Performance Assurance for Contingency Planning

By Mike Matchett

To plan for contingencies, companies need to determine how to ensure availability while balancing performance and cost issues. BMC Software offers a best-practices approach called Performance Assurance®, which helps companies deliver quality service efficiently and cost-effectively with fewer resources. This article describes how companies can apply the process of Performance Assurance for contingency planning.

IT departments everywhere are reviewing their disaster recovery strategies and planning for new possibilities while the high-tech economy struggles to recover. Customers are demanding increasingly competitive service quality and performance. Because companies need to ensure business continuance, they must be smarter and more proactive in their decisions about hosting plans and resource investing.

These types of decisions form the heart of capacity planning, which is embodied in the BMC Software® Performance Assurance® approach using PATROL® Perform and Predict products on Dell™ PowerEdge™ servers. Performance Assurance helps companies balance availability, performance, and cost for capacity planning as well as contingency planning. Performance Assurance addresses questions such as:

- How much spare system is necessary for crisis peaks?
- How powerful should the hot site server be to handle the minimum requirements in different contingency scenarios?
- What kinds of performance degradation in throughput and response time will occur during cutover situations?

Applying Performance Assurance to contingency planning

Performance Assurance is based on a best-practices methodology that advises companies to first visualize historical data, then analyze current behavior, and finally predict future capability. The following steps show the basic process of Performance Assurance as applied to contingency planning.

1. Create a performance baseline

Create an application and system performance baseline by collecting and visualizing historical performance data. This data helps companies understand the operations in their production environments today.

In particular, identify the normal operating behavior of applications and services in system resource terms (CPU utilization, I/O, memory, and so on). For example, Figure 1 shows a graph produced by the Visualizer component of PATROL Perform and Predict that gives a snapshot of monthly system CPU utilization. Figure 2 shows another Visualizer graph that tracks the utilization of seven hard disks.

Companies should retain enough history to understand past behavior, but they also should remember the dynamic nature of current systems and applications and not get caught up in the past.

2. Characterize the workload

Analyze and elevate the raw data into a higher level view of the workloads. Characterize the workloads to correspond to modules, applications, services, service groups, users, and user groups. During contingency planning, companies usually should

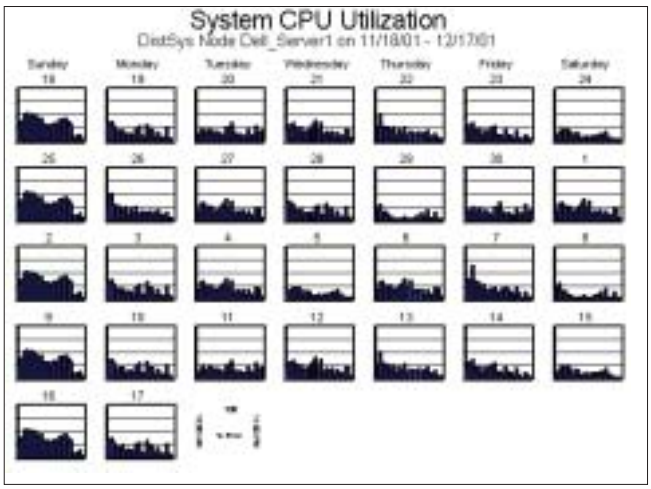

Figure 1. A monthly performance baseline

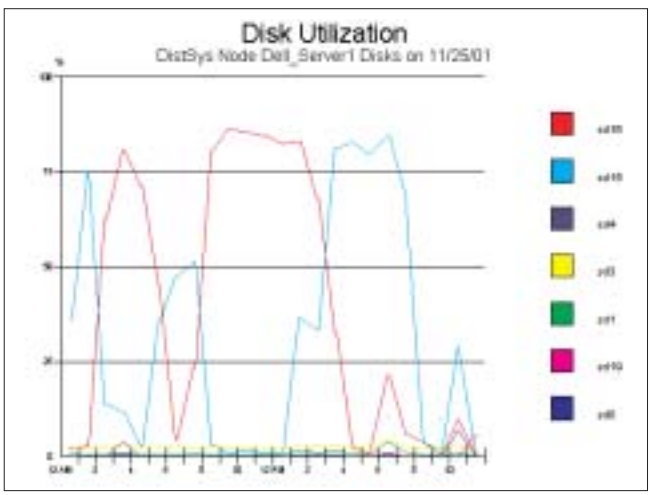

Figure 2. Disk utilization during weekend batch runs

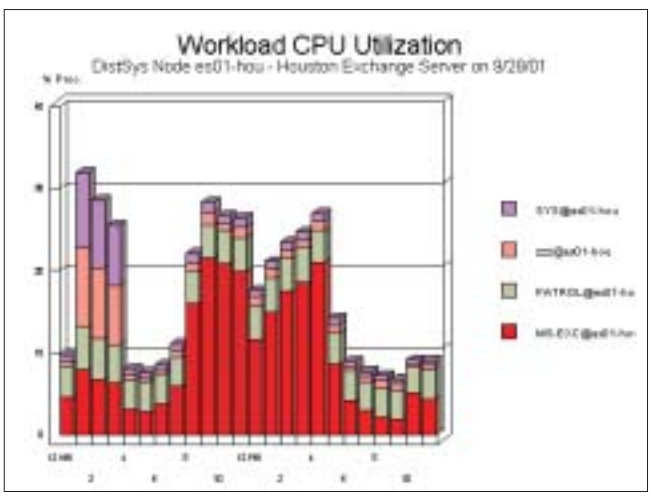

Figure 3. Total CPU utilization by workload on one node

characterize workloads at an application level so that they can easily model alternative scenarios for re-hosting those applications. Workload characterization can not only include the processes and users that work on a system, but also define the computing resource demands of multiple-tier client-server systems, relational database management system (RDBMS) servers, attached storage and RAID (redundant array of independent disks), and network traffic.

