DYNES: A Nationwide Dynamic Network System

1 Introduction

This NSF-funded project (grant number 0958998) will develop and deploy the **Dynamic Network System (DYNES)**, a nationwide cyber-instrument spanning about 40 US universities and 14 Internet2 connectors. A collaborative team including **Internet2**, **Caltech**, **University of Michigan**, and **Vanderbilt University** will work with regional networks and campuses to support large, long-distance scientific data flows in the LHC, other leading programs in data intensive science (such as LIGO, Virtual Observatory, and other large scale sky surveys), and the broader scientific community.

For the latest announcements concerning DYNES, subscribe to the dynes@internet2.edu mailing list.

- http://lists.internet2.edu/sympa/subscribe/dynes.

To ask questions of the DYNES project team, email dynes-questions@internet2.edu.

- http://lists.internet2.edu/sympa/subscribe/dynes-questions

The purpose of this document is to provide an overview of the DYNES objectives and architecture. Additional documents are provided for Regional Networks and End-Sites who may be interested in participating in the DYNES project. These documents are as follows:

- DYNES: Regional Network and End-Site Participation Requirements
- DYNES: Criteria for Site Selection
- DYNES: Frequently Asked Questions (FAQ)
- DYNES: Application Package

This document and all of the above are on the DYNES web site:

- http://www.internet2.edu/dynes

2 DYNES Overview

2.1 Motivation: Meeting the Network Requirements of Data Intensive Science

In order to meet the science requirements, Internet2 and ESnet, along with several US regional networks, US LHCNet, and GEANT2's AutoBAHN in Europe, have developed a strategy (developed at a meeting in CERN, March 2004) based on a dual or 'hybrid' network architecture, where traditional IP network backbone is paralleled by a circuit-oriented core network reserved for large-scale science traffic. The major examples are Internet2's ION Network (formerly known as the Internet2 Dynamic Circuit Network (DCN)) and ESnet's Science Data Network (SDN), each of which provide:

1) ***Increased bandwidth capacity and reliability of network access,*** by mutually isolating the large long-lasting flows (on the DCN and/or SDN) and the traditional IP mix of many small flows
2) ***Guaranteed bandwidth as a service*** by building a system to automatically schedule and implement virtual circuits traversing the network backbone, and
3) ***Improved ability of scientists to access network measurement data*** for all the network segments end-to-end that are critical to their science through the perfSONAR monitoring infrastructure.

By integrating existing and emerging protocols and software for dynamic circuit provisioning and scheduling, in-depth end-to-end network path and end-system monitoring, and higher level services for management on a national scale, DYNES will allocate and schedule channels with bandwidth guarantees to several classes of prioritized data flows with known bandwidth requirements, and to the largest high priority data flows, enabling scientists to utilize and share network resources effectively. DYNES is dimensioned to support many data transfers which require aggregate network throughputs between sites of 1-20 Gbps, rising to the 40-100 Gbps range. This capacity will enhance researchers' ability to distribute, process, access, and collaboratively analyze 1 to 100 TB datasets at university-based Tier2 and Tier3 centers now, and petabyte-scale datasets once the LHC begins operation.

2.2 Benefits to Campuses and Connectors: An Advanced Research & Training Infrastructure

DYNES will bring the benefits of schedulable, reliable bandwidth to a much broader scientific community in addition to leading science projects by instrumenting approximately 40 campuses and 14 regional networks with the necessary switching and inter-domain circuit provisioning equipment at low cost, and making high throughput and end-to-end network monitoring services available to many scientists throughout the campus that need them, in addition to the thousands of physicists using the LHC Tier2 and Tier3 centers.

