Global R&E Network Architecture Program

Authors:
Prof. Jianping Wu (CERNET) & H. David Lambert (Internet2)

With contributions from:
Prof. Xing Li (CERNET), Jim Williams (Internet2/Indiana U), Rob Vietzke (Internet2), Erik-Jan Bos (Internet2), Greg Bell (ESnet)

Version: 1.0 – Date: April 30, 2013

1. Executive Summary

The globalization of education, science, research and information technology services "above the network" is creating at a global scale the same imperatives to build advanced global networking capabilities that drove many of the world's leading nations to create domestic R&E networks. With critical science projects emerging in underserved portions of the world and increased collaboration between all regions of the world imperative, a framework to ensure an advanced global set of resources can emerge to create a cohesive global fabric of network capabilities to support advanced science, research and education is essential. This paper offers a framework through which the global leaders of National Research and Education Network (NREN) and funding agencies can begin the process of implementing a next generation community-owned global research and education network.

The purpose of this project is to 'create a model for engineering and funding a scalable and persistent global infrastructure that meets the advanced networking needs of an increasingly globalized research, education and science community, while also providing opportunities to advance the state of the Internet itself and to provide a global platform for the instantiation of advanced networking technologies. As an advanced network project at its core, certain principles must be established and managed to assure that investments are routinely, flexibly and deterministically positioned to always provide advanced

services and to never become intractably burdened by incumbent business needs.

2. How is global advanced networking provided today?

The Internet and the global Internet in particular is a relatively recent development. Because of the nature of the technology, the development path and the various funding mechanisms, the current global R&E network architecture is only loosely coordinated. There are continental network architectures that have been built to national boundaries. Intercontinental service is funded in a variety of ways by individual NRENs, governmental funding bodies (for example the European Commission, US National Science Foundation) and sometimes science projects themselves, the LHC activities being one example.

Government funded science agency programs contribute general purpose capacity within the direction of their programs, however the structure of their funding limits their ability to make long term investments. Additionally, governments may place somewhat (or very) restrictive use policies on their network investments.

Among the several dozen individual capacity contributors, there is coordination – often bilateral, sometimes trilateral attempting to stitch together a common set of capabilities. This works fairly well between continents, but becomes less predictable and reliable, as users of these networks increasingly need to traverse one or more of these bilateral arrangements to reach a collaborator beyond a single coordinated circuit pair. Due to procurement or national restrictions, coordination is largely in the form of informal agreements of "reciprocity" (the bilateral parties agree to procure and deploy roughly equal amounts of bandwidth to one another) and some procurement coordination (mainly to promote diversity in path and carrier, with only a few examples to achieve pricing advantages). The NSF-EC supported ACE Project is an example of a successful bilateral coordination that has generated significant value.

A feature of the current global R&E network architecture is the existence of "Open Exchange Points" (OEPs). Open Exchange Points promote policy-free, i.e. the OEP does not add policy, interconnection between R&E networks that have presence in the same region. There is some engineering-level coordination among the various operators of the OEPs, but largely they are located where an organization has decided to invest. Typically, this has been in a location where several networks are already present or have expressed an interest in having an OEP. However, as these OEPs have arisen organically, there is no real coordinated global discussion currently about where OEPs should and could most effectively be located or what services and service levels should be offered.

OEPs will play a critical role in the development of the GNA. In many cases, the GNA will require that OEPs change their role and responsibilities. This topic is complex and requires NREN CEO discussion and merits a separate discussion paper.

3. The challenges of the existing infrastructure

OEPs
There are also a wide variety of implementations of the exchange points. Whereas some OEPs run under excellent operations regimes with substantial network capacity and advanced protocols, other OEPs sometimes are substantially less robust. As many of the OEP capabilities were set up to serve local needs, the diversity of OEP capabilities poses a challenge when viewed globally as one weak link in a chain of interconnected resources diminishes the value of all the others. See the topics in the NREN CEO Summary and NREN CEO discussion documents.

Timeliness Challenge: At present, major intercontinental links purchases are driven by 1-3 year short term procurements of funding agencies. In contrast, many leading continental networks include long term investments in flexible infrastructure and can evolve quickly within planned funding. The GNA must avoid this timing mismatch and link more closely NREN procurement cycles with international bandwidth.

