The ARRA ANI Network Testbed Project

Feb 2, 2010

Supporting Advanced Scientific Computing Research • Basic Energy Sciences • Biological and Environmental Research • Fusion Energy Sciences • High Energy Physics • Nuclear Physics

Brian L. Tierney
bltierney@es.net
Overview

• Project Start Date: September, 2009
• Funded by ARRA for 3 years
  – Hopefully will continue beyond ARRA
• Designed, built, and operated by ESnet staff
• 1 of 3 ARRA “Advanced Network Initiative” (ANI) projects in the DOE
  – ANI 100G Prototype
  – ANI Network Testbeds
  – 4 ANI research projects
ANI 100G Prototype Network

- The “100G Prototype” network will connect DOE Supercomputer Centers at NERSC, Argonne, Oak Ridge, and MANLAN at 100Gbps
  - Not for production ESnet traffic
- Testbed will use 100G waves from the prototype network
- Testbed will use 100G routers from the prototype network
2 Testbeds Components

- Network Research Testbed
  - Phase 1: “tabletop” testbed at LBNL
  - Phase 2: Deployed on ESnet dark fiber in Long Island

- Application / Middleware Testbed
  - ANI 100G network connecting Magellan resources
Relation to the Magellan Project

- High-speed storage resources for the ANI Application Testbed will be provided by the ARRA-funded Magellan Project
  - DOE computing project funded at $33M
- Magellan is:
  - a research and development effort to establish a nationwide scientific mid-range distributed computing and data analysis testbed.
  - Consists of two sites
    - NERSC / LBNL
    - ALCF / ANL
  - Will provide multiple 10’s of teraflops and multiple petabytes of storage, as well as appropriate cloud software tuned for moderate concurrency.
Nationwide 100G Prototype Network

Sunnyvale

Magellan

NERSC

ALCF / ANL

Magellan

Chicago

NYC

OLCF / ORNL

Nashville
ARRA/ANI Testbed Goals

• Enable network, middleware and application end-to-end R&D at 100G

• Configurable

• Breakable
  – Isolated from the production network

• Reservable

• Easy to reset to known state

• A community network R&D resource
  – open to researchers and industry to conduct experiments
Testbed Requirement:
Support Network Research

• Support research on:
  – Data Plane
  – Control Plane
  – Management Plane
  – AA Plane
  – Service Plane

• Ability to do multi-domain hybrid networking
  – Support R&D on multi-layer and multi-domain control and signaling at layers 1-3

• Ability to reserve and manage each components

• Support protection and recovery research

• Support interoperability testing of multi-vendor 100G network components

• Provide access to all available monitoring data
Testbed Requirement:
Support Middleware/Application Research

- Ability to support end-to-end experiments at 100 Gbps
  - Connect to Magellan systems at NERSC (Oakland, CA) and ALCF (Argonne, IL)
Network Testbed Components

• Network Testbed will consist of:
  – DWDM devices (Layer 0-1)
    • GMPLS-enabled
  – Layer 2 switches supporting Openflow
  – Layer 3 Routers
    • Juniper M7i
    • OSCARS compatible, MPLS-enabled
  – Test and measurement hosts
    • Virtual Machine based test environment
    • 2x10G test hosts initially
      – Eventually 100G from Acadia 100G NIC project
Phased Approach

• Phase 1:
  – “Tabletop Testbed” at LBL
  – Ability to simulate multiple domains
  – MPLS enabled routers
  – Secure access via a gateway node from anywhere
  – Scheduled via simple calendar
  – Available to researchers this summer
Tabletop Testbed Configuration
Sample Configuration: Multi-Domain Multi-Layer Protection Testing

North Domain
- Openflow Switch
- north-wdm1
- north-wdm2

South Domain
- south-wdm1
- south-wdm2

East Domain
- east-wdm1
- east-wdm2

10G Tester

Sample Configuration:
- Multi-Domain
- Multi-Layer Protection Testing

Test inter-domain optical protection schemes

Test inter-domain higher layer (> 1) protection schemes
Phased Approach

• Phase 2:
  – Nodes are deployed on Long Island MAN
    • This is where ESnet will have dark fiber
  – 1 or more 10G Ethernet circuits between each node
    • Hope to upgrade segments to 100G as hardware becomes available
  – Ability to simulate multiple domains
  – MPLS enabled routers
  – Scheduled using OSCARS-like system
  – Available Early 2011
ANI Prototype Network Architecture – current plan

Key
- **Layer 1 device (vender TBD)**
- **Testbed Layer 1 device (Infinera DTN)**
- **Site Router**
- **ESnet Router**
- **Exchange Point Router**
- **100GE**
- **ESnet managed fiber**
- **Testbed Node (test and measurement hosts)**
Prototype or Testbed?

