Applications Breakout Report
Day #1

BDEC Workshop

Barcelona

29 January 2015

Role of Applications Breakout

- Define the scientific customers for exascale
  - Identify the ready, eager communities
  - Raise expectations for computational scientists and engineers who are "satisfied" with current scale
  - Discover new applications with new capabilities
- Interface outward to stakeholders
  - Importance of predictive science, whether from first principles or statistical tools
- Interface inward to algorithm developers
- Ensure scientific accountability
  - Validation and verification improved at the new scale and integrated with observation
  - What did we do with the last nine orders of magnitude? (importance of real data)

What do we add to the discussion?

- Call attention beyond the Google-type problems already optimized by commercial applications
- In scientific applications, exploratory searches are followed by confirmatory simulations
  - Correlation **and** causation

Application types (1/2)

- Third paradigm
  - PDE-based models
  - Particle-based models
  - Linear algebra-based models (e.g., DFT)
  - Image processing
- Fourth paradigm
  - Archiving and retrieving from massive data sets
  - Clustering
  - Searching

Application types (2/2)

- Combinations of Third and Fourth paradigms
  - Fourth informs Third
    - Inverse problems
    - Data assimilation
    - Visualization and computational steering
  - Third informs Fourth
    - Design of experiments
  - Both paradigms in a virtuous loop

Algorithmic Requirements

- Pipeline between simulation and analytics
  - Common data structures
  - *In situ* use of the data
- More efficient, scalable tools for the "inner" problems
  - Adaptive, scalable solvers
  - SVD, SVM,
  - Search algorithms
  - Graph algorithms
- Efficient I/O and strategic checkpointing for ever-increasing data sets
- Combination of the usual "-omics" with spatial dimensions

Programming environment requirements

- Portability (and performance portability)
- Usability
- *In situ* use of the data
- …

In/adequacy of existing national strategies and plans (1/3)

- Value-chain orientation focuses "big data" on the MapReduce-type problems
  - May miss some of the major scientific opportunities by being driven by the commercial opportunities
  - Commercial engineering drivers are growing, but not yet as exascale-intensive as the science drivers
- Big data and HPC are disconnected at the science policy and industry level in many countries
  - Big data is oriented towards social/business data
  - HPC is oriented towards scientific challenges
  - Potential of machine learning is underexploited in science and engineering

In/adequacy of existing national strategies and plans (2/3)

- Some calls are worded in ways that make scientists wary of bridging from HPC to big data
- Observational campaigns are sometimes inadequate (e.g., in resolution) to fully exploit or provide checks on simulational campaigns
  - And *vice versa*
  - Missed opportunities from lack of coordination or lack of synchronization in time or between nations

In/adequacy of existing national strategies and plans (3/3)

- Some calls are too narrowly posed to dynamically respond to scientific opportunities
- Funding for sustainability – code development and data curation – is rare compared to funding for research
- What lessons from the co-design centers for HPC can influence the big data campaigns, e.g., the EU Centers of Excellence

Comments to kick back to plenary

- Want to save full-resolution simulations, since the newest phenomena are in the smallest scales
- Cannot current cache the data stream

Questions to kick back to plenary

- Will the exascale machine look just like the petascale machines?
- Will the “killer apps” be something completely different from what we are planning on today?

Possible Case Studies emphasizing time to solution

- Complex simulations
- Merging of simulation with sensors
- Real-time aspects
- Examples
  - Weather forecasts
  - Wildfires
  - Earthquakes
  - Hazardous release evacuation
  - Reactor failure
  - Radiation treatment planning

Profile of the group

- Nations represented
  - EU: France, Germany, Spain, UK
  - Asia: Japan, Korea
  - USA
- Disciplines represented
  - PDE-based: combustion, meteorology, fusion
  - Observation-based: astronomy, biomedical informatics
  - Tools-based: PDE frameworks, runtime systems