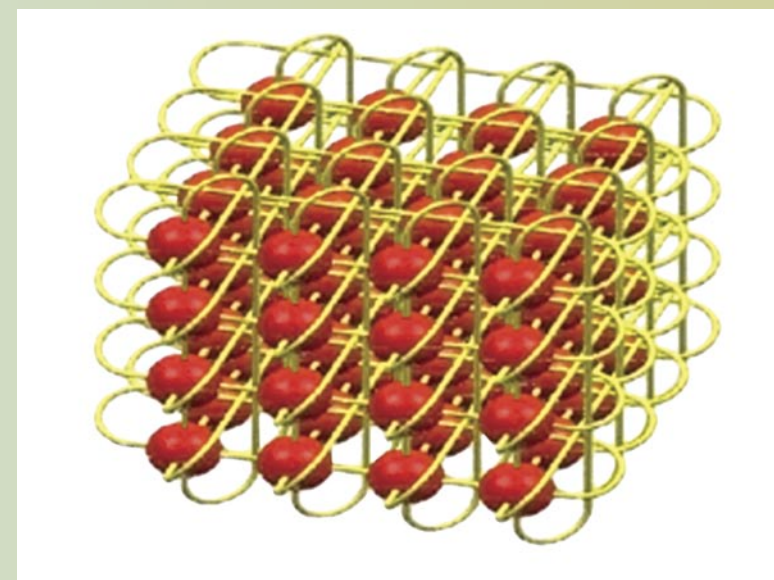


From its pivotal role in the NSFNET backbone and the Simple Network Management Protocol (SNMP) to recent key contributions to grid technologies as well as pioneering work in autonomic and on demand computing, IBM has a notable history in distributed computing. IBM's research spans a wide range of topics — clusters, distributed middleware, hardware and software fault tolerance, and distributed systems management — and addresses key challenges in transforming today's business IT environments to be more effective and responsive. These challenges include creating greater flexibility by integrating disparate business and IT processes, reducing costs and increasing responsiveness through automation, and improving capital and asset utilization through virtualization of resources. The projects in this area share a common, unified goal: providing manageable, reliable, and secure technologies that can deal with the increasing size and complexity of the information processing systems that businesses increasingly rely upon for everyday tasks. Our work has had significant impact on IBM middleware products, internal IBM IT processes, and services engagements with major clients.

LIFECYCLE MANAGEMENT OF SERVICES AND POLICIES

The construction, deployment, and management of large-scale distributed services is becoming ever more challenging. New paradigms for the structure and management of complex systems in which Research has played a significant role, such as Service-Oriented Architecture (SOA), autonomic computing, and on demand computing, require novel technologies for introspection, the separation of concerns, and adaptability. These technologies involve such concepts as model-driven development, management via Service Level Objectives (SLOs), and policy-based computing. IBM researchers are contributing to standards for expressing SLOs and to protocols for establishing agreements between entities within a system. In addition, researchers are contributing to standards for expressing policies and developing technologies for checking interactions between policies when they are authored, deployed, and executed. Among the operations provided are: dominance checking, coverage checking, conflict detection and resolution, and business value management.

Toroidal Interconnection in Blue Gene/L



SCALABILITY OF HIGHLY PARALLEL SYSTEMS

With 65,536 compute nodes and 1024 I/O nodes, Blue Gene/L sets a new bar of scalability for parallel computers. Current state-of-the-art approaches to the management of such large collections of computing resources have had only limited success on systems with more than 1,000 nodes. IBM's strategy with respect to system management in this environment is based on a hierarchical organization of the machine. This keeps the basic structure simple, allowing levels of scalability not previously attained. The Blue Gene/L job management system is responsible for dynamically allocating partitions and scheduling jobs to optimally use the machine's resources (compute nodes). This external mapping approach is thus able to consider more details of system operation in the optimization.

DISTRIBUTED MESSAGING SYSTEMS

In large distributed systems there may be millions of pieces of information and millions of consumers of that information who require it in a timely fashion. IBM researchers are focusing on developing advanced technology for messaging middleware for getting information to the right place at the right time. There are projects for developing highly scalable and available middleware for the publish/subscribe messaging model; for developing smart messaging capabilities for simpler and more efficient application integration; for extending the reach of asynchronous messaging to Web browser-based client platforms; and for developing techniques that support complex message flows based on service level agreements.

GOAL-DRIVEN CONFIGURATION AND PLANNING IN AUTOMATED SERVICE DEPLOYMENT

Distributed systems consist of multiple heterogeneous hardware and software components over which services are deployed. Creating instances of services requires the configuration of both the resources and the management servers that support them. Models have been developed for representing application components, a set of deployment goals and constraints, and a knowledge repository of best practices and deployment patterns. Using model transformation techniques, a set of deployment topologies satisfying all requirements and constraints can be automatically constructed. With these topologies, the deployment tasks themselves can be automated by dynamically creating workflows to control the sequence of activities performed while optimizing business objectives.

AUTONOMIC MANAGERS DYNAMICALLY ADJUST RESOURCES BASED ON BUSINESS GOALS

