RELIABILITY AND MAINTAINABILITY ASSESSMENT FACTORS FOR RELIABLE FAULT-TOLERANT SYSTEMS

Salvatore J. Bavuso
NASA Langley Research Center
Hampton, Virginia

First Annual NASA Aircraft Controls Workshop
NASA Langley Research Center
Hampton, Virginia
October 25-27, 1983

ABSTRACT

A long-term goal of the NASA Langley Research Center is the development of a reliability assessment methodology of sufficient power to enable the credible comparison of the stochastic attributes of one ultrareliable system design against others. This methodology, developed over a 10-year period, is a combined analytic and simulative technique. An analytic component is the Computer-Aided Reliability Estimation capability, third generation, or simply CARE III. A simulative component is the Gate Logic Software Simulator capability, or GLOSS.

This paper focuses on the numerous factors that potentially have a degrading effect on system reliability and the ways in which these factors that are peculiar to highly reliable fault-tolerant systems are accounted for in credible reliability assessments. Also presented are the modeling difficulties that result from their inclusion and the ways in which CARE III and GLOSS mitigate the intractability of the heretofore unworkable mathematics.

RELIABILITY ASSESSMENT GOAL

A long-term goal of the NASA Langley Research Center is the development of a reliability assessment methodology of sufficient power to enable the credible comparison of the stochastic attributes of one ultrareliable system design against others (fig. 1). This methodology, developed over a 10-year period, is a combined analytic and simulative technique.

OBJECTIVE: DEVELOP A CAPABILITY TO ASSESS THE RELIABILITY OF ANY FAULT-TOLERANT DIGITAL COMPUTER-BASED SYSTEM, INCLUDING THE SYSTEM EFFECTS OF SOFTWARE

Figure 1

COMBINED ANALYTIC SIMULATIVE METHODOLOGY

The methodology for performing reliability assessments is based on the utilization of an analytic model that accounts for the long time constants of hardware and/or software failures and a separate analytic model that tracks the short time constants of system fault-handling mechanisms. These models, which are embodied in computer programs, in conjunction with a simulative model, make possible the reliability assessment of large, practical fault-tolerant systems (fig. 2).

The CARE III computer program (codeveloped by the Raytheon Company and the Langley Research Center (ref. 1)) provides an analytic capability. The GLOSS is a simulative capability that provides CARE III with stochastic fault-handling data. The GLOSS concept was demonstrated by application to the CPU of an avionic processor. A generalized GLOSS that provides a user-friendly hardware description language interface is currently being developed. The GLOSS was codeveloped by the Bendix Corporation and the Langley Research Center (refs. 2, 3).

