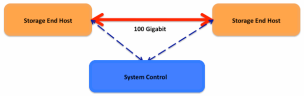
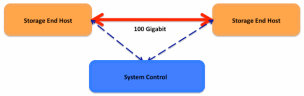


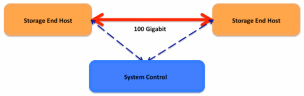
An Advanced Network and Distributed Storage Laboratory (ANDSL) For Data Intensive Science

Principal Investigator: Miron Livny University of Wisconsin-Madison
Contributing Senior Personnel:
Frank Würthwein, University of California San Diego
Matt Crawford, Ruth Pordes, Fermi National Accelerator Laboratory
Brian Tierney, Lawrence Berkeley National Laboratory



Overview

- One of the 4 ANI research projects.
  - 2 FTEs for testing and research.
- Project Start Date: September, 2009
- Funded by ARRA for 3 years
- Testing to the data storage and access profiles of the applications.
- Using scientific data patterns of OSG Communities.



Scope

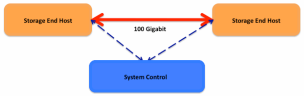
We will focus our effort on the software aspects – workload generators, monitoring tools and integration with the VDT -of the proposed program of work.

We will use storage and compute capabilities deployed by other related efforts as these capabilities become available and connected to the 100Gb testbed.

We expect that such “end-point” capabilities will be deployed at ANL and LBL and connected to the 100 Gb testbed.

We will also explore opportunities to leverage storage capabilities at FNAL.

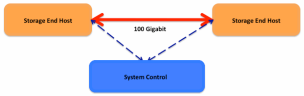
We will use our close ties with the ESnet team at LBL to coordinate the ANDSL work with related ESG activities.



Objectives

We will work closely with the OSG stakeholders to profile a broad set of real-life data intensive science applications as well as experiment with advanced storage and data management technologies.

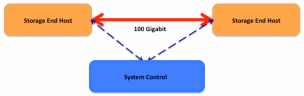
The goal is to close the adoption gap, decreasing the time between availability of production 100Gbps links and the end-to-end use of such a network to enable scientific discovery.



Plan SD = Start Date +x in months;

Table 1: High Level Deliverables and Milestones

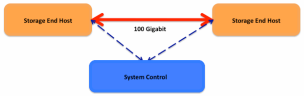
Deliverable	Milestone
Documented analysis of available relevant software	SD+2
Document of science community use cases to be studied.	SD+3 and incrementally every 6 months.
Documented design of and plan for software release of data and job workload generators, profiling and analysis tools.	SD+3 and incrementally every 3 months.
Release of first ANDSL software for studies	SD+6
Report of results of studies	SD+6 and incrementally every 3 months..
Documented plan for annual program of work	SD+11 and SD+23
Final report of the studies and publications of the Laboratory	SD+36



