PCS - 5733

35

Design and Analysis of Fault-Tolerant Digital Systems

Barry W. Johnson

University of Virginia, Charlottesville

Addison-Wesley Publishing Company

Reading, Massachusetts • Menlo Park, California • New York
Don Mills, Ontario • Wokingham, England • Amsterdam • Bonn
Sydney • Singapore • Tokyo • Madrid • San Juan

TABLE 4.2 Input patterns capable of detecting faults (test vectors) in the circuit of Fig. 4.3

Fault	Number of test vectors	Test vectors ABC
A_0	2	100, 101
A_1	2	000, 001
B_0	2	010, 111
B_1	2	000, 101
C_0	2	011, 111
C_1	2	010, 110
D_0	1	101
D_1	2	000, 001
E_0	1	101
E_1	1	111
F_0	0	—
F_1	2	000, 101
G_0	0	—
G_1	1	111
H_0	1	010
H_1	1	000
I_0	1	111
I_1	1	101
J_0	2	010, 110
J_1	2	011, 111
K_0	1	010
K_1	2	011, 111
L_0	1	101
L_1	4	000, 001, 011, 111
M_0	0	—
M_1	4	000, 001, 011, 111
N_0	1	010
N_1	4	000, 001, 011, 111
Z_0	4	010, 100, 101, 110
Z_1	4	000, 001, 011, 111

Several important points should be made about the estimation of coverage. First, the estimation of fault coverage requires the definition of the types of faults that can occur. Stating that the fault detection coverage is 0.9, for example, is meaningless unless the types of faults considered are identified. For example, the fault detection coverage for the circuit of Fig. 4.3 is 0.9 for all stuck-at-1 and stuck-at-0 faults, but the fault detection coverage may decrease substantially if stuck-open faults are included.

A second important point about the fault coverage is that it is typically assumed to be a constant. It is easy to envision applications in which the probability of detecting a fault, for example, increases as a function of time, after the occurrence of the fault. However, to simplify the analysis, the various fault coverages are normally assumed to be constants.

4.3 Reliability Modeling

Reliability is perhaps one of the most important attributes of systems. Almost all specifications for systems mandate that certain values for reliability be achieved and in some way proved. We have seen in the previous sections that reliability can be determined experimentally if a set of N systems is operated over a period of time and the number of systems that fail during that time period is recorded. One problem with the experimental approach is the number of systems that would be required to achieve a level of confidence in the experimental results. This is particularly a problem when costs limit the number of systems that can be built. For example, the space shuttle program could not afford to build 1000 of its on-board processing systems such that reliability could be experimentally verified.

A second problem with the experimental approach is the time required to run such experiments. Many systems today are being designed to achieve reliabilities of 0.9_7, or higher, after ten hours of operation. Using the exponential failure law, a reliability of 0.9_7 corresponds to a failure rate of 10^{-8} failures per hour. Therefore, on the average, we would have to wait approximately 100 million hours, or approximately 11,416 years for the first failure to occur. Clearly, we need alternatives to the experimental approach.

The most popular reliability analysis techniques are the analytical approaches. Of the analytical techniques, combinatorial modeling and Markov modeling are the two most commonly used approaches.

4.3.1 Combinatorial Models

Combinatorial models use probabilistic techniques that enumerate the different ways in which a system can remain operational. The probabilities of the events that lead to a system being operational are calculated to form an estimate of the system's reliability.

The reliability of a system is generally derived in terms of the reliabilities of the individual components of the system. The two models of systems that are most common in practice are the series and the parallel. In a **series system,** each element of the system is required to operate correctly for the system to operate correctly. In a **parallel system,** on the other hand, only one of several elements must be operational for the system to perform its functions correctly.

In practice, systems are typically combinations of series and parallel subsystems. Once we have discussed both the series and parallel structures, we will examine techniques for modeling systems that contain *both* series and parallel subsystems.

Series Systems

The series system is best thought of as a system that contains no redundancy; that is, each element of the system is needed to make the system function correctly. For example, a digital filter that contains a microprocessor, an analog-to-digital converter, and a digital-to-analog converter needs each of these elements to perform the digital filtering function; if any one of the three elements fails, the system fails. One way of representing the series system is through the use of **reliability block diagrams.** The reliability block diagram can be thought of as a flow diagram from the input of the system to the output of the system. Each element of the system is a block in the reliability block diagram and, for the series system, the blocks are placed in series to indicate that a path from the input to the output is broken if one of the elements fails.

For example, the generalized reliability block diagram of a series system that contains N elements is shown in Fig. 4.4. Each of the N elements is required for the system to function correctly. The reliability of the series system can be calculated as the probability that none of the elements will fail. Another way to look at this is that the reliability of the series system is the probability that all of the elements are working properly.

