slots. The total number of m-p is constrained by the propagation delay in the cable and nodes. At 10Mhz, each 100 metres of cable causes a delay of 450 nanoseconds and so may be thought of as storage of 4.5 bits. Each node has a delay of bits Each m-p requires 38 bits so a system with 12 nodes for example, allows a maximum of 2 m-ps with a gap of 14 bits. [(38x2) + 14 = 90]. This itself represents a great deal of wasted bandwidth. To improve bandwidth, a shift register in the Monitor can superficially lengthen the ring to adjust the number of gap bits.

The Monitor is central to the ring's operation, it initialises and subsequently maintains the m-p structure. At power on, the Monitor enters start mode and the frame structure is established. During start mode m-ps with the first two bits set are circulated. These full m-ps ensure that already synchronised nodes do not attempt to transmit until all nodes are synchronised. After ten seconds, assuming no error conditions exist, run mode is entered. The ring is now operational.

A node wishing to transmit waits until an empty slot arrives; it then marks it as full, inserts the destination and source addresses, the data, and finally initialises the response bits. The transmitter may only transmit one m-p at a time. The transmitted m-p then goes round the ring to its destination where the control bits are set on-the-fly to indicate busy, rejected, ignored, or accepted. In the latter case the data are copied into the destination node. The m-p now returns to the source where the full/empty bit is reset to zero; thus emptying the slot. This returned m-p is also checked with the original m-p transmitted to ensure that no errors have occurred during the transmission. If there are no errors the response bits are noted, the m-p freed for use by downstream nodes, and the host device informed of the m-p's return.

The response bits are used to carry flow control information back to the transmitter.

Bit:	37	38	Information
	0	0	busy - the destination node acknowledges but has not read the m-p because the interface unit has not yet signalled the node that it has finished with a previous m-p.
	0	1	Accept - the destination has read the m-p.
	1	0	Not selected - the destination acknowledges but has not read the m-p because its source selector is not set to 255 or this address does not exist.
	1	1	Ignored - no node acknowledged the destinnation address.

Whenever a transmitter receives a response other than accepted, it is not allowed to transmit immediately but forced to wait for the ring structure to cycle around. The second and further unsuccessful transmission tries cause the transmitter to be backed off for 15 ring cycles. This prevents the ring being swamped with useless traffic. The round robin scheduling puts an upper limit on this delay while the variable backoff produces a system in which efficiency improves under load.

3.2 Reliability and Maintainability Aspects of the Cambridge Ring

The designers of the Cambridge Ring system have realised the potential reliability problems, and have included in the basic design, built-in capabilities for localising errors and failures. Although this is no solution, it does help to locate the faults quickly and thereby enables an engineer to track and resolve the faults more easily.

3.2.1 Parity checks

First, consider the m-p structure. Every m-p includes a parity bit that is continously checked and maintained by all nodes. Each node computes the parity of every passing m-p and if the generated parity does not match the old, a fault has occurred. This is corrected in-situ and a fault message is transmitted in the next empty m-p to destination zero. This message contains the address of the sending node and so indicates the section of the ring where the fault occurred. The fault message may itself become corrupted giving rise to further valid fault messages; nevertheless, the indicators reaching the Monitor will at least be correct for the nearest faults. The Monitor indicates the fault by an error count indicator and an error source indicator.

3.2.2 Maintenance functions of the Monitor

The m-p includes a bit, the Monitor Pass bit that is set by a transmitter when it fills a m-p. This bit is always cleared by the Monitor on passing m-p. If the Monitor detects a m-p that has this bit cleared but is still marked full, it marks the m-p empty. It is thus impossible for a fault to cause a m-p to become permanently full.

The Monitor is able to detect errors which interfere with the permanent m-p structure and then to rapidly re-instate the correct structure. A large number of errors such as one caused by a power dip may cause the m-p structure to lose synchronisation. This causes the Monitor to re-initialise the network automatically to its original state.

During operation, the Monitor injects random data into empty m-p and checks that it returns uncorrupted unless the m-p has been used in the meantime. Thus error checks are performed continously even if there is no user data on the network. In these ways the Monitor keeps the performance of the ring under continous surveillance and can give warning of incipient faults.

3.2.3 Checking of returning packets

A transmitter counts slots when a m-p is transmitted. Thus, a returning m-p is recognised even if its source address has become corrupted; the packet is then compared bit by bit with a copy of the one transmitted. If any discrepancy is detected, a transmission comparision error is reported to the interface unit.

3.2.4 Ring breaks

The schemes so far locates transient errors or node failures. They can also be used to detect ring breaks, providing that the repeater continues to operate with no signal on the ring input cables. In fact, the ring repeater has a phase locked loop which continues to operate in the centre string of zeroes, and the node is made to transmit a repeated fault message packet. Therefore a break in the cable is detected by the next active node downstream.

3.2.5 Others

Precautions are taken to reduce the probability of repeater failures. The part of the repeater which needs to operate is made as simple as possible. Power is also supplied to repeaters around the ring from a power supply unit, so that operation does not depend on power supply from the station.

3.3 Errors and Actions Taken

A response error, for example if a busy response is changed to accepted may cause that m-p to be lost. These errors and others may be induced by transient errors. In this case a checksum may be used at a higher level to detect and correct it. Another transient induced error may corrupt the addresses of the m-p. The transmitter detects this by counting slots. Similarly, data errors may occur and this will be detected in the same way. All these errors should cause a parity fault to be detected at the next node which in turn sends a fault message to the Monitor. Thus most errors are recorded.

Another class of errors is the framing error. The loss of the first bit of a m-p or the change of a gap digit is called a framing error. It causes nodes after the error and before the Monitor to become unsynchronised and usually causes consequential errors. The Monitor will re-enter start mode until framing errors cease so that the ring is rapidly resynchronised, and synchronised nodes are inhibited from using the ring until that is complete. The slot structure is automatically re-created for one complete cycle and enables all the other nodes to recover. After 128 framing errors have occurred, the Monitor re-executes the complete initialisation cycle. This is particularly useful if a 'burst' of framing errors occur.

CHAPTER FOUR

PRELIMINARY INVESTIGATIVE STUDIES

4.1 Review of Literature

It is a well known fact that ring networks are vulnerable to faults in the transmission path (Liu 84, Clark 81). Yet the Cambridge Ring makes no real attempt to avoid reliability problems (Wilkes 79, CR 82). This is not confined to the Cambridge Ring, in fact Moore, Geer and Graf (Geer 84) suggested that few if any existing networks could survive wide-scale disasters. They felt that most of the current fault tolerant networks are:

(1) vulnerable to becoming fragmented in a hostile environment,

(2) incapable of handling problems when link outages occur, or

(3) incapable of ensuring message delivery in a hostile environment.

Other authors made similar arguments to the above three problems. Kirk (Kirk 84) suggesting that a fault tolerant design must ensure:

(1) no single failure disrupts all communication,

(2) no latent failure remains undetected and

(3) can be extended without disrupting all communications.

In his paper he stressed the importance of a reliable network stating that "if LANs are to be acceptable as a media for integrated communications in the office or factory, steps must be taken to enhance its reliability." Also, another serious problem he raised was that of medium failure, for little protection can be given against accidental damage to the cable and repair is likely to be time consuming. Kirk's third point is interesting. One reason for installing a LAN is the ease of expansion. To add a new node, the entire ring installation would have to be shut down before physical work could be carried out. Only when this has been completed can ring operation be restored. This procedure may take a few hours or a few days and during the entire period, no computing service would be available to users. Clearly this is both inconvenient and undesirable.

These problems have been taken a step further by Saltzer and Pogran (Saltzer 8Ø) in their work at Massachusetts Institute of Technology (MIT). They stated that a more significant factor to take into account when selecting a network may have little to do with the issue of the best technology. Instead some site-wide networks of say 1ØØØ computers will consider more mundane issues such as which technology is easiest to install, reconfigure and maintain. If a basic ring is used to connect up a large number of nodes, a single failure might result in a

"Christmas tree" effect, that is, locating a burnt out bulb requires checking each.

Fortunately, the effects of many faults can be reduced by adding fault tolerance to the basic ring topology. The following paragraphs will present a brief survey of the current state of research in the fault tolerant ring network field. Section 4.2 will describe and evaluate each one these techniques in greater detail.

Falconer (Falconer 84) suggested that a catastrophic ring failure will occur if a fault occurs in one or all of the following:

- the node
- the ring's central controller
- the transmission cable

Protection against non-catastrophic failures would in most applications be uneconomic. The highest risk is usually considered to come from node fault or cable break, and it here that the greatest number of options lies.

One of the earliest solutions was the use of Bypass Relays (Kirk 84), provided at each repeater. Should any repeater fail, the relay associated with that repeater would switch it out of the ring. This technique however does not take into account link breaks.

Another early design was the Dual Ring (Falconer 84). In contrast to the bypass relay approach, this method is worthwhile only when the cable is the most probable source of failure. Should a link be severed, operation would continue on the spare link cable. However no provision is made for repeater failures.

Duplication of the Ring (Weitzman 80) is perhaps an improved technique. In this method, the entire ring is duplicated so that a fault detected on the primary ring will cause a switch of operation to the secondary ring. This is unfortunately the extent of its use since another fault will disable the ring. Further levels of duplication are possible but they are economically unrealistic.

Perhaps a more elegant approach is provided by the Self-Heal ring (Zafiropilo 74). This technique uses spare links passing all the way around the ring with data transmission in the opposite direction to the normal links. Repeaters are able to detect when there is a break in the ring upstream to them, if one is detected, the repeater will 'loop back' i.e. take its input from the spare link entering it from the opposite direction. The repeater must then signal to the next repeater downstream from the break to loop back in a different way such that its output is sent back along the spare link. Racal's Planet Cambridge Ring system is an application of this design.

So far the techniques reviewed are restricted in that they allow only a limited number of failures. The next few examples presented here address this problem.

The Hierarchical Multi-Loop system forces the partition of a network into several sub-sections. In this way, faults will be isolated to the sub-loop where they occur rather than affecting the entire installation. J.R. Pierce (Pierce 72) proposed a three-stage network to improve its reliability.

Availability may be further enhanced by mesh-like topology, developed by Hafner called a Braid (Hafner 76). Besides the ring connections, each repeater have one or more extra links connecting it to other repeaters. Thus, a signal may have a choice of several paths when the main ring is down. It does however require much more cabling, and to be effective the extra links should be routed through separate ducts. It is also known as the Mesh network.

The design originating from MIT (Saltzer 80) reduces maintenance problems. The Star-shaped Ring uses the concept of a wire-centre, where all inter-repeater cables must be arranged so that they always loop back through a single room. It can be visualised as a flower with petals for every ring node. This improves maintenance since troubleshooting is centred in only one area. A node failure is isolated by bypassing the petal to which it is attached. Its main disadvantage is the great length of cabling required.

Several other techniques have been designed but they are variations, combinations or improvements on the established approaches presented here. Gridnet (Graf 84) is one such example, designed for a large network where network survival is the primary aim. The architecture reflects both dual loops and mesh topologies. The salient feature of this network is its adaptive routing technology which uses distributed processing to establish alternate routes between pairs of nodes. Another recent design originating from the Marconi Research Centre (Kirk 84) has elements of the Star-Ring and Mesh topologies. It has been designed to overcome multiple faults and to ease network expansions.

4.2 An Evaluation of Fault Tolerant Rings

This section analyses a range of fault tolerant ring designs. Section 4.2.1 presents the evaluation criteria. Section 4.2.2 surveys and compares the various approaches, culminating in the results and conclusions presented in section 4.2.3.

4.2.1 Evaluation Criteria

Various design decisions and goals were highlighted from the brief review in section 4.1. They are summarised below. The following properties will be used as the basis for comparisons:

(a) Availability

(b) Degree of resilience (including fragmentation)

(c) Expandability

(d) Ease of installation

(e) Ease of maintenance

(f) Cost

Reliability requirements are often expressed in terms of system availability : a system will be expected to perform with a certain maximum allowable time out of service during a particular time period. Two factors are important for system availability :

(1) how rapidly the fault can be repaired

(2) how often the system fails

Consider the first point. Before the fault can be repaired, the cause must be isolated. Unless the system is well designed (and documented) this can be a lengthy and tedious process. The second point implies that system failures should be kept to a minimum in the first place.

A well designed system increases system availability because it is easier to maintain. Further, if system availability is to be increased, it must be designed to provide maintenance personnel with as much diagnostic information as possible in the event of failure.

The degree of resilience depends very much on the application. Resilience is likely to be less significant when the network is installed in a University environment where its use is less critical and users can tolerate occasional faults. However resilience becomes more critical in the factory or in defense (eg. battleship) applications. Here lives may be endangered or vast sums of money may be lost due to any stoppages. Perhaps it is most significant in situations where maintenance is almost impossible, such as a spacebound vehicle. The degree of resilience depends on the number of failures tolerated before the entire system is brought down. Factors such as fragmentation of the network and the ability of the system to withstand multiple failures are significant.

Expandability is the ease with which an existing system may be enlarged. Ideally, when installing new nodes into the network, the ongoing operation of the network is not disrupted. In contrast, an undesirable situation arises when the entire system has to be shut down completely for new nodes to be installed and tested before restoring network services. Non-stop operation would be preferable.

Installing a network depends to a large extent on the cabling requirements. Various questions would need to be asked. Does it require separate cable ducts? How much cabling is required? Are there any restrictions on how they are laid out? These are some factors which will affect the implementation of a network. For example, if a LAN is to be installed in an older building, it is unlikely to have cable ducts designed for todays needs. The building may have been designed only for power cables. Although an extra cable can be accommodated, the ducts might not fit in more. This might rule out certain network topologies. Obviously the less cabling a topology requires, the easier it is to install. Interference from the power cables must also be considered. This might necessitate the installation of a separate duct.

Cost is naturally one of the most important considerations. Designing reliability into a system will certainly incur an additional cost. Cost must be acceptable, although this must be examined in the light of the application.

4.2.2 Description and Evaluation

4.2.2.1 Bypass Relay

A simple technique to allow for repeater failure is to provide bypass relays at each repeater. See below.

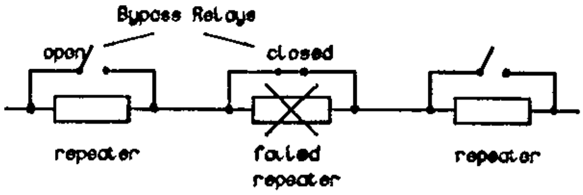


Fig. 6 Bypass Relays

These relays must be actively held on by each repeater to bring the repeater into the ring. If a repeater should malfunction, this active signal is removed causing the ring to bypass that repeater. Because the bypass relay is controlled directly by the repeater, a possibility exists whereby a faulty repeater fails to switch the relay or even worse it may affect the operation of the network. Also, as repeaters fail, arbitrarily long lengths of links without repeaters are found in the ring. Since a repeater drives a limited length of link, the performance of the ring will deteriorate. Its major disadvantage is that this technique is only confined to repeater failures; link breaks will disable the ring. Thus it only provides a partial solution. However bypass relays are simple, cheap and easy to install. Also they allow for multiple repeater failures.

4.2.2.2 Dual Ring

This is one of the simplest fault tolerant enhancement approaches. See Fig. 7.

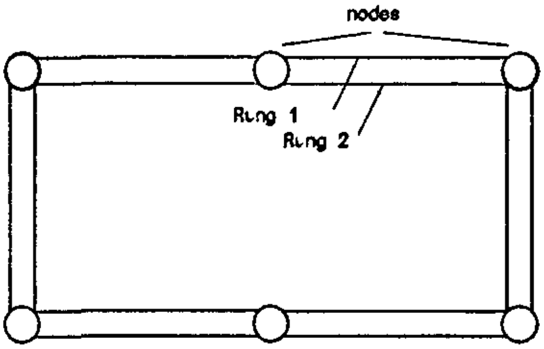


Fig. 7 Dual Ring

An extra link is laid between any two repeaters. If one of these links is severed, the second link is brought into service. However a failure of both cables between any two repeaters causes loss of service. Failure of any repeater will also bring down the network. This technigue is worthwhile only when the cable is the most probable source of failure. Also, to be effective, each cable should follow a separate physical path, thus installation can be a problem. However since only the cables are duplicated, cost is minimal and operation is simple.

4.2.2.3 Duplicating the Ring

Both the bypass relay and dual ring techniques solve only one primary failure. By duplicating the complete ring (Fig. 8), both link break and repeater failure may be rectified. The secondary ring remains in the standby mode until the occurrence of a fault.

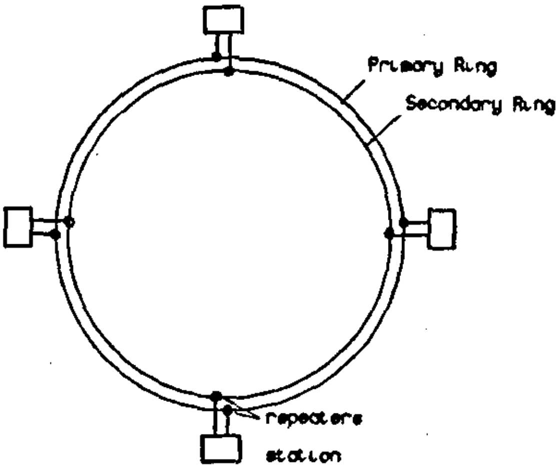
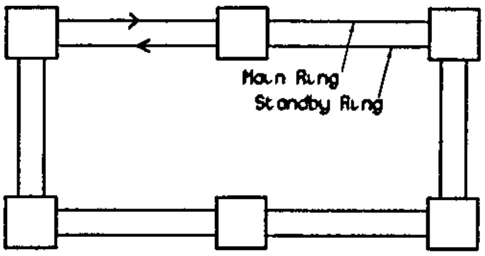


Fig. 8 Duplicate Ring

In the event of failure, operation is switched from the primary to the secondary ring. Because of complete duplication, this is an expensive approach costing approximately twice as much. Moreover it can only manage one fault, a second fault will disable the network. For the same reason as with the dual ring, the second ring should follow a separate physical path. Also since the second ring is not completely independent from the first, faulty nodes might cause unforeseen errors.

4.2.2.4 Self-Heal Ring

The Self-Heal ring is based on a bidirectional double ring structure. A standby cable is installed alongside the main transmission path but is designed to support transmission in the opposite direction. Fig. 9A shows the self-heal ring. Fig. 9B illustrates its action when a link break occurs - repeaters on either side of the break detects the breakage causing them to switch relays, in effect forming a loop-back to isolate the broken link. Similarly, Fig. 9C shows the action taken when a faulty node is detected.



□ REPEATER

Fig. 9A Self-Heal ring

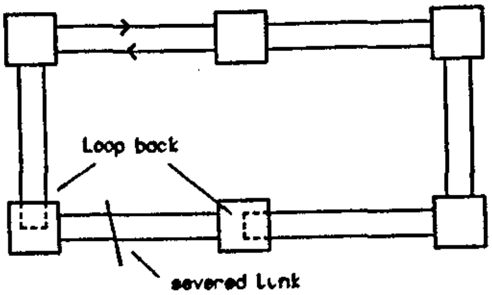


Fig. 9B Bypassing a severed link

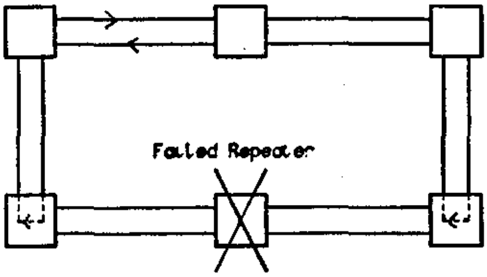


Fig. 9C Bypassing a Failed Repeater

A major disadvantage is the potential partition problem when more than one fault occurs. The network will be divided into isolated segments (Fig. 10). This may be acceptable for a ring which does not rely on a central controller, but in the case of the Cambridge Ring which does, only the segment which has it will continue to operate. The result is a severe loss of service. Another problem is the increased cable length under fault conditions. The maximum distance between any two nodes will be half that of a normal ring. Also, the fault detection and reconfiguration circuitry need to be more complex. Finally, because of incomplete autonomy of fault-tolerant equipment from the ring (usually built into and controlled by the nodes), unforeseen problems may arise with a faulty node. An advantage is the ease of cable installation. The cable pair can be installed in the same duct without any consequence.

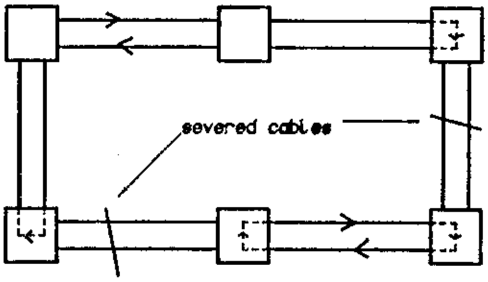


Fig. 10 Fragmentation problem with a self-heal ring: two isolated rings form when two faults occur

4.2.2.5 Hierarchical Multi-Loop System

The Hierarchical Multi-Loop system has several levels of interconnections. As shown in Fig. 11, there is a main loop where the central controller resides. Other loops are then attached to this main loop to which computing devices are connected. One example is a two-stage network, the Collins C-System. This consists of first-level loops linking ring stations. These first-level loops themselves are connected to a single second-level loop. Node failures are thus constrained to the loop concerned, and not the entire network.

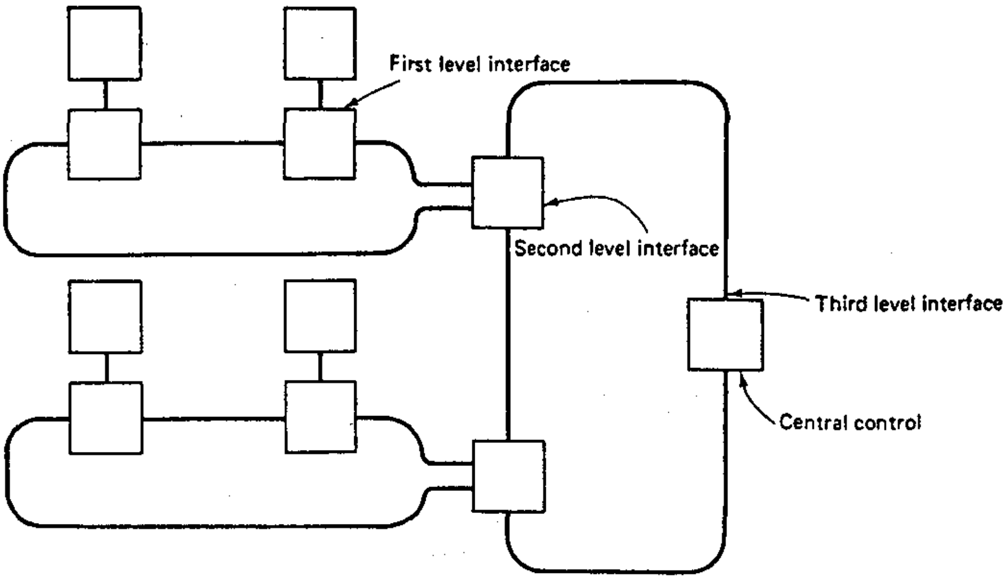


Fig. 11 The two-stage Collins C-System

The number of levels may be further extended. Pierce (Pierce 72) proposed a three stage network. The different levels of loops are connected by special interfaces called C boxes .that transfer blocks from one loop to another. These C boxes also supervise the routing of messages through the various loops to their destination.

In general, reliability improvements obtained by increasing the number of stages beyond three do not warrant the added network complexity or cost.

The obvious drawback of this approach is its complexity - installation can be a problem. Cost is likely to be higher than other methods. Also if any C box should fail, then the entire network will be disabled.

4.2.2.6 Braid or Mesh Network

The Braided network provides a substantial increase in ring availability. See Fig. 12. The idea is to provide higher reliability by introducing link redundancy.

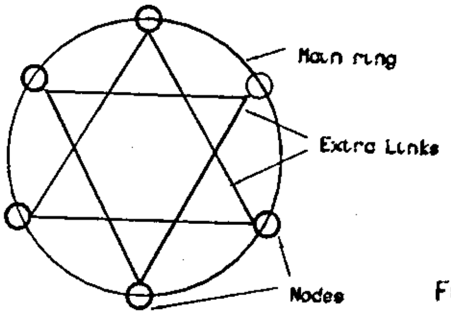


Fig. 12 Braid or Mesh Ring

The outer path is the main ring containing the nodes. Each node may have more than one input and/or output to allow for several alternative transmission paths should any faults develop. Thus this design allows for multiple faults, the maximum number depending on the level of braiding. However, to be effective the different paths should be routed in separate ducts. Thus installation and maintenance may be a problem. Also, this technique requires much more cabling compared to some of the other techniques. The maximum braid length depends on the maximum unrepeated cable length, reducing the maximum distance permissible between nodes. Another disadvantage is that a failure of the main cable may result in a functioning node being bypassed.

4.2.2.7 Forward Loop Backward Hop (FLBH) Network

The FLBH network is an enhanced variation of the mesh network. In this class of network, each node has a forward link connecting to its neighbour and a backward link connecting to a node at some distance s, where s is called the skip distance. Variations of FLBH are obtained by choosing different values of s. In the optimal FLBH network, the parameter s is selected such that the diameter is minimized (Gerla 85).

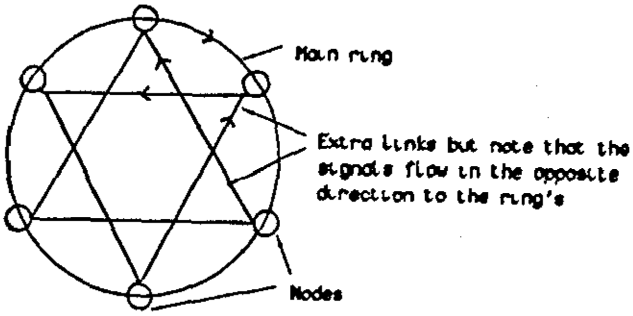

Main ring
Extra links but note that the signals flow in the opposite direction to the ring's
Nodes

Fig. 13 The Forward Loop Backward Hop Network

Both forward and backward links are active, and several paths exist from a source to a destination. This network can tolerate several link and node failures, before becoming partitioned. It improves delay and reliability since the skipping of several nodes creates "short cuts." It has drawbacks similar to the mesh since its topology is really the same.

4.2.2.8 Star-Shaped Ring Network

The Star-Shaped ring suggested at MIT is unique in that it achieves reliability by topological design rather than technological. Referring to Fig. 14, the entire ring is arranged such that inter-repeater cables always loop back through a single room called the wire-centre. The result is a ring network in the shape of a star. The bypass relays are activated to bypass any loop where a node fault or cable break is detected.

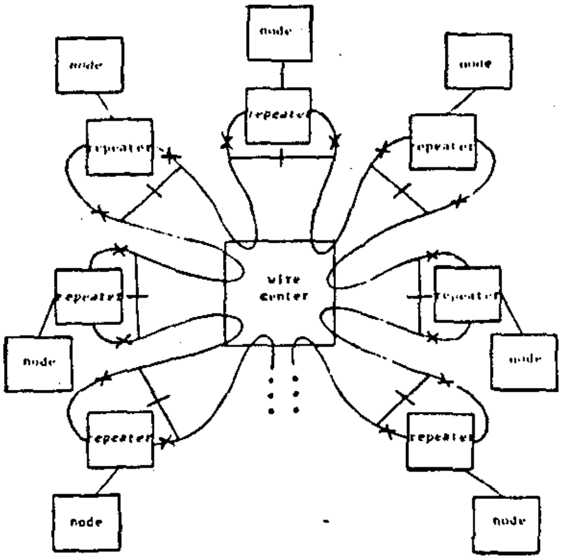


Fig. 14 The Star-Shaped ring

The aim of the designers was to improve maintenance and thus make serviceability easier, achieved through the wire-centre concept. This idea has the advantage of simplicity and the ability to tolerate any number of faults. However, if a link is severed, an operational node is bypassed. But its main disadvantage is the great length of cable required, and the potential installation problem with a large network.

4.2.2.9 Gridnet

Gridnet was designed to survive wide-scale disasters - it overcomes problems of network fragmentation. The approach adopted was to develop many alternate routes for data transmission and by using distributed processing for route selection and communication control. Routing is accomplished independently of any single node or link.

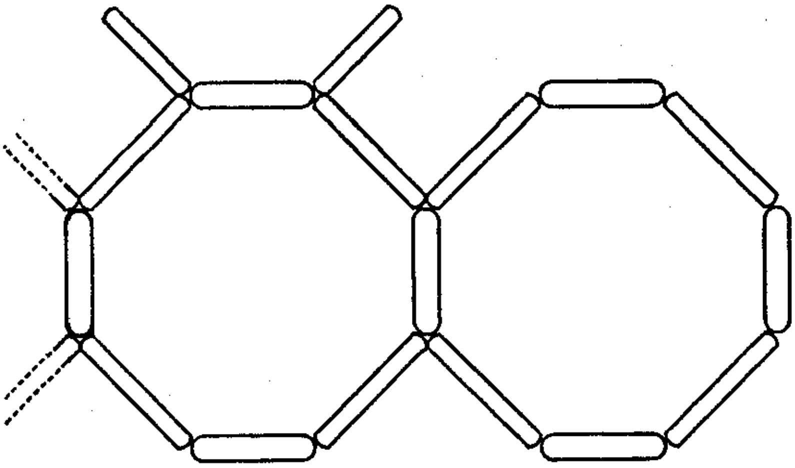


Fig. 15 Gridnet consists of an interconnection of Loops

The architecture is formed by interconnecting a number of dual loops. Each loop within the network is connected to a maximum of four adjacent loops by "gateway" stations. Gateway stations make routing decisions based on their knowledge of the operational status of other loops in the local neighbourhood. By using an adaptive routing technology, alternate paths can be established between distant pairs of nodes despite simultaneous interruptions to the continuity of mutiple loops.

This approach is highly reliable, able to withstand numerous simultaneous faults. However the technique to achieve this is complex and thus expensive. Also, installation may be a problem with its complex topology which also consumes a large amount of cable.

4.2.2.10 <u>Others</u>

IBM proposed an architecture based on the star-shaped ring (Bux 82). Instead of one wire-centre, several are distributed throughout an installation. See Fig. 16. Physically, the ring consists of a set of interconnected distribution panels and lobes radiating from the panels. Wiring from the distribution panels to the nodes is star-shaped. Contained in the distribution panel are bypass relays used to cut inactive or malfunctioning nodes out of the ring. Although this configuration reduces cabling requirements it also removes the protection of link breaks between distribution panels. IBM however is working on techniques of routing and reconfiguration to overcome the link failure problem (Dixon 83).

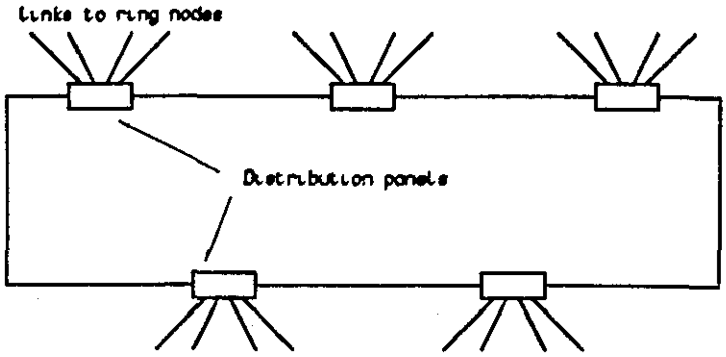


Fig. 16 IBM Token Ring Configuration

Another fault tolerant ring developed at Marconi Research Centre has a mesh-like topology. The network consists of "intelligent" repeaters linked together by an arbitrary mesh of links (see Fig. 17),the only constraint being on the number of links attached to each repeater - in practice three. These links are configured into a ring which passes in both directions along each functional link by a distributed reconfiguration procedure carried out in the "intelligent" repeaters. The result is a network which can tolerate multiple faults. Another feature is the possibility of adding or removing links and repeaters with only momentary disruptions to the network operation. It allows flexibility in layout and expansion.

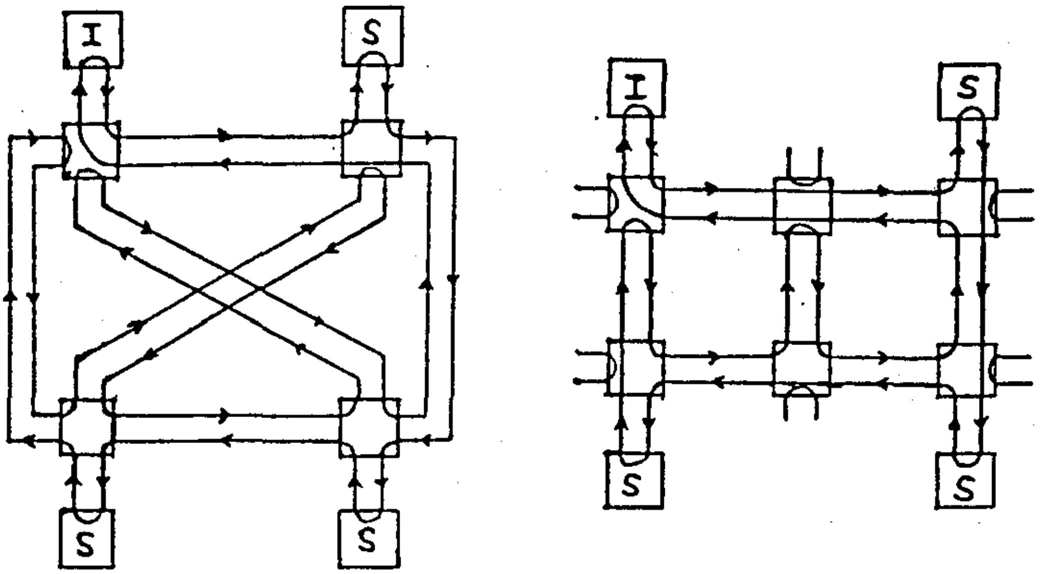


Fig. 17 Two possible configuration of the Marconi Research Fault Tolerant ring.

