CENTER FOR RELIABLE COMPUTING

CURRENT RESEARCH

by

Edward J. McCluskey
John F. Wakerly
Roy C. Ogus

October 1975

Technical Report No. 100

DIGITAL SYSTEMS LABORATORY

STANFORD ELECTRONICS LABORATORIES

STANFORD UNIVERSITY · STANFORD, CALIFORNIA
CENTER FOR RELIABLE COMPUTING

CURRENT RESEARCH

by

Edward J. McCluskey
John F. Wakerly
Roy C. Ogus

October 1975

Technical Report no. 100

DIGITAL SYSTEMS LABORATORY
Department of Electrical Engineering  Department of Computer Science
Stanford University
Stanford, California
ABSTRACT

This report summarizes the research work which has been performed, and is currently active in the Center for Reliable Computing in the Digital Systems Laboratory, Stanford University.
# TABLE OF CONTENTS

1. **INTRODUCTION.** ............................................. 1

2. **THEORY OF FAULTS IN LOGIC SYSTEMS** .................... 3
   2.1 Reliability Modeling. ......................................... 3
      2.1.1 Static Redundancy ....................................... 5
      2.1.2 Dynamic Redundancy ..................................... 10
      2.1.3 Hybrid Redundancy ...................................... 14
      2.1.4 Comparisons ........................................... 18
   2.2 Signal Reliability. ........................................... 19
   2.3 Self-diagnosing Computer Design ............................. 21
      2.3.1 Error-detecting Codes .................................. 23
      2.3.2 Self-checking Circuits. ............................... 23
      2.3.3 Fail-safe Circuits. ................................... 24
   2.4 Multiple Fault Studies. ..................................... 24
      2.4.1 Fault Masking ......................................... 25
      2.4.2 Iterative Networks. .................................. 25
   2.5 References. .............................................. 26

3. **STUDY OF MAINTAINABLE COMPUTERS** .......................... 37
   3.1 Applying Triple Modular Redundancy to Small Computers 37
      3.1.1 Effects of Transient Failures in Sequential Modules. 37
      3.1.2 Reliability of TMR Memory Systems .................. 39
      3.1.3 A TMR Microcomputer System ......................... 41
   3.2 Self-Checking-Circuits. .................................... 53
      3.2.1 Checked Binary Addition Using Checksum Codes and Check Symbol Prediction .... 53
      3.2.2 A Self-Checking Checker for Periodic Signals. .... 54
      3.2.3 Fail-Safe Sequential Machines ...................... 57
   3.3 Self-Diagnosing Computer Design ............................ 59
3.3.1 Data Path Checking .......................... 60
3.3.2 Control Unit Checking. ...................... 65
3.3.3 Diagnosis. ................................ 68
3.3.4 Peripheral and Input/Output Checking ....... 69

3.4 References Prepared under Previous NSF Grants. .... 71
3.5 Additional References. ........................... 73

4. INVESTIGATION AND EVALUATION OF DUAL COMPUTER CONFIGURATIONS

4.1 Introduction .................................... 75

4.2 Summary of Technical Areas in the Project. ........ 76

4.2.1 The SIRU Computer System .................... 76
4.2.2 Error Recovery Techniques in Computer Systems. 78
4.2.3 System Reliability Modeling. ................. 79
4.2.4 Signal Reliability of a Circuit. ............... 80
4.2.5 Software Reliability .......................... 81

4.3 References ..................................... 82

5. BIOGRAPHICAL SKETCHES OF PRINCIPAL RESEARCHERS IN CENTER FOR RELIABLE COMPUTING .......................... 83

6. LISTING OF PERSONNEL IN CENTER FOR RELIABLE COMPUTING. .... 85

7. PREVIOUS RESEARCHERS IN CENTER FOR RELIABLE COMPUTING. ... 87

8. CENTER FOR RELIABLE COMPUTING BIBLIOGRAPHY ........... 89

8.1 Journal Papers ................................. 89
8.2 Conference Papers. ............................. 91
8.3 Technical Papers ............................... 94
0.4 Technical Notes. ............................... 97
1. INTRODUCTION

This report summarizes the research work which has been performed and is currently active in the Center for Reliable Computing (CRC) in the Digital Systems Laboratory, Stanford University. The report covers three projects in the CRC, and presents recent results obtained by researchers in the group, as well as an indication of the current status of each project.

The first project concerns the theory of faults in logic systems. Theoretical studies of the effects of faults on logic networks have been conducted and techniques for the modelling and evaluation of the reliability of redundant systems have been developed. Other areas involve the design of ultra-reliable systems, self-diagnosing computer systems, and the testing of digital circuits. Studies of fail-safe and self-checking circuits have also been carried out. The summary of the research in this project is presented in Section 2.

The second project involves the study of maintainable computers. Techniques for implementing fault-detection and diagnosis of small scale processors at a low cost are being investigated, as well as the study of improving the reliability of I/O controllers and peripheral devices. The design of redundant microprocessor systems and implementation of a low-cost maintainable minicomputer are current studies being performed. The description of the research in this project is found in Section 3.

The third project deals with the study of dual computer
configurations. A dual processor system used in a guidance and navigation application was implemented for NASA-Ames Research Center and the CRC has been evaluating the performance of the system.

Section 5 presents a short biography of the principal researchers on the three projects, and Section 6 describes some previous researchers in the CRC.
2. THEORY OF FAULTS IN LOGIC SYSTEMS

Summary of Previous Work and Work in Progress

2.1 Reliability Modeling

A large variety of redundancy techniques -- TMR, Standby Sparing, Coding, Multiprocessing, etc. -- exists for improving the reliability of a digital system. When a system is being designed for a specific mission with given reliability requirements, it is necessary to:

(1) decide whether use of a redundancy technique is required,

(2) choose the optimum redundancy technique (if any is required) for the given application, and

(3) evaluate the resulting system reliability to determine whether the design requirements have been met.

The objective of research into reliability modeling is to develop tools to permit the previous three steps to be carried out as accurately as possible. Inaccuracy can result in systems which either (1) are more expensive than necessary, or (2) (unknowingly) fail to meet the reliability requirements.

Because of its serious practical importance, reliability modeling has received a great deal of attention [Borgerson, 1975; Bouricius, 1971; Rennels, 1973; Mathur, 1971].*

However, this earlier work results in only partially satisfactory results due to the inherent complexity of the problem. Straightforward combinatorial techniques usually have a computational

*The references for each section are compiled at the end of the particular section.
complexity too great to be useful so that more sophisticated approaches are necessary.

Redundancy techniques are usually classified as being either static, dynamic or hybrid. In static redundancy, also called masking or massive redundancy, the effect of a fault is masked instantaneously by the non-faulty components of the system. No alteration in the system interconnection pattern is required for this error correction to take place. The principal static redundancy technique is triple modular redundancy (TMR) which was used in the Saturn V launch vehicle computer [Dickenson, 1964] and has been proposed for use in other flight control systems [Masreliez, 1975]. Another important form of static redundancy is interwoven redundant logic [Pierce, 1955]. Finally, the possibility of achieving static redundancy by replication of circuit components [Creveling, 1956; Lewis, 1963] should be mentioned. Interwoven logic is limited in its application since it must be applied at the gate level which can be an important restriction when the trend is towards higher levels of integration with consequent implications for non-independence of gate failures. TMR is more generally applicable since it can be applied at the gate, subsystem, or entire system level.

Dynamic redundancy, also called selective, stand-by or sparing redundancy, is characterized by the detection of a fault and the implementation of a corrective action. Fault detection is either concurrent or periodic. In concurrent detection, use is made of a
coding technique relying on redundant signals [Kautz, 1962]. An important special case is the duplication of the system to generate two copies of the outputs with fault detection caused by a mismatch between duplicate signals [Wachter, 1975]. Periodic detection involves stopping the normal operation of the system at intervals to allow diagnostic tests to be carried out. Correction can be implemented by error-correcting circuitry if an error-correcting code is being used. Otherwise it is necessary to resort to some diagnostic routine to determine the proper signal values and fault-free system components and often to initiate a reconfiguration which eliminates the faulty subsystem.

In hybrid redundancy [Mathur, 1970; Siewiorek, 1973; Ogus, 1974], use is made of a static redundancy technique -- probably TMR -- to provide both error masking and error detection. Hybrid redundancy differs from static redundancy in that a reconfiguration action is taken when an error is detected. The major types of hybrid redundancy are TMR + spares [Ogus, 1974] and self-purging redundancy [Pierce, 1962; Chandy, 1972; Losq, 1975A].

2.1.1 Static Redundancy. A TMR system which consists solely of three copies of the desired circuit such as that shown in Figure 2.1, followed by a rank of voters from which the system output is derived is called a one-level TMR system. A system in which an input signal must pass through more than one vote before reaching the output, as in Figure 2.2, is called a multi-level TMR system.
Figure 2.1  One-level TMR System
One-level TMR Modeling. It is customary to model the reliability of a one-level TMR system by assuming that the system will perform correctly as long as no more than one of the three copies of the circuit has failed. In other words, it is assumed that for any failures affecting two or more copies, the system will fail. This assumption is much too pessimistic since there are many two-circuit failure situations which do not result in system failure. For example, consider the situation where one of the circuits has its output stuck-at-1, shown in Figure 2.3a, a second circuit has its output stuck-at-0, and the third circuit is fault-free. The system output from the voter will be correct. Another drastically over-simplified situation is shown in Figure 2.3b. Here the circuit being protected by TMR is a two-input AND gate. The figure shows the situation which occurs if an input (x) to one of the copies of the

Figure 2.2 - Multi-level TMR Systems
Figure 2.3- One-level TMR Circuits with Multiple Faults
gate is stuck-at-1 and the other input (y) to a second copy is also stuck-at-1. The system output is \( z = xy \) as required for correct operation.