The specific benefits of DYNES include:

- **Guaranteed bandwidth service:** User-specified bandwidth may be requested and managed in a Web Services framework. This service can be accessed by individual users, system or site managers, and in the latter DYNES project stages also autonomous agents, based on role-based (PKI certificate) authentication and authorization.
- **Traffic isolation and traffic engineering** to provide for high-performance TCP stacks and other transport mechanisms; including the option for non-standard mechanisms that cannot co-exist with commodity TCP-based transport. This will also enable the engineering of explicit paths to meet specific requirements, for example to bypass congested links, using lower bandwidth, lower latency paths.
- **Secure connections:** The circuits are "secure" to the edges of the network (the site boundary) because they are managed by the control plane of the network that is isolated from the general traffic.
- **End-to-end (cross-domain) connections of several types:** between Labs or other experiment sites and institutions, as well as to sites hosting computational resources (Tier2 or Tier3, or supercomputer sites to support some scientific fields), and among collaborating sites for shared simulations, visualization or collaborative work on data analysis.
- **Reduced cost of handling high bandwidth data flows,** since the most highly capable routers are not necessary when every packet goes to the same place

DYNES will crystallize the benefits of the hybrid network architecture by integrating the circuit networks and provisioning services into an end-to-end system architecture encompassing storage systems, network interfaces and generally available high-throughput applications.

2.3 Benefits to the National and International Research and Education Community

Impact on Researchers

DYNES services researchers' needs by allowing dynamically scheduled bandwidth across multiple domains. It provides this not only for the high-energy physics groups, but for **any** campus researcher in need of large amounts of scheduled bandwidth.

Impact on Students

Students in the sciences will benefit from greater access to large scientific data sets that can be used for analysis or enhance instruction. This may include LHC data incorporated into a physics lecture series or a software/analysis tutorial that lives on the disk array at a Tier 3 or Tier 2 campus. Computer science and networking students will also benefit from analyzing empirical performance of the network, scheduling algorithms, path-finding methods and security models of the IDC on a large scale.

Impact on Network Professionals

Network professionals at regional networking sites and campuses will gain valuable experience running dynamic networks. Networks that have frequently evolving configurations present new issues related to planning and debugging networks. Procedures such as looking at data provided by perfSONAR to gain a view of the network may be new to many network operators who have worked in primarily static environments. Having hands-on access to DYNES and associated documentation will enable these individuals to learn new skills. This exposure will develop an expertise among U.S network professionals as it pertains to dynamic networking. Network architects and designers will benefit from real-life experience with networks and large long-range flows among storage pools viewed as a ***dynamic system.***

Impact on Universities

Individual researchers, student, and network professionals will not only be impacted but also the institutions to which they belong. By using DYNES to access large data sets and learn how to operate and manage global-scale networked systems, U.S. educational institutions will remain on the leading edge of research necessary to attract the best students and faculty from around the world. DYNES is especially important to smaller sites that may not have access to crucial network and/or storage resources otherwise.

3 DYNES Architecture

DYNES' packet and circuit architecture is composed of Internet2's ION service and extensions over regional and state networks to US campuses. It will connect with transoceanic (IRNC PRONET, USLHCNet), European (GÉANT), Asian (SINET3) and Latin American (RNP and ANSP) R&E networks. It will build on existing key open source software components that have already been individually field-tested and hardened in part by the PIs: the DCN Software Suite (OSCARS/DRAGON), perfSONAR, the UltraLight Linux kernel (Michigan), and Caltech's Fast Data Transfer open source application suite: FDT, FDT/dCache and FDT/Hadoop.

Figure 1 shows two typical transfers that DYNES supports: one between a Tier 2 and a Tier 3 and another between a Tier 2 and a Tier 1 site. The clouds represent the network domains involved in such a transfer. A DYNES instrument must provide two fundamental capabilities at the Tier 2, Tier 3, and regional networks:

1. Network resource allocation such as bandwidth to ensure performance of the transfer
2. Monitoring of the network and data transfer performance

