

Support for Domain Research Communities with Data Intensive Applications

Eric Boyd

October 13, 2008

Internet2 Fall Member Meeting



Definition: Data Intensive Applications

- Data transfer performance generally not available from out-of-the-box host configurations
- Network capacity to support sustained transfer rates measured in the hundreds of megabytes a second and greater peak rates
- Network capacity to move a few terabytes of data in hours instead of days
- Network loss far lower than would be well-tolerated by pedestrian use

Internet2 and the LHC (1)

- Several years ago, Internet2 embarked on a strategic outreach to the LHC Community
- Rich Carlson named as LHC ambassador
 - Eric Boyd, Jeff Boote, John Volbrecht, others supported outreach
- LHC was seen as being in the vanguard of domain research reliant on networking
- Goals were twofold:
 - Understand LHC community needs
 - Support the LHC effort

Internet2 and the LHC (2)

- Internet2 and LHC identified several common areas
 - Performance Monitoring (initial focus)
 - Dynamic Circuit Networks (emerging focus)

Internet2 and the LHC (3)

Outcomes in the Performance Arena:

- Close partnership with ESnet and DANTE
- LHC OPN, US LHC Atlas community, and US LHC CMS community have adopted perfSONAR
- Specific directives to deploy perfSONAR at Tier 1s and US Tier 2s
- US LHC Best Practices Guide for Performance Monitoring

Internet2 and the LHC (4)

Outcomes in the Dynamic Circuit Arena:

- Close partnership with ESnet and DANTE
- US LHC Atlas community and US LHC CMS community have expressed a strong interest in DCN
- Lambda Station and Terapaths have been modified to dynamically provision circuits

LHC Conclusions (1)

- Tight engagement with an application outreach community led to rapid adoption of advanced network cyberinfrastructure
- Network cyberinfrastructure has the potential to provide real impact on network experience
- Ongoing R&D efforts focused by engaging community need
- Degree of engagement is not sustainable over multiple virtual organizations (VOs)

LHC Conclusions (2)

- Paraphrasing LHC Participant:
 - “If we had this level of engagement from Internet2 a few years ago, we would have used the Internet2 Network rather than building our own.”

Other Ongoing Engagements with VOs

- The Physics community and the Large Hadron Collider (LHC) project with the large data transfers from CERN to US research sites
- The Radio Astronomers and the Very Long Baseline Array (eVLBI) with the large real-time data flows
- The Laser Interferometer Gravitational-Wave Observatory (LIGO) and the data flows from their detectors
- The Network for Earthquake Engineering Simulation (NEES) and their network needs for data transfer, video streams and collaboration between their distributed sites
- The Open Science Grid to create regular performance evaluation as they provide services to researchers

Relevant Strategies from Strategic Plan

- II. Provide researchers and scholars with the tools and support they need to envision and execute the next generation of collaborative discovery.
- III. Maintain vigorous partnership, outreach, and advocacy programs to jointly develop, promote, and share the transformational power of cyberinfrastructure throughout the communities in which we work and live.

Open Questions

- Which application communities should we target
 - Criteria for engagement?
 - Number of communities (breadth vs. depth)?
- How do we maximize availability of scarce resources in staff and community?

Strawman: Targeting Communities

- Moving forward, Internet2, will seek to positively impact a broader set of researchers and communities. Through this effort Internet2 will seek to engage research communities that:
 - Benefit from better support for data intensive applications
 - Use the Internet2 network for their research
 - Offer experience and opportunities that have potential for benefiting a broader community (e.g. multiplier effect)
 - Contribute to reducing the performance gap
 - Expand the scope of performance management (i.e. greater deployment of perfSONAR)

Strawman: Maximize Impact

- Internet2 will seek to effectively balance its limited resources in this area by focusing on:
 - Identifying and effectively addressing common performance inhibitors
 - Improving the awareness of the performance gap in scientific communities, so that they understand the need to organize around this issue
 - Encouraging the expansion of common infrastructure (i.e. perfSONAR and DCN) to effectively understand and improve network performance
 - Engaging communities that have sufficient resources to offer a greater return

www.internet2.edu

INTERNET[®]
2