LBNE Technology Demonstrator in a Nutshell

LBNE S&C Workshop

Feb 3sd, 2014

Maxim Potekhin

potekhin@bnl.gov

Overview

Definitions:

- “technology” in this context applies to distributed computing, i.e. processing + data
- “demonstrator” is not meant to represent the full production scale, but rather be a limited, but self-contained and operational system. It’s not a final or definitive prototype

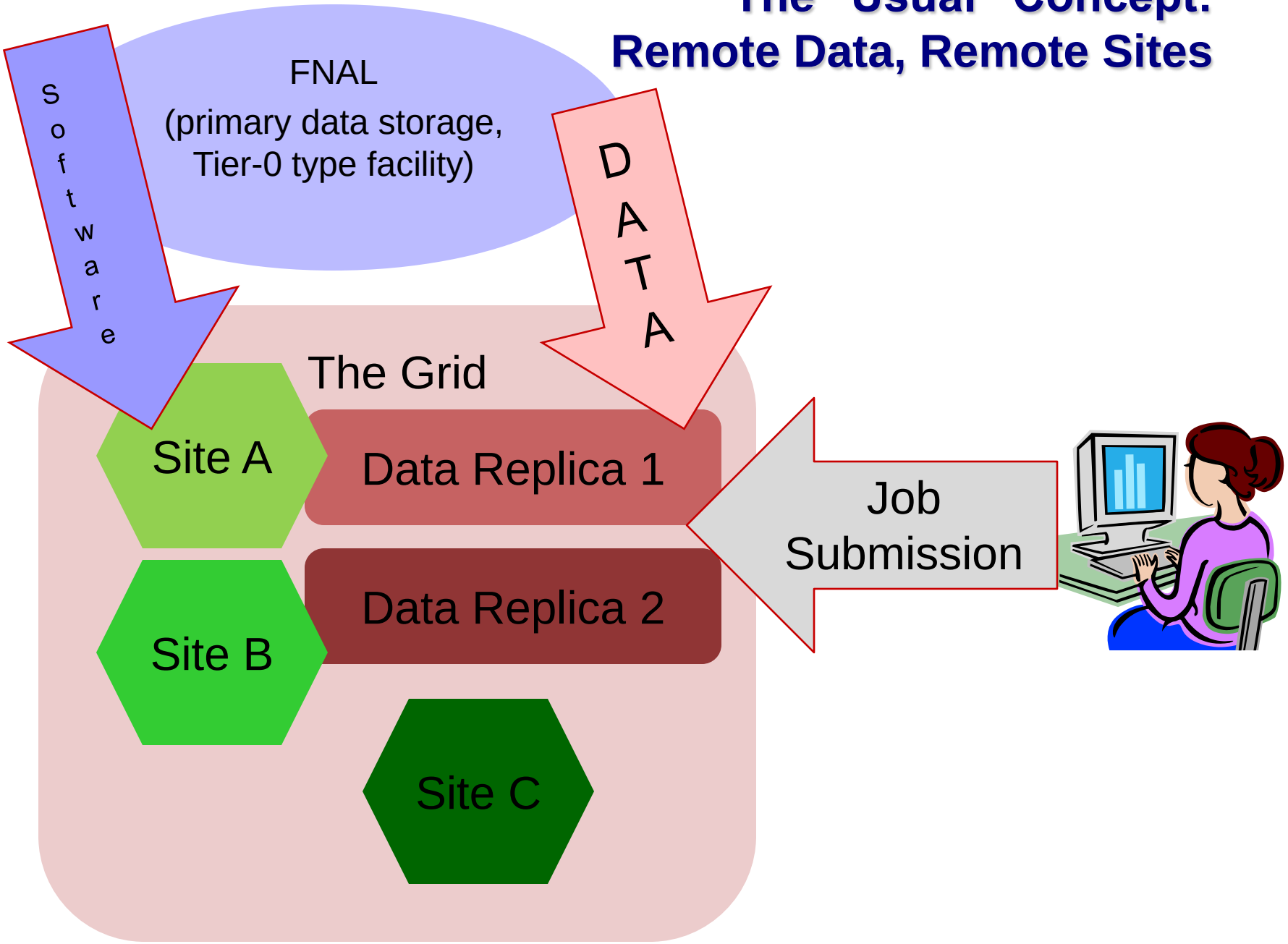
Why do we need it?

- Distributed computing is the norm nowadays, however in present initial stages of the LBNE S&C capability development, we still have not designed or implemented anything to meet the requirements as we currently define them (according to the S&C Plan and the Requirements document that is being finalized)
- A prototype of such system will help us identify limitations in our software organization, capabilities of participating institutions and to create momentum in moving towards meeting the goals of the distributed model

Is it something new?

- Of course not! It’s application of Grid Technology to meet the needs of LBNE

The "Usual" Concept: Remote Data, Remote Sites



What do we have now, what can we use?

Can go in a few directions:

- Open Science Grid software stack for most of functionality
- FIFE/jobsub – set of FNAL tools (already used?)
- Plain Condor-G with some wrappers etc

Most important progress areas in HEP computing in the last 2 years – at least two important systems (which of course existed for while) became robust and scalable enough to serve as a real foundation for our infrastructure

- CVMFS
- xrootd

Open Science Grid:

- Has experience with both technologies
- Has completed similar projects (to the LBNE demonstrator) in recent past, allowing for a lot of leverage in this new exercise

Conclusion

- A lot of technical details are still not defined and we need to address that, e.g.
  - Precisely defined MC payload and its requirements
  - Specific ways to ship and distribute data, modes of access to data
  - 100% clear understanding of CVMFS issues (OASIS etc)
  - Computation beyond MC, i.e. 35t data – code readiness for distributed operation?
- There is a lot of value in pursuing this project, on many levels