Suppose we let $C_{iw}(t)$ represent the event that component C_i is working properly at time t, $R_i(t)$ is the reliability of component C_i at time t, and $R_{series}(t)$ is the reliability of the series system. Further suppose that the series system contains N series components as shown in Fig. 4.4. The reliability at any time t is the probability that all N components are working properly. In mathematical terms,

$$R_{series}(t) = P(C_{1w}(t) \cap C_{2w}(t) \cap \cdots \cap C_{Nw}(t))$$

Assuming that the events, $C_{iw}(t)$, are independent, we have

$$R_{series}(t) = R_1(t)R_2(t)\cdots R_N(t)$$

or

$$R_{series}(t) = \prod_{i=1}^{N} R_i(t)$$

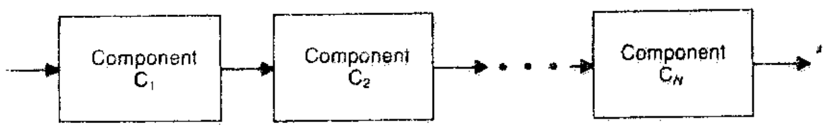

Fig. 4.4 The reliability block diagram of a series system—each element of the system must operate correctly for the system to operate correctly.

An interesting relationship exists in a series system if each individual component satisfies the exponential failure law. Suppose that we have a series system made up of N components, and each component i has a constant failure rate of λ_i. Also assume that each component satisfies the exponential failure law such that the reliability of each component is $R_i(t) = e^{-\lambda_i t}$. The reliability of the series system is given by

$$R_{series}(t) = e^{-\lambda_1 t}e^{-\lambda_2 t}\cdots e^{-\lambda_N t}$$

or

$$R_{series}(t) = e^{-\sum_{i=1}^{N}\lambda_i t} = e^{-\lambda_{system} t}$$

where $\lambda_{system} = \sum_{i=1}^{N}\lambda_i$ and corresponds to the failure rate of the system. In other words, the failure rate of a series system can be calculated by adding the failure rates of all the components that make up the series system.

As an example of a series system, consider the simple aircraft control system shown in Fig. 4.5. This system contains sensors that are used to measure the roll, pitch, and yaw positions of the aircraft; sensors to measure the crew's desired roll, pitch, and yaw positions; actuators that are used to control the roll, pitch, and yaw; and computers that perform the computations required of the flight control system. The computers receive the sensor values and supply the actuator commands over a serial data bus that connects the sensors, actuators, and the computers. A special high-speed data bus interconnects the computers for the purpose of data transfer among the computers. Each element of the system is required if the system is to perform correctly; there is no redundancy in the system. For example, the failure of any one sensor or computer renders the system unable to operate correctly.

The reliability block diagram of the flight control system is shown in Fig. 4.6. The reliability block diagram illustrates the series nature of the system. For simplicity, assume that all six sensors have the same reliability $R_s(t)$, each of the three actuators has the reliability, $R_{act}(t)$, and each computer has the reliability $R_c(t)$. Also, let the computer interconnection bus have the reliability $R_{bus1}(t)$ and the primary control bus have the reliability $R_{bus2}(t)$. By taking the product of the element reliabilities, we find that the reliability of the system is given by

$$R_{system}(t) = R_s^6(t)R_{act}^3(t)R_c^3(t)R_{bus1}(t)R_{bus2}(t)$$

Because the failure rates can be added in a series system to obtain the failure rate of the system, we can write

$$\lambda_{system} = 6\lambda_s + 3\lambda_{act} + 3\lambda_c + \lambda_{bus1} + \lambda_{bus2}$$

where λ_s is the failure rate of one sensor, λ_{act} is the failure rate of one actuator, λ_c is the failure rate of one computer, λ_{bus1} is the failure rate of the com-