• Some confusion on these terms:
  – Prototype: 100G network connecting 3 Supercomputer Centers and MANLAN
  – Network Testbed: Network devices and hosts that researchers can configure
    • 1st in Tabletop, later in LIMAN (Long Island MAN)
  – Application/Middleware Testbed: Magellan connected by prototype network
Testbed Access

• Currently there are 9 DOE-funded projects that will be the first to get access

• Testbed will be accessible to anyone:
  – DOE researchers
  – Other government agencies
  – Industry

• Must submit a short proposal to the testbed review committee
  – Committee will be made up of members from the R&E community and industry
Timeline: Phase 1 (Tabletop Testbed)

- Project start
- Network device selection
- Hardware arrives; begin install and config
- Application testbed (Magellan) available, 10G
- Simple management software developed
- Tabletop Testbed available to DOE-funded researchers
- Start accepting Testbed proposals
- Hold Testbed Workshop
Timeline: Phase 2

- **Final WAN deployment planning based on ANI Prototype awards**
  - Oct 2010

- **Testbed available to 1st group of non-DOE researchers**
  - Jan 2011

- **Magellan-based Application Testbed available at 100G**
  - **Long Island dark-fiber based network testbed available to researchers**
  - Apr 2011

- **Testbed available to 1st group of non-DOE researchers**
  - Jul 2011

- **Accept 2nd round of Testbed Proposals**
  - Oct 2011

- **Deploy v2 of Testbed Management Software**
  - Jan 2012

- **2nd round Testbed users get access**
For More Information

• Contact: bltierney@es.net

• https://sites.google.com:443/a/lbl.gov/ani-100g-network/
  – (temporary location)

• http://twitter.com/100Gnetwork
EXTRA SLIDES
Testbed to support DOE ARRA-funded Research Projects

- **Climate 100 (LLNL, LBL, ANL)**
  - Project Goal: Scaling the ESG to 100G
  - Testbed role: provide interconnect between “Magellan project” resources at ANL and NERSC

- **Advanced Network and Distributed Storage Laboratory (OSG: U Wisc, FNAL, etc)**
  - Project Goal: enhance Virtual Data Toolkit (VDT) data management tools to effectively utilize 100G networks
  - Testbed role: provide interconnect between “Magellan project” resources at ANL and NERSC

- **100G NIC (Acadia and Univ New Mexico)**
  - Project Goal: produce a host NIC capable of 100 Gbps
  - Testbed role: Provide test environment for this device

- **100G FTP (BNL/Stony Brook University)**
  - Project Goal: design 100G transport protocols and tools to enable 100G FTP server to client data transfer
  - Testbed role: Provide a 10G test environment
Testbed to support DOE Network Research Projects

• “Resource optimization in hybrid core networks with 100G links” (Univ Virginia)
  – Project Goal: Design a hybrid network architecture that scales to multi-Tb/s and supports both IP-routed and dynamic-circuit services efficiently
  – Testbed role: Provide a test environment for validating these methods
• “Integrating Storage Resource Management with Dynamic Network Provisioning for Automated Data Transfer” (BNL, LBL)
  – Project Goal: Integration of dynamic virtual circuits into BestMAN (Berkeley Storage Manager / SRM v2.2)
  – Testbed role: Provide a 100G test environment for verifying this work
• “Provisioning Terascale Science Apps using 100G Systems” (UC Davis)
  – Project Goal: Advanced path computation algorithms for efficiency and resiliency
  – Testbed role: Provide a control plane test environment for these algorithms using OSCARS
• Virtualized Network Control (ISI, ESnet, Univ New Mexico)
  – Project Goal: multi-layer, multi-technology dynamic network virtualization
  – Testbed role: provide a control plane test environment for experiments using hybrid networking

• “Sampling Approaches for Multi-domain Internet Performance Measurement Infrastructures to Better Serve Network Control and Management” (Ohio State Univ)
  – Project Goal: Develop new network measurement sampling techniques and policies
  – Testbed role: Provide a network environment for initial deployment
100G Prototype Network Design Final Baseline Design
Intra-Domain Protection Testing

Test optical protection (i.e. GMPLS link/node protection), path selection efficiencies
Intra-Domain Multi-Layer Circuit Provisioning

Test multi-layer circuit provisioning (e.g. IP over Ethernet over Optical)
Multi-Domain Multi-Layer Circuit Provisioning

North Domain

- Openflow Switch
- north-wdm1
- north-wdm2
- 10G Tester
- 1GE
- 10GE
- 2x 10GE

South Domain

- south-wdm1
- south-wdm2
- Openflow Switch
- 1GE
- 10GE

- Optical (layer 1) inter-domain data plane exchange

East Domain

- east-wdm1
- east-wdm2
- 10G Tester
- 1GE
- 10GE

- 10GE

- Test inter-domain multi-layer circuit provisioning

- Ethernet (layer 2) inter-domain data plane exchange

- 1GE
Intra-Domain Flow Selection Switching

“Regular” traffic traversing best-effort provisioned (i.e. shared) path

Openflow switch determines path of “regular” and “selected” traffic based on input from App host

“Selected” traffic traversing guaranteed-service provisioned path

FS/BS/App host