the Availability Digest

Calculating Availability – Heterogeneous Systems Part 3

June 2008

In Parts 1 and 2 of this series,[1] we reviewed some elementary concepts of probability theory and applied them to the analysis of the availability of active/active systems and active/standby systems. We considered not only user downtime due to dual-system failures but also user downtime due to failover times and failover faults. We extended these results to include the case in which the nodes in the system have different availabilities.

Here in Part 3, we show how to calculate the availability of a complex system comprising parallel and serial-node configurations. This is a step-wise analysis in which parallel subsystems and serial subsystems are iteratively reduced to single nodes until only one – the system node – remains.

But first, let us briefly review where we left off.

A Review of Parts 1 and 2

State Probabilities

When considering system availability, it is convenient to think of a system being in one of several states. For instance, the system may be up, it may be down, or it may be failing over to a backup system. More to the point, we are not specifically interested in system availability; we are interested in *user* availability. We take this view because not all failures affect all users. For instance, in an *n*-node active/active system, should a node fail, only 1/*n* of the users on the average need to be failed over to surviving nodes.

We are therefore interested in the probability that a system will be in any one particular state (the states are mutually exclusive as the system can only be in one state at a time). System states can be defined by three Boolean functions – AND, OR, and NOT. For instance, a two-node system is up if Node 1 OR Node 2 OR both are up. Alternatively, the system is up if Node 1 AND Node 2 are NOT down.

Letting p(x) = the probability that the state x is true, then

- The AND operator implies multiplication. The probability that *x* AND *y* are true is $p(x)p(y)$.
- The OR operator implies addition. The probability that *x* OR *y* is true is $p(x) + p(y)$.[2]
- The NOT operator implies the complement. The probability that *x* is NOT true is $1 - p(x)$.

[1] Calculating Availability – Heterogeneous Systems Part 1, *Availability Digest*; March 2008.
Calculating Availability – Heterogeneous Systems Part 2, *Availability Digest*; May 2008.

[2] This is true if *x* and *y* are mutually exclusive. If they are not, the probability that x OR *y* is true is $p(x) + p(y) – p(x)p(y)$.

Applying these relationships to the previous statements, let *a* be the availability of either node in a two-node system. That is, *a* is the probability that a node is up. The probability that a node is down is the probability that it is NOT up (the complement function):

probability that node is up = *a*
probability that node is down = (1-*a*)

Considering only the case of a dual-node failure taking down the system, the probability that the system is up is the probability that Node 1 and Node 2 are up OR the probability that Node 1 is up AND Node 2 is down OR the probability that Node 1 is down AND Node 2 is up:

probability that system is up = $a^2 + a(1-a) + (1-a)a = a^2 + 2a(1-a)$

Likewise, the system is up if Node 1 AND Node 2 are NOT down:

probability that system is up = $1 - (1-a)^2$

These two expressions are identical as a little algebra will show.

System Availability

We consider that a dual-node system is down under three conditions:

a) Both nodes have failed.
b) One node has failed, and the users are being failed over to a backup node.
c) One node has failed, and the failover has failed (a failover fault).

Let

mtbf	be the mean time before failure for a node.
mtr	be the mean time to repair a node.
a	be the availability of a node = mtbf/(mtbf+mtr).
mtfo	be the mean time for a failed node to failover to a backup node.
d	probability that a failover will fail (the probability of a failover fault).
A	system availability (probability that system is up).
F	probability that system is down = 1-*A.*

In Part 2, we showed that

$$F = (1-a)^2 + (1-a)\frac{\text{mtfo}}{\text{mtr}} + (1-a)d \qquad (1a)$$

Though not done in Part 2, Equation (1a) can be extended to the case of an *n*-node system with *s* spares:

$$F = f(1-a)^{s+1} + (1-a)\frac{\text{mtfo}}{\text{mtr}} + (1-a)d \qquad (1b)$$

where

f = number of ways that *s*+1 nodes can fail out of *n* nodes.