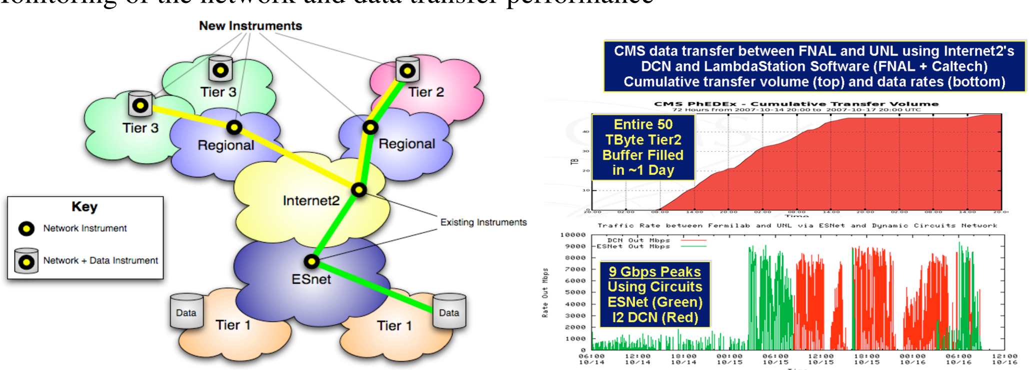

Figure 1 DYNES example of data transfers from a Tier 3 to a Tier 2 and a Tier 2 to a Tier 1 across a dynamically allocated circuit. (Right) A living example: restoring a 50 TByte dataset from FNAL to UNL over the Internet2 DCN (Red) and ESnet (Green).

All networks in the path require the ability to allocate network resources and monitor the transfer. This capability currently exists on backbone networks such as Internet2 and ESnet, but is not widespread at the campus and regional level. Also, the Tier 2 and Tier 3 sites have an additional requirement:

3. Hardware at the end sites capable of making optimal use of the available network resources

This additional requirement for Tier 2 and Tier 3 sites as compared to regional networks leads to the need for multiple configuration options depending on site type. **Figure 1** categorizes these options as "Network + Data" Instruments on the Tier 2/Tier 3 networks and "Network" instruments on the regional network.

Despite the success of dynamic circuits in production environments, including Internet2's ION, ESnet's Science Data network, and US LHCNet, participating sites only have had a limited number of locations connected to the dynamic infrastructure. Almost all of the connections have statically provisioned pieces at either campus or regional level. DYNES enables more sites to participate by providing the necessary equipment and extending dynamic capabilities into the regional and campus networks. It does so with the same field-tested software components used in current deployments.

