Extreme Scale Resilience Home

The Argonne National Laboratory/MCS/Extreme Scale Resilience group covers fault tolerance and resilience for HPC simulations and data analytics at extreme scale

Lead: Franck Cappello, ANL

Topics and people

- **Multi-level Checkpoint / Restart**: Bogdan Nicolae, Leonardo Bautista Gomez (Postdoc now at BSC), Franck Cappello
  - Main project: VeloC (ECP)
- **Lossy compression**: Sheng Di, Franck Cappello.
  - Main projects: EZ (ECP), CODAR (ECP)
- **Silent soft errors/data corruptions detectors and compression**: Sheng Di, Franck Cappello
  - Main project: Aletheia (NSF)
- **Failure characterization and prediction**: Sheng Di, Rinku Gupta, Franck Cappello
  - Main project: Catalog (DOE ASCR)
- **Failure modeling and fault tolerance optimizations**: Sheng Di
- **Fault tolerance protocols**: F. Cappello

Main collaborators: Marc Snir (ANL and UIUC), Jon Calhoun (Clemson), Bill Kramer (UIUC), Bogdan Nicolae (IBM Dublin), Thomas Ropars (EPFL), Amina Guermouche (UVSQ), Frederic Vivien (Inria), Yves Robert (LIP), Satoshi Matsuoka (Titech), Mitsuhisa Sato (U. Tsukuba), Omer Subasi (BSC), Osman Unsal (BSC), Leonardo Bautista Gomez (BSC)

Tools and software

- **SZ** (Error Bounded Lossy Compressor for floating point data sets)
- **Z-checker** (An lossy data compression assessment tool)
- **AID** (Adaptive Impact-Driven Detection) library for SDC detection
- **FTI** (operational prototype): Fault Tolerance Interface for multi-level checkpoint/restart (in memory checkpointing, checkpointing on remote nodes, erasure encoding, etc.)
- **HELO/ELSA** (operational prototypes): System event clustering and Failure predictor
- **MPICH-HFT** (prototype under development): Fault tolerant MPI with hierarchical fault tolerant protocol

Main collaborative activities

- Illinois-Inria-ANL-BSC-JSC-Riken-UTK Joint Laboratory on Petascale Computing

Recent Publications (from 2013)

- W. He, H. Guo, T. Peterka, S. Di, F. Cappello, HW Shen, Parallel Partial Reduction for Large-Scale Data Analysis and Visualization, in The 8th IEEE Symposium on Large Data Analysis and Visualization (IEEE LDAV) in conjunction with IEEE VIS 2018, Berlin, Germany, October 21, 2018.
- C. Wang, N. Dryden, F. Cappello, and M. Snir, Neural Network Based Silent Error Detector, in IEEE CLUSTER 2018, 2018, [best paper award (in the programming and system software track)]
- S. Di, F. Cappello, Fast Error-bounded Lossy HPC Data Compression with SZ. IEEE IPDPS 2016
- L. Bautista Gomez and F. Cappello, Detecting Silent Data Corruption for Extreme-Scale MPI Applications, EuroMPI 2015
- L. Bautista-Gomez and F. Cappello, Detecting and correcting data corruption in stencil applications through multivariate interpolation. FTS 2015 workshop at IEEE Cluster 2015
- L. Bautista-Gomez and F. Cappello, Exploiting Spatial Smoothness in HPC Applications to Detect Silent Data Corruption. IEEE HPCC 2015
- L. Bautista-Gomez, Franck Cappello, et. al. GPGPUs: How to Combine High Computational Power with High Reliability (Embedded Tutorial). Design, Automation & Test in Europe, DATE’14
- S. Di, S. Bouguera, L. Bautista Gomez, F. Cappello, Optimization of Multi-level Checkpoint Model for Large Scale HPC Applications, IEEE IPDPS 2014
- L. Bautista Gomez, F. Cappello, Improving Floating Point Compression through Binary Masks, Proceedings of IEEE BigData 2013
- S. Di, D. Kondo, F. Cappello, Characterizing Cloud Applications on a Google Data Center, short paper, Proceedings fo ICPP2013