

HRG Insight:

Duration and Frequency of Unplanned Outages

Harvard Research Group provides custom research to its clients in select areas including High Availability and Fault Tolerant systems. One of the on-going activities in this practice to survey users of higher availability systems and determine their needs, preferences, and the manner in which these systems are being utilized to support their business.

Harvard Research Group is pleased to share with you some of the results of these surveys:

The respondents represented a statistical sampling categorized as:

- ✓ Corporate demographics: in 1998-1999 the mid two-thirds of the sample had approx. 300 to 20,000 employees and a total annual corporate revenue of \$60 to \$10 billion. In 2000, the mid two-thirds of the sample had approx. 250 to 15,000 employees and a total annual corporate revenue of \$50 million to \$10 billion.
- ✓ Moderate to larger IT installations, i.e. at least 50% have ten or more servers. (Some of these may be divisional/subsidiary operations)
- ✓ All running production systems, minimal development efforts included
- ✓ Predominant server manufacturers were IBM, HP, Compaq, Sun, and Dell (not necessarily ordered by representation)
- ✓ In 1998 and 1999, operating system representation was approximately UNIX 30%, NT 30%, and proprietary O/S 35%. In the 2000 results, UNIX was approaching 40%, NT remained relatively constant, and proprietary systems dropped to approximately 28%.
- ✓ Diverse usage across industries and applications
- ✓ Geographically dispersed across the U.S.

Analysis of Unplanned Server Outages

	1999	2000
Ave. Outages per Year	2 - 3	2 - 3
Outages per Year Central 50%	1 - 5	1 - 5
Average Duration of Outage (min.)	60	60
Range of Central 80% (min.)	10 - 240	5 - 480
'Acceptable' Downtime at 50% of Users (min.)	<120	<120
'Acceptable' Downtime at 33% of Users (min.)	< 30	< 30

Note: Average Duration of Outage expressed as the mode, i.e. the most common response.

Key Observations and Conclusions

During the observation period 1999 – 2000, the results of the survey were extremely consistent with most respondents having experienced two to three unplanned outages during the past year. The average duration of each was approximately 60 minutes.

From prior surveys, the vast majority of respondents, approximately 75%, indicated that they measure outage times for servers. A slightly lower percentage, but still a majority, measured other outages i.e. network.

As one might expect, there were wide variations in defining what constituted a ‘tolerable’ outage. From prior surveys, nearly fifty percent of the respondents indicated an outage of approximately two hours or more would severely impact operations. Nearly a third indicated that an outage of 30 minutes would be significantly disruptive. With the ever increasing dependence upon systems and interdependence of servers, one should anticipate the ‘level of tolerance’ continue to diminish.

Discussion

Harvard Research Group conducts surveys with users of high availability systems on a regular basis. The surveys cover a wide range of applications and industries across the U.S. These respondents indicated a relatively consistent level of unscheduled downtime (system outages) over the past two years. The average user has experienced two to three outages with an average duration of approximately 60 minutes.

Approximately a third of the respondents indicated that an outage in excess of thirty minutes would be disruptive to operations. In looking at the duration of outages, the only central eighty percent was considered in order to eliminate the skew. The range was from just a few minutes to four or more hours with an average/mode of sixty minutes. Consequently, one can conclude that the average outage is causing some significant level of disruption to operations. Furthermore, the average respondent is experiencing this level of disruption two to three times per year.

The implications and recommendations are seemingly clear:

1. In general, it is most likely that some additional expenditures on high availability are appropriate based upon the above findings.
2. Outages should be diligently monitored and recorded in all aspects, i.e. duration, users and applications affected, cause and remedy applied, as a component in a substantiated cost/benefit analysis to upper management toward incremental enhancements.
3. With the interdependence of systems including major portions of the customer base, supply chain and internal organizations, the cost/benefit analysis should reflect the global enterprise and its external contacts, not merely the immediate user community.
4. IT environments continue to grow dramatically as a consequence of the influence of the Internet. One must take a proactive approach to the analysis, i.e. how will the user community, applications, and dependencies change in the intermediate term in order to plan more effectively.

The total cost of downtime will be further discussed in a future paper.