Networking for LHC

David Foster CERN IT-CS

(Slide – Opening)

Being able to communicate, send data, between two computers anywhere on the planet is a fundamental concept of the grid. It is assumed that this will work, seamlessly, reliably and perhaps most of all quickly. Indeed, this can be said now of the expectations of the increasing number of people who use the internet daily and expect to be able to fetch information to their computer from where it is stored, anywhere on the planet.

But the requirements for the LHC are larger, much larger, than todays home user and as such require more network capacity than todays internet can provide. For this reason, we have had to push the envelope of what can be provided by the worlds dedicated research and education networks and stretch to the limit the technology available, in order to deal with the enormous data volumes of the LHC. Indeed, we will continue to adopt emerging technologies and push the limits of international networking for many years to come. There is no end in sight of what we will demand of local and wide-area networks.

(Slide - Partners)

So, what is the network infrastructure the LCG uses? Well, firstly, we are talking about networks which are provided by many different parties. Here I show some, I hope most, if not all, of the organisations who directly contribute to the LHC data movement from CERN to the Tier-1s. The national research networks (NRENs) of many countries as well as dedicated projects such as the EU funded GEANT project in Europe and the US Department of Energy funded project, USLHCNet, are all part of a large and dynamic network collaboration. These collaborative efforts provide the connectivity we need to distribute the LHC data reaching to to all Tier-1 centers, wherever they are, including Canada via CANARIE and the Academia Sinica inTaipei.

(Slide - OPN)

The key delivery mechanism for this primary data movement is the so-called LHC Optical Private Network. This connects CERN to each of the 11 Tier-1 centers with a dedicated circuit of 10Gbps. This slide also indicates the interconnections between Tier-1 centers that helps provide redundancy and resilience to the overall network infrastructure.

This capacity of 10Gbps to each center is enough to send around 400 ITunes songs, per second, to each of the Tier-1s in parallel. It total, this is also equivalent to transferring around 10, 3h, standard definition movies per second. The end to end path between CERN and the Tier-1 centers is constructed by “stitching” together 10Gbps circuits provided by a number of different parties which leads to technical and operational challenges but optimises the available resources.

(Slide – GEANT)

In Europe the path may be constructed from the infrastructure provided by the GEANT project, shown in this slide, and a particular NREN. For connectivity to the US the end to end path from CERN may consist of a commercial transatlantic circuit and a circuit provided by the DOE funded Energy Sciences network, ESNet. To reach Canada and Taipei the paths are constructed from a variety of sources. These are just a few examples and many combinations are actually used in practice. This federated infrastructure has driven the need for innovations in areas such as automated path construction for end-end circuits and in operational coordination. Indeed the concept of dynamic, mission critical, end-end circuit construction is a new type of service that has been adopted by the research and education networking community in response to the requirements of applications such as the LCG requires. This is an excellent example of how global networks are constructed in practice, using the best facilities available from whoever can provide them in a global federated collaboration.

(Slide – Architecture)

But certainly this is not all. The Tier-2 and Tier-3 centers that participate in the data analysis must have access to the data stored in the Tier-1 centers. Whilst it is crucial that data is moved from CERN to the tier-1 centers quickly and efficiently, it is quite a predictable activity in scope. The requirements of data analysis tend to follow more opportunistic and therefore chaotic patterns. For this, in addition to dedicated circuits, extensive use is made of the general purpose, but high capacity, IP networks provided by the research and education network community. You can consider this as a “research and education internet” which interconnects all the research and education activities world-wide through many organisations, such as the National Science Foundation (NSF) funded, Internet2 network.

The growth in these research and education networks has been astounding. The Energy Sciences Network of the United States reports it has seen a 10x increase in its IP traffic every 47 months since 1990 with increasingly traffic being dominated by the large science experiments.

With disk usage at grid centers expected to double within a few years of LHC startup, the demands on the network, as data are re-distributed across the grid, can certainly be anticipated to increase substantially. The emergence of 40Gbps single wavelength capability and ultimately 100Gbps should enable the research and education networks and the LHCOPN to respond to the inevitable demand for increased capability to move data around.

(Slide – GLIF)

These networks are not static. You should have a feel now for the fact that networking is an organic, changing activity which leverages new opportunities as they present themselves and consequently we must always be flexible to create, and sieze, new opportunities for connecting scientific collaborators together in order to provide access to scientific data and enable participation. What you can see, perhaps not very clearly, here is a collection of circuits that has been put in place by various organisations for use by the scientific community.

I would like to highlight one example. Together with commercial partners and the open lightpath exchange of the Dutch national research network, Surfnet, called Netherlight, we were able to provide an unprecedented level of connectivity between CERN and the Tata Institute of Fundamental Research (TIFR) in Mumbai at 1Gbps. This arrangement has been essential to ensure that the TIFR can collaborate as a first-class partner in the LHC Computing Grid (LCG). Other examples exist and together with USLHCNet and partners we have been experimenting with connectivity between CERN and institutes in South America at 1Gbps also. Africa hosts a tier-2 center in capetown but network capabilities have been limited so far. Nevertheless, bridging the digital divide in the interests of scientific collaboration has been gaining increasing recognition in recent years through active participation of both the scientific community and the governments.

(Slide - Harvey)

The trend we can anticipate is that with lowering cost and increasing size and increasing performance of discs this will put more and more pressure on the networks to be able to deliver scientific data “into the hands” of the scientist. The wide area networks discussed above, and in particular USLHCNet shown here, will scale with the emergence of 40Gbps and 100Gbps networks in the coming years but the real pressure will be on campus networks to deliver the data at reasonable rates to the end user. With a typical 1-2TB working dataset per person this will require of the order of 4 hrs to transfer with a dedicated capacity of 1Gbps. Most campus networks are far from being able to deliver this level of service today but expectations will increase as more powerful laptops become available. Perhaps a counter-force may be the rise of "cloud computing” solutions where the data could be co-located with the processing elements in large data centers? In these scenarios the physicist would work in an interactive way with a remote “virtual” machine from her laptop. Whilst this might alleviate the pressure to move data around in an asynchronous way, it will increase the pressure for highly responsive, interactive networks that can guarantee a high quality of service in real-time.

(Slide – Growth)

The scientific community may be able to take advantage of new solutions appearing as a side effect of the direction that business is taking. Although in the commercial space the highest margins are made from providing relatively low bandwidth applications, to very large markets, the emergence and delivery of new types of services to internet users from very large data centers need to be considered. The technical paradigm shift that is making this possible and driving commercial interest is the growth of the number of users that have direct access to cost-effective high performance networks. Although well below the capacities required to deliver TB's of data to a laptop, the scientific community may nevertheless have an opportunity to exploit these new infrastructures designed to deliver interactive content. It remains to be seen how these developments will play out in the long term and how these market forces will drive innovation but ubiquitous networking is providing tremendous opportunity to explore new ways of performing science.

However, what is important here is the ability to provide access to scientific data at CERN, delivering it to scientists in many parts of the world and engaging them in the journey of scientific discovery. The network partnerships and infrastructures for research and education are fundamental to being able to make this happen. CERN and the LCG are playing a key role in helping to catalyse the investments necessary to connect our collaborators world-wide and grid technologies are providing the platform to truly "think globally but act locally".