

IT Managers Weigh In On Application Performance

Peter Sevcik and Rebecca Wetzel
NetForecast, Inc.

Business Communications Review, April 2006

With trends such as datacenter consolidation and multimedia convergence, IT managers are more concerned than ever about whether their networks are actually delivering application performance that satisfies their end users. Lots of new technologies have come on the market promising improved application performance, but it's difficult to know what solutions work in the real world, and which technologies solve bona fide user problems.

So to cut through the hype and get to the real skinny about what enterprises are thinking and doing about networked application performance, we surveyed BCR readers about their satisfaction with networked application performance, their ability to measure and/or improve it, and their self-perceived effectiveness on the application performance front.

The survey offers up a “from-the-trenches” view of four aspects of networked application performance: availability, accessibility, response time, and efficiency. Availability reflects network connectivity and service continuity, as well as server uptime and recovery speed, while accessibility reflects users' ability to access applications through the network and the application's ability to operate over the network (e.g., not to hang when accessed remotely or return 400-series Web errors, etc.). Response time reflects the end-user's “thumb twiddling time” while waiting for applications to reply to user interactions, and efficiency reflects such factors as WAN circuit and server utilization.

Respondents were asked to assess their ability to report on these four performance functions, and to rate their organization's effectiveness for each function. We also asked a number of other performance-related questions.

In all, 117 individuals completed the online survey between December 19, 2005 and March 6, 2006. Of those individuals, 68 percent list networking as their primary job responsibility, and 32 percent are primarily responsible for applications or application servers. The enterprises that employ the

respondents run the gamut in size from fewer than 50 to more than 40,000 employees, and from less than \$1 million annual revenue to more than \$1 billion. Ninety-five percent of the enterprises in the sample use a mix of private and public (Internet) network access and infrastructure, and four out of five rely primarily on a private corporate network.

Satisfaction With Application Performance

Overall, about two-thirds of respondents are satisfied with networked application performance, while one-third are dissatisfied. Half of the enterprises that rely on both private networks and the public Internet (directly or using a VPN) find that the Internet provides substantially less satisfactory performance for their critical applications than their private networks.

A bare majority says they're delivering on the promise, and can back up that claim with measurement statistics. The rest are still struggling.

Those who measure performance are consistently more satisfied and confident in overall performance than those who do not (more on this later), and applications managers are consistently more satisfied than are their network management

counterparts. Seventy-nine percent of applications managers describe themselves as somewhat or very satisfied with application performance, while only 60 percent of network managers are similarly satisfied with performance.

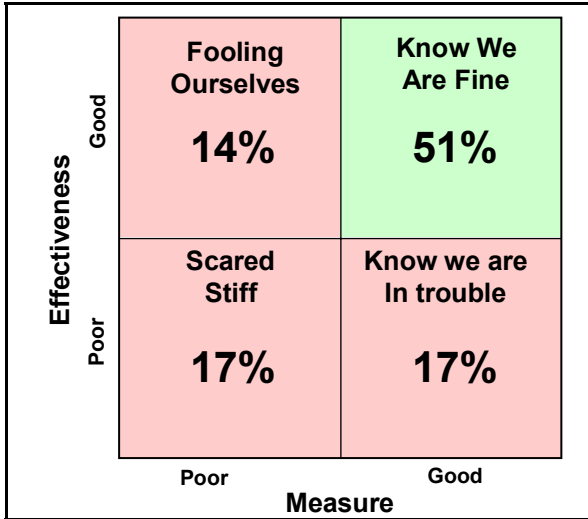
How Well Are Enterprises Doing?

To assess how well enterprises are doing, we asked respondents to rate their organizations' ability to report on each of the four performance functions (availability, accessibility, response time and efficiency), and rate their functional effectiveness in each. To show these survey results at a glance, we created a matrix summarizing respondents' “good” and “poor” ratings of their ability to measure and report on the four functions, versus their ability to effectively deliver on those functions. Figure 1 shows the matrix with its four quadrants and the distribution of survey responses within the quadrants.

We determined that the upper right quadrant is desirable because it is “safe,” while the other three

quadrants are “dangerous.” You will notice that on the matrix we characterize the likely mental state of IT managers in each of the quadrants as “fooling ourselves,” “scared stiff,” “know we are in trouble,” and “know we are fine.” Here is how we arrived at those characterizations.