Table 1 summarises the major features of the fault tolerant rings.

Features	Bypass Relays	Dual Ring	Duplication of Ring	Self-Heal	Braid or Mesh	Hierarchical Multi-Loop	Star-Shaped Ring	Cricket	IBM Token ring	Marconi Research Fault Tolerant Ring
Faults tolerated : Links	NO	YES	YES	YES	YES	YES	YES	YES	LIMITED	YES
Nodes	YES	NO	YES	YES	YES	YES	YES	YES	YES	YES
Tolerate Multiple faults	YES	LIMITED	NO	LIMITED	YES	YES	YES	YES	LIMITED	YES
Network Fragmentation	NO	NO	IRRELEVANT	YES	NO	SOME	NO	NO	YES	SOME
Ease of Installation	EASY	EASY	EASY	EASY	DIFFICULT	MODERATE	DIFFICULT	DIFFICULT	MODERATE	DIFFICULT
Is normal ring operation possible during expansion?	NO	NO	NO	YES	YES	CERTAIN CASES	YES	YES	YES	YES
Is normal ring operation possible in the presence of a fault (s)?	YES	YES	YES	YES	YES	YES	YES	YES	LIMITED	YES
Extra cabling required	NO	X 2	X 2	X 2	>3	>1	>4	>2	>1	>4
Comments	Multiple faults apply only to repeater failures	Multiple faults apply only to link breaks and limited to one per section	Can survive only one fault	Multiple faults will cause fragmentation of network, thus limited	Certain faults will cause healthy nodes to be bypassed	Multiple faults tolerable only if a fault does not occur on the main loop. A single fault will disable a group of nodes on the same loop	Cable installation problematic since it is star structured. A link break will cause isolation of corresponding node	Complex adaptive routing method. Complex topology	Link break between Distribution panels is catastrophic. Ring operation not possible if ring cable needs to be lengthened	COMPLEX TOPOLOGY

TABLE 1 : Summary of Fault Tolerant Rings Comparison

4.2.3 Discussion

The survey has revealed several general points concerning the various design approaches.

(1) All the designs were based on ring, star, mesh topology or combinations of them.

(2) They were designed primarily to solve problems of node failures and/or link breaks.

(3) The levels of fault tolerance are quite different, each has strengths in specific applications.

(5) The more resilient designs have more complex topologies

Less obvious points were also raised. One is the significance of operational independence of the fault tolerant components from the ring equipment. This prevents the ring equipment from causing unforeseen actions on the fault tolerant devices which in turn may cause uncalled actions.

Ring, star and mesh topologies offer different advantages. Ring networks were introduced originally in part to replace star networks. The rationale being that a star network relies on a central controller and thus it is inherently unreliable. Also, cabling requirements of a star network are more substantial which in turn give rise to installation and maintenance problems. Likewise mesh type networks which form the topology of telecommuniation networks were frowned upon by LAN designers for similar reasons.

However those were the days when LAN technology was only just emerging; where technicalities and performance were all important. As the technology matured, other user oriented factors were becoming more significant. Installation, maintenance and reliability became more important factors to be considered.

The basic reliability issues of networks were questioned. Although fully distributed ring LANs such as token rings do not rely on a central controller, they do rely on a single length of cable. An analogy can be drawn between this single transmission path and reliance on a central controller. In the light of this, the basic star topology offers unparalleled advantages as far as the cable factor is concerned. Any number of cable breaks leads only to those nodes attached being disabled. Faulty nodes can similarly be removed from the network with simple algorithms.

The present trends in fault tolerant designs can be seen to reflect the above argument. Increasingly, star and mesh-like topologies are being used to supplement a basic ring structure. MIT's Star-Shaped ring, IBM's token ring and Gridnet are typical examples. Some designs are extremely resilient but correspondingly they have their drawbacks. Gridnet and Marconi Research Fault-Tolerant ring are examples. Their mesh-like cabling requirements are very high while employing complex routing algorithms to manage communication and reliability. These complexities translate to higher cost.

4.3 Objectives for the Development of a Reliable Cambridge Ring

The basic objective was to resolve the reliability problem of the Cambridge Ring system. The following design goals were established to guide the development process.

4.3.1 Design Goals

(1) Fault tolerance

The ring must not fail when a node fails or a cable is severed.

(2) Degree of Resilience

The ring must tolerate a large number of faults. This includes the situation when a fault has not been rectified before another occurs.

(3) Reduce Fragmentation

When there are several faults in the ring at any one time, they should not cause partition of the network. This is particularly significant for the Cambridge Ring since it relies on a central controller.

(4) Ease of Installation

The way the system is installed, in particular the layout of cables, is significant. The number of cables and how they are laid depends on the network topology. Other questions which have to be considered include :

(a) Can the cables be installed in existing cable ducts?

(b) Does it necessitate a false floor being built?

(5) Ease of Maintenance

When a fault occurs, the ease of finding and repairing it quickly makes the network easier to maintain. Since the ring consists of serial active components, the problem when one fails is like finding the burnt-out Christmas tree light. Trouble shooting may require visiting each node with test equipment. Clearly this is undesirable.

(6) Ease of Reconfiguration and Expansion

In the future the network will probably undergo expansion. Preferably this task should not cause too many disruptions to users of the ring. The same should apply to reconfiguration of the network when required. Facilities should be made available to ease the tasks. Proper documentation of the installation is essential.

(7) Operational Independence

Operation of the fault tolerant component should be independent from the ring equipment to reduce the probability of unforeseen interactions.

(8) Take into account Future Developments

LAN technology is evolving fast. To prevent obsolescence the design should take into account future developments such as higher transmission speeds and the use of optical fibres.

4.3.2 Design Constraints

(1) Cost

The fault tolerant enhancement cost to the ring should be within reasonable limits.

(2) Minimal modifications to existing Cambridge Ring equipment

The design should require none or minimal modifications to the ring equipment.

4.3.3 Design Features

From the objectives, the design should incorporate the following features.

(1) Bypass technology

Relays should be used to bypass faulty sections of the ring. This way faults are isolated completely until they are repaired and put back on-line.

(2) Automatic operation

Faults should be isolated without the need for human intervention. This way, faults are corrected in real time and eases maintenance by allowing the servicemen to re-instate the faulted system to its fully operational state at a later time.

(3) Off-line redundancy

Off-line redundancy is a technique whereby equipment is initially quiescent until it is required. Mathematically, this improves reliability by 23% (this proof can be found in many standard texts on reliability, see for example Frankel 84).

(4) Operator control

It should be possible for an operator to control the system. Thus network configuration may be altered, for example a node may be bypassed, and removed for attention before replacing it back on-line. Likewise a section of the ring may be isolated for expansion. All these should be carried out without too many disruptions to normal ring operation.

(5) Reporting facilities

Summary reports of the status of the network such as which nodes or links are faulty and/or bypassed are essential for the management of the network. This may help with future expansion and maintenance of the installation.

CHAPTER FIVE

DEVELOPMENT OF THE HIERARCHICAL RING-STAR SYSTEM

5.1 Summary

A new fault tolerant design has been developed for the Cambridge Ring based on topological enhancements. It has been designed for a high degree of resilience, able to tolerate multiple faults without any significant partition problem. In addition, provisions are made to ease installation, maintenance, and expansion of the network in a way to facilitate non-stop operation. Installation in particular requires only minor modifications to the present range of Cambridge Ring equipment.

The proposed Ring-Star concept has a topology which can be visualised as a ring with a star structure superimposed within it. It is this star structure, with extra links connecting to each node which provides fault tolerance to the ring. In effect, the star centre contains bypass relays which allow alternative paths to be switched dynamically according to a fault algorithm to isolate broken links or faulty nodes. It must be pointed out that while this is being carried out any ring data circulating at that moment will inevitable be lost due to the relay switching operation. Higher level protocols should be able to recover any lost data.

This basic Ring-Star concept has evolved a step further as a result of the acknowledged cable installation problem with any star-structured network. By distributing several Ring-Stars throughout the installation and connecting them up in a "star-within-a-star" topology, the advantages of the basic Ring-Star concept can be applied practically to a larger installation. The Hierarchical Ring-Star forms the final proposal as a result of this research.

This design is not limited to the Cambridge Ring. It can be adapted for any other ring network, and is suitable for installation in buildings or clusters of buildings within a site.

5.2 Introduction

Although a wide variety of fault tolerant ring configurations have been developed, few have been designed for the Cambridge Ring. Racal's Planet self-heal ring is one and there are simple implementations of the star-shaped ring (CAMTEC is one commercial organisation which suggest such a configuration to clients).

Chapter 4 details the disadvantages of the self-heal design and in particular the severe partition problems with a ring controlled by a master station. Because of its limitations with multiple faults, a better technique is sought. The star-shaped ring implementation of the Cambridge Ring although alleviating the partition problem is deemed unsuitable because of likely installation difficulties.

Section 5.3 describes the conceptual development of the Ring-Star. Technical design of the Ring-Star was carried out in two phases, discussed in section 5.4, 5.5 and 5.6. The first developed a prototype Ring-Star, which after evaluation led to the development of a second design. The key difference is the evolution of a single unit Ring-Star into a multiple Ring-Star architecture.

5.3 Conceptual Development

Chapters 2 and 3 have provided the background information, in particular the technology of the Cambridge Ring. Of relevance here is the reliability and maintanability aspects, the likely errors and the way the Cambridge Ring has been designed to cope with them. In brief, the Cambridge Ring has been designed to reduce the occurrence of bit errors, and when catastrophic faults occur, to provide information pinpointing the source quickly. It makes no attempt to correct the fault.

The literature research in chapter 4 has identified key design issues which resulted in a set of objectives for the development of a suitable resilient ring. It is restated here for convenience.

Objectives & Constraints:

- resilient to repeater and cable faults
- tolerate a large number of faults
- reduce fragmentation problem
- ease of installation, maintenance, expansion
- operational independence
- take into account future developments
- cost
- minimal modifications to existing Cambridge Ring equipment

The last point needs clarification. The question of designing a fault tolerant system to operate on existing equipment or instead to design an ideal system even if it requires that existing equipment has to be re-engineered. The second approach is scientifically more exciting with freedom to explore new ideas without restrictions and perhaps come up with the "best" design. The first approach however makes more commercial sense but restricts ideas and techniques to fit existing equipment. Considering the investment already made in the Cambridge Ring equipment, the first approach was decided upon.

Current statistics on the reliability of Cambridge Ring installations (see Binns 82 for example) indicate very reliable Cambridge Ring equipment. Failures of repeaters have been almost non existent, instead most major sites

(Kent University and University College, London) reported faults with the devices connected into the ring and human errors such as the accidental severing of ring cables. There were intermittent errors but no major failures. Another irritation was that installations have to be shut down everytime routine maintenance, expansion or contraction of the network was required.

Thus a suitable design should give special considerations to cable break problems and maintenance/expansion facilities. In particular, when work is required on the ring, the entire installation should not need to be shut down completely.

5.4 Developing the Basic Ring-Star

The ring topology has a very weak structure, because it has one and only one transmission path. Therefore, to improve the reliability, its topology must be enhanced.

The approach taken was to develop first a topology to satisfy the primary objective of fault tolerance, then to add features to complete the other objectives.

5.4.1 Topology

The merits of ring, star and mesh topologies have been explored in chapter 4. Although the star network has the disadvantage of a central node, it does have an inherent property that any device connected into it may fail without any consequence to the rest of the network. Similarly no consequential problems exist when the cable is cut. That is to say, the most resilient structure as far as the interconnection of devices is concerned is the star.

A ring network on the other hand will fail completely with just a single failure but it offers performance improvements in data communication. In contrast to star networks, the weak point is the transmission cable which incidentally is the strong point of the star. By combining the best features of both structures, a highly resilient topology will be created while retaining the efficient communication advantages of the ring.

Fig. 18 shows the proposed Ring-Star topology, so named because a star structure is superimposed onto the ring. Note that this star structure is there only to enhance the reliability of the ring, data communication is always carried out in the ring until the occurrence of a fault.

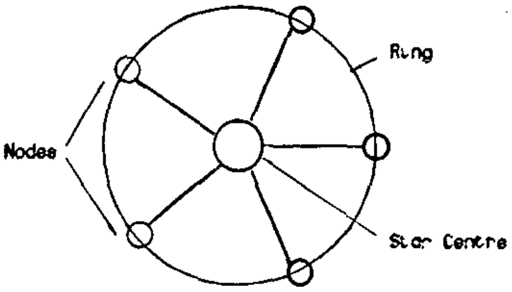


Fig. 18 The Basic Ring-Star topology

It could be argued that for the Cambridge Ring, it will still rely on a central controller but unfortunately this is a design fact. However the basic weakness of the ring, its topology has now been strengthened. As far as the development of the topology is concerned, the central controller will be treated as any other device on the ring. The problems of the central controller will be dealt with later. It must be noted that for a ring which does not rely on a central controller such as the IBM token ring, the Ring-Star must be one of the most resilient structures.

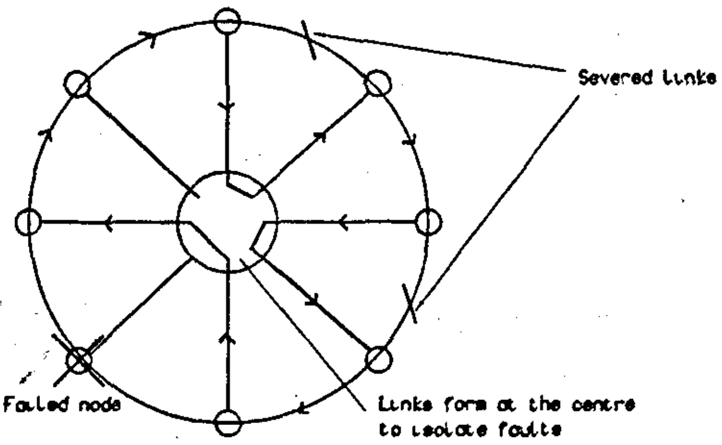


Fig. 19 The star structure satisfies the key objective of resilience to multiple faults without partition problems

5.4.2 Theory of Operation

Recall that the steps necessary to achieve fault tolerance are fault detection, location, diagnosis, isolation and finally recovery. Four major components will achieve this:

(a) Ring nodes

(b) Monitor station

(c) Error Logger

(d) Centre Switching Unit (CSU)

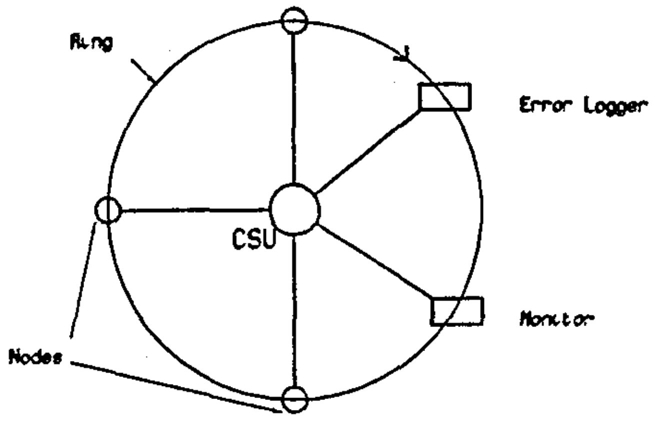


Fig. 20 The Ring-Star with its major components

As explained in chapter 3, each node on the ring continously checks every passing m-p. Errors detected will result in the node sending a fault message to ring address zero. This message details the type of fault and the node's own address - thus fault detection and location.

In the Cambridge Ring, both the Monitor and the Error Logger must be set to address zero. Therefore errors reported by the nodes will be received by the Monitor and the Error Logger. The Error Logger acts as a database storing all errors which have been reported. By analysing the contents of the database, the Error Logger can detect the occurrence of critical faults. For example, if any node in the ring should fail, the result is a stream of fault messages being transmitted to the Error Logger by the next functioning node downstream. Using a timer, the Error Logger will within a certain time period detect this as a critical fault - thus fault diagnosis. Note however that this detection process is only partial. Although in theory it ought to diagnose both node failures and link breaks, in practice link breaks might not always be successfully identified. A better link break detection technique will be described in detail in the next section.

On diagnosing the occurrence of a critical fault, the Error Logger will dispatch a message to the CSU. The CSU is a simple device being basically a microprocessor controlled relay circuit. It functions as a switching centre responsible for directing network traffic flow through alternative paths. All nodes on the ring have a connection into the CSU. The message with the fault type and location information is processed by a fault algorithm. The result is a set of relay switching patterns relating the location of the fault to its position in the CSU. On activation of the relays, the ring fault will be bypassed by redirecting ring traffic around it - thus fault isolation is achieved.

The Monitor functions as the central controller in the Cambridge Ring system but in the Ring-Star, its role is to detect ring errors and in particular to resynchronise the ring after a fault has been isolated. These features have been designed into the Monitor and they operate automatically - thus network recovery need not be incorporated into the Ring-Star design.

5.4.3 Link Break Signal Generation and Detection

A link break is characterised by the detection of a stream of parity faults from a single source. Unfortunately when tests were carried out, this expected result did not arise. Instead the stream of errors appeared to have come from several sources, some of which were not even known ring addresses!

This odd result contradicted statements in the Cambridge Ring 82 standards and advice was sought. After discussions with Dr. Andy Hopper of Cambridge University (one of the key designers of the original Cambridge Ring) and Dr. Steve Wilbur of University College, London (whose research includes error logging for the Cambridge Ring), it was clear that the CR82 document stated an ideal situation. Indeed the statement would be true if every node on the Cambridge Ring has its own power supply! The cost makes this highly unrealistic. If the nodes takes its power direct from the ring (as is the usual case), then when it is severed, it cannot possibly transmit the correct fault messages due to the power loss. This was exactly what happened so an alternative solution must be sought.

The technique adopted to reliably provide a signal makes use of the repeater. When a link is broken the repeater immediately downstream generates a signal internally. This signal was identified and after filtering and amplification was brought out to connect directly into a latch in the

CSU. A unique position in the latch correspond to a unique ring node address. Since a low level signal is generated, the CSU was designed to read the latch periodically to detect this. Thus the CSU can ascertain if a link break has occurred and to locate its source. See Fig. 21. This technique requires a small modification to the ring repeaters.

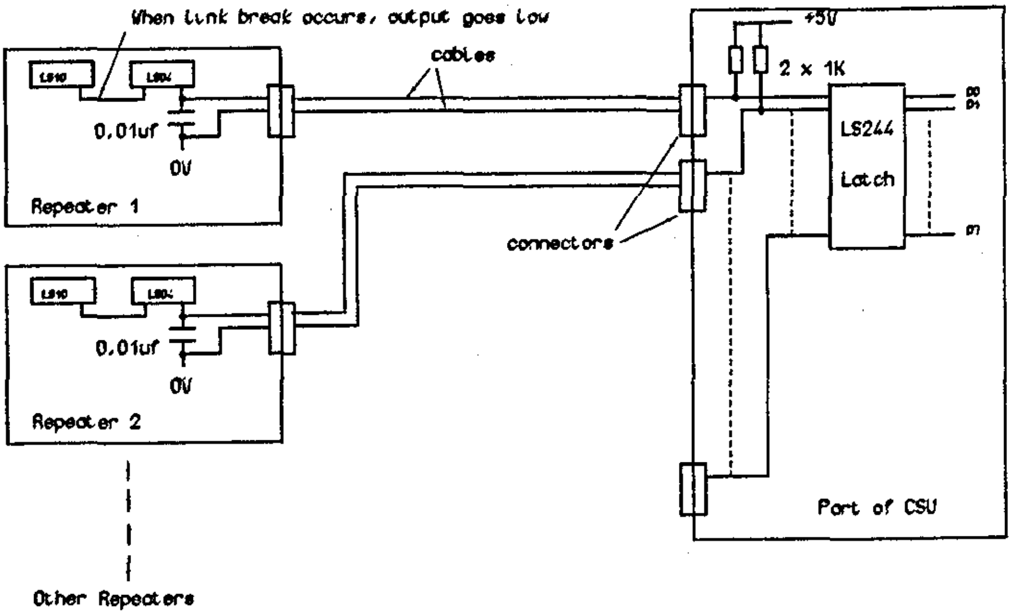


Fig. 21 A modification to allow the detection and location of link breaks

5.4.4 Off-Line Redundancy

In this concept, the system responsible for fault tolerance is normally quiescent until the occurrence of a fault. On detection of a fault, it is automatically activated to carry out its task.

There are two reasons for adopting this technique. First, off-line redundancy improves system reliability by up to 23% Second, if it is not adopted there is a possibility that the fault tolerant device might malfunction to cause unwanted actions on the system it is trying to protect. For example, it may switch on some bypass relays unintentionally, thereby disrupting normal ring operation.

5.4.5 Configuration Design

Having decided on a star structured architecture, the next step is to design the configuration in detail. Three variations were conceived, illustrated in Fig. 22A, 22B, and 22C. Ideally any design should minimise the number of relays per node to reduce cost and complexity. Version A (Fig. 22A) requires only two relays per connection, but it has the disadvantage that when a link break occurs, a healthy node will always be bypassed as well when the link is isolated. Version B was implemented in the first prototype of the experimental Ring-Star. Although this design requires less cabling, it has the same drawback as version A. It is possible that a perfectly operational node is bypassed unintentionally. This happens when more than one fault occurs. In Fig. 23, node 1 initially fails but before it can be repaired, link B is severed. Node 2 is unintentionally isolated too. This may not at first hand sound too significant but what if that node happens to be the Monitor? Version C was adopted to avoid the problems of Version A and B.

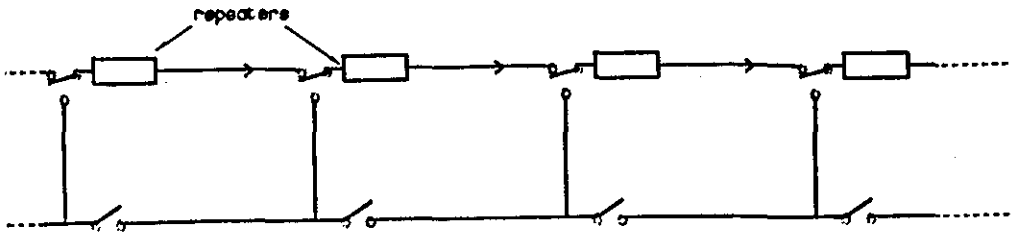


Fig. 22A

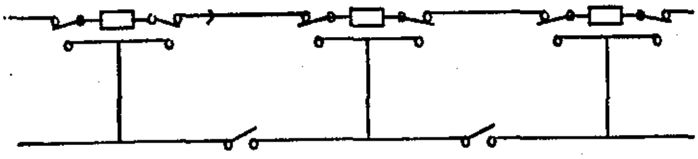


Fig. 22B

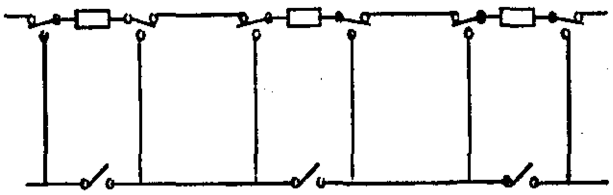


Fig. 22C

Next, the physical location of the relays must be decided. Two alternatives are possible, either to design bypass switching into the repeater logic circuit itself or to use external relays. The former was rejected because it was realised that the repeater may fail in such a way as to affect its operation, for example, failing to bypass the repeater when it should. Recall also that one of the key design objective was operational independence between the ring system and its fault tolerant component. Thus a decision was made to use external bypass relays controlled by an independent device ie. CSU.

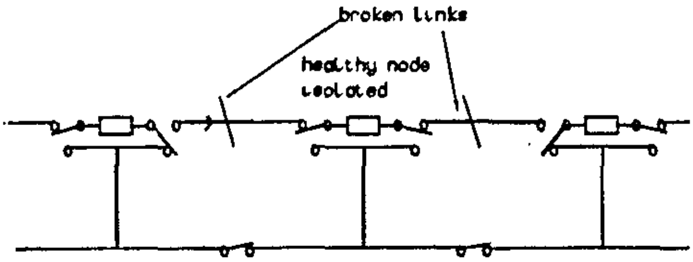


Fig. 23

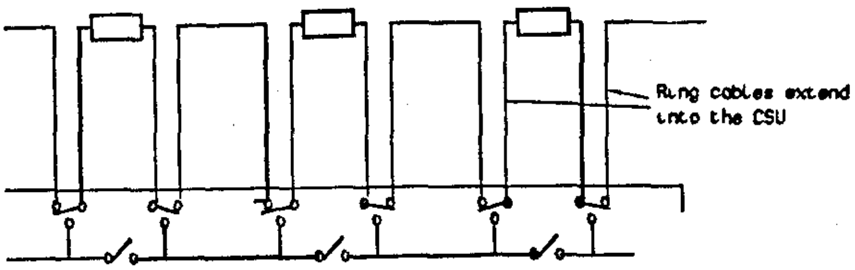


Fig. 24 Physical implementation of the relays. Relays are located in the CSU

5.5 Evolution of the Ring-Star

It was realised during the configuration design of the Ring-Star that there was a further problem. The concept of the Ring-Star may be sound but practically its star structure can limit its applications. Cable installation would be a major problem. This section describes the evolution of the single Ring-Star structure into a multiple Ring-Star architecture.

Configuration of a network topology depends on several factors:

- architectural layout of the installation, ie. number of rooms, floors, buildings, and how they are arranged.
- the availability and number of cable ducts, false ceilings or floors
- type of cable used eg. normal or flat
- level of resilience required
- cost

The ideal is a single Ring-Star, but it is likely to be suitable only in small installations with a small number of nodes. Flat cables may be used under the carpet to connect nodes to the CSU across the room rather than along walls as is the case when normal cables are used (unless a false floor is installed). With a large installation, both methods of laying the cables will be unsuitable. Imagine maintaining an installation with hundreds of cables running along the walls!

The solution to this problem is to use multiple CSUs distributed throughout the installation. Fig. 25 is an example of this technique. To prevent isolation when the section of ring cable between nodes of two CSUs is severed, the CSU-to-CSU links are added. Instead of having cables running across the installation, they are now mostly localised into groups.

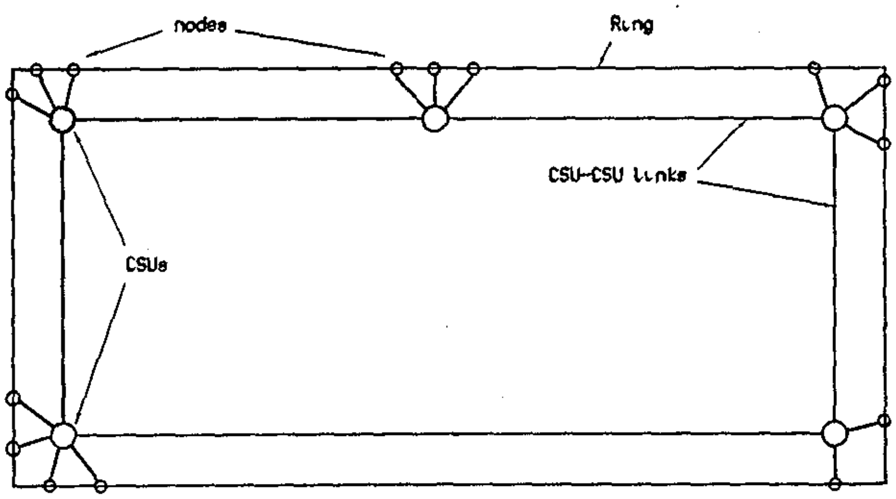


Fig. 25 By distributing CSUs over the installation the cabling problem is reduced

This arrangement is quite natural in that it reflects the architecture of the building. Computing equipment is usually located in clusters in rooms separated by corridors or other rooms. This point is especially relevant with Cambridge Ring equipment. They are supplied in racks, each holding several nodes and thus forms a natural cluster arrangement. Installation is therefore simpler.

However this layout now resembles the self-heal ring inheriting the disadvantage of partition problems when multiple faults occur. But fortunately, the advantage of clustering can be retained by evolving the system further into what is called a Hierarchical Ring-Star.

5.6 The Hierarchical Ring-Star Architecture

Referring to Fig. 26, it can be visualised as a star-within-a-star topology arranged in levels. The first connects nodes to a number of CSUs. These CSUs are themselves connected to another level of CSU which in turn are linked into a central CSU. This example illustrates a 3-level architecture but in practice it can have less or more levels. This approach is an extention of the basic Ring-Star concept and it will therefore retain the advantages of the Ring-Star concept for a wider area architecture.

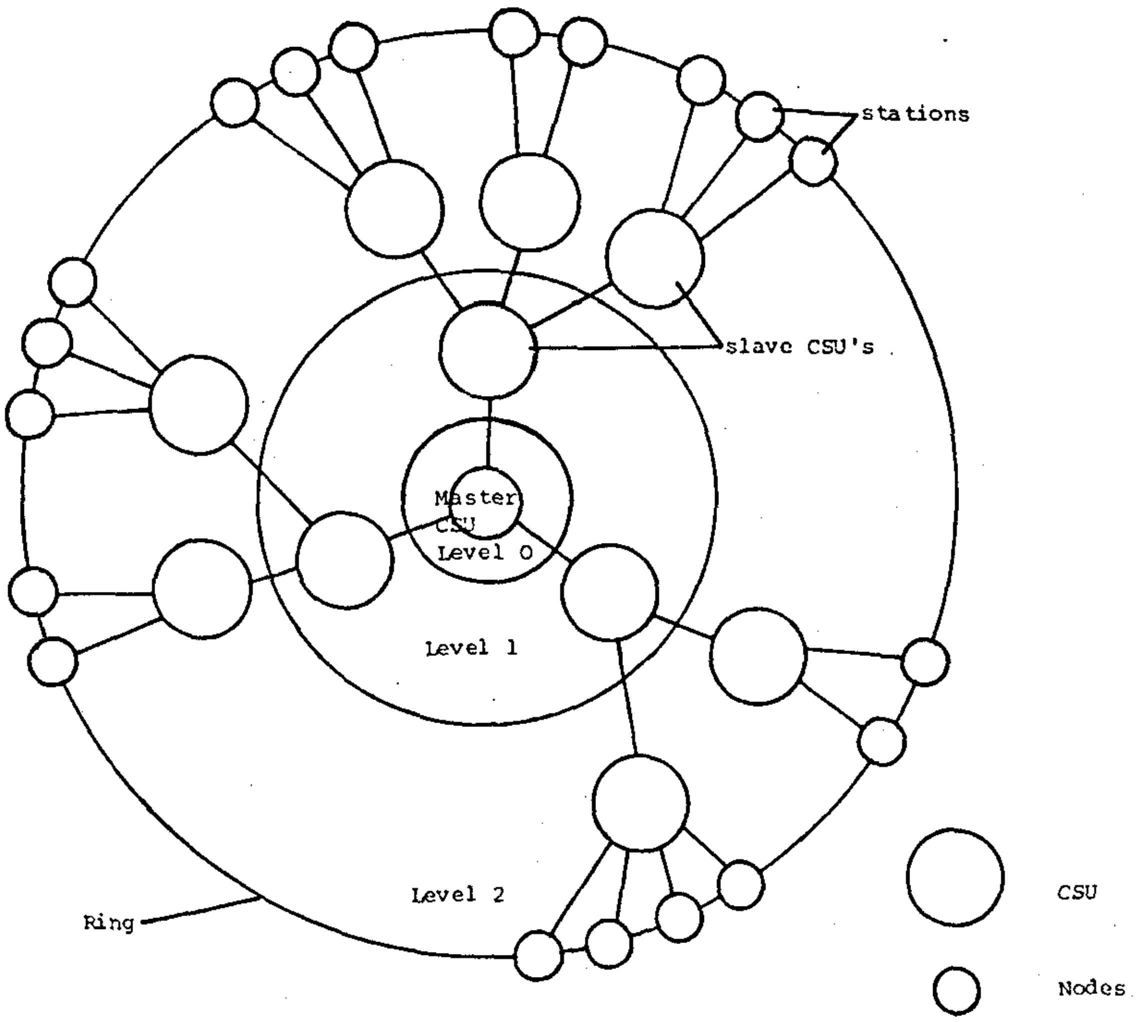


Fig. 26 A 3-Level Hierarchical Ring-Star Architecture

The central CSU serves to coordinate the other CSUs, thus they are called the Master CSU and the Slave CSU respectively.

Although cable management is reduced, with a much larger installation it can still be a problem. To further reduce cabling requirements, more levels can be added. This can be achieved by arranging the topology to fit into the architecture of the building or site. In general, one level is allocated to the room, the second to a floor, a third to a building and the fourth to connect between buildings within a site. The number of levels should be selected on the basis of the total number of nodes, how they are grouped, and the spread between them. This approach is highly structured and modular, using the basic Ring-Star topology as a building block to construct large networks.

CHAPTER SIX

IMPLEMENTATION OF THE HIERARCHICAL RING-STAR SYSTEM

6.1 Introduction

This section discusses the practical design details made. The hardware, software, algorithms, and operation of the Hierarchical Ring-Star system are described.

