Integrated Environments for Extreme-Scale Collaborative Science

Dec 1, 2011

Gabriele Garzoglio, Parag Mhashilkar, Ruth Pordes, Scientific Computing Facilities, Fermilab.

In this paper we posit that to fulfill the computational challenges to deliver extreme-scale science, computer science research must address the following needs for collaboration environments:

- A tighter integration across and between many tiers of computational platform (from leadership-class computing through to mobile devices)
- New paradigms for connecting data to processing software to exploit diverse classes of networks including terabit networks.
- New algorithms to transform the status of all infrastructural components into summary information that will enable decisions on resource allocation and usage.

Our arguments encompass large scale and medium scale experimental science, modeling and theoretical simulations, and are informed by the customer base we work with and support at Fermilab. These include the US CMS LHC collaboration, the Intensity Frontier collaborations, and Fermilab scientists in astrophysics collaborations – SDSS, DES, and emerging, LSST. The customers also include collaborative research in theoretical physics (LQCD), computational cosmology (Astrophysics Center) and accelerator simulations (SciDAC-2 COMPASS). The characteristics of these communities span:

- Large- to small-scale collaborative research across administrative domains, globally distributed in space and time zones.
- Long-life time collaborative projects with data and code being used for research of longer than a decade (and sometimes approaching two decades).
- Distributed computing and data systems that involve a significant diversity of types of computational resources for the end-to-end delivery of scientific results and publications.
- Ever larger simulation applications that fine-tune existing theoretical models with increasingly high precision and include fast turn-around to influence real time steering of devices based on the results.

A key element is the value of providing integrated environments across unique leadership-class computing, traditional high performance (multi-core) systems, national and international high throughput grids and clouds, local organizational campus infrastructures and faculty clusters, the individual researcher laptop and an increasing variety of mobile devices.

A Collaborative Environment for the Integrated Seamless use across these Resources is a Computer Science Research challenge and gap.

Characteristics of the challenges in collaborative tools that work across an integrated environment include:

- Mechanisms for data storage and management, workflow pause and restart, while waiting for resource allocation on the high end and unique capability systems.
- Mechanisms for simulation workflows to effectively migrate across computing tiers in different administrative domains on-demand to achieve better turnaround, higher precision levels and immediately synthesized results.

- Novel tools for the management of resource and workflow components allocation to achieve optimal effectiveness in total throughput in an environment of shared, bulk resources of similar performance and capability e.g. clouds and grids.
- Data transfer, storage and access capabilities that serve dynamically located and available mobile and laptop "personal" resources.
- A blueprint and principles that define an "engineered backbone" to merge, integrate and interface all the tiers. These should include interface specifications and best practices of integration to drive infrastructural fault tolerance and resilience.

In simple terms, the goal of the end user is the ability to get the most work done in as little time as possible. The functionality gap to achieve this goal can be filled by an increasingly complete knowledge about and availability of required data and the ability to steer and schedule the computations (jobs) easily, efficiently and successfully.

A key challenge to this seamless integration is due to the mix of network characteristics that will continue, and expand in scope, over the next decade. The tiers of computing above will be connected through these different classes of networks, from terabit backbones, to 100Gigabit local networks, all the way to dynamically attachable mobile devices (5G etc.). Research is needed to achieve this integration for:

- Seamless connection of data availability and job processing software.
- Asynchronous, co-scheduling of data movement and processing. Currently, the typical paradigms to serve data to processing software consist in transferring the data to pre-allocated computing resources (data pre-staging or job-initiated request). In the era of terabit-class networks, processing software and data placements will be handled asynchronously; the connection between the two will be handled by instantiating just-in-time on-demand data streams and registering processing software to them.
- Interpretation of information available from conditions and states from all components. For example, data from the networks, computing and data systems must be transformed to usable knowledge and affect in real time the allocation decisions across the computational infrastructure. Distilling and making available the knowledge on the condition of the network can affect where data streams and processing software are instantiated. [1]
- Management of resource reservation, usage and allocation based on "workflow" status and need. Requires tighter integration of platform tiers e.g. to execute each component of a single "workflow" on the most appropriate tier via a single user interface.
- The "engineered backbone". Interface definitions are essential to support the integration across the resource types and provide an environment where services can be continually enriched without impacting the currently available collaborative environments.

[1] We have a local paper (CortexNet) that proposes research activities to address some of these issues.