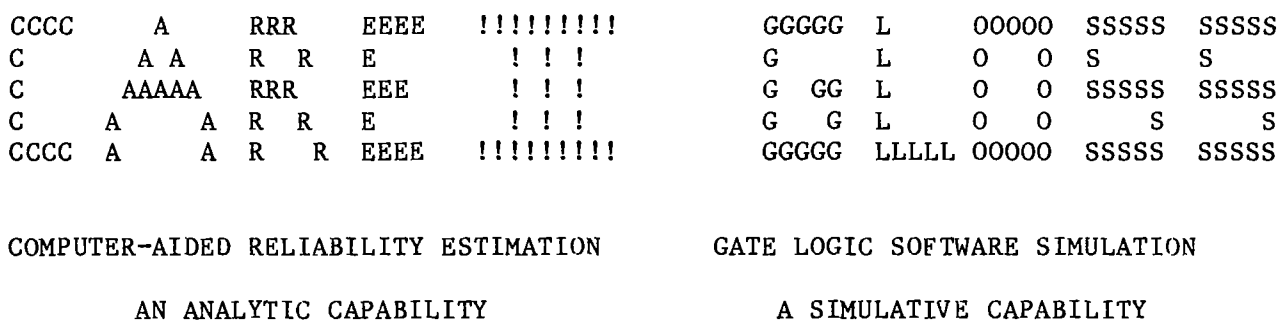


Figure 2

TECHNOLOGICAL DEVELOPMENT LEADING TO CARE III

The motivation for developing the combined analytic simulative methodology dates back to 1973. The long-term development of CARE III is depicted in figure 3. State-of-the-art reliability evaluators were typical of CARE, a computer program developed by the Jet Propulsion Laboratory, and TASRA (Tabular System Reliability Analysis), developed by Battelle Memorial Laboratories. The Raytheon Company and Langley jointly developed the CARE II, which provided a superset CARE model with an extensive fault-handling model. Langley was also involved in the development of CAST (Combined Analytic Simulative Technique), which provides the current Langley modeling concept. CAST was developed by the Ultra Systems, Inc. CARSRA (Computer-Aided Redundant System Reliability Analysis) was a spin-off from the Boeing ARCS (Advanced Reconfigurable Computer System) study. Langley has also been involved in numerous technology development studies, some of which are depicted in the figure. This long-term involvement has culminated in the development of the CARE III.

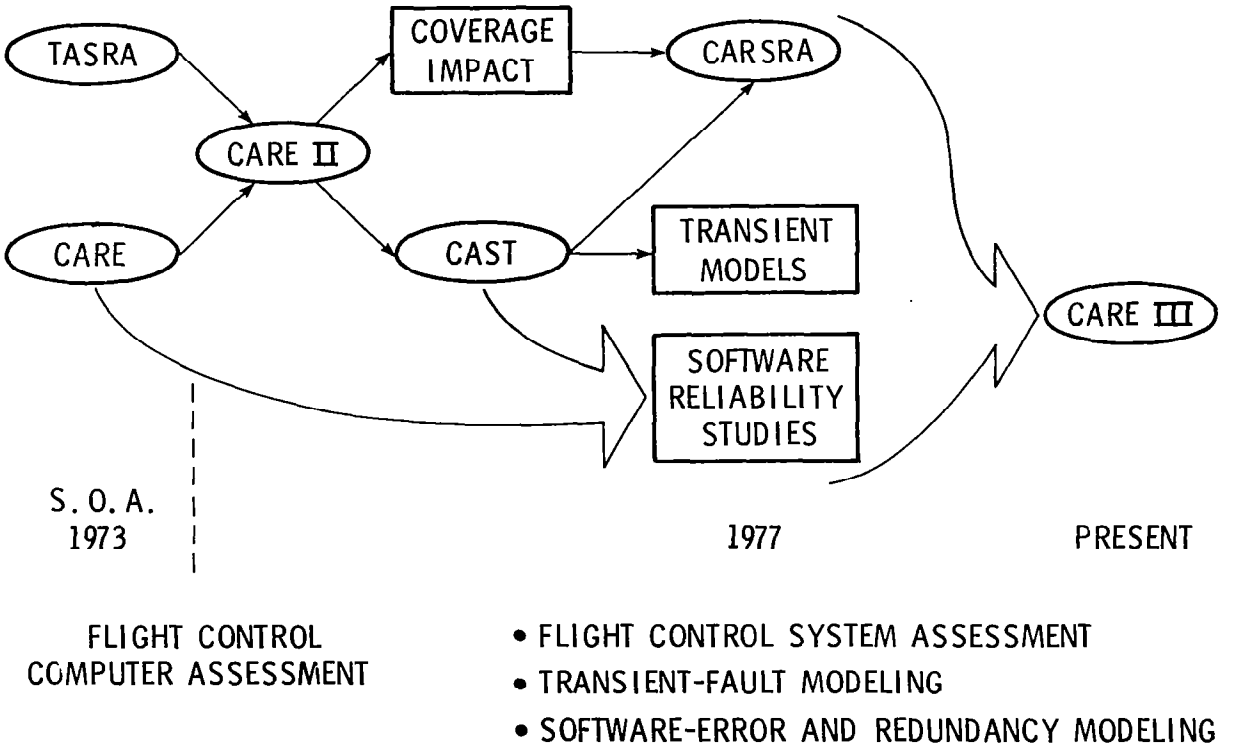


Figure 3

PROFOUND OBSERVATIONS

On our way toward developing the specifications for CARE III, we found that for ultrareliable systems certain factors that previously were of little interest to the reliability analyst now potentially have a significant effect (fig. 4). This is particularly true of systems with a flight crucial probability of failure of less than 10^{-9} in a 1- to 10-hour mission. An example of this observation is the latent (undetected) fault. We also realized that even complex assessment capabilities must be user-friendly; this is always a difficult task for complex capabilities.

