ECONOMICS OF DEPENDABILITY FOR COMPUTER SYSTEMS

Herbert Hecht
SoHaR Incorporated
5731 W. Slauson Ave.,
Culver City, California
Voice: 310.338.0990 xt 110 Fax 310.338.0999
herb@sohar.com

ABSTRACT

The paper introduces the computer systems developer and integrator to an array of economic trade-off techniques for deciding on:

- How to set overall reliability or availability goals for a computer system
- Where to employ redundancy (and the best configuration)
- When to investigate prognostic based maintenance

The emphasis is on balancing the cost of dependability improvement against the avoidance of cost of failure.

INTRODUCTION

Computers and other digital components are increasingly used in critical applications in which failure can cause large economic loss. The title of this conference highlights the need for highly dependable computing and presentations describe methods of achieving this. In all this the need for economic evaluation should not be overlooked. Have you ever asked how a reliability or availability goal was established? Have you wondered whether active or passive (static) redundancy is more suitable for your application? And are you looking for a test port that can best indicate whether a circuit board is nearing the end of its useful life? If these are subjects that

hold your interest you will find this paper useful. Not all of these problems will be solved, but you may find pointers towards the right answer for your application.

The first part of this paper establishes guidelines for setting reliability or availability goals for a system. When the initial design fails to meet the established goals one possible solution is to improve the dependability by making a critical part or the entire system redundant. The second section describes the economic consequences of several forms of redundancy. A recurrent problem is to reduce the cost of maintenance, and this is addressed by examining prognostic techniques that reduce the need for unscheduled maintenance. The concluding section urges the reader to make the principles of economic evaluation a part of the design routine.

SETTING RELIABILITY AND AVAILABILITY GOALS

The term *system* is used for everything within the bounds of the current design decisions. *Components* are parts of the system and can be hardware, software, or even skinware (operators). The system furnishes *services* that may include control of the *plant.*

The goal of reliability efforts is to reduce the occurrence of a loss, *L,* that may arise from serice or plant shut-down, equipment damage, or personal injury (1,2). The loss may also include contingent liabilities, e. g., claims by a customer for non-delivery of contracted products. Design decisions for possible avoidance of this loss must be made well in advance of the time of the occurrence of the loss but that does not prevent assigning a reasonable value to *L.* The decision to reduce *L* typically involves an increase in system cost (for additional reliability) that should be traded off against a reduction in the expected cost of failure

$$\mathbf{E}[V_f] = f \times L \qquad \text{Eq. 1}$$

The failure probability, *f*, can be reduced by improving the quality or robustness of the product, but that reliability improvement will increase the cost by an amount ΔV_r. A given reliability improvement will be economically desirable as long as

$$-\Delta \mathbf{E}[V_f] \geq \Delta V_r \qquad \text{Eq. 2}$$

The equation says that a cost increment for reliability will be considered beneficial if it is not more than the decrement in the cost of failure. The relationship between expenditures for reliability and avoidance of cost of failure can also be explored by postulating that both cost elements must ultimately be borne by the user of the service or the owner of the plant. If the user's expenditures affected by a dependability decision are designated as V_u, then

$$V_u = V_r + \mathbf{E}[V_f] \qquad \text{Eq. 3}$$

It is understood that improvement decisions always involve increments and the Δ symbols have therefore been dropped. The user's total cost and its two constituents are graphed in Figure 1.