Equation (1a) can be written as

$$F = (1-a)\left[1-\left(a-\frac{\text{mtfo}}{\text{mtr}}-d\right)\right] = (1-a)(1-a') \qquad (2)$$

© 2008 Sombers Associates, Inc., and W. H. Highleyman

where

$$a' = a - \frac{mtfo}{mtr} - d \qquad (3)$$

This leads to an important observation. The system acts as a heterogeneous system with the first node to fail having an availability of *a* and the surviving node having a reduced availability of *a*':

> *Following the failure of one node, failover time and failover faults cause the system to behave as if it comprises a remaining node with decreased availability.*

Heterogeneous Systems

If the two nodes in an active/active system have different availabilities, say a_1 for Node 1 and a_2 for Node 2, then Equation (1a) becomes

$$F = (1-a_1)(1-a_2) + \left(1 - \frac{a_1 + a_2}{2}\right)\frac{mtfo}{mtr} + \left(1 - \frac{a_1 + a_2}{2}\right)d \qquad (4)$$

Complex Systems

Systems can be more complex than the parallel systems considered in Part 2. There may be a network of subsystems in a serial/parallel configuration. A parallel subsystem comprises a set of nodes that act as backups to each other. It will continue to function even if one or more nodes should fail (depending upon the level of sparing). For a dual-node, singly-spared parallel subsystem in which the two nodes have availabilities of a1 and a2 respectively, the availability of the parallel subsystem is the probability that Node 1 AND Node 2 will NOT be down:

Availability of dual-node, singly-spared parallel subsystem = 1 – (1-a1)(1-a2) (5)

A serial subsystem requires that all nodes in the subsystem be operational in order for the subsystem to be operational. Should any one node fail, the subsystem will fail. For instance, in a two-node serial subsystem in which the nodes have availabilities of a1 and a2 respectively, the availability of the serial subsystem is the probability that Node 1 AND Node 2 will be up

Availability of dual-node, serial subsystem = a1 x a2 (6)

The availability of a complex system can be analyzed by first calculating the availability of each of the parallel subsystems in the complex and by replacing each with a single node with the equivalent availability. Then any series of subsystems are each replaced with a single node with the equivalent availability. More parallel subsystems may be created that are resolved followed by more serial subsystems. This process continues until the system has been reduced to a single node with its calculated availability, which is the system availability.

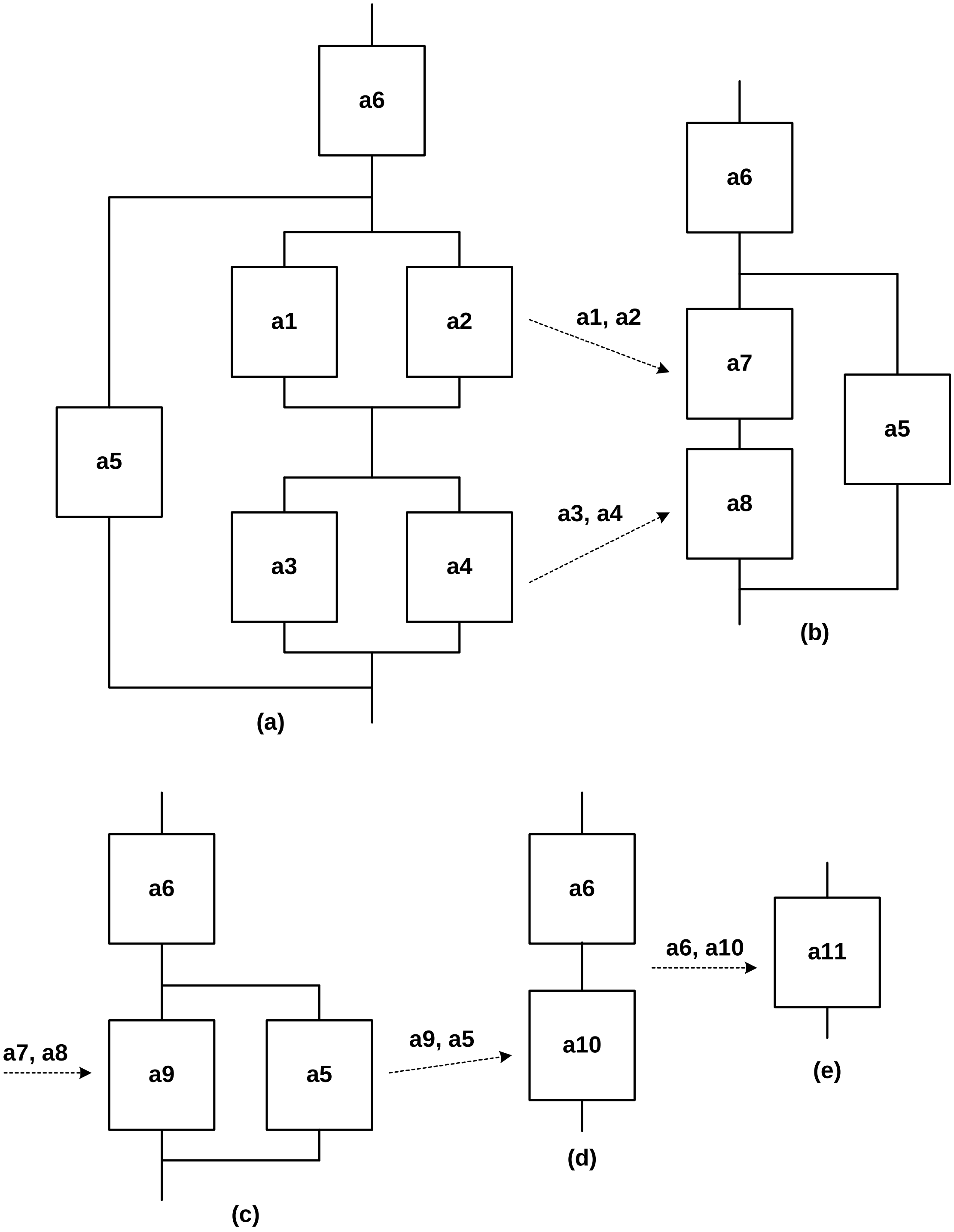
For instance, consider the system of Figure 1a. It comprises six nodes with availabilities of a1 through a6. We start by noting that there are two parallel subsystems of two nodes each. The availability of the a1/a2 parallel subsystem is

