

Workshop on Resiliency in High Performance Computing (Resilience) in Clusters, Clouds, and Grids

Stephen L. Scott^{1,2} and Chokchai (Box) Leangsuksun³

¹ Tennessee Tech University, TN, USA

² Oak Ridge National Laboratory, TN, USA

³ Louisiana Tech University, LA, USA

Clusters, Clouds, and Grids are three different computational paradigms with the intent or potential to support High Performance Computing (HPC). Currently, they consist of hardware, management, and usage models particular to different computational regimes, e.g., high performance cluster systems designed to support tightly coupled scientific simulation codes typically utilize high-speed interconnects and commercial cloud systems designed to support software as a service (SAS) do not. However, in order to support HPC, all must at least utilize large numbers of resources and hence effective HPC in any of these paradigms must address the issue of resiliency at large-scale.

Recent trends in HPC systems have clearly indicated that future increases in performance, in excess of those resulting from improvements in single-processor performance, will be achieved through corresponding increases in system scale, i.e., using a significantly larger component count. As the raw computational performance of these HPC systems increases from today's tera- and peta-scale to next-generation multi peta-scale capability and beyond, their number of computational, networking, and storage components will grow from the ten-to-one-hundred thousand compute nodes of today's systems to several hundreds of thousands of compute nodes and more in the foreseeable future. This substantial growth in system scale, and the resulting component count, poses a challenge for HPC system and application software with respect to fault tolerance and resilience.

Furthermore, recent experiences on extreme-scale HPC systems with non-recoverable soft errors, i.e., bit flips in memory, cache, registers, and logic added another major source of concern. The probability of such errors not only grows with system size, but also with increasing architectural vulnerability caused by employing accelerators, such as FPGAs and GPUs, and by shrinking nanometer technology. Reactive fault tolerance technologies, such as checkpoint/restart, are unable to handle high failure rates due to associated overheads, while proactive resiliency technologies, such as migration, simply fail as random soft errors can't be predicted. Moreover, soft errors may even remain undetected resulting in silent data corruption.

The goal of this workshop is to bring together experts in the area of fault tolerance and resilience for HPC to present the latest achievements and to discuss the challenges ahead.

Resilience 2012 is the follow-on workshop to the successful Resilience 2011 held in conjunction with EuroPar 2012 in Bordeaux - France, Resilience 2010 held with GGCRID 2010 in Melbourne - Australia, Resilience 2009 in Munich - Germany, and the earlier Resilience 2008 held in conjunction with CCGrid in Leon - France.

Program

The Resilience 2012 workshop program included presentations of contributed peer-reviewed papers, as well as, several invited talks.

Contributed Papers

- *“High Performance Reliable File Transfers Using Automatic Many-to-Many Parallelization.”* Paul Kolano. NASA Ames Research Center (USA).
- *“A Reliability Model for Cloud Computing for High Performance Computing Applications.”* Thanadech Thanakornworakij, Raja Nassar, Chokchai Leangsuksun, and Mihaela Paun. Louisiana Tech University (USA).
- *“The Viability of Using Compression to Decrease Message Log Sizes.”* Kurt Ferreira, Rolf Riesen, Dorian Arnold, Dewan Ibtesham, and Ron Brightwell. Sandia National Laboratories (USA), IBM Research (Ireland), and University of New Mexico (USA).

Short Papers of Invited Talks

- *“Chaotic-identity Maps for Robustness Estimation of Exascale Computations.”* Nageswara Rao. Oak Ridge National Laboratory (USA).
- *“Programming Model Extensions for Resilience at Extreme Scale.”* Saurabh Hukerikar. University of Southern California (USA).
- *“User Level Failure Mitigation in MPI.”* Wesley Bland. University of Tennessee (USA).

Invited Talks

- *“Does Partial Replication Pay Off?”* Dorian Arnold. University of New Mexico (USA).
- *“Resiliency: Going Forward.”* Chokchai Leangsuksun. Louisiana Tech University (USA).