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

**News**

**July, 2016** F. Cappello (and George Bosilca) Tutorial on Resilience for HPC at DSN 2016
**Mar. 2016** F. Cappello plenary session, CODA 2016, Santa Fe
**Jan 2016:** SZ lossy compressor available here

**Sept 2015** F. Cappello Keynote on Trust, Cluster 2015/FTS 2015: Fault Tolerance System workshop, Chicago

**May 2015** F. Cappello Keynote on Trust in results of numerical simulations and scientific data analytics at HIPEAC CSW & BR, Oslo (cancelled)

**April 2015** White paper on Trust in results of numerical simulations and scientific data analytics

**April 2015** AID (Adaptive Impact-Driven Detection) library for SDC detection

**Jan. 2015** Using Data Analytics to Detect Corruptions in Numerical Simulations, BDEC Barcelona, short presentation

**Nov. 2014** FTI 0.9.5 release

**Nov. 2014** 2nd workshop of the Joint-Laboratory on Extreme Scale Computing

**Nov. 2014** FTI Demo on the DoE booth at SC14

**Nov. 2014** FTI presented at Journée thématique «Impact des nouveaux calculateurs pour l’océan et l’atmosphère » by Julien Bigot (CEA – Maison de la Simulation)

**July 2014** Our journal article "Toward exascale resilience: 2014 update" is presented in HPC wire

**April 2014** Inria-UIUC/NCSA-ANL-BSC Join-lab workshop, June 9-11

**March 2014** Presentation of the G8 ECS project results at Kobe

**Feb. 2014** BDEC workshop at Fukuoka (Japan): The Need for Resilience Research in Workflows of Big Compute and Big Data Scientific Application

**Dec. 2013** Data compression algorithm based on masking presented at the DoE SC Associate Director's meeting

**Nov. 2013** Mini workshop on Resilience at the 10th workshop of the INRIA-Illinois-ANL joint laboratory on Petascale computing

**Nov 2013** Large presence of the ESR group at SC13: 2 Papers, 1 Panel, 2 bird of feather (BOF), 1 Emerging Technology Demo + Chairing the Test of time award

**Oct 2013:** PUF (Partnership University Fund) project “Preparing for Next Generation Numerical Simulation Platforms” accepted

**Oct. 2013:** F. Cappello Invited talk at the Extreme Scale CoDesign Meeting in China October, 2013

**Oct 2013:** The Paris project on Silent soft errors/data corruptions detection funded for 3 years

**Topics and people**

- **Multi-level Checkpoint / Restart:** Franck Cappello, Sheng Di (Postdoc ANL), Leonardo Bautista Gomez (Postdoc now at BSC)
- **Lossy compression for checkpointing:** Sheng Di (Postdoc ANL), Jon Calhoun (UIUC), Franck Cappello.
- **Silent soft errors/data corruptions detectors and compression:** Sheng Di (Postdoc ANL), Pierre-Louis Guhur (ANL), Franck Cappello
- **Failure characterization and prediction:** Sheng Di (Postdoc ANL), Rinku Gupta (ANL), Ana Gainaru (Ph. D. UIUC, now at Mellanox),
- **Failure modeling and fault tolerance optimizations:** Sheng Di (Postdoc Inria)
- **Fault tolerance protocols:** F. Cappello (ANL)

Main collaborators: Marc Snir (ANL and UIUC), Bill Kramer (UIUC), Bogdan Nicolae (IBM Dublin), Thomas Ropars (EPFL), Amina Guermouche (UVSQ), Frederic Vivien (Inria), Yves Robert (LIP), Satoshi Matsuoka (Titech), Mitsuhsa 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)**
- **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 Joint Laboratory on Petascale Computing**
- **G8 Extreme Scale Climate Simulations** with France, Germany, Spain, Japan and Canada
- **EESI2: European Exascale Software Initiative 2, EESI2 Resilience Working Group**
- **Partnership University Fund “Preparing for Next Generation Numerical Simulation Platforms”**
Recent Publications (from 2013)

- D. Tao, S. Di, Z. Chen and F. Cappello. **Exploration of Pattern-Matching Techniques for Lossy Compression on Cosmology Simulation Data Sets.** Proceedings of the 1st International Workshop on Data Reduction for Big Scientific Data (DRBSD1) in Conjunction with ISC’17, Frankfurt, Germany, June 22, 2017.
- S. Di, F. Cappello. **Fast Error-bounded Lossy HPC Data Compression with SZ.** IEEE IPDPS 2016
- J. Bautista Gomez and F. Cappello. **Detecting Silent Data Corruption for Extreme-Scale MPI Applications.** EuroMPI 2015
- L. Bautista-Gomez and F. Cappello. **Exploiting Spatial Smoothness in HPC Applications to Detect Silent Data Corruption.** IEEE HPGC 2015
- S. Di, E. Berrocal, K. Helsey, L. Bautista-Gomez, R. Gupta, F. Cappello, **Towards Effective Detection of Silent Data Errors for HPC Applications.** Poster, IEEE/ACM SC14
- 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 and F. Cappello, Detecting Silent Data Corruption through Data Dynamic Monitoring for Scientific Applications, Poster, to appread in Proceedins of ACM PPoPP 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
• B. Nicolaie, F. Cappello, AI-Ckpt: Leveraging Memory Access Patterns for Adaptive Asynchronous Incremental Checkpointing, Proceeding of ACM HPDC 2013
• M. S. Bouguerra, A. Gainaru, F. Cappello, L. Bautista Gomez, N. Maruyama and S. Matsuoka, Improving the computing efficiency of HPC systems using a combination of proactive and preventive checkpointing, Proceedings of IEEE IPDPS 2013
• M. El Mehdi Diouri, O. Gluck, L. Lefevre, F. Cappello, Towards an energy estimator of fault tolerance protocols, Poster, in Proceedins of ACM PPoPP 2013