Failure State Diagrams - Repair Strategies

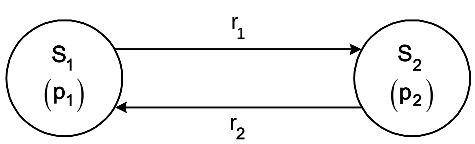
October 2007

Where We Left Off

In our September, 2007, Geek Corner article, [Calculating Availability – Failure State Diagrams](), we introduced failure state diagrams as a formal way to derive availability relationships. A failure state diagram contains nodes that represent various states of failure of a system and connectors that indicate ways in which one state may transition to another.

Each connector is characterized by the rate at which its source state will transition to its target state. In the steady state, the transition rates exiting a state must equal the transition rates entering a state.

For instance, Figure 1 shows a system with two states, S_1 and S_2. The system will be in state S_1 with a probability of p_1, and it will be in state S_2 with a probability of p_2. When in state S_1, the system will transition to state S_2 at a rate of r_1. Likewise, when the system is in state S_2, it will transition to state S_1 at a rate of r_2. Therefore, the total transition rate from state S_1 to S_2 is p_1r_1; and the total transition rate from state S_2 to S_1 is p_2r_2.



Simple Failure State Diagram
Figure 1

What are the probabilities that the system will be in either state?

Since the exiting transition rate from a state must be equal to its entering transition rate in the steady-state condition, then

$$p_1r_1 = p_2r_2 \qquad \text{for state } S_1$$
$$p_2r_2 = p_1r_1 \qquad \text{for state } S_2$$

Note that for these two nodes, there are two equations; but they are not independent (in this trivial case, they are, in fact, the same). In general, for an n-node system, there will be n equations, but only n-1 of them will be independent. Since we want to solve for n variables (the n probabilities), we need one more equation. This equation is simply the observation that the probabilities must add up to one since the system will always be in one and only one state. Thus,

$$p_1 + p_2 = 1$$

© 2007 Sombers Associates, Inc., and W. H. Highleyman
www.availabilitydigest.com

These equations can now be solved for the probabilities that the system will be in state S_1 or in state S_2:

$$\text{probability that the system will be in state } S_1 = p_1 = \frac{r_2}{r_1 + r_2}$$

$$\text{probability that the system will be in state } S_2 = p_2 = \frac{r_1}{r_1 + r_2}$$

Application of the Failure State Diagram to Repair Strategies

Intuitive Analyses

In our November, 2006, article entitled Calculating Availability – Repair Strategies, we intuitively derived the availability of a system for sequential repair and for parallel repair:

sequential repair	There is only one repair person. Should multiple nodes fail, they are repaired one at a time by the single repair person. Each node is returned to service as soon as it is operational, at which time the repair person moves on to the next failed node.
parallel repair	There are multiple repair personnel available. Should multiple nodes fail, they are repaired simultaneously by different repair people. Each node is returned to service as soon as it is operational.

Our intuitive arguments led to the following expressions for system availability for an n-node system with one spare:

$$F = 1 - A = n(n-1)(1-a)^2 \qquad \text{for sequential repair} \qquad (1)$$

$$F = 1 - A = \frac{n(n-1)}{2}(1-a)^2 \qquad \text{for parallel repair} \qquad (2)$$

where

$$a = \frac{mtbf}{mtbf + mtr} \qquad (3)$$

and

A	is the availability of the system = (1-*F*).
F	is the probability of failure of the system.
n	is the number of nodes in the system.
a	is the availability of a node.
mtbf	is the mean time before failure for a node.
mtr	is the mean time to recover a node.

Let us now derive these expressions formally using failure state diagrams.[1]

Sequential Repair

The failure state diagram for sequential repair is shown in Figure 2. The system comprises *n* nodes. There is one spare node in the system. Therefore, the system will survive the failure of any one node. Should any two nodes fail, the system is down.

[1] Failure state diagrams for several cases are analyzed in Appendix 3, Failover Fault Models, in the book entitled *Breaking the Availability Barrier: Survivable Systems for Enterprise Computing*, by Dr. Bill Highleyman, Paul J. Holenstein, and Dr. Bruce Holenstein, AuthorHouse; 2004.