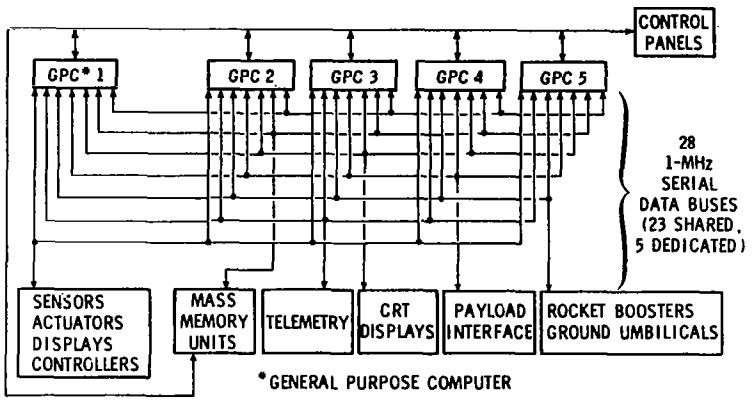
PROBLEMS:

1. EVERYTHING IMPORTANT WHEN $P_F < 10^{-1}$
2. PROGRAM VERSATILITY vs CONVENIENCE AND EFFICIENCY

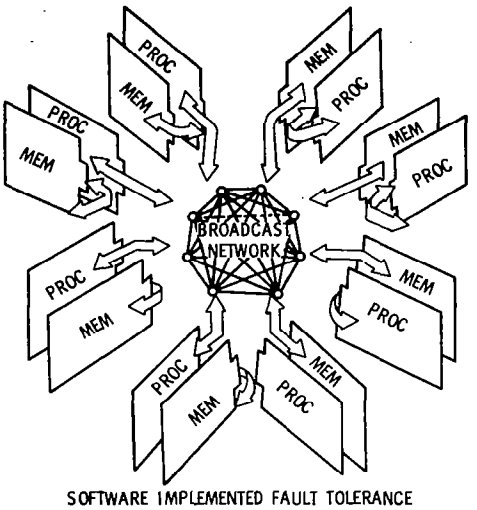
Figure 4

HIGHLY RELIABLE FAULT-TOLERANT SYSTEMS TO WHICH CARE III IS APPLICABLE

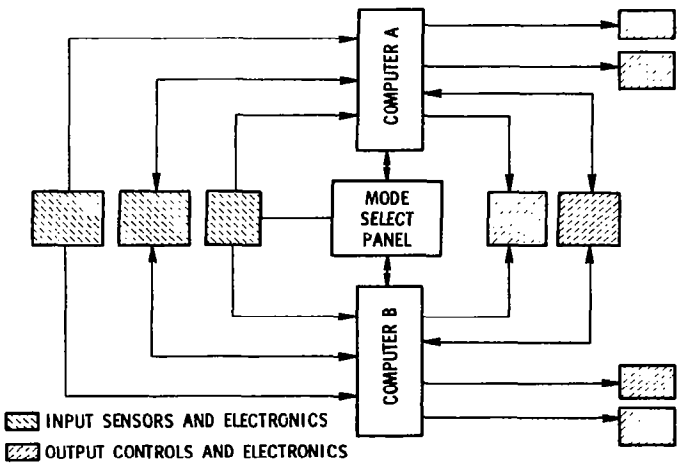
The class of fault-tolerant systems of most interest currently utilizes off-the-shelf processors or computers (fig. 5). These systems rely heavily on the ability of the processors to detect system faults/errors, to identify the fault/error to the smallest reconfigurable unit, and to effect recovery.