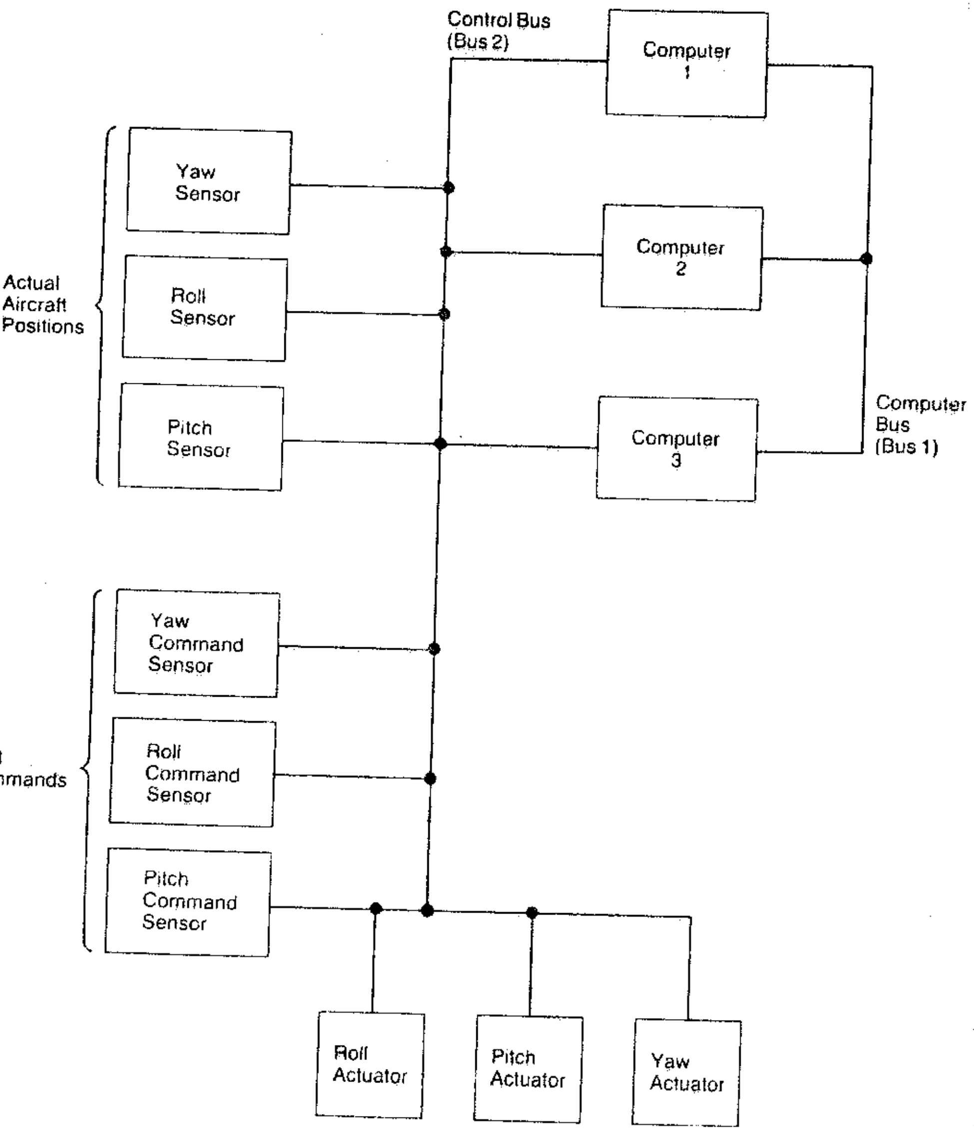

Fig. 4.5 An aircraft control system designed as a series system.

...puter interconnection bus, and λ_{bus2} is the failure rate of the primary control bus. λ_{system} is the failure rate of the system. If the failure rates of the system are

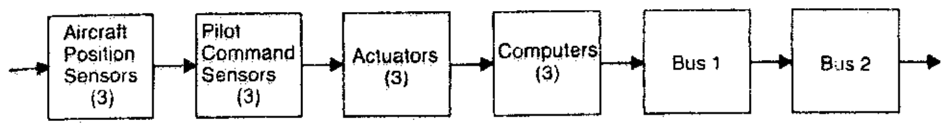

Fig. 4.6 The reliability block diagram of the system in Figure 4.5 illustrates the series nature of the system.

$$\lambda_s = 1 \times 10^{-6} \quad \text{failures per hour}$$
$$\lambda_{act} = 1 \times 10^{-5} \quad \text{failures per hour}$$
$$\lambda_c = 4 \times 10^{-4} \quad \text{failures per hour}$$
$$\lambda_{bus1} = 1 \times 10^{-6} \quad \text{failures per hour}$$
$$\lambda_{bus2} = 2 \times 10^{-6} \quad \text{failures per hour}$$

the system failure rate will be

$$\lambda_{system} = 1.239 \times 10^{-3} \quad \text{failures per hour}$$

The reliability after five hours for this system is approximately 0.995.

Parallel Systems

The distinguishing feature of the basic parallel system is that only one of N identical elements is required for the system to function. For example, many families have two or more cars when one is, in many cases, sufficient to meet the family's needs. The probability of having at least one car working can be determined by modeling the multiple-car family as a parallel system.

The reliability block diagram of the basic parallel system that contains N identical elements is shown in Fig. 4.7. As can be seen, a path exists in the reliability block diagram from input to output as long as one of the N identical elements remains operational. The unreliability of the parallel system can be computed as the probability that all of the N elements fail. Suppose that we let $C_{if}(t)$ represent the event that element i in the parallel system has failed at time t, $Q_{parallel}(t)$ be the unreliability of the parallel system, and $Q_i(t)$ be the unreliability of the i^{th} element. $Q_{parallel}(t)$ can be computed as

