Measuring Availability Targets

By: Harris Kern's Enterprise Computing Institute

\To accurately measure systems availability as experienced by end users, you must first thoroughly understand the system's configuration. These include all the components and resources used by the application, both local and remote; and the hardware and software components required to access those resources. The next step is to monitor all these components for outages, then calculate end-to-end availability. In this article we will focus on how to do these calculations.

Quantifying Availability Targets

To quantify the amount of availability achieved, we must perform some calculations:

Committed hours of availability (A). This is usually measured in terms of number of hours per month, or any other period suitable to your organization.

Example: 24 hours a day, 7 days a week = 24 hours per day x 7 days x 4.33 weeks per month (average) = approximately 720 hours per month

Outage hours (B). This is the number of hours of outage during the committed hours of availability. If high availability level is desired, consider only the unplanned outages. For continuous operations, consider only the scheduled outages. But for continuous availability, you should consider all outages.

Example: 9 hours of outage due to hard disk crash, 15 hours of outage for preventive maintenance

Next you can calculate the amount of availability achieved as follows:

Achieved availability = ((A-B)/A)*100%)

For the statistics in the examples above, here's each calculation:

High availability = ((720-9)/720)*100% = 97.92% availability

Continuous operations = ((720-15)/720)*100% = 98.75% availability

Continuous availability = ((720-24)/720)*100% = 96.67% availability

When negotiating an availability target with users, make them aware of the target's implications. Here is a table of availability targets versus hours of outage allowed for a continuous availability level requirement.

Continuous Availability Target	Hours of Outage Allowed Per

	Month
99.99%	0.07 hours
99.9%	0.7 hours
99.5%	3.6 hours
99.0%	7.2 hours
98.6%	10.0 hours
98.0%	14.4 hours

It is important to recognize that numbers like these can be difficult to achieve, since time is needed to recover from outages. The length of recovery time correlates with the following factors:

Complexity of the system. The more complicated the system, the longer it takes to restart it. Hence, outages that require system shutdown and restart can dramatically affect your ability to meet a challenging availability target. For example, applications running on a large server can take up to an hour just to restart when the system was shut down normally - longer still if the system was terminated abnormally and data files must be recovered.

Severity of the problem. Usually, the greater the severity of the problem, the more time is needed to fully resolve the problem, including restoring lost data or work done.

Availability of support personnel. Let's say that the outage occurs after office hours. A support person called in after hours could easily take an hour or two simply to arrive to diagnose the problem. You must allow for this possibility.

Other factors. Many other factors can prevent the immediate resolution of an outage. Sometimes an application may have an extended outage simply because the system can't be put offline while applications are running. Other cases may involve the lack of replacement hardware by the system supplier, or even lack of support staff. We have seen many availability targets missed simply because a system supplier could not give due attention to the problem, and no backup system supplier existed.

Be aware, you won't get precise measurements of *every* user's availability experience. That's not realistic. But recognize that users do have availability requirements to which you must pay attention. Don't get too dependent on technical measurements for rating your performance. In the end what matters most is that users are happy with the service that the IT organization provides.