Hardware was kept simple by designing the circuits around a microprocessor, supported with standard ssi, msi and lsi components. The first prototype was wire-wrapped but the final design was produced on printed circuit boards.

Two different circuit boards were required : the Master CSU (MCSU) and the Slave CSU (SCSU). They are similar in design but for a different configuration of relays.

Section 6.2 describes the MCSU-SCSU interconnection scheme in the Hierarchical Ring-Star structure. A description of the MCSU is covered in section 6.3 followed by that of the SCSU in section 6.4 and the Error Logger in section 6.5. Section 6.6 has details of the Node Dictionary, a central concept in the design. Finally, operation of the

Hierarchical Ring-Star is explained in section 6.7 to complete the chapter. Section 6.7 also includes the algorithms employed in the design, paying particular attention to the fault algorithm.

6.2 MCSU-SCSU Interconnections

The SCSUs are connected to the MCSU in a star structure, similar to the way ring nodes are connected to the SCSU. In general, the Hierarchical Ring-Star can have any number of levels, dictated by the architecture of the installation. But to illustrate the interconnection scheme, a simple 2-level architecture will suffice. See Fig. 27A. The original idea was to use a single link between the MCSU and the SCSU as shown in the diagram but this has a serious drawback. A situation may arise whereby a perfectly operational SCSU is bypassed, for example when the ring cables on either side of that SCSU are broken as shown in Fig. 27B. Several nodes on the ring will be isolated as a result - a form of partition problem. To satisfy the objective of minimal partition, a modification is made by adding a second link as illustrated in Fig. 28.

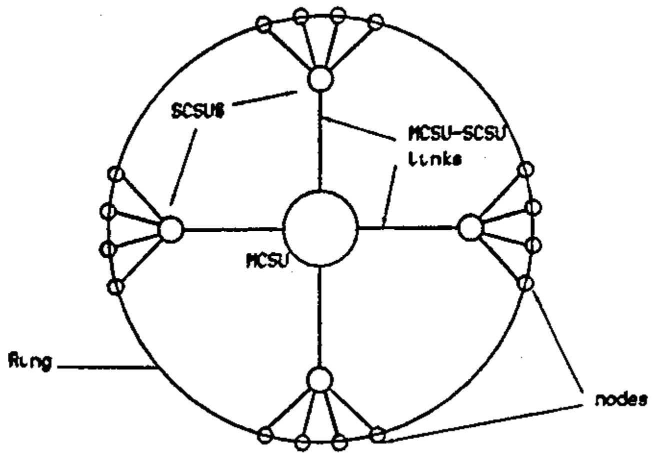


Fig. 27A A simple MCSU-SCSU Interconnection Scheme

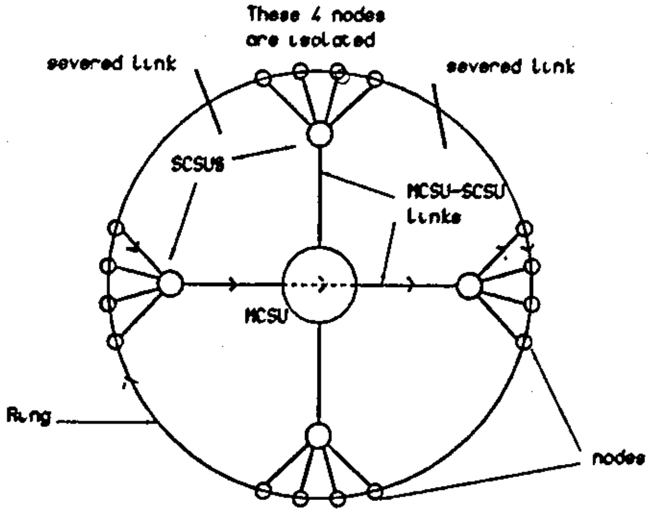


Fig. 27B Partition problem with the simple MCSU-SCSU Interconnection Scheme

The number of relays required in the MCSU are doubled but the second pair of cable is more significant. Ring cables are not cheap and since the SCSU may be physically placed a fair distance from the MCSU, this may add quite substantially to the overall cost. So the question is, "Is this partition problem in practice really that critical?"

This can only be answered in the context of an application. However for the experimental system, the ideal model shown in fig. 28 was designed.

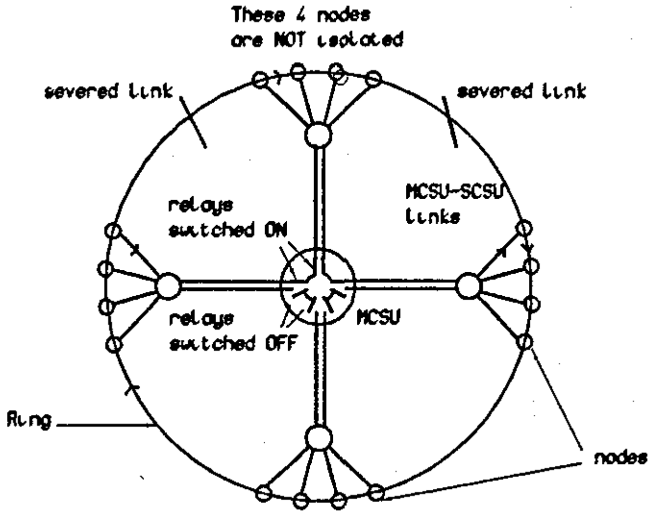


Fig. 28 Improved Interconnection Scheme to prevent Partition Problems

It was also decided that a 2-level architecture be adopted for two reasons. Firstly, this project on completion will be adopted for the Cambridge Ring installation in the department of Electronics Engineering. Secondly, a 2-level architecture is the minimum Hierarchical Ring-Star configuration possible and therefore the cost is minimised.

Most of the ring nodes in the department are situated in several rooms within a single floor of the building. If a SCSU is allocated to each room, they can in turn be connected to the MCSU within the same floor. Thus a 2-level architecture suffices. Also, although most rooms have only one or two nodes, a 4-node SCSU was decided upon.

The spares are there for future uses. For the same reason, the MCSU has been designed with 8 relay ports.

6.3 The Master Centre Switching Unit (MCSU)

6.3.1 Introduction

The MCSU functions as the central controller of the Hierarchical Ring-Star system. Conceptually, it is positioned in the middle of the installation responsible for controlling the network structure. It does this by coordinating SCSUs to alter the ring's transmission path. The MCSU is also responsible for storing information concerning the installation by maintaining a database called the Node Dictionary. The term dictionary is used because it holds complete details of every segment of the ring. For example, it knows if a particular node or link has been bypassed or not, and whether it is faulty. This Node Dictionary is dynamically maintained to reflect the real-time status of the network.

6.3.2 Basic hardware

The MCSU was designed as a conventional microcomputer to which was added extra circuitry required to control bypass relays. The block diagram is shown in Fig. 29. The detailed circuitry can be found in appendix 3.

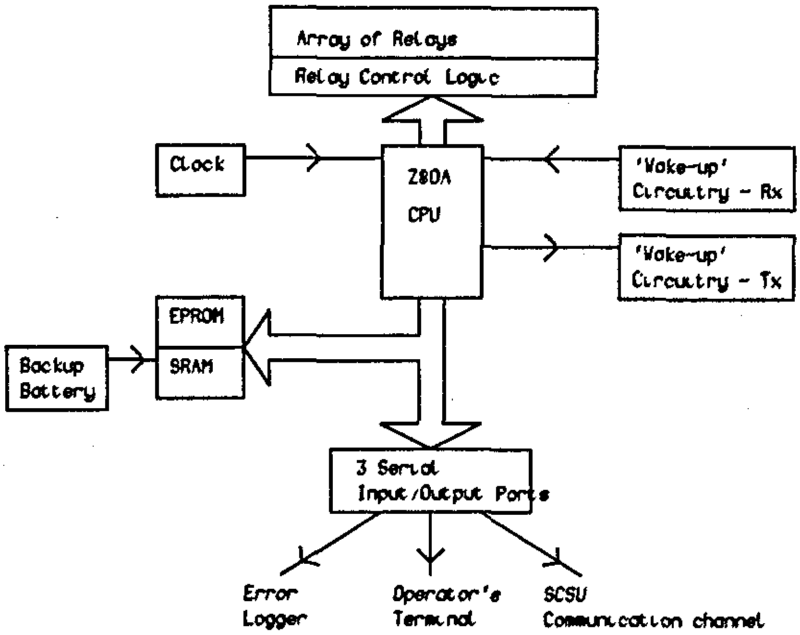


Fig. 29 Block diagram of the MCSU

The unit consists of a Zilog Z-80A microprocessor, 8 kilobytes of EPROM, 8 kilobytes of static RAM, three serial input/output ports together with the necessary supporting circuitry. This configuration provides the means to control an array of relays, arranged to support eight relay ports. Each relay port connects a single SCSU into the system.

6.3.3 Relay configuration

Each relay port (RP) consists of two relays connected back to back, shown in Fig. 30 in the off position.

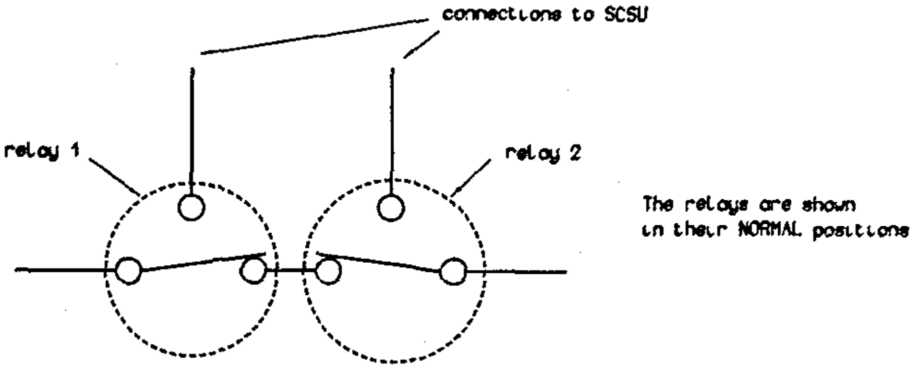


Fig. 30 Relay configuration of a Relay Port in the MCSU

When the MCSU is first powered up, an algorithm configures the relays so that if a RP is not connected to a SCSU, it is left in its normal position. This is because any unused RP must provide a through path in case a signal travels between two SCSU separated by one or more unused RPs. Fig. 31 should further clarify this.

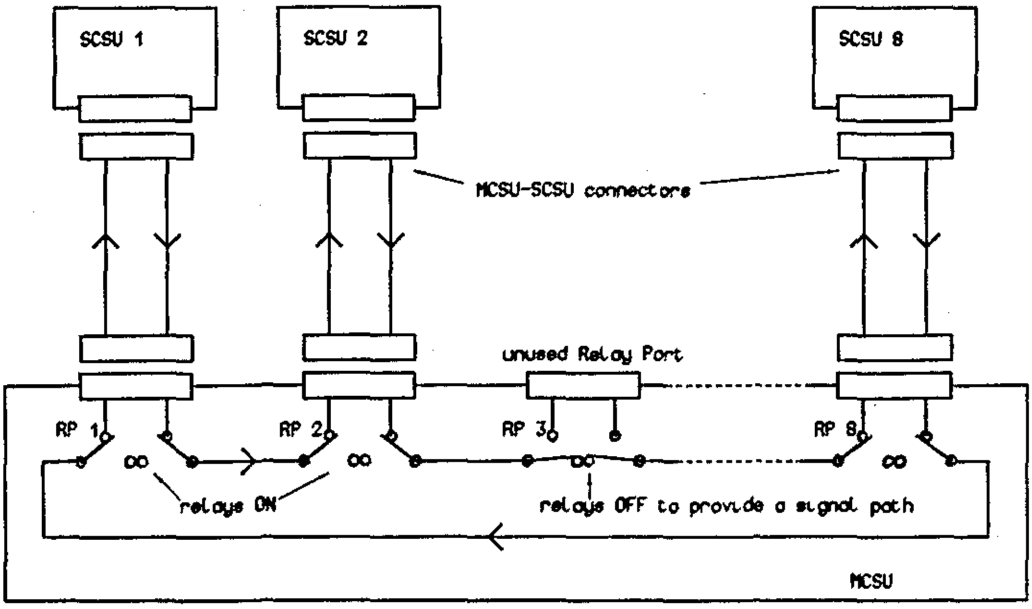


Fig. 31 Physical implementation of the Hierarchical Ring-Star

6.3.4 Communication channels

Two serial input/output ports are provided by a Zilog Z80A SIO; one for communicating with the Error Logger and the other with a terminal. It is through the terminal that an operator can manipulate the MCSU either for configuration control or for status information. The Error Logger channel is required for passing fault messages between the Error Logger and the MCSU. An Intel 8251A provides a third serial port for communication with the eight SCSUs. Note that instead of having 8 serial ports (one for each SCSU), only one was implemented to reduce cost. All the SCSUs are wired onto the same multidrop line and by using time multiplexing together with protocol techniques, messages can be exchanged reliably. This is satisfactory because communication between the SCSU and the MCSU is infrequent and short. It is only required when faults are detected or when operator controlled configuration of the system is required. The physical connections are illustrated Fig. 32.

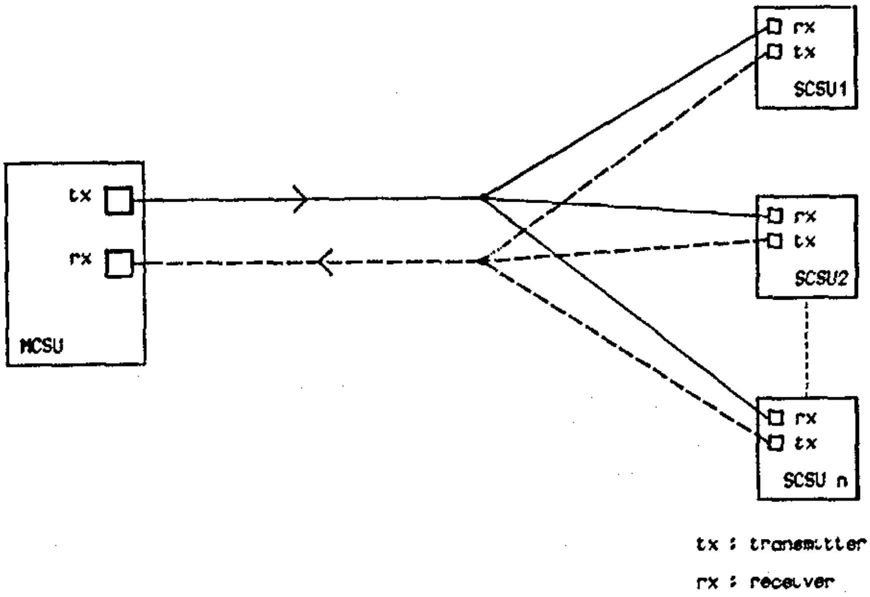


Fig. 32 Multidrop line technique adopted for MCSU-SCSU Communication

6.3.5 Basic tasks

When the MCSU receives a message which requires switching operation (eg. fault message received from the Error Logger), a relay pattern generation algorithm is executed. The resulting relay switching pattern activates the appropriate relays to accomplish the task.

The relay pattern generation algorithm, and the MCSU control software are contained in the EPROM while the RAM provides buffer space and workspace for the microprocessor. Also stored within the static RAM is the Node Dictionary. The Node Dictionary is a file which stores all data relating to the status of all the nodes, links, and SCSUs in the network. More significantly it records the node's ring address-to-relay port relationship and the SCSU's address-to-relay port relationship so that it knows which relay port to activate in order to isolate a faulty node, for example. This information is fed into the MCSU when the system is first installed, after which it is dynamically maintained.

Since the MCSU is usually inactive, a battery has been added to the static RAM to retain the contents for ten years. This is required to implement off-line redundancy.

6.3.6 Off-line redundancy

To implement off-line redundancy, the "wake-up" concept is adopted. Put simply, the normally quiescent device is activated by a "wake-up" signal when its operation is required. Two "wake-up" circuits are employed; one used to activate the MCSU itself while the other is used by the MCSU to activate SCSUs.

The MCSU can be activated by three sources:

(a) The operator

The operator is expected to set up the system initially and when required, to control the network ie. reconfiguration. Summary reports of the current network status may also be requested. A manual switch is provided to "wake" the MCSU up.

(b) The Error Logger

When node failures are detected by the Error Logger a message must be sent to the MCSU for corrective actions. Before this message is dispatched, a "wake-up" signal must be sent ahead to activate the MCSU.

(c) The SCSU

Link breaks are detected by the SCSU. Once corrected, a message must be sent to the MCSU for logging. Again a "wake-up" signal must precede the message.

The following circuit achieves the task.

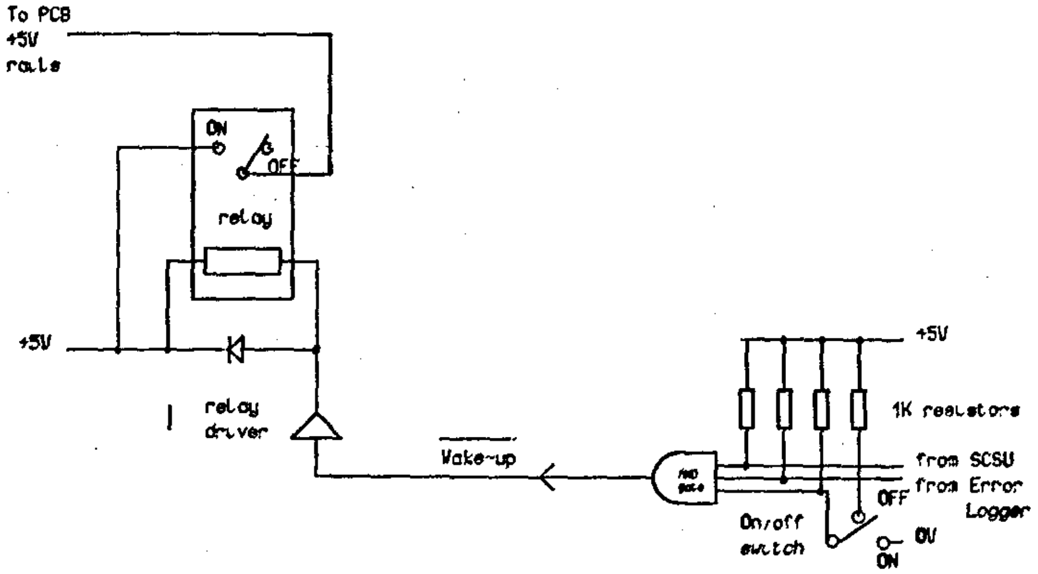


Fig. 33 "Wake-up" Circuit. All inputs to the AND gate are normally high. A low on any input activates the MCSU

The "wake-up" signal is generated by an AND gate. Normally the output of the AND gate is high. When an operator requires the service of the MCSU, the manual switch provided is set on. This forces the output of the AND gate low generating the "wake-up" signal. This signal next drives a relay driver to switch on the relay completing a circuit to provide the MCSU with power. Similarly the

Error Logger achieves the same effect by driving a low signal into the AND gate. Note that the logic has been so designed that if any one of the inputs is accidentally cut, the MCSU is automatically activated.

When required, the MCSU can activate any one or more of the SCSUs. For example, to bypass a node attached to a particular SCSU, only that SCSU needs to be operational. The rest are not required and thus left in their quiescent state. Again a simple high/low signal is enough. A latch functions as a parallel interface to supply the "wake-up" signal. Under the control of the microprocessor, any pattern may be sent to the latch to activate the selected SCSU. Once activated, it remains in that state until reset by a complementary pattern.

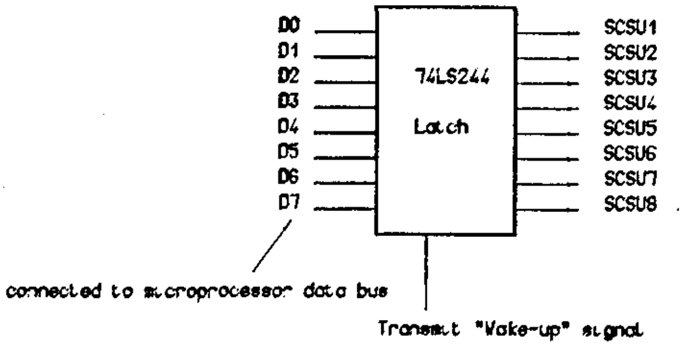


Fig. 34 "Wake-up" Circuitry - Transmit. The latch is normally initialised to binary 1. When any SCSU is to be 'woken up', a 0 is written into the latch corresponding to that bit. eg. if SCSU2 is required, the pattern sent to the latch is binary 10111111

6.4 The Slave Centre Switching Unit (SCSU)

6.4.1 Basic hardware

The hardware is similar to the MCSU. With the exception of the relay configuration, the SCSU is really a simplified MCSU. It has been devolved into a peripheral device largely controlled by the MCSU, thus slave CSU.

The circuitry differs from that of the MCSU by having only one serial port, less memory and a latch which detects link breaks. The block diagram is shown below. Appendix 4 has the detailed circuit.

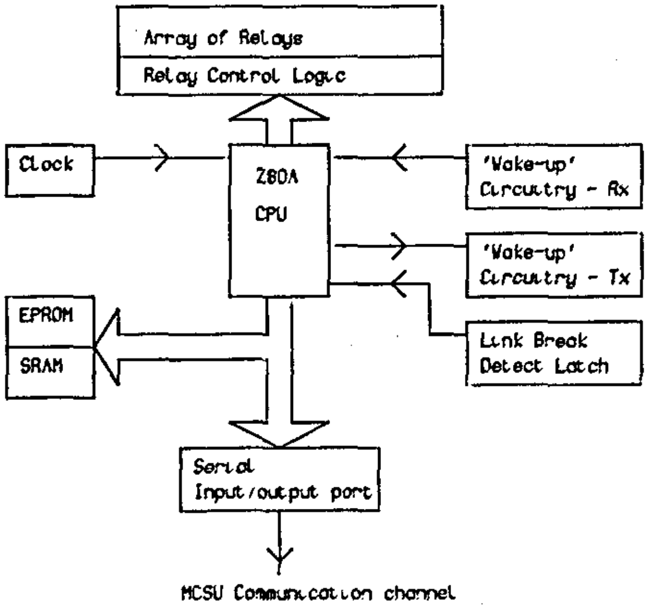


Fig. 35 Block Diagram of the Slave CSU

The serial port communicates with the MCSU, passing fault messages and receiving switching commands. The link break latch records the status of the links attached to nodes to which the SCSU is connected. By reading this latch periodically, the SCSU can detect a broken link.

6.4.2 Relay configuration

Each repeater has three relays to control ring signal path. Two are on either side of the repeater. Their design are discussed first.

Consider the relay positions in Fig. 36.

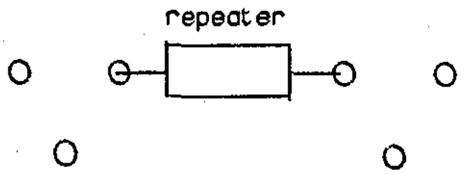


Fig. 36 Relay configuration

Ideally, each relay should be able to switch in three ways thus:

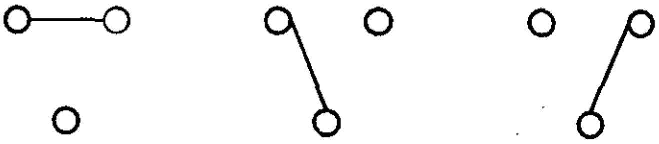


Fig. 37 The three possible relay positions

The reasons are illustrated in Fig. 38A, Fig. 38B and Fig. 38C.

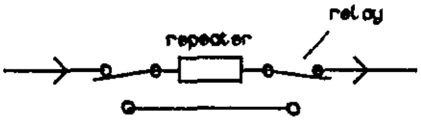


Fig. 38A Normal position of relays

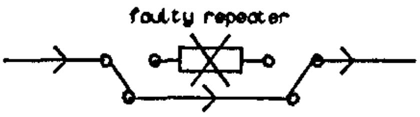


Fig. 38B Position of relays when bypassing a faulty repeater

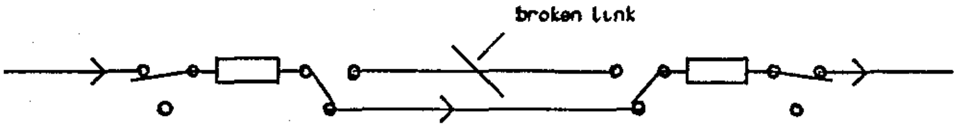


Fig. 38C Position of relays when bypassing a broken link

However no such relay exists. A three position relay can be implemented but it requires two relays. This is uneconomic considering the number of such relays required and moreover, it can be avoided. Consider a section of the Ring-Star structure in perspective. It can be shown that position B can be left out by adding a third relay (instead of four if 3-position relays are adopted). Relay position A is a must since this is the initial and normal position for ring signals when no faults exist. This is also the normal position of the relay when no power is applied to the relays.

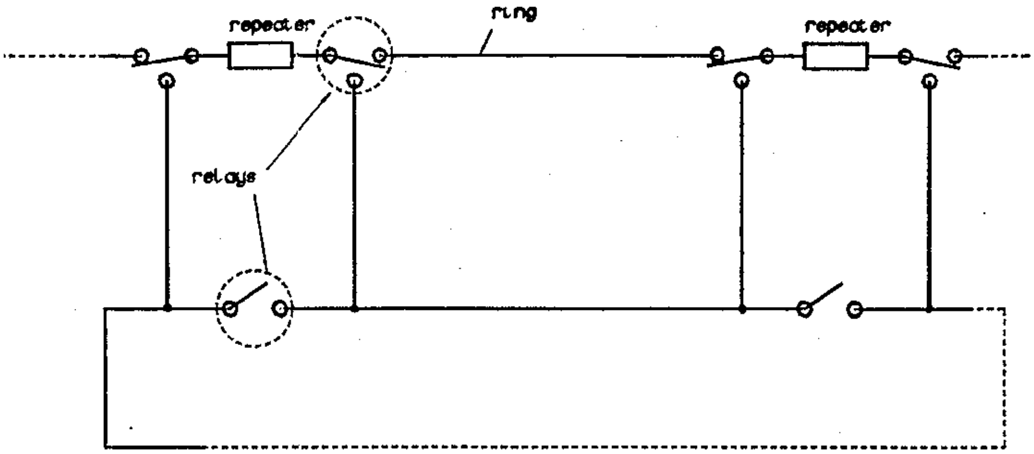


Fig. 39 Normal positions of relays when ring is operating normally

By observing Fig. 40A and Fig. 40B, relay position B is altogether not required.

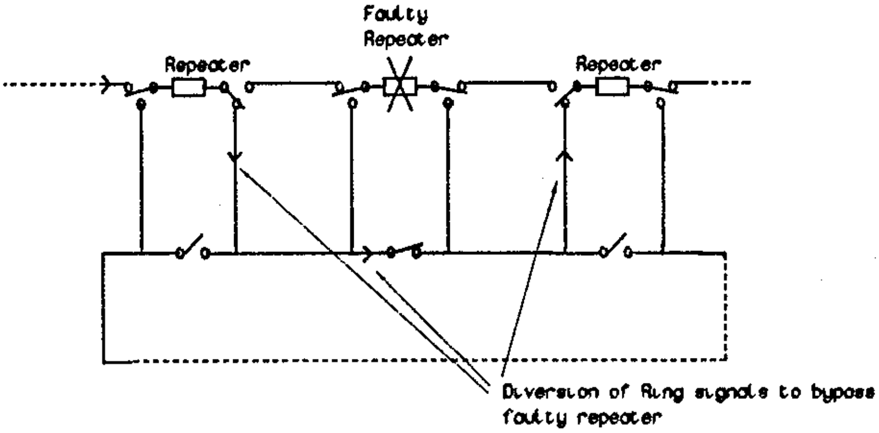


Fig. 40A Configuration to bypass a faulty repeater

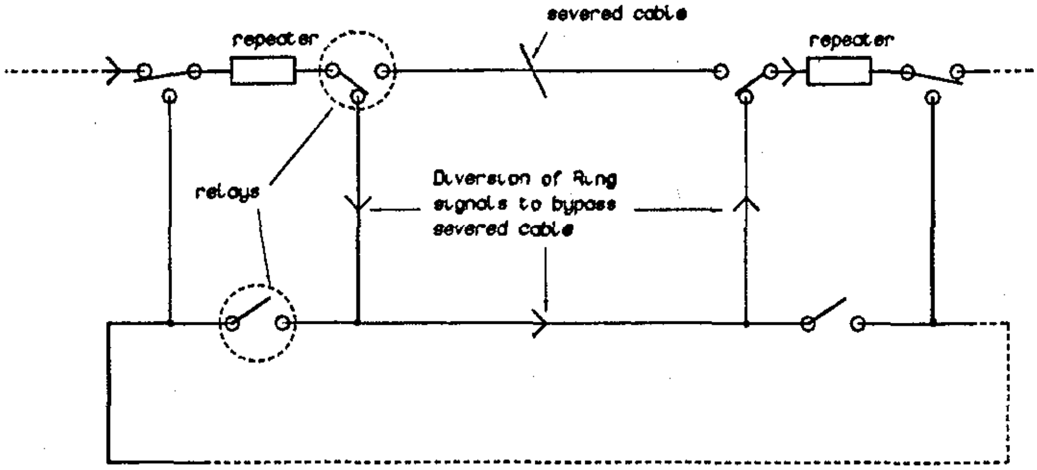


Fig. 40B Configuration to bypass a severed cable

This configuration therefore requires three relays for every node. The next question is, "Where are the relays physically located?" To realise another objective which requires that existing Cambridge Ring equipment can be used with minimal modification, the configuration shown in Fig. 39 is unacceptable. It requires relays to be built into the repeaters itself. To avoid this, the relays are designed into the SCSU with connectors linking them to the ring and the repeaters as illustrated in Fig. 41.

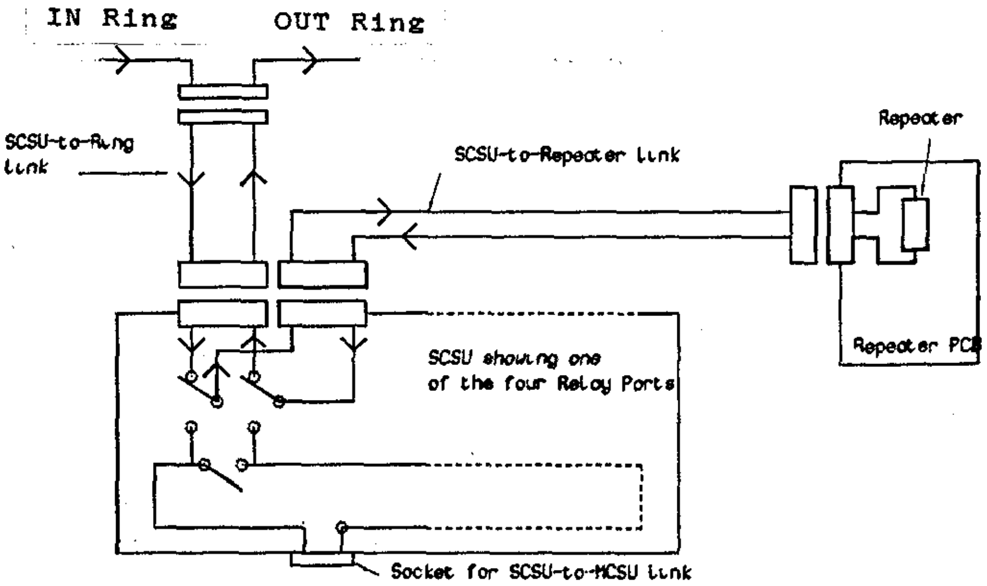


Fig. 41 Physical implementation of the SCSU

In a normal ring implementation, connectors are used to link each repeater into the ring. Now, instead of that, connectors from the ring and repeaters are brought into the SCSU where the link is then made through the relays. On the SCSU this is called a relay port.

6.4.3 Basic Tasks

The SCSU has two major tasks:

(a) to detect link breaks

(b) to accept and execute switching commands received from the MCSU

When either of the above two tasks are required, the SCSU is activated from its normally quiescent state by a "wake-up" circuit. This circuit is controlled by two sources; the MCSU and the node.

