Invited Presentation to SOS-10 :

"**Cooperative Computing**": Defining the 3rd Domain

Thomas Sterling
Louisiana State University
and
California Institute of Technology

March 8, 2006

Center for Computation & Technology

AT LOUISIANA STATE UNIVERSITY

“terminology” is a tool; it is how we use it

- Our field creates many new terms or adapts and changes old terms on a continuous basis as required
- There are no absolutes in terminology, only conventional usage
- There are underlying ideas
- Terminology helps us manipulate and interrelate those ideas, if used and applied effectively
- I think there are three ideas here
- They are getting scrambled by the conventional usage of only two terms
- I’d like to suggest we increase the power of our terminology to more effectively reflect the ideas
- There are no absolutes: these are only my opinions

The Three Ideas

1. A class of concurrent computing in which lots of tasks are performed simultaneously but are entirely independent except, perhaps but not necessarily, at their beginning and end points.
2. A class of concurrent computing in which a loose confederation of related subtasks are performed mostly independently but periodically coordinate and exchange intermediate results before proceeding, scaling as the size of the problem scales.
3. A class of concurrent computing in which the parallelism of a fixed size problem allows the time of execution to be reduced as additional parallel hardware resources are applied over at least a range of scale of the parallel hardware.

A Tale of Two Terms

- Capacity Computing
  - Throughput computing
  - Aggregate sustained performance to cost
- Capability Computing
  - Works on a single problem
  - Reduction in time to solution
- Casually derived, defined, and applied by the HPC community
- Taken on a quality attribute of their own
- But, most people feel demeaned if you refer to their machines as anything other than a capability system

The Middle Way: Cooperative Computing
also “coordinated computing”

- Concurrent execution of processes that operate primarily separately with periodic global (barrier) synchronizations and data exchanges
- Scales as problem size scales potentially over wide range – sustained performance per unit cost
- Scales for fixed size problems for narrower range limited by ratio of useful work critical path to overhead – reduction in execution time for single problem

And: I was wrong in the absolute sense

- There is a difference between capability/capacity computing and capability/capacity machines
- Its clear that a capability machine can perform a capacity workload
- A machine optimized for throughput workloads can achieve fixed problem size scaling over some range of system size
- And, both can engage in cooperative computing at least to some degree
- But optimization for any one degree of computing workload may detract from measure of merit for other two, either by providing inadequate functionality or mechanisms that are too expensive

Advances Needed

- For capacity computing
  - Minimize cost
  - Reduced power
  - RAS
- For cooperative computing
  - Faster per node performance and efficiency
  - Global bandwidth
  - Global barrier synchronization
- For capability computing
  - Low overhead mechanisms for managing fine grain parallelism and shared name space
  - Very high local memory bandwidth
  - Local and global latency hiding
  - Efficient fine grain messaging

But What I Really Think Is:

- Its not capacity vs. capability (vs. cooperative)
- Its about:
  - Memory capacity and bandwidth (local and global)
  - Memory access latency sensitivity
  - Parallelism management overhead
- It's a balance problem
  - Intrinsic to the application
  - Characteristic of any given machine
  - Exhibiting cost, power, and floor space
  - Mission time constraints

Questions to the Panelists

- What are the principal factors that distinguish capacity systems from other forms of computing systems?
- What is the role and impact of capacity computing for current and future scientific problems?
- What technical challenges confront the continued growth of capacity computing performance?
- What will be the dominant directions for future generation capacity computing and system types?
- Will the current division between capacity and capability computing be retained over the next decade or will there emerge a different useful distinction in form and function?

Panelists

- Thomas Sterling – Louisiana State University
- Jim Ang – Sandia National Laboratory
- Paul Durham – Daresbury Laboratory
- Bill Gropp – Argonne National Laboratory
- Andrew Jones – University of Manchester
- Steve Scott – Oak Ridge National Laboratory