A technique for calculating one-level TMR reliability which accurately models the effects of multiple failures which do not cause system failure was developed under a previous NSF grant* [Siewiorek, 1975]. It was shown that neglect of the multiple fault phenomenon just described can result in mission time predictions which are as much as 30% lower than the more accurate value obtained using Siewiorek's method.

**Multiple-level TMR Modeling.** A great deal of attention has been devoted to the problem of calculating the reliability of a multiple-level TMR network [Brown, 1961; Teoste, 1962; Rhodes, 1964; Longden, 1966; Rubin, 1967; Lyons, 1962; Gurzi, 1965; Jensen, 1964]. None of these approaches produced a technique for determining the exact reliability of a multiple-level TMR network; only bounds were obtained. The difficulty with obtaining an exact solution stems from the necessity for taking into account the complex interactions among the various stages of the network. Thus any straightforward combinatorial approach leads to a technique of computational complexity. In [Abraham, 1974], an algorithm for calculating the exact reliability of a multiple-level TMR network is described. The algorithm was developed by Abraham and Siewiorek

---

* NSF grant GJ 27527
under a previous NSF grant*. The success of the Abraham-Siewiorek algorithm depends on their technique of associating voters with circuit inputs and then partitioning the network into subnetworks which have the property that the overall network reliability can be calculated as the product of the subnetwork reliabilities.

**Interwoven redundant logic modelling.** In an interwoven redundant logic network the tasks of error correction and calculation of the output function are not separated as they are in TMR networks so that one-stage interwoven systems are of no importance. The problem of modelling interwoven networks was studied in [Jensen, 1963] and [Teoste, 1961] but only approximate techniques were discovered. The exact analysis was considered too difficult and costly [Teoste, 1964; Goldberg, 1966]. Our earlier success in modelling multiple level TMR networks led to the hope that a similar approach would be useful for interwoven networks. This turned out to be true, and a technique for exact modelling of interwoven redundant logic was developed under our present NSF grants# and is reported upon in [Abraham, 1975].

2.1.2 **Dynamic Redundancy.** The accurate modelling of dynamic redundant systems is difficult because the system reliability is very sensitive to the reliability of the mechanisms for fault detection and reconfiguration, and it is usually quite difficult to

* NSF grant GJ 27527

# NSF grant GJ 40286 and DCR 7307973
model the reliability of these mechanisms accurately. In [Bouricius, 1969] a technique for estimating dynamic system redundancy is presented in which the problem of deriving detection and recovery mechanism reliabilities is avoided by introducing the concept of a coverage factor. The coverage factor is defined as the probability of system recovery given that a failure has occurred. The concept of coverage is very important for dynamic redundant systems. It has been used to demonstrate the extreme sensitivity of these systems to the recovery mechanisms [Arnold, 1973] and thus to focus attention on these mechanisms as being the critical aspect of such systems. The major problem with making use of coverage in modelling is the practical difficulty of determining the coverage factor for a specific system and evaluating the effect of the number of spares on the value of the coverage factor. As the number of spare modules is increased, so also is increased the number of module failures which the system can withstand and still continue to function. However, an increase in the number of spares also causes an increase in the complexity of the detection and reconfiguration mechanisms and thus a decrease in their reliability.

Of major importance in modelling dynamic redundant systems is the ability to determine the optimum number of spares. This requires that the interrelationship between more spares giving better protection against module failures but worse detection and reconfiguration reliability be explicitly accounted for. Bouricius et al. attempt to take this interrelationship into account by letting the module
failure rate (\( \lambda \)) depend on the number of spares. While this allows them to show that there exists an optimum number of spares for any coverage factor less than one, it is a completely artificial strategem which is hard to relate to the details of a specific system.

In dynamic redundant systems there are two techniques possible for replacing a failed unit by one of the standby spares: logic switching or power switching. In logic switching power is applied to all the units and reconfiguration consists of substituting the outputs from a standby unit for those of a failed on-line unit. In power switching spares are unpowered until the time at which they are switched on-line. Power switching has two advantages: (1) the savings in power consumption by providing power only to on-line units and (2) the possibility of a reduced failure rate for the non-powered spares [Nerber, 1965]. The ratio of the powered device failure rate to the unpowered device failure rate is called the dormancy factor (usually assumed \( \geq 1 \)).

Under the present grant, a very detailed analysis of stand-by systems [Losq, 1975 B; Losq, 1975 C] has been carried out. Rather than relying only on the coverage factor parameter, the analysis is carried out in terms of the parameters:

\( V_0 \), the rate of fail-safe failures, those failures which result in discarding a fault-free module but successfully replacing it with a spare module;

\( V_1 \), the rate of unsafe failures, those failures which result from a failed module not being successfully replaced (either because the failure is not detected or because a detected failure does not result in a successful reconfiguration).
This study has produced the following results:

(i) A technique for determining the optimum number of spares for a given system design and reliability specification.

(ii) A proof that for extremely short mission times systems with one spare are optimum.

(iii) A proof that for mission times which are not extremely short, but which are less than one-tenth of the mean lifetime of a single unit, the optimum number of spares is still small -- five or fewer for most systems.

(iv) The demonstration that it is possible to calculate a parameter, $\tau$, which specifies the useful life of a stand-by system. The system reliability is very high for mission times less than $\tau$ and drops sharply towards zero for mission times greater than $\tau$.

(v) The effects of imperfect fault detection mechanisms -- mechanisms which are designed to catch only a fraction of all possible module errors -- have been studied. A simple technique for determining the optimum number of spares for such systems and calculating their reliability is given.

(vi) If fail-safe techniques [Usas, 1975B; Mine, 1967] are used to design the fault-detection and recovery mechanisms, it is possible to design a stand-by system whose reliability is always greater than a system having no spares and whose reliability increases monotonically with the number of spares in the system.
2.1.3 **Hybrid Redundancy.** Two general types of hybrid redundancy have been proposed: standby hybrid redundancy and self-purging redundancy, see Fig.2,4. In **standby hybrid redundancy** [Mathur, 1970; Siewiorek, 1973A; Ogus, 1974A] the system is initially placed into operation with three modules (for TMR, N for NMR) active and connected to the Voter from which the system output is derived. The remaining modules are designated as spares and are actively connected to the voter only when one of the on-line modules has failed and been disconnected from the voter. A major advantage of this form of redundancy is the possibility of keeping the spare modules unpowered until they are placed on line. It is thus possible to take advantage of the lower failure rate of unpowered modules [Nerber, 1965]. **Self-purging redundancy** [Pierce, 1962; Chandy, 1972; Losq, 1975A] has all of the modules initially connected to the voter. Only when a failure has been discovered is a module disconnected from the voter. This form of redundancy suffers from the necessity of keeping power on all non-failed modules. It has the advantage of not requiring a complex interconnection network and associated control as for standby hybrid redundancy. The mechanism for disconnecting a failed module from the self-purging system voter is a very simple device which is local to the module.

**Standby Hybrid Redundancy.** Work was carried out under a previous NSF grant* to arrive at an efficient design for implementing

---

*GJ-27527
(a) Standby hybrid redundancy

(b) Self-purging redundancy

Figure 2.4
standby hybrid redundancy. The resulting design, the iterative cell hybrid redundant system, was presented at the FTC/2 conference* [Siewiorek, 1972] and published in [Siewiorek, 1973A]. This design was compared with the only other "published" implementation of standby hybrid redundancy, the status register hybrid redundant system, [Roth, 1967B; Goldberg, 1966] and it was shown that the iterative cell design requires substantially less equipment than the status register design, thus making the iterative cell design both less costly and more reliable. A study was also made of various strategies for choosing which spare to use to replace a failed on-line module [Siewiorek, 1973B]. It was shown that a strategy in which each spare is used for only a subset of all voter inputs has as good reliability as a scheme in which any spare can be used to replace any failed on-line module.

Since the reliability of any hybrid redundant system depends critically on the mechanisms for failure detection and reconfiguration, an investigation was carried out to modify the iterative cell design so as to incorporate redundancy into these critical portions of the design. Two designs were arrived at -- one of which uses TMR and the other which uses fail-safe techniques [Mine, 1967]. These designs were both shown to provide about the same substantial improvement in reliability over the unprotected iterative cell design and to require approximately the same amount of additional hardware. This work was started under a previous *---

NSF grant* and finished under the present grant#. It was presented at FTC/3† [Ogus, 1973] and published in [Ogus, 1974A].

Self-purging redundancy. Although self-purging systems have very good reliability properties, they have previously received very little attention. Work begun under a previous NSF grant* and continued under a present grant# has resulted in a detailed design of the switching and retry mechanisms for self-purging systems as well as the development of techniques for determining both the exact reliability or very tight bounds on the reliability [Losq, 1975A].

In analysis of redundant systems it is standard practice to make the assumption that only single-module failures occur. However there are two situations in which this assumption is invalid:

(i) Any application in which power is not always applied to the system as, for example, in a long space mission in which there are periods of time during which power is conserved by turning off the computers. Since there is a probability of failure associated with unpowered equipment, more than one module can fail during the power-off time without the failure being detected.