When a link break occurs, the "wake-up" circuit will detect a high to low transition transmitted from the repeater which diagnosed the break. This switches power into the SCSU, "waking" it up. The SCSU can ascertain where the cable is broken by the very same signal. These signals are also fed into a latch. By scanning this latch the SCSU can determine the location of the break. This has been explained in detailed under the section "Link break signal generation and detection." Two alternative actions are next carried out depending on where the break is. The break may occur on a <u>normal</u> link or an <u>edge</u> link. These two types of links are shown below:

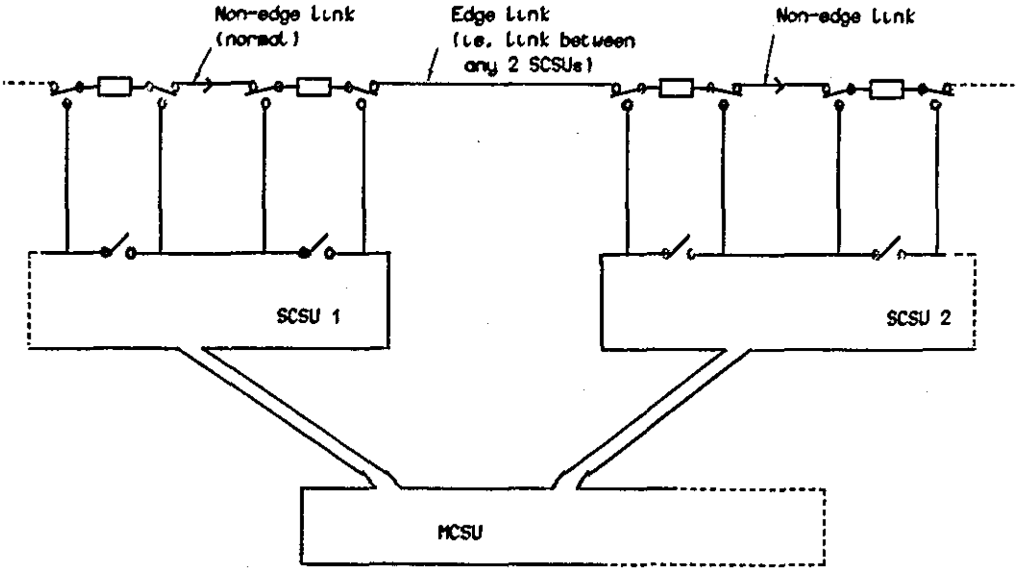


Fig. 42 To illustrate the two types of links. Edge links are links which connect two nodes separated by two SCSUs. Normal or Non-edge links connect nodes within one SCSU.

If it is a normal link, it is immediately isolated as illustrated in Fig. 43.

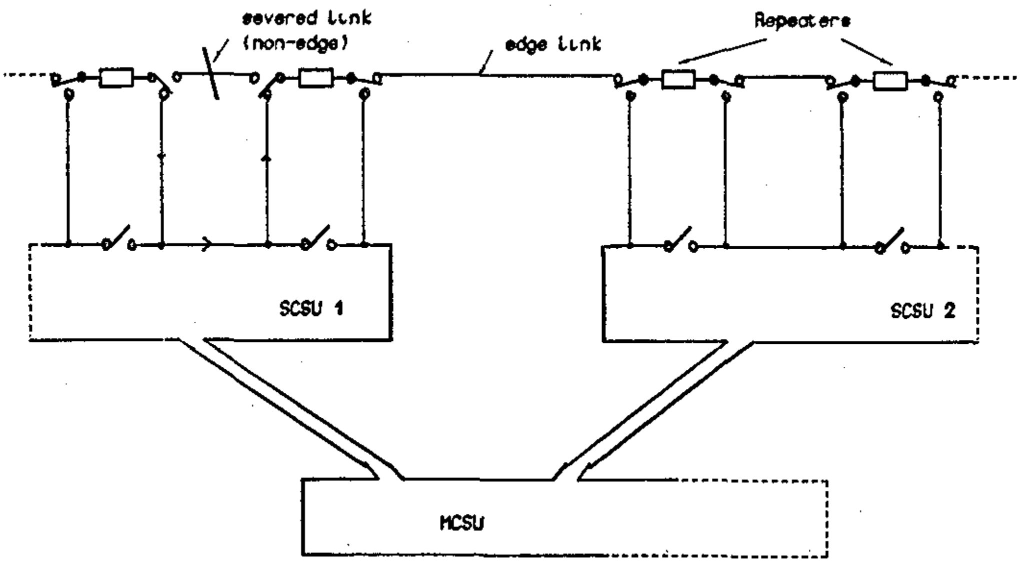


Fig. 43 Isolating a non-edge link requires action of only 1 SCSU

A message is then sent to inform the MCSU of the fault. The Node Dictionary will be updated to reflect the new status. Again a "wake-up" signal must precede this message to activate the MCSU.

If it is an edge link, no switching action will be carried out. Instead a message will be sent to the MCSU. It awaits the return message which will contain the necessary commands to isolate the fault. The reason is that in this case, the actions of two SCSUs are required. The MCSU is responsible for coordinating this task.

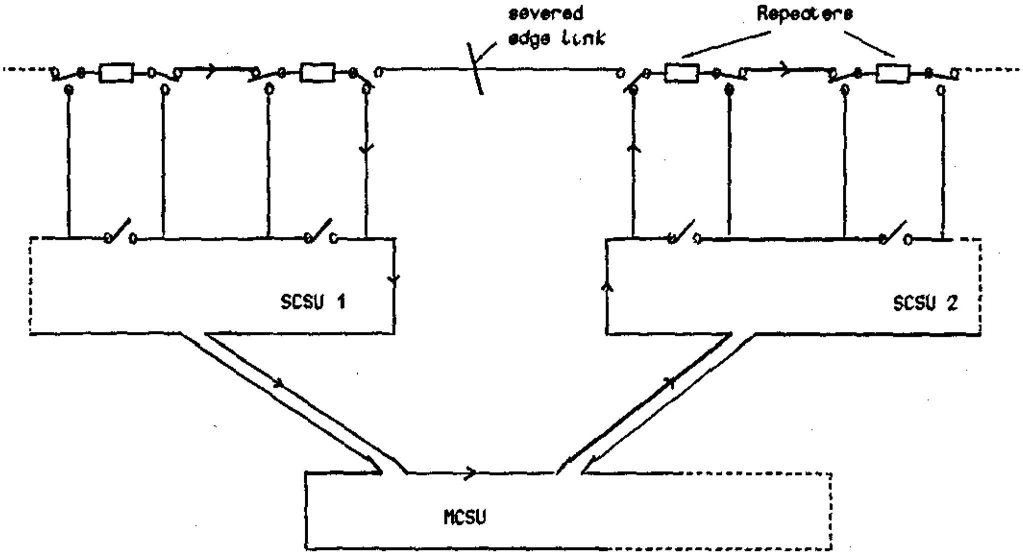


Fig. 44 Isolating an edge-link requires the actions of 2 SCSUs and the MCSU

The second type of task originates from the MCSU itself. It may be an operator's command to bypass a node or a link. Likewise it could be the corresponding reset commands. Operator commands allow the ring configuration to be altered. Typical examples are when maintenance or expansion are required.

In both cases, the SCSU will remain in the active state until the occurrence of another stimuli. If the links have been physically repaired, the MCSU can be instructed to de-activate the particular SCSU by removing the "wake-up" signal. It will then return to its quiescent state.

Finally, each SCSU has a unique address, set up (on an 8-way switch) during installation. This is important since it identifies the SCSU which is communicating with the

MCSU. Recall that a single communication channel is shared between all SCSUs.

6.5 Error Logger

The primary function of the Error Logger is to log ring errors. It forms part of the Hierarchical Ring-Star system responsible for fault detection.

The Error Logger is an active device connected directly into the ring, and is positioned immediately upstream of the Monitor. Since all error messages are sent to ring address zero, the Error Logger must be installed with that address to receive them.

It has four main tasks:

(a) monitor the ring for errors, to log them and their sources

(b) keep track of error occurrences

(c) inform the MCSU when node faults are detected

(d) provide summary report of ring errors

Before describing the tasks above, all possible error conditions are presented.

6.5.1 Ring Errors

The following errors are caused by faulty nodes, faulty ring cables or intermittent noise induced faults.

(a) Empty - a packet has been received by the Monitor with an illegal leader sequence. This could lead to an empty chain error. This type of packet is deleted by the Monitor.

(b) Parity - The parity of a packet entering a node was incorrect.

(c) Ø to 1 - On checking a returned unused packet which was filled with data it is found that a bit which was transmitted as a zero has changed to a one.

(d) 1 to Ø - On checking a returned unused packet which was filled with data it is found that a bit which was transmitted as a one has changed to a zero.

(e) 2nd time full - Indicates that a full packet is making its second pass through the Monitor. When a full packet enters the Monitor, the "Monitor passed" bit is cleared. Therefore, if that packet re-enters the Monitor with that bit uncleared, an error has occurred.

(f) Lost leader - Indicates that the "Start of packet" bit is not 1. This is a framing error and may cause the ring to lose synchronisation.

(g) One in gap - The gap should be an "all zero" sequence. If not, this causes a framing error.

All the above errors cause error messages to be sent to ring address zero.

6.5.2 Hardware

The Error Logger is based on a standard Z80 Small Server, with two serial ports added.

In total, there are three ports. One interfaces into the Cambridge Ring through a node, allowing ring packets to be read or written to. Fault messages sent by ring nodes are received through this port. The second port provides a serial link for communication with the MCSU. The third connects a terminal to the Error Logger allowing commands to be entered by an operator or error information to be displayed.

6.5.3 Functions of the Error Logger

Since the Error Logger is set to addressed zero, it automatically monitors the ring for errors. These are logged together with their source in a file. Thus a dynamic record of the ring's error status is maintained.

Each type of fault and their occurrence rate are recorded. In particular, it tracks the occurrence frequency so that if any single source causes X number of errors over a time period Y, a critical fault has occurred. Note that it does not distinguish between the different types of errors as long as they originate from the same source. A message will be sent to inform the MCSU. From this information, the MCSU is expected to bypass the fault. In the experimental system, X is set to 20 and Y set to 1 second. Critical faults are recorded in the file together with the other errors. If required, the Error Logger may be issued a command to print out an error report. This is simply a summary of all errors recorded with their frequency and source.

6.5.4 Fault Message Packet Format

Fault messages are transmitted on the first passing empty packet with the following format.

Bits	Description
1 - 3	set to one
4 - 11	bit address set to 0
12 - 19	address of error source
20 - 27	error count
28	one in gap error
29	lost leader error
30	2nd time full error
31	1 to 0 error
32	0 to 1 error
33	parity error
34	empty error
35	fault packet received
36 - 37	set to one
38	determined by fault message packet parity

6.5.5 Error Logger to MCSU Communication Protocol

A simple protocol is employed for communication with the MCSU. Since messages are short (a few bytes), an asynchronous transmission technique is adopted with the following characteristics : eight bits, odd parity, one start bit, and one stop bit.

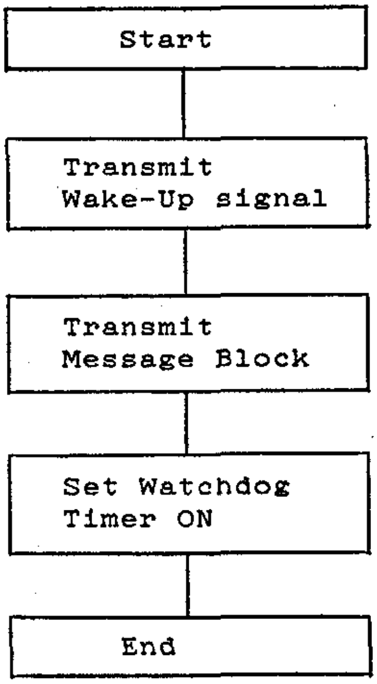
Messages are exchanged in blocks with the following format:

Code for Node Fault ie. 11H
Relay Port number
Checkword, Byte 1 (lsb)
Checkword, Byte 2 (msb)

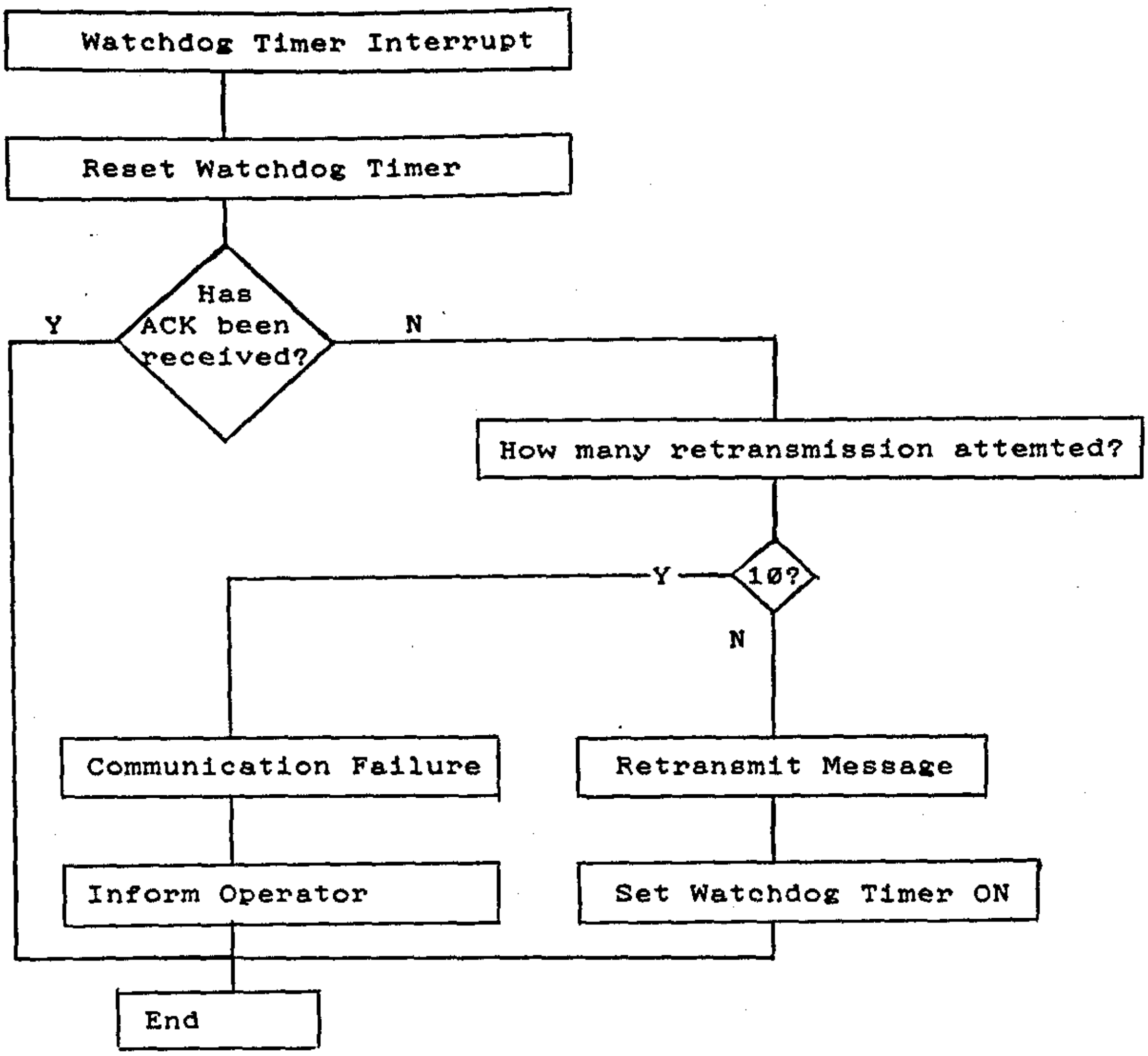
Fig. 45 Message Format for Error Logger to MCSU Communication

The first byte contains the fault code ie. 11H for node fault, followed by the Relay Port number of the node which detected the fault. The checkword ensures error-free exchange of messages. Normally checksums are used in such transmission protocols but in this case of short 2-byte messages, they are not. Instead the checkword duplicates the message completely. The receiver checks this against the original message and accepts it as correct if they match. This technique reduces the chances of undetected bit errors. An acknowledgement is then returned to the sender. If the message received is erroneous, it is simply rejected. The sender is not informed of this but instead is expected to detect the problem.

On transmission of a message, the sender sets a watchdog timer. If an acknowledgement is not received within the time period set, it automatically re-transmit the same message. A total of ten re-transmissions are allowed, after which the task is abandoned. A "transmission failed" message is then sent to the operator's terminal. Fig. 46 illustrates the protocol in two flow charts.



Transmission



Acknowledgement management

Fig. 46 Error Logger - MCSU communication protocol.

6.5.6 Other Facilities

The terminal linked into the Error Logger provides an operator with error information as follows:

(a) as ring errors are detected, they are immediately displayed on the terminal.

(b) a summary error report can be requested through the menu driven user interface (entering control A on the keyboard) and choosing the "Print Error Information " option.

6.6 The Node Dictionary

The Node Dictionary is central to the operation of the Hierarchical Ring-Star. It is a database containing configuration information relating to the architecture of the installation as follows:

- the node's ring addresses
- SCSU addresses
- status
- address-to-relay port relationships

The last class of information is the most important. It allows fault messages received from the Error Logger and SCSU (containing the type of fault with the corresponding address of the node which detected them) to be translated into the corresponding relay set on the SCSU. Only then can the faulty component be isolated. This information must be entered into the Node Dictionary when the system is first installed.

Nodes, SCSU addresses and status information are all eight bit binary quantities stored in a strict sequence in the file to reflect the address-to-relay port relationship.

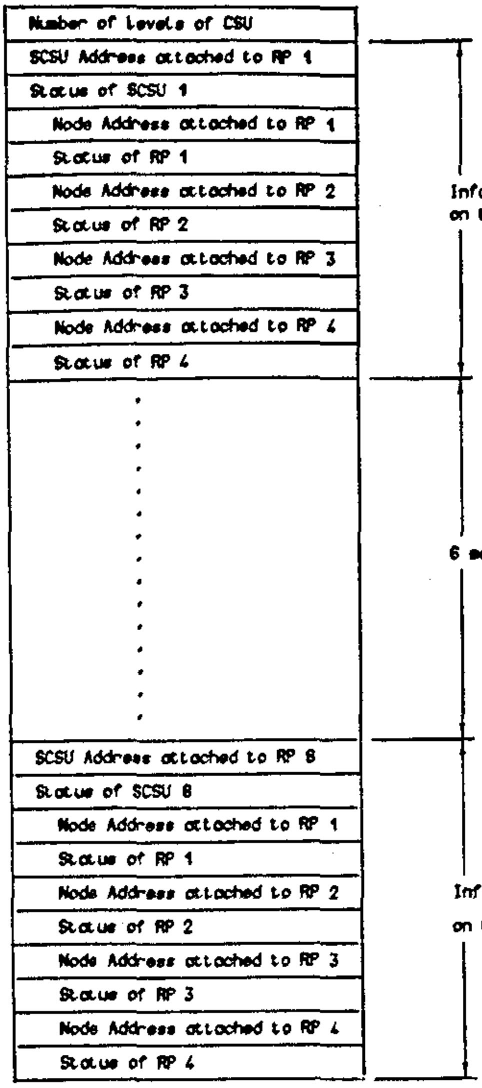
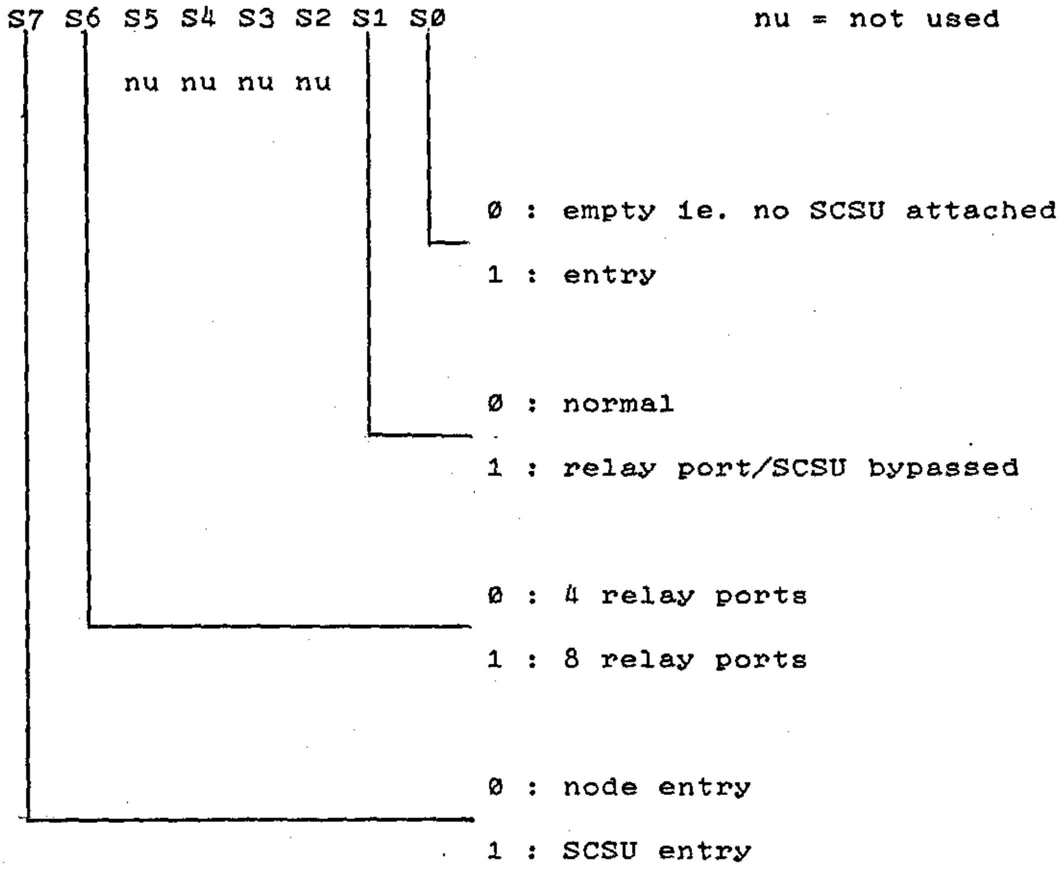

Number of Levels of CSU
SCSU Address attached to RP 1
Status of SCSU 1
Node Address attached to RP 1
Status of RP 1
Node Address attached to RP 2
Status of RP 2
Node Address attached to RP 3
Status of RP 3
Node Address attached to RP 4
Status of RP 4
Information relating to 1st RP on the MCSU
6 more sets of RP
SCSU Address attached to RP 8
Status of SCSU 8
Node Address attached to RP 1
Status of RP 1
Node Address attached to RP 2
Status of RP 2
Node Address attached to RP 3
Status of RP 3
Node Address attached to RP 4
Status of RP 4
Information relating to Last RP on the MCSU

Fig. 47 Node Dictionary File Structure

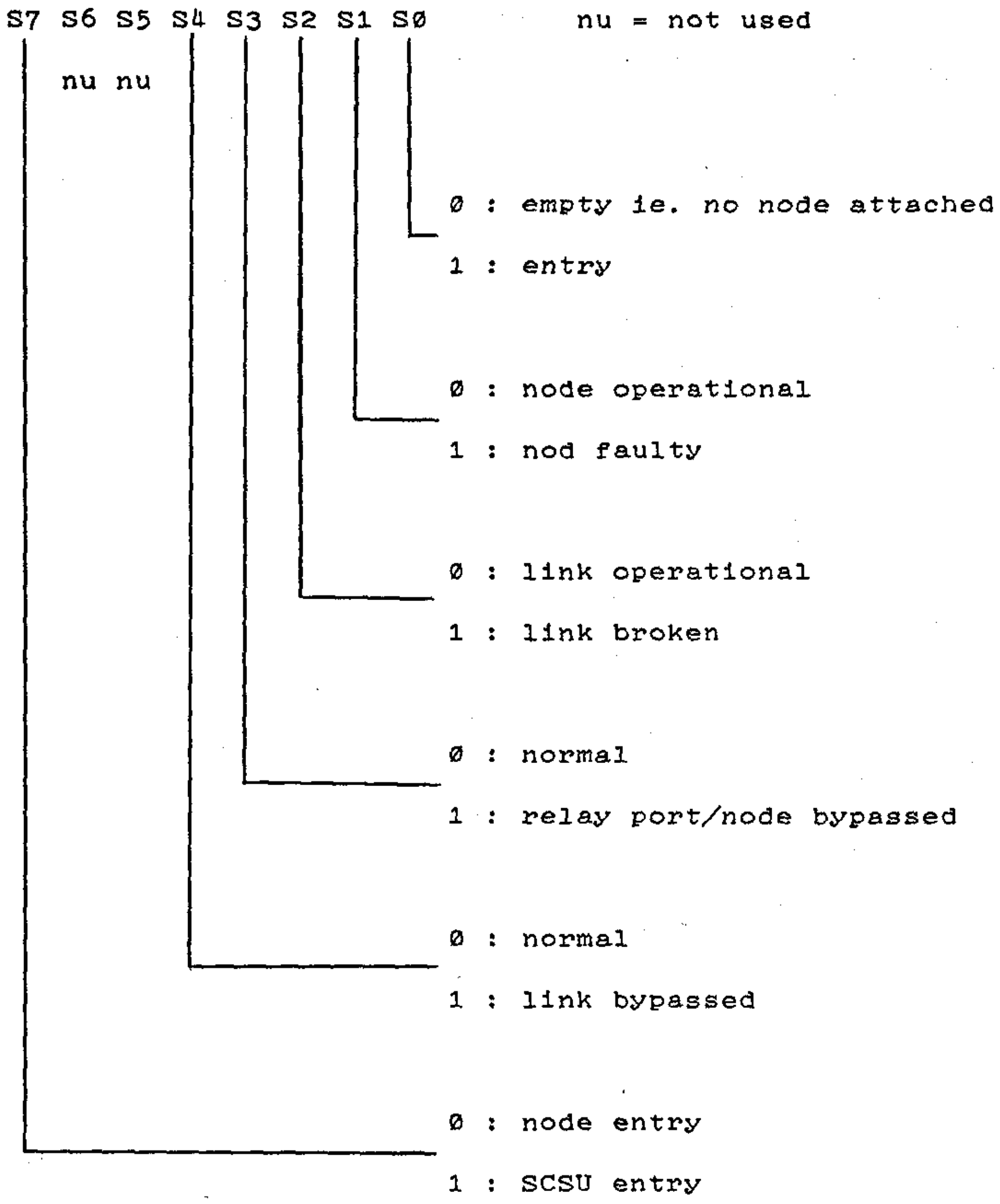
The first field, "number of levels of CSU" indicates the total number of levels in the Hierarchical Ring-Star installation, in this case, two.

The rest of the file is divided into records, each relating to one relay port of the MCSU. The address and status of the SCSU attached to this port are recorded into the first field. The status field stores the following flags:



Note that if any of the relay ports does not have a SCSU attached, an entry <u>must</u> still be made. Status bit SØ reflects this. S1 indicates whether the relay port (or SCSU if one is attached) has been bypassed or not. Bits S6 and S7 are there only to aid the Node Dictionary searching algorithms. An "8 relay ports" field implies that it is within a MCSU record. Otherwise, it is within a SCSU record. A "node entry" field implies the relay port has a node attached while a "SCSU entry" implies a SCSU attached to the relay port.

The next four fields in turn stores information relating to the four relay ports of the SCSU - addresses of nodes and their status. The status field stores the following flags:



Bit SØ indicates whether the relay port has a node entry or not and whether it has been bypassed in S1. Likewise the status of links attached to the node is contained in S4. The condition of the node and the link attached to it are shown in S3 and S2 respectively. S7 is there to speed up algorithms used for searching the Node Dictionary.

During the operation of the Hierarchical Ring-Star, the contents of the Node Dictionary may change as the system changes state. For example, more ring nodes may be added into the network. Likewise, as faults occur, the relevant status fields alters accordingly. These are carried out dynamically.

6.7 Operation of the Hierarchical Ring-Star

6.7.1 Introduction

Operational tasks of the major components of the Hierarchical Ring-Star have been described in the sections preceding this. This section serves to put them together in a coherent form to present the complete operation of the system.

The MCSU is central to the operation of the Hierarchical Ring-Star, responsible for controlling and coordinating the entire installation. In turn, the MCSU depends critically on the Node Dictionary to provide it with the necessary information. This information is dynamic, changing when and as faults occur or as the configuration alters. Node failures and breaks in ring cables are reported by the Error Logger and SCSUs respectively.

In contrast to node failures, link breaks are usually resolved in-situ by the SCSUs responsible for detecting them. Node failure messages on the other hand must first be processed by a fault algorithm in the MCSU. Only then will a set of suitable relay switching patterns be issued to the relevant SCSUs to isolate the faulty node. Similarly, an operator may request for a particular node or section of ring to be bypassed, and subsequently to be reset. In this case, both the node and the link will have to go through an algorithm in the MCSU in order to generate the appropriate switching patterns.

In the latter case, a user interface is provided to ease the interactive process necessary with the MCSU.

6.7.2 <u>User Interface</u>

A menu driven user interface allows an operator access to facilities provided to control the Hierarchical Ring-Star. The menu driven technique has been implemented because it offers an uninitiated user one of the easiest means to use the system. A set of commands is presented on the terminal display from which a user can select a choice. Whenever data entry is required, the user will be prompted to enter them.

The system provides the following functions:

(a) Set up Node Dictionary

(b) Display system status

(c) Bypass node

(d) Bypass relay port

(e) Bypass link

(f) Bypass SCSU

(g) Reconnect node

(h) Reconnect relay port

(i) Reconnect link

(j) Reconect SCSU

(k) De-activate SCSU

(l) Disable Error Logger

(m) Enable Error Logger

When the system is first installed, data must be entered into the Node Dictionary by invoking the Set up Node Dictionary command. This has been explained in section 6.6. Suffice to say that addressing information is entered.

To see what have been entered, the Display system status command will display contents of the Node Dictionary in a suitable decoded form. If alterations are required, the Set up Node Dictionary command must again be invoked and the entire procedure repeated. Editing facilities are limited. No alterations can be made to the file after coming out of the command. However once complete, the Node Dictionary will be dynamically maintained, either as faults occur or when an operator reconfigures the system.

Reconfiguration of the network structure is usually carried out when expansion or maintenance is required. Commands (c) to (m) have been designed for this purpose. For example a node may need to be removed for testing. By entering the Bypass node command, the system will request for the node's ring address to be input. Once done, and after receiving an acknowledgement, the node can be removed without disrupting the ongoing ring operaton. When that node has been returned to the ring, a Reconnect node command will put it back into operation. Likewise, any link or SCSU can be removed and then reconnected by selecting the appropriate commands from the menu.

For ring expansion, the Bypass link command is normally used to isolate the section of the ring where the extension is required. Once acknowledged that this has been carried out, the isolated link can be cut and extended. When completed the new section can be brought into operation by issuing the Reconnect link command.

Bypass relay port is a command allowing any relay port on the system to be isolated from the ring. This may either be in the MCSU or in any SCSU. In the case of the SCSU, the relay port may have a node attached or it may not. In fact this command was designed to support the latter, and particularly to facilitate the addition of a new node to the ring. Recall that unused relay ports have a loopback plug attached to provide a path for active ring signal. So to add a new node, this path must be isolated first. Once done, the command Reconnect relay port is issued to bring the node into operation. In a similar way, Bypass SCSU and Reconnect SCSU allows a new SCSU to be installed.

De-activate SCSU command has been implemented to support the off-line redundancy technique adopted. When a SCSU is activated to "repair" a fault, it is left in that state even after the fault have been physically rectified and the Reconnect command issued. To comply with the off-line redundancy technique, that SCSU must now be de-activated. The De-activate SCSU command accomplishes this.

In all cases, disruptions to normal ring operation are minimal and should only result in a momentary loss of service. Thus all the commands can be carried out while the ring is in active operation.

A guide for operating the Hierarchical Ring-Star system can be found in Appendix 2. By going through and explaining every item on all the menus, a user is shown, by example, the entire operation of the system.

6.7.3 Operation

The basic tasks of the system have been described in the previous section. To support them, several underlying operations are next described. The sequence of operation is:

(a) receive and decode commands

(b) search the Node Dictionary and update its status

(c) compile a list of switching commands

(d) coordinate and control the distribution of tasks

(e) inform the operator (via the terminal) of the tasks carried out

(f) send a message to the Error Logger to update the fault file

Command codes may be received from three sources: the Error Logger, SCSU and the system operator's terminal.

The operator interacts with the MCSU via the terminal, entering commands to carry out various tasks. Node fault messages are received from the Error Logger while the SCSU sends the MCSU link break informaton. Messages received from the SCSU will contain both the relay port number and its own address. The Error Logger will supply the node's ring address, while the user interface prompts the operator to supply all the necessary addressing information for any particular command.

From the information received, the MCSU searches the Node Dictionary. The Node Dictionary acts as a road map for the MCSU recording the entire network configuration. Once the node, link or SCSU have been located, status flags are updated and depending on this and the status of neighbouring entries, a list of switching actions are compiled. The reason for this is that some faults may require the nodes next to them to switch in tandem in order that they may be bypassed. For example, if node (n) is reported faulty and node (n+1) is found to be previously faulty (and thus bypassed), then the switching operation must divert ring signals between node (n-1) and (n+2) instead of (n+1).

The next step is to distribute the switching tasks. Some tasks are centred on a single SCSU alone but taking the example above, if node (n-1) reside in a different SCSU from that of node (n), then obviously two different sets of switching tasks are required for two SCSUs.

Protocol are employed to distribute the tasks reliably. Switching actions are contained in what is called a CSU Switching Block (CB). It encapsulates the complete list of switching commands. These are sent to and received by the relevant SCSU for immediate action.

On successful completion, the operator is informed by a message displayed on the terminal. Likewise the Error Logger is informed so that its error file may be updated. In all cases, if any switching commands are not carried out, the status flags are reverted to their original state. This usually refers to the situation when the communicaton protocol fails to deliver a CB successfully.

All the above techniques employed will be discussed in detail in the following sections.

6.7.4 Codes

These codes are the common language used by the various units within the Hierarchical Ring-Star for communication. They are divided into four classes, some are command codes while the rest are informative. Command codes are issued by the MCSU to the SCSU requesting switching actions. Information codes are issued by the MCSU, SCSU and Error Logger for passing information between them.