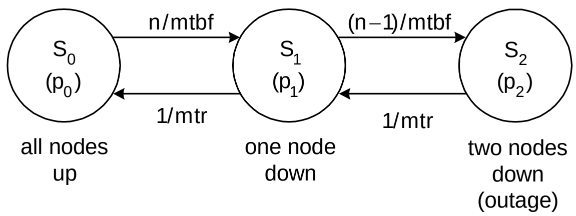
© 2007 Sombers Associates, Inc., and W. H. Highleyman
www.availabilitydigest.com

There are three states in this system. In state S_0, all nodes are operational. In state S_1, one node has failed; but the system continues in operation. In state S_2, two nodes have failed; and the system is down.

A node failure in state S_0 will cause a transition to state S_1. Each node fails at the rate of 1/mtbf (i.e., if the nodal mtbf is 1,000 hours, each node will fail once every 1,000 hours). Therefore, in state $S_{0,}$ in which *n* nodes are operational, the nodal failure rate will be *n*/mtbf. This is the rate at which the system will transition from state S_0 to state S_1.

When in state S_1, it will take a repair time of mtr to repair the node. Thus, the system will transition from state S_1 back to state S_0 at a rate of 1/mtr.

Alternatively, while in state S_1, a second node could fail. In this case, the system will enter state S_2, at which point the system is down. Since there are *n*-1 nodes operational in state S_1, the transition rate to state S_2 is (*n*-1)/mtbf. While in state S_2, one of the two downed nodes will be repaired. When it is repaired, the system will return to state S_1 and will once again be operational. Since the nodal repair time in state S_2 is still mtr, the transition rate from state S_2 is also 1/mtr.



Sequential Repair Failure State Diagram
Figure 2

In the steady state, the transitions into and out of each state must be equal. This leads to the following state transition probabilities:

State Transition Equations

State 0 $\quad p_0 n/\text{mtbf} = p_1/\text{mtr}$

State 1 $\quad p_1[1/\text{mtr} + (n-1)/\text{mtbf}] = p_0 n/\text{mtbf} + p_2/\text{mtr}$

State 2 $\quad p_2/\text{mtr} = p_1(n-1)/\text{mtbf}$

Solving for the state probabilities in terms of p_0, we have

State Probabilities (in terms of p_0)

$$p_0 = p_0$$

$$p_1 = \frac{n(\text{mtr})}{\text{mtbf}} p_0 = \frac{n(1-a)}{a} p_0$$

$$p_2 = \frac{n(n\text{-}1)\text{mtr}^2}{\text{mtbf}^2} p_0 = \frac{n(n-1)(1-a)^2}{a^2} p_0$$

where we have used the fact that

© 2007 Sombers Associates, Inc., and W. H. Highleyman
www.availabilitydigest.com

$$\frac{mtr}{mtbf} = \frac{1 - \frac{mtbf}{mtbf + mtr}}{\frac{mtbf}{mtbf + mtr}} = \frac{1 - a}{a}$$

To solve for the state probabilities, we must now make use of the fact that the sum of the probabilities must be 1:

$$p_0 + p_1 + p_2 = 1$$

This lets us solve for the state probabilities as a function of nodal availability, *a*:

State Probabilities (in terms of *a*)

$$p_0 = a^2 / D$$
$$p_1 = na(1 - a) / D$$
$$p_2 = n(n - 1)(1 - a)^2 / D$$

where

$$D = a^2 + na(1 - a) + n(n - 1)(1 - a)^2$$

The system is down if it is in state S_2. Thus, the probability of failure of the system, *F*, is the probability that the system will be in state S_2, or

$$F = n(n - 1)(1 - a)^2 / D \qquad (4)$$

and

$$A = 1 - F = n(n - 1)(1 - a)^2 / D \qquad (5)$$

These are the true expressions for the availability of a single-spared n-node system that uses sequential repair. Note that they are equivalent to our intuitive Equation (1) except for the denominator *D*. To the extent that the value of *D* departs from one, *D* represents the error in our intuitive approach.