(ii) Even when the system is powered, it is possible for a failure to occur but for the failure not to cause an error until a later time. This phenomenon is called error latency and is discussed in [Shedletsky, 1975A] which reports on work carried out under a present NSF grant#. In this paper it is shown that there are situations for which the error latency (time between occurrence

---

*GJ 27527  #GJ 40286
of a fault and its detection) is comparable to the mean time between failures. Thus it is possible for a second fault to occur before the first fault has been discovered. This has the same effect as the occurrence of a multiple fault.

In work carried out under a present NSF grant* two designs of self-purging systems which withstand multiple failures have been developed [Losq, 1974]. These designs resulted from a detailed theoretical study of the requirements for multiple-fault tolerance in redundant systems. The reliability analysis of the designs shows that their reliability is equal to that of a standby hybrid system with powered spares for single-module failures. While the standby hybrid systems can easily have a system failure as a result of a multiple fault, the designs of [Losq, 1974] are shown to have a very high probability of recovering from a multiple failure.

2.1.4 Comparisons. One general objective of reliability modeling is the development of techniques for choosing a particular redundancy scheme and configuring the system to satisfy a particular design objective. As a result of the research just summarized, it is possible to provide the following guidelines for choosing a redundancy technique:

(i) For extremely short mission times masking redundancy is optimal.

(ii) For mission times whose duration is of the same order of magnitude as the simplex (non-redundant) system mean life self-purging redundancy provides the best performance.

*NSF GJ 40286
(iii) For very long mission times standby redundancy is best, unless unpowered modules have the same failure rate as powered modules -- in which case self-purging redundancy should be used.

2.2 Signal Reliability.

Almost all of the literature on reliability modelling of digital systems has been concerned with functional reliability, the probability that the system realizes the desired design function. We have found that many useful results can be obtained by using a different reliability measure, the signal reliability, which is defined as the probability that the given signal is correct. The signal reliability differs from the functional reliability in that it allows for situations in which a signal will take on the correct value for a given input even in the presence of a fault. Signal reliability was discussed in [Amarel, 1962] in which it was called input-output reliability but has been neglected since.

Fundamental to the study of signal reliabilities is the ability to calculate the probability that a circuit output will take on a given value (usually taken to be 1) when the probabilities that the inputs take on given values are known. Research carried out under a present grant* has resulted in techniques for calculating these probabilities for combinational circuits [Parker, 1975A; Parker 1975B; McCluskey, 1974] and for sequential circuits [Parker, 1975C]. These techniques were then used to develop a method for calculating the signal reliability of a combinational circuit [Ogus, 1975].

* NSF grant GJ 40286
An important application of the signal reliability parameter is in calculating the error latency of faults in digital circuits. The error latency of a fault is defined as the amount of time which elapses between an occurrence of the fault in a circuit and the first appearance of an error. Techniques have been developed for calculating error latencies and are reported in [Shedletsky, 1974A; Shedletsky, 1975A] for combinational circuits and in [Shedletsky, 1974B] for sequential circuits.

The concepts of signal reliability and error latency are applicable to situations in which the signals being applied to the circuit inputs can be modelled as random variables. There are two general types of situation for which such a model of the inputs is valid. One occurs when the circuit inputs are being generated by a random (or pseudo-random) source as in random testing or random test set generation. The other situation arises when the inputs are, in fact, deterministically generated but the generation mechanism is sufficiently complex so that it is not possible to characterize it more simply than as a random variable. An example of the second situation might be an adder circuit contained within a computer's arithmetic unit; the inputs to the adder are deterministic but the generation mechanisms are too complex to model directly.

In random test set generation a test set for a given circuit is constructed by simulating the fault-free circuit in parallel with circuits containing all faults to be tested for [Agrawal, 1972]. Inputs for the simulated circuits are determined by some random mechanism. The input sequence is continued until all of the faulty circuits have
produced at least one output which differs from the output of the fault-free circuit or until a sufficiently high percentage of all the faulty circuits have met this condition. It is usually not possible to achieve detection of all the faults; in a practical circuit some of the faults may be untestable because they involve (inadvertently) redundant equipment or it may be uneconomic to insist on 100% testing.

In random testing randomly-generated inputs are applied directly to the physical circuit to be tested and also to a reference circuit. If the two circuits' outputs do not differ during the application of the input sequence, the circuit under test is accepted as good.

Both in random test set generation and in random testing, questions arise concerning the length of the required input sequence to guarantee that a fault is detected with a given probability, the best statistics to use for the input source, etc. Such questions were addressed in [Agrawal, 1972; Parker, 1973; Rault, 1971; Schnurmann, 1975; David, 1975]. The techniques developed for studying error latency have been found to be directly applicable to these problems of random testing and random test set generation. The error latency approach has produced much more precise results concerning these questions [Shedletsky, 1974A; Shedletsky, 1974B; Shedletsky, 1975A; Shedletsky, 1975B].

2.3 Self-diagnosing Computer Design.

Decreasing hardware costs and increasing reliability requirements have led us to investigate the possibility of designing a low-cost self-diagnosing computer. Techniques have been demonstrated for detecting failures in all of the subsystems of a typical computer processor [Wakerly, 1973A] as well as in a large variety of peripheral
and I/O subsystems [Usas, 1975c]. A simplified design example of a 16-bit self-diagnosing computer was given in [Wakerly, 1973] and a proposal to carry out the detailed design and construction of a 32-bit self-diagnosing computer has been submitted to NSF by John Wakerly. These computer designs incorporate the results of theoretical studies in the areas of error-detecting codes, self-checking circuits and fail-safe circuits.

In designing a self-diagnosing computer it is necessary to choose the error-detecting codes, to design the checking circuits which provide error indications, and to design the translation circuits necessary to convert from one code to another if more than one type of code is used. The translation and checking circuits should be self-checking so that failures in these circuits do not override the fault-detection capability of the computer.

The use of error-detecting codes in computers has received a great deal of theoretical attention and is actually fairly widespread in actual systems. However, most of the previous use of error-detecting codes has been confined to one subsystem of the computer. In the self-diagnosing computer with which we have been concerned it is necessary to use error-detecting codes throughout the entire computer system. Thus a study was conducted of error-detecting codes from the point of view of their applicability for use for error-detection in all of the parts of the computer system. This study has led to some new results on properties of error-correcting codes.
2.3.1 Error-detecting Codes. One of the most attractive error-detecting codes for use in arithmetic units is the low-cost residue code [Avizienis, 1971]. Study of this code led to the demonstration that it is also quite effective in detecting unidirectional multiple errors in mass storage devices and repeated-use multiple faults in byte-serial data transmission [Wakerly, 1975]. Rules for using residue codes for checking various arithmetic and non-arithmetic shift operations were also developed [Wakerly, 1973].

Checksum codes were studied and techniques for check symbol prediction were developed, thus making the checksum codes competitive with residue codes for checking arithmetic operations and data transmission and storage in some situations [Wakerly, 1974A].

2.3.2 Self-checking Circuits. In a computer arithmetic unit it is necessary to check both arithmetic and logical operations. The logical operations are difficult to check with the same codes which are effective for arithmetic operations. In studying this problem a new class of circuit called "partially self-checking circuits" was conceived which alleviates this problem [Wakerly, 1974B].

A number of new designs for self-checking circuits and checkers were developed. They are presented in [Wakerly, 1973] along with efficient practical MSI and LSI implementations.

It may be desirable to use different error-detecting codes for main memory and for the processor. For use in such a situation totally self-checking interfaces between codes were developed [Wakerly, 1973].
A persistent problem in designing error-detection circuitry for digital systems has been the design of a circuit to detect failures in the timing signal. A circuit to provide such a capability was discovered in connection with the studies of failure detection techniques for peripherals [Usas, 1975A].

2.3.3 Fail-safe Circuits. Our studies of failure-detection techniques for input-output systems has led to the conclusion that fail-safe circuits (circuits which produce unidirectional errors) are useful for controlling errors in such systems [Usas, 1975B]. Previous realization techniques for fail-safe circuits have assumed delay line memory elements [Diaz, 1973; Sawin, 1974] or have assumed that JK flip-flops which always have complementary outputs (even in the presence of failures) are used [Tohma, 1974]. Our studies of fail-safe circuits have led to a realization technique which uses D flip-flops and does not require the assumption that failures do not destroy the complementarity of their outputs [Usas, 1975C].

2.4 Multiple Fault Studies.

The bulk of the theoretical work on reliable computation has been concerned with situations in which it is assumed that only single faults occur. This is justified on the basis of independence among faults. Two phenomena act to invalidate this single-fault assumption: the trend towards higher levels of integration and the error latency property discussed in Section 2.2.
2.4.1 **Fault Masking.** A test set for single faults may fail to detect all multiple faults because of the property of fault masking: the presence of one fault may mask the effect of a test in detecting another fault. This property of fault masking has been studied and a technique for determining the existence of fault masking with respect to a given test has been developed [Diaz, 1975A]. This study also results in a method for extending a test set which detects all single faults but not all multiple faults into a test set for all multiple faults.