(a) MCSU to SCSU

Commands codes

Command	Code (in hexadecimal)
Bypass node	11
Reconnect node	21 !
Bypass link	12
Reconnect link	22 "
Switch left relay	13
Reset left relay	23 #
Switch right relay	14
Reset right relay	24 $
Switch centre relay	15
Reset centre relay	25 %

These codes directly manipulate relays in the SCSU. Relays are organised into relay ports, each consisting of three relays. Fig. 48 illustrates this: note the left, right and centre relays.

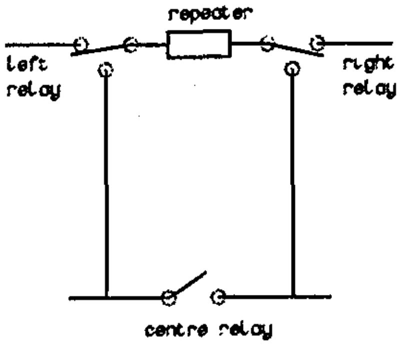


Fig. 48 Relay Port

Bypass node command causes a particular ring node to be bypassed. Bypass link command does the same for a section of the ring. Likewise, the rest of the commands manipulate each relay in a relay port. The Reconnect or Reset command switches the relays back to its normal reset state. Taking an example, a node reported faulty will initially be bypassed and removed. When it has been returned repaired, the reset command will bring it back into operation.

(b) SCSU to MCSU

Information codes

Information	Code (in hexadecimal)
Link break (bypassed)	31
Link break (not repaired)	32

These codes inform the MCSU of link breaks. The first code indicates to the MCSU that the broken link has been bypassed so the MCSU merely needs to update the Node Dictionary. The second code informs the MCSU that it has not been repaired so the MCSU is expected to issue command codes back to the SCSU for the necessary corrective actions as well.

The reason for the two codes is that not all links are under the control of any one SCSU. If a link break occurs between two SCSU, then only one will detect it but both are required to switch together to isolate the link. Since SCSU operations are all independent of each other, the only way is to let the MCSU control the situation.

(c) Error Logger to MCSU

Information codes

Information	Code (in hexadecimal)
Node fault	11

When the Error Logger detects node faults on the ring, these codes convey the information. The MCSU will take corrective actions.

(d) MCSU to Error Logger

Command codes

Command	Code (in hexadecimal)
Disable error detection process	52
Enable error detection process	53

During the period when relays are switched, momentary breaks in the ring cable will occur. Thus the Error Logger may receive false error messages. The first code informs the Error Logger of the impending action, so that error messages are simply ignored till the second code re-enables it. These codes are used, for example, when the operator requests the MCSU to isolate a node.

Information code

Information	Code (in hexadecimal)
Link break occurred	51

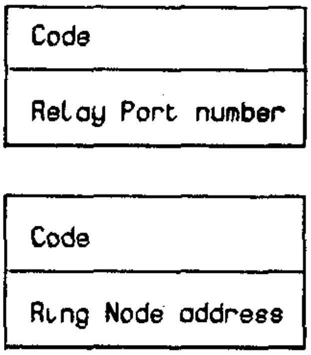
This code informs the Error Logger of a link break so that it can update its error file.

it can update its error file.

6.7.5 Action Blocks

The codes themselves will only inform the recipient of commands or information. Without addressing information, they will be useless. Thus the codes are usually never sent on its own but are, instead, embedded into an Action Block.

An Action Block is a two byte block containing both a code and an address. The address may be a relay port number or a ring node address. These locate switching operations to a particular location in the ring.



Action Block format for MCSU-to-SCSU messages

Action Block format for MCSU-to-Error Logger messages

Fig. 49 Action Block Format

The only exception is when the MCSU informs the Error Logger to switch on or off its error detection process - this requires only a single byte.

6.7.6 CSU Block (CB)

At a higher level is the concept of CSU Block (CB). Depending on whether the receiver is the SCSU or the MSCU, each CB holds the necessary information either required for a single SCSU for all its switching tasks, or for the MCSU.

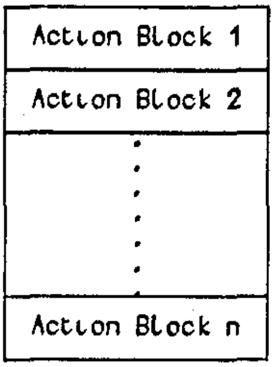


Fig. 50 Switching Block Format

As shown above, the CB is essentially made up of Action Blocks. Each Action Block is responsible for a single operation. Depending on the task, the CB may contain one or more Action Blocks.

6.7.7 MCSU-SCSU Communication Protocol

A communication protocol is employed to pass CBs between the MCSU and SCSU. Although the protocol is kept as simple as possible, it employs techniques to ensure reliable communication.

The simplicity originates from hardware design decisions. First, the transmit and receive channels are physically separated onto two cables so that there can be no collision between transmissions.

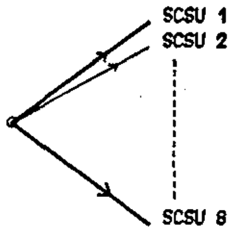


Fig. 51A The MCSU-to-SCSU communication is one to many

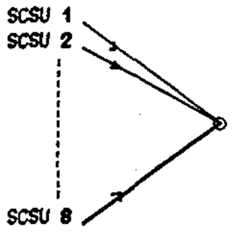
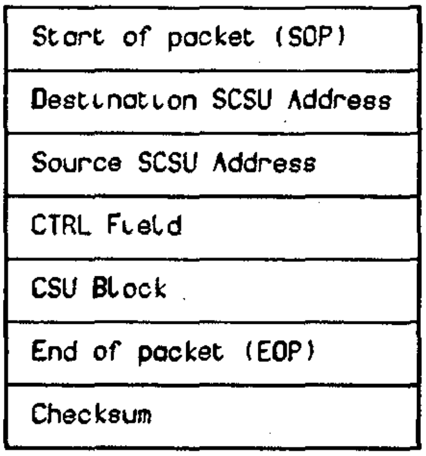


Fig. 51B The SCSU-to-MCSU communication is many to one

As far as transmission of CB between the MCSU to SCSU is concerned, no collision is possible. But the converse is unfortunately true when the transmission is from the SCSU to MCSU. It was deemed uneconomic to provide a separate channel for each SCSU.

The decision taken to implement such a simple technique was based on the small number of SCSU involved and the very different nature of system operation. Complex protocol techniques such as CSMA/CD and tokens are justified for networks with a large number of nodes. The overheads are not justified for a small number of nodes i.e. SCSUs. Also, the application is different. In a computer network, throughput on the channel can be very high but in the Hierarchical Ring-Star, the channels are not even used except when faults are detected and rectified. Also, the only possibility for collision is when multiple faults occur simultaneously. This is not very likely but the protocol employed takes this into account nevertheless.

The communication packet structure and the receiving process is first discussed followed by protocol techniques and finally the transmitting process.



Start of packet (SOP)
Destination SCSU Address
Source SCSU Address
CTRL Field
CSU Block
End of packet (EOP)
Checksum

Fig. 52 Packet structure

SOP and EOP serves to synchronise the transmission processes. If a collision occurs, part of one packet may be destroyed. The receiving process may as a result accept the second packet as part of the first. SOP prevents this. The receiving process always search for the SOP and when found, resets itself to read in the rest of the packet. An earlier packet even if it has not been completely received will be immediately discarded. The sender will eventually discover this by timing out. A watchdog timer is set on transmission and is reset only when an acknowledgement is received. If it is not received within a time period, a re-transmission is initiated.

The packet is read byte by byte till it detects the EOP which signals the end of the packet. The receiving process will then read in the Checksum field. As the bytes are read, a double precision checksum is continuosly computed. When complete, this checksum value is compared to the value given in the Checksum field. If they match, the packet is assumed to be received correctly. An Acknowledgement packet is then returned to the sender. If the checksums do not match, transmission error is assumed and the packet is discarded. The sender will not be informed, but instead expected to re-transmit the same packet when its watchdog timer times out. Note that the same protocol is employed for the MCSU-SCSU communication as for the MCSU-Error Logger communication.

Assuming the packet is received correctly, the receiver will next process the addressing field. This is different for the MCSU and the SCSU. In the case of the SCSU, the receiving process compares the Destination Address field to its own address. If it matches, the CB is meant for this SCSU and is then forwarded to the next process which will execute the Action Blocks enclosed. If it does not match, the CB is simply discarded. The Source Address provides the receiver with the sender's address for the Acknowledgement packet.

In the case of the MCSU, the Destination Address field is obviously irrelevant since it is always zero. The Source Address field is more important since replies must be sent to the proper sender.

In all cases, when a packet is ready for transmission, a "wake-up" signal is first sent to activate the destination CSU. After a short delay the packet is dispatched, setting on a watchdog timer simultaneously. If an acknowledgement is received before the time expires, the transmission is deemed successful. Otherwise, a re-transmission is initiated. A total of ten transmissions is allowed before the whole procedure is abandoned. A message will then be displayed on the terminal to indicate this.

In some instances, for example, when a link break occurs between two SCSU, two CBs have to be sent to each of them. To accommodate this possibility, a one bit sliding window protocol has been implemented for acknowledgement management. This uses a stop-and-wait method, since the sender transmit a packet and then waits for the acknowledgement before sending the next. This helps to synchronise the communication process.

Summarising the protocol, the transmitting process :

(a) encapsulates the CB with the SOP, Source and Destination address, EOP and Checksum to form a packet

(b) sends a "wake-up" signal to the destination

(c) transmit the packet after a short delay

(d) set on the watchdog timer

(e) waits for an acknowledgement

(f) if acknowledgement is received, transmission is complete

(g) otherwise, a re-transmission is initiated

(h) if ten transmissions have been attempted without success, the transmission is abandoned.

(i) a message is displayed - in the case of the MCSU a message is displayed on the operator's terminal. The SCSU will display the pattern 11110000 on the array of light emitting diodes (1=off, 0=on).

And, the receiving process :

(a) searches for the SOP; if found

(b) read in the rest of the packet, otherwise continues searching

(c) while reading the packet, it simultaneously looks for the EOP and sums all bytes

(d) when EOP is found, the Checksum field is extracted and compared to the internally generated value

(e) if they match, the packet is deemed successfully received

(f) and an acknowledgement is returned to the sender

(g) if they do not match, the packet is ignored

6.7.8 Relay Pattern Generation Algorithms

These algorithms produce a set of relay switching patterns according to the type of fault or task required. The algorithms view the entire relay configuration completely when generating the switching patterns. It sees the network configuration represented in the Node Dictionary. Thus the algorithms can be superficially compared to a situation of "navigating through the map of the system." The contents of the switching patterns depends directly on the status of the system.

Depending on the situation, one of the following algorithms are employed:

(a) Link bypass

(b) Link reset

(c) Node bypass

(d) Node reset

(e) SCSU bypass

(f) SCSU reset

Faulty nodes or links are bypassed according to algorithms (c) and (a) respectively. Likewise, commands can be issued to bypass nodes, links or SCSU and to reset them. Since the switching actions required for commands to bypass nodes or links are similar to actions required for isolating faults, the algorithms are equally applicable. Also, edge and non-edge cases are handled in the same way.

The algorithms are described by using flowcharts. The following conventions are used :

Refer to Fig. 53,

The normal position of the two state relay is its initial state.

Switch to Ø implies switching to the other position.

Switch to 1 implies switching back to the normal position.

Subcripts refer to the relative position of relays on each relay port.

( )1 is to the left

( )2 is to the right

( )3 is the centre relay

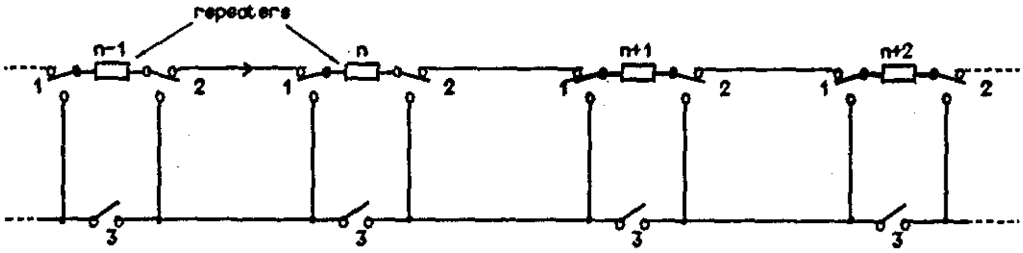


Fig. 53 Conventions used. Relays shown in their NORMAL positions

Either, the nth relay port is the port where the faulty node is attached to or it is the relay port whose ring node reported a link break.

The positions of other relay ports are relative to the nth relay port.

(n+1),(n+2), etc refers to relay ports to the right of this relay port.

(n-1), (n-2), etc refers to relay ports to the left of this relay port.

In both cases, the larger the number the further it is from the reference relay port, n.

All algorithms are based on an analysis of conditions on either side of the item to be bypassed.

RP = Relay Port

6.7.8.1 Node/RP BYPASS algorithm

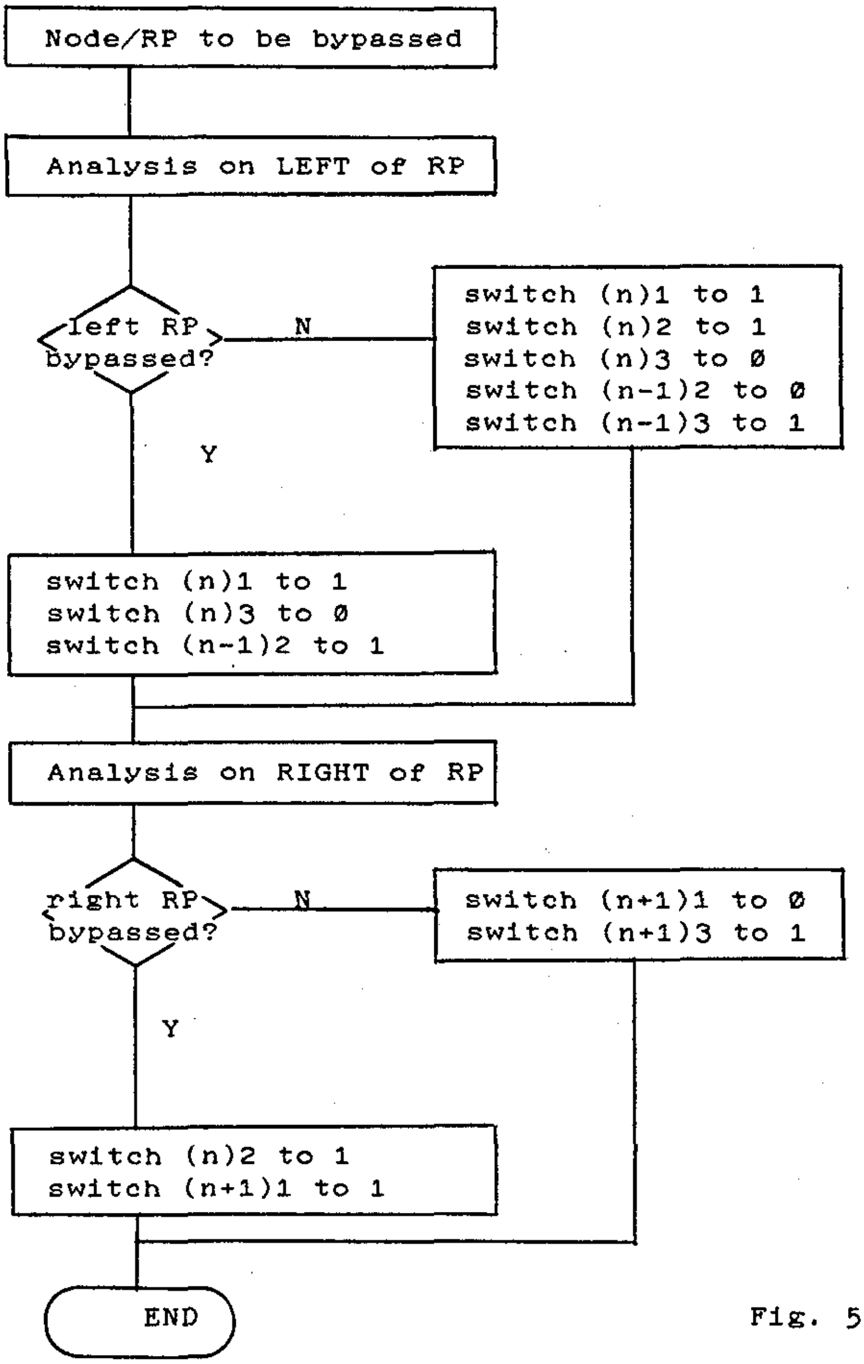


Fig. 54

6.7.8.2 Node/RP RESET algorithm

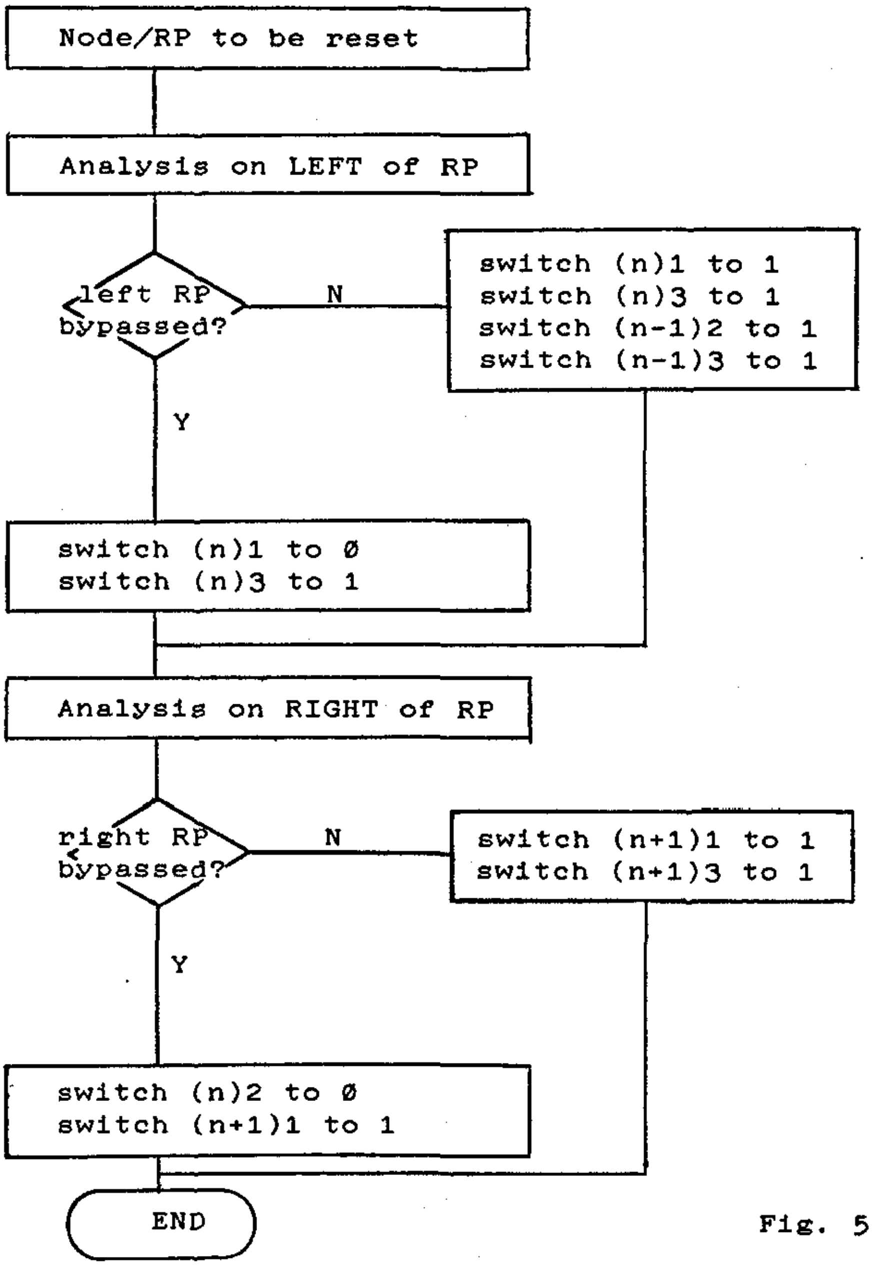


Fig. 55

6.7.8.3 Link BYPASS algorithm

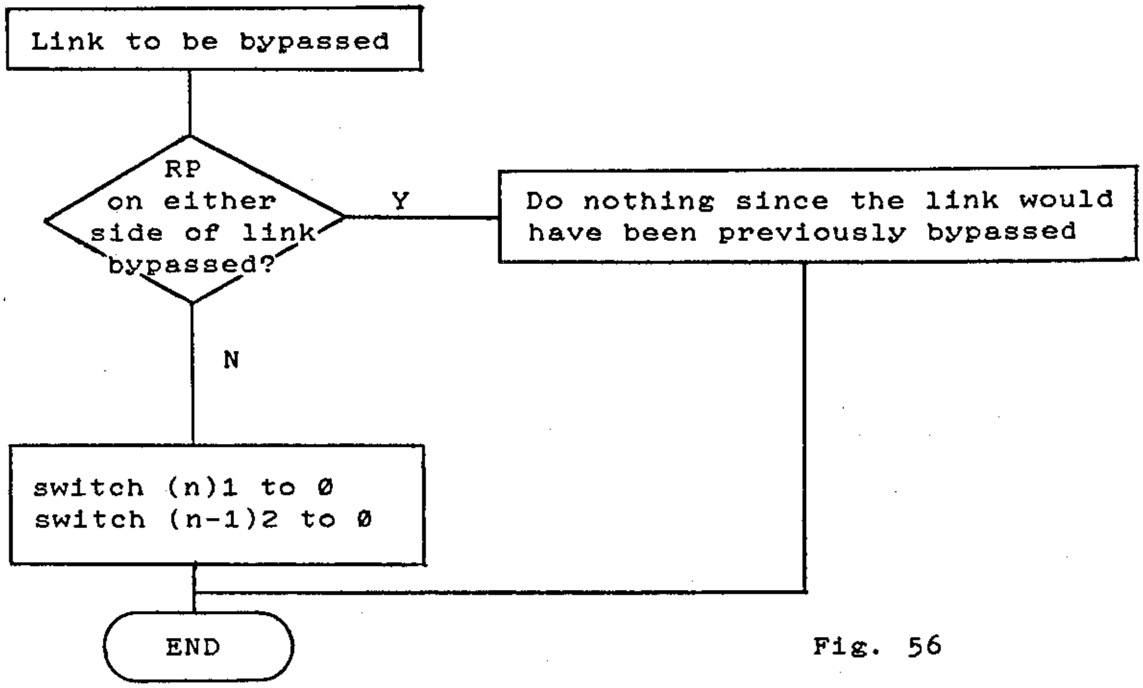

Link to be bypassed
RP on either side of link bypassed?
Y
Do nothing since the link would have been previously bypassed
N
switch (n)1 to Ø
switch (n-1)2 to Ø
END

Fig. 56

6.7.8.4 Link RESET algorithm

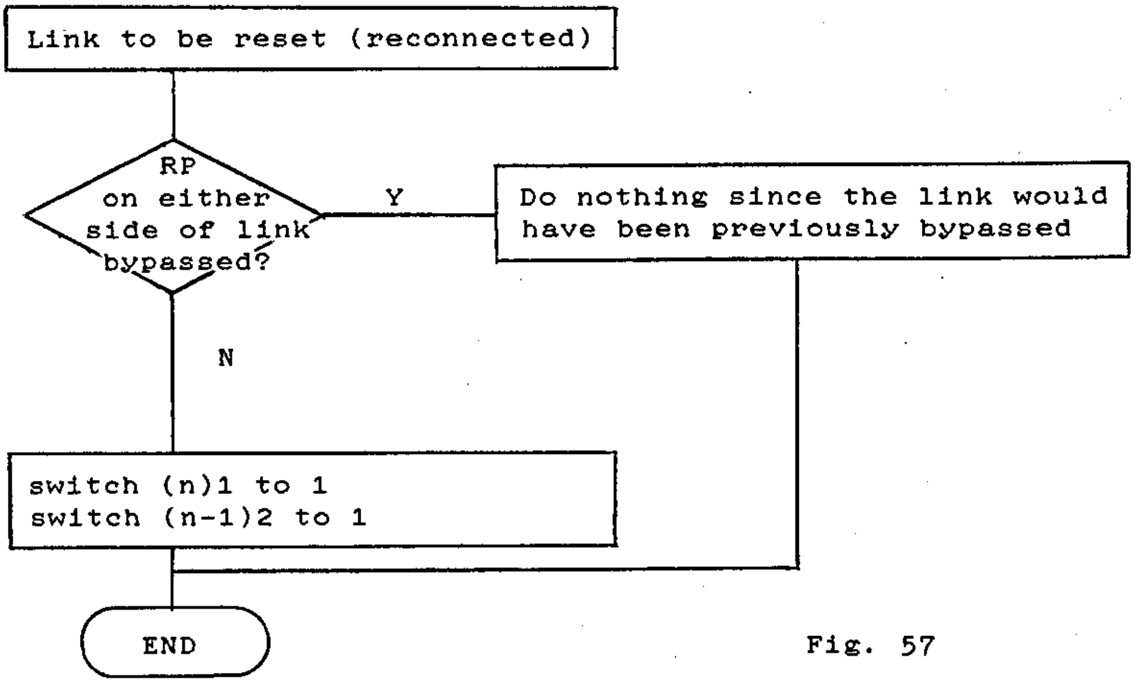

Link to be reset (reconnected)
RP on either side of link bypassed?
Y
Do nothing since the link would have been previously bypassed
N
switch (n)1 to 1
switch (n-1)2 to 1
END

Fig. 57

6.7.8.5 SCSU BYPASS algorithm

More conventions have to be introduced to explain this algorithm.

Referring to Fig. 58,

The first RP on any SCSU is the first RP immediately downstream between two SCSUs.

The last RP on any SCSU is the first RP immediately upstream between two SCSUs.

Relay m1 is the left relay on the MCSU: Relay m2 is the right relay on the MCSU

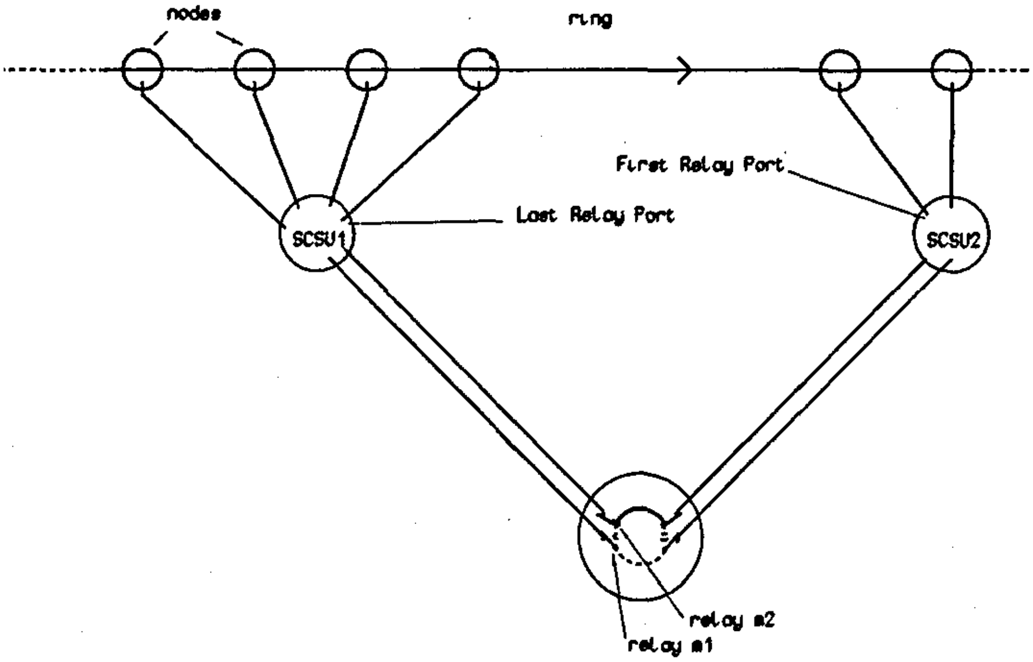


Fig. 58 Conventions

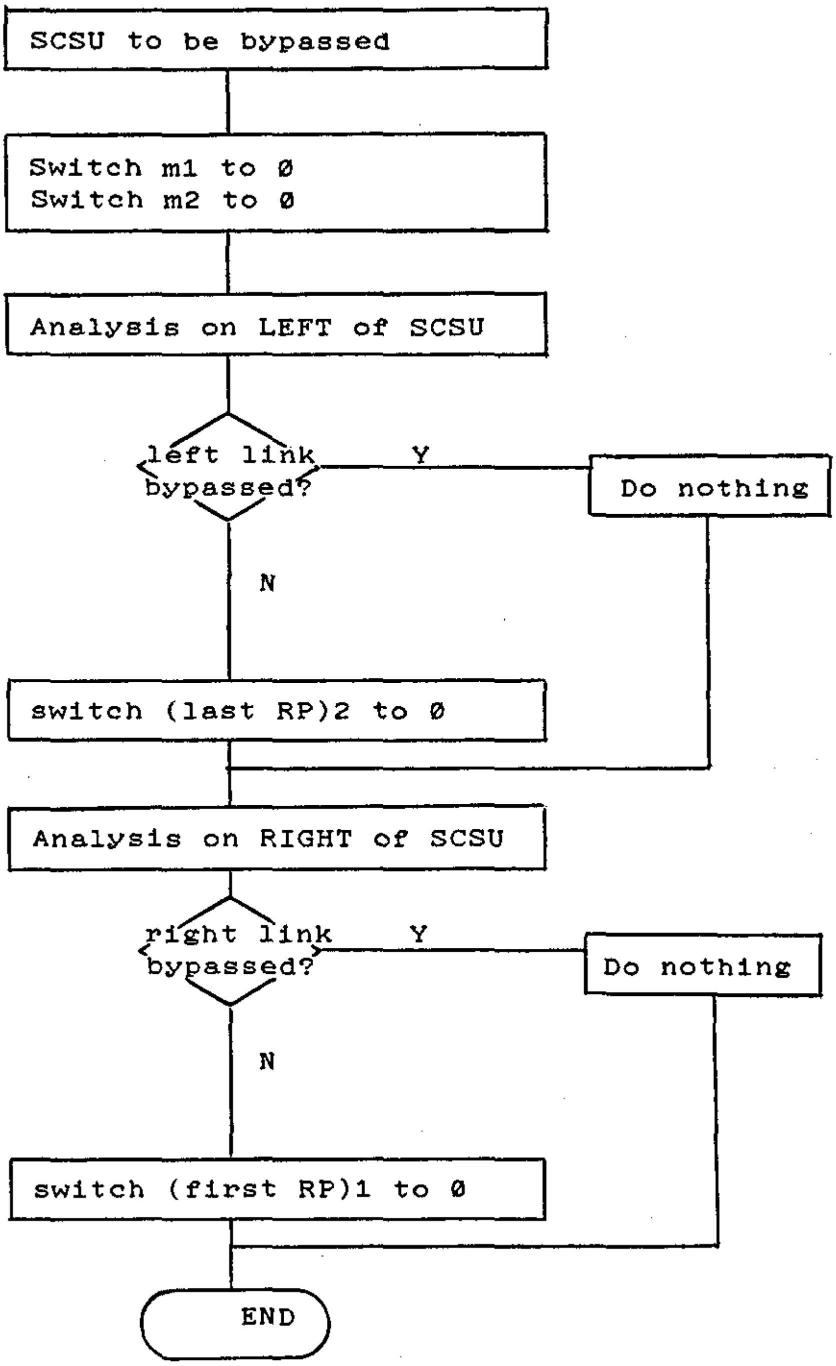

SCSU to be bypassed
Switch m1 to Ø
Switch m2 to Ø
Analysis on LEFT of SCSU
left link bypassed?
Y
Do nothing
N
switch (last RP)2 to Ø
Analysis on RIGHT of SCSU
right link bypassed?
Y
Do nothing
N
switch (first RP)1 to Ø
END

Fig. 59

6.7.8.6 SCSU RESET algorithm

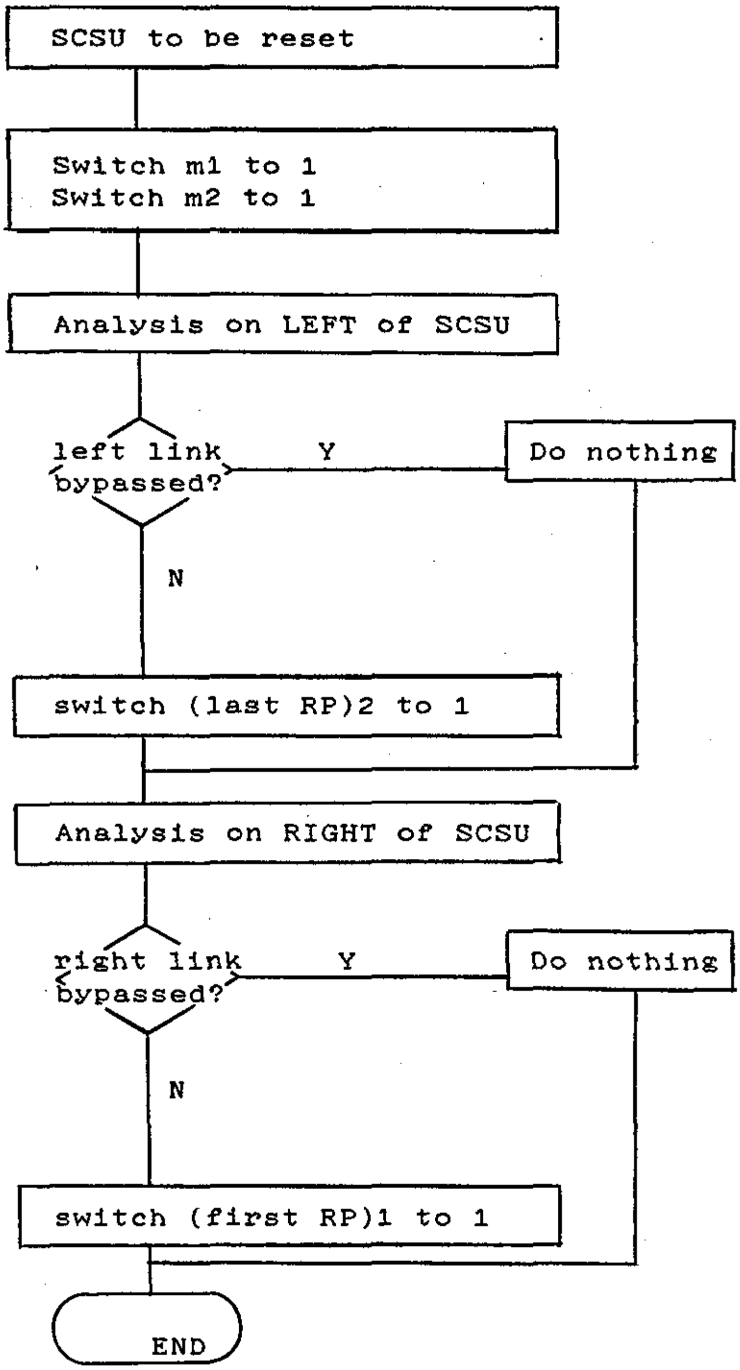


Fig. 60

6.7.8.7 Unique Cases

So far the algorithms have ignored repeater faults. The Cambridge Ring fault detection technique is based on a node, made up of a repeater and an access logic. It is the access logic which detects faults. The repeater by itself does not have the capability to detect faults..