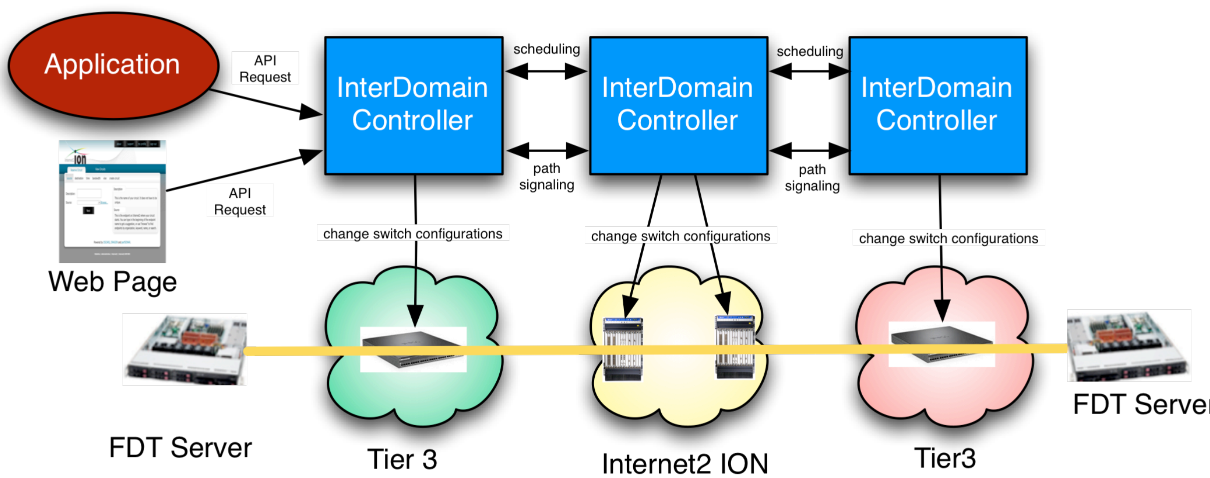

Figure 2 InterDomain Controller Software Architecture

As shown in **Error! Reference source not found.**, there is a software component know as InterDomain Controller (IDC) which enables the dynamic capabilities across multiple domains. The IDC software handles user requests, inter-domain interactions, and the scheduling of resources. The IDC utilizes the InterDomain Controller Protocol (IDCP) which defines the format of messages exchanged between IDCs and client applications. The user may request a circuit using (1) a web page interface or (2) through an external application that speaks the IDCP (such as LambdaStation or TeraPaths). A path is then calculated and the request forwarded down the chain until each network has had a chance to verify the resources. Communication between domains is secured using X.509 certificates to identify the networks. The DCN Software Suite contains the IDC software, which is based on the OSCARS and DRAGON software. The DCN Software Suite is designed to accommodate multiple network element types and topologies and the exact configuration of the IDC software will be dependent on the underlying network design. In all configurations the OSCARS software will be required. In many configurations DRAGON software will also be needed. The exact configuration of the IDC software is best determined as part of a review of individual site architecture drawings and plans.

As shown in a DYNES (sub-)instrument at a Tier2 or Tier3 site consists of the following hardware, each item of which has been carefully chosen to combine low cost and high performance:

1. An Inter-domain Controller (IDC)
2. An Ethernet switch
3. A Fast Data Transfer (FDT) server. Sites with 10GE throughput capability will have a dual-port Myricom 10GE network interface in the server.
4. An attached disk array with an LSI or equivalent Serial Attached SCSI (SAS) controller capable of several hundred Mbytes/sec to local storage.

The Fast Data Transfer (FDT) server connects to the disk array via the SAS controller and runs the FDT software developed by Caltech. FDT is based on an asynchronous, flexible multithreaded system that automatically adjusts I/O and network buffers so that the network achieves maximum utilization. The disk array stores the datasets to be transferred among the sites in some cases. The FDT server serves as an aggregator/throughput optimizer in this case, feeding smooth flows over the WAN and LAN directly to the Tier2 or Tier3 clusters.

The configuration also includes an Inter-domain Controller (IDC) server to handle the allocation of network resources on the switch, interactions with other DYNES instruments related to network provisioning, and network performance monitoring. The IDC creates virtual LANs (VLANs) dynamically between the FDT server, local campus, and wide area network. Sites may choose to add additional connections to the switch (not displayed in the diagram) via one of the switch's other ports.