- they believe they are doing fine but are perilously uninformed;
- they believe performance is poor but cannot prove it; or
- they have reports that show performance is poor.



Respondents in these dangerous quadrants have reason to worry and change what they are doing.

Figure 2 shows the function that respondents most often placed on each dangerous quadrant. Availability was most often placed in the safe quadrant. It appears that both networks and servers have high availability and enterprises measure this function well, so there is little need to improve availability relative to the other functions.

The functions that landed in the three dangerous quadrants constitute the primary pain points and, if addressed, are likely provide a significant benefit to enterprises. The results show where remediation needs to be applied.

Figure 1 – How Well is Performance Managed?

In the upper right “safe” quadrant, respondents perceive that they are doing a good job of measuring as well as delivering on the performance functions, and therefore are confident that they know they are fine. Respondents in the “dangerous” quadrants, on the other hand, have reason for alarm.

For example, respondents who perceive their enterprise is good at measuring yet poor at delivering performance know they are in trouble. Respondents who perceive they are good at delivering performance but are poor at measuring it cannot be fully confident that they are actually doing well, and therefore risk fooling themselves. Finally, those who perceive themselves as poor at measuring performance as well as poor at delivering it are apt to be scared stiff.

Our analysis shows that 51 percent of survey respondents are in the safe quadrant because they measure performance and believe they are delivering acceptable or better results. Conversely, the remaining 49 percent are in dangerous quadrants because:

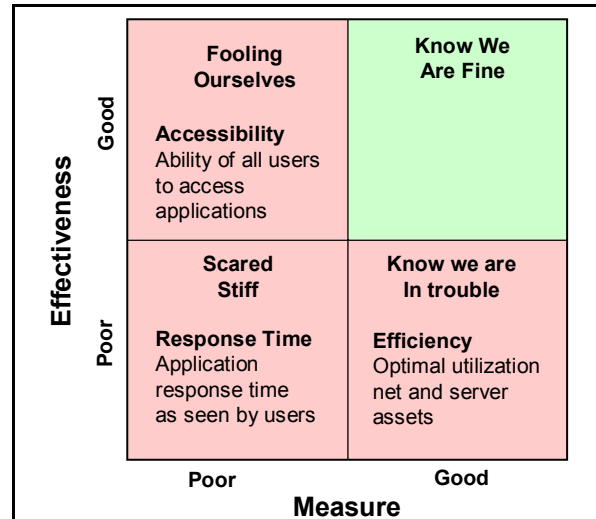


Figure 2 – What Needs Improving?

How To Move Out Of The Danger Zones

In the “Fooling Ourselves” quadrant, the primary pain point is the inability to measure accessibility (defined as the capacity for all users to access applications). Many enterprises in the survey deploy applications and networks with the assumption that the users simply are there and thus can reach whatever application they require. That is

how they can get away with claiming good accessibility without ever measuring it. This is a bad assumption. Users are often locked out of applications through infrastructure configuration errors and desktop changes. Furthermore, even though the assumption may have been true in the past, when IT knew exactly where and how every user was accessing the system, IT now knows less. Users may no longer just be on-site employees but contractors, partners, and employees outside the traditional office. IT can no longer safely assume that all the users are getting access when and where they need it. Enterprises must reach out to the users and actively measure accessibility.

The primary pain point in the “Scared Stiff” quadrant is the ability to effectively deliver good response time. User-level response time appears in this quadrant if a survey respondent admits to not measuring, yet knows there is a problem because users are complaining about poor performance, or the respondent fears poor performance is likely.

Enterprises with response time pain in this quadrant must first institute proper measurement practices that would move them to the right on the matrix. They may also need to invest in performance acceleration technology to move them up in the matrix, and into the safe quadrant. Lastly, the main pain point in the “We Know We Are In Trouble” quadrant is the ability to make

efficient use of resources. This group suffers from the inability to optimize server and network use despite the fact that they measure utilization well. Their response to this issue, where they admit the need to improve, indicates that they understand the problem. These enterprises need a strategy that includes management that will improve resource efficiency. It may also require negotiating with suppliers to purchase resources in a more granular way.

