Transferring Scientific Data: Opportunities and Best Practices

Presented by ESnet / Berkeley Lab

Introduction

LHC scientists move hundreds of petabytes of data every year to and from collaborators, institutions, and storage and analysis facilities around the world. To successfully enable data transfers of this magnitude and scale, international computer networks need to be specifically engineered from end to end—from Tier 0 at CERN to any Tier 3 institution. Figure 2 shows the general path of all LHC data transfers between the LHC tier sites. Much of the LHC's data is carried from CERN to the US Tier 1 sites and onward via ESnet.

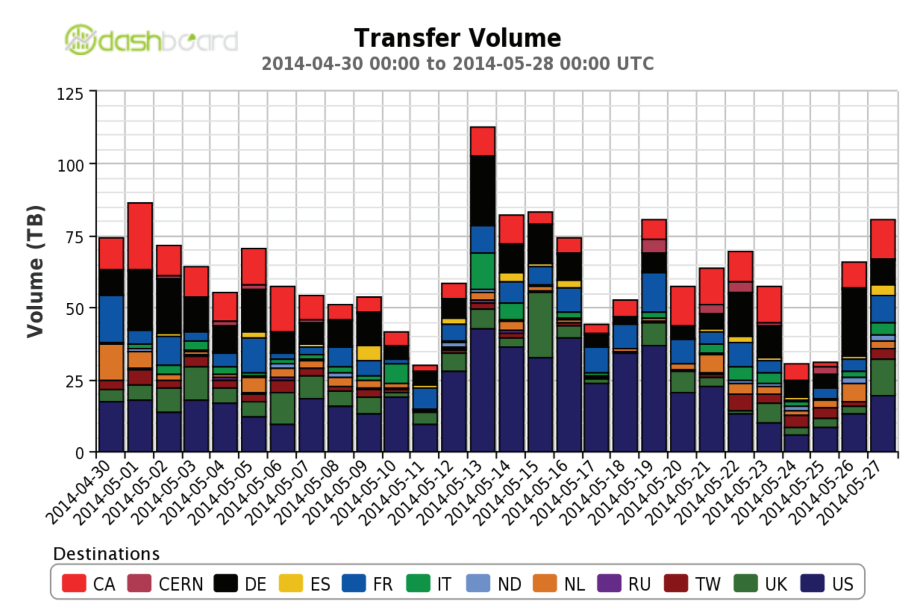

Figure 1: Data transfers from just the ATLAS experiment over a 28-day period. The volume of data transfers are noted in terabytes vs. days. Note the US data transfers are marked in dark blue.

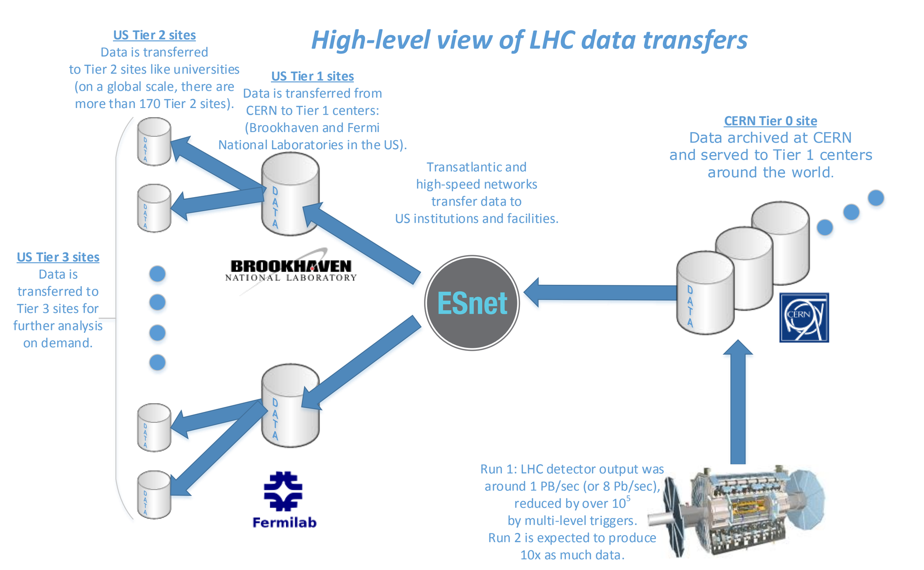

Figure 2: General overview of how data will flow from an LHC detector to a Tier 3 institution during Run 2. Data is transferred from a detector to some on site storage. At CERN, the data rate is reduced. The reduced data set is then distributed internationally. Run 2 is expected to increase the LHC's data production by 10-fold.

