

The Importance of Being Proactive The Necessity of System Performance Reviews

By Laura R. Bryngelson

Almost every CIO, IT Director, or System Manager will, at some point in his or her career, be in a situation where system performance becomes an all-encompassing issue. All of the sudden the system is brought to its knees, management and the interactive community are panicking, and the entire company is blaming the IT department, or worse, they're all blaming you!

What's wrong with this picture? Unless there is a catastrophic hardware failure, system performance doesn't go from great to horrible overnight. Often, system performance degrades very subtly at first – response times increase imperceptibly, programs are hindered, and batch processes take longer to complete. Then, once several performance pulse points begin increasing and building upon each other, you've hit the “knee in the curve”, and performance degradation increases exponentially.

How can this be avoided? With regularly-scheduled proactive system reviews, an IT department can see where their system currently is with respect to performance, where they will need to be in the immediate future, and identify any forthcoming significant events that will need to be addressed. These reviews should be done on a yearly basis at the very least in static shops, and in very dynamic environments, should be done on a monthly-to-quarterly basis.

A regular review of the system should include information on the hardware and software technology, CPU loading and balance, memory, disk I/O, user response, and batch performance. You should know ahead of time whether operating systems, disk drives, upgrades, or processors will be sold and supported in the near future, and if and when they will be obsolesced. The last thing you want to hear from the support person on the other end of the phone is “Your environment is currently unsupported due to the obsolescence of hardware or software. Our solution will require migration to currently supported technology.”

The CPU is the hub of activity on your system, and regularly reviewing the performance indicators associated with the CPU can help avert bottlenecks. You need to know how much of a load is placed on your CPU and whether the system has any bandwidth available for added users or applications. You should be able to identify peak times of usage, whether processes are batch or interactive, and if your queue structure is optimized for your needs.

Having adequate memory can make or break a system, and “adequate” depends on many factors. The number of users, jobs, disks, the version of the operating system, the make and model of disk arrays, and the number and type of applications all help to determine how much memory you should have installed on the system. A performance review should investigate industry-accepted indicators that point to inadequate memory – memory manager percentage, read hit percentage, and page faults.

The physical act of reading from or writing to a disk is thousands of times slower than reading information stored in memory. Hence, many performance issues are caused by disk I/O bottlenecks. If memory is inadequate, then the CPU is forced to perform a disk I/O in order to find the information it needs. However, other factors can lead to bottlenecks as well. If a disk or file is fragmented, then the chunks of data the CPU brings into main memory on the assumption that the data is contiguous (“prefetching”) will actually only contain part of the information needed by the user or application, and will result in at least one more physical disk I/O to gather all the information. If the majority of accessed data resides on the same logical device, this can also cause a disk bottleneck, as the CPU can only access the disk once at a time, through the one controller that leads to that single disk. If one logical device is being accessed consistently, then

a disk I/O queue will begin to form. These are processes waiting in the request queue for a particular drive. If the disk I/O queue length begins to increase, it is an indication of several possibilities: low memory, file imbalance among disk drives, or disk/data fragmentation.

Although several other pulse points previously mentioned are just as likely to indicate performance issues, you are more likely to hear about performance issues from your users – vociferously! Regular reviews of your system can help to pinpoint potential problems before you hear about them. Monitoring the response times and CPU queue length can give the watchful system administrator an idea of how long users are having to wait for their processes to complete.

Finally, observing the batch performance usually indicates whether batch processes are able to complete in a timely manner or if they impact interactive users. Interactive users generally should have priority over batch because, let's face it, people get paid to sit and wait for the CPU; batch processes don't. If your system is configured in such a way that batch processes are fighting interactive users during the primary day shift, but nothing is running overnight, you are not using your system to its full capacity. Many batch processes can be scheduled to run during low-utilization hours, freeing up valuable (and expensive!) bandwidth during the day for interactive use.

So... you've decided to do regular reviews of your system and have identified a need for a change. Because of your diligence, you have been able to plan for another year of productivity and make better business decisions. Perhaps your system needs more memory, or you want to partition your system volume set to aid in disaster recovery, or maybe your disks and databases are heavily fragmented, and you plan to defragment and repack them. How do you show the powers-that-be that your hard work and resources were not wasted in an unnecessary upgrade? The simple fact that a performance review of the system indicated a problem in the first place is part of the equation, but you need to show results. After any significant change to your system, a complete review on the system is again necessary - using the same performance indicators and similar data (same time of day, week, or month, with similar processing).

Having the information is just not enough, however. You need to analyze it and report the findings to the appropriate people in charge. Do you simply take the numbers into a meeting and start rattling them off, trying to ignore the glazed look in your co-workers eyes? Of course not – visual cues and succinct synopses will impart more important information much more quickly. Your review should include graphs (preferably in color so that finer points can be distinguished); a summary page which distills the individual analysis of every pulse point into an overview for the entire system resource and offers possible causes and solutions; and, for those who only want the nitty-gritty facts (and they wanted them yesterday); a report card page that simply lists the pulse point, a very brief (one sentence) summary, and whether that pulse point rated good, bad, or in between. This way, all your bases are covered, and for one blissful moment in time, your entire department is on the same page.

Of course, we all know that system administrators have absolutely nothing better to do with their time than to constantly keep abreast of system performance, surf multiple web-sites for discontinuance notices, analyze 20-25 different performance indicators, and write up lengthy reports, including full color graphics. However, if you find yourself a little short on time, you may want to drop by the Lund Performance Solutions booth this week and let us make you look like a hero to your IT department with Annual Reviews, Before-and-After Reviews, and System Performance Outsourcing Contacts. You *can* proactively monitor your system performance... and still make your 3:00 PM tee time!