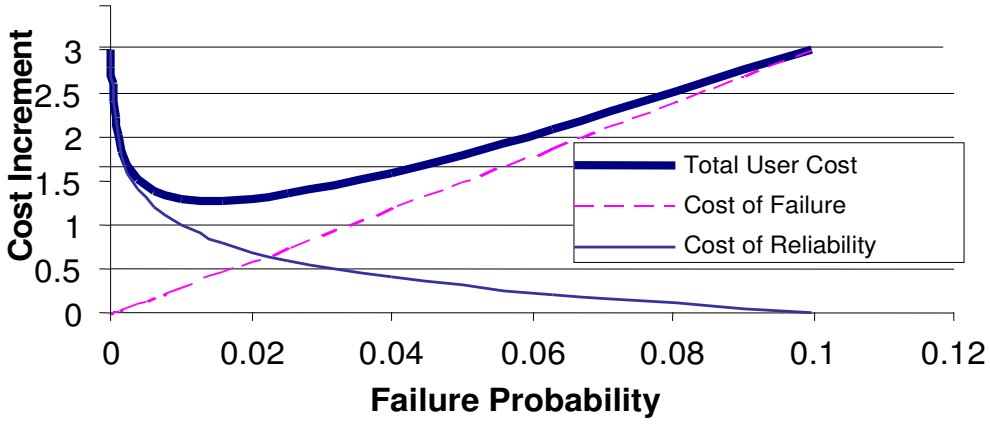

FIGURE 1. COST VERSUS RISK

The values shown for failure probability and cost increment are arbitrary. The cost of failure rises linearly with failure probability as is evident from Equation 1. The non-linear relation

between the cost of reliability and failure reduction will be discussed in next section. A decision maker using this criterion will first institute the improvements that have the shallowest slope (right end of the cost of reliability curve in Figure 1) and then consider those with successively steeper slope. Thus, at the low failure probability side of the graph only the costliest improvement opportunities remain, resulting in the very steep slope. Zero failure probability can be approached but not reached.

The most significant aspect of this graph is that the total user cost has a minimum, and the failure probability at which the minimum is reached represents the optimum reliability under economic motivation. The minimum is at the point at which the slope of the cost of reliability curve has the exact negative value of the slope of the cost of failure line. Since the latter is a constant, it is usually not too difficult to locate the minimum. Also, since the minimum is rather shallow, the economic results are not sensitive to small errors in the quantities used in Figure 1.

The curve has some interesting implications for selecting the quality and reliability of computer components. Consider an office computer for a small carpentry shop. If it fails it is easily replaced or repaired. The loss associated with the failure is small and no special reliability requirements are specified for the computer. If the same computing function is required for a supermarket the cost for replacement or repair will still be small but the market will have to shut down or be on manual cash register operation at a cost of several thousand dollars. Finally, the computer may control traffic signals in a metropolitan area. Replacement will disable all signals, require dispatch of traffic officers to important intersections and call-in of specialists for re-initialization, probably at a cost of several tens of thousands of dollars. As the loss increases in these examples, so does the slope of the cost of failure line. And as the slope increases, the minimum shifts to the left and results in a higher total user cost. The increase in the loss and cost

of failure makes it desirable to use higher quality computers and redundancy, thereby reducing the failure probability and the expected cost of failure (Equation. 1).

Most of the above applies to availability as well as reliability. The decision maker's primary tool for reliability improvement is the reduction of the failure probability. Availability can also be improved by reduction of the downtime. The latter can be achieved by logistics (placement of spares, technician availability, etc.) as well as by prognostic measures that are discussed later in this paper. Therefore economically optimum improvement of availability will usually require a number of trade-offs between these alternatives.

REDUNDANCY

Structural parts can be made more reliable by selection of stronger materials or by increasing sensitive dimensions; over-design (derating) can be used to reduce the failure probability of many components, including some discrete electronic parts. But for most computing components a demonstrable increase in reliability can only be achieved by redundancy. Therefore a few forms of redundancy and their economic implications are now presented.