INFRASTRUCTURE

Another key factor to enabling data transfers for traffic to/from LHC tier sites is having the appropriate network infrastructure, hardware, and configurations at the destination sites, like the Tier 3 institutions. There are three components that contribute to setting up a successful workflow for these destination sites.

1. Have the right dedicated systems or hardware; which include
   - a Science DMZ (see Figure 3, also http://fasterdata.es.net/science-dmz/);
   - dedicated hosts that are configured for data transfers (called data transfer nodes); and
   - appropriate amount of local storage.
2. Configure the appropriate tools or software on the dedicated system or host (i.e., with PanDA or PhEDEx packages).
3. Have the local IT or support team tune the systems to ensure the network has no packet loss—this could be network tuning or host tuning (see Figure 4 for an example of a network with packet loss).

All three of these factors will lead to a successful science workflow at any institution.

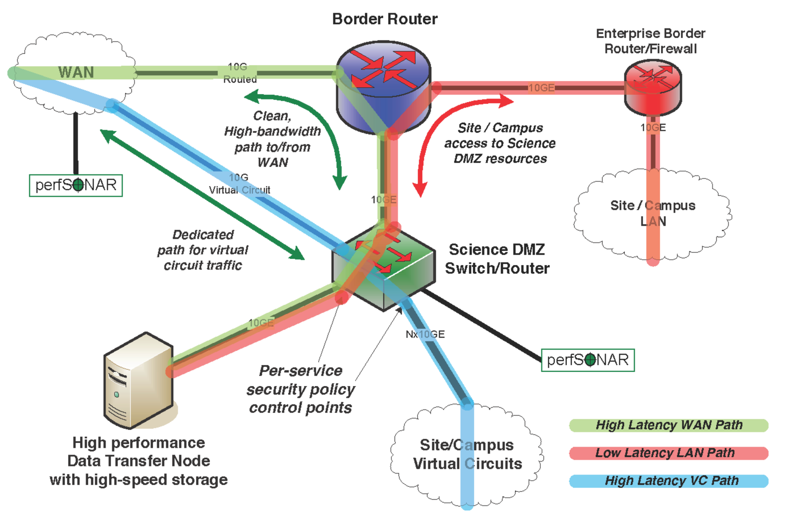

Figure 3: General Science DMZ diagram. Data will travel from its original source, onto the wide area network (WAN) to the data transfer node, where there is some temporary storage available. From there, the scientist can pull the data onto their campus local area network (LAN) for analysis.

DATA TRANSFER PROTOCOLS AND TOOLS

It is important to use the right tools and protocols for data transfers. Different protocols behave differently on the networks and some protocols perform better than others. Many LHC researchers use PanDA or PhEDEx for data transfers. However, other good protocols can be:

- FDT (Fast Data Transfer) tool from Caltech: http://monalisa.cern.ch/FDT/
- bbcp: http://www.slac.stanford.edu/~abh/bbcp/
- lftp: http://lftp.yar.ru/
- axel: http://axel.alioth.debian.org/
- GridFTP from ANL has features needed to fill the network pipe
- Globus: https://www.globus.org

For benchmarking, the following table shows approximately how long it should take to transfer 1 TB of data across varying network speeds:

10 Mb/sec network	300 hrs (12.5 days)
100 Mb/sec network	30 hrs
1 Gb/sec network	3 hrs
10 Gb/sec network	20 minutes

These describe the throughput for data transfers based on end-to-end network bandwidth noted in megabits and gigabits per second (Mb/sec and Gb/sec).
(See http://fasterdata.es.net/fasterdata-home/requirements-and-expectations/ for more information.)

WHEN DATA TRANSFERS DON'T WORK WELL...

There are many causes of poor network performance, but one of the most common causes is packet loss. Packet loss causes the most common network protocol, TCP, to adjust its rate of data transmission creating poor network performance. The graph in Figure 4 shows the large performance impact on data transfers with even a very small amount of packet loss.