$$Q_{parallel}(t) = P(C_{1f}(t) \cap C_{2f}(t) \cap \cdots \cap C_{Nf}(t))$$

or

$$Q_{parallel}(t) = Q_1(t)Q_2(t)\cdots Q_N(t) = \prod_{i=1}^{N} Q_i(t)$$

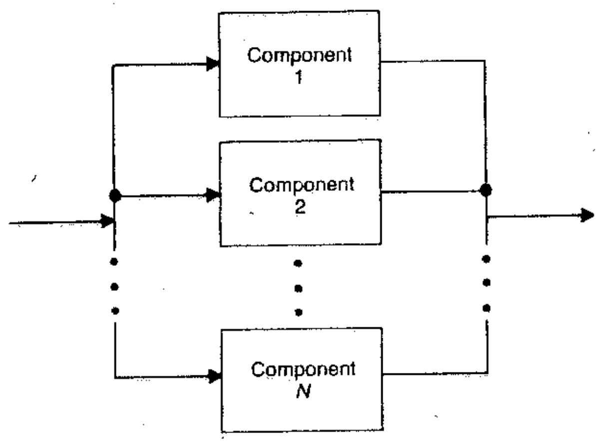

Fig. 4.7 The reliability block diagram of the parallel system—only one of N components must operate correctly for the system to operate correctly.

The reliability of the parallel system can now be computed because we know that the reliability and the unreliability must add to 1.0. Mathematically, we must have $R(t) + Q(t) = 1.0$ for any system. Consequently, we can write

$$R_{\text{parallel}}(t) = 1.0 - Q_{\text{parallel}}(t) = 1.0 - \prod_{i=1}^{N} Q_i(t) = 1.0 - \prod_{i=1}^{N}(1.0 - R_i(t))$$

Note that the equations for the parallel system assume that the failures of the individual elements that make up the parallel system are independent. For random hardware failures, the independence of failures is a good assumption; however, for failures that are the result of items such as external disturbances, the independence assumption is not very good. Therefore, the combinatorial modeling techniques are most often applied to the analysis of random failures in a system's hardware.

To analyze a system that has a parallel structure, consider the system shown in Fig. 4.8. The architecture of the system in Fig. 4.8 is commonly found in aerospace applications. The system consists of two identical computers, two identical interface units, two identical display devices, and two identical communication buses. The system requires that at least one of each unit work properly for the system to perform its functions. Once a particular unit has failed, it is assumed that the second unit of that type automatically assumes the functions of the failed unit.

One important point about the system of Fig. 4.8 is that it has both a series and a parallel structure. It is parallel in the sense that only one of the

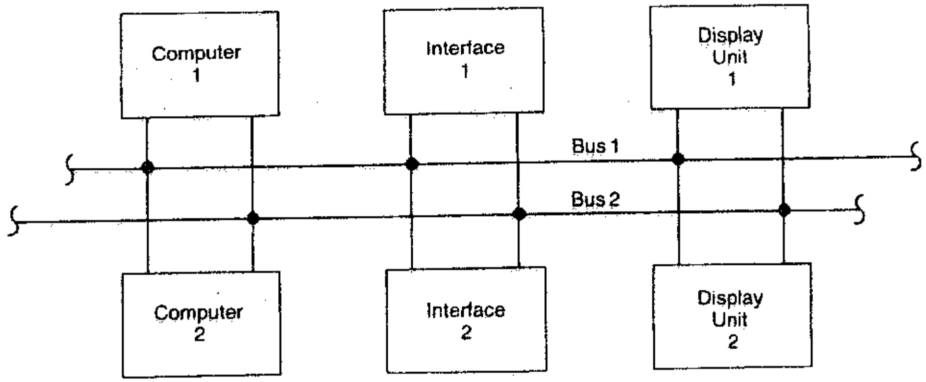

Fig. 4.8 An example computer system having a structure that is a combination of series and parallel.

two computers, for example, must function for the system to function. It is series in the sense that one computer, one interface unit, one display device, *and* one bus must operate for the system to operate. The reliability block diagram of the system of Fig. 4.8 is shown in Fig. 4.9. Note that a path from the input of the diagram to the output exists if and only if enough elements are functioning to allow the system to operate properly.

A reliability block diagram that contains both series and parallel structures can be reduced to a single series diagram by replacing each of the parallel portions of the system with an equivalent, single element that has the same reliability as the parallel structure. For example, we know from the

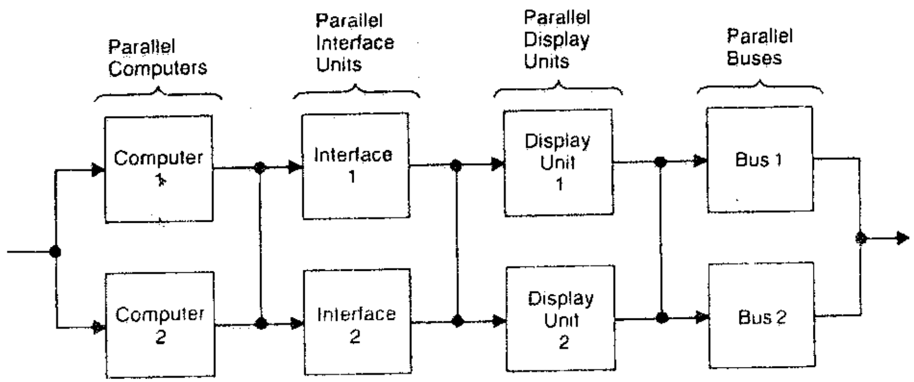

Fig. 4.9 Reliability block diagram of the series/parallel system of Fig. 4.8.