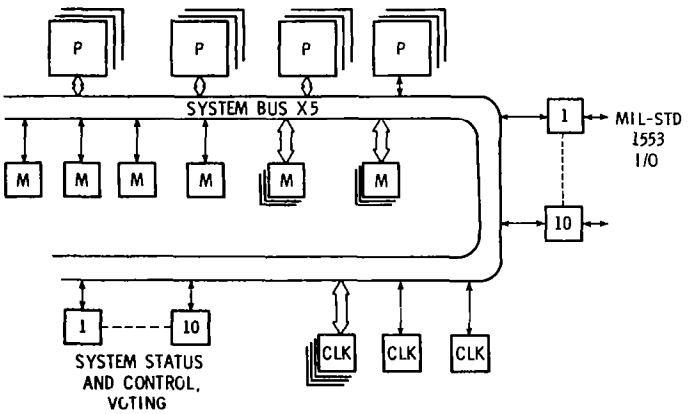
SPACE SHUTTLE SYSTEM



SOFTWARE IMPLEMENTED FAULT TOLERANCE (SIFT)



DC-9-80 DIGITAL FLIGHT GUIDANCE SYSTEM



FAULT-TOLERANT MULTIPROCESSOR

Figure 5

COVERAGE - A MAJOR RELIABILITY DRIVER

In ultrareliable fault-tolerant systems, the inability of a system to achieve perfect fault/error handling is often the dominant cause of system failure (fig. 6). The major contributor of diminished fault/error handling is the latent fault/error. The long-term (latent) accumulation of faults/errors poses a severe threat to the system's ability to detect and mask out anomalies. The modeling of fault/error handling adds a tremendous amount of additional complexity to the reliability assessment task.

THE PREDOMINANT CAUSE OF FAILURE IN ULTRARELIABLE DIGITAL SYSTEMS HAS BEEN SHOWN TO BE ATTRIBUTED TO FACTORS OTHER THAN HARDWARE SPARES DEPLETION

COVERAGE - MEASURE OF SYSTEM'S ABILITY TO HANDLE FAULTS $\Longrightarrow$ SYSTEM

- FAULT DETECTION
- FAULT ISOLATION
- RECONFIGURATION AND RECOVERY

UNDETECTED FAULT - LATENT FAULT

Figure 6

DELINEATION OF HARDWARE AND SOFTWARE FAILURE AND ERROR MODELS

The increased complexity is indicated by the number of additional fault/error models that now must be considered. The increase in the number of fault/error models that must be accounted for is largely attributed to use of the digital computer (which possess extensive memory capability) and very high system reliability requirements. An extensive memory capability is a two-edged sword in that not only are computational capability and flexibility enhanced, but the likelihood of latent faults and errors occurring is also increased. Ultrareliability necessitates the consideration of design errors, which previously were considered to be insignificant. Each branch in the trees in figure 7 represents a fault/error model. Faults are hardware generated, whereas errors are caused by a fault or by software design anomalies. Either one may be permanent or may appear to be transient or intermittent. The common piece-parts reliability analysis is shown as a permanent random hardware failure.