In any installation, it is likely for a ring to include several repeaters without any access logic cards. These repeaters will not detect faults but will instead propagate them downstream to the next repeater or node. If this is a node, there should be no problem. The fault will be detected and the faulty repeater isolated correctly. But if it is another repeater, the fault will be passed on to the next unit downstream. However if this unit is a node, the node bypass algorithm will incorrectly isolate the second repeater leaving the faulty repeater in the ring. Therefore the algorithm will fail if there are more than one repeater between any two nodes in the ring. This problem can be solved by adding an additional watchdog algorithm on top of the basic fault algorithm. The rule is: "If fault messages from the same source continue to be received after corrective action has been carried out, the next upstream relay port should be bypassed. This is recursively executed until no more faults are detected."

An example will clarify this algorithm. Fig. 61A shows a section of a ring with two repeaters between two nodes.

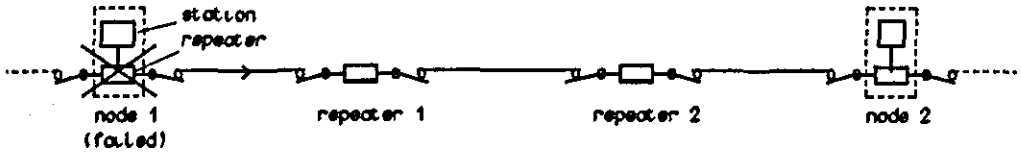


Fig. 61A

If node 1 fails, the fault will propagate through the two repeaters to be detected by node 2. Consequently, repeater 2 will be bypassed and at the same time error messages will continue to be received by node 2. The watchdog algorithm will next bypass repeater 1 and finally node 1.

Next refer to Fig. 61B and consider what happens if the loopback plug LP1 is accidentally removed. Node 1 sees a link break, but again a false one. The link break algorithm will wrongly isolate link L while node 1 continues detecting a link break.

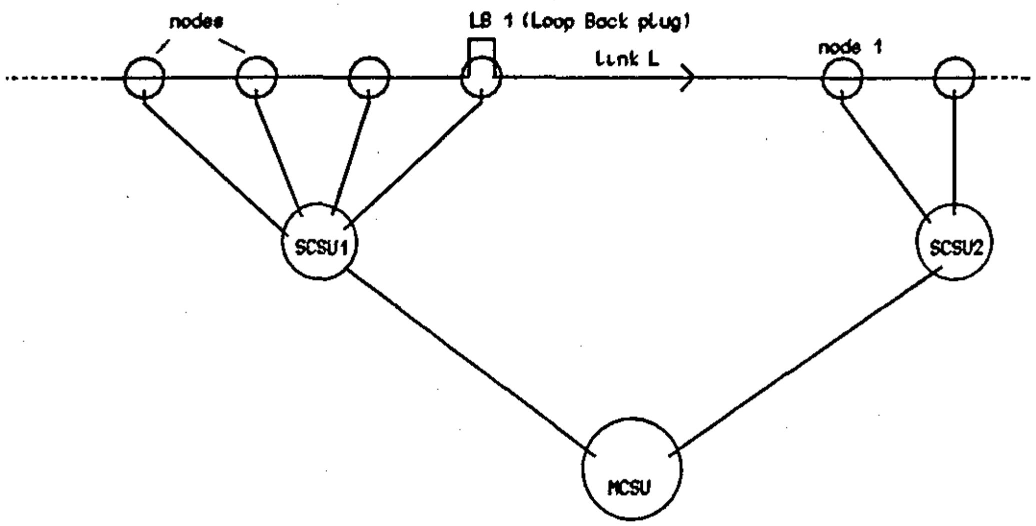


Fig. 61B

Now, consider the case when there are several loopback plugs in use, and LP1 is removed unintentionally.

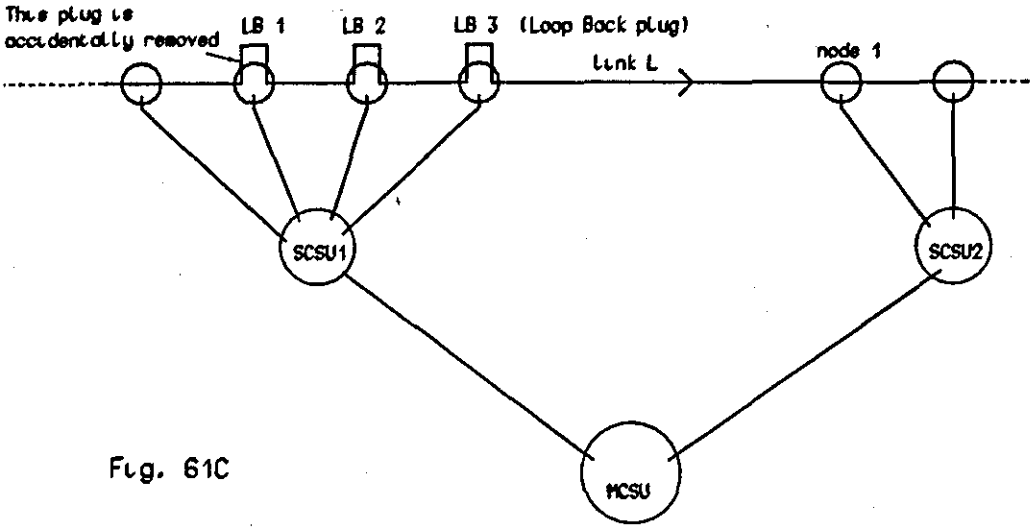


Fig. 61C

These two examples can be solved by applying the watchdog algorithm.

The same algorithm can again be applied to solve the problem of a <u>broken node-to-SCSU cable</u>. Refer to Fig. 61D. If the connection between node 2 and the SCSU is used, say to isolate the link between node 2 and 3, and is then severed, node 3 will detect a link break signal. The watchdog algorithm will bypass node 2 to bypass the break.

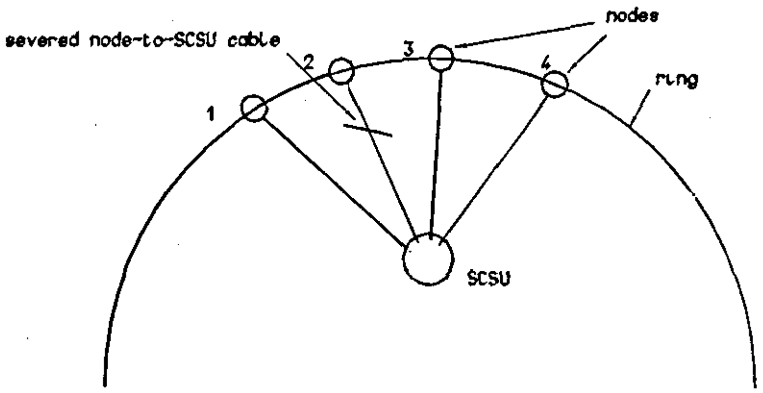


Fig. 61D

However if, instead of a broken node-to-SCSU cable, the SCSU-to-MCSU cable is broken, then the watchdog algorithm will fail.

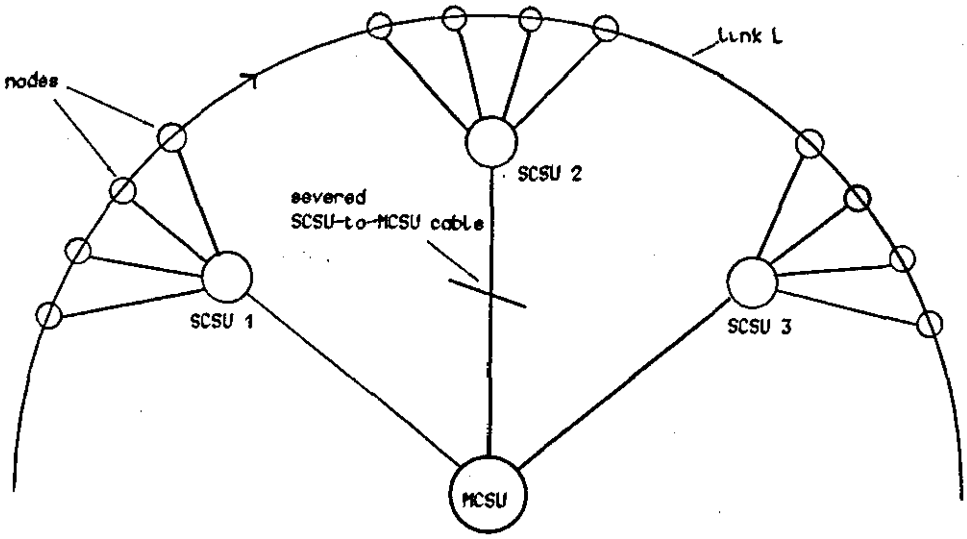


Fig. 61E

To bypass link L in this case will require SCSU2 to be isolated. The watchdog algorithm have been modified to detect and solve this.

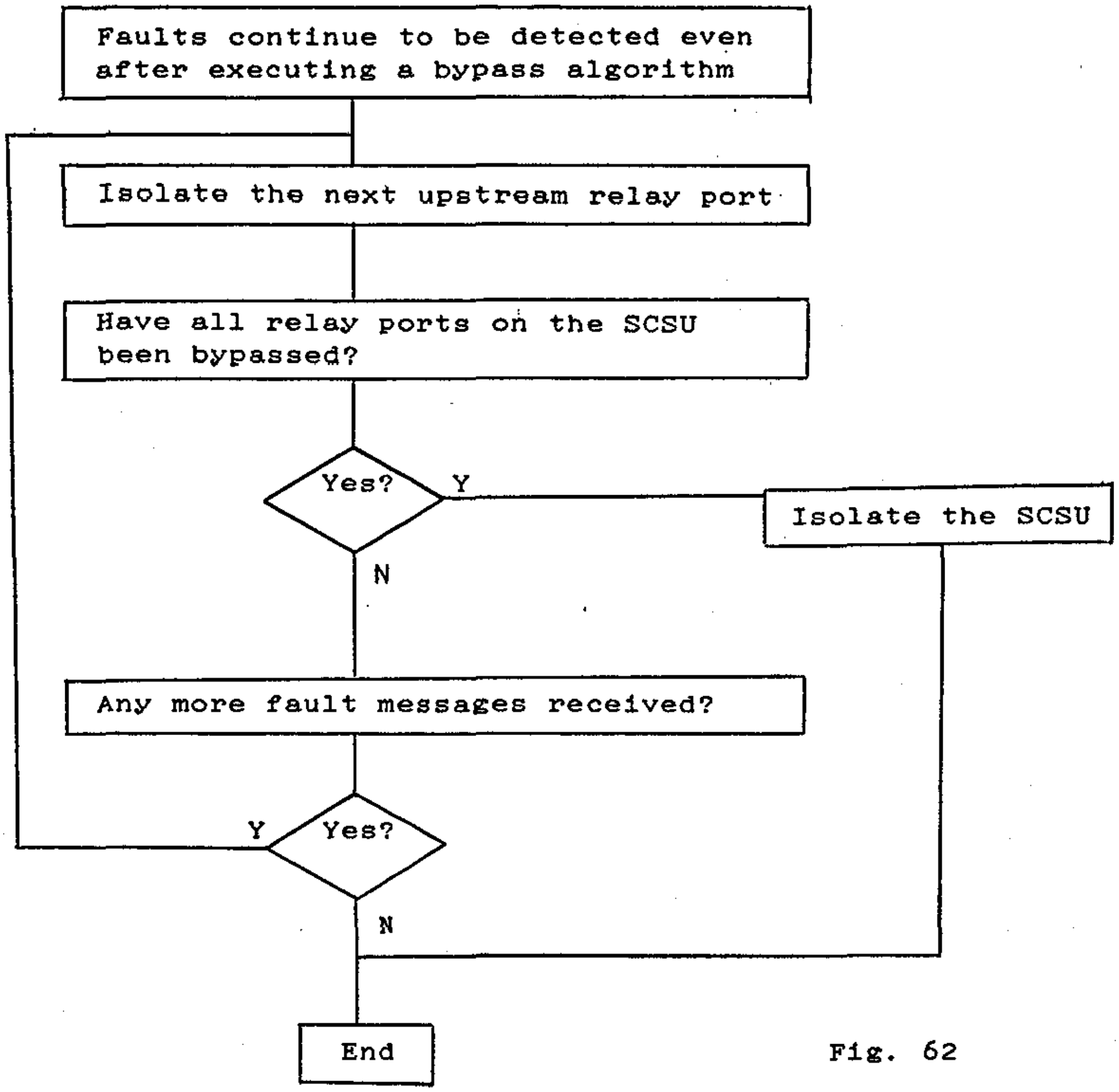

Faults continue to be detected even after executing a bypass algorithm
Isolate the next upstream relay port
Have all relay ports on the SCSU been bypassed?
Yes?
Y
Isolate the SCSU
N
Any more fault messages received?
Y
Yes?
N
End

Fig. 62

6.7.9 Installation Procedure

The connections between the different units are made as follows. The cable which normally links a repeater into the ring must be redirected into one of the D type connector (marked "To Ring") of a Relay Port on the SCSU. The second of the D type connector (marked "To Repeater") must be connected into the D type connector on the repeater. This procedure is carried for all the repeaters on the ring. All Relay Ports must be connected up. Thus if a RP is not actually connected into the ring (spare RP on SCSU), a loop back plug must be used. Each of the SCSU must then be connected into the MCSU by means of the MCSU-SCSU cable. Details of the interconnection scheme and cabling information can be found in Appendix 5.

All SCSUs and ring nodes must have unique addresses, chosen between the range Ø1H to ØFEH. However ØFØH and ØFFH cannot be used as they are reserved for other purposes. The MCSU must be set to address Ø.

The Error Logger must be located immediately upstream of the Monitor. The nodes and SCSUs can be placed anywhere but it would be helpful to cluster a SCSU to every four nodes to reduce cabling requirements. A terminal should be attached to both the Error Logger and the MCSU through the relevant connectors. The Error Logger-MCSU connection must also be made. These three use the RS232 scheme, i.e. 25 pin D type connectors with the following connections : pin

2 to pin 3, pin 3 to pin 2 and pin 7 to pin 7.

Once installed, the next stage is to set up the Node Dictionary. The command "SET UP NODE DICTIONARY" must be issued to the terminal attached to the MCSU. An interactive dialog will be initiated to prompt the operator to enter all the configuration information (see Appendix 2). It would be helpful if the layout of the entire installation is first done on a piece of paper before entering the configuration.

CHAPTER SEVEN

EXPERIMENTAL STUDY OF THE PROTOTYPE HIERARCHICAL RING-STAR SYSTEM

7.1 Experimental Study

The constituent components of the Hierarchical Ring-Star system were continuosly tested during the entire development period. The Error Logger, SCSUs, MCSU and all the communication protocols function correctly at this stage. In particular the concepts and algorithms involved were tested to ensure they worked in practice.

In summary, the following major components and concepts were tested :

(a) Faults - to ensure all link breaks and node failures, either occuring singly or in multiples, are correctly isolated.

(b) Error Logger - to ensure all errors are received and logged correctly and when node failures are detected, to send a message to the MCSU.

(c) SCSU - to ensure that it detects and isolates link breaks and then informs the MCSU of their occurrences. It was also tested to see if it successfully executed switching commands received from the MCSU.

(d) MCSU - to ensure that all its functions are carried out correctly.

(e) Operator controlled tasks - commands to isolate nodes, links and SCSUs were thoroughly tested.

(f) Algorithms - to ensure all algorithms are executed correctly.

(g) Communication protocols - to ensure the protocols deliver messages successfully.

A brief description of the testing procedure follows:
Being a major component of the Ring-Star, the Error Logger is responsible for collecting, logging and detecting ring faults. Testing requires the use of faulty nodes but since none was available, it would have to be physically created. Two problems arose from this - because there were not enough data available on such faults it was difficult to create realistic conditions, and this might prove expensive. A better solution was to emulate a whole range of faults, by programming a node to transmit known error messages to the Error Logger. This way, the whole range of possible faults could be tested, and since the error types were known beforehand the results from the Error Logger could be verified.

A standard Cambridge Ring Z80 Small Server was programmed to function as the Fault Message Generator. This unit was used as a basis for testing the Error Logger and subsequently to test the Ring-Star's ability to tolerate node faults.

One early Error Logger test carried out was to ensure that all ring errors were received and logged correctly and when node failures were detected, to output a message to a terminal. Faulty nodes are diagnosed if x number of errors are detected within a time period y, and in the tests, x and y are variable quantities. The Fault Message Generator generates a fault message with the following format:

Destination Address (DADD) = 0 (i.e. Error Logger address)
Source Address (SADD) = variable to simulate a range of faulty nodes
Error Flags = variable to simulate a range of ring errors

The quantities x, y, SADD and Error Flags were varied. By observing the output of the Error Logger on the terminal attached, results can be evaluated. For example when the following settings were made : x = 20, y = 40 seconds, and several fault messages with SADD set to 02 and error flags totalling more than 20 errors transmitted, the output displayed on the terminal was "node fault detected with address 02." Now, when it was arranged for the messages to be transmitted outside the 40 seconds time limit, no such message was displayed. In all cases, the fault messages generated were logged correctly in the Error Logger. This observation can be made by invoking the "Print Error

Information" command.

Other commands were similarly tested - known inputs were initiated and their outputs monitored. Obviously the tests did not proceed as smoothly as described, numerous problems were encountered. Most of these were engineering problems, and they were resolved as development progressed. Suffice to say the concepts works; an experimental study of the Hierarchical Ring-Star in its entirety will demonstrate that they work in practice but more significantly proves that the Ring-Star concept enhances reliability of the Cambridge Ring.

The configuration below was set up to assess the potential of the Hierarchical Ring-Star system.

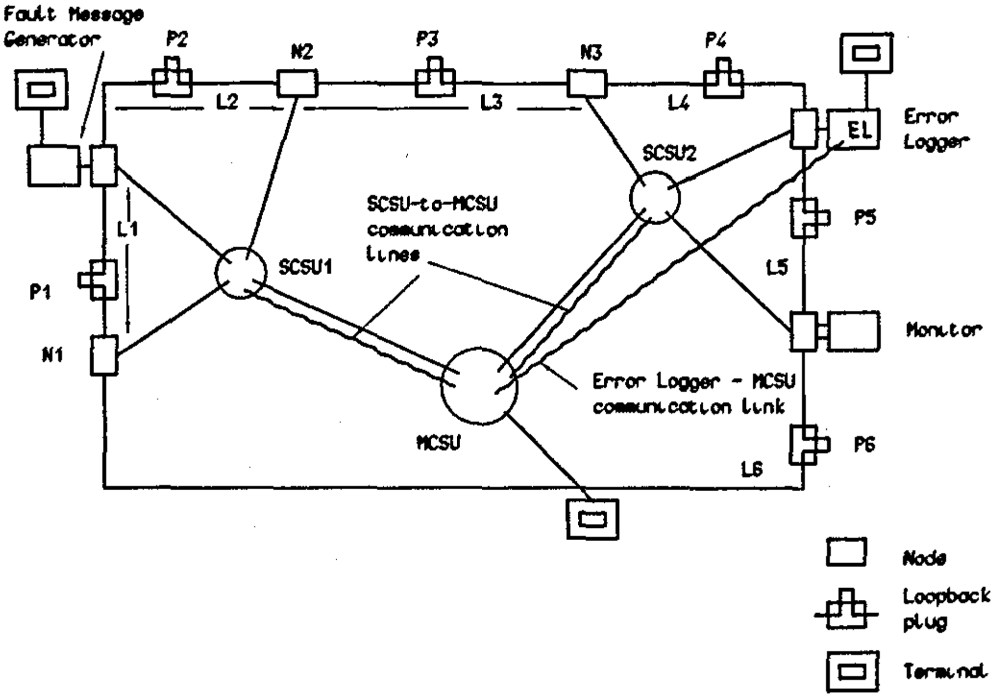


Fig. 63 Configuration of Experimental system

The plugs on the ring cable allow link breaks to be tested. By removing the plugs, one or more breaks in the cable can be affected. One plug was installed to every segment of the links, labelled L1, L2, L3, L4, L5 and L6. To emulate node failures, the Fault Message Generator allows any type of ring error to be transmitted to the Error Logger by entering commands and data through the terminal attached. The terminal connected to the EL displays error information as they are received. Also, error file contents may be displayed by entering a command. The Hierachical Ring Star system can be directly controlled by an operator via the terminal attached.

To illustrate the Ring-Star resilience to cable breaks, plug P1 was removed. This immediately caused link L1 to be bypassed, and a message displayed on the operator's terminal to indicate the fault. By invoking the "Display Node Dictionary" command, the configuration status change can be seen. P2, P4, and P5 have also been tried with success.

Next P3 was removed. Note that this is an edge link. Again the system successfully re-configurated the ring, this time through SCSU1, MCSU and SCSU2.

Finally to evaluate further its effectiveness, P1 was removed before powering up the ring. Then when power was applied to bring the ring into operation, link L1 was automatically isolated. This ability can be very useful in a practical installation. In large networks, it is quite possible for technicians to overlook unconnected segments of the links or loose connectors. The Ring-Star system will in these situations allow the ring to operate but more important, it displays the problem on the terminal and points to its location.

Node failures were emulated with the help of the Fault Message Generator. By instructing the Fault Message Generator to transmit a series of fault packets with the source address set to N1, the Error Logger successfully received and diagnosed the fault. Consequently, the MCSU isolated the node. Again a message was displayed on the terminal and the Node Dictionary updated. An edge node, N2 was next tried. It was isolated through SCSU1, MCSU and SCSU2 successfully. The isolation of faults was carried out almost instantaneously. A delay of a few seconds was however experienced with node isolation (and when longer links were bypassed) due to the re-synchronising process of the Monitor.

The situation when two links broke simultaneously was also examined. By removing L1 and L2 in quick succession, the system was found to repair L1 and L2 in that sequence in succession.

By invoking the "Configuration Mode" from the main menu, operator commands were attempted. Nodes, links, relay ports, and SCSUs were isolated from the ring. These were carried out individually and then several in sequence in various combinations. For example, after node N2 was bypassed, a second command to bypass link L1 was issued. To check if these two commands were executed, plug P1 and node N2 were physically removed from the active ring. The continuing operation of the ring confirms this. They were replaced for the next test, the reset/reconnect commands were executed. N2 was "reconnected" into the ring first. This was visually confirmed by a light emitting diode (led) on the node (when a ring repeater is active on the ring, the led lights up). Next link L1 was reconnected successfully. The Node Dictionary now displays a fully operational ring.

7.2 Discussion

Again the experimental tests did not always work as expected. Various problems did surface. One of the most trying, was line matching problems between the SCSUs and the MCSU communication link. It was finally resolved for the experimental set-up but it cannot be guaranteed for larger configuration. Fortunately, this is an engineering problem which can be resolved with careful installation. Note that this is a common problem with most communication circuits, usually resolved in-situ during installation.

However, once the design faults had been ironed out, the Hierarchical Ring-Star system was shown to work.

CHAPTER EIGHT

EVALUATION

8.1 Cabling requirements

Appendix 1 has derived the formula to calculate cable length requirement for the Hierarchical Ring-Star system (L_H). The Mesh (L_m), Star-shaped ring (L_S), and the Self-Heal ring (L_{SH}) have been included for comparison.

$L_m = 6rR$ where R = Radius of ring

$L_S = 2nR$ n = number of nodes

$L_{SH} = 4\pi R$

$L_H = R[2\pi + 31n/80]$

The plot of cable length, L against the number of nodes, n is shown in Fig. 64.

Several observations can be made. The mesh and self-heal rings' length depends on the size of the installation irrespective of the number of nodes. The cabling requirement of the Star-Shaped ring and the Hierarchical Ring-Star depends directly on the number of nodes and as a result increases with larger n.

The Star-Shaped ring gives the worst results because of its single star centre. Not surprisingly, the Hierarchical Ring-Star shows similar results but they are much less pronounced. The reduction is due to the technique of distributing a number of star centres throughout the ring. For installations with up to 50 nodes, the Hierarchical Ring-Star is comparable to the Mesh and Self-Heal rings but deteriorates after that.

From the above analysis of cable requirements, the Hierarchical Ring-Star may not be acceptable for large networks. But this should not be looked at in isolation.

First, are most installations large? Currently many installations do not contain more than forty nodes. Second, the resilient requirement must be included in the evaluation. Some organisations might be able to tolerate their network being down for a few hours or even a few days. Others might not.

The environment will also affect the network. In most cases, one would not envisage the occurrence of several simultaneous faults. In these situations, a network which could tolerate single faults may be suitable. However if the fault is not easily accessible, it may be left unrepaired. If a second fault then develops, the network will fail. In this case it might be better to implement a more resilient network. Now consider a network installed in a warship. In time of war, the network must be able to tolerate multiple simultaneous failures due to explosions. In such an application, the degree of resilience must be extremely high and the cost involved will be insignificant. This would apply to any applications which involve human lives including for example a nuclear power plant or a chemical factory.

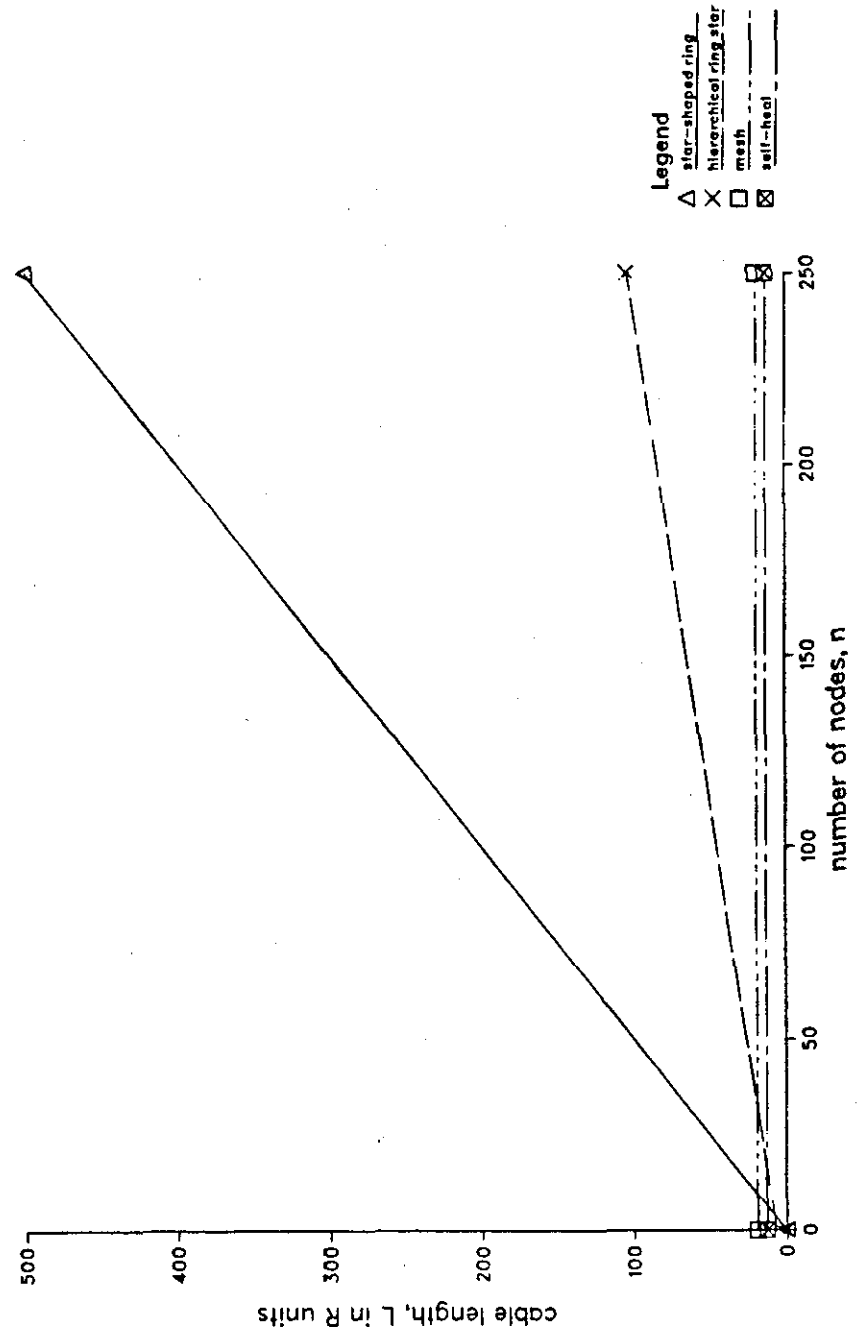


Fig 64. Cabling requirements.

8.2 Cost

The following cost estimate is made for a small quantity of units at 1985 prices.

Master CSU	
Components	£150
PCB	£ 50
Total	£200

Slave CSU	
Components	£140
PCB	£ 50
Total	£190

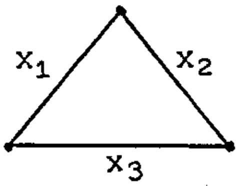
Consider the experimental configuration with 16 nodes. This will require one Master CSU and four Slave CSUs (4 nodes per Slave CSU), costing a total of (4x£190) + £200 = £960. Cost per node = £960/16 = £60. This estimate does not include the extra cabling cost which depends on the size of the network, and it does not take into account normal commercial 'mark-up' on these basic costs.

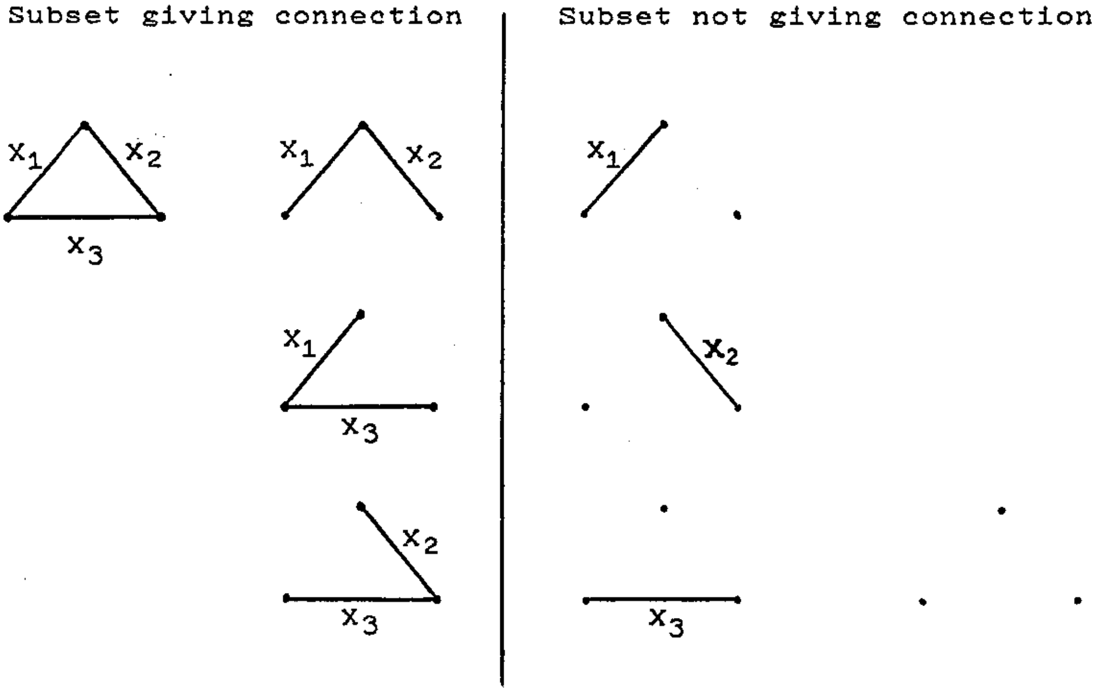
8.3 Reliability Evaluation of the System

The overall reliability of the experimental Hierarchical Ring-Star system is evaluated and compared to the ring without any fault tolerant enhancement.