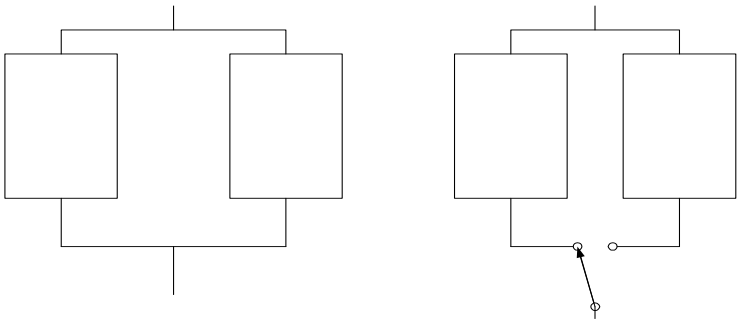

FIGURE 1. PASSIVE AND ACTIVE REDUNDANCY

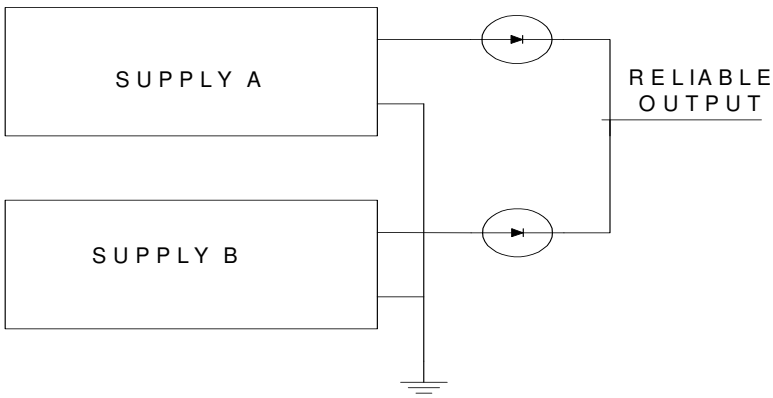

FIGURE 3. DIODE ISOLATION-REDUNDANT POWER SUPPLIES

Passive redundancy, shown on the left in Figure 2, usually needs isolators to prevent short circuit failures in one element from disabling the other and to avoid echo problems. Monitors to indicate that both units are operative are frequently provided. Active redundancy requires switching and the switch actuation by an operator or automatic detection mechanism. The figure shows switching at the output of the elements but power switching may be substituted for or used together with output switching. The reduced power consumption, the lower failure rate achievable for unpowered elements, and the greater application flexibility frequently favor switched redundancy.

A typical use of passive redundancy is in power converters with DC output where diode isolation can be used as shown in Figure 3. Each of the supplies must be capable of carrying the entire load. At times a more economical arrangement is to employ three supplies, each rated for one-half of the required load. After a first failure each of the two remaining supplies can still furnish enough power for the entire load.

Triple modular redundancy (TMR) with voting is another example of the passive type(3). The block diagram shown in Figure 4 is deceptively simple but significant problems may be encountered in the implementation. One of these is that the communication between the processors and the voter is frequently much slower than the transmission rate of the internal bus.

This dictates that only significant outputs be voted. Also, unless the processors run on a common fault-tolerant clock (a non-trivial design problem), the voter must tolerate the delay for receipt of

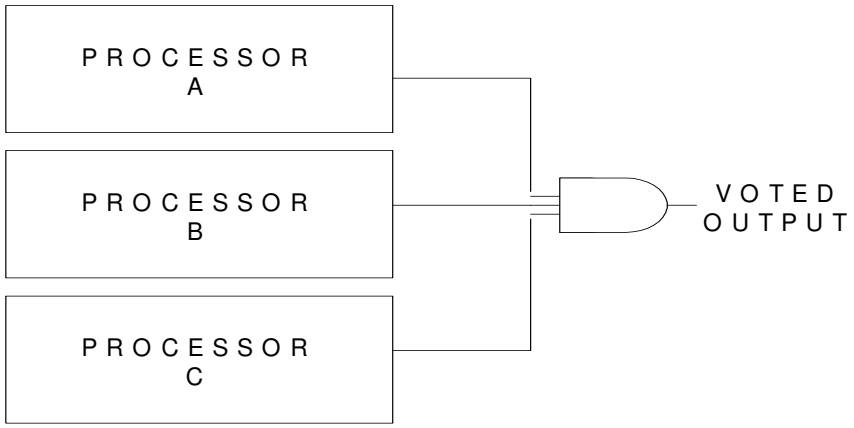

FIGURE 4. TRIPLE MODULAR REDUNDANCY WITH VOTING

the result from the slowest processor. TMR offers very high fault coverage without the need for anomaly detectors. The reliability can be assessed by two criteria: maintenance frequency (failure of the first processor) for which the failure probability is:

$$F_1 = 3 \times f \qquad \text{Eq. 4}$$

and plant outage (second failure) for which the failure probability is:

$$F_2 = 3 \times (1-f) \times f^2 \qquad \text{Eq. 5.}$$

In both equations the failure probability of a single processor is denoted by *f*. The cost of TMR is more than three times that of the single processor. Completely assembled TMR configurations are offered by some vendors. Because of the higher initial and maintenance cost and the throughput reduction TMR may be reserved for the most critical applications, including supervisory functions.

