

Panel Discussion
Data Intensive vs. Scientific Computing:
Will the Twain Meet for Parallel Processing?

Moderator: Vipin Kumar, University of Minnesota

Panelists

Tilak Agarwal, *IBM*

David Bailey, *NASA*

Jim Gray, *Microsoft*

Olaf Lubeck, *Los Alamos National Laboratory*

Bob Lucas, *ARPA*

Tom Sterling, *Caltech*

Rick Stevens, *Argonne National Laboratory*

Hans Zima, *University of Vienna*

Much of recent research in practical parallel computing appears to be driven by the demands of the scientific and engineering domains. Grand Challenge Applications and, more recently, the DOE's ASCI project have provided a great impetus to the development of very high performance parallel architectures, programming environments, and applications for performing large-scale floating-point intensive simulation of physical phenomena.

But the market for such scientific applications appears small and rather limited. In contrast, the market for data intensive applications (e.g., data mining and decision support) is growing very rapidly, and such applications can potentially become the biggest consumers of parallel computing. The panel will address this dichotomy between traditional scientific computing and data intensive applications in the context of parallel processing.

What are the algorithmic features that make data intensive applications different than scientific applications? What demands do these features place on the architectures and environments of parallel computing? Are the desired features of parallel architectures and environments for data intensive applications fundamentally different than used for traditional scientific computing applications? Are these two sets of applications impacted differently by the deep memory hierarchies in current and future generation architectures (serial and parallel)?