a7 = 1 – (1-a1)(1-a2)

The availability of the a3/a4 subsystem is

a8 = 1 – (1-a3)(1-a4)

© 2008 Sombers Associates, Inc., and W. H. Highleyman
www.availabilitydigest.com



A Complex System
Figure 1

© 2008 Sombers Associates, Inc., and W. H. Highleyman

We replace these two parallel subsystems with single nodes with availabilities of a7 and a8, as shown in Figure 1b. This now exposes a two-node serial subsystem with availabilities of a7 and a8. Its availability is

$$a9 = a7 \times a8$$

The serial subsystem is replaced with a single node with availability a9, as shown in Figure 1c. This leads to another two-node parallel subsystem with availabilities of a5 and a9. The availability of this parallel subsystem is

$$a10 = 1 - (1-a5)(1-a9)$$

Replacing this parallel subsystem with a single node with availability a10 gives the configuration shown in Figure 1d. This again is a two-node serial subsystem in which the nodes have availabilities of a6 and a10. Its availability is

$$a11 = a6 \times a10$$

This has reduced the complex system to a single node, and a11 is the availability of the entire system of Figure 1a.

This example has ignored one complexity, and that is failover. If in a parallel subsystem, there is a failover time following the failure of a node, the failover downtime must be accounted for when calculating the availability of the parallel subsystem. Availability in the presence of failover is the subject of Equations (1) through (4) above.

We will consider failover in a complex system in the example that we analyze in Part 4 of this series. This example is that of an active/active system backed up by a remote standby system. It is a particularly important configuration when active/active systems use synchronous replication via network transactions and are therefore limited in the distance that may exist between the active/active nodes due to application latency. For disaster tolerance, what is often provided is a backup system a great distance away using asynchronous replication.

Summary

By breaking down a complex system into a series of parallel and serial subsystems, the availability of the overall system can be easily determined. A parallel subsystem comprises a set of nodes that can withstand the failure of one or more nodes. A serial subsystem will fail if any node in the series fails.

The first step is to resolve the availability of parallel components in which one or more may fail and yet the system is still operational. If there is a failover time required, the downtime during failover and the probability of a failover fault must be considered. Each parallel subsystem is replaced with a single node with the availability of the parallel subsystem.

This may result in the presence of one or more serial subsystems. Each of these is analyzed and replaced with an equivalent single node. Once again, further parallel subsystems may result and then further serial subsystems as the parallel subsystems are resolved. This procedure continues iteratively until the system has been reduced to a single node.

In Part 4, we will apply this technique to analyze by example an active/active system backed up by a hot standby system.

© 2008 Sombers Associates, Inc., and W. H. Highleyman