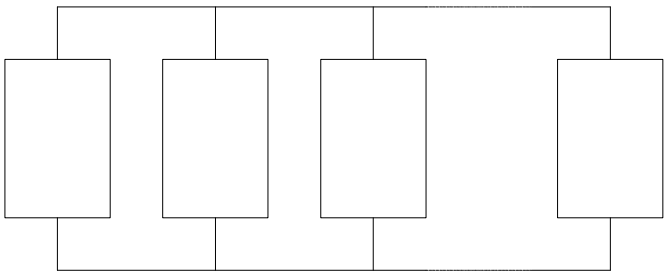

FIGURE 5. MULTIPLE REDUNDANCY

Both passive and active redundancy can be used with multiple copies as shown in Figure 5. The cost of reliability will vary directly with the number of redundant units. The failure probability of the redundant structure is given by:

$$F = f^n \qquad \text{Eq. 6}$$

where F denotes the failure probability of the redundant structure, f the failure probability of the individual elements and n the number of elements. The reduction in failure probability gained by adding one more element is:

$$\Delta F = f^n - f^{n+1} \qquad \text{Eq. 7}$$

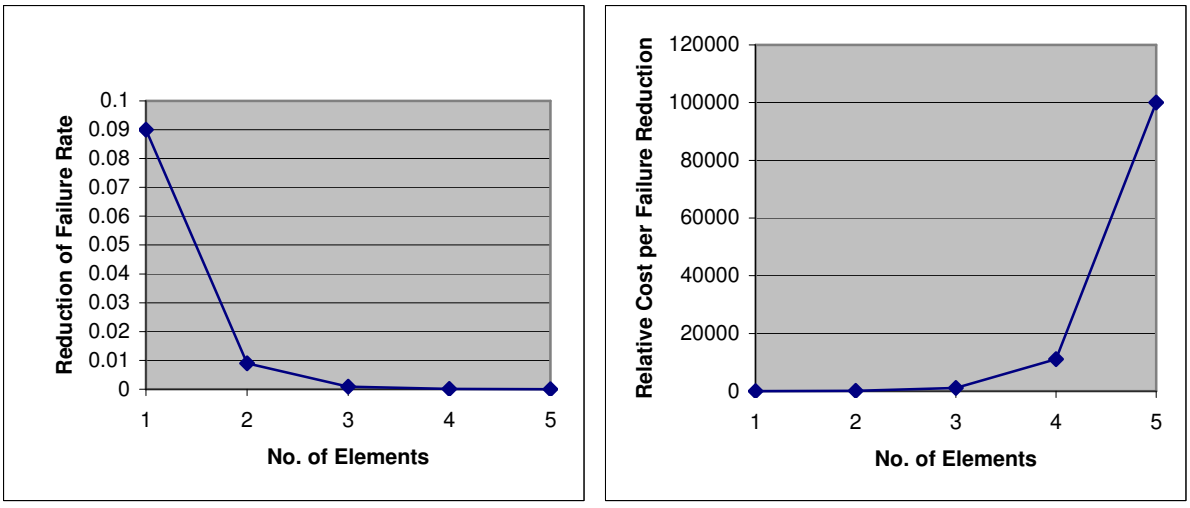

FIGURE 6. ADDING REDUNDANT ELEMENTS

Figure 6 shows the effect of adding one element to the number shown on the abscissa for $f = 0.1$. The failure rate reduction is graphed in the left part of the figure and the relative cost of failure rate reduction is shown in the right part. The very limited reliability benefits and very steeply rising cost of implementing high degrees of redundancy are clearly seen. This relation accounts for the non-linear cost of reliability curve in Figure 1.

The increased cost to achieve a higher reliability increment by redundancy can be avoided if other reliability improvement techniques can be brought to bear, e. g., derating or stress screening. Because each improvement method operates along its own cost curve it is economically desirable to use multiple failure reduction techniques to achieve a reliability objective(4). Thus, instead of moving to the left along a single cost curve in Figure 1, the required reliability increment is obtained from the shallow (right) end of another cost curve.