2.4.2 **Iterative Networks.** An important special class of combinational circuit is the (unilateral) iterative network or iterative logic array. Because of the importance of this class of circuit, special testing techniques have been developed [Kautz, 1967; Menon, 1971; Landgraff, 1971; Friedman, 1973]. These methods all make the assumption that at most one cell of the network is faulty. We have been able to develop a testing method which does not require this assumption but instead permits any number of cells to be faulty [Diaz, 1975B]. A test procedure is given whose length is independent of the number of cells in the network. This procedure is applicable for any iterative network whose basic cell is reduced and has strongly connected components. Many practical networks satisfy these conditions. For those that do not, a **simple modification** is developed to make them testable with a constant number of tests.
2.5 References


3. STUDY OF MAINTAINABLE COMPUTERS

Summary of Previous Results and Work in Progress

3.1 Applying Triple Modular Redundancy to Small Computers

The decreasing cost of computer hardware is increasing the feasibility of triple modular redundancy (TMR) as a means of providing fault tolerance in small computer systems. In a TMR system, each module is triplicated and majority voters are used at the interfaces between modules to mask the effects of single module failures.

In order to discover any special problems that might occur in applying TMR to small systems, we began the design of a TMR microcomputer system. Our first discovery was that applying TMR to arbitrary sequential modules requires special consideration of the effects of transient failures (section 3.1.1). Next we examined the reliability of triplicated memory systems and found that triplication is a better choice than coding for small systems (section 3.1.2). Finally we constructed models of several different TMR microcomputer system configurations and derived the reliability of each (section 3.1.3). We have shown that careful use of TMR can improve the mission time of small microcomputer systems by a factor of 3 to 10.

3.1.1 Effects of Transient Failures in Sequential Modules

Triple modular redundancy was first proposed by von Neumann [1956] as a means of masking the effects of transient component failures in a system. Applying redundancy at the component level was also proposed by Moore and Shannon [1956] as a means of masking transient failures in relay networks.
Subsequent researchers showed that TMR could be used to mask the effects of permanent failures in a system [Dickinson and Walker, 1958], and that TMR could be applied at higher levels than the component level [Flehinger, 1958; Lyons and Vanderkulk, 1962; Brown, Tierney and Wasserman, 1961].

Recent investigations into the reliability of TMR systems have concentrated on the effects of permanent failures [Gurzi, 1965, Rubin, 1967; Abraham and Siewiorek, 1974]. It has been assumed that most transients are masked and leave no permanent effect in TMR systems [Lyons and Vanderkulk, 1962]. That is, the effect of a transient failure is masked by the voter during the short period of time that it is present, and the effect disappears with the transient. Once a transient disappears, another can be tolerated. The only transients that were recognized to cause system failures were those that affected more than one member of a replicated module trio at the same time [Avizienis, 1967].

That multiple transients over a period of time could be tolerated was in fact shown to be true by von Neumann [1956] and by Moore and Shannon [1956] for redundancy schemes applied at the component level. In systems that apply redundancy at a higher level, we have demonstrated that this is not always true, that the effect of a single transient failure in one module can be permanent [Wakerly, 1975a]. The transient has a permanent effect when the affected module is a sequential machine that is never re-synchronized.

The need to re-initialize modules after transient failures has been long recognized for self-repairing systems with selective redundancy [Avizienis, 1967; Avizienis, 1971]. However, the re-synchronization problem in TMR systems has been neglected because redundancy has in the past been
applied at a level low enough that the problem did not occur [von Neumann, 1956; Moore and Shannon, 1956; Dickinson, 1964]. The increasing complexity of integrated circuits is continually increasing the minimum level at which TMR may be applied in systems using standard MSI and LSI components. For example, in a microcomputer system, voting must be applied at the processor level. Reliability analysis of such systems may indicate that TMR should be applied at a level even higher than the minimum [Gurzi, 1965; Abraham and Siewiorek, 1974; Longden, 1966]. Yet as redundancy is applied at higher levels, transients become more likely to have permanent effects. Thus the designer must be aware of the effects of transients in specifying the application of redundancy to a system.

We have proved [Wakerly, 1975a] that multiple transients can be tolerated by a TMR system if and only if a synchronizing sequence is applied to the system periodically during normal operation. We have shown system structures that provide for easy synchronization, and we have suggested ways of modifying systems that do not normally receive synchronizing sequences.

3.1.2 Reliability of TMR Memory Systems

A TMR system can be partitioned into a number of cells so that the system reliability is the product of the cell reliabilities [Abraham, 1974]. The simplest type of cell, shown in Fig. 3.1, has one triplicated module with voters on each input. (There is a separate voter circuit for each input bit of the module.) The cell tolerates any single module or voter failure', since errors produced by such a failure will be corrected by the voters at the inputs of the next cell. Assuming for simplicity perfect voter reliability,
Fig. 3.1 A TMR voter-module cell.

Fig. 3.2 An n-bit by w-word semiconductor memory module.
the cell reliability $R_{\text{cell}}$ is a function of the module reliability $R_m$,

$$R_{\text{cell}} = R_m^3 + 3R_m^2(1-R_m). \quad (3.1)$$

The semiconductor memory module of a small computer system can be modeled as shown in Fig. 3.2. There is some shared address decoding and driving circuitry, an array of memory chips, and perhaps some shared output circuitry. The memory array consists of $n_s$ 1-bit by $w$-word memory chips arranged in an $n \times n_s$ matrix to form the $n$-bit by $w$-word array. If the memory chip reliability is $R_C$ and the reliability of the common circuitry is $R_d$, then module reliability is $R_m^{n_s} R_d$, and it would appear from (3.1) that the reliability of a TMR memory system is

$$R_{\text{sys}} = (R_m^{n_s} R_d)^3 + 3(R_m^{n_s} R_d)^2 (1-R_m^{n_s} R_d). \quad (3.2)$$

The above analysis neglects the organization of the memory array. Assuming that there is a voter for each bit of the memory output, the system fails only if there is a simultaneous error in a single bit position of two of the triplicated memory modules. Consideration of the memory array structure hence leads to the more accurate reliability formula,

$$R_{\text{sys}} = R_d^3 (3R_c^2 - 2R_c^3)^{n_s} + 3R_d^2 (1-R_d) R_c^{2n_s}. \quad (3.3)$$

The reliability expression above always produces a reliability value greater than or equal to (3.2). The improvement obtained by using (3.3) decreases as the reliability of the memory array ($R_c$) relative to the common circuitry ($R_d$) increases. For example, if $R_c=1$ the formulas are identical.
But for typical semiconductor memory systems, the common circuitry comprises only about 10-15% of the total, and so the reliability value obtained by considering the structure of the memory array (3.3) is significantly higher than that obtained by simple analysis (3.2). A typical example is shown in Fig. 3.2.

The above discussion is intended only to give an indication of the nature of our results on memory systems. The actual memory system analysis is somewhat more complex, taking into account voter reliability, the placement of voters for the memory system inputs, the possibility of having different chip types within the memory array, and a solution to the problem of multiple pattern-sensitive failures. The effectiveness of the TMR memory system organization and a system using a single-error-correcting code have been compared. While both systems are guaranteed to correct all single failures in the memory array, analysis has shown that the TMR system is more reliable because it corrects a larger number of multiple failures than the coded memory. Also the TMR system corrects all single failures in the common circuitry (Fig. 3.2) while the coded system does not unless a copy of the common circuitry is provided for each of the memory bit slices. For an 8-bit memory system, coding requires 4 redundant memory bits per word while TMR requires 16. On the other hand, coding requires a decoder that is more complex than the simple TMR voters, especially if the decoder itself is to be fault-tolerant. For small fault-tolerant memory systems that are to be interfaced to a TMR processor, TMR appears to be a much better choice than coding.
8K by 8-bit memory
64 1K by 1-bit memory chips
8 decoder and driver chips
Reliability for all chips:
\[ R_c(t) = e^{-\lambda t}, \lambda = 10^{-6} \text{ failures/hour} \]

Fig. 3.3 TMR memory system reliability.
3.1.3 A TMR Microcomputer System

The thought of applying TMR to microcomputer systems raises some interesting questions. First of all, since a microprocessor is just a single chip, there is some question whether reliability can really be increased in a system that must use many voter chips constructed from the same unreliable technology as the microprocessor itself. Secondly, a microprocessor is a rather complex sequential machine with only limited access to its internal state, and so special care must be taken if the system is to tolerate multiple transient failures.

We will use the simple model of a microcomputer system shown in Fig. 3.4. The system consists simply of a microprocessor and memory, with data, address, and control lines going from the microprocessor to memory and data lines going from the memory to the microprocessor. Connections to peripherals are ignored; for the TMR system it is assumed that each peripheral interface has voters which monitor the I/O commands given by all three triplicated processors.

A typical LSI microprocessor is the Intel 8080 [Intel, 1974]. The 8080 is an 8-bit processor in a 40-pin package. It has 16 address lines, an 8-bit bidirectional data-bus, and 9 control lines entering and leaving the chip. The data bus must be externally split into two one-way buses for voting to be applied, and hence there are a total of 41 lines in an 8080 system that could be voted on. Since three voter circuits can be placed on a single 14-pin package, it is conceivable that a TMR 8080 system could have 3 8080 packages and 41 voter packages (triplicated voters) or 14 voter packages
Fig. 3.4 Microcomputer system model.

Fig. 3.5 Mission time improvement factor for a TMR microcomputer system.
(nontriplicated voters). Since a large percentage of integrated circuit failures are related to problems in packaging and I/O pins rather than circuit complexity, it is quite conceivable that the total voter unreliability in a TMR microcomputer system could approach or even exceed the microprocessor unreliability. In such a system the use of TMR could actually decrease the overall system reliability.

