Running Head: UNDERSTANDING FAILOVER AND RECOVERY TIMES

Increasing System Availability: Understanding Failover and Recovery Times

Jered McClure

Walden University

Increasing System Availability: Understanding Failover and Recovery Times

 Taking into account the failover time of any system will greatly increase the accuracy of the availability index of that system. In other words, without knowing the true time it will take to failover to a given redundant system, the availability of that system is questionable in the overall scheme of things. Given the knowledge of how fast each node will fail over to its redundant counterpart, the real estate system’s (from assignment two) availability can be more accurately defined. This in turn will allow the organization to better utilize its resources to increase the overall system availability.

Below, in figure 1, is a diagram outlining each node’s availability in the real estate system, specifically mentioning the mean time to recover (mtr), the mean time to fail over (mtfo) and the probability of a failover fault (d). Note that it now has a redundant LAN which was missing before:



Figure

 With this data we are able to begin the process of determining how the failover time, in conjunction with the time to recover each node, affects the overall system availability. This is begun by first reducing the server redundancy on the left. The failure rate of the primary server is multiplied by the failure rate of the secondary server, and this value is subtracted from one. However, the failure rate of the second server is its availability, minus the quotient of the mean time to fail and the mean time to recover, minus the probability of the failover failing (e.g. the primary fails, then the failover fails as well).

 Below, in figure 2, shows this reduction of the redundant servers and their total availability. Moreover, the LAN has been reduced. As its MTFO and MTR are both zero, this is a simple matter of finding the availability of the redundant pair:



Figure

From here, both redundant systems are reduced. Note that the mtfo, d, and mtr are carried up into the reduced phase as their values will directly affect the availability in the next reduction:



Figure

 Just as the two servers were reduced, taking into account failover and recover times, the overall system redundancy must also take these into account when failing over the WAN. This is where the true nature of knowing the failover time frames can make a difference in where system changes can be made. Because the failover time is two hours, the availability of the redundant system is actually reduced quite drastically, from 0.9965 to 0.5535. In other words, the secondary redundant passive system has an availability index less than one nine (approximately 1180 hours a year, or 49.2 days).



Figure

 Bringing this back around to the overall system availability, the above metric drops the availability from a previously calculated four nines to just over three nines. This is a massive change! Furthermore, the previous index was based on the lack of dual LAN configuration, without the knowledge of the mtfo and mtr, the organization has wasted quite a bit of money on implementing the additional LAN, when it should have put capital into reducing the mtfo of the WAN. Essentially, the knowledge of failover and recovery times is absolutely vital in locating failure points. Without this knowledge, an organization is simply making a blind guess as to where their real problems exist.



Figure

Reference

Availability Digest. (2010, October). *Simplifying Failover Analysis - Part 1*. Retrieved May 2, 2014, from Availability Digest: http://www.availabilitydigest.com/public\_articles/0510/failover\_analysis.pdf

Stern, E. M. (2003). *Blueprints for High Availability* (2nd ed.). Indianapolis: Wiley Publishing, Inc.