the Availability Digest

Calculating Availability – Heterogeneous Systems Part 2

May 2008

In Part 1 of this series,[1] we reviewed the elementary probability concepts that apply to calculating availability. In this Part 2 of our series, we first briefly review these concepts; and we then apply them to redundant systems in which the nodes do not have the same availability (an assumption that we have made up to this point). We consider the impact on system availability of non-symmetrical failover times, failover faults, and dual-node failures.

A Review of Probability Fundamentals

When analyzing availability, we are concerned with binary states. Either a system is up, or it is down.

A binary state can be defined by three logical Boolean functions – AND, OR, and NOT. For instance, a two-node system is up if Node 1 is up OR if Node 2 is up. This is equivalent to saying that a two-node redundant system is up if Node 1 AND Node 2 are NOT down.

The following statements summarize the Boolean probability relationships. Let p(*x*) be the probability that *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*).

For instance, consider a two-node system with each node having an availability of *a*. That is, the probability that a node will be up is *a*. The probability that the system is up is the probability that Node 1 AND Node 2 are up OR that Node 1 is up AND Node 2 is down OR that Node 2 is down and Node 1 is up. Using the above rules, this is

$$a^2 + a(1-a) + (1-a)a = a^2 + 2a(1-a) \qquad (1)$$

Alternatively, the probability that the system is up is the probability that both nodes are NOT down. This can be expressed logically as NOT (Node 1 is down AND Node 2 is down), which results in

$$1 - (1-a)(1-a) = 1 - (1-a)^2 \qquad (2)$$

A little algebraic manipulation will show that Equations (1) and (2) are equivalent.

[1] Calculating Availability – Heterogeneous Systems Part 1, *Availability Digest*; March 2008.
[2] This is true if *x* and *y* are mutually exclusive. If they are not, then the probability that x or *y* is true is p(*x*) + p(*y*) – p(*x*)p(*y*).

One must be careful when making simplifications. In Equation (1), it is tempting to say that the second term can be ignored if *a* is very close to one; and therefore (1-*a*) is very small. However, this leads to a system availability of a^2. This is an absurdity since it says that adding a spare node reduces the system availability.

Application to Heterogeneous Systems

Let us now apply these concepts to some cases of heterogeneity. For simplicity's sake, we will limit the examples to dual-node, single-spared configurations. That is, the system remains up in the presence of any single-node failure but is down if both nodes fail.

Failover

When we consider failover, we realize that there are two cases in which users will be without service:

- if both nodes fail, as we analyzed above.
- when one node fails until the other node is brought into service. This is the failover time.

We analyzed the impact of failover time extensively in our article entitled Calculating Availability – Failover in our February, 2007, issue. Let us analyze failover for a dual-node, single-spared system using probability principles.

We define the following:

p(up) = *a* = the probability that a node will be up (node availability).
p(down) = (1 - *a*) = the probability that a node is down (that is, it is not up).
p(failover) = the probability that the system is down during a failover period.

We need to know what p(failover) is. Let

MTFO = the mean time to fail over (the average time that users are down while the backup is being brought into service).
mtbf = the mean time between failures for a node.
mtr = the mean time to repair a node.
A = system availability (the probability that a user is receiving service).

Assume that we are dealing with a two-node active/active system, and therefore there will be two node failures on the average during each mtbf period. Thus, the probability that the system will be down during failover is

$$\frac{\text{MTFO}}{\text{mtbf}/2}$$

However, from the users' viewpoint, only half of the users are affected. Therefore, the effective availability is half that above, and

$$\text{p(failover)} = \frac{\text{MTFO}}{\text{mtbf}} \qquad \text{for an active/active system} \qquad (3a)$$

Noting that *a* = 1 – mtr/mtbf, we can express mtbf as

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

and

$$\text{p(failover)} = \frac{\text{MTFO}}{\text{mtbf}} = \frac{\text{MTFO}}{\text{mtr}}(1-a) \qquad (3b)$$

The probability that the system will be down is the probability that Node 1 AND Node 2 have failed OR that a failover is in process:

$$\text{p(user being down)} = (1-a)^2 + \frac{\text{MTFO}}{\text{mtbf}} \quad \text{for active/active systems} \qquad (4a)$$

Note that this is true for active/active systems. For an active/standby system, only the failure of the primary system will cause a failover. Therefore, the factor of mtbf/2 becomes mtbf. However, now all users are affected, so that the probability of user downtime is increased by a factor of 2. The result is that

$$\text{p(user being down)} = (1-a)^2 + \frac{\text{MTFO}}{\text{mtbf}} \quad \text{for active/standby systems} \qquad (4b)$$

The probability of user downtime is the same for active/active systems as it is for active/standby systems, though for different reasons.

Let us look at some practical cases. Consider a dual-node, singly-redundant system with nodes that have an mtbf of 4,000 hours (about 6 months) and an mtr (mean time to repair) of four hours. This results in a nodal availability of three nines (1 - 4/4000 = 0.999). According to Equation (2), this dual-node system will have an availability of six nines if failover time is ignored.

