Infrastructure Optimization through Architecture Validation

HyPerformix, Inc.
4301 Westbank Drive
Building A, Suite 300
Austin, TX 78746
www.hyperformix.com

Abstract
This whitepaper discusses a proven approach to managing system performance by combining simulation modeling and a process called 'Performance Budgeting' – to validate, very early in the system lifecycle, that architectures are optimally tuned to make use of infrastructure resources.

Introduction

The last decade has seen unprecedented increases in computing power, network capability, as well as software and systems complexity. As technology has increased, so has the ability to use this technology to provide more comprehensive and increasingly varied services to customers. This has resulted in a corresponding increase in the number of systems and applications that can be considered mission critical.

While providing the optimum user experience is now considered a requirement of mission critical systems, this has to be balanced against the cost of building and maintaining these systems. If done incorrectly, ongoing expenditures on hardware and software become as enormous as they are unnecessary. Currently, these ongoing expenditures are reduced only after systems reach production through painful and expensive system consolidation efforts, while the cost of corrective action during test is assumed to be a necessary part of the system lifecycle.

With increasing pressure to manage budgets, these costs need to be controlled from the beginning of a project, not from the end. It is imperative that companies work smarter – early in the lifecycle – to deploy systems providing competitive functionality at competitive costs.

This paper discusses a proven approach, combining simulation modeling and a process called *Performance Budgeting*, to validate very early in the system lifecycle that architectures are optimally tuned to make use of infrastructure resources. The solution leads to systems that perform better, cost less to maintain, and are more stable than systems created in the traditional manner.

The Problem

The costs of correcting system architecture, capacity, and performance deficiencies increase dramatically through the systems lifecycle. While the most cost-effective time to correct these issues is early in the development lifecycle, this is also the time when the least amount of measurable information is available. Because of this, companies have traditionally released systems to production or integration test before addressing infrastructure optimization issues. This has traditionally resulted in:

- Overspending on development costs. Anecdotal evidence abounds of instances where design mistakes are uncovered in test and production leading to late stage development work to correct. This, in spite of numerous studies indicating that correction costs increase by an order of magnitude for each change in stage (from requirements to design, from design to development, from development to test, and from test to production).
- Overspending on assets. Companies often purchase more assets than needed; convinced that proper assurance of the performance of the system prior to production is impossible. In a recent article in InformationWeek, the Gartner Group estimates that companies wasted more than $1 billion between 1998 and 2000 on high-end servers,[1] and the trend is getting worse. The same article estimates that this number could grow to $2 billion between 2001 and 2003.
- Extreme testing costs. Mission critical systems and applications can't be left to chance, so companies often turn to the only technology they know to solve the problem; mirrored systems. Aside from the sheer cost of such an effort in hardware and software, mirrored systems also require significant investments in maintenance and create a willingness to accept significant delays and information limitations in investigating the impact of system changes.

[1] InformationWeek, August 22, 2001.

Simulation Solution

What is simulation modeling? Simulation modeling has been used for decades in industries such as manufacturing, software and hardware design, chemical process engineering and aeronautical engineering to efficiently and accurately predict performance and capacity early in the lifecycle. Simulation provides the ability to create a virtual, computer-based representation of the entire environment from end-to-end, complete with hardware, software, and user processes. These models are used to efficiently and effectively perform what-if analyses to understand the aggregated effect of the interactions of these components (the user experience). However, simulation models are generally only as good as the data used to create them.

So, what about the data problem? As stated, the difficulty in validating architectures and applications early in the development lifecycle is that little quantitative information is known about the performance of the application. This issue is addressed through the process of Performance Budgeting.

A Pragmatic Approach

The goal of performance budgeting is to ensure new architectures and applications are designed to meet performance requirements early in the lifecycle. The process begins with information gathering; collecting the specific data points needed to understand the end-to-end performance and capacity of the system. Some of this information will be known early in the lifecycle, much will not. For data points that are not currently known, budgetary numbers are assigned and documented in the performance budget. The known information and the performance budget are used to build a model of the system. This model is used to develop an understanding of the integrated effects of the components on end-to-end response time and overall system capacity and utilization. Simulation modeling provides an environment to tune the architecture, infrastructure allocation, and performance budget until an optimal (maximum performance and stability, lowest cost) solution is reached. At this point, the performance budget is established as a set of requirements for the component measurements.

Also at this time, it is possible to run sensitivity and bottleneck analyses. A sensitivity analysis is performed by varying numbers in the performance budget by some percentage. The goal is to determine which components have the largest overall effect on the user experience.

The performance budgeting process applies well to information systems since there is typically an incredible amount of reuse, and therefore known information, in these systems. Hardware components are commercially available, and their performance is documented. Commonly available applications can also be benchmarked.