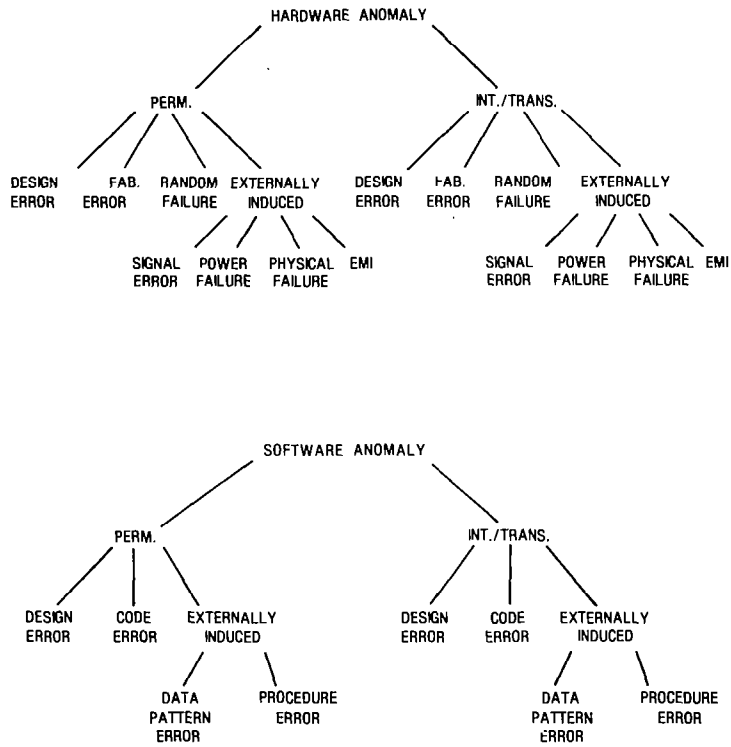


Figure 7

FUNCTIONAL DEPENDENCY TREE FOR A NEAR-FUTURE PROPOSED FLIGHT CONTROL SYSTEM

Ultrareliable fault-tolerant systems increase the system reliability by employing redundancy, which further compounds the modeling task. A typical proposed advanced reconfigurable flight control system would utilize triple voting of units for the sensors, processor memories, and actuator electronics (fig. 8). In this example, the number of units increased from 22 for a nonredundant system to 64 for the fault-tolerant architecture.

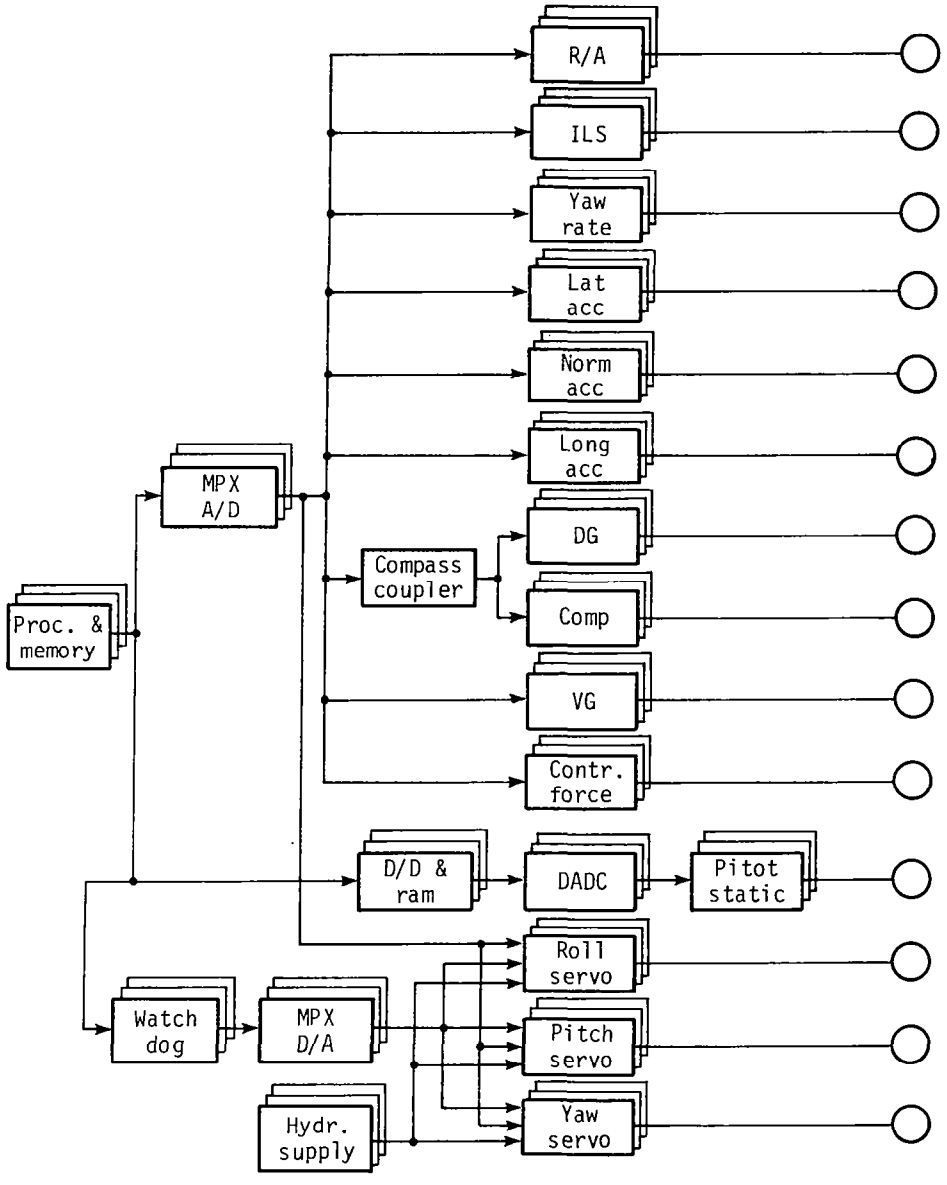
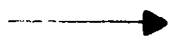