We can use a very simple model to justify the above thesis. Suppose that the reliability of a microprocessor module is $R_m$ and the voter reliability is $R_V$. If $n$ voters are required for each replicated module in a TMR system, then the total voter reliability is $R_V^n$. The total voter reliability can be related to the module reliability by a factor $k$ such that $R_V^n = R_m^k$. The factor $k$ can be interpreted to mean that the total failure rate of the voters is $k$ times the failure rate of the microprocessor module.

For a system with several voter packages for every microprocessor package, $k$ could be in the range 0.1 (very reliable voters) to 2 or more (voter reliability comparable to microprocessor reliability). The reliability improvement obtained by using TMR for various values of $k$ is shown in Fig. 3.5. The figure of merit used is the mission time improvement factor (MTIF), that is, the ratio of the mission times of the TMR system and the corresponding nonredundant system. (The mission time is the amount of time it takes for the system to degrade from its initial reliability of 1 to some terminal reliability $R_t$.) For the perfect voter case ($k=0$), the theoretical maximum MTIF is obtained, 2.84 for $R_t = .95$ or 2.08 for $R_t = .90$; but for imperfect voters ($k>0$), the MTIF can be much less. If module and total voter reliabilities are equal ($k=1$), the MTIF is only 1.42 for $R_t = .95$, and about 1.0.
for $R_t = .90$ (the mission times of the TMR and the nonredundant systems are equal). If the total voter failure rate is twice the module failure rate ($k=2$), the TMR system has a shorter mission time than the nonredundant system for either value of $R_t$.

Of course the above model is an oversimplification because it neglects the reliability of memory and other support circuits that are present in typical microcomputer systems. Including these components as part of the module would increase voter reliability relative to module reliability and hence reduce the value of $k$. However, the improvement is not dramatic and we have found that for practical systems it is often still worthwhile to try to increase total voter reliability by decreasing the number of voters.

Recalling that there are 41 lines that might be voted on in an Intel 8080 system, the most drastic reduction in the number of voters would be obtained by voting on none of the lines. The system would consist of three identical microcomputer/memory systems, each initialized to the same starting state and operating synchronously from a common fault-tolerant clock. Peripherals would have their own internal voters and they would perform operations as dictated by the majority of the replicated address, data, and control lines of the three identical systems. The problem with this scheme is that there is no mechanism for a microprocessor/memory system to be resynchronized after a transient failure, since there is no coupling among the replicated systems. Even if transient failures do not occur, we shall see that this system is less reliable than the next system we describe.

Suppose that voters are placed at the master reset input and the 8 data inputs of each microprocessor, as shown in Fig. 3.6. The address, data out, and control lines of each microprocessor go directly to the corresponding
Fig. 3.6 Minimum TMR microcomputer configuration for resynchronization.
memory module without any voting. We have proved that this configuration has the minimum number of voters needed to provide re-synchronization after transient failures. For example, suppose a transient failure causes several registers of one microprocessor and several words in the corresponding memory module to contain incorrect data. Each of the incorrect registers is resynchronized with correct data if it is loaded from memory, since the voters insure correct memory output regardless of any possible errors in the state of one of the memories. Once the microprocessor is resynchronized, the memory is resynchronized by loading the incorrect memory words from the microprocessor.

Of course, it is possible that a transient failure can affect not only the register state but also the program state of a microprocessor. In general the microprocessor can attain any erroneous state and before being resynchronized it can create arbitrary errors in the corresponding memory module. It is this possibility that necessitates a voter on the master reset line of the microprocessors. Associated with each microprocessor is some interface circuitry that can be instructed by the software to initiate a hardware reset. Periodically the software would cause such a reset to occur, and since the reset line is voted on, a completely unsynchronized microprocessor must still obey the reset command. The reset command causes the microprocessor to begin executing a routine at some fixed location. The routine in this case is a synchronizing routine that first initializes all of the processor registers from memory, and then corrects any possible errors in a single memory module by sequentially reading and then rewriting every word in the memory.

Resynchronization after transient failures is also possible if voters
are placed in the data lines going to the memory rather than coming out of the memory. However, placing the voters in the output lines results in better system reliability, since the memory reliability improvement for semiconductor memories discussed in section 3.1.2 is applicable. This improvement does not occur when voters are placed on the memory input only, since a single bit error in the memory output can produce several erroneous bits when the affected data is processed by the microprocessor.

The reliability analysis of the system in Fig. 3.6 is similar to the analysis of memory systems in section 3.1.2. In fact, equation (3.3) can be used to find the system reliability neglecting voters by simply considering the microprocessor to be part of the common circuitry of the memory module \( \text{R}_d \). In a similar way the reliability of the scheme with no voters and the scheme with voters on the data inputs only can be derived by use of equation (3.2). The exact reliability including voters for a nonredundant system and the three TMR system configurations is shown in Figs. 3.7 and 3.8 for two typical sets of parameters. It can be seen that for these two cases TMR increases the mission time by a factor of 3 to 4 for a terminal reliability of .95.

The work in this and the previous section is still in progress and a technical report is in preparation.
Fig. 3.7  TMR microcomputer system reliability.
Fig. 3.8  TMR microcomputer system reliability.
3.2 Self-Checking Circuits*

The theory of self-checking combinational circuits was studied by the proposed principal investigator under previous NSF grants (GJ-27527 and GJ-40286). Recent work has focused on applications of self-checking combinational circuits and on the theory and applications of self-checking sequential circuits. Self-checking adders using checksum codes have been studied (section 3.2.1); a self-checking checker for periodic signals such as clocks has been designed (section 3.2.2); and a design approach for fail-safe sequential machines using realistic fault assumptions has been developed (section 3.2.3).

3.2.1 Checked Binary Addition Using Checksum Codes and Check Symbol Prediction

The code words of a checksum code are vectors of b-bit bytes with a single check byte that is the modulo $2^b$ sum of the data bytes. A checksum code can detect any error that affects only a single byte of a code word, and hence these codes are quite effective for detecting failures in data transmission and storage in byte-sliced systems. In such systems the circuits that handle the data are partitioned into b-bit byte slices, and hence a single component failure always results in a detectable single-byte error.

Checksum codes are not arithmetic codes, so that the check symbol of the sum of two code words cannot be derived from only the check symbols of the given code words. Hence, addition of two code words cannot be

* The work in this section has been supported by both NSF grants GJ-40286 and GK-43322.
checked by simply adding the check symbols, as in arithmetic codes. However, we have shown that check symbol prediction techniques can be used effectively in the design of self-checking adders for checksum code words. Such techniques had been developed earlier for simple parity-check codes [Sellars, 1968], requiring all of the inter-bit carry circuits of an adder to be duplicated. We have developed a similar technique for checksum codes which requires only the inter-byte carries to be duplicated [Wakerly, 1975b]. While duplicating the inter-bit carries requires an overhead of over 50% in the adder circuitry, duplicating the inter-byte carries requires only about 12% extra circuitry in a conventional MSI adder chip.

Although a self-checking adder for checksum code words is still slightly more expensive than a self-checking adder for words in a residue code (an arithmetic code), the checksum code still enjoys some advantages. As mentioned earlier, a checksum code detects all errors in a single b-bit byte. The residue code detects all single-byte errors except all 0's changed to all 1's or vice versa. Hence the single-byte error-detecting capability of a checksum code is superior to a residue code. For applications where single-byte error detection is of primary importance, a checksum code can be used to provide the required error detection capability; checked addition can still be performed at a cost only slightly higher than with residue codes, and much lower than with parity-check codes.

3.2.2 A Self-Checking Checker for Periodic Signals

Periodic signals have a known behavior, and deviations in their waveforms may indicate failures in the signal source. Monitoring these signals
can be a valuable technique for detecting both hardware and software failures in a computer. Many circuits have been presented in the literature to check for errors in these signals (for example, see [Koczela, 1971] or [Chang, 1973]). These circuits, however, share a common weakness in that they are all susceptible to undetected internal failures; that is, they are not self-testing.

A totally self-checking periodic signal checker has been designed [Usas, 1975a]. As shown in Fig. 3.9, the checker consists of two monostable multivibrators M1 and M2. M1 is triggered on the leading edge of the input signal and produces an output pulse of fixed width equal to \( t_{on} \) the expected on-time of the input signal. M2 is triggered on the falling edge of the input and produces a pulse of duration \( t_{off} \) the expected off-time of the input.

The waveforms in Fig. 3.9 show that the output of M1 is a regeneration of the input while M2 produces the complement of the input. Hence during correct operation the checker output is always either 10 or 01. If the input signal becomes stuck on or off, the checker output becomes 00, and if the on-time or off-time vary more than a few percent from their expected values the checker output becomes either 00 or 11. It has been shown [Usas, 1975a] that failures in the monostables themselves produce either the 00 or the 11 output, and the checker is totally self-checking.

There still remains the issue of monitoring the checker output and sounding an alarm when the 00 or 11 output appears. This issue has been thoroughly discussed in [Usas, 1975d], and fail-safe solutions have been presented.
Fig. 3.9 Totally self-checking clock checker.
3.2.3 Fail-Safe Sequential Machines

Self-checking circuits have two properties, self-testing and fault-secureness. The self-testing property guarantees that all faults are tested in normal operation, while fault-secureness guarantees that any fault produces either the correct output or an error indication. The fault-secureness property for self-checking circuits is equivalent to the definition of "fail-safe" given by Tokaoka and Ibaraki [1972]. For historical reasons the name "fail-safe" has been applied to fault-secure sequential machines.

