Defining The Application Performance Index

Now a single metric can tell you if your applications are performing satisfactorily

Progress comes in two ways—new technology or new methods. Now is the time to shift the focus towards methods over technology, and here is why.

Every year, *CIO* magazine polls IT industry leaders to determine their burning issues. The most recent "State of the CIO" survey, published last October, lists the top dozen management priorities. The most interesting thing about the list is that it does not specifically call for any new technology. Not even a technology refresh. Top managers are looking for new methods by which to measure, quantify and improve IT.

The top three goals focus on how well IT performs:
■ Increase business efficiency through IT-enabled process improvement.
■ Align IT and business goals.
■ Improve internal customer satisfaction.

Given that the most critical aspect of performance is how users view application performance, wouldn't it be great if there were a new method to quantify application performance so that the CIO's goals could be achieved before next year's survey? Such a new method is coming.

First, we need to review where IT management has gotten us with technology innovations. There are at least 20 leading vendors of application performance measurement tools, each with a unique way to instrument and gather information on how well IT is running. These vendors often compete on the accuracy of their data along with the level of detail they can supply.

Many enterprises rely on more than one vendor and then add several home-grown tools, such that they are now swimming in numbers. Worse yet, the variety of numbers fuels arguments over accuracy and relevance, rather than helping form the insight requested by the CIO.

So now is the time to stop adding new numbers, but rather creating new methods for reducing the data at hand to meaningful information. The best approach to defining such a methodology would be to create it as an open standard rather than a proprietary solution.

Last fall, NetForecast organized a group of vendors to develop and then specify a new way to report on performance based upon measurement capabilities that already exist. The result is the Application Performance Index or Apdex.

The Apdex

The Apdex is a numerical measure of user satisfaction with the performance of enterprise applications. It converts many measurements into one number on a uniform scale of 0 to 1 (0 = no users satisfied, 1 = all users satisfied). This metric can be applied to any source of end-user performance measurements. If you have a measurement tool that gathers timing data similar to what a motivated end user could gather with a stopwatch, then you can use this metric. The Apdex fills the gap between timing data and insight by specifying a uniform way to measure and report on the user experience.

The index translates many individual response times, measured at the user-task level, into a single number. A Task is an individual interaction with the system, within a larger process. Task response time is defined as the elapsed time between when a user does something (mouse click, hits enter or return, etc) and when the system (client, network, servers) responds such that he/she can proceed with the process. This is the time during which the human is waiting for the system. These individual waiting periods are what define the "responsiveness" of the application to the user.

How The Apdex Works

The tools that support the Apdex will conform to a specification currently under development that will become publicly available. It specifies a process that Apdex-compliant tools and services will implement. A key attribute of the process is that it is simple. Here is a basic overview.

The process starts with defining a Report Group that the index value will represent. This is the first step in reducing the vast number of measurement samples into a meaningful subset. Some example Report Group parameters are application, user group and time of day.

The index is then based on three zones of application responsiveness:
■ **Satisfied**—The user is fully productive. This represents the time value (T seconds) below which users are not impeded by application response time.
■ **Tolerating**—The user notices performance lagging within responses greater than T, but continues the process.

Use BCR's Acronym Directory at www.bcr.com/bcrmag

Frustrated—Performance with a response time greater than F seconds is unacceptable, and users may abandon the process.

So the two thresholds of T and F define three performance buckets into which all the samples of a Report Group can be placed: 0-T, T-F, and >F. The index calculation is a weighted sum of the percentages of samples that fall into each of the performance zones.

Defining the target time T is a fundamental part of the Apdex process. All Apdex values are based upon this basic reference goal for each application. This is what grounds the index in a business need and gives the values a clear reference. There are several methods for determining T, and many more will be learned as the index is implemented.

A lot of research has been done in human-computer interaction to determine when applications are fast enough (T), and too slow (F). There is little research on the ground between T and F because people have found little need or value in subdividing the tolerating zone. The good news is that the value of F is a function of T.
■ Application usability guru Jakob Nielsen defines "reasonably fast operations, taking between 2 and 10 seconds" as the range between T and F (*see reference 1*).
■ In 1997, when the typical Web page loaded in 10 seconds, Judith Ramsay, et al (*reference 2*) found that users significantly changed their perception of how interesting the content was when they had to wait 41 seconds and longer.
■ Nina Bhatti, et al (*reference 3*) ran controlled experiments where users configured and purchased a PC on line. The experiments showed a definite shift of ratings from good to poor at 10 seconds, and users rated performance as unacceptable if pages loaded in more than 39 seconds.
■ NetForecast has conducted observations of users in various business environments. We have found several examples of production users such as insurance claims processors—who needed a 1-second response—suddenly abandoning the process at 4 seconds. A financial services firm operated well below 3 seconds but started to lose business above 12 seconds. Finally, an international supply chain management system had users working productively at less than 5-second response time, and complaints that affected business started at 15 seconds.

The above examples indicate ratios of 3:1, 4:1 or 5:1 between the two thresholds, with a preponderance of 4. Thus Apdex defines F to be 4 times T, and the three performance zones are defined on a base value of T seconds.