Figure 8

POSSIBLE STOCHASTIC MODELING APPROACHES

Until recently, the reliability analyst was forced to compromise the analysis of such large systems either by modeling sections of the problem at a time and/or by making simplifying assumptions to keep the size of the reliability model tractable (fig. 9). The difficulty in this approach is that it is time consuming and complex. Perhaps more important, it is prone to error and is often unreproducible. Reliability models for the advanced reconfigurable system example shown in figure 8, which would include the details previously discussed (fig. 7), would require on the order of millions of states in the Markov modeling sense. For each state, there exists an ordinary differential equation. Thus, a Markov model for this system would require the solution of millions of differential equations, a task that is expensive, if not impossible.



- MARKOV (CAST, ARIES, CARSRA, SURF)
- COMBINATORIAL (CARE, CARE II)
- KOLMOGOROV (CARE III) (REF. 4)

Figure 9

ALTERNATE STOCHASTIC MODELING APPROACHES

Aside from using the popular Markov technique, two other approaches come to mind. The combinational method is the traditional piece-parts technique (fig. 10). In applying this technique to a fault-tolerant system with a reasonable degree of complexity, one soon learns, as in the development of CARE II, that the computational aspects become unmanageable and involve nested integrals four or more deep. The Kolmogorov method, in conjunction with a state aggregation technique, overcomes the computational difficulties of both the Markov and combinatorial techniques.

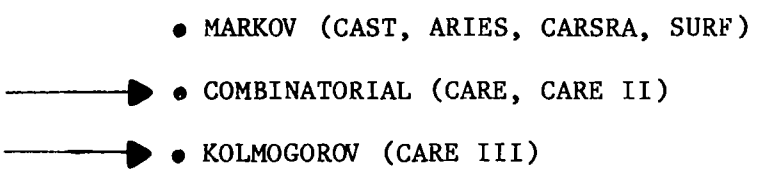


Figure 10

THE CARE III APPROACH

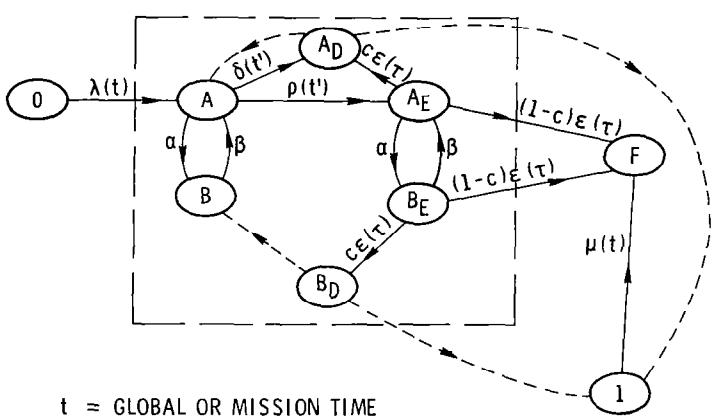
The ability of CARE III to provide extensive fault occurrence and fault-handling models is largely attributed to its ability to cope with large state spaces and is made possible by the observation that the time constants associated with fault occurrence are on the order of 10^4 hours, whereas the time constants of the fault-handling model are on the order of 10^{-5} hours. This wide time separation allows the fault occurrence model to be treated as being independent of the fault-handling model. Thus, the fault-handling model is evaluated without regard to fault occurrences (fig. 11). The results of the fault-handling model are then combined with the fault occurrence model to produce the desired reliability outputs. The fault occurrence model is solved using Kolmogorov's forward differential equations. The Kolmogorov technique is used ecause the state reduction process discussed above necessarily requires the solution of a nonhomogenous (time-dependent failure rates) Markov process.

