ESnet On-Demand Secure Circuits and Advance Reservation System (OSCARS)

GridNets 2006
Oct 1-2, 2006

Chin Guok (chin@es.net)
ESnet Network Engineer

David Robertson (dwrobertson@lbl.gov)
DSD Computer Software Engineer

Lawrence Berkeley National Laboratory

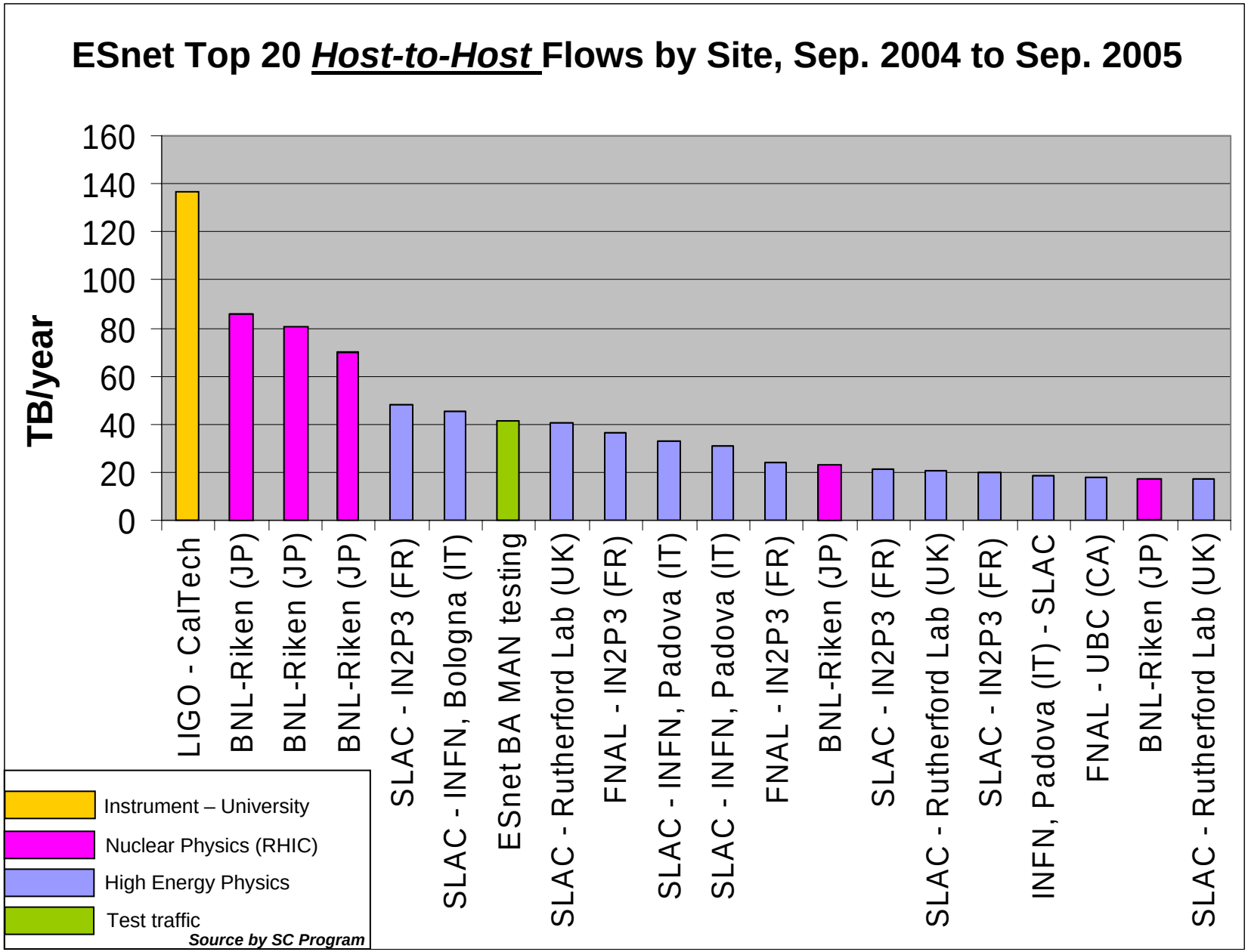
Outline

- Requirements for Virtual Circuit Services
- OSCARS Architecture
- Inter-Domain Reservations: Tough Problem
- OSCARS Collaborative Efforts
- OSCARS: Guaranteed Bandwidth VC Service for SC Science