Reliability calculation is based on a technique called the m-level hierarchical clustering(MHC) [SOI 85]. This technique is especially applicable in that it decomposes the structure of the network into a set of multiple-level hierarchical clusters. Thus it lends itself naturally to the hierarchical Ring-Star topology.

The problem is to determine the reliability of a network comprising a number of nodes and interconnections. For the network to function all nodes must be connected but it assumes that the network has more connections than would be necessary if each connection is completely reliable. In order to determine the probability of all nodes being connected, we must determine the subset of network graphs which maintain connectivity. Alternatively we could obtain the subset of network graphs which do not give connection. Consider a simple 3-node network as an example.




Subset giving connection

Subset not giving connection

If we assume the nodes are reliable, and the probabilities of each connection being sound is p, the probability of not being connected is q = (1-p), then we can proceed to determine the reliability.

For the example given;

$P(x_1, x_2, x_3) = p^3$

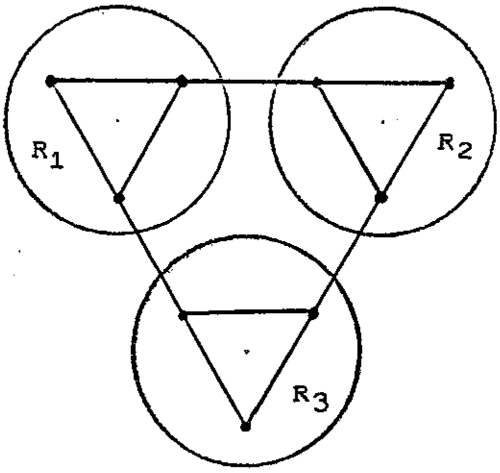
$P(\bar{x}_1, x_2, x_3) = p^2q$

$P(x_1, \bar{x}_2, x_3) = p^2q$

$P(x_1, x_2, \bar{x}_3) = p^2q$

$$\begin{aligned}\text{Total Probability} &= p^3 + 3p^2q \\ &= p^3 + 3p^2(1-p) \\ &= 3p^2 - 2p^3\end{aligned}$$

The task is fairly simple for this case but it gets impossibly tedious for a complicated network. The MHC technique reduces this problem and although not completely accurate, it does provide a conservative answer. This method divides the network into clusters and then treat each cluster as a node with a given probability of working. For large networks, this process can be repeated giving a hierarchy of clusters. For example, the 9-node network below can be divided into three clusters; R_1, R_2 and R_3.



The reliability of each of the networks inside the cluster can be determined as before i.e.

$$R_1 = R_2 = R_3 = 3p^2 - 2p^3$$

The reliability of the whole network, R is then given by $R_1R_2R_3$ times the reliability of the network using clusters as nodes,

$$\text{i.e.} \quad R = R_1R_2R_3(3p^2 - 2p^3) = (3p^2 - 2p^3)^4$$

The key issue is how the nodes should be clustered. The clustering technique is generally defined as one of finding natural groupings of a set of nodes. This involves two separate issues:

(i) the similarity (or nearness) between two nodes

(ii) how to partition a set of nodes into clusters

As overall reliability decreases with an increase in the diameter of the network (SOI 85), it is obvious that clusters should be chosen to correspond to highly connected sets of nodes, which result in a small diameter. Further, since reliability evaluation is now dependent on spanning trees internal to the cluster, the cluster must contain the shortest paths between its nodes.

The Hierarchical Ring-Star fits naturally into the above two criteria since its topology has itself been designed to form clusters. Nodes within clusters are chosen to minimize distances between them and each node-to-SCSU cluster is highly connected. Thus each SCSU forms a natural cluster with the nodes connected into them.

Levels of clusters must be formed. Basically, an m-level hierarchical clustering of a set of nodes consists of grouping the nodes into 1st level clusters, which in turn are grouped into 2nd level clusters, etc. This operation continues in a bottom-up fashion, finally grouping the m-2nd level clusters into m-1st level clusters, whose union constitutes the mth level cluster. The mth level cluster is the highest level cluster and as such it includes all the nodes of the network.

8.3.1 Numerical evaluation of the system

The 16 nodes 2-level Hierarchical Ring-Star is clustered into four levels as shown in Fig. 64 for the evaluation.

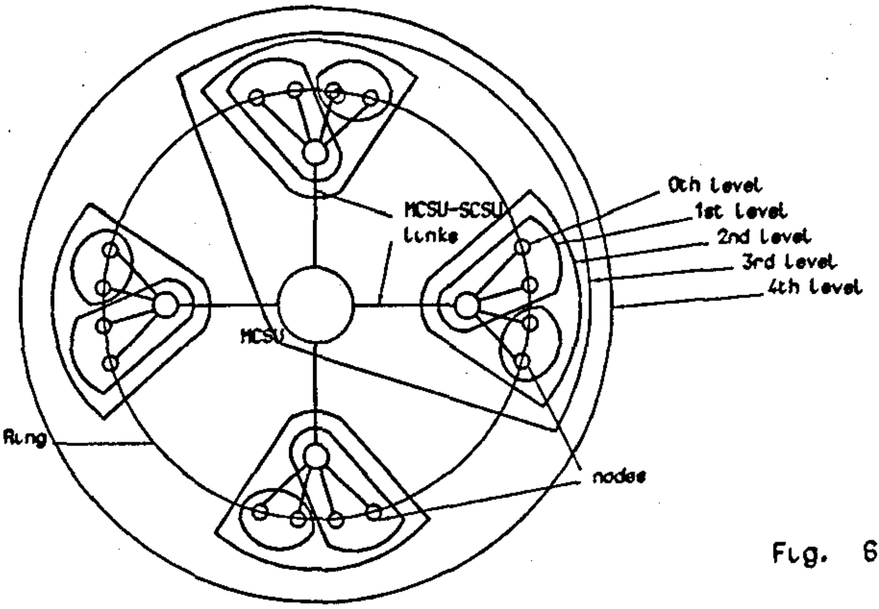
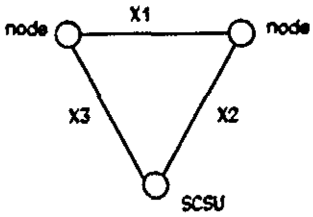


Fig. 64

Assume all links have the same reliability p, and unreliability q, all nodes have the same reliability n, all SCSUs have the same reliability s, and the MCSU has a reliability of m. These assumptions are made for the sake of mathematical simplicity.

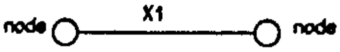
<u>Step 1</u>: evaluate the reliability of the 1st level cluster.



The spanning trees are X1X2, X1X3, and X2X3.

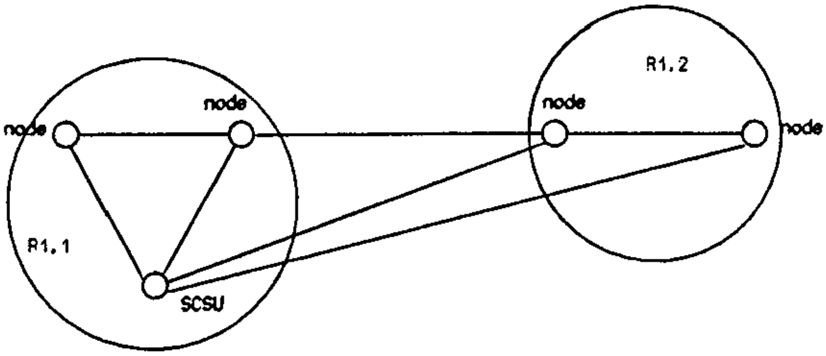
$$\begin{aligned}\text{Reliability (links only)} &= \Pr\{X1X2 \cup X1X3 \cup X2X3\} \\ &= p^2 + p^2q + p^2q \\ &= p^2 + 2p^2(1-p) \\ &= 3p^2 - 2p^3\end{aligned}$$

$$\begin{aligned}\text{Overall Reliability, R1.1} &= n.n.s(3p^2 - 2p^3) \\ &= n^2s(3p^2 - 2p^3)\text{............(i)}\end{aligned}$$



$$R_{1.2} = n^2p \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(\text{ii})$$

Step 2: evaluate the reliability of the 2nd level cluster.



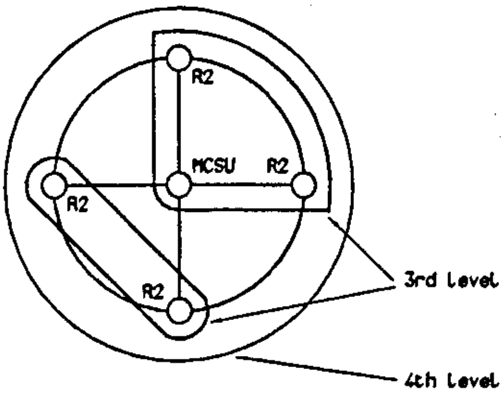
Each second level cluster is really two nodes with reliability, $R_{1.1}$ and $R_{1.2}$ interconnected by 3 links in parallel.

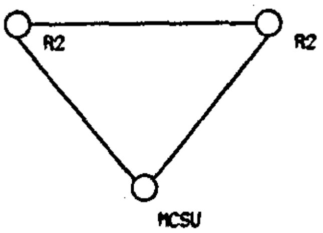
$$R_2 = R_{1.1}[1 - (1-p)^3]R_{1.2}$$

$$= [(3p^2-2p^3)n^2s][n^2p][3p - 3p^2 + p^3]$$

$$= n^4s[9p^4 - 15p^5 + 9p^6 - 2p^7]\ldots\ldots\ldots\ldots(\text{iii})$$

Step 3: evaluate the reliability of the 3rd level cluster.

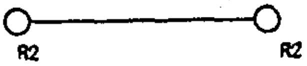




This structue is similar to (i) except that the SCSU is replaced with a MCSU.

Thus,

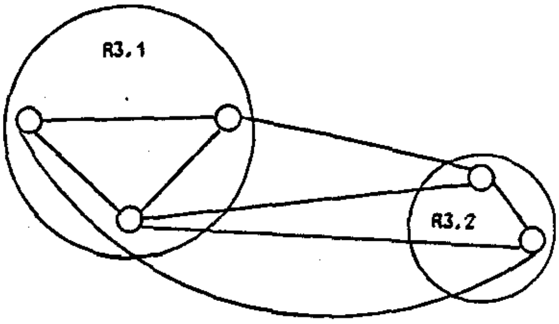
$$R_{3.1} = (3p^2 - 2p^3)n^2m \text{ .... (iv)}$$



This structure is similar to (ii)

$$R_{3.2} = R_2R_2p \text{ .... (v)}$$

Step 4: evaluate the reliability of the 4th level cluster.



$R_{3.1}$ and $R_{3.2}$ are interconnected by four links in parallel.

$$R = R_{3.1}R_{3.2}[1 - (1 - p)^4]$$

$$= R_{3.1}R_{3.2}[4p - 6p^2 + 4p^3 - p^4] \text{ .... (vi)}$$

Substituting (iii), (iv), and (v) into (vi), the reliability of the 16 nodes 2-Level Hierarchical Ring-Star system is,

$$R = [(3p^2 - 2p^3)n^2m][R_2R_2p][4p - 6p^2 + 4p^3 - p^4]$$

where,

$$R_2 = n^4s[9p^4 - 15p^5 + 9p^6 - 2p^7]$$

With off-line redundancy, the value of R is improved further by approximately 23%

Not enough data is available to enumerate values for p,m,n and s in the Cambridge Ring environment. Instead, a range of possible values is substituted to evaluate the reliability of the experimental Hierarchical Ring-Star system (HRS). The results are depicted in Table 2. The reliability of the HRS with and without off-line redundancy is evaluated. For comparison, the same ring without any fault tolerance is included. In this case the ring is simply a serial network with reliability $p^{16}n^{16}$. The improvement factor (R2/R3) is computed and presented in Table 2.

The evaluation shows that with less reliable components, the overall reliability of the Hierarchical Ring-Star is low. As component reliability improves, the incremental improvement in reliability of the Hierarchical Ring-Star is higher till they are equal at unity. The opposite is true with the improvement factor R2/R3. It goes lower as component reliability improves. This is obvious since with 100% reliable parts, a machine must in whole be 100% reliable. Thus the Hierarchical Ring-Star is most effective when ring nodes and/or links are less reliable. But in practice, this may be questionable. At realistic values of node and link reliability (0.98 or above), the reliability improvement with the Hierarchical Ring-Star is not too significant at approximately 50% better. However the most important consideration is the overall reliability. In this case, it approaches 90%.

p	m	n	s	R1 without off-line redundancy	R2 with off-line redundancy	R3 basic ring without fault tolerance	R2 —— R3
0.90	0.90	0.90	0.90	0.17	0.21	0.03	7.00
0.91	0.90	0.93	0.91	0.25	0.31	0.07	4.40
0.92	0.91	0.93	0.92	0.27	0.34	0.08	4.25
0.93	0.93	0.95	0.92	0.36	0.45	0.14	3.2
0.94	0.92	0.95	0.92	0.38	0.46	0.16	2.88
0.95	0.95	0.95	0.95	0.43	0.53	0.19	2.79
0.95	0.93	0.96	0.93	0.45	0.55	0.23	2.39
0.96	0.93	0.97	0.92	0.51	0.62	0.32	1.94
0.96	0.92	0.97	0.94	0.52	0.64	0.32	2.00
0.97	0.92	0.97	0.95	0.55	0.68	0.38	1.79
0.97	0.96	0.98	0.95	0.64	0.79	0.44	1.8
0.98	0.93	0.98	0.95	0.64	0.79	0.52	1.52
0.98	0.94	0.98	0.96	0.66	0.82	0.52	1.58
0.99	0.96	0.98	0.97	0.72	0.88	0.62	1.42
0.99	0.99	0.99	0.99	0.85	1.05(i.e. 1)	0.72	1.39
1.00	1.00	1.00	1.00	1.00	1.23(i.e. 1)	1.00	1.00

TABLE 2

CHAPTER NINE

DISCUSSION, CONCLUSION AND SUGGESTIONS FOR FURTHER WORK

9.1 General Comments

The Ring-Star concept has been developed to enhance the Cambridge Ring with fault tolerance, and experiments have demonstrated its effectiveness. However it must be pointed out that the Ring-Star deals only with the communication channel. In this sense although the Monitor is an integral unit of the Cambridge Ring it has not been considered. Thus faulty nodes and broken ring cables should not disrupt ring operation but an imperfect Monitor will.

One solution is to apply fault tolerant techniques such as Standby Replacement or Triple Modular Redundancy (described in chapter 2) to the Monitor. These techniques are expensive and bearing in mind the Monitor itself is reliable, their use should be cost justified. This is an area for further experimentation.

This single point failure problem extends to the CSU (SCSU and MCSU) and the Error Logger. Once the system has been brought into operation to 'heal' a ring fault, reliability is determined by the probability that the CSU and to a lesser degree the Error Logger fails within the time period required to rectify the fault. However, since normally this may take only a few minutes (e.g. replacing the faulty node with another while it is being repaired), high reliability can be achieved over long periods of time. Dependence on the CSU is not as catastrophic as was first thought. Using off-line redundancy, the CSU does not even play a part in normal operation. Similarly, the Error Logger is not absolutely crucial. Instead of totally relying on the Error Logger to detect all ring faults as was originally envisaged, it is now only responsible for detecting node faults. The task of link break detection has been delegated to the SCSUs.

The same point must also be made about the Name Server and the Boot Server. Although strictly not required in the operation of the Cambridge Ring, their use has been built into most protocols. For example, the Single Shot Protocol requires a device to obtain from the Name Server the ring address of the device it wishes to communicate with before proceeding to establish a link. Therefore a totally secure Cambridge Ring system requires a fault tolerant Name Server and a Boot Server too. Fault tolerant network servers are another potential area for further work.

Several general points can be made of the Hierarchical Ring-Star:

(a) Distributed fault detection

Critics might argue that the Hierarchical Ring-Star is based on a single MCSU. Although central to its operation in that the MCSU stores all the configuration information and allows an operator to control its topology, its primary role is different. It is only directly responsible for controlling node failures. Link breaks which are probably the more likely of the two faults are controlled by SCSUs. The exception is edge cases.

(b) Enforces record keeping

In any installation, keeping documentaton up to date may present a problem. Personnel in charge may adopt a blase attitude, resulting in an inaccurately kept record. Future maintenance or expansion may as a result be difficult. With the Ring-Star concept, control is enforced since the Node Dictionary must be set up during installation. (This is then dynamically maintained.) Since the Node Dictionary is a "road map" of the network, configuration information is then always available.

(c) Performance

Peformance of the Cambridge Ring will alter as nodes or links are bypassed. The reason is that slot structure may change due to changes in the ring delay. As a result the number of minipackets in circulation may increase or decrease. It has been shown that this affects ring peformance (Blair 83).

(d) Network size

The maximum inter-node cable length will be reduced compared to a basic ring because the signal path must divert into the CSU. This is a consequence of the star structure.

9.2 Independent of Technology

The Hierarchical Ring-Star has been designed for wide applications. Because of the way it has been designed (based on topology), it can easily be adapted for use with any other type of ring technology. Thus token rings, register-insertion rings or any other future proprietary rings may take advantage of the methodology presented in this thesis.

Similarly, data transmission speed is irrelevant. It makes no difference to the Hierarchical Ring-Star if the speed is 1K baud or 1ØØM baud. However the higher speeds may necessitate the use of optical fibre as the transmission medium. In this case, the mechanical relays in the CSUs should be replaced by their fibre optic equivalent.

9.3 Conclusions

This thesis has presented a unique technique to overcome the potential reliability problems of the Cambridge Ring. The Ring-Star concept in the form of the Hierarchical Ring-Star is proposed as a fault tolerant enhancement for the Cambridge Ring. Experiments have shown that it works.

9.3 Conclusions

This thesis has presented a unique technique to overcome the potential reliability problems of the Cambridge Ring. The Ring-Star concept in the form of the Hierarchical Ring-Star is proposed as a fault tolerant enhancement for the Cambridge Ring. Experiments have shown that it works.

In the early stages of the project, a literature review was initiated to carry out a detailed study of fault tolerant rings, with the aim of identifying key design issues. They were analysed, and taking into account the Cambridge Ring technology, a set of objectives was produced. The Hierarchical Ring-Star system has been developed to meet these objectives. How have these objectives been achieved?

Realising that ring networks have a weak structure, the first phase of the work focussed on the topology. The idea of the Ring-Star arose because it was realised that rings and star topologies each have unique strengths which complemented one another. Rings were invented for their performance improvements in data communication but at the expense of topological reliability. Star topology on the other hand has a stong structure in that any device connected into it may fail without any consequence to the rest of the network. Combining the two topologies creates a network which is both efficient and reliable.

This topology and the ensuing design satisfies the objectives. First and foremost, node and link breaks would not bring down the entire network, as only the affected sections are isolated. Dynamic reconfiguration ensures this. Second, it has a high degree of resilience in that multiple faults are tolerated and thirdly, fragmentation of the network is minimised. This in fact is the key factor in favour of the Ring-Star, many of the other fault tolerant rings can only tolerate one fault. A second fault will either bring down the network or cause the network to be divided into two isolated segments. The objective of operational independence between the fault tolerant component and the ring was however not totally achieved. To ensure the effective detection of broken cables, a modification had to be made to the repeater. This is necessary because the Cambridge Ring was found to be less effective in detecting link breaks.

Installing the system into an existing Cambridge Ring is relatively easy. It is a matter of redirecting the node-to-ring cable from the ring into the SCSU and connecting a cable from the SCSU into the ring. And once installed, future expansion is relatively easy.

The system supports maintenance in two ways. First, it provides diagnostic information through the Node Dictionary. The Node Dictionary stores and updates network configuration details as changes take place. For example if a node should fail, its location is recorded so that a technician can easily find the faulty node instead of having to trace through the ring for it. Second, once located, the node can simply be removed without having to worry about its effect on the ring. It is this latter feature that makes expansion easy. The section of the ring to be extended is first isolated, new nodes are then added before a command is issued to the MCSU to bring them into operation.

This convenience extends to users. Whenever reconfiguration of the network is carried out, users would only experience a delay of a few seconds. Compare this to a basic ring which requires the network to be taken out of service, and the fault rectified before normal service can be restored. This may take days. In short, the Hierarchical Ring-Star allows almost non-stop operation. Finally, by keeping the fault tolerant component of the system independent from network technology, it is better protected from obsolescence arising from future developments. For the same reason, the Hierarchical Ring-Star is not limited to the Cambridge Ring. It can easily be adapted for any other ring.

By adopting the Hierarchical Ring-Star, the Cambridge Ring could resolve its greatest disadvantage - topological unreliability. In fact any potential network implementor should consider reliability problems seriously. The impact may not be felt until it is too late. Perhaps the following quotation best sums it up:

"For many applications fault tolerance is no longer regarded as a bonus - it is essential."

- Rob Summerfield

Minicomputer News Vol 8 no 10 October 1985

REFERENCES

AVIZIENIS 85 A. Avizienis, C.S. Raghavendra and M. Gerla "Reliable Loop Topologies for Large Local Computer Networks", IEEE Transactions on Computers, Vol C-34, no.1, Jan. 1985

BINNS 82 S.E. Binns, I.N. Dallas and E.B. Spratt "Further Developments on the Cambridge Ring Network at the University of Kent", Local Area Networks, Proceedings of the IFIP TC6 International In-depth Symposium on Local Computer Networks, Florence, Italy, April 1982

BLAIR 82 G.S. Blair "A Performance Study of the Cambridge Ring", Computer Networks, vol 6, 1982

BUX 82 Bux et al "A Local Area Communication Network Based on a Reliable Token Ring System", Local Computer Networks, IFIP 1982.

CACCETTA 84 L. Caccetta "Vulnerability of Communication Networks", Networks Vol 14 1984

CHEN 85 T.N. Chen "Ring Network and a Fault-Tolerant Cambridge Ring Architecture", International Conference on Network and Electronic Office Systems, London, September 1985, IERE Publication no 63

CLOSS 81 F. Closs and R.P. Lee "A Multi-Star Broadcast Network for Local Area Communication", Local Networks for Computer Communication, IFIP 1981

COOPER 84 R. Cooper and A. Hart "Sweeping Cables Under the Carpet", Designer's Journal, Feb. 1984

CR 82 Cambridge Ring 82 Interface Specifications, SERC and Joint Network Team, 1982

DAMSKER 82 D. Damsker "Totally Distributed, Redundantly Structured Hardware and Software Local Computer Controlled Network", Local Computer Networks, IFIP 1982

DAS 83 S.K. Das, K.W. Blackmun and M.C. Mak "Name Serving, Error Logging and Interfacing on City University's Cambridge Ring Network", International Conference on Networks and Electronic Office Systems, Reading University, Sep. 1983. IERE Publication no 57

DAVIES 84 P.A. Davies and F.A. Ghani "The Application of Optical Fibres to Local Area Ring Network", Ring Technology Local Area Networks, IFIP, 1984

DIXON 83 R.C. Dixon, N.C. Strole and J.D. Markov "A Token Ring Network for Local Data Communications", IBM Systems Journal, vol 22 nos 1/2, 1983

FALCONER 84 R.M. Falconer "A Study of Techniques for Enhancing the Reliability of Ring Local Area Network (LAN's)", Ring Technology Local Area Networks, IFIP, 1984

FRANKEL 84 E.G. Frankel "Systems Reliability and Risk Analysis", Martinus Nijhoff Publishers, 1984

GRAF 84 H.A. Graf, N.F. Geer and R.T. Moore "Gridnet : An Alternative Large Distributed Network", Computer, April 1984

HAFNER 76 E. Hafner "Enhancing the Availability of a Loop System by Meshing", International Zurich Seminar on Digital Communications, 1976

HOPPER A. Hopper and D.J. Wheeler "Maintenance of Ring Communication Systems", Internal report, Computer Laboratory, University of Cambridge

JOHNSON 84 R. Johnson "Network Reliability and Acyclic Orientation", Networks Vol 14 1984

B JOHNSON 84 B.W. Johnson "Fault-Tolerant Microprocessor-Based System" IEEE Micro, December 1984

KIRK 84 P.R. Kirk, A.N. Slater and S.R. Forman "A Reconfigurable Fault-Tolerant Ring Network", Project Universe Conference, Loughborough University, 1984

LIU 84 M.T. Liu and D.M. Rouse
"A Study of Ring Networks", Ring Technology Local Area Networks, IFIP 1984

LOSQ 76 J. Losq
"A Highly Efficient Redundancy Scheme : Self-Purging Redundancy", IEEE Trans. Computers, Vol C-25, no 6, June 1976

MEISER 82 N.B. Meiser
"Methodology for Assessing the Robustness of a Local Network Based Computer System", Local Computer Networks, IFIP 1982

MAURICE 79 M.V. Maurice
"The Cambridge Digital Communication Ring", Local Area Communications Network Symposium, Boston, M.A., May 1979

NEEDHAM 79 R.M. Needham
"Systems Aspects of the Cambridge Ring", Proceedings of the Seventh Symposium on Operating Systems Principles, Pacific Grove, California, December 1979

O'CONNOR 85 P.D.T. O'Connor
"Practical Reliability Engineering", John Wiley & Sons, 1985

PERSONICK 85 S.D. Personick
"Switches take to Optics", Electronics Week, March 18, 1985

PIERCE 72 J.R. Pierce
"Network for Block Switching of data", Bell System Technical Journal, July/Aug. 1972

J PIERCE 72 J.R. Pierce
"How Far Can Data Loops Go", IEE Transaction Communication, vol COM-20, June 1972

POTVIN 71 J.N. Potvin, et al
"A Star-Ring: A Computer Intercommunication and I/O System", Information Processing, 1971

RAGHAVENDRA 84 C.S. Raghavendra
"Fault Tolerance in Regular Network Architecture", IEEE Micro, December 1984

ROGERS 84 A.S. Rogers
"The Performance of a Fault-Tolerant Ring Communication System", Performance of Computer Communication Systems, IFIP, 1984

SOI 85 I.M. Soi and K.K. Aggarawal "Overall Reliability Evaluation for Large Computer Communication Networks: An MHC Approach", Microelectronic Reliability, vol 25, no 2, 1985

SPRATT 80 E.B. Spratt "Operational experiences with a Cambridge Ring local-area-network in a University environment", Proceedings of IBM/IFIP International Workshop on Local Area Networks, Zurich, Aug. 1980

SALTZER 80 J.H. Saltzer and K.T. Pogran "A Star-Shaped Ring Network with High Maintainability", Computer Networks 4, 1980

SALTZER 81 J.H. Saltzer, D.D. Clark and K.T. Pogran "Why a Ring", Proc. IEEE, 1981

SERLIN 83 O. Serlin "Fault Tolerant Computers", Data Processing, vol 25 no 10, December 1983

STRUBE 85 A.R. Strube "The Challenge of Reliability", IEEE Circuits and Devices Magazine, May 1985

TANENBAUM 81 A.S. Tanenbaum "Computer Networks", Prentice Hall, 1981

WEARDEN 85 T. Wearden "New York Fibre Optics Firm Sets Its Sight on U.K.", Electronic Times, 16 May 1985

WEITZMAN 80 C. Weitzman "Distributed Micro/Minicomputer System", Prentice Hall, 1980

WILKOV 71 R.S. Wilkov "Reliability considerations in Computer Network Design", Information Processing 71, North Holland, 1971

ZAFIROPULO 74 P. Zafiropulo "Performance Evaluation of Reliability Improvement Techniques for Single-Loop Communication System", IEEE Trans. Communications, Vol COM-22 no 6, June 1974

Z80A-CPU, Z80A-SIO, Z80A-CTC Technical Manuals Zilog Inc.

The Interface Circuits Data Book
Texas Instruments

The TTL Data Book
Texas Instruments

Transring 2000 Series Systems User Manual
Scientific and Electronic Enterprises Ltd.,
3 Young Square, Brucefield Industrial
Estate, Livingston, West Lothian, Scotland

INT 1 "Notes on the Station Design version 3", Internal Report, Computer Laboratory, Cambridge University

APPENDIX 1

To derive formulae for cable length calculations

The following formulae were developed for comparing the Mesh, Self-Heal, Star-Shaped ring and the Hierarchical Ring-Star topologies in their cabling requirement.

The layout of any installation will depend on the total number of nodes, their spread and in particular the architecture of the building or site. This obviously varies greatly. Thus an accurate comparision of the four topologies is impossible, and to make any comparisons at all, somewhat unrealistic assumptions have to be made.

The following are the main assumptions made:

- a ring is installed in a circle with radius R and diameter D.
- the topology models the ideal structure of each design.

n = total number of nodes

L = total length of cable required

Mesh Network

Various configurations of the Mesh is possible but for the comparison, one where alternate nodes are linked with an extra cable is modelled. From Fig. A1, it can be observed that such a topology is really three rings connected in parallel.

Thus, $L = 3 \times 2\pi R$

$= 6\pi R$

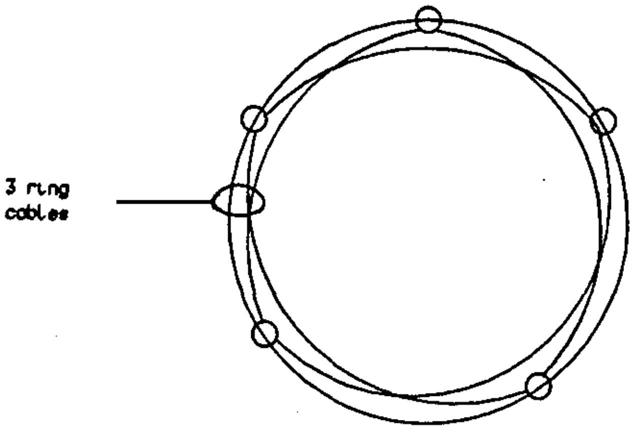


Fig. A1 The Mesh network is really 3 rings in parallel

Self-Heal Ring

Since it is a double loop,

$L = 2 \times 2\pi R$

$= 4\pi R$

Star-Shaped Ring

This ring is shaped into a star structure with a pair of cables connecting up each node. Each cable pair is equivalent to length 2R.

$$L = n \times 2R$$
$$= 2nR$$

Hierarchical Ring-Star

It is likely that in the future, the most common network would wire up a building. Thus, a 3-level Hierarchical Ring-Star network will be modelled, with a rigid structure for simplicity.

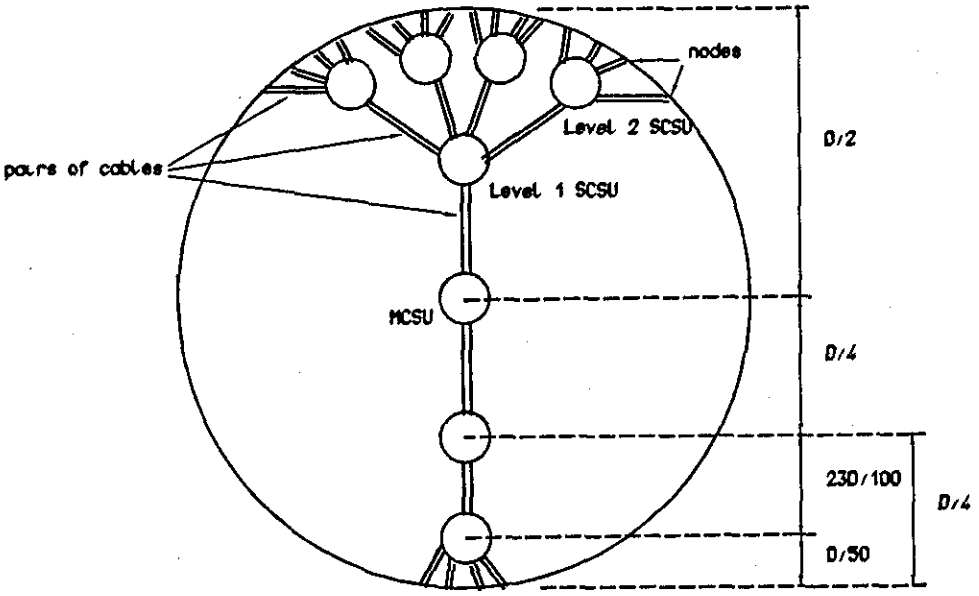


Fig. A2 A 3-Level Hierarchical Ring-Star Structure

It is further assumed that the distance between CSUs (MCSU and SCSUs) are spaced equally, thus D/4. Relative to this distance, the length of cable connecting a node to the SCSU is small. Assuming a building with an average length of 50m, a node-to-SCSU cable of length 1m will give D/50. Also, each SCSU is assumed to have 4 relay ports and there are a multiple of 4 nodes in the network.