A number of researchers have recently been concerned with the design of fail-safe (or fault-secure) sequential machines. Most of the contributions to a solution of the problem have assumed a simple delay element as the realization of the required memory function [Diaz, 1973; Sawin, 1974]. Tohma [1974] has suggested a design approach using more practical and familiar memory devices such as JK flip-flops. However, he assumes that the double-rail outputs of a faulty flip-flop are always complementary. That is, if the fault causes one output line to be an erroneous 1(0), then the other line is always an erroneous 0(1). Unfortunately this is not the case if one considers the circuit implementation of a JK flip-flop and common integrated circuit failure modes. Usas [1975b] has presented a design approach that is based on the use of D flip-flops as memory elements and that assumes a more realistic model of the faulty behavior of the flip-flop.

Analyzing the circuit implementation of a conventional D flip-flop, Usas [1975b] shows that there do indeed exist faults such that the flip-
flop outputs are identical, but that no fault results in both erroneous 1's and 0's on the flip-flop outputs. That is, if a fault $f$ produces an erroneous 1(0) on a flip-flop output for some input value, then there is no input that will produce an erroneous 0(1) on that output in the presence of $f$. This permits the erroneous value to be held in the flip-flop by a suitable choice of next-state mapping for illegal states. The design procedure results in a machine such that any flip-flop failures cause the machine to go to and remain in a "trap" state. Trap states can easily be detected in order to produce an error signal.

For the detection of single flip-flop faults, the design technique requires only one extra flip-flop over the normal design. However, the combinational excitation circuitry is more complex because the technique requires each flip-flop to have a separate excitation circuit; sharing is not allowed. A second design procedure is given that uses more extra flip-flops, two for machines with up to 126 states and three for up to about $2^{37}$. This procedure allows shared logic realizations of the flip-flop excitation functions and usually results in a lower overall cost for the machine.
3.3 Self-Diagnosing Computer Design

The object of this research has been to develop design principles for self-diagnosing computer processors, machines with extra hardware for automatic detection and diagnosis of internal failures. This work began under NSF grants GJ-27527 and GJ-40286, and since July 1974 has been supported by GK-43322.

Our previous efforts were documented by a report [Wakerly, 1973] and a summary [Wakerly and McCluskey, 1974]. A research monograph based on this work is also being written. Our current efforts are towards applying the general results of the previous studies to the design and actual implementation of a self-diagnosing computer processor. Several aspects of such a design were not dealt with by the previous studies and remain to be explored.

The planned processor has 32-bit data paths and a microprogrammed control unit. The instruction set has a full complement of arithmetic, logical, and shift operations, with both register-to-register and register-to-memory operations available. The execution time for a typical register-to-register instruction will be about 1.2 microseconds. The data paths and control unit of the processor have been designed and the microcode is currently being written.

There have been a number of previous efforts of self-checking computer processor designs like our self-checking design; these designs used error-detecting codes and some duplication to detect hardware failures. However, our design is different from previous designs in several respects. First
of all, we are designing a general-purpose processor with an instruction set and execution times comparable to typical minicomputers. Our processor can be contrasted with the JPL STAR [Avizienis, 1971] which was a byte-serial machine, and Bell Labs ESS [Chang, 1973] which had no add instruction and required two microcycles for every register-to-register transfer. Secondly, the cost of checking in our design will be less than 25% of the total hardware cost for a 32-bit processor. Thirdly, our design makes efficient use of standard MSI integrated circuits, and the data paths and control unit are designed to allow straightforward integration as LSI circuits comparable to those available commercially today. Finally, although previous machines claimed a high degree of self-checking for single gate faults, in our design we can guarantee detection of 100% of the failures that affect a single integrated circuit package.

3.3.1 Data Path Checking

The previous studies [Wakerly, 1973] showed how error-detecting codes and self-checking circuits could be used to detect errors in the data paths of a typical processor. The processor data paths are byte-sliced, so that each register, adder, multiplexer, etc., that handles a b-bit byte of data is a separate component. The data paths of an nb-bit processor can then be implemented by n b-bit slices operating in parallel, with appropriate interconnections for arithmetic carries, shifts, etc. In such a system, a single component failure affects only a single byte, and errors can be detected by the use of a b-bit residue code. The residue code associates with each datum
in the processor a \( b \)-bit check symbol which is the residue of the datum modulo \( 2^b - 1 \). An extra \( b \)-bit check slice is thus required to store and process the check symbols of the data residing on the data path slices. In our previous work we showed that the extra check slice is sufficient to detect all single byte errors regardless of the actual data path width, be it 3\( b \) bits, 8\( b \) bits, or more. However, it was not clear at that time whether the check slice could be the same as the data path slices in a practical design or whether the check slice would have to be different.

Our current work has shown that it is possible to design a standard byte slice that serves for both data and check symbols, at a cost comparable to a byte slice for an unchecked machine. Hence for an \( nb \)-bit machine using \( b \)-bit slices, the cost of data path checking is about \( 1/n \), or 13\% for a 32-bit machine using standard 4-bit components.

Our studies have shown that the only essential difference between a check slice and a data slice is that the check slice must have a provision for modifying the check symbol for certain arithmetic operations and re-loading it for non-code-preserving operations. Except for this the check symbols are processed exactly the same as corresponding data in the data slices. Hence we have designed a 4-bit slice, a simplified version of which is shown in Fig. 3.10. The byte slice has a 16-word general register file, 8 scratch pad registers, memory address and data registers, shifting and byte-swapping logic, and ALU. A processor that uses this slice can efficiently perform all of the operations of a typical minicomputer: loading
Fig. 3.10 Processor data path slice.
and storing data, arithmetic and logical operations, shifts and rotates, microprogrammed multiply and divide.

The only difference between the slice of Fig. 3.10 and a conventional slice is that the data path from the ALU output back to the register inputs is broken; in a conventional slice this connection is usually already made. The broken connection allows the system implementation shown in Fig. 3.11. For the data path slices, the ALU output is connected directly to the register inputs. For the check slice, the ALU output goes through combinational logic that modifies the check symbol as required before returning it to the registers. At the end of each machine cycle, the output of the data and check slices is loaded into a check register and checked for validity. For non-code-preserving operations it is possible to reload the check slice with a new check symbol generated from the non-code result.

The byte slice suggested by Fig. 3.10 has been designed in detail. Using conventional TTL integrated circuits, the slice can perform one operation every 200 nanoseconds. In our current design the combinational logic external to the check slice is implemented with a small, fast read-only memory and adds 30 nanoseconds to the maximum cycle time. Hence checking requires a 15% overhead in the cycle time of the machine.

The data path slice of Fig. 3.10 is similar to currently available LSI data path slices [Monolithic Memories, 1974; Rattner, 1974]. In fact, the slice has been designed with LSI in mind; it should be possible to fabricate the entire slice on a single integrated circuit chip. Our data path slice design requires no more circuitry and only four more pins than a comparable slice that cannot be used as a check slice.
Fig. 3.11 Configuration of byte slices for 32-bit processor.
3.3.2 Control Unit Checking

In our previous work [Wakerly, 1973] we showed how error-detecting codes can be used to detect errors in the microprogram storage of a microprogrammed control unit. If the storage is implemented using $b$-bit by $n$-word ROM chips, then only one extra ROM chip is needed to hold $b$ check bits for each word and insure detection of all single package failures. In our present processor design we expect to have 1K of 44-bit microprogram words implemented with 11 1K by 4-bit chips. One extra chip is required for checking and hence the overhead for checking the microprogram storage is $1/11$, or about 9%.

In addition to microprogram storage a microprogrammed control unit requires registers to hold the instruction and the microprogram address and a fair amount of combinational logic for computing the next microprogram address. The next address could be computed for example as the next sequential address, as a subroutine return address, or as a function of internal flags (carry, overflow, etc.). In [Wakerly, 1973] we indicated that the only way to detect errors in some of the control unit functions was by duplication, while errors in other functions could be detected by clever use of the existing coding and by taking extra microprogram steps for checking. In our subsequent studies we have found that the "trick" methods are not sufficiently general for efficiently implementing a typical processor, and so we have rejected them in our design. Instead we plan to duplicate the instruction register, the microprogram address register, and the next address logic.
As shown in Fig. 3.12, there will be two identical "microprogram control" modules. The modules operate in parallel, so that both load their instruction register from the processor data bus, test processor flags, and sense microprogram jump conditions and addresses. The output of a module is a 10-bit microprogram address. The output of one of the modules is used to address the microprogram memory while the output of the second is compared with the first for equality.

Of course, the overhead for duplication of the control module is 100%, far from our overall system goal of 25%. However, the control module is only a part of the control unit, and the rest of the control unit (i.e., the microprogram memory) requires an overhead of only 9%. The fraction of the total control unit cost attributable to the duplicated control module depends on how that module is implemented. The present control module is designed with LSI in mind, so that the entire module could be fabricated on a single chip comparable to existing commercially available LSI control units chips [Rattner, 1974]. Only that single LSI package would be duplicated.
Fig. 3.12 Microprogrammed control unit structure.
3.3.3 Diagnosis

After an error is detected in the self-checking computer processor, diagnosis must be initiated to isolate the failure to one replaceable module. Diagnosis will be performed by special microprograms called microdiagnostics. Although these microprograms have not been written, the hardware features necessary to support diagnosis have been considered in the present processor design. The goal is to minimize the amount of hardware that must be working for successful diagnosis to take place.