➢Requirements for Virtual Circuit Services

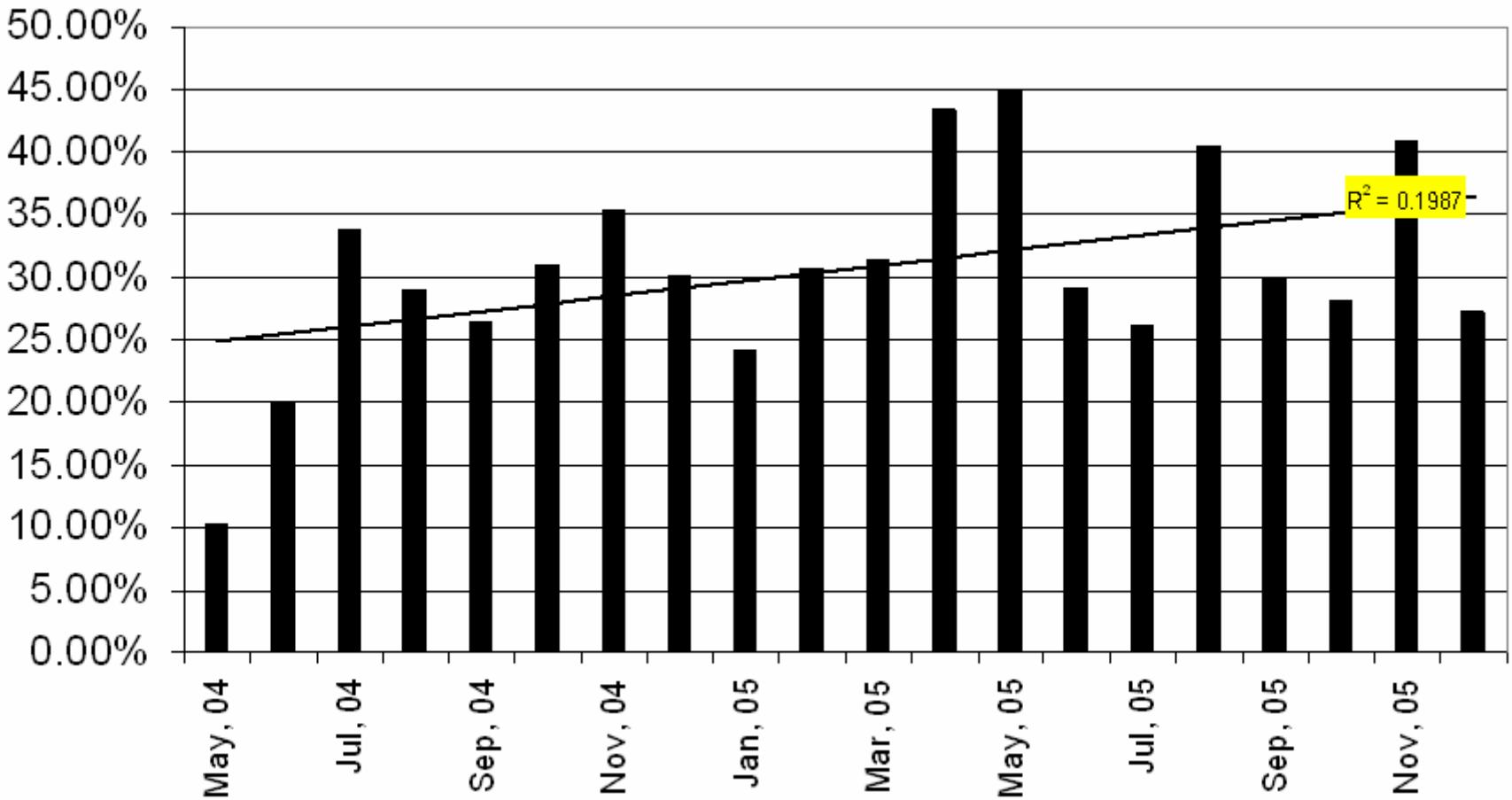
- **Identified as one of the two most important new network services by the 2002 High-Performance Networks Planning Workshop sponsored by the U.S Department of Energy, Office of Science** (Ref-1) ***(the other being end-to-end performance monitoring)***
- **Today**
  - Primarily to support bulk data transfer with deadlines
- **In the near future**
  - Support for widely distributed Grid workflow engines
  - Real-time instrument operation
  - Coupled, distributed applications
- **To get an idea of how circuit services might be used to support the current trends, look at the one year history of the flows that are currently the top 20**
  - Estimate from the flow history what would be the characteristics of a circuit set up to manage the flow