Whereas there are thus economic disadvantages to the use of higher degrees of redundancy, the use of fractional redundancy is frequently quite advantageous. An example of fractional redundancy is the use of half-capacity power supplies that was discussed in connection with Figure 3. Another common example is a single spare computer that can replace of any one of the active computers in a network. A limitation on fractional redundancy is that the failure probability of the individual elements must be low to preclude the possibility of a second failure before redundancy is restored.

PROGNOSTICS TO REDUCE THE COST OF MAINTENANCE

In many applications, maintenance contributes much more to the cost of ownership than the original equipment acquisition. For mechanical and electromechanical equipment the major practices were scheduled and unscheduled maintenance. Unscheduled maintenance incurs a cost

that is many times that of scheduled maintenance because the right skill levels, test equipment and spare parts are unlikely to be available when and where needed. If the maintenance is performed at a fixed schedule it has to be timed well ahead of the event that might cause failure in the majority of the population (exhaustion of lubricant, abrasion, corrosion, etc.), and this can entail considerable waste. To avoid this, the practice shifted to reliability centered maintenance (RCM)(5) in which the scheduling is based on observation of precursors of the failure causing events, such as acoustics or bearing temperature rise in rotating machinery. As long as the decision levels for maintenance were empirically set they, too, had to err on the conservative side. Through the concepts of prognostics and life consumption monitoring (LCM)(6) the empiricism is being replaced by studies of the service time remaining before an irreversible progression to failure sets in. Because the assumption of exponential failure probability showed no benefit for replacement other than after failure. Also, the high rate of innovation in the computer field caused hardware to be replaced before wear-out became significant. Thus, RCM and LCM were not employed in the computer field.

But the increasing use of digital components in critical applications and the scarcity (and correspondingly high cost) of competent repair facilities puts new emphasis on avoiding unscheduled maintenance even though the approach must be different than that followed with mechanical components. Some of the reasons are shown in Figure 7 and essential differences in factors that affect the implementation of prognostics are analyzed in Table I. The failure modes for mechanical devices have been researched for decades or centuries and in many cases eliminated. Those that remain must be dealt with by maintenance and prognostics are an effective indicator of the need for such action.

TABLE I. FACTORS THAT AFFECT THE IMPLEMENTATION OF PROGNOSTICS

Factor	Mechanical	Electronic
No. of failure modes	Few	Many
Typical MTBF	1,000 hrs	>100,000 hrs
Failure mechanisms	Abrasion, fatigue, corrosion	Thermal fatigue, corrosion, threshold shift, response time, electromigration, etc.
Time to failure models	Established	Research required
Size ratio: sensor/part	Small	Can be very large

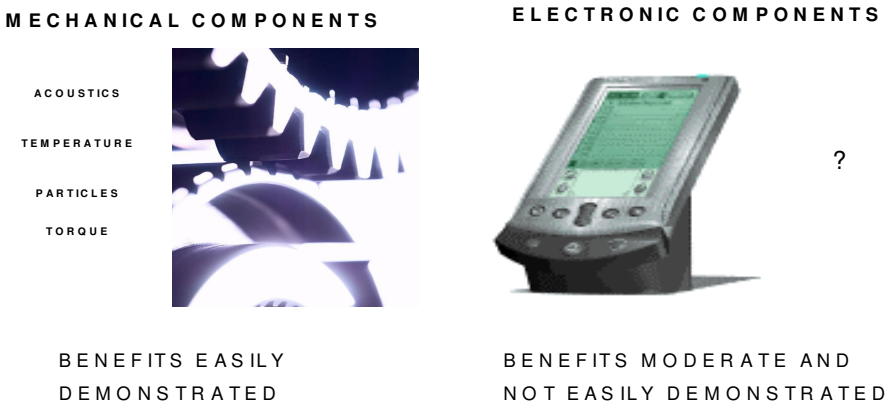

FIGURE 7 PROGNOSTICS FOR ELECTRONICS IS NOT EASY

By contrast, the design of electronic parts is constantly undergoing changes, prevention of failure modes is an active research area, and where there is little prospect of reducing high failure rates the device type may be replaced, e. g., high capacitance ceramic capacitors replacing tantalum electrolytic ones. Where wear-out is observable in electronic components its time scale is usually twenty to 100 times longer than that of mechanical components. This together with the factors listed in Table I mean that prognostic techniques used in the mechanical area, addressing individual parts and their failure modes, are not likely to be economically beneficial. However, broad spectrum prognostics that cover a number of parts and part types and multiple failure modes may deserve to be examined, particularly when they can be installed at little cost.

