PERFORMANCE AND ARCHITECTURE LABORATORY

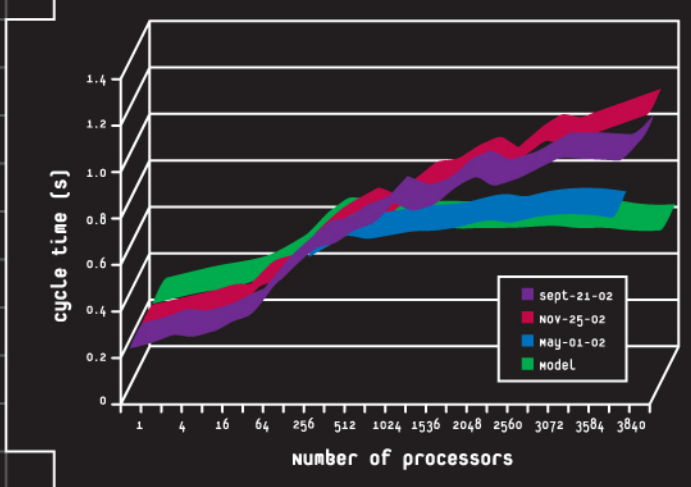
DEFINING THE FUTURE OF SUPERCOMPUTER SYSTEMS

The complexity of extreme-scale supercomputers and of the applications running on them is now comparable to, but less understood than, the physical systems those supercomputers simulate. Therefore, deeper understanding of computing systems has become crucial to the Laboratory mission, which relies to a great extent on large-scale, high-fidelity simulations.

The Performance and Architecture Laboratory (PAL) is a team of computer scientists dedicated to developing a theory of supercomputer performance. From that theory, we have developed a set of tools and techniques, as well as the associated system software, to guide the design of the next-generation extreme-scale supercomputing systems and applications.

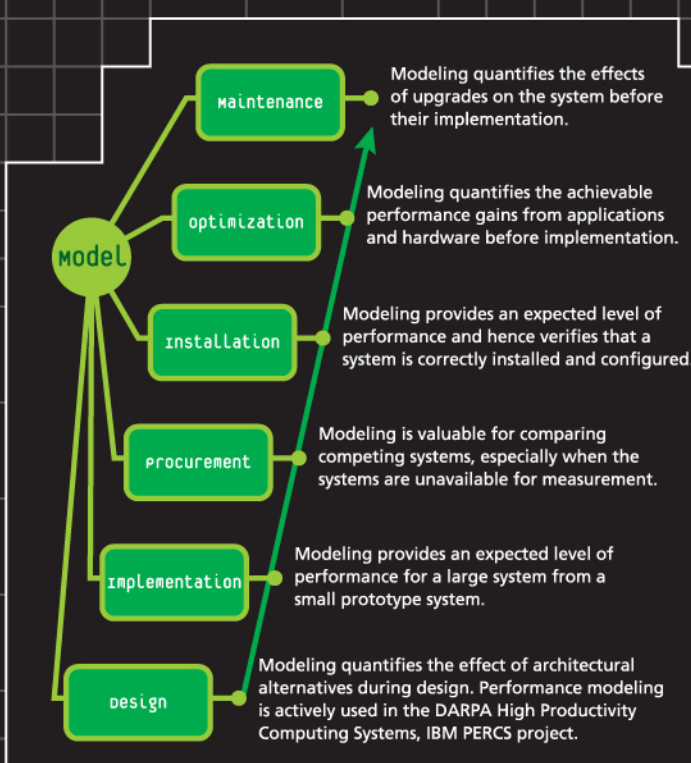
By studying and characterizing computer components, system software, applications, and their interdependencies, we provide insight into optimal designs for supercomputers. We discover performance-limiting factors that emerge only in large-scale systems, mitigate the negative effects of those factors and of the increasingly critical problem of system reliability, and develop performance models, our particular forte. The effectiveness of PAL's approach was validated by application of this methodology to the fastest supercomputer in use in the United States, the Q machine at Los Alamos, by improving the performance of important applications running on it by as much as 70 percent.

Our performance analysis and modeling work is the basis for another of PAL's major research thrusts: defining and building a system software infrastructure that optimizes the performance of future-generation supercomputers for the Laboratory's key applications. PAL is also a major partner in the High Productivity Computer Systems Project, sponsored by the Defense Advanced Research Projects Agency, which is chartered to build a petaflop-scale system by 2008.



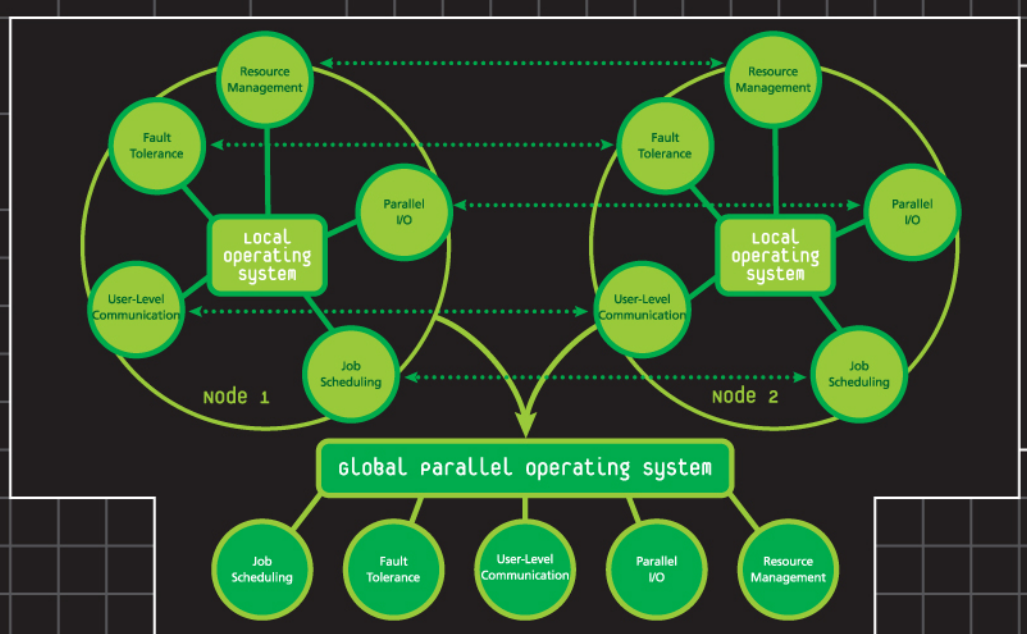
Improvement in Q Machine Performance

The green line represents the modeled runtime of an important application on the Q machine. The red, purple, and blue lines represent the performance of this application as optimizations to the system were applied. The difference represents performance improvement.



Applications of Performance Modeling

Performance modeling can be applied to all stages in the life cycle of a supercomputer to accurately analyze and predict system and software performance, even on supercomputers that do not exist yet.



Global Parallel Operating System

Large-scale, clustered supercomputers use a local operating system such as Linux, enhanced with a set of distributed cluster services. Because these services are independent of each other, redundancy and performance problems occur often. PAL is developing a parallel operating system that handles all the communications between nodes and contains the necessary infrastructure for all the cluster services.

our fastest supercomputer

This is a view of the Q machine, the fastest supercomputer in the United States.

Adolfy Hoisie, CCS-3
hoisie@lanl.gov, 505-667-5216

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36.
LALP-04-080

NNSA

The World's Greatest Science Protecting America

Los Alamos NATIONAL LABORATORY ESTABLISHED 1943