The Traffic Mix

As Figure 3 shows, client-server, email, and Web-enabled applications comprise the bulk of traffic in most of the respondents’ networks (between 15 and 20 percent on average). Web services, file transfer, voice over IP, peer-to-peer, and terminal-to-host make up the next tier in the traffic mix, each contributing 5-10 percent of traffic on average. Streaming video, videoconferencing and streaming audio are at the low end of the traffic mix, each contributing less than 5 percent of traffic to the average mix.

It’s surprising that despite a decade-long push to convert all applications to Web-based implementations, these still represent less than a quarter of the total traffic, and traditional client-server applications still constitute the largest single traffic class.

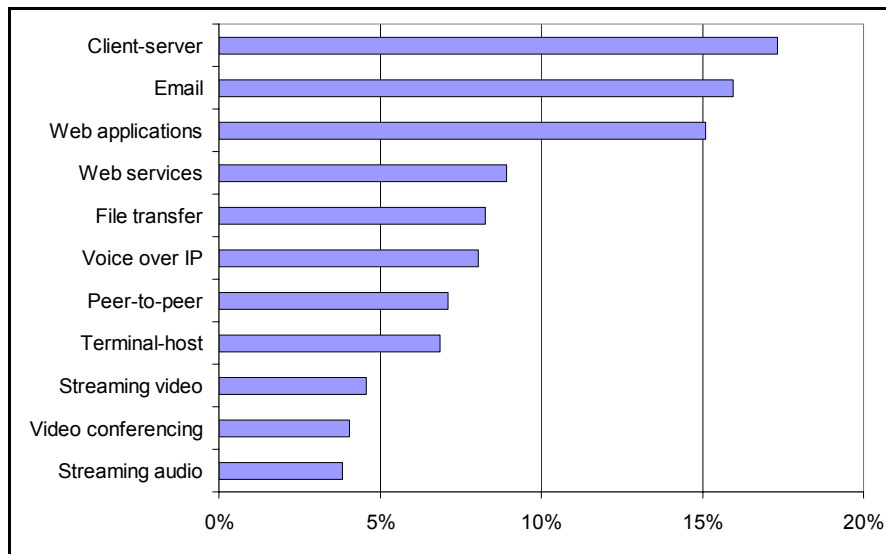


Figure 3 – Distribution of Application Traffic

Improvement Techniques In Use

Ninety percent of survey respondents use one or more techniques to improve performance. Figure 4 shows the techniques most used by respondents to improve performance

Several initial conclusions emerge from this data. Techniques for controlling performance (e.g., WAN traffic management and QoS) are more widely deployed than acceleration techniques (e.g., content caching, compression, and protocol optimization). Enterprises appear to favor solutions with the broadest applicability. Acceleration techniques mirror the types of protocols in use. Because Web applications comprise a relatively small percentage of respondent traffic, vendors who offer pure Web performance improvement solutions risk minor player status in the bigger performance picture.

To our surprise, we found no correlation between the number of technologies brought to bear on aspects of the performance problem and corresponding satisfaction with or confidence in performance. We suspect that despite the fact that many enterprises have purchased multiple performance enhancing products, most of these accelerate only a limited subset of critical traffic,

making their overall contribution marginal. Also many of these techniques are still in limited deployments as companies assess their effectiveness and deployment ramifications.

It is likely that control techniques (e.g., traffic management and QoS) are more widely deployed than other solutions because they are better understood, have a longer track record, and provide benefit to all applications. We note that control solutions can wring more out of existing assets (i.e., they make better use of bandwidth or server assets), so these devices are well suited to optimize efficiency.

We asked survey participants whether they would prefer to purchase performance measurement and improvement solutions in the form of products, services, or a mix of both. Fifty-eight percent of respondents prefer to buy products and install and operate them, 12 percent prefer to buy service only solutions, and 30 percent prefer a service/product mix. A combination of product and service is currently not offered and represents a very large overlooked market: about one third of the enterprises!