The basic economic relation for the benefit, B, obtainable from use of prognostics is

$$B = (M_u - M_p) \times N - C \qquad \text{Eq. 8}$$

where M_p and M_u are the maintenance cost (including charges for non-availability of the affected system), N is the total number of successful prognostic events during a given time frame, and C is the cost of providing the prognostics. The M_u, M_p *and* C terms are very application dependent and will be discussed later. But the N factor can be further expanded and this will give us some clues as to where prognostics might be applicable. Thus,

$$N = \lambda \times t_H \times e \qquad \text{Eq. 9}$$

where λ is the failure rate for the part and failure mode, t_H is the time horizon (the interval of failure events to be considered, with discounting for future failures), and e is the effectiveness of the prognostic measures. This decomposition together with some assumptions is used in Table II to find the economically most desirable assembly level at which the prognostic measures should be applied. The $\lambda \times t_H$ product at the LRU (line replaceable unit) level has been arbitrarily set to unity. A typical LRU (or black box) is assumed to consist of five PCBs (printed circuit boards), with each furnishing one-fifth of the failure probability. Each PCB is assumed to contain four high-failure rate parts (a total of twenty for the LRU), thus fixing the part failure probability at 0.05.

TABLE II. ASSEMBLY LEVEL EFFECTS

Level	$\lambda \times t_H$	e	N
LRU	1	0.1	0.1
PCB	0.2	0.3	0.06
Part	0.05	0.9	0.045

The *e* factor at the LRU level has been set at 0.1, a fairly pessimistic assumption, and thus *N* at the LRU level is also 0.1. Because the individual circuit boards may contain a smaller number of part types (thus permitting more specific prognostics) the *e* factor has been increased to 0.3. At the part level an extremely optimistic assumption allows the effectiveness to be increased to 0.9. In spite of the conservative assumptions at the LRU level and increasingly optimistic ones in the lower rows, it is seen that a greater number of failures will be covered at the highest assembly level. Thus, as a general guideline, the development of prognostics should start at the LRU level.

Equation 8 shows that the benefit is proportional to the $(M_u - M_p)$ term, the cost difference between unscheduled and prognostic initiated maintenance, and thus the most suitable applications for prognostics are where this difference is large. Therefore computers installed in inaccessible locations are prime targets. So are specialty components that are difficult to obtain and for which therefore the expected replacement time may be large, and components that require special calibration and certification before the service can be resumed. The benefit will be diminished by the cost term, *C*, and this should therefore be kept low. There may be some opportunities that carry zero or negligible cost: the dissipation in power converters, the degradation in output of some functions, and the rate of memory errors in digital devices. Each one of these will be briefly described.

The vast majority of power conversion components make use of semiconductor switches(7). If these turn on and off instantaneously, producing a true square wave output, there is no power loss at all. Practical switches have a brief gradual transition between on and off states, sometimes called *rounding*, that represents an equivalent resistance in series with the switch and causes dissipation. Aging will cause more rounding and the resulting increasing power dissipation at the

switch produces a temperature rise that causes further deviation from ideal switching. Measuring the difference between input and output power is a direct way of capturing the dissipation and also permits compensation for the load state (normally dissipation is less at light loads than at heavy ones). Input and output voltage are routinely recorded in most applications and in many, so is at least the output current. Where additional current measurement is required, small and inexpensive current sensors can be incorporated into the connectors to the power converter.

Many analog circuits that are associated with computers in control applications employ feedback to keep the controlled variable (pump speed, belt tension, etc.) constant even as the driving power of the controller deteriorates. Capturing the feedback voltage (or current) provides a clue to possible near term failure of the controller or the controlled device. Access to the feedback may require modification of the controller, but the cost for this is usually quite small compared to the potential savings due to preventive maintenance.