analysis of a parallel system that the parallel organization of the two computers in Fig. 4.9 has a reliability given by

$$1.0 - (1 - R_c(t))^2$$

where $R_c(t)$ is the reliability of one computer. Therefore, the parallel organization of the computers can be replaced by a single element having a reliability of $1.0 - (1 - R_c(t))^2$, as is illustrated in Fig. 4.10. The transformation of a parallel system into an equivalent series system is a very common technique used to reduce reliability block diagrams. The reliability block diagram of Fig. 4.9 can be reduced to a series diagram by applying the reduction concept to each parallel structure. The reduced block diagram is shown in Fig. 4.11 where $R_c(t)$ is the reliability of one computer, $R_{if}(t)$ is the reliability of one interface unit, $R_d(t)$ is the reliability of one display unit, and $R_b(t)$ is the reliability of one bus.

The reliability for the reduced block diagram of Fig. 4.11 can be written as

$$R_{system}(t) = [1 - (1 - R_c(t))^2][1 - (1 - R_{if}(t))^2][1 - (1 - R_d(t))^2] \cdot [1 - (1 - R_b(t))^2]$$

As an example, the reliability of the system after one hour given $R_c(1) = R_{if}(1) = R_d(1) = R_b(1) = .9$ will be

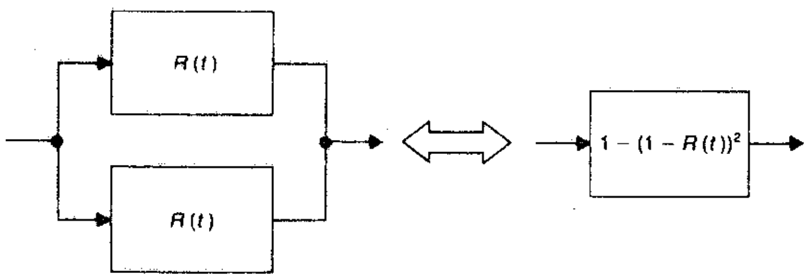

Fig. 4.10 A parallel system can be reduced to a series element with the proper reliability function.

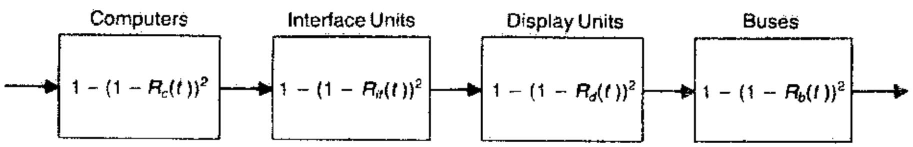

Fig. 4.11 Reduced reliability block diagram for the system of Fig. 4.8.

$$R_{system}(1\text{ hour}) = [1 - (1 - .9)^2][1 - (1 - .9)^2][1 - (1 - .9)^2][1 - (1 - .9)^2] = .96$$

Now that we have several tools for investigating the reliability of systems, we can compare the reliability benefits that redundancy can offer. For example, the redundant system just analyzed had a reliability of 0.96 after one hour. The nonredundant system containing one computer, one display device, one bus, and one interface device has a reliability equivalent to the product of the individual element's reliabilities because the nonredundant system is a simple series system. Therefore, the nonredundant system has a reliability of 0.6561 after one hour. As is seen, the incorporation of redundancy has significantly improved the system's reliability.

Note, however, that redundancy does not always improve a system's reliability. Whether or not the reliability is improved depends on the amount of redundancy employed and the reliability of the elements used to construct the system, as well as other factors. For example, if each element of the redundant system in Fig. 4.8 has a reliability of 0.1 at the end of one hour, the redundant system has a reliability of 0.0013 at the end of one hour, and the nonredundant system has a reliability of 0.0001 at the end of one hour. We certainly would hope that elements with a reliability as poor as 0.1 would never be used in a system, but this example does show that the redundancy does not significantly improve the reliability. As we stated in the first chapter, reliability and fault tolerance are not one in the same. As we shall see when we begin to analyze more complex systems, the distinction between fault tolerance and reliability becomes even clearer.