The Apdex formula is the number of satisfied samples plus half of the tolerating samples plus none of the frustrated samples, divided by all the samples:

$$Apdex_T = \frac{Satisfied + Tolerating\ /2}{Total\ Samples}$$

So it is easy to see how this ratio is always directly related to the users' perception of satisfactory application responsiveness. To understand the full meaning of the ratio, it is always presented as a decimal value with a sub-script representing the target time T. For example, if there are 100 samples with a target time of 3 seconds, where 60 are below 3 seconds, 30 are between 3 and 12 seconds, and the remaining 10 are above 12 seconds, the Apdex is:

$$\frac{60 + 30\ /2}{100} = 0.75_3$$

Apdex Benefits

There are several benefits to using the Apdex. It is the first user experience metric that is comparable across all transactional applications—a value of $0.85T$ means the same thing in all applications even with different values of T. Thus the enterprise manager can have a common way to compare performance across applications or other reporting groups he or she defines.

This is the one-number metric that senior management can easily understand and use to manage IT across many applications. Managers can easily see which applications need improvement or investment—i.e., those that have a low Apdex value but are important to the business.

Apdex also lets enterprises measure the effectiveness of performance improvement investments. An Apdex value should improve with a performance-driven upgrade. This is a good way to determine which applications need help; identify remedial investment; and then track if the investment paid off.

But the greatest benefit of the Apdex methodology is its ability to quickly show the alignment of application performance to the needs of the business—one of the top CIO goals. Imagine the following simple exercise: A CIO is managing a portfolio of several major business applications from order processing to corporate email. The CIO gets consensus among the business managers on a ranking of the applications by importance to the business. Presumably order processing will be high and email low. Then the CIO just has to rank the same applications by the Apdex value they deliver during the business day.

If the rankings match, the applications and business needs are properly aligned. If, on the other hand, email has a high Apdex while order processing has a significantly lower Apdex, then the applications are out of alignment. The CIO knows where he stands and can direct change and track the success of the change until proper alignment is achieved.

Of course a real business alignment exercise would be more complex, but using the Apdex as a tool for discovery and remediation will be a central part of the strategy. For example, ensuring that

Ideally, your most important applications should have the highest Apdex scores

Apdex values meet corporate objectives is also important. We expect enterprises will use the index as a tool in various management approaches customized to their own needs.

The Apdex also helps with the other two top CIO goals. This is a process improvement technique that can increase business efficiency; it is all process, and not new technology. Finally, it is directly a tool for improving internal customer satisfaction. It is not often that you can make significant headway on your top three goals with such little cost.

Open Movement
The best part about this initiative is that it is being defined as an open standard within the Apdex Alliance, whose membership currently includes:
- Adlex
- FineGround
- NetForecast
- NetQoS
- Network Physics
- Packeteer
- Peribit
- Swan Labs
- WildPackets

The vendors on this list are committed to putting this reporting feature into their products. They are further committed to helping enterprises make full use of this capability as a foundation for continuous process improvement in managing the performance of their IT infrastructure.

The open collaborative approach provides many benefits. Clearly, the soon-to-be-released specification will be just the first step in a dialogue among enterprises, vendors, consultants and the analyst community on how to improve the methodology over time. Vendors will produce better products thanks to more input, review and evaluation from their peers. Furthermore, the alliance will certify the implementations of members' products to ensure compliance with the specification(s).

Finally, the group is starting an enterprise advisory board to get direct input for developing best practices around the Apdex. A few enterprises have already agreed to participate.

Invitation To Join
The current alliance members have invested resources to make Apdex a success. But we still need more help. I encourage additional vendors to join. This methodology is applicable to a broad range of products or services. For example, half of the current members are traditional measurement companies, while the other half are performance enhancement vendors. Member companies can

shape the future specifications and management methodologies of Apdex.

We also invite enterprises to participate in the advisory board to both help guide the work of the alliance and to learn best practices from each other. The larger the group, the better the product.

This is the start of a grand change in how we manage technology. It will shift the dialogue from technology pushing its way into enterprises to making technology accountable in support of the business. Come be part of the revolution!□

References
1 "*Usability Engineering,*" by Jakob Nielsen, published by Morgan Kaufmann, San Francisco, 1994

2 "*A Psychological Investigation of Long Retrieval Times on the World Wide Web,*" Judith Ramsay, Alessandro Barbasi, Jenny Preece, published in Interacting With Computers, Elsevier, March 1998.

3 "*Integrating User-Perceived Quality Into Web Server Design,*" by Nina Bhatti, Anna Bouch, Allan Kuchinsky of HP, published in the proceedings of the 9th International World Wide Web Conference Amsterdam, May 15–19, 2000

The Apdex Alliance will have an enterprise advisory board for input and development of best practices

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.

Companies Mentioned In This Article

Apdex Alliance (www.apdex.org)

Adlex (www.adlex.com)

FineGround (www.fineground.com)

NetForecast (www.netforecast.com)

NetQoS (www.netqos.com)

Network Physics (www.networkphysics.com)

Packeteer (www.packeteer.com)

Peribit (www.peribit.com)

Swan Labs (www.swanlabs.com)

WildPackets (www.wildpackets.com)