ESnet Large-Scale Science Flows by Site



ESnet Top 100 Flows as Fraction of Total

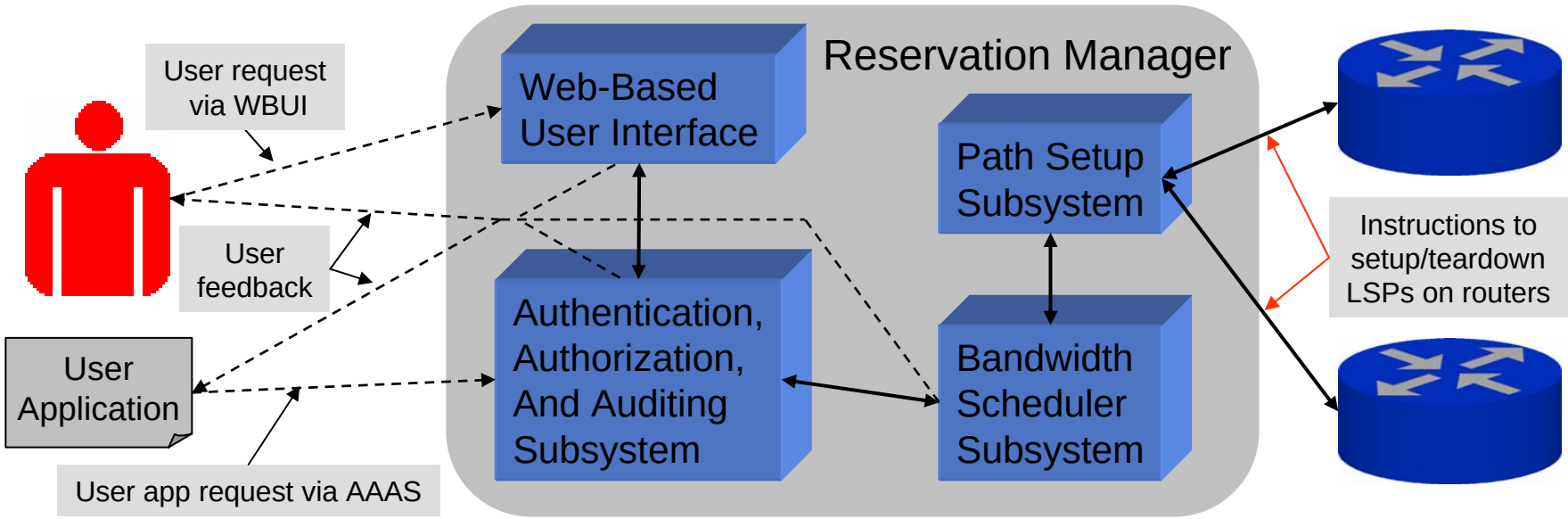
- Plot of the top 100 flows, by month, as a % of total traffic
  - This does not include production LHC flows
  - A steady increase



➢OSCARS Architecture

Reservation Manager (RM) Components:

- Web-Based User Interface (WBUI) will prompt the user for a username/password and forward it to the AAAS.
- Authentication, Authorization, and Auditing Subsystem (AAAS) will handle access, enforce policy, and generate usage records.
- Bandwidth Scheduler Subsystem (BSS) will track reservations and map the state of the network (present and future).
- Path Setup Subsystem (PSS) will setup and teardown the on-demand paths (LSPs).