Semiconductor memories are inherently subject to occasional errors, usually called single event upsets or SEUs, that convert a 1 into a 0 or vice versa. Error detecting and correcting code (EDAC) to deal with SEUs is a standard feature in many commercial devices (8). The rate at which corrections are being made is dictated by the radiation environment seen by the memory chip and follows a Poison distribution, the parameter of which is constant at a given location. Over time the error rate increases due to micro-structural changes in the silicon. Persistent changes can be very sensitively detected. Table III shows the statistical significance of observing consecutive levels of errors during an interval for which 10 errors are expected. Entering this table for 12 SEUs shows that the probability of two consecutive intervals of 12 or more errors due to chance events is only 0.0434. When this is observed there is over 95% certainty that

degradation has occurred. If three consecutive intervals of 12 or more SEUs are observed, the probability that this is due to degradation is over 99%. More sophisticated discrimination techniques, based on the rate of increase in errors observed in successive time intervals, may also be employed.

TABLE III. SIGNIFICANCE LEVEL FOR CONSECUTIVE EVENTS

No. of	Consecutive Events = or > N			
SEUs (N)	1	2	3	4
10	0.4170	0.1739	0.0725	0.0302
11	0.3032	0.0919	0.0279	0.0085
12	0.2084	0.0434	0.0091	0.0019
13	0.1355	0.0184	0.0025	0.0003
14	0.0835	0.0070	0.0006	0.0000
15	0.0487	0.0024	0.0001	0.0000

Replacement based on this prognostic may not only prevent a catastrophic memory failure but may also be an indicator of general ageing of other semiconductor devices. Again, the memory error correction rate is either already being logged or can be made available at little cost and it is therefore a desirable element in the prognostic tool kit.

CONCLUSION

Computers and other digital components are increasingly used in critical applications in which failure can cause large economic loss. You have listened to presentations that describe methods

of achieving the required high dependability. And in this paper we hope to have motivated you to look at the economic effects of your design decisions. Three aspects of economics of dependability have been discussed here:

- Setting economic goals for dependability improvement – when should we stop
- Comparing benefits of diverse dependability improvement methods – and learning where mixing of methods is advantageous
- Including maintenance approaches in the economic analysis – exploring prognostics for converting the need for unscheduled maintenance into schedulable maintenance

Redundancy is the most universally available reliability improvement technique for electronic components and the cost factors associated with redundancy were investigated in the second section. The economic benefits of fractional redundancy were pointed out.

Because the cost of maintenance, including service or plant outage during repair, are becoming an increasingly important component of the cost of ownership of electronic components, techniques for replacing maintenance after failure (unscheduled maintenance) with schedulable maintenance prior to failure were described in the last section. Prognostics for electronic components are a recent concept and are inherently different from those employed for mechanical devices. An economically motivated approach to identifying opportunities for prognostics for electronics has been provided, and three examples for cost-effective implementation of prognostic techniques were described.

REFERENCES

(1) Jutila, S. T., *Economics of Reliability: A Key Point for System Users,* University of Toronto, June 1971, AD-D875339
(2) Hecht, H., *Economics of Reliability for Spacecraft Computers,* The Aerospace Corporation, October 1971, AD0736760
(3) Sheble, N., *More is Always Better when it's Critical,* ISA InTech, 1 October 2003

(4) Larsen, R. S., *A Comment on Machine Reliability and Availability,* Stanford Linear Accelerator Center, July 2002, LCC-0084
(5) Moubray, J.. *Introduction to Reliability-centered Maintenance,* Maintenance Resources, Terre Haute IN, 2003
(6) NATO Research and Technology Organization, *Recommended Practices for Monitoring Gas Turbine Engine Life Consumption,* April 2000, AD A380949
(7) Gottlieb, I. M., *Power Supplies, Switching Regulators, Inverters and Converters,* TAB Books, Blue Ridge Summit PA, 1994, ISBN: 0830644059
(8) *Simple Error Detection and Correction Codes,* http://www.ee.washington.edu/conselec/CE/kuhn/cdmulti/95x7/error.htm