A Capital Effectiveness and Computing Resources Capabilities and Forecasting Model

An Advanced Capacity Planning Model

A Design Perspective

Scianta Intelligence

1289 North Fordham Blvd. Suite A312
Chapel Hill, NC 27517

(919) 678-0477
www.scianta.com

Scianta Capacity Planning and Capital Effectiveness

Précis

The Computing Services Effectiveness model integrates the enterprise's capitalization and discounted cash investments in its IT infrastructure with the cost of maintaining a sustainable computing resource pool to produce a risk adjusted return on investment. The model provides load sensitivity and goal seeking capabilities tied directly to the capital and expense model for each scenario. Our effectiveness model provides a tight coupling between the organization's capacity and performance capabilities (through the model's own distributed capacity sub-model) and the idea of economic utility functions associated with the organization's investment in its application and infrastructure frameworks. The model itself is based on a tier of functionalities – thus allowing an organization to start small and grow as their needs dictate.

Modern corporations and government agencies rely on their increasingly distributed computer infrastructures – application, web, and database servers as well as routers, load balancers, and firewalls - to support a wide and often unpredictable load of clients and applications. In the modern interconnected world, the infrastructure is the backbone of the organization's mission capabilities. Scianta also recognizes that the enterprise's local and distributed infrastructure is intimately and ineluctably tied to business strategy decision and that the quality of these decisions is based on a reliable and sound application of business intelligence. Hence, without a sustainable, extensible, and elastic infrastructure, an organization has increasingly limited abilities to effectively accomplish its fundamental missions.

Thus businesses and government agencies at all levels are faced with two critical challenges: maintaining cost effective capacity in their computing resources and planning for future growth in an environment of high unpredictability. Together, these conflicting objectives comprise a major management and capital budgeting problem. To solve this problem, organizations have turned to several alternatives: outsourcing their computer support, increasing their own computing capabilities, and providing excess computing resources. Each of these options provides clear benefits and also significant disadvantages. Evaluating these alternatives to make an informed decision means more than simply calculating the costs of new machines or the monthly service fees charged by an Applications Service Provider. Making an informed decision means understanding the opportunity and utility costs of computing resources, accurately forecasting future demands on that resource, assessing the risks of performance restrictions, and gaining an insight into the potential and extended load bearing capabilities of an organization's current resources.

This coupling of economic, risk, and capacity capability models into a comprehensive Service Effectiveness and Forecasting Model provides a completely new approach to linking infrastructure with the enterprise's application framework. Figure 1 illustrates these bi-directional dependency relationships.

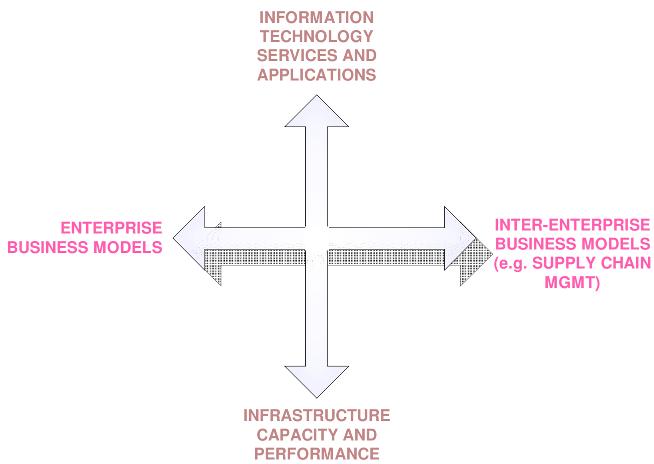

Figure 1. Coupling Infrastructure and Business Models

The intersection of the enterprise's computing technologies and its business models lies at the crux of a maintainable business architecture (and, naturally, this applies to local, state, and federal government agencies involved with law enforcement, policy making, and over-sight). At the junction of computing resources and business models lies the Scianta effectiveness and economic modeling system. Our analysis combines mathematical and statistical analysis with deep impact and configuration analysis based the evolution of multiple virtual machines. By applying machine learning capabilities to both the performance and the transaction space of web, application, and databases servers as well as other infrastructure components, the effectiveness model builds up a model of the computing resources. Its ability to learn the long term periodicity of both supply and demand, provides the model with a high dimensional predictive capability. On top of this virtual model of an organization's machine systems, the model lies a comprehensive risk assessment, economic, and demand sensitivity model. From this fusion of technologies the service effectiveness model computes the over-all capabilities of the organization's computing resources (current and predicted based on estimated growth), provides load impact analysis and prediction, and, given investment and expense figures, calculates a risk-biased return on investment.

Figure 2 shows the high level architecture of Scianta's Capital Effectiveness and Capabilities model. This is a very high level overview of the model itself. From this perspective you can see that the model rests on a multi-tiered architecture beginning with a capacity planning model and extending to a complete economic assessment model.