4.3.2 Fault Coverage and Its Impact on Reliability

As defined earlier, fault coverage is a measure of a system's ability to recover from faults. For example, a system with redundant computers that requires reconfiguration before the redundancy can be used depends heavily on good fault coverage. During the analysis of the parallel system, we assumed that the fault coverage was perfect; if we had three computers and needed only one to operate, the reliability was calculated solely as the probability that one of the three computers was operational. Unfortunately, the assumption of perfect fault coverage does not consider that the system may not be able to use the redundancy because it cannot identify that a unit is faulty, remove that faulty unit, and replace it with a fault-free one.

To illustrate the problem, consider a simple parallel system consisting of two identical modules and having the reliability block diagram shown in Fig. 4.12. Assume that module 1 is the primary module and that module 2 is a spare module that is switched on-line in the event of failure of module 1. In other words, the system uses the concept of standby sparing. Under ideal circumstances, the standby sparing system functions correctly as long as one of the two modules functions correctly. In reality, however, the failure

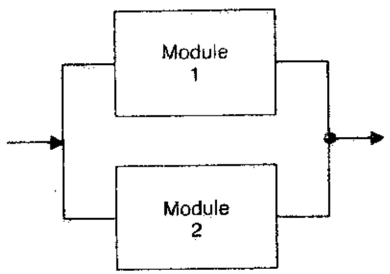

Fig. 4.12 Reliability block diagram of a simple parallel system to illustrate the impact of fault coverage.

of the primary module (module 1, in this case) must be detected and correctly handled before the second module can be used. In other words, the parallel system with two modules functions correctly as long as one of the following two conditions exist:

1. Module 1 is functioning correctly.
2. Module 2 is functioning correctly, module 1 has failed, and the failure was detected and appropriately handled.

The probability that one of these two events will exist can be written in terms of the reliabilities of the modules and the fault coverage as

$$R_{system}(t) = R_1(t) + (1 - R_1(t))C_1R_2(t)$$

where C_1 is the fault coverage associated with module 1, $R_1(t)$ is the reliability of module 1, and $R_2(t)$ is the reliability of module 2. The reliability equation enumerates all of the working states of the system. If the reliabilities and the coverage factors of the two modules are identical, the reliability expression reduces to

$$R_{system}(t) = R(t) + R(t)C(1 - R(t))$$

where $R(t)$ is the reliability of one module and C is the fault coverage. Note that if the fault coverage C, is 1.0, the reliability expression reduces to

$$R_{system}(t) = 2R(t) - R^2(t) = 1 - (1 - R(t))^2$$

which is the reliability of the perfect parallel system. Also note that if the fault coverage is 0.0, the reliability expression reduces to simply the reliability of one module; therefore, the primary module must function correctly for the system to function correctly.

It is interesting to study the impact of the fault coverage in the parallel system with two modules. Figure 4.13 shows the reliability of the system as a function of fault coverage for a module reliability of R. Note that the

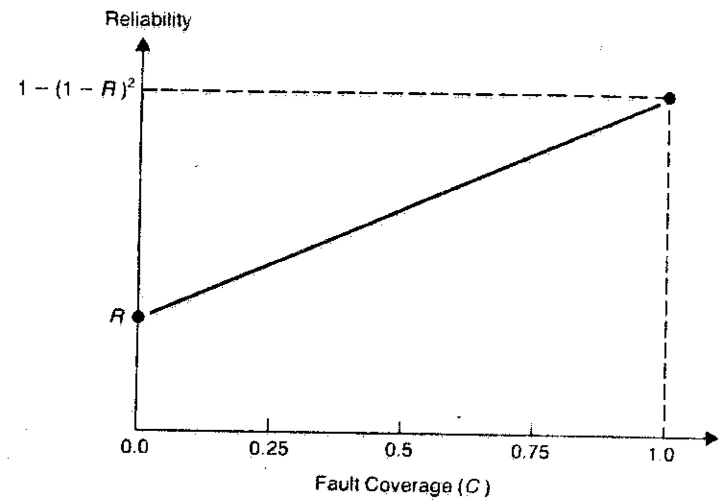

Fig. 4.13 Reliability versus fault coverage for a parallel system with two modules and using standby sparing. Each module has a reliability of R and a fault coverage of C.

reliability of the parallel system with two modules is a linear function of the coverage factor. At a coverage of 1.0, the reliability is $1 - (1 - R)^2$, which is the reliability of the perfect parallel system. At a coverage of 0.0, the reliability is R, which is simply the probability that the primary module will not fail.

One important point about the above analysis is that the failure of the second module is unimportant unless it has replaced the first module. In other words, module 1 is the primary module and as long as it functions correctly, the system functions correctly, even if module 2 fails. In many systems, this may not be true. For example, consider a duplex system that performs comparisons between the two modules as one form of fault detection. Once a fault is detected, the two modules go into more detailed fault analysis routines, often called self-diagnostics, in an attempt to identify which of the two modules is faulty. If the faulty module is identified, the system continues to operate with the one fault-free module, and the comparison mechanism is disabled. If the faulty module cannot be identified, the system discontinues its operation. Therefore, an undetected fault in either of the modules causes the system to fail and must be accounted for in the reliability analysis.

