Launching into the Future of Commodity Cluster Computing

Dr. Thomas Sterling
Center for Advanced Computing Research
California Institute of Technology
and
High Performance Computing Group
NASA Jet Propulsion Laboratory

Abstract

It is observed that once you've finally made it to the North Pole, no matter what direction your next step takes, you can't go anywhere but South. To some, it appears that Commodity Cluster computing is in somewhat of a similar position. Having firmly established itself as the leading approach to scalable high performance computing, commodity clusters such as Beowulf-class systems have no where to go. Yes, they will continue to track Moore's exponential growth in peak performance and storage capacity, and system software will become incrementally better. But in terms of strategic advances, commodity clusters are largely corralled in a paradigm cul-de-sac.

In some sense there is a kind of truth in this assessment: the fundamental premise of commodity clustering is invariant; Existing, i.e. predeveloped and standalone, systems (uninformatively referred to as "nodes") are integrated by means of a commercially available interconnect network and supporting software. However, the flaw of this defeatist perspective is that the nodes look like more of the same, only faster. The future of commodity cluster computing will be defined, not solely by Moore's Law, but by the revolutionary structures embodied by the nodes and the innovative ways in which they will be employed.

This plenary presentation will explore the innovations in both hardware and software that are likely to drive the future of commodity cluster computing throughout the rest of this decade and toward the trans-Petaflops performance regime of performance. In so doing, we will examine current projections of device technology to anticipate the performance, capacity, power, size, and cost curves of future commodity clusters. Perhaps of more impact are the changes anticipated in hardware architecture including blade technology, system and SMP on a chip, processor in memory architecture, and anticipated advances in networking including Infiniband and optical switching. As new uses are found for clusters with a rapidly expanding customer base including commercial and business communities, a combination of open source and commercial software tools will be developed for ease of use and high productivity as well as resource management and fault recovery. As system scale explodes even for moderate cost systems, the software tools to manage them will take on new responsibilities alleviating much of the burden experienced by today's practitioners. The talk will conclude with a look at some more bizarre possibilities made driven by other market and product trends. The future of clusters can show that even at the North Pole, with the right technology and vision, you can also go straight up.

Proceedings of the IEEE International Conference on Cluster Computing (CLUSTER'02)
0-7695-1745-5/02 $17.00 © 2002 **IEEE**