Argonne NATIONAL LABORATORY

Taking on Exascale Challenges: Key Lessons and International Collaboration Opportunities Delivered by European Cutting-Edge HPC Initiatives

Franck Cappello
Argonne National Laboratory &
University of Illinois
cappello@anl.gov

U.S. DEPARTMENT OF ENERGY

Exascale Challenges

- Load balancing
  - Static reasons: process variation-gate length, wire width, etc.,
  - Dynamic reasons: power management, Turbo modes, etc.
  - Heterogeneous Applications (multi-physics, workflows)
- Resilience
  - not only for process crash, DUE and radiation induced SDC
  - But also for Bugs that could lead to systematic errors
- Plateauing of the storage bandwidth
  - Memory keeps increasing (X10 PB), storage I/O is not (1TB/s).
  - In situ data analytics, Burst Buffers, Compression (error bound lossy compression)

→ A lot of pressure on runtime systems and application developers

- **G8 ECS** (Toward Exascale Climate Simulation)
  - Important collaboration between NCAR and BSC, JSC (performance tools)
- **JLESC**: UIUC/NCSA, Inria, ANL, BSC, JSC, Riken Google: joint-lab (first link)
  - Research topics on
    - Computational sciences (Application, benchmarking, mini-apps)
    - Applied Mathematics (AD, Num. Libraries, Optimization)
    - Resilience
    - I/O, Storage, Visualization
    - Programming models and runtime
    - HPC Clouds
  - 2 workshops per year + student/researchers visits (1W, 1Y+)
  - 22 Bilateral projects. Outcomes: papers + software
  - 6 multi-institutions topics. Outcomes:
    - White papers (research challenges, methodology)
    - Repositories (mini-apps, mini workflows, benchmarks, data sets for experiments, etc.)
    - Events: a BOF as SC15 (Characterizing Extreme-Scale Computational and Data-Intensive Workflows)

- Example of collaborations:
  - Scalability Enhancements to FMM for Molecular Dynamics Simulations
  - Reducing Communication in Sparse Iterative and Direct Solvers
  - Energy efficiency and load balancing
  - Overlap Communication and Computation with Hybrid MPI+OmpSs
  - Algorithms for detecting and correcting silent errors
  - New Techniques to Design Silent Data Corruption Detectors
  - Programming Model Extensions for Resilience
  - Hybrid resilience for MPI + tasks codes
  - In situ visualization
  - Object-based storage system for HPC systems
  - Towards Interference-aware scheduling in HPC systems
  - Integration of BSC and Inria Tools (for performance/power modeling)
  - Extreme-Scale Workflow Tools: Swift, Decaf, Damaris, and FlowVR