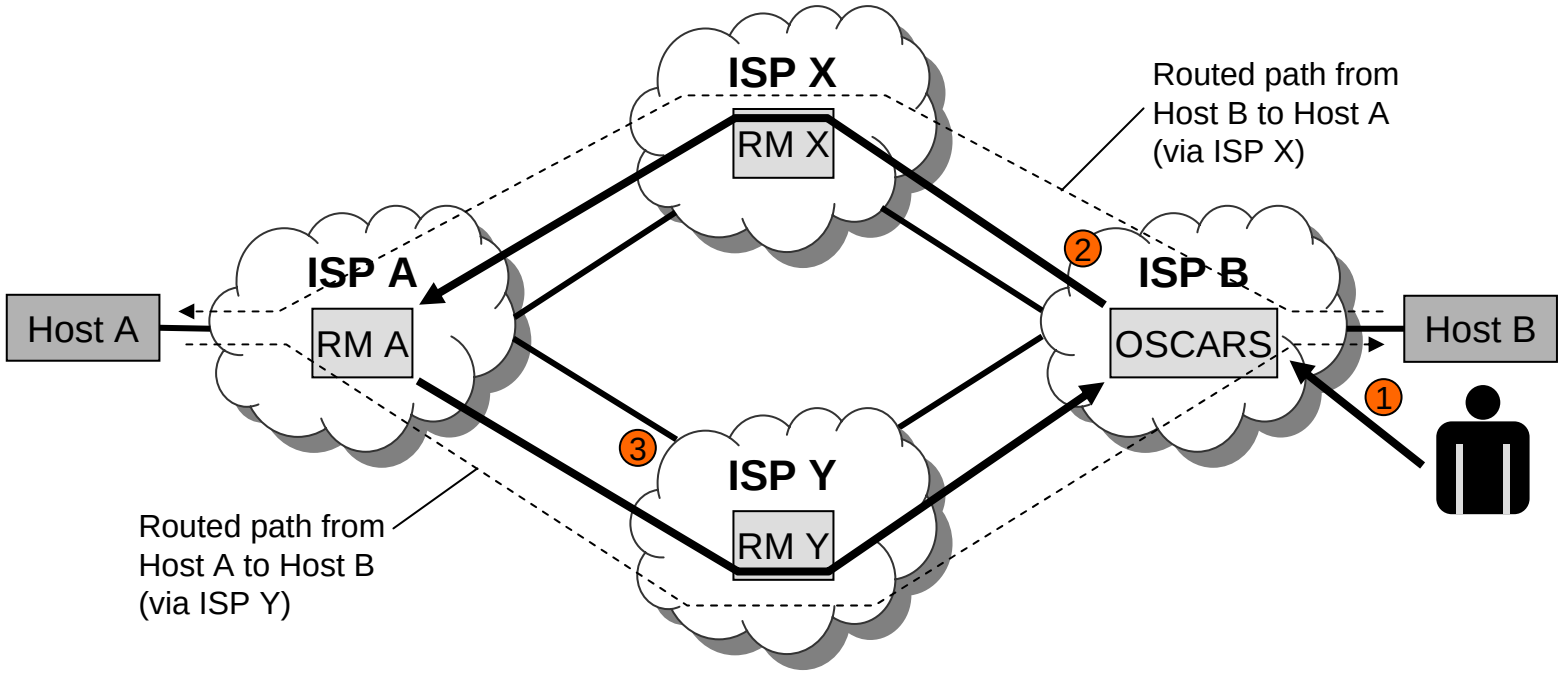
OSCARS Reservations

1. A user submits a request to the RM specifying start and end times, bandwidth requirements, and the source and destination hosts
2. Using the source and destination host information submitted by the user, the ingress and egress border routers, and the circuit path (MPLS LSP) is determined
3. This information is stored by the BSS in a database, and a script periodically checks to see if the PSS needs to be contacted, either to create or tear down the circuit
4. At the requested start time, the PSS configures the ESnet provider edge (PE) router (at the start end of the path) to create an LSP with the specified bandwidth
5. Each router along the route receives the path setup request via the Reservation Resource Protocol (RSVP) and commits bandwidth (if available) creating an end-to-end LSP. The RM is notified by RSVP if the end-to-end path cannot be established.
6. Packets from the source (e.g. experiment) are routed through the site's LAN production path to ESnet's PE router. On entering the PE router, these packets are identified and filtered using flow specification parameters (e.g. source/destination IP address/port numbers) and policed at the specified bandwidth. The packets are then injected into the LSP and switched (using MPLS) through the network to its destination (e.g. computing cluster).
7. A notification of the success or failure of LSP setup is passed back to the RM so that the user can be notified and the event logged for auditing purposes
8. At the requested end time, the PSS tears down the LSP