Consider once again the parallel system with two modules as shown in Fig. 4.12, but now perform comparisons between the two modules as one means of fault detection. Assume for now that the comparison is perfect and detects all faults. Once the comparison process detects a fault, the system implements self-diagnostics to attempt to determine which of the two modules is faulty. If the fault can be located successfully, the fault-free module begins to perform the functions of the system. The system functions correctly as long as both modules work or the fault has been detected and handled correctly. The reliability of the system can be written by enumerating the working states of the system.

In mathematical terms, we have

$$R_{system}(t) = R_1(t)R_2(t) + R_1(t)(1 - R_2(t))C_2 + (1 - R_1(t))C_1R_2(t)$$

where $R_1(t)$ is the reliability of module 1, $R_2(t)$ is the reliability of module 2, C_1 is the fault coverage of the self-diagnostics of module 1, and C_2 is the fault coverage of the self-diagnostics of module 2. If the reliabilities and fault coverages of the two modules are identical, the reliability reduces to

$$R_{system}(t) = R^2(t) + 2R(t)C(1 - R(t))$$

For perfect fault coverage, we obtain the same as before; that is, the system has the reliability of the perfect parallel system. If the fault coverage is 0.0, the system has a reliability of $R_{system}(t) = R^2(t)$, which is simply the probability that both modules operate correctly. Figure 4.14 shows the reliability of the system as a function of fault coverage.

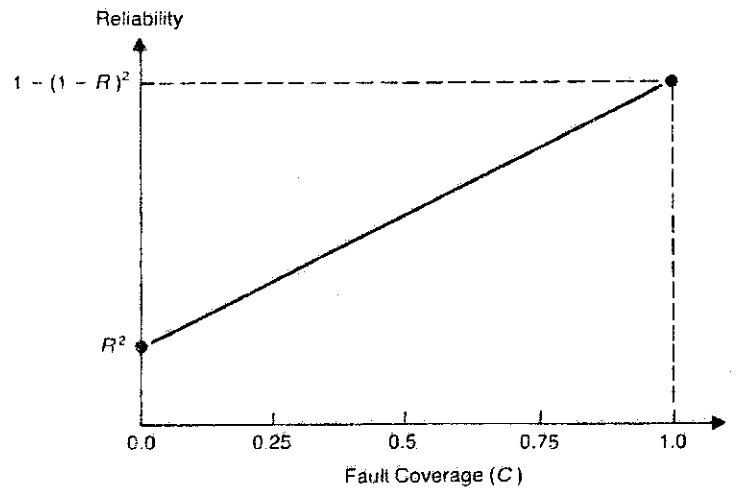

Fig. 4.14 Reliability versus fault coverage for a parallel system with two modules and using comparisons between the two modules. Each module has a reliability of R and a fault coverage of C.

4.3.3 *M*-of-*N* Systems

***M*-of-*N* systems** are a generalization of the ideal parallel system. In the ideal parallel system, only one of N modules is required to work for the system to work. In the M-of-N system, however, M of the total of N identical modules are required to function for the system to function. A good example is the TMR configuration where two of the three modules must work for the majority voting mechanism to function properly. Therefore, the TMR system is a 2-of-3 system.

Consider as an example the TMR system. As seen in the previous sections, we can write the reliability of a system by enumerating all of the possible states in which the system can be functional. Suppose that we have a TMR system with modules 1, 2, and 3 connected in a majority voting arrangement. As long as two of the three modules are functioning correctly, the system will perform correctly. Ignoring the reliability of the voter, the reliability of the TMR system can be written as

$$R_{TMR}(t) = R_1(t)R_2(t)R_3(t) + R_1(t)R_2(t)(1 - R_3(t)) + R_1(t)(1 - R_2(t))R_3(t) + (1 - R_1(t))R_2(t)R_3(t)$$

where $R_i(t)$ is the reliability of the i^{th} module. If $R_1(t) = R_2(t) = R_3(t) = R(t)$, the reliability of the TMR system reduces to

$$R_{TMR}(t) = R^3(t) + 3R^2(t)(1 - R(t)) = 3R^2(t) - 2R^3(t)$$

Now that we have the expression for the reliability of the TMR system, it is interesting to examine the reliability improvements that can be obtained through the use of TMR. Figure 4.15 shows a plot of the reliability of a TMR arrangement as a function of the reliability of the modules that compose the TMR system. In other words, Fig. 4.15 simply shows a plot of the equation $R_{TMR} = 3R^2 - 2R^3$ versus R. As can be seen, there is a point at which the reliability of the TMR system and the reliability of the single module cross. The crossover point is easily found by setting the reliability of the TMR system equal to the reliability of the single module and solving the resulting quadratic equation.