Let P be the total number of level 2 CSU

Let Q be the total number of level 1 CSU

$$P = n/4 \text{ and } Q = P/4 = n/16$$

L = circumference of ring + MCSU-to-level 1 CSU links + level 1 CSU-to-level 2 CSU links + level 2 CSU-to-node links

$$= 2\pi R + (2XD/4) + (2x23DY/100) + (2Dn/50)$$

$$D = 2R$$

$$L = 2\pi R + (2x2Rn/16) + (2x23x2Rxn/1600) + (4Rn/50)$$

$$= 2\pi R + Rn/4 + 23Rn/400 + 2Rn/25$$

$$= 2\pi R + 155Rn/400$$

$$L = R(2\pi + 31n/80)$$

APPENDIX 2

A Guide for Operating the Hierarchical Ring-Star System

When the MCSU is first switched on, only the prompt ">" will be displayed. This indicates where commands are typed in. In fact there is only one command, Control-A (hold down the CTRL key and press the "A" key) to get into the Master Menu. From here onwards, the user will be prompted to enter data or commands selected from menus. If a character is typed wrongly, it can be deleted by using the "BACKSPACE" or the "RIGHT CURSOR" key. The operation will be explained by going through the contents of the Master Menu and sub-menus step by step.

By entering Control-A, the Master Menu shown below will be displayed.

```
MENU:

  ENTER CONFIGURATION MODE  1

  SET UP NODE DICTIONARY    2

  DISPLAY NODE DICTIONARY   3

  EXIT                      6
```

```
What do you want to do?

Choice = _
```

If the choice = 1, the next menu is displayed. Note : in all cases, a Carriage Return must be typed to enter the command.

```
CONFIGURATION MODE


Enter command for task required


  Task          Bypass  Reset

Node              Ø       A

Link              1       B

Relay Port        2       C

SCSU              3       D

Display EL        X

Enable EL         Y

Exit              E


Choice = _
```

From here, to assist with the explanation the format below will be adopted. The contents of the screen display (roughly) are shown on the left with comments on the right.

Display	Comments
Choice = Ø	Node Bypass command selected.
Node Address = _	Enter node address (to be isolated from ring) in 2 characters e.g. 1A, Ø5, F2. Hexadecimal notation assumed.
Response (Done, Failed or Address does not exist	The system will response in one of 3 ways. If the command is carried out successfully, 'Done' is displayed, otherwise, 'Failed'. If the address cannot be found in the Node Dictionary the third response is made.
Choice = 1	Link Bypass command selected.
Address of Node to to which Link is connected into =	As above
Response (Done,	As above
Failed or Address	
does not exist	As above
Choice = 2	Relay Port Bypass command selected.
SCSU Address = _	Address of the SCSU on which Relay Port is contained.
Port no. = _	The specific Relay Port on the SCSU
Response (Done, Failed or Address	

does not exist	As above
Choice = 3	SCSU Bypass command selected.
SCSU Address = _	Address of the SCSU.
Response (Done,	
Failed or Address	
does not exist	As above
Choice = X	Disable Error Logger command selected.
Response (Done	
or Failed)	
Choice = Y	Enable Error Logger command selected.
Response (Done	
or Failed)	
Choice = E	Exit command selected.
>	Prompt indicating exit from
	Configuration Mode.
Choice = A	Node Reset command selected.
Node Address = _	As in choice = Ø
Response (Done,	

Failed or Address

does not exist

Choice = B	Link Reset command selected.
Address of Node to	
to which Link is	
connected into =	As in choice = 1
Response (Done,	
Failed or Address	
does not exist	

Choice = C	Relay Port Reset command selected.
SCSU Address = _	As in choice = 2
Port no. = _	
Response (Done,	
Failed or Address	
does not exist	

Choice = D	SCSU Reset command selected.
SCSU Address = _	As in choice = 3
Response (Done,	
Failed or Address	
does not exist	

The "Disable EL" command can be used by an operator to reconfigure the network without causing 'false' error detection. Recall that relay actions will cause artificial ring breaks. Thus this command should be used before carrying out reconfiguration and reset by the "Enable EL" command after.

If choice = 2 is selected from the Master Menu, the "SET UP NODE DICTIONARY" MODE is entered. This is normally the first task required when the system is first installed. Addressing information is entered into the Node Dictionary. In essense, the system prompts the operator (or network administrator) to enter addresses of devices connected into the Relay Ports. In the case of the SCSU, the actual nodes' ring addresses are entered. For the MCSU, the addresses of SCSUs are entered.

<u>NOTE:</u> If any RP does not have a device attached, an 'N' <u>must</u> be entered.

<u>Display</u>	<u>Comments</u>
How many levels of CSU are there in the Network Configuration?	
No. of levels = _	In this case, enter Ø2 for a 2-level architecture.
Enter Address of Slave CSU attached to Relay Port 1	The first of 8 Relay Ports on the MCSU. Enter a 2

Slave CSU Address = _	character response as before.
Enter Network Address of Node attached to Relay Port 1 Network address = _	This is similar to the above but for Relay Port 1 on SCSU1.
Enter Network Address of Node attached to Relay Port 2 Network address = _	As above but for RP 2
Enter Network Address of Node attached to Relay Port 3 Network address = _	As above but for RP 3
Enter Network Address of Node attached to Relay Port 4 Network address = _	As above but for RP 4
Enter Address of Slave CSU attached to Relay Port 2 Slave CSU Address = _	The second of 8 Relay Ports on the MCSU.
Enter Network Address of Node attached to Relay Port 1 Network address = _	Relay Port 1 on the SCSU above.
Enter Network Address of Node attached to Relay Port 2 Network address = _	As above but for RP 2

Enter Network Address of Node attached to Relay Port 3 Network address = _	As above but for RP 3
Enter Network Address of Node attached to Relay Port 4 Network address = _	As above but for RP 4
Enter Address of Slave CSU attached to Relay Port 3 Slave CSU Address = _	The third of 8 Relay Ports on the MCSU.

.

.

etc

.

.

Enter Address of Slave CSU attached to Relay Port 8 Slave CSU Address = _	The last of 8 Relay Ports on the MCSU.

If choice = 3 is selected from the Master Menu, the "DISPLAY NODE DICTIONARY" MODE is entered. This simply prints out the contents (address and status) of the Node

Dictionary. A sample screen is shown.

Format is :

nth Level 1 Slave CSU Address = : Status

CSU Port = Node Address : Status

```
1 = Ø2H : ENTRY,
  A = ØØH : EMPTY
  B = 22H : ENTRY, Node ok, Link ok
  C = B3H : ENTRY, Node ok, Link broken, Link bypassed
  D = 24H : ENTRY, Node ok, Link ok
2 = ØØH : EMPTY,
  A = ØØH : EMPTY,
  B = ØØH : EMPTY,
  C = ØØH : EMPTY,
  D = ØØH : EMPTY,
3 = Ø9H : ENTRY,
  A = 14H : ENTRY, Node ok, Link ok, Link bypassed
  B = F3H : ENTRY, Node faulty, Node bypassed, Link ok
  C = 43H : ENTRY, Node ok, Link ok, Link bypassed
  D = Ø4H : ENTRY, Node ok, Link ok
4 = Ø5H : ENTRY, SCSU bypassed
  A = ØØH : EMPTY
  B = 22H : ENTRY, Node ok, Link ok
  C = B3H : ENTRY, Node ok, Link ok
  D = 24H : ENTRY, Node ok, Link ok
     .
     .
     .
```

8 = ØØH : EMPTY,

A = ØØH : EMPTY,

B = ØØH : EMPTY,

C = ØØH : EMPTY,

D = ØØH : EMPTY,

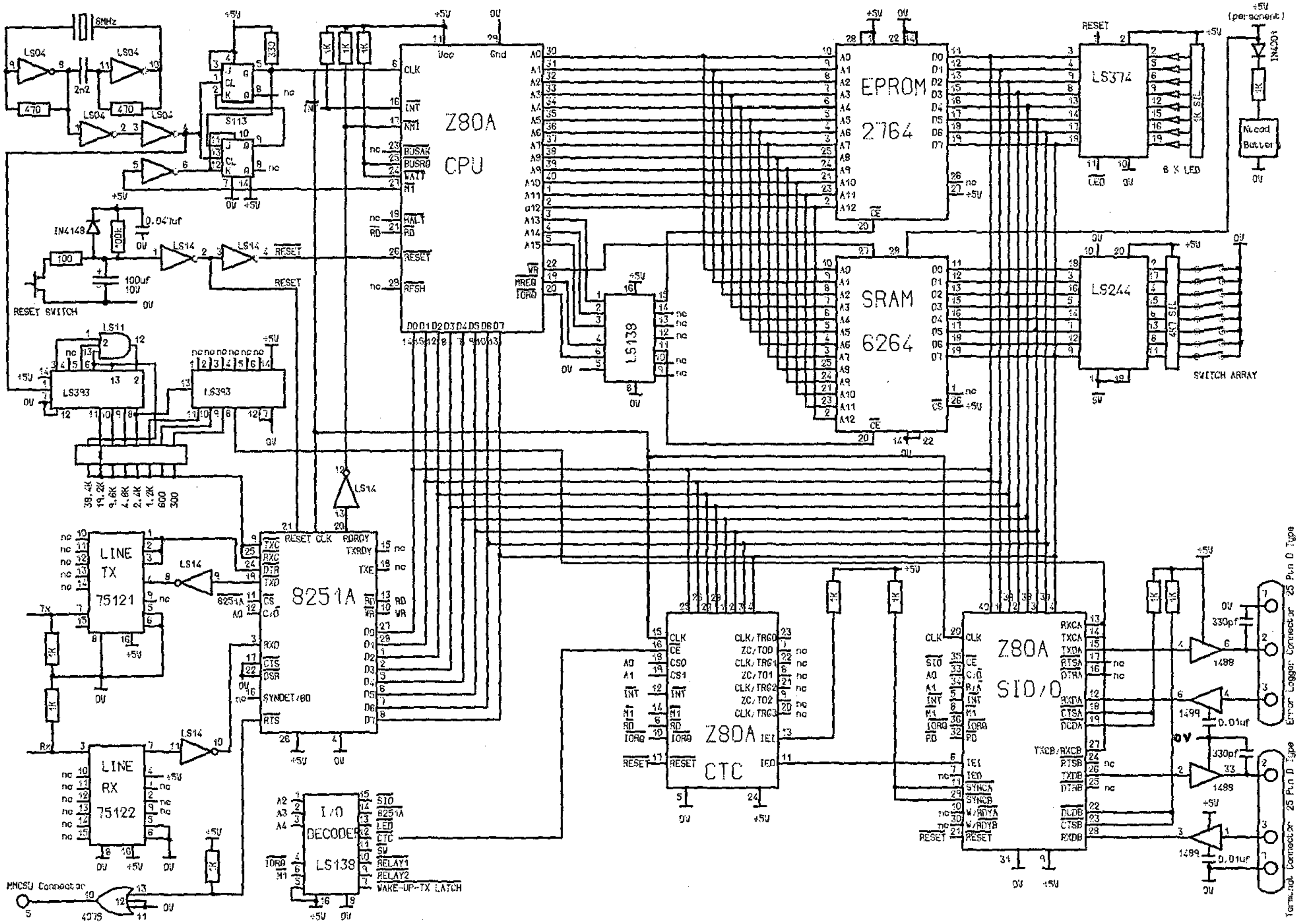
Notes

The status for the SCSU Relay Ports can be combinations of: Node ok, Link ok, Node faulty, Link broken, Node/Relay Port bypassed, or Link bypassed.

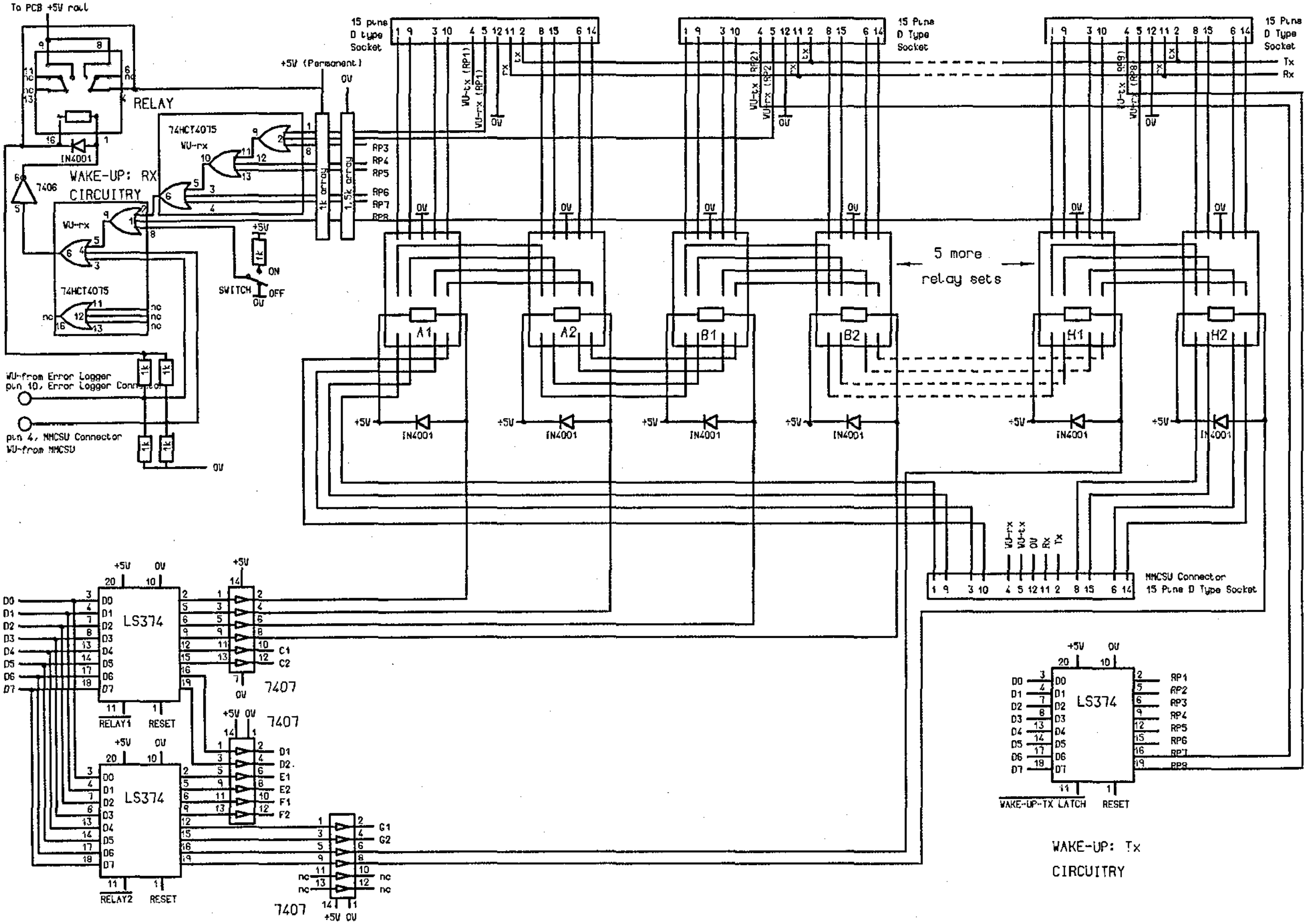
When a Node or Relay Port is bypassed, the links on the Relay Ports on either side of the Relay port are bypassed too.

APPENDIX 3

Circuit diagram of the Master CSU

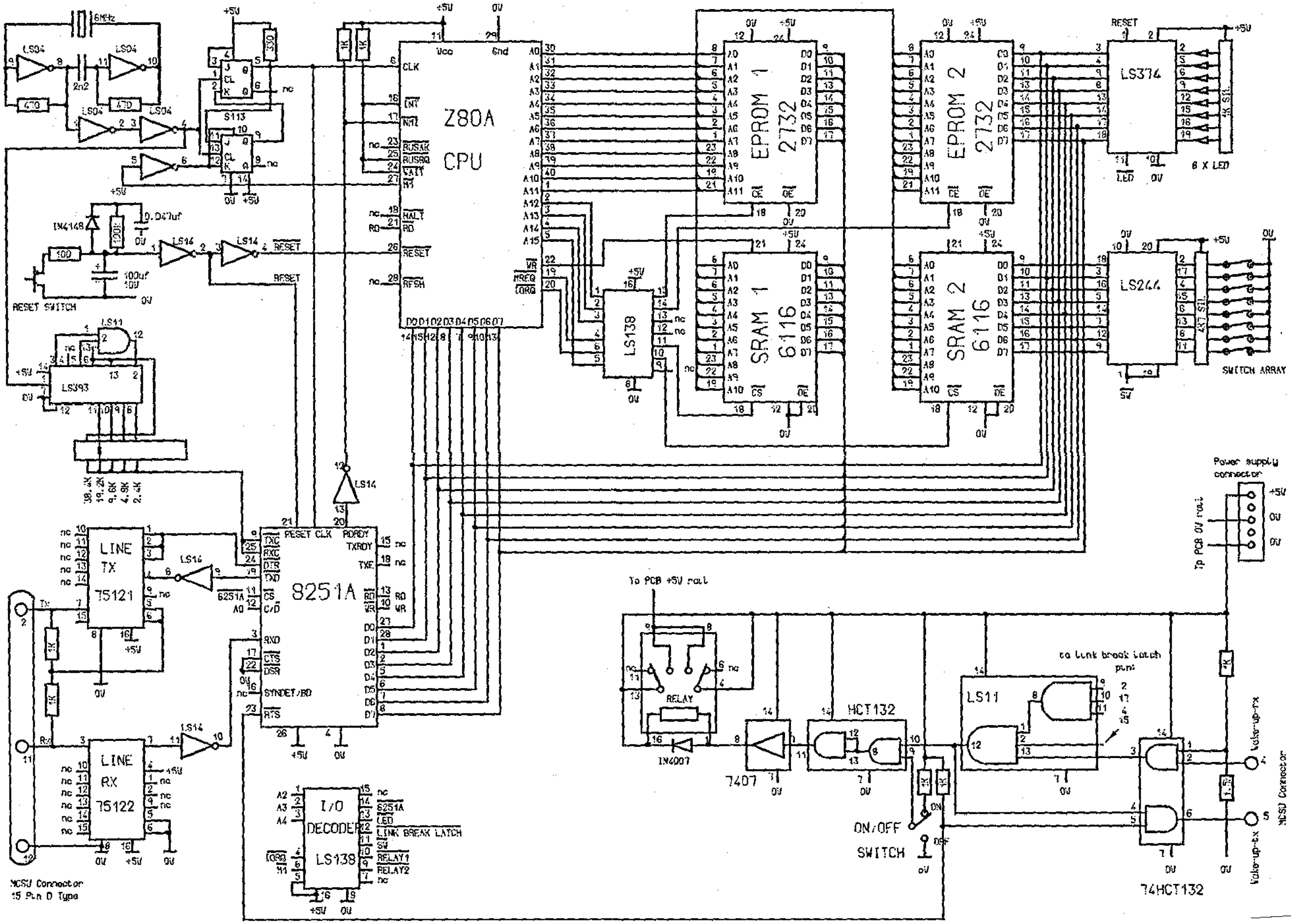


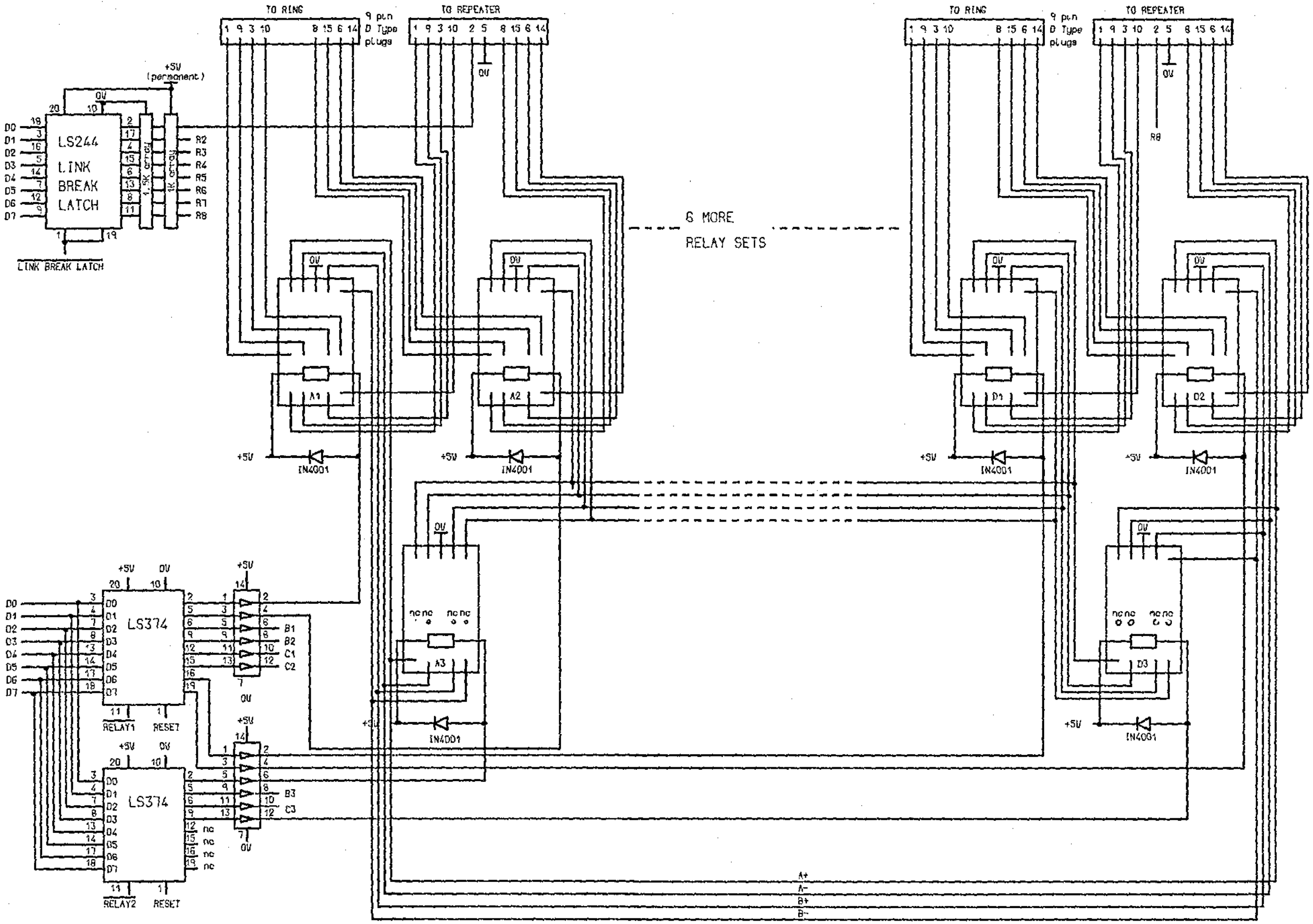
8MHz
LS04
2n2
470
S113
330
1K
Z80A
CPU
Vcc
Gnd
CLK
INT
NMI
BUSAK
BUSRQ
WAIT
M1
HALT
RD
RESET
RFSH
WR
MREQ
IORQ
IN4148
100K
0.047uf
100
100uf
10V
RESET SWITCH
LS14
LS11
LS393
39.4K
19.2K
9.6K
4.8K
2.4K
1.2K
600
300
EPROM
2764
SRAM
6264
LS138
LS374
LS244
8 X LED
SWITCH ARRAY
1K SIL
4K7 SIL
+5V (permanent)
IN4001
Nicad Battery
LINE TX 75121
LINE RX 75122
8251A
TXD
RXC
DTR
CS
C/D
RXD
CTS
DSR
SYNDET/BD
RTS
RDRDY
TXRDY
TXE
Tx
Rx
I/O DECODER LS138
SIO
8251A
LED
CTC
SW
RELAY1
RELAY2
WAKE-UP-TX LATCH
MMCSU Connector
4075
Z80A CTC
CLK/TRG0
ZC/TO0
CLK/TRG1
ZC/TO1
CLK/TRG2
ZC/TO2
CLK/TRG3
IEI
IEO
Z80A SIO/0
RXCA
TXCA
TXDA
RTSA
DTRA
RXDA
CTSA
DCDA
TXCB/RXCB
RTSB
TXDB
DTRB
DCDB
CTSB
RXDB
SYNCA
SYNCB
W/RDYA
W/RDYB
1488
1489
330pf
0.01uf
Error Logger Connector 25 Pin D Type
Terminal Connector 25 Pin D Type
+5V
0V
nc


To PCB +5V rail
RELAY
IN4001
7406
WAKE-UP: RX
CIRCUITRY
74HCT4075
WU-rx
74HCT4075
SWITCH
ON
OFF
0V
+5V
1k
+5V (Permanent)
0V
1k array
1.5k array
RP3
RP4
RP5
RP6
RP7
RP8
15 pins
D type
Socket
WU-tx (RP1)
WU-rx (RP1)
rx
tx
WU-tx (RP2)
WU-rx (RP2)
WU-tx (RP8)
WU-rx (RP8)
Tx
Rx
A1
A2
B1
B2
H1
H2
5 more
relay sets
IN4001
WU-from Error Logger
pin 10, Error Logger Connector
pin 4, MMCSU Connector
WU-from MMCSU
MMCSU Connector
15 Pins D Type Socket
WU-rx
WU-tx
0V
Rx
Tx
LS374
D0
D1
D2
D3
D4
D5
D6
D7
RELAY1
RESET
RELAY2
7407
C1
C2
D1
D2.
E1
E2
F1
F2
G1
G2
nc
WAKE-UP-TX LATCH
RP1
RP2
RP3
RP4
RP5
RP6
RP7
RP8
WAKE-UP: Tx
CIRCUITRY

APPENDIX 4

Circuit diagram of the Slave CSU


6MHz
LS04
2n2
470
LS04
+5V
330
S113
nc
Z80A
CPU
CLK
INT
NMI
BUSAK
BUSRQ
WAIT
HALT
RD
RESET
RFSH
WR
MREQ
IORQ
Vcc
Gnd
+5V
0V
1K
IN4148
100k
0.047uf
100
100uf 10V
RESET SWITCH
LS14
RESET
LS11
LS393
38.4K
19.2K
9.6K
4.8K
2.4K
EPROM 1
2732
EPROM 2
2732
SRAM 1
6116
SRAM 2
6116
CE
OE
CS
LS138
LS374
8 X LED
4K SIL
LED
LS244
4K7 SIL
SWITCH ARRAY
SW
LINE
TX
75121
LINE
RX
75122
Tx
Rx
8251A
TXC
RXC
DTR
TXD
CS
C/D
RXD
CTS
DSR
SYNDET/BD
RTS
RDRDY
TXRDY
TXE
RD
WR
A0
I/O
DECODER
LS138
A2
A3
A4
IORQ
M1
nc
8251A
LED
LINK BREAK LATCH
SW
RELAY1
RELAY2
MCSU Connector
15 Pin D Type
To PCB +5V rail
RELAY
IN4007
7407
HCT132
LS11
to link break latch pins
ON/OFF
SWITCH
ON
OFF
74HCT132
Power supply connector
To PCB 0V rail
1.5K
wake-up-rx
wake-up-tx
MCSU Connector


TO RING
9 pin
D Type
plugs
TO REPEATER
1 9 3 10 8 15 6 14
1 9 3 10 2 5 8 15 6 14
0V
+5V
(permanent)
LS244
LINK
BREAK
LATCH
LINK BREAK LATCH
1.5K array
1K array
D0
D1
D2
D3
D4
D5
D6
D7
R2
R3
R4
R5
R6
R7
R8
6 MORE
RELAY SETS
A1
A2
D1
D2
A3
D3
nc nc
IN4001
+5V
LS374
RELAY1
RESET
RELAY2
RESET
B1
B2
C1
C2
B3
C3
nc
A+
A-
B+
B-

APPENDIX 5

Hierarchical Ring-Star Interconnection Scheme

and

Cabling Details

Interconnection Scheme and Cabling

The interconnection scheme between the ring, repeater, SCSU and MCSU is shown below.

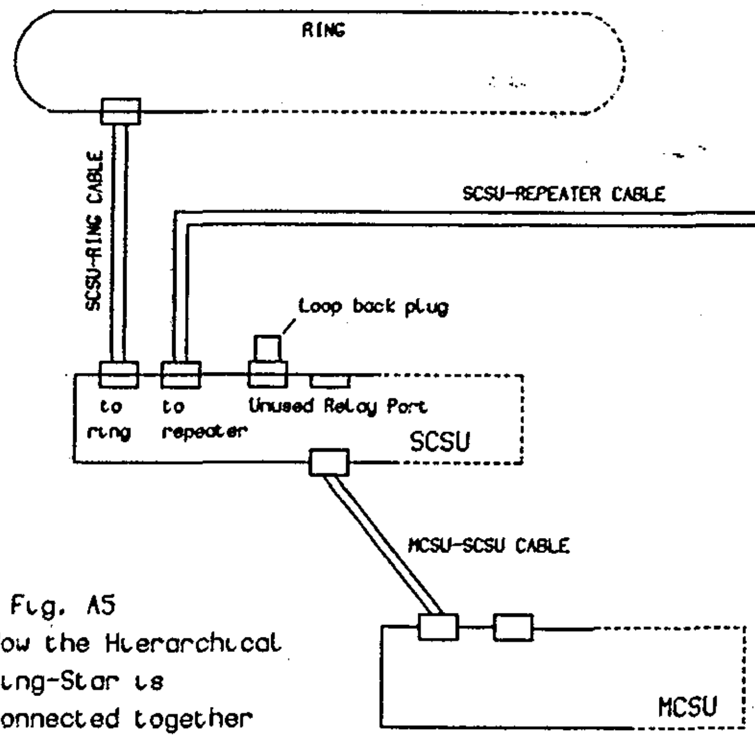
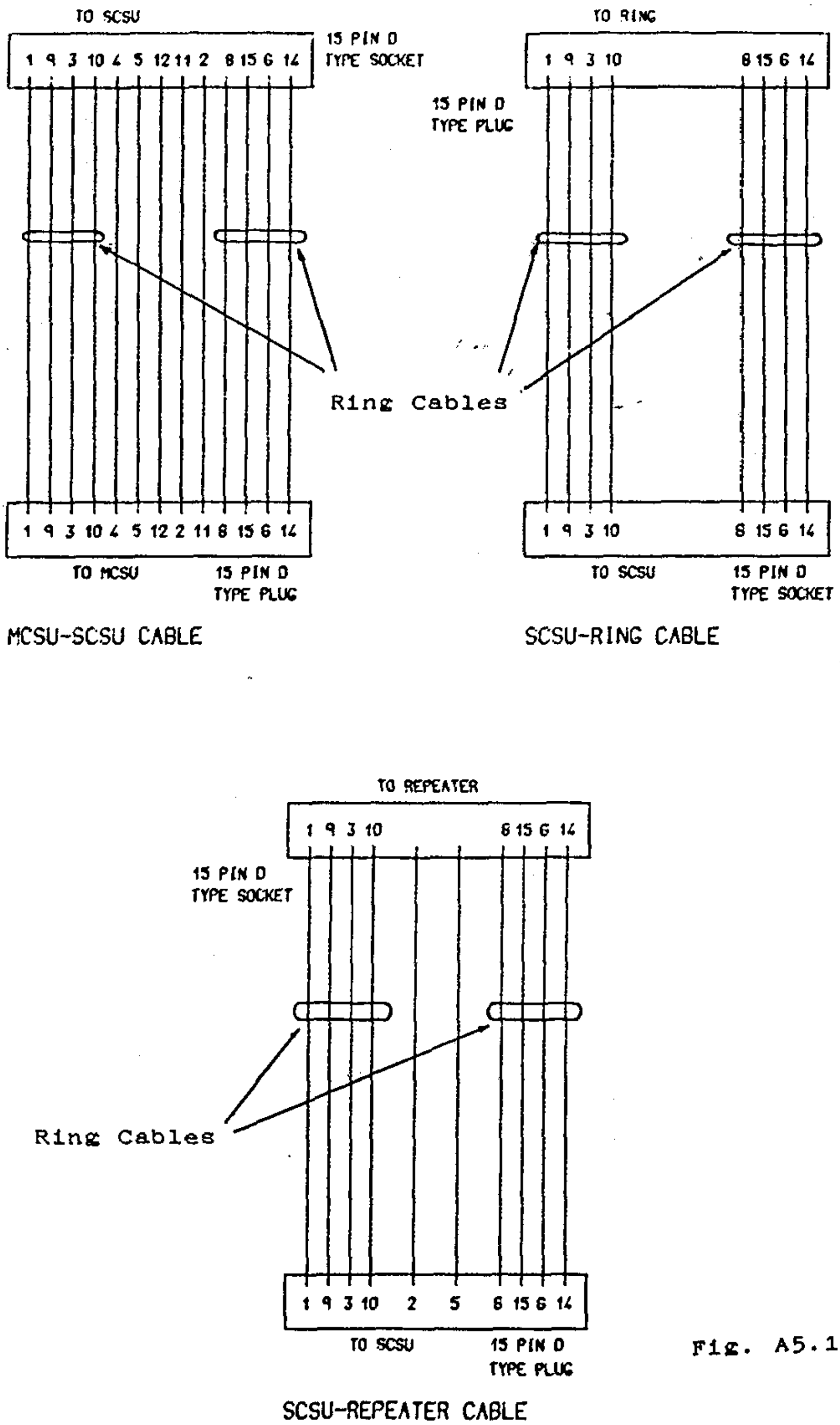


Fig. A5
How the Hierarchical Ring-Star is connected together

The various cables must be made up as shown in Fig. A5.1. The cables marked 'Ring Cables' must conform to the CR82 standard. The other cables can be any twisted pair telephone cables.



MCSU-SCSU CABLE

SCSU-RING CABLE

SCSU-REPEATER CABLE

Fig. A5.1