BSC RS: MANA-2.0: A Future-Proof Design for Long-Running MPI-based Computations in HPC

Objectives

Abstract: Support for long-running computations on supercomputers has long been a pain point. To maintain scheduling flexibility, sysadmins set a maximum resource allocation (e.g., 48 hours) for HPC jobs. Sysadmins also often offer short-duration queues at a discount (e.g., 2 hours at 75% discount) in order to make use of idle cycles. Transparent checkpointing offers the dream of robust, fault-tolerant long-running jobs at scale, that can be employed in either of the two types of queues.

MANA-2.0 (MPI-Agnostic, Network-Agnostic checkpointing) is an effort to achieve this dream. Like the original MANA academic prototype, MANA-2.0 operates over any MPI implementation and network interconnect that supports the MPI API standard. MANA-2.0 is also future-proof, in the sense that it runs independently of the underlying MPI and network libraries. Details of new algorithms required for its robustness will be presented.

MANA-2.0 is being tested on: (i) NERSC's Cori supercomputer (proprietary Cray MPI and Cray GNI network); (ii) NERSC's Perlmutter (#5 supercomputer; proprietary Cray MPI and HPE Cray Slingshot network); and (iii) CentOS Linux for other HPC sites. Like all large projects, this has been a years-long collaboration that is only now coming to fruition. The many participants will be credited in the talk.

Short Bio: Professor Cooperman currently works in high-performance computing. He received his B.S. from the University of Michigan in 1974, and his Ph.D. from Brown University in 1978. He came to Northeastern University in 1986, and has been a full professor there since 1992. His visiting research positions include a 5-year IDEX Chair of Attractivity at the University of Toulouse/CNRS in France, and sabbaticals at Concordia University, at CERN, and at Inria/France. In 2014, he and his student, Xin Dong, used a novel idea to semi-automatically add multi-threading support to the million-line Geant4 code coordinated out of CERN. He is one of the more than 100 co-authors on the foundational Geant4 paper, whose current citation count is 34,000. Prof. Cooperman currently leads the DMTCP project (Distributed Multi-Threaded CheckPointing) for transparent checkpointing. The project began in 2004, and has benefited from a series of PhD theses. Over 150 refereed publications cite DMTCP as having contributed to their research project.

Speakers

Speaker: Gene Cooperman, Professor of Computer Science, Northeastern University
Host: Leonardo Bautista, Computer Architecture for Parallel Paradigms – CS Visitor Researcher, BSC

Barcelona Supercomputing Center - Centro Nacional de Supercomputación