Case 1: Cold Standby

Following a node failure, a cold standby system must have its database and applications loaded and the applications started. Let us assume that it takes four hours to bring up a cold standby. In this case,

$$\text{p(user being down)} = .001^2 + \frac{4}{4000} = 10^{-6} + .001 \approx .001$$

$$A \approx 0.999$$

In this case, failover time is predominant. It has reduced a system with six 9s availability to one with three 9s availability, just that of a single node.

Case 2: Hot Standby

A hot standby system is kept ready to take over by keeping its database synchronized with the active system via data replication. Let us assume that it takes six minutes (0.1 hours) to bring the hot standby online. Then

$$\text{p(user being down)} = .001^2 + \frac{0.1}{4000} = 10^{-6} + 2.5\text{x}10^{-5} = 2.6\text{x}10^{-5}$$

$$A = 0.999974$$

Failover time is still predominant, but the system availability is now more than four 9s, much better than that of a single node.

Case 3: Active/Active

In an active/active configuration, both nodes are actively running the application. Therefore, all that is required following a node failure is to switch users from the failed node to the surviving node, a process that can be accomplished in seconds.

Let us assume that switchover takes three seconds. That means that following a node failure, half of the users will be down for three seconds. In this case,

$$\text{p(user being down)}=.001^2 + \frac{3/3600}{4000} \approx 10^{-6} + .2\text{x}10^{-6} = 1.2\text{x}10^{-6}$$

$$A \approx 0.9999988$$

The system availability is a little less than six 9s, close to what it would be if failover time were zero. In this case, failover time is not a major factor in system availability.

This is the philosophy behind active/active systems – let it fail, but fix it fast.

Note that if HP Nonstop systems with four 9s availability were to be used for the active/active nodes, the failure probability due to a dual-node failure would be 10^{-8}. In the above failover case, the system availability would be $2.1\text{x}10^{-7}$, leading to a system availability of a little less than seven 9s. Failover time in this case is important and can be the limiting factor in system availability.

Failover Faults

A major problem in redundant systems is failover faults. A failover fault is the failure of the backup system to take over. Failover is a very complex process and is difficult to test. Therefore, the probability that a failover will fail may not be insignificant.

If failover faults are considered, there will be a dual-system failure in an active/active system if one node fails AND if the second node fails while the first node is down OR if one node fails and there is a failover fault. As is the case with failovers, in the event of a failover fault, only half of the users will be effected. Let

p = the probability that a failover will fail (a failover fault).

The probability that one of the two nodes will fail is $2(1 - a)$. The probability that there will be a failover fault is therefore the probability that there will be a single node failure followed by a failover fault, or $2(1 - a)p$. The probability that there will be a dual node failure is $(1 - a)^2$. Noting that only half of the users will be affected by a failover fault, the probability that users will be down is

$$\text{p(user being down)} = (1-a)^2 + (1-a)p = (1-a)p \qquad (5)$$

As argued above, for an active/standby system the result is the same since failover faults are a problem only if the active system fails, but all of the users are affected.

Let us consider a dual-node, singly-redundant system of nodes with three 9s availability and with a failover-fault probability of 1% (that is, one out of 100 failovers will fail). In this case, from Equation (5),

$$p(\text{user being down}) \approx 0.001 \text{ x } .01 = 1\text{x}10^{-5}$$

The system availability due to dual-node failures has been reduced from six 9s to a five 9s. A 1% probability of failure has made the system ten times less reliable! Following the failure of one node, Equation 5 states that the effective availability of the surviving node is only (*1 – p*) rather than *a*.

Combining the above results, we obtain the relationship expressing the probability of user failure, *F*, when dual-node failures, failover time, and failover faults are considered:

$$F = (1-a)^2 + \frac{MTFO}{mtbf} + (1-a)p = (1-a)^2 + \frac{MTFO}{mtr}(1-a) + (1-a)p \qquad (6a)$$

where Equation (3b) was used. The first term is the probability of a dual-node failure. The second term is the probability that the system will be down during failover. The third term is the probability that the system will be down due to a failover fault.

We can rewrite Equation (6a) as

$$F = (1-a)\left[1-\left(a - \frac{MTFO}{mtr} - p\right)\right] = (1-a)(1-a') \qquad (6b)$$

The failure probability of the first node is (1 – *a*). The probability of failure of the second node is (1 – *a*'), where *a*' is the nodal availability *a* reduced by the failover time and failover fault probability:

$$a' = a - \frac{MTFO}{mtr} - p$$

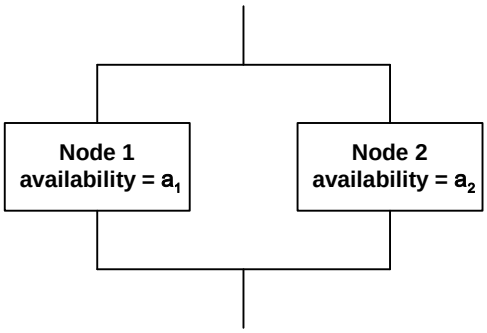
Thus, the system acts as a heterogeneous system with the first node to fail having an availability of *a* and the other node having a reduced availability of *a*'. This leads to the following observation:

> *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.*

This same reasoning can be extended to an *n*-node system.[3]

Heterogeneous Nodes

In many cases, the nodes in a system are not homogeneous. They may be different systems, or they may be affected by different environmental conditions. In any event, they have different availabilities.