Figure 3 shows the application breakdown among Microsoft® Exchange®, system utilities, and overhead. Note the spike in CPU utilization during the nightly backup from 1:00 A.M. to 3:00 A.M. This type of data allows companies to identify and isolate important workloads for planning purposes.

3. Profile the workload (or analyze performance)

After obtaining historical data and a workload view, companies need to understand how the application workloads vary across time, in terms of cycles, intervals, peaks, interference patterns, variances, and exceptions to the norm. Determine what behavior is normal and which time intervals represent peak application volumes of demand on the systems. Correlate workload transactions to units of business so that companies can make technical recommendations in terms that are informative and actionable to the business management.

4. Track variance

Identify new, unexpected, and abnormal behavior in the applications and systems through an automatic analysis of variance over time. This analysis, which is not the same as real-time thresholding and alerting on metric values, is a sophisticated multivariate adaptive statistical filtering (MASF) of new performance behavioral data compared to the identified baseline.

Visualizer uses MASF to produce graphs that display current system performance compared to a baseline (see Figure 4, for example). Companies can obtain exception reports when any metric in the performance database begins to deviate beyond the tolerance for variance from the baseline.

5. Build baseline models

Build baseline models of systems and workloads from the data, choosing representative behavior as determined in step 3. Companies can now understand the response times of the service and wait components of a workload, easily identifying any resource bottlenecks. Run the baseline models through predictive workload growth scenarios to evaluate future performance degradation under crisis. To understand the performance and capacity consequences of various growth scenarios, study the throughput and response time curves versus the business transaction amounts.

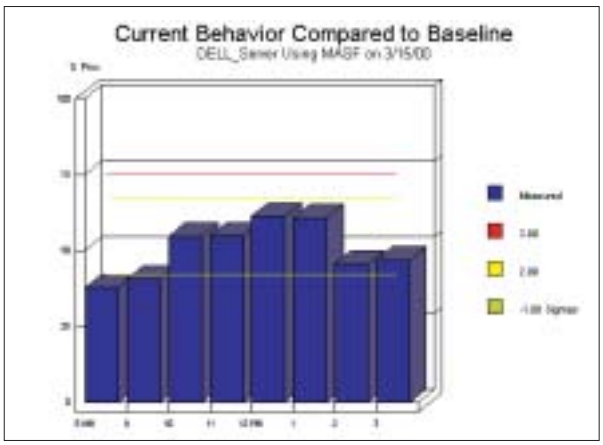

Figure 4. System CPU utilization compared to a baseline of variance

PATROL Predict provides a breakdown of response time by workload over the various resources. For example, Figure 5 shows that wait times are queuing bottlenecks when some Internet server work is failed over to an existing Exchange server.

6. Create models of contingency scenarios

Use the baseline models to quickly evaluate models of different cutover and hosting scenarios. PATROL Predict lets companies cut and paste workloads from model to model and rapidly explore all options. Similar to server consolidation planning, modeling for contingency planning explores hosting workloads on alternative servers or combining workloads in fallback positions on remaining cluster servers.

The models show relative changes in performance and throughput, which can indicate the effects of upgrading or downgrading host CPU, memory, I/O, and RAID resources. Using the models, companies also can evaluate the effects of limiting transaction arrival rates, changing relative process priorities, implementing new fair-share scheduling schemes, or even dropping non-critical work from a server.

7. Account for future growth

Best-practices planning should account for future growth possibilities. Although companies cannot always provide solid estimates far into the future, they can max out current models using Predict's built-in scenario planner and provide guidance on the maximum capabilities of the planned systems.

Figure 6 shows the Predict Scenario Planner used to evaluate successive periods of 10 percent growth in workload. An evaluation could be made by observing CPU utilization, which increases linearly with workload, but the real story is found here in the response time, which degrades nonlinearly with workload growth. Companies need to plan for performance, not just utilization.

Figure 5. Response time overview

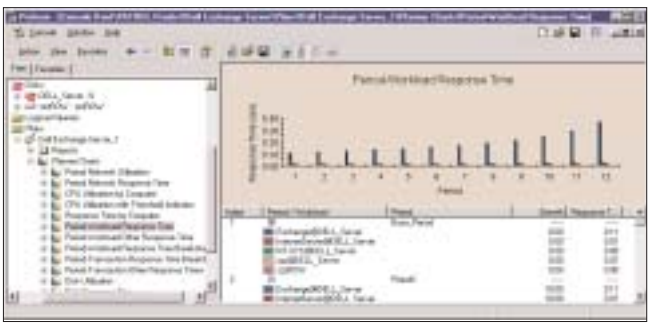

Figure 6. Period-workload response time

A strategy for assuring availability

The BMC Software Performance Assurance products, PATROL Perform and Predict, run on Dell PowerEdge servers and are integrated with PATROL Availability solutions that monitor and proactively manage infrastructure and application components during production operations. The Performance Assurance products are also integrated with PATROL Service Level Management, which ensures that customers receive the service quality they require. In the end, Performance Assurance is about lowering risks and controlling costs while ensuring performance. Contingency planning shares these same goals.

Mike Matchett *(mike_matchett@bmc.com) is a senior performance solutions product manager at BMC Software, a provider of enterprise management solutions. He is working on new requirements for future PATROL Performance Assurance solutions. Mike received a B.S. in Electrical Engineering from the Massachusetts Institute of Technology.*

FOR MORE INFORMATION

For more information, service, and support, visit

BMC Software: http://www.bmc.com

BMC Performance Assurance (planning demos available under "Interactive Tools"): http://www.bmc.com/solutions/performance