➢Inter-domain Reservations: Tough Problem

- Motivation**:**
  - For a virtual circuit service to be successful, it must
    - Be end-to-end, potentially crossing several administrative domains
    - Have consistent network service guarantees throughout the circuit
- Observation:
  - Setting up an intra-domain circuit is easy compared with coordinating an inter-domain circuit
- Issues:
  - Cross domain authentication *and* authorization
    - A mechanism to authenticate and authorize a bandwidth on-demand (BoD) circuit request must be agreed upon in order to automate the process
  - Multi-domain Acceptable Use Policies (AUPs)
    - Domains may have very specific AUPs dictating what the BoD circuits can be used for and where they can transit/terminate
  - Domain specific service offerings
    - Domains must have way to guarantee a certain level of service for BoD circuits
  - Security concerns
    - Are there mechanisms for a domain to protect itself (e.g. RSVP filtering)

Inter-domain Path Setup



1. On receiving the request from the user, OSCARS computes the virtual circuit path and determines the downstream AS (ISP X).
2. The request is then encapsulated in a message forwarded across the network (ISP X) towards Host A, crossing all intervening reservations systems (RM X), until it reaches the last reservation system (RM A) that has administrative control over the network (ISP A) that Host A is attached to.
3. The remote reservation system (RM A) then computes the path of the virtual circuit, and initiates the bandwidth reservation requests from Host A towards Host B (via ISP Y). This can be especially complex when the path back (from Host B to A) is asymmetric and traverses AS's (e.g. ISP Y) that were not traversed on the forward path, causing the local OSCARS to see the path originating from a different AS than it originally sent the request to.

➢OSCARS Collaborative Efforts

- **To ensure compatibility, the design and implementation is done in collaboration with the other major science R&E networks and end sites**
  - Internet2: Bandwidth Reservation for User Work (BRUW) (Ref-2)
    - Development of common code base
    - Successful inter-domain VC reservation and setup. X.509 signed soap messages over SSL used for inter-domain communication.
  - GEANT: Bandwidth on Demand (GN2-JRA3), Performance and Allocated Capacity for End-users (SA3-PACE) and Advance Multi-domain Provisioning System (AMPS) (Ref-3) Extends to NRENs
    - Instance of AMPS inter-domain manager installed in ESnet testbed.
    - Successful inter-domain reservation (no setup) between AMPS inter-domain manager at GEANT and ESnet.
    - Developing OSCARS service WSDL description to model that of the GEANT2 PACE project
  - BNL: TeraPaths - A QoS Enabled Collaborative Data Sharing Infrastructure for Peta-scale Computing Research (Ref-4)
    - Interoperability tests between OSCARS and Terapaths utilized WSDL description modeled from the GEANT2 PACE project
  - GA: Network Quality of Service for Magnetic Fusion Research (Ref-5)
  - SLAC: Internet End-to-end Performance Monitoring (IEPM) (Ref-6)
  - USN: Experimental Ultra-Scale Network Testbed for Large-Scale Science (Ref-7)

➢OSCARS: Guaranteed Bandwidth VC Service For SC Science

- **ESnet On-demand Secured Circuits and Advanced Reservation System (OSCARS)** (Ref-8)
- **In its current phase this effort is being funded as a research project by the U.S. Department of Energy, Office of Science, Mathematical, Information, and Computational Sciences (MICS) Network R&D Program**
- **A prototype service has been deployed as a proof of concept**
  - To date more then 25 accounts have been created for beta users, collaborators, and developers
  - More then 200 reservation requests have been processed

Footnotes

Ref-1 Report of the High Performance Network Planning Workshop http://www.es.net/pub/esnet-doc/2-3high-performance_networks.pdf

Ref-2 Internet2 BRUW Project: http://discvenue.internet2.edu/wordpress

Ref-3 GEANT PACE Project: http://pace.geant2.net

Ref-4 BNL TeraPaths Project: http://www.atlasgrid.bnl.gov/terapaths

Ref-5 General Atomics QoS Project: http://www.fusiongrid.org/network

Ref-6 SLAC IEPM Project: http://www-iepm.slac.stanford.edu

Ref-7 UltraScienceNet Testbed: http://www.usn.ornl.gov

Ref-8 ESnet OSCARS webpage: http://www.es.net/oscars