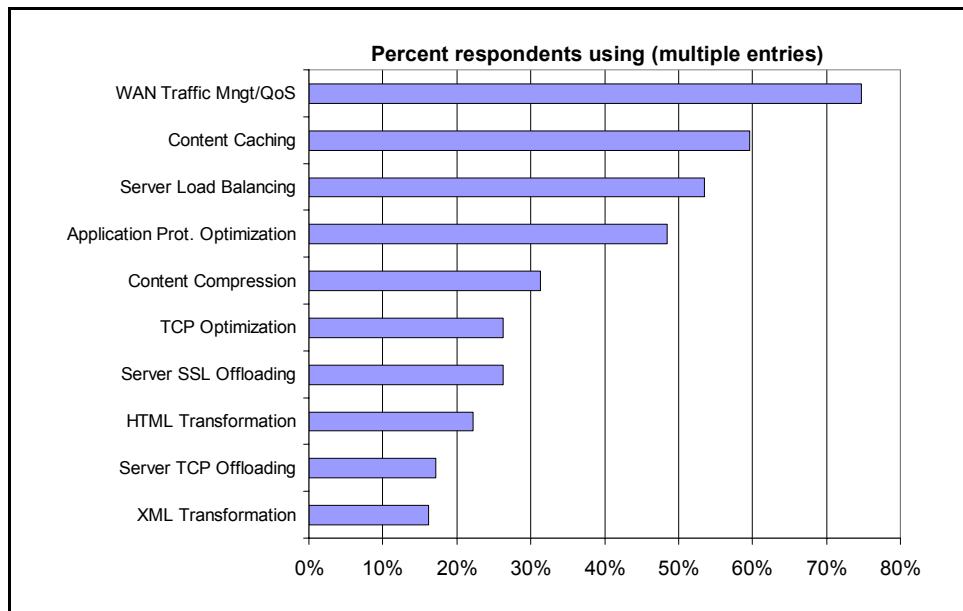


Figure 4 – Performance Technologies in Use

To Measure Is To Improve

Although we found no correlation between the number of techniques used to improve performance and corresponding satisfaction with and confidence in performance, we did find a clear correlation between measurement and effective application performance. Companies that measure and report well have 71 percent higher ratings for effective performance. We posit that this is because measurement and reporting fosters good processes for handling performance at all management levels. Once aware of performance issues, people are more likely to address them.

The importance of measurement and reporting to achieving better performance is reinforced by the fact that not only are enterprises who measure much more satisfied with performance overall, they are also more likely to find themselves in the safe quadrant.

We note that end-user response time is the least effectively measured of the four performance functions, a shortcoming that savvy enterprises would be wise to address.

Conclusions And Recommendations

The survey results show that measurement and reporting are valuable first steps in managing and improving all aspects of application performance. We suggest that enterprises will be best served by a comprehensive reporting system, not one that is tied to any single control or acceleration device in a piecemeal solution.

Unfortunately, such a comprehensive reporting system is unavailable today. We recommend that control and acceleration vendors collaborate with measurement and reporting vendors to meet this need. Enterprises need more alliances and fewer products that compete on “me too” measurement and reporting features— acceleration vendors should concentrate on their unique primary added value, and should partner with measurement vendors to do the rest.

NetForecast helps change delivery systems to improve the performance of networked applications. This includes advising enterprises on how to evaluate, improve and manage the performance of business applications, as well as advising vendors about customer requirements, technology issues, and adoption trends.

The relatively minor role of Web-based applications may explain why many of the acceleration techniques used by our survey respondents did not help performance as much as we expected. Vendors of acceleration technologies would be smart to support more applications (Web-only is not enough).

This is just one example where current solutions target select performance functions and applications when broader solutions are needed for best long term performance results. We recommend that, given this environment of narrow solutions, enterprises educate themselves on the benefits of available control and acceleration systems. They can then engineer a comprehensive system to improve the effectiveness of all four performance functions.

The bad news from the survey is that enterprises have bought lots of performance-enhancing technologies that have yet to deliver broad benefits. The good news is that there is a clear opportunity for vendors to do more to measure, report on, and improve the performance of the complete range of enterprise applications—and to collaborate to integrate individual solutions into a comprehensive application delivery system that will improve application performance across the board.

NetForecast and BCR would like to thank those BCR readers who took the time to complete our survey.

Peter Sevcik is President of NetForecast and is a leading authority on Internet traffic, performance, and technology. Peter has contributed to the design of more than 100 networks, including the Internet, and holds the patent on application response-time prediction. He can be reached at peter@netforecast.com.

Rebecca Wetzel is an Associate of NetForecast and a 20-year veteran of the data networking industry, with unparalleled inside knowledge of the Internet service and product markets. She works with network product vendors and service providers to develop and implement strategies for new IP-based offerings. She can be reached at rebecca@netforecast.com.