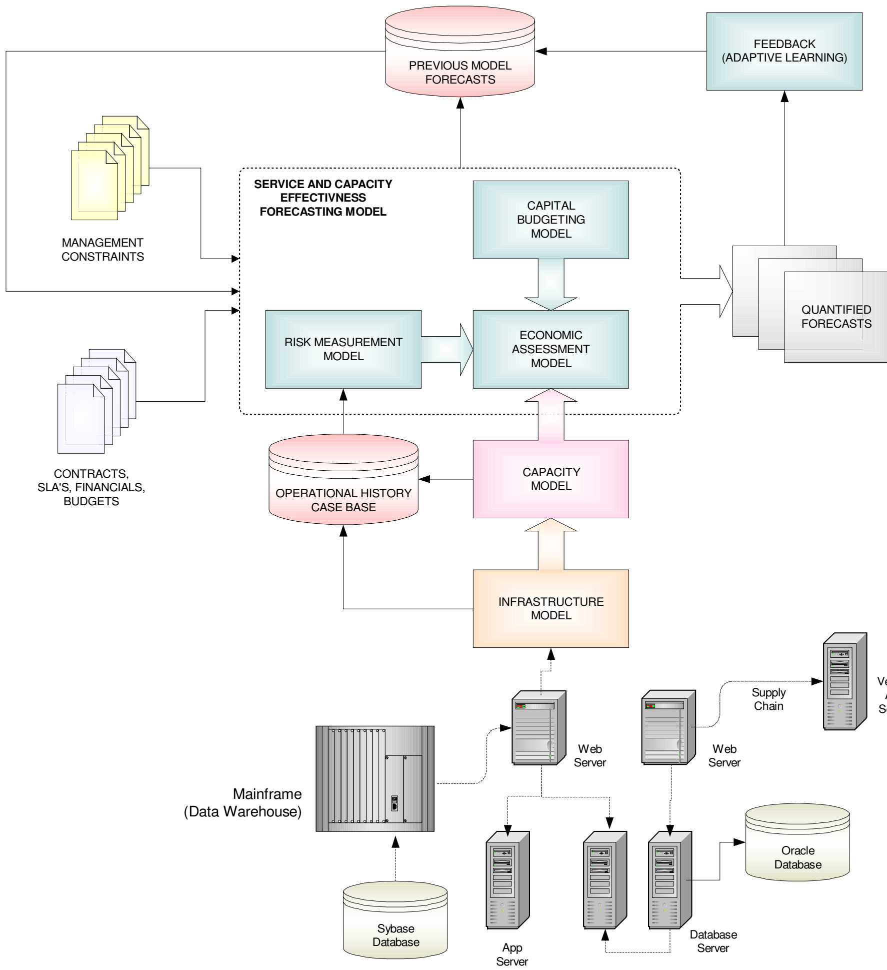

Figure 2. The eService Effectiveness Model Architecture

How do the risk and econometric models determine the cost effectiveness of a particular computing architecture and how does the impact and load forecasting model predict the net present value (depreciated) costs of the service given various anticipated application and customer loads? This is accomplished through the underlying capacity and performance analysis model. Only by understanding the deep, time-varying interplay between computing resource capabilities and on-going demand can a model provide an effectiveness model with accurate supply and demand information. Figure 3 illustrates this architectural layer.

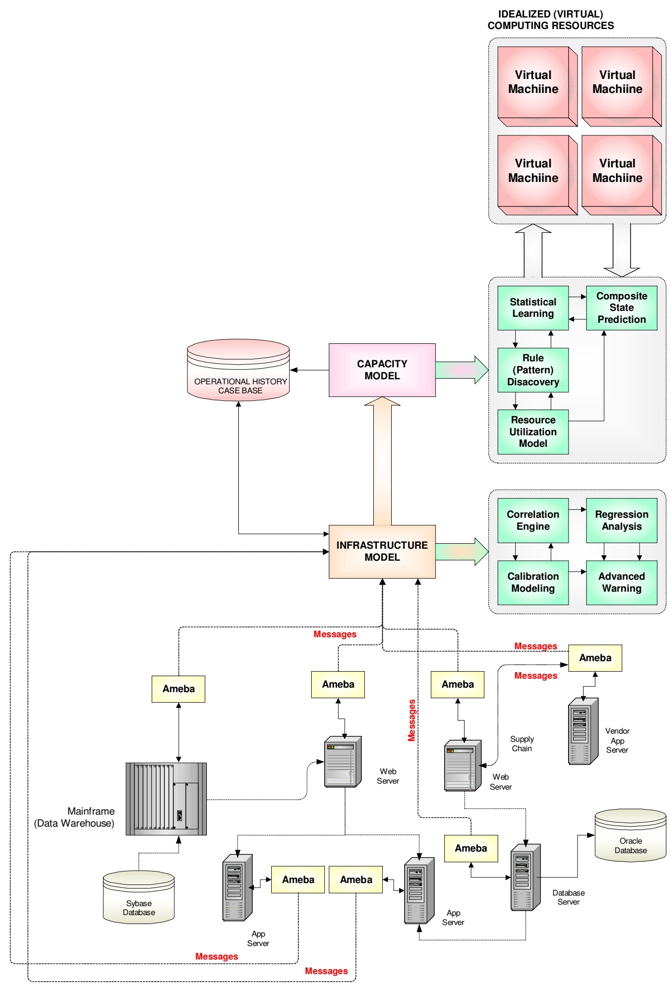

Figure 3. The Capacity Planning Model

Our capacity planning model relies on two advanced capabilities: distributed agents and computational intelligence. The distributed agents (called an AMEBA, the autonomous multi-explorer Behavior Agent) reside within the machine application and operating system space – their task is to gather, filter, transform, package, and send data as well as information messages to the machine intelligence facilities of the capacity model. The capacity model consists of two components – first, a low level infrastructure performance and calibration model that learns the periodicity of the various machine data streams and thus isolates anomalous behaviors, and, second, a multiple machine computing model that builds, tests, and runs a series of virtual machines based on the learned capacity of the underlying, real computing layer. It is this top level model, built on a simulation of the underlying physical architecture, that provides the capitalization and computing capabilities model with its powerful and robust abilities.