With this approach, the need for corrective action and rework due to performance issues can be reduced or eliminated by developing performance understanding in the design phase. Further, the cost of maintenance is reduced by using a developed model of the system from design and development in maintenance and support. Key benefits of this approach include:

- Aiding decision-making during architecture/application design.
- Ensuring proposed architectures/designs perform as promised.
- Optimizing architectures/applications to run on infrastructure.
- Testing the application on the infrastructure before either exists.

The task doesn't end here, however. With a model in hand, a refined understanding of our assumptions, and a performance budget, we can continue to monitor these assumptions. As the system is developed, and measured data becomes available, it is checked against the performance budget (with special attention paid to sensitive system areas and bottlenecks). Measurements that do not correspond to the budget are caught as early in the lifecycle as possible and immediately addressed at the lowest cost possible.

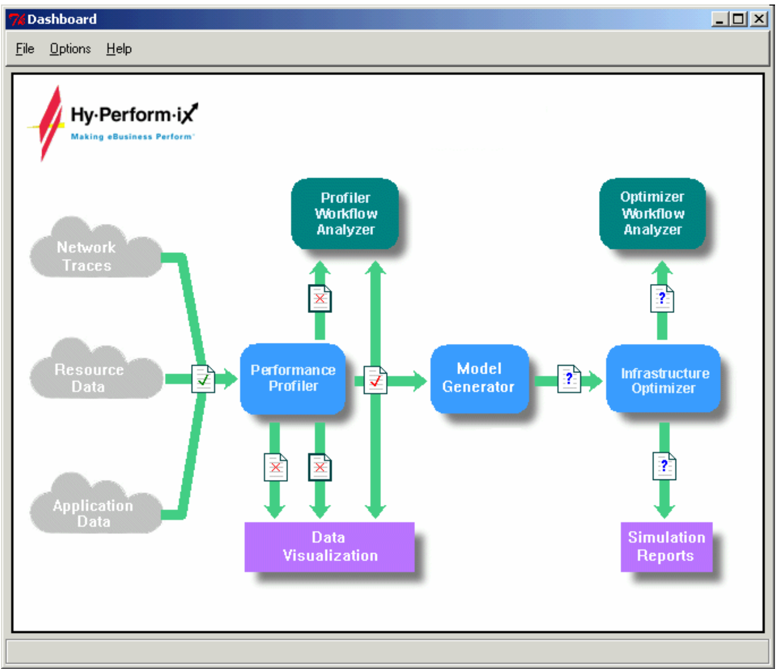
The Implementation

HyPerformix developed a product family, the HyPerformix Integrated Performance Suite™ (IPS), which allows for the rapid, automated creation of application models based on data collected in a test or production environment. IPS includes the HyPerformix Performance Profiler™ (Profiler), the HyPerformix Application Model Generator™ (AMG), and the HyPerformix Infrastructure Optimizer™ (Optimizer).

IPS also includes an extensive library of commercially available hardware and network components (clients, servers, routers, switches, gateways, etc.).

This family of products (shown in figure 1) provides an excellent environment to evaluate the what-if questions that will drive the success of your IT environment and your business.

Figure 1 – HyPerformix Integrated Performance Suite



The Value

Costs are saved in three main areas. The first, and most significant, is through reductions in effort devoted to system design corrections during test and production. With the ability to validate system performance early in the design and development phases, these efforts don't get pushed into testing and production, where corrective action is the most expensive. In addition to the costs saved solely from the reduction in effort, lost revenue due to downtime and poor system performance is reduced as well.

The second area is in the reduction of redundant and unnecessary asset costs. With increased visibility into the true performance of the system, effort can be devoted to meeting the system needs, not towards purchasing enough of a buffer to cushion against bad forecasts and estimates.

Finally, this approach provides an ongoing benefit in the construction of an accurate system model. The analysis of performance impacts of multiple alternate hardware or software configurations that are done during design and development is just as important once the system reaches production. The resulting model can, and should, exist and mature along with the system itself. In this way, it can provide ongoing, cost-effective decision support regarding the impact of future changes such as increased business load, the addition of functionality, and changes to the infrastructure.

Conclusion

In this time of reduced budgets it is no longer possible to rely on the old process of 'capacity planning by checkbook'. Companies that could rely on brute force means to manage capacity in the cash-rich days of the 90's must now constantly strive to work smarter, to minimize infrastructure costs, and to drive process efficiencies that will minimize resource costs.

Simulation coupled with Performance Budgeting is the most efficient, most effective process on the market today to achieve these goals and guarantee the success of your IT organization. With the process described above we're striving to:

- Minimize application rework due to performance issues by more than 50%.
- Reduce the infrastructure cost of supporting the given load by more than 20%.