APPROACH	o DEFINE SYSTEM STATE ONLY IN TERMS OF NUMBER OF EXISTING FAULTS
	o INDEPENDENTLY EVALUATE TRANSITION PARAMETERS AS A FUNCTION OF DISTRIBUTION OF POSSIBLE FAULT TYPES AND STATES
	o DETERMINE RELIABILITY USING KOLMOGOROV'S FORWARD DIFFERENTIAL EQUATIONS
TASK	o NUMBER OF STATES DRASTICALLY REDUCED, TRANSITION RATES NECESSARILY TIME DEPENDENT

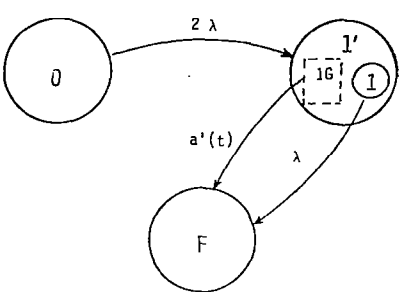
Figure 11

A MIXED MARKOV MODEL AND ITS STATE-REDUCED AGGREGATED RELIABILITY MODEL

An illustration of the state reduction technique can be seen by observing the reliability model of a two-unit system (fig. 12(a)). States 0, 1, and F are the fault occurrence states. The states enclosed in the dashed lines are the fault-handling states. The two-unit model is a mixture of a nonhomogenous and a semi-Markov model, which is the type of model CARE III was designed to evaluate. The model that CARE III actually evaluates is the aggregated reliability model shown in Figure 12(b). The aggregated model is a nonhomogeneous Markov model. CARE III approximates the mixed process with a nonhomogeneous Markov process and can do so because of the wide separation in time constants in the fault occurrence and fault-handling models. In the aggregate model, the states are strictly fault occurrence states (defines number of failed units). The fault-handling model information contained in the dashed box of the two-unit system is mapped into the time-varying transition rate a'(t). The nonhomogeneous aggregated Markov model is solved using the Kolmogorov solution technique to produce time-varying probabilities of being in states 0, 1', and F (the failure state) over the desired mission time. Although the state reduction wasn't too dramatic for this simple example, in practical assessments, state reductions of 6 orders of magnitude have been estimated.



(a)



(b)

Figure 12

CARE III FAULT-HANDLING MODEL

The ability of CARE III to model the fault/error models delineated in figure 7 is made possible by CARE III's single- and double-fault models through the judicious selection of the appropriate transition rates and/or state holding probability density functions. The double fault model accounts for critically coupled coexisting failures, which are user defined. The critically coupled failures, when they exist, are defined by certain combinations of pairs of states in the single-fault model (e.g., failure of two critically coupled units each in state A will cause system failure). The structure of the single-fault model can be grasped by referring to figure 13, the reliability model of a two-unit system.

Initially, the system is in state 0 and has experienced no failures. When a failure occurs, the system enters state A, the active latent state. This arrival is governed by the arrival density $\lambda(t)$. Depending upon the nature of the failure (i.e., permanent, transient, intermittent, etc.), the fault-handling model will be defined differently. For example, if the failure is intermittent, $\lambda(t)$ would be the probability density function (pdf) for the arrival of an intermittent, and states A and B define the intermittent model where α and β are constant transition rates into and out of state B. When the system is in state B, the benign state, the failed unit appears to have healed itself, that is, the manifestation of the failure, a fault, vanishes. However, when the failed manifestation is once again resumed (the fault reappears), the system enters state A, where the failure looks like a permanent failure. It could be detected by a self-test program with pdf $\delta(t')$, and the system would enter state A_D, the active detected state. If a spare exists, the system will purge the faulty unit and switch in the spare (dashed arc to state 1). Alternatively, while in the active state, the fault could generate errors with pdf $\rho(t')$. The system then will enter A_E, the active error state. The intermittent failure could manifest its intermittent state again, and the system would then enter state B_E, the benign error state. Although the failure is benign, the error may not be benign and may cause system failure which is denoted by the B_E to F transition $(1-c)\varepsilon(\tau)$.