Let the availabilities of the two nodes in a singly-redundant system be a_1 and a_2, respectively. Then, if

[3] See Rule 24 in Chapter 5, The Facts of Life, *Breaking the Availability Barrier: Survivable Systems for Enterprise Computing*, AuthorHouse; 2004.

only dual-node failures are considered, the system availability is

$$A = 1 - (1 - a_1)(1 - a_2) \quad (7)$$

For instance, if the availability of Node 1 is four 9s, and if the availability of Node 2 is three 9s, the availability of the system is

$$A = 1 - 10^{-4} \times 10^{-3} = 1 - 10^{-7}$$

or seven 9s.

A useful application of this result is in the analysis of the impact of environmental faults. An environmental fault is any happenstance external to the computing system that will cause a node outage. Environmental faults can range from computer-room fires to hurricanes, earthquakes, or terrorist attacks. These faults are asymmetric in that they affect nodes in a geographically distributed system differently. Thus, the nodes are heterogeneous in that they have different overall availabilities.

Environmental-fault calculations are inherently unreliable because, for instance, it is hard to state a probability that an earthquake will occur at a given site. However, one can take educated guesses that should err on the conservative side. For instance, one might reasonably assume that a hurricane will damage a data center in Florida once every twenty years. Then the probability that the data center will be damaged in the next year is 5%.

In addition, we must know how long it will take to recover from such a failure. Recovery may include finding another building, ordering and installing equipment, and bringing up the system. This might take days, weeks, or more.

Let us define the following:

mtbf = mean time between node failures due to hardware or software faults.
mtr = mean time to repair a node following a hardware or software fault.
mtbe = mean time between environmental faults.
mte = mean time to restore a node to service following an environmental fault.

Then the probability of a node failure, *f*, is the probability that it will be down due to a system fault OR that it will be down due to an environmental fault:

$$f = (1 - a) = \frac{mtr}{mtbf} + \frac{mte}{mtbe} \quad (8)$$

For instance, consider a node with an availability of three nines (that is, mtr/mtbf = 10^{-3}). Assume that the node is in a hurricane region in which it is estimated that a hurricane strong enough to do significant damage will occur every 160,000 hours (about 20 years). Furthermore, it will take 700 hours (about one month) to return the site to service. Then the availability of the node is reduced from three nines to

$$f = 10^{-3} + \frac{700}{160{,}000} = 10^{-3} + 4.4 \times 10^{-3} = 5.4 \times 10^{-3}$$

$$a = .9946$$

The other node in the redundant system is in an area not impacted by any environmental fault. It is an identical system with three 9s availability. The system availability is then, from Equation (7),

$$A = 1 - (5.4x10^{-3} x10^{-3}) = .9999946$$

or a little over five nines rather than six 9s as it would be in the absence of environmental faults.

This analysis can be extended to failovers and failover faults in active/active systems by noting that either of these occurs only after a first node has failed. The probability that a first node will fail is the probability that Node 1 will fail OR that Node 2 will fail. In a homogeneous system, this probability is 2(1 – a).[4] In a heterogeneous system, Node 1 will fail with a probability of (1 – a_1); and Node 2 will fail with a probability of (1 – a_2). Therefore, the probability that Node 1 OR node 2 will fail is (1 - a_1)+(1 – a_2). The average probability of a first node failure is then

$$\frac{1}{2}[(1 - a_1)+(1 - a_2)] = \left(1 - \frac{a_1 - a_2}{2}\right)$$

This relationship replaces the term (1 – *a*) in Equation (6a), which can then be rewritten as

$$F = (1 - a_1)(1 - a_2) + \frac{MTFO}{mtr}\left(1 - \frac{a_1 + a_2}{2}\right) + \left(1 - \frac{a_1 + a_2}{2}\right)p \quad \text{for active/active systems} \quad (9)$$

The previously referenced Chapter 5 in *Breaking the Availability Barrier* expands this result for heterogeneous nodes to cover an *n*-node system and to include the recovery time of the system for those cases that require recovery activities following the repair of a node, such as database resynchronization.

Summary

The calculation of system availability goes beyond just multiple node failures. The time that users are denied service while the system is failing over to a backup node must be considered, and this time is often a predominate factor. Furthermore, the failover process itself may fail, resulting in a dual-node failure even though one of the nodes is perfectly good.

The analysis is made more complex if the nodes are heterogeneous, each having a different availability. However, the extension of the availability equation to cover this case is straightforward.

In our final article on this topic, we will extend this analysis to cover systems that can fail if a single component fails. For instance, a set of redundant computers might be accessing a single database, which itself might be redundant.

[4] Since these events are not mutually exclusive, the probability that Node 1 OR Node 2 will fail is really 2(1 – *a*) – (1 – *a*)^2 (see footnote 2). For (1 – *a*) << 1, which is the case in which we are interested, the expression holds.