

## Panel:

### Revisions to the IEEE 754 Standard for Floating-Point Arithmetic

Organized by:  
Dr. Eric Schwarz, *IBM Corporation*

Almost twenty years ago the IEEE 754 binary floating-point standard was adopted. Since then almost every microprocessor as well as many programming languages have defined the floating-point arithmetic to be IEEE 754 compliant. From the many years experience in implementing the standard in hardware and writing floating-point programs, there have been numerous suggestions for revisions. All IEEE standards must undergo a review process every 5 years or be dropped as an active standard. For past reviews this standard was extended without much discussion. But finally in January 2001 an in-depth review was started. A committee was formed and over the past two years many revisions have been evaluated. The most extensive change to the standard is to adopt formats for decimal floating-point arithmetic. This proposal creates decimal floating-point data formats for 32, 64, and 128 bits. Decimal floating-point arithmetic provides an exact representation of displayed numbers and provides a precise round at decimal radix point. This type of arithmetic is required in financial calculations. Some experts argue that decimal will replace binary due to its ability to represent decimal numbers exactly, while others think that binary will remain the key floating-point format due to its speed of execution and its more regular spacing of intervals. Another once controversial proposal is the addition of fused multiply-add. This operation only causes one rounding error, while in most implementations, provides twice the performance of separate operations. Other additions to the standard include a quadword format and many predicate functions such as comparison operators like greater than. Also operators for maximum and minimum have been accepted that after hours of arguing now favor a numeric result over a NaN. There are also deletions such as the single extended and double extended formats. And there are some items that are deleted in one meeting and resurrected in the following meeting such as signaling NaNs. Over the past two years of committee review there has been many proposals discussed. This panel discussion will enlighten the audience to the additions, deletions, and some of the current controversial proposals. The panel will consist of :

- David Hough, Sun Microsystems, Editor of the Standard – Overview
- Mike Cowlshaw, IBM Corp., Decimal Floating-Point Software Advocate
- David Bailey, Lawrence Berkeley Lab., Quadword Precision Advocate
- David Matula, Southern Methodist University, Academics / Industry Consultant
- Eric Schwarz, IBM Corp., Decimal Floating-Point Hardware – Panel Chair