Assuming single package failures, any failure in the data paths of the machine can be easily isolated by microdiagnostics to one replaceable slice. Since there is one extra data path slice for checking, an operator can manually reconfigure the machine for continued operation by replacing the failed slice with the check slice and disabling data path checking until repairs are completed.

A failure in a single microprogram memory ROM package can be detected, but using the present single-error-detecting code diagnosis is not possible unless another copy of the ROM contents are available for comparison. This would mean duplicating the entire microprogram memory or having a copy available on auxiliary storage. If the ROM failure is not catastrophic, then it is possible to load a copy of the microprogram from external storage and perform the check. However, a catastrophic failure of a ROM package may affect all microinstructions and prevent this diagnosis from taking place. An alternative currently under study is to use a more sophisticated code to provide both detection and location of ROM failures. For memory
words of up to 60 bits, two 4-bit check symbols in a 2-redundant  
b-adjacent code [Bossen, 1970] are sufficient to detect and locate (or  
even correct) all 4-bit errors.

A failure in one of the two duplicated control modules (Fig. 3.12)  
is detected by the equality checker monitoring their microprogram address  
outputs. Once a failure is isolated to one of the two modules, the other  
can manually be made the primary module (if it isn't already), and system  
operation can continue with checking disabled. The problem then is to  
determine which of the two modules is the faulty one. There are currently  
manual procedures for making this determination, but no automatic procedures  
short of triplication and voting have yet been devised.

The work in this and the previous two sections is still in progress  
and a technical report is in preparation.

3.3.4 Peripheral and Input/Output Checking*

This work has been directed toward the problems of detecting errors  
and performing fault diagnosis in the components associated with the input/  
output system of a computer. Such units as CPU interfaces, channels, bus  
controllers, device interfaces, device controllers, and devices have been  
studied with the goals of guaranteeing an indication of the occurrence of  
the error (self-checking) and providing expeditious repair (self-diagnosing).

Checking the validity of data transferred between a CPU and peripherals  
is fairly simple -- the same codes used to check memory data or CPU data

*The work in this section has been supported by both NSF grants GJ-40286  
and GK-43322.
can be used to check the data transfers over I/O buses. Problems and solutions in interfacing between different codes for processors and peripherals have been given [Wakerly, 1973]. However, there remain unsolved problems in controlling the transfers of data between processors and peripherals, problems that do not occur in the design of a synchronous CPU. Since I/O buses usually have at least some asynchronous signals and some "hand-shaking", there are many new problems to consider. The sequential nature of I/O control led to the more fundamental studies of self-checking sequential machines described in sections 3.2.2 and 3.2.3. With these new techniques, self-checking controller designs are possible. In addition, specifications for signaling in digital buses to insure error detection have been outlined.

Error detection in peripheral devices is a difficult problem because of the diversity of devices in common use. However, the problem has been approached by considering device classes (disks, tapes, card equipment, etc.), and a list of applicable methods has been compiled. Also, the design of a CPU-controlled probe for peripheral diagnosis has been completed and operational guidelines are being assembled.

This first phase in research work on input/output fault detection and diagnosis is near completion and will be documented by a Ph.D. thesis [Usas, 1975c].
3.4 REFERENCES PREPARED UNDER PREVIOUS NSF GRANTS

* GJ-27527
† GJ-40286
+ GK-43322


3.5 ADDITIONAL REFERENCES


4. INVESTIGATION AND EVALUATION OF DUAL COMPUTER CONFIGURATIONS

4.1 Introduction

This report summarizes the current status of the project which is concerned with the investigation and evaluation of dual computer configurations.

For the past two years, work has been underway under the sponsorship of NASA Ames Research Center to study reliable computer systems for use in Short Takeoff and Landing (STOL) aircraft. A prototype system has been developed by the Charles Stark Draper Laboratories (CSDL) in Cambridge, Mass., for this purpose, and was delivered to the NASA Ames Research Center in mid-1975 for a series of flight tests under the direction of personnel from NASA. This system, known as the SIRU system [1], consists of a set of navigational instruments [the Strapdown Inertial Reference Unit (SIRU)], and a digital computer complex to process information from the SIRU and display this information to the pilot of the aircraft. The SIRU navigation package is fault-tolerant and it was desirable to develop a reliable computer system that would match the high reliability of the SIRU instrument system. The computer complex used is a dual processor, real-time system using two Honeywell H316 central processing units, with a special purpose arbiter component to evaluate the operation of the two Honeywell processors. Both processors execute the same algorithms, and the arbiter designates one processor as master and one as backup for each basic operational cycle.

The initial work on the project involved the study of the actual SIRU system itself. Research work is ongoing in this area, and we hope to
use the SIRU system as a basis for a more general study of dual computer configurations. Theoretical studies have been carried out of both the hardware and software aspects of the dual computer system, and these investigations are continuing. A second thrust of the present research is to thoroughly evaluate the prototype system both using the theoretical results as guidelines, as well as through experimentation on the actual system. Based on these studies we hope to make recommendations which could improve the prototype system performance.

4.2 Summary of Technical Areas in the Project

4.2.1 The Siru computer system

CSDL carried out the initial design of the computer system [1], and the design was studied by the Digital Systems Laboratory (DSL). It was decided that closer collaboration between DSL and CSDL would prove fruitful and, as a result, meetings were held at CSDL in November 1973, May 1974 and December 1974 during which the progress of the design was discussed and reviewed. Thereafter, the final system design was decided upon, and CSDL carried out the actual construction of the system.

A very important component of the SIRU dual computer system is the arbiter unit. In the prototype system, the arbiter forms an opinion as to the functional status of the system and communicates this information to the pilot; in later versions of the system, however, the arbiter might be responsible for the control of actuators in the aircraft and thus would become a very critical unit. In the latter case, the arbiter would have to be a highly reliable module.
A reliability model has been developed to study the effect of the actual arbiter reliability on the overall system reliability. In the prototype system the arbiter is the only component which is not redundant and the model assumes that an arbiter failure will result in system failure. Results have been obtained which indicate the minimum reliability value the arbiter could have before the redundant system becomes less reliable than the nonredundant (simplex) system. Other aspects, such as the effect of the arbiter on the system mission times, have been studied, and we hope to generalize these studies to enable more general systems to be modeled.

We are also studying the possibility of improving the arbiter for future systems. The first aspect is the improvement of the arbiter's reliability. Since the arbiter has an important role in the system and will become even more crucial in later versions of the system, it will become necessary to realize the unit using some type of redundant structure. Redundant implementations of the arbiter are being studied to determine which redundancy structure would be the most suitable for the arbiter from cost, reliability and speed viewpoints.

A second aspect is the increasing of the arbiter capabilities. A more powerful arbiter would be able to form a more accurate evaluation of the two processors, and make a better decision as to their possible malfunctioning. The idea of using a programmed microprocessor seems to be an attractive way of achieving this goal. The microprocessor would not only be more powerful than the present hardwired arbiter, but would also have an increased amount of flexibility. In addition, improving the reliability of microprocessor systems through redundancy is at present being studied elsewhere in DSL, and promising results have been forthcoming [6].
An important aspect of the use of microcomputers in aviation systems is the necessity of obtaining microprocessors having military specifications. At present only one such device seems to be commercially available [7].

A final area of consideration is the possibility of replacing the H316 processors themselves with microprocessors. This would greatly reduce the size and weight of the system, factors which are important in aviation systems. The speed and capabilities of current microprocessors and semiconductor memory systems are presently being studied to determine the feasibility of their replacing the H316 processors.

4.2.2 Error recovery techniques in computer systems

A survey was carried out of the various types of error recovery techniques employed in current computer systems [2]. Some of the principal systems which use these techniques were described [2], and a bibliography containing some of the principal papers in the area of recovery techniques was also included in [2].

A significant aspect of highly reliable computer systems is their recovery capabilities. It is very important that some type of automatic recovery procedure must be implemented in these systems.

Failures in digital systems are generally classified as being solid or transient. The effects of solid faults remain in the system until removed by the repair facility, which is very often a manual operation; the effects of transient failures may disappear from the system. Recovery procedures tend to differ for transient and solid faults, and several techniques have been characterized. Generally, transient failures can be
resolved by retry or refresh methods; solid faults usually involve some type of reconfiguration to nullify their effects, and this causes some degree of system-performance degradation. A more comprehensive description of these points can be found in [2].

4.2.3 System reliability modeling

The SIRU system is one example where duplication is used to achieve higher computer system reliabilities. Such dual redundant systems depend upon a scheme for detecting a faulty module and giving control to the properly functioning unit. Various techniques exist to allow system recovery from transient errors. In order to determine the effectiveness of these error detection and recovery schemes, mathematical models of the computer systems are used.

These mathematical techniques are particularly applicable to the analysis of dual redundant computer systems such as SIRU. Such systems have a small number of states, making the mathematical manipulations feasible. The error detection and recovery methods can be modeled to first order by a simple Bernoulli process with probability of success, \( p \). Such a model incorporates a state for the error detection or recovery technique which has probability \( p \) of returning to non-degraded system performance and probability \( 1-p \) of entering some degraded (or failure) state. A refinement of such a model incorporates:

(1) the probability of successful recovery without further system degradation, i.e., the recovery effectiveness; and

(2) the probability that the system does not immediately enter a
failure state given that a fault occurs, i.e., coverage. As an example of such an analysis, consider the four-state Markov chain with states:

1 -- 2 computers functioning (full performance)
2 -- error recovery / one computer failed
3 -- 1 computer functioning / 1 computer failed (degraded performance)
4 -- both computers failed (system failure).