Effort Allocation

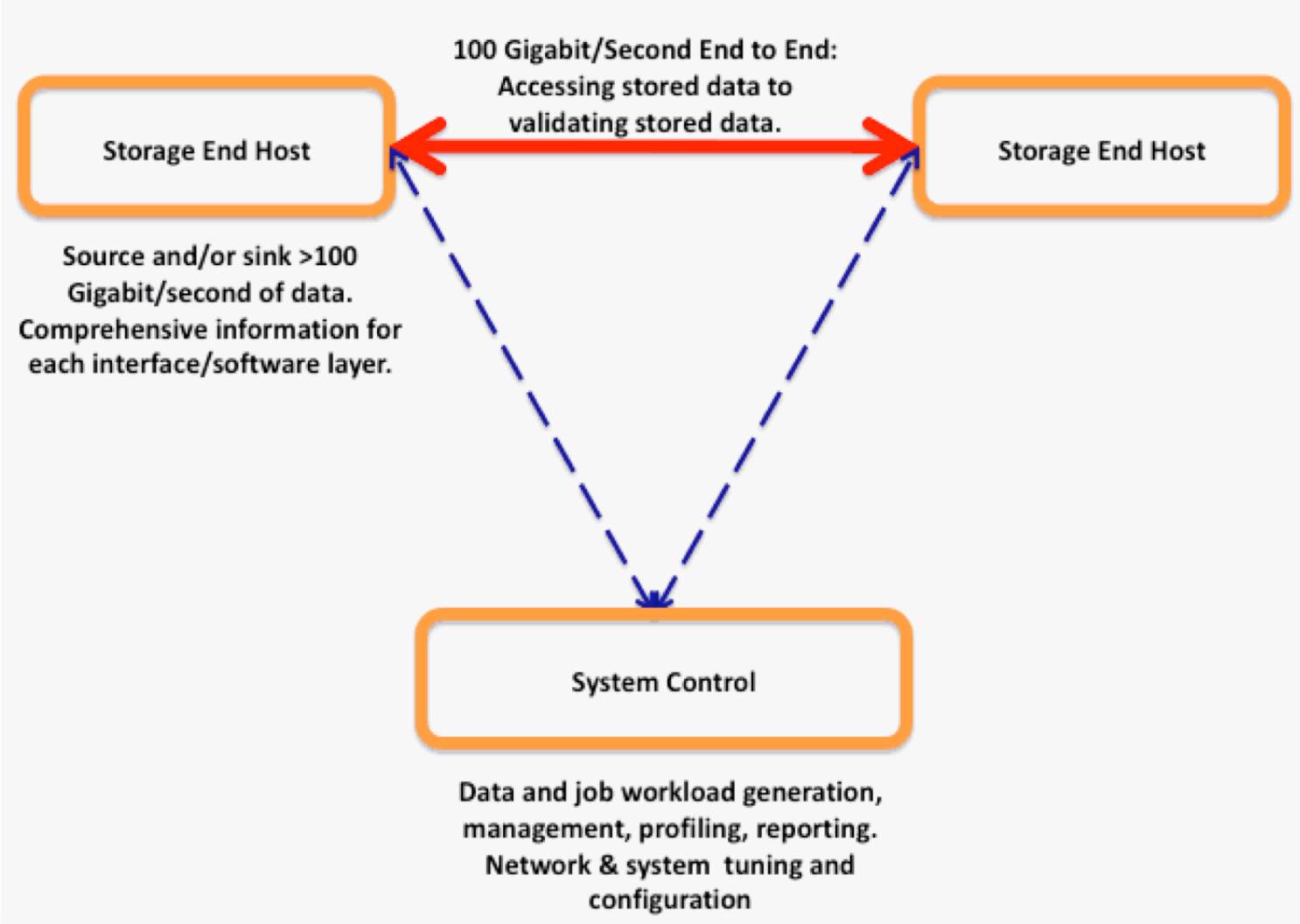
Table 2: Effort Profile

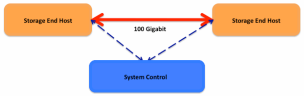
Area of Work	FTE Months
Determination, measurement and documentation of the data access throughput, methods and patterns of the science stakeholders.	3
Development and packaging, in phases, of parameter driven work and data load generators to test and profile the data and job access, management and throughput needs.	12
Extensions of the middleware, file systems, I/O libraries, including working with the applications communities.	9
Development of profiling and performance measurements tools for use with the generators.	9
Running of the generators on the 100 Gigabit prototype network and other new capabilities as new services, capacities, end-point hosts and configurations become available.	12
Analysis of the results and iterative tuning of the tests.	10
Reporting and publication of the results,	10
Coordination with stakeholders, fabric and resources providers, partners, rest of OSG project. Project tracking, reporting and planning. Architecture and software design.	7
Total FTE months.	72



Rely on 2 Magellan storage end hosts + others as they are available.

Figure 1: ANDSL Implementation Architecture





Program of Work

Develop tools for testing data workloads on a variety of network and data source/sink topologies.

Test end-to-end data transfer systems interfaced to aggregated 20, 40, 50 and 100 Gigabit capacity networks.

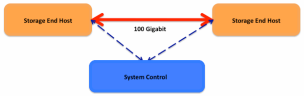
Make test harnesses useful for profiling existing production systems.

Typical university LHC Tier-2 is ~250 TB disk storage, ~1000 cores; and activities ~4,000 jobs a day and ~3,000 of file transfers a day.

Outside of the program of work:

No development of the software in the software stack; no device driver, GridFTP or other middleware development

Profiling local (posix) data I/O rates and access patterns. We will focus on bulk data movement.

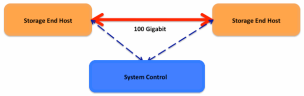


Phased Approach given timeline for hardware being made available

Initial work already underway:

Identify bottlenecks in current 10 G system at UCSD with 700 TB data store in Hadoop.

Extend tests and profiles to ~40G locally at UCSD in collaboration with SDSC, covering different data movement protocols (including Caltech FDT), hardware configurations, and using realistic workloads from the CMS experiment.



Also

Rely on CEDPS for any GridFTP extensions needed; Work with FNAL dCache and LBNL Bestman for any parochial GridFTP changes needed.

Work with ANI Testbed group on profiling and monitoring tools at the operating system and network interfaces layer. (Will need access to the bare operating system for tuning and analysis.)