The landscape of Parallel Computing Research: A view from Berkeley

by Yunqi Zhang
04/05/2013

1. Applications
   Trying to better express a parallel computation with 13 dwarfs.
2. Hardware
   a. Processors: building smaller processor building blocks
      i. energy-efficient
      ii. dynamic voltage scaling and power down
      iii. heterogeneity
   b. Memory: much larger fraction of total system silicon for memory
   c. Interconnection networks: linearly scale
      i. collective - small message size
      ii. point-to-point - large message size
      iii. exception: 3-D FFT
   d. Communication Primitives: Synchronization
      i. transactional memory
      ii. full-empty bits in memory
      iii. message passing
   e. Performance and energy counters:
      i. programmer and compiler are explicitly responsible for performance
3. Programming models
   a. psychological research (really???!!!)
   b. independent of the number of processors -> scalability
   c. rich set of data sizes and types
   d. support of proven styles of parallelism
4. System software
   a. Autotuners
      i. larger search space by parallelism
   b. Deconstructing the operating system support
      i. embedded systems increase in functionality
      ii. virtual machine: protection against attacks, failures don't interfere other programs
5. Metrics of success
   a. Maximizing programmer productivity
   b. maximizing application performance