D can be written as

$$\begin{aligned} D &= a^2 + na(1 - a) + n(n - 1)(1 - a)^2 \\ &= [a^2 + 2a(1 - a) + (1 - a)^2] + (n - 2)a(1 - a) + [n(n - 1) - 1](1 - a)^2 \\ &= [a + (1 - a)]^2 + func(1 - a) \\ &= 1 + func(1 - a) \\ &\approx 1 \end{aligned}$$

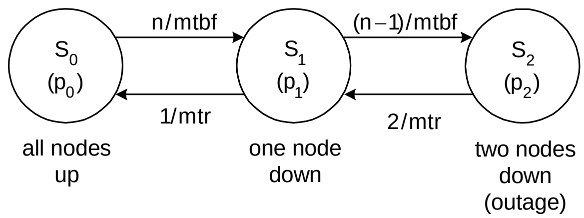
D is equal to 1 plus some function of (1-a). If (1-*a)* is small, func(1-*a*) can be ignored and *D* is approximately equal to one.. Thus, we find that the accuracy of the intuitive approach of Equation (1) depends upon the nodal availability *a* being very close to one and therefore (1-*a*) being very small. It also depends upon there being a modest number of nodes in the system. (More specifically, it depends upon the quantity (*n*-2)a(1-*a*) being very small.) This is certainly the case for the high-availability systems in which we are interested. Under these assumptions, *D* is approximately one; and the intuitive relationship holds.

If nodal availability is not close to one (say less than two 9s), or if there are hundreds of nodes in the system, the more accurate Equations (4) and (5) should be used.

© 2007 Sombers Associates, Inc., and W. H. Highleyman
www.availabilitydigest.com

Parallel Repair

The analysis of parallel repair follows closely the above analysis of sequential repair. The failure state diagram for this case is shown in Figure 3. The only difference is that there will be two repair people working on the two failed nodes when in state S_2, the system failure state. Thus, the transition rate out of state S_2 will be twice as fast, or 2/mtr. That is, if one repair person can repair a node in an average of four hours, two repair people working independently on two failed nodes will generate on the average two repairs in four hours or an average of one repair every two hours.



Parallel Repair Failure State Diagram
Figure 3

Following the sequential repair analysis, we have:

State Transition Equations

State 0 $\quad p_0 n/\text{mtbf} = p_1/\text{mtr}$

State 1 $\quad p_1[1/\text{mtr} + (n-1)/\text{mtbf}] = p_0 n/\text{mtbf} + 2p_2/\text{mtr}$

State 2 $\quad 2p_2/\text{mtr} = p_1(n-1)/\text{mtbf}$

State Probabilities (in terms of p_0)

$$p_0 = p_0$$
$$p_1 = \frac{n(\text{mtr})}{\text{mtbf}} p_0 = \frac{n(1-a)}{a} p_0$$
$$p_2 = \frac{n(n-1)\text{mtr}^2}{\text{mtbf}^2} p_0 = \frac{n(n-1)(1-a)^2}{2a^2} p_0$$

State Probabilities (in terms of *a*)

$$p_0 = \frac{a^2}{D}$$
$$p_1 = n\frac{a(1-a)}{D}$$
$$p_2 = \frac{n(n-1)}{2}\frac{(1-a)^2}{D}$$

where

$$D = a^2 + na(1-a) + \frac{n(n-1)}{2}(1-a)^2$$

© 2007 Sombers Associates, Inc., and W. H. Highleyman
www.availabilitydigest.com

Again, the system is down if it is in state S_2. Thus, the probability of failure of the system, *F*, is the probability that the system will be in state S_2, or

$$F = \frac{n(n-1)}{2}\frac{(1-a)^2}{D} \qquad (6)$$

and

$$A = 1 - F = 1 - \frac{n(n-1)}{2}\frac{(1-a)^2}{D} \qquad (7)$$

These are the true expressions for the availability of a single-spared n-node system that uses parallel repair. Note that they are equivalent to our intuitive Equation (2) except for the denominator *D*. *D* represents the error in our intuitive approach.

As discussed in the analysis of sequential repair, the accuracy of the intuitive approach of Equation (2) depends upon the nodal availability *a* being very close to one and upon the quantity *n*(1-*a*) being very small. Under these assumptions, *D* is approximately one; and the intuitive relationship holds.

Note that for a two-node system (*n*=2), *D* is, in fact, equal to one; and the intuitive Equation (2) is accurate.

Summary

Our intuitive expressions for system availability under different repair strategies are valid provided that the nodal availability is high and that the number of nodes in the system is modest. This is certainly the case for the redundant systems with which we are concerned.

Should these approximations not be valid, then Equations (4) and (5) should be used to calculate the availability of a single-spared system under sequential repair. Equations (6) and (7) should be used to calculate the availability of a single-spared system under parallel repair.

© 2007 Sombers Associates, Inc., and W. H. Highleyman
www.availabilitydigest.com