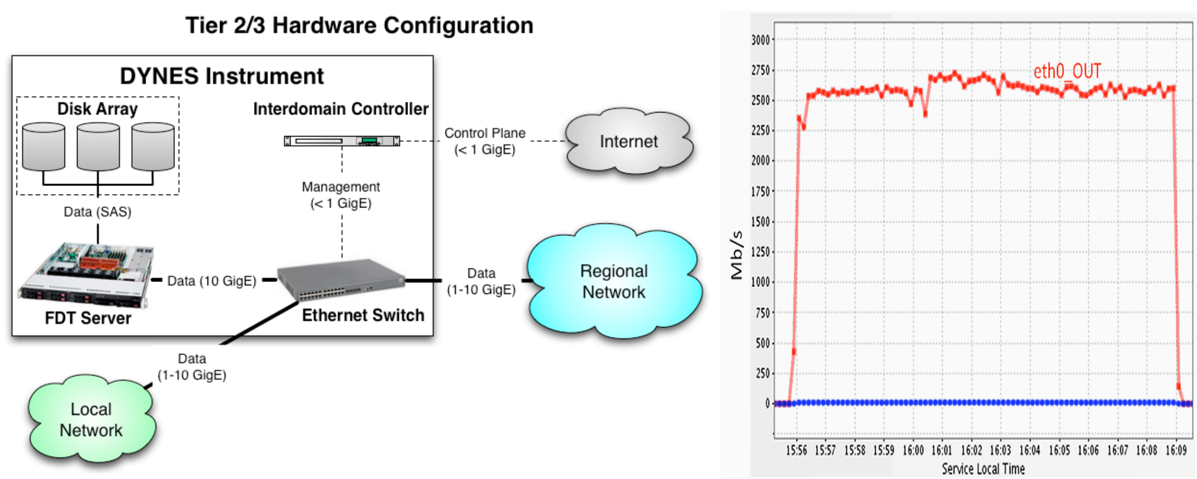

Figure 3 (Left) A DYNES instrument at a Tier 2 and Tier 3 campus. (Right) A living 2007 example, of an FDT transfer between disk servers at CERN and Caltech. A smooth 2.6 Gbps (or 2.0 TByes/hr) is achieved with one RAID controller, and 5 Gbps with two RAID controllers.

As shown in **Figure 4,** Regional networks require (1) an Ethernet switch and (2) an Inter-domain Controller (IDC). The configuration of the IDC consists of OSCARS, (optionally) DRAGON, and perfSONAR just as in the Tier 2/Tier 3 cases. A regional network does not require a disk array or FDT server because they are providing transport for the Tier 2 and Tier 3 data transfers, not initiating them. This allows the regional network to provision resources on-demand through interaction with the other instruments.

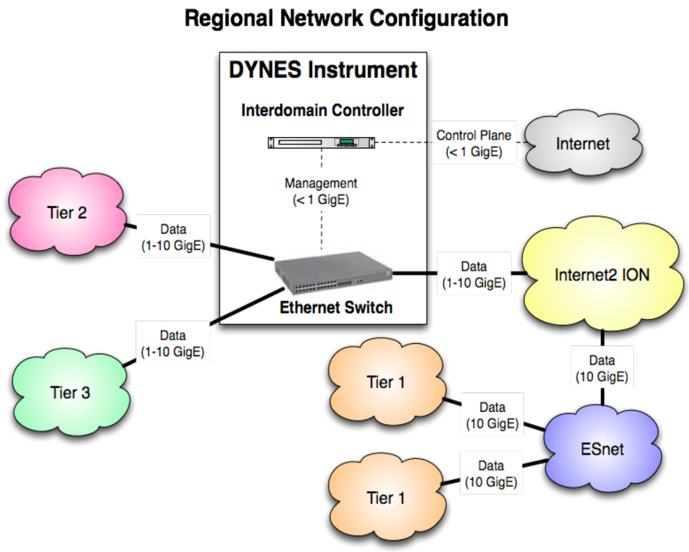

Figure 4 A DYNES Regional Network Instrument

4 Summary

The DYNES team will partner with the LHC and astrophysics communities, OSG, and Worldwide LHC Computing Grid (WLCG) to deliver these capabilities to the LHC experiment as well as others such as the LIGO, VO and eVLBI programs, broadening existing Grid computing systems by promoting the network to a reliable, high performance, actively managed component.

Future science programs in HEP, astrophysics and gravity wave physics, and other data intensive disciplines, will be facilitated by DYNES' technologies and worldwide network partnerships. Working with CHEPREO and similar education and outreach efforts targeting under-served communities both in the US and overseas, DYNES will reach a wide variety of students at collaborating institutes including underrepresented groups and minorities. This will lower the barriers, and enable individual graduate students, undergrads, postdocs and faculty to use DYNES to achieve high throughput in support of their research in many data intensive fields.