

IBM Research has a distinguished history in the theory and practice of performance analysis, modeling, and optimization. Fundamental contributions have been made in a number of important areas, including seminal work on product-form queueing networks, such as the convolution algorithm (Hisashi Kobayashi and Martin Reiser) and the mean value analysis algorithm (Steve Lavenberg and Martin Reiser), on fast simulation of rare events (Phil Heidelberger), and on the theory of effective bandwidth (Cheng-Shang Chang). These research results have played a vital role in understanding important problems in the design, development, and management of complex systems.

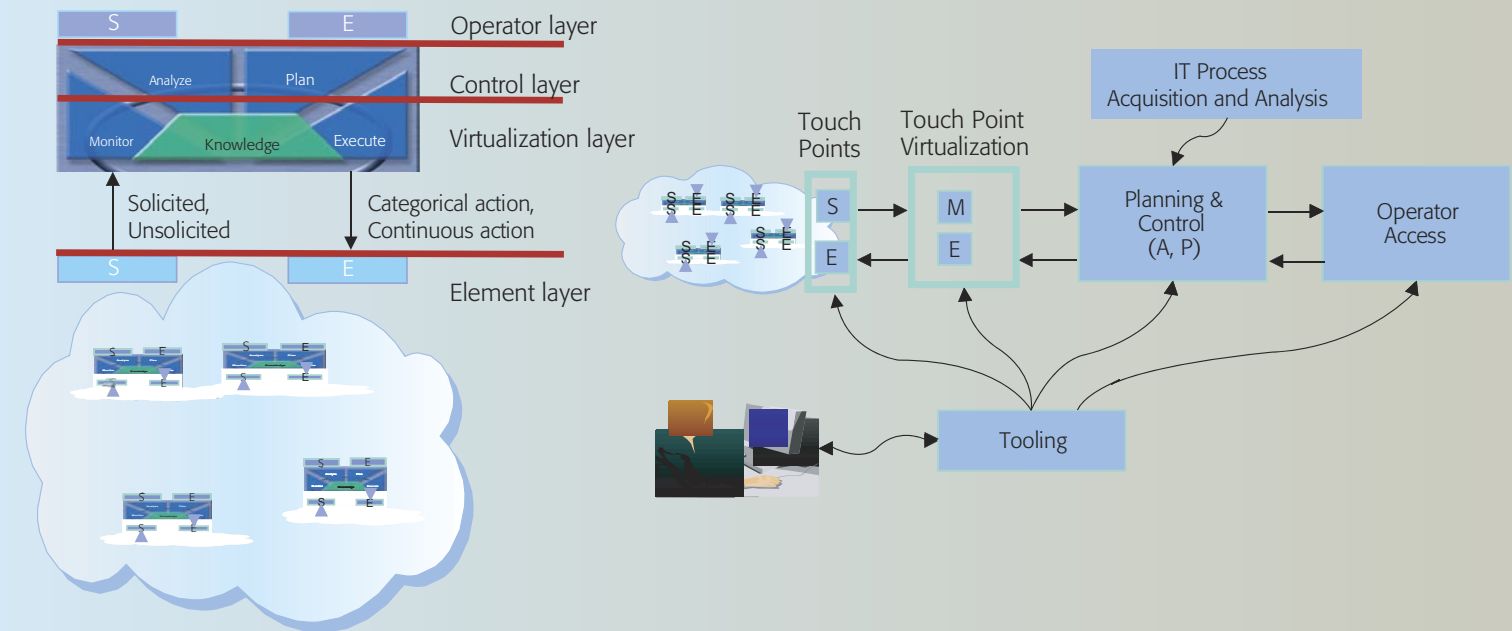
The rapid growth of technology brings many challenges to the performance community. With increased system capability and connectivity, more and more complex applications are developed and run on large-scale IT infrastructures, such as grid computing and on demand data centers. Performance modeling and analysis is crucial to the capacity management of such complex systems. Quantitative models are needed for hardware components and software applications, as well as for IT infrastructures and business processes.

#### STOCHASTIC MODELS AND QUEUEING THEORY

A key focus is stochastic models and queueing theory. IBM Research's analysis of data from a wide range of systems demonstrates complex arrival and service patterns that include long-range dependencies, heavy-tail distributions, and non-stationary effects. While the significant performance degradation from such behaviors is difficult to analyze with traditional methods, researchers have used approximate methods and asymptotic analysis to obtain analytical results for stochastic models and queueing networks with these complex characteristics. The results have delivered satisfactory application performance for IBM clients and allowed them to make better capacity planning decisions.

#### CONTROL AND OPTIMIZATION

Another focus is the control and optimization of performance measures in stochastic queueing networks. Recent and emerging applications, such as those related to the Internet, have raised issues regarding differentiated service and quality-of-service contingent on non-traditional metrics, such as service-level-agreements based on distributions as opposed to averages. These, together with the complex arrival and service processes found in practice, make many of the optimal control and resource allocation problems extremely difficult. Nonetheless, IBM researchers have developed provably optimal control policies for various queueing networks, in some cases based on asymptotic solutions.



#### PERFORMANCE TOOLS

Developing performance tools is an effective approach to deliver advanced research results and technologies to practitioners. IBM researchers have developed a suite of tools in many different areas, including traffic generation and benchmarking, automatic model building and calibration, capacity planning, workload and performance forecasting, resource control and management, as well as network and server design. These tools have helped IBM Global Services practitioners in performance testing, application sizing, and capacity planning. As computer systems and applications grow increasingly efficient and technically complex, accurately modeling and observing real-world system characteristics and behavior are critical facets of product development. To this end, researchers have developed a sophisticated full-system simulation infrastructure, which facilitates the precise study of processor and system behavior, and allows for a complete evaluation of future system designs. Widely used by IBM and its partners, this simulation infrastructure delivers a powerful and strategically important solution in several high-profile initiatives that are shaping the computing landscape, by enabling aggressive development and refinement of hardware designs, early evaluation platforms, and incorporation of vital feedback from software development teams.