In mathematical terms, we have

$$R_{TMR} = 3R^2 - 2R^3 = R$$

or

$$3R - 2R^2 = 1$$

which implies the quadratic equation

$$R^2 - \frac{3}{2}R + 0.5 = 0$$

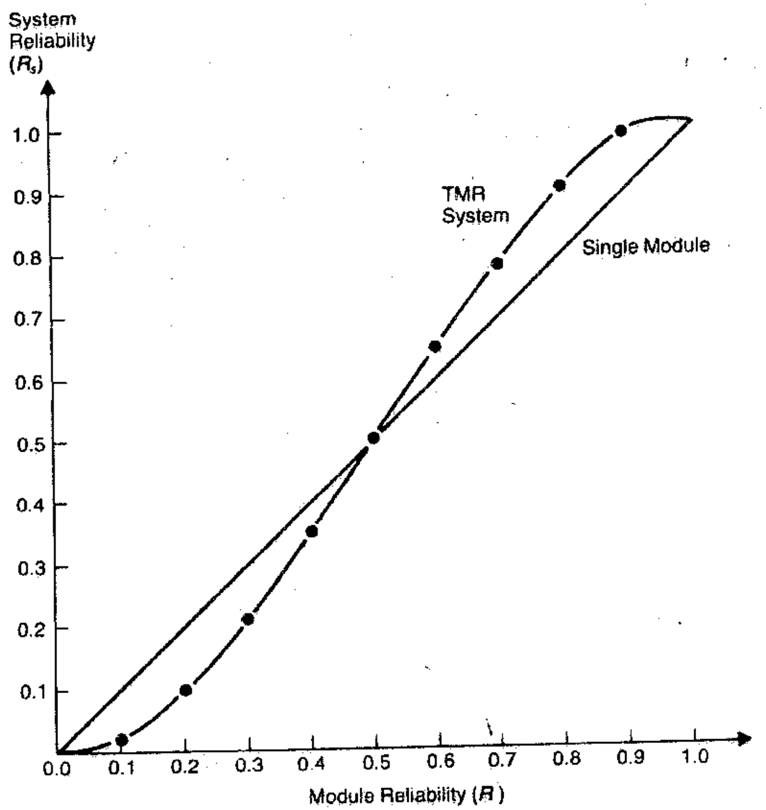

Fig. 4.15 Comparison of the reliability of a TMR system composed of three identical modules with the reliability of a single module.

The two solutions to the quadratic equation are 0.5 and 1.0, which implies that the reliability of the TMR system is equal to that of the corresponding nonredundant system when the reliability of the single module is 0.5 or the module is perfect ($R = 1$).

This further illustrates a point that we made when we defined fault tolerance and reliability. A system can be tolerant of faults and still have a low reliability. For example, a TMR system constructed from modules that have individual reliabilities of 0.5 can tolerate a fault in one of those modules, but the reliability of the TMR system is the same as the reliability of a single module. Conversely, a system can achieve a high reliability without being fault tolerant. Certainly, a system that consists of a perfect module will have the highest possible reliability but will not possess, or need, the attribute of fault tolerance. This is, of course, an unrealistic example, but, in general, as the reliability of the components of a system increases, the reliability of the system also increases. It is possible for the reliability of a nonredundant system to approach that of a redundant system constructed from the same modules. The nonredundant system, however, will not be fault tolerant.

In many cases, we have systems that are of the M-of-N structure but are not TMR; the general NMR system is a good example. In general, if there are N identical modules and M of those are required for the system to function properly, the system can tolerate $N - M$ module failures. The expression for the reliability of an M-of-N system can be written as

$$R_{M\text{-of-}N}(t) = \sum_{i=0}^{N-M} \binom{N}{i} R^{N-i}(t)(1 - R(t))^i$$

where

$$\binom{N}{i} = \frac{N!}{(N-i)!\,i!}$$

For example, the TMR system reliability is given by

$$R_{TMR}(t) = \sum_{i=0}^{1} \binom{3}{i} R^{3-i}(t)(1 - R(t))^i$$

which reduces to

$$R_{TMR}(t) = 3R^2(t) - 2R^3(t)$$

which is identical to the expression derived earlier.

4.3.4 Markov Models

The primary difficulty with the combinatorial models is that many complex systems cannot be modeled easily in a combinatorial fashion. The reliability block diagrams can be extremely difficult to construct, and the resulting reliability expressions are often very complex. In addition, the fault coverage that we have seen to be extremely important in the reliability of a system is sometimes difficult to incorporate into the reliability expression in a combinatorial model. Finally, the process of repair that occurs in many systems is very difficult to model in a combinatorial fashion. For these reasons, we often use **Markov models.**

The purpose of the presentation in this text is not to delve into the mathematical details of Markov models but to understand how to use Markov models. For more explicit mathematical details, refer to the references ([Shooman 1968] and [Trivedi 1982]). The discussions here will provide sufficient mathematical background to apply the Markov model but will not pursue various techniques for solving the models.