The error detection density is $\varepsilon(\tau)$, and 1-C is the proportion of errors from which the system is unable to recover. While in state B_E, the error could be detected and corrected. In this event, the system enters state B_D (benign detected) by transition $c\varepsilon(\tau)$. At this point, the system may choose to do nothing further with the detected and corrected error and so move to the benign state, or the system may choose to reconfigure out the module containing the error and therefore move to state 1. The dashed arcs are instantaneous transitions. The other transition out of state A_E is to state F, the single-point failure transition $(1-c)\varepsilon(\tau)$. This transition is similar to the B_E to F transition. In a well-designed fault-tolerant system, $(1-c)\varepsilon(\tau)$ should be near zero. If $\lambda(t)$ is the pdf for the arrival of a transient, α would be set to a value greater than zero and β would be equal to zero. The pdf $\lambda(t)$ for the arrival of a permanent failure would be defined so that $\alpha = \beta = 0$. The dashed arc going from state A_D to A enables the analyst to include the effects of the system decision that the detected fault which took the system from state A to A_D was, in fact, a transient. In this regard, the system would not reconfigure out a nonfailed module.

The reader will note that the reliability model has three measures of time associated with it, which necessarily makes the model a semi-Markov process. This added complexity is required because the behavior of the system is dependent on the onset of the various fault-behavior events. The availability of data for the fault-handling models is unfortunately still poor at best and is often nonexistent altogether. The creation of the data is the subject of a considerable amount of current research. The GLOSS capability alluded to in figure 2 was used to estimate $\delta(t')$ and $\varepsilon(\tau)$ for permanent faults in the CPU of an avionic miniprocessor. (See fig. 13.)

Although the literature has often reported that transient faults are by far the most frequently occurring anomaly, virtually no test data exist that can be used for modeling transient occurrences or transient fault handling. Test data for intermittent faults are also sparse (ref. 5).

In view of the extreme sensitivity that reliability assessments of ultrareliable systems show to best-guess transient and intermittent failure occurrence data, one can only wonder why such data are not abundant.

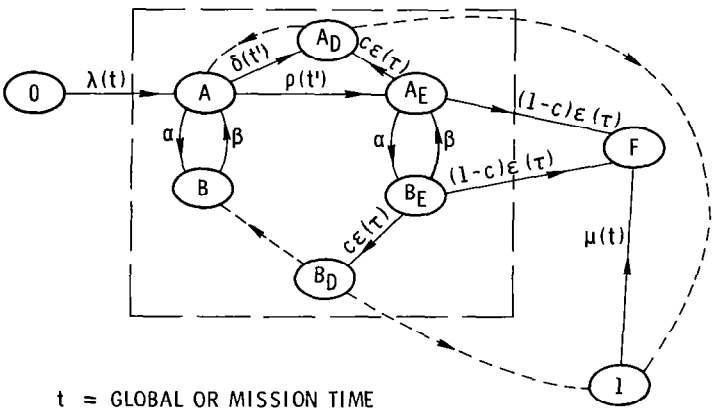


Figure 13

CONCLUDING REMARKS

The reliability assessment of ultrareliable fault-tolerant systems adds new dimensions of complexity to the assessment methodology (fig. 14). New tools are emerging to assist the reliability analyst to cope with the additional modeling complexities.

The availability of data for these novel tools is, however, slow in coming and will no doubt stunt the progress of developing ultrareliable fault-tolerant systems.

- NOVEL POWERFUL ASSESSMENT METHODOLOGIES ARE EMERGING: CARE III AND GLOSS
- AVAILABILITY OF DATA IS SPARSE
- LACK OF SUFFICIENT DATA WILL STUNT THE GROWTH OF ULTRARELIABLE DIGITAL SYSTEMS

Figure 14

REFERENCES

1. Bavuso, S. J.: Advanced Reliability Modeling of Fault-Tolerant Computer-Based Systems. Electronic Systems Effectiveness and Life Cycle Costing, J. K. Skwirzynski ed., Springer-Verlag, 1983, pp. 279-302.

2. McGough, J. G.; and Swern, F. L.: Measurement of Fault Latency in a Digital Avionic Processor Part II. NASA CR-145371, 1983.

3. Bavuso, S. J.; McGough, J. G.; and Swern, F. L.: Latent Fault Modeling and Measurement Methodology for Application to Digital Flight Controls. Advanced Flight Control Symposium, USAF Academy, Colorado Springs, Colo., August 1981.

4. Feller, William: Introduction to Probability Theory and Its Applications. Third ed. Wiley and Sons, 1968.

5. O'Neill, E. J.; and Halverson, J. R.: Study of Intermittent Field Hardware Failure Data in Digital Electronics. NASA CR-159268, 1980.