Analysis of this Markov model can determine values of recovery effectiveness and coverage necessary to achieve desired reliabilities for the computer system [3].

It is also possible to extend the model of the error detection or recovery process to include several states, thus realizing a more accurate analysis of the merits of such a process. This is especially useful when these processes are implemented with arbiters and/or microprocessors.

4.2.4 Signal reliability of a circuit

The fact that faults present in a circuit will not always cause the output of a circuit to be incorrect leads to interest in the probability that the output of a circuit is correct. This measure is called the signal reliability of the circuit output. Two methods for evaluating signal reliability have been devised [4]; the first uses the concept of fault-equivalence classes, and the second employs the probabilistic model for combinational circuits.

The measure has application in fault-latency studies [5]. It is known that faults occurring in a circuit may not immediately manifest themselves as errors on the output of the circuit but that, generally, some time elapses before this manifestation takes place. This interval between
the occurrence of a fault and its manifestation as an error on the output is called the "latency" of a fault, and this latency is an important parameter because it provides an indication as to how far to roll back a program when an error occurs. In addition, if the latency of a fault is large, a second fault may occur before the first is detected as an error, and this would violate the often-used single-fault assumption.

The signal-reliability measure is also useful in modeling intermittent faults in a circuit because it measures the instantaneous probability of a correct output of the circuit.

4.2.5 **Software reliability**

This research has been concerned with methodologies for testing and proving correct parallel software systems. Abstract modeling formalisms such as Petri nets, vector addition systems, Presburger logic and first order logic have been studied as possible representations of component interactions in parallel systems, and their ability to express correctness criteria has been investigated. A simple 3-process single operating system has been modeled as a Petri net, and by means of a reduction procedure several correctness questions concerning activity and history sequences in that system have been answered. In relation to the testing of parallel systems, formalisms for error description, detection and injection of faults into these systems are being investigated.

As a case study for applying the results of this research, the SIRU inertial navigation computer, a dual redundant real time system, has been chosen. A low level description of the system software in a PL360-like
language has been obtained, and is presently being used in studying the robustness of the current software, as well as techniques for fault injection and detection in real time operating systems. A high level representation in PASCAL of an operating system for the SIRU dual redundant computer that incorporates the recommendations from the above study will be obtained. Hopefully the techniques obtained in dealing with real time systems might be extended to more complex operating systems.

4.3 References


5. BIOGRAPHICAL SKETCHES OF PRINCIPAL RESEARCHERS IN CENTER FOR RELIABLE COMPUTING (CRC)

Thomas H. Bredt, Assistant Professor of Electrical Engineering (on leave), Ph.D. (CS), Stanford University, 1970, is currently performing research on parallel computer systems including operating systems and asynchronous logic networks. Studies of operating systems include resource allocation design methodologies, system structure, logical correctness, and error detection and recovery. Hazards in asynchronous systems and mathematical models of parallel systems are also being investigated.

Jacques Losq, Research Associate in Electrical Engineering, Ph.D. (EE), Stanford University, 1975, has been working on modelling and analyzing redundant digital systems. His general interest is focused on the determination of the degree and type of redundancy that provide the best trade-off between cost and performance for specific missions. This includes the study of massive, stand-by and mixed redundancy along with careful analysis of the more general survivable systems, in particular multi-micro (or mini) computers.
Edward J. McCluskey, Professor of Computer Science and Electrical Engineering, Sc.D. (EE), Massachusetts Institute of Technology, 1956, is Director of the Digital Systems Laboratory. He is widely known for his contributions to switching theory, and is the author of a standard text on the subject. His current research efforts are directed toward the area of fault-tolerant computing, including testing and diagnosis of digital circuits, redundancy schemes, and reliability modeling and evaluation. He is also involved in investigating multiprocessor systems, particularly those using microprocessors.

Roy C. Ogus, Research Associate in Electrical Engineering, Ph.D. (EE), Stanford University, 1975, has been project leader of the dual computer systems project, which is involved with the evaluation of a dual computer system used in a guidance and navigation application. His current research interests include the areas of computer reliability, fault-tolerant computing, computer architecture and microprocessor systems.

John F. Wakerly, Assistant Professor of Electrical Engineering, Ph.D. (EE), Stanford University, 1974, has been active in the development of the Digital Systems Laboratory's hardware laboratory. His research interests are in the areas of computer hardware reliability and maintainability, computer architecture, and microcomputer systems. His current projects include the design and construction of an ultra-reliable triplicated microcomputer system, a multi-microprocessor network, and a self-diagnosing minicomputer system.
6. LISTING OF PERSONNEL IN CENTER FOR RELIABLE COMPUTING

DIRECTOR: Edward J. McCluskey

ASSISTANT DIRECTORS: Thomas H. Bredt
                        John F. Wakerly
                        Roy C. Ogus

SENIOR RESEARCH STAFF: Jacques Losq
                         Andre Verdillon

STUDENT RESEARCHERS:  R. Betancourt
                        D. M. Beaudry
                        M. L. Blount
                        D. Davies
                        M. Hadidi
                        A. F. Hunter
                        S. G. Kolupaev
                        P. J. LeVine
                        D. J. Lu
                        J. R. McClure
                        K. P. Parker
                        J. V. Phillips
                        J. Savir
                        J. J. Shedletsky
                        P. A. Thompson
                        W. A. Wallach
7. PREVIOUS RESEARCHERS IN CRC

Previous researchers in the Center for Reliable Computing
together with their current locations and Ph.D. dissertation titles
are listed below.

Thomas H. Bredt - Hewlett-Packard Computer Division, Cupertino, CA.
Ph.D. 1970, "Control of Parallel Processes."

Frederick W. Clegg - Hewlett-Packard Computer Division, Cupertino, CA.

Donald J. Chesarek - IBM Systems Development Division, Los Gatos, CA.
Ph.D. 1972, "Fault Detecting Experiments for Sequential Machines."

Donald P. Siewiorek - Carnegie-Mellon University, Departments of
Computer Science and Electrical Engineering, Pittsburgh, Pa.
Ph.D. 1972, "Fault-Tolerant Computers Using Self-Diagnosis and
Hybrid Redundancy."

Raymond T. Boute - Bell Telephone Manufacturing Co., Antwerp, Belgium.
Ph.D. 1973, "Faults in Sequential Machines: Algebraic Properties
and Defection Methods."

Jacob A. Abraham - University of Illinois, Department of Electrical
Engineering, Champaign, IL.
Ph.D. 1974, "Reliability Analysis of Digital Systems Protected
by Massive Redundancy."

Hajime Mitarai - Canon Industries, Tokyo, Japan.
Ph.D. 1974, "The Use of Semiconductor Read-Only Memory for Logic."

John F. Wakerly - Digital Systems Laboratory, Stanford University,
Stanford, CA.

David T. Wang - IBM Systems Products Division, Endicott, NY.
Ph.D. 1974, "An Algorithm for the Generation of Test Sets for
Combinational Logic Networks."

Jacques Losq - Digital Systems Laboratory, Stanford University, Stanford, CA.

Kenyon C. Y. Mei - Hewlett-Packard Computer Division, Cupertino, CA.
Ph.D. 1975, "Dominance Relations of Stuck-at and Bridging
Faults in Logic Networks."

Francisco J. O. Dias - University of Sao Paulo, Sao Paulo, Brazil. Ph.D. 1975, "Multiple Fault Analysis in Combinational Logic Circuits."

In addition, the following scholars have visited the CRC and performed research in the listed areas.

Pawel N. Kentopf - Polish Academy of Sciences, Warsaw, Poland. Design of universal logic modules.

Yoshihiro Tohma - Tokyo Institute of Technology, Tokyo, Japan. Redundancy techniques for constructing ultra-reliable systems.


Andre Verdillon - University of Grenoble, Grenoble, France. Fault properties.
8. CENTER FOR RELIABLE COMPUTING BIBLIOGRAPHY

8.1 JOURNAL PAPERS

Prepared Under Previous NSF Grants


* NSF grant GJ 165
† NSF grant GJ 27527
‡ NSF grant GJ 40286


---

† NSF grant GJ 27527
ψ NSF grant GJ 40286
Ω NSF grant GJ 43322
8.2 CONFERENCE PAPERS

Prepared Under Previous NSF Grants


---

* NSF grant GJ 165
+ NSF grant GJ 27527
Ψ NSF grant GJ 40286


---

† NSF grant GJ 27527
ψ NSF grant GJ 40286
⊗ NSF grant GJ 43322


---

♀ NSF grant GJ 40286

♀ NSF grant GJ 27527
8.3 TECHNICAL PAPERS

Prepared under Previous NSF Grants


---

* NSF grant GJ 165
† NSF grant GJ 27527


- - - - - -

✦ NSF grant GJ 27527
ψ NSF grant GJ 40286

8.4 TECHNICAL NOTES

Prepared under Previous NSF Grants


---

* NSF grant GJ 165
† NSF grant GJ 27527
‡ NSF grant GJ 40286


---

† NSF grant GJ 27527
¶ NSF grant GJ 40286