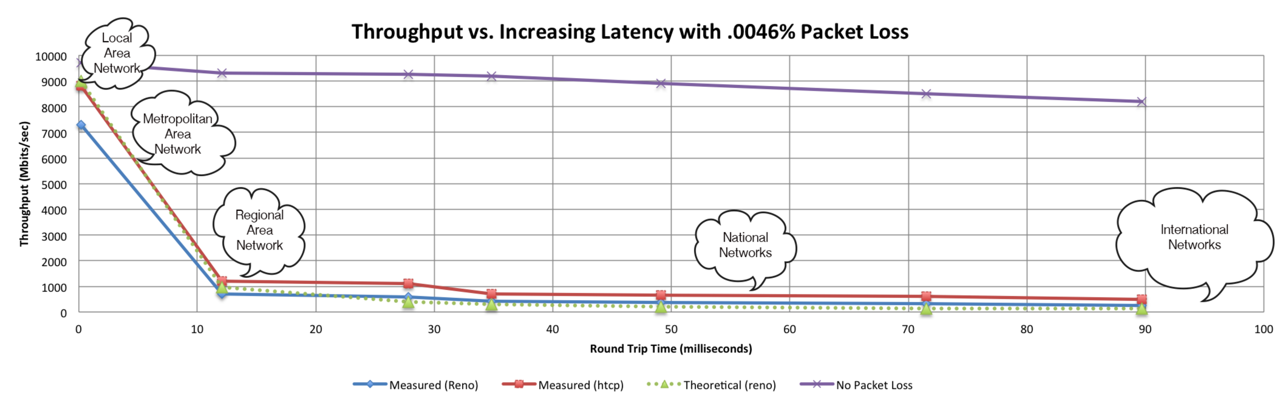

Figure 4: Graph that displays the behavior of network throughput when it is dropping 1 in 22,000 packets. With .0046% packet loss, data that traverses network paths more than 10 ms away will see extremely low data transfer performance.

The LHC and ESnet

Network traffic on ESnet has been growing exponentially since the early 1990s, and LHC traffic has played a significant role in driving that growth for the past decade (see Figure 5). ESnet has a vision of making it possible for scientists to access their data as needed without regard to geography—across the network, productively, and on human-reasonable time scales.

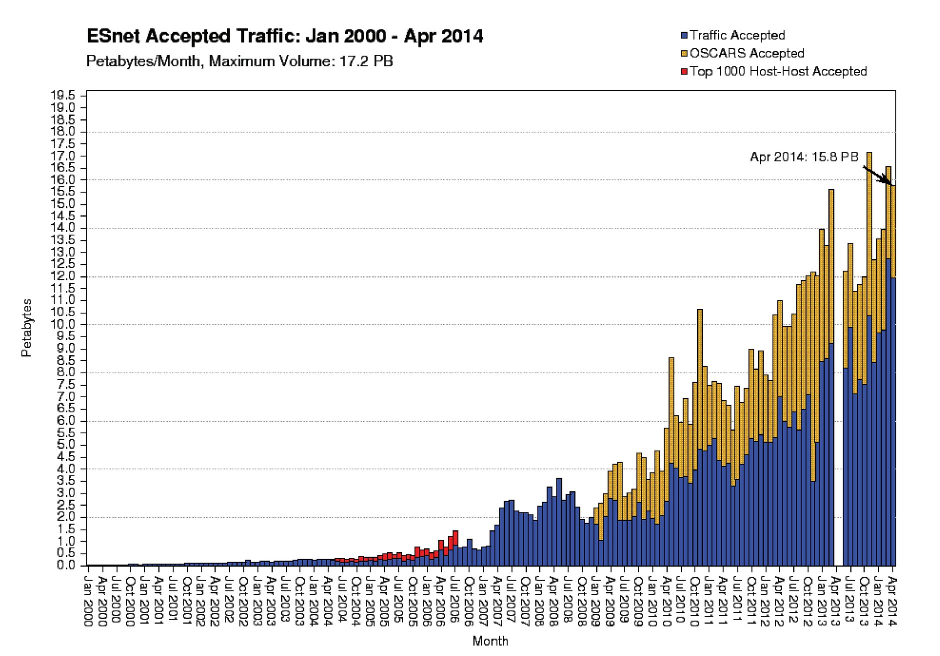

Figure 5: ESnet's monthly traffic volume reached 15. 8 PB in April 2014. Much of this data is due to LHC-related traffic as noted from the start of the LHC testing from November 2004 to April 2010 and the beginning of LHC operations after April 2010. The yellow bar graphs show traffic carried by OSCARS, an on-demand circuit reservations system, which is dominated by LHC data.

About ESnet

ESnet is the Department of Energy's high-performance, high-bandwidth network that provides reliable network connections for scientists at national laboratories, universities and other research institutions, enabling them to collaborate on some of the world's most important scientific challenges including energy, climate science, and the origins of the universe. ESnet is funded by DOE Office of Science and located within the Scientific Networking Division at Lawrence Berkeley National Laboratory.

Contacts and More Information

Call your local support or IT team if you are seeing any problems!

For other questions or more help, email: engage@es.net

More information on transfer protocols and best practices visit Fasterdata Knowledge Base: fasterdata.es.net

ESnet
ENERGY SCIENCES NETWORK

BERKELEY LAB

CSO 014-ESN-0579