Acceptable Use Challenge: The current approach also features loosely coordinated limits on acceptable uses of each link. Some links are restricted by funding and policy to a particular domain science. Others are restricted to only certain funded projects, while others share R&E traffic with fully commercial activities. In at least one nation, it appears that traffic that might transit through that nation may be filtered based on domestic information management policies.

In contrast, the GNA will be a single network with a single AUP administered at the OEPs. The exact form of the AUP is a topic for discussion among the NREN CEOs (again this refers to the supplementary documents provided). The mechanism by which the AUP is administered at the OEPs is a technical topic which needs separate study.

Multi-hop Challenge: Even as major science infrastructure emerges in Chile, Australia and South Africa, areas of the world like Africa, South America and the Middle East often cannot be reached by major collaborators in Europe, North America and Asia through uniformly scaled capacity or capabilities. The resulting experience is the lowest common denominator regardless of what capacity might be available or contributed along the path. In contrast, most national NRENs have arranged funding to provide uniform capacity on an end-to-end basis.

Uneven Exchange Point Challenge: The exchange points that do exist and that are run well form a critical piece of the current infrastructure; however they are unevenly placed and sometimes unevenly constrained by policies that may limit even provisioning of global capabilities. There are also unevenly operated and unevenly equipped exchanges at critical intersections of contributed capacity.

The development and structure of the next generation of OEPs is critical to the GNA. This issue deserves separate detailed technical and policy analysis in a distinct document.

4. A Program for designing an Architecture for Global R & E Collaboration

A new global optical exchange network, on par with the most advanced NREN infrastructures, must be built with a common set of core principles that allow it to have a consistent set of services on a global scale and that allows it to remain relevant as a long term leadership investment of the participating parties. As a differentiated network from single-party or commercial offerings, its operating principles and priorities must include a flexible, future driven set of principles that assure the investment remains responsive to the most current and forward thinking needs.

We seek to establish a framework that:

- Establishes a global architecture for a new network capability with the highest standards to which all NRENs can contribute and participate and which reaches all regions of the world.

- Establishes a governance structure for the new network capability that enables further growth and feature richness in those areas that desire this, with opportunities for all contributors. The notion of a lowest common denominator set of features is explicitly avoided.

- Establishes a platform for finding sustainable funding sources that are capable of contributing to this new network capability, exceeding the usual short term funding cycles. Innovative procurement procedures will be developed that can make the best use of these sustainable funds when available.

- Assures long term capability to scale capacity globally in a way that assures all NRENs can predictably reach all other NRENs at an optimal level of performance.

- Assures an agreed upon, consistent set of standards for acceptable use based resolutely on open and inclusive definitions that assure broad use of end to end capabilities.

Design principles

The new network capability will adhere to a number of well established design principles, which could include:

- Simplicity, as this design principle has proven to lead to scalable, flexible and robust infrastructures.
- Sustainability, as this infrastructure and the services layered upon it are designed to last and evolve over at least 15 to 20 years. Evolve is the key concept. This directs the focus to policy rather than technology.
- High performance, the new global R&E network architecture seeks to deliver speeds and functionality on par with the world's most advanced NRENs.
- Inclusiveness, all NREN-like organizations globally that are willing and able, will be able to participate based on fair and open principles.

Global Scalable Optical Capacity

The global network will include a combination of owned facilities and commonly operated equipment staged throughout the world to maximize connectivity in both current and future areas of interest to the global R&E community. The network will be architected to provide both next generation shared services like a global distributed open exchange at layer-2 and the ability for participants to acquire and allocate dedicated incremental capacity (at layer 1 or layer 2) among any of the points of presence on the network.

On domestic scale, leading NRENs have settled on owned fiber IRUs as an essential underpinning of their ability to continuously advance their network offerings. IRU-based networks are the only approach that have provided both the scalability necessary to meet capacity needs, while also providing the cost and flexibility profile necessary to afford the growth. IRU-based networks are also generally the only networks that can have their underlying transport technology rapidly refreshed as technologies evolve.

On the new global scale, IRUs on undersea cables are available that provide most of the benefits of terrestrial cables, including economic and capacity benefits over circuits. However, due to the physical nature of such cables, careful consideration for increased mean-time-to-repair and upgradability must be considered. To design around such constraints, collaborators in the new global infrastructure will need to consider investment in multiple rings between each continent and creation of the ability of each major world-region to connect to at least two of these rings. Such an investment must be carefully coordinated so that as collaborators are identified to fund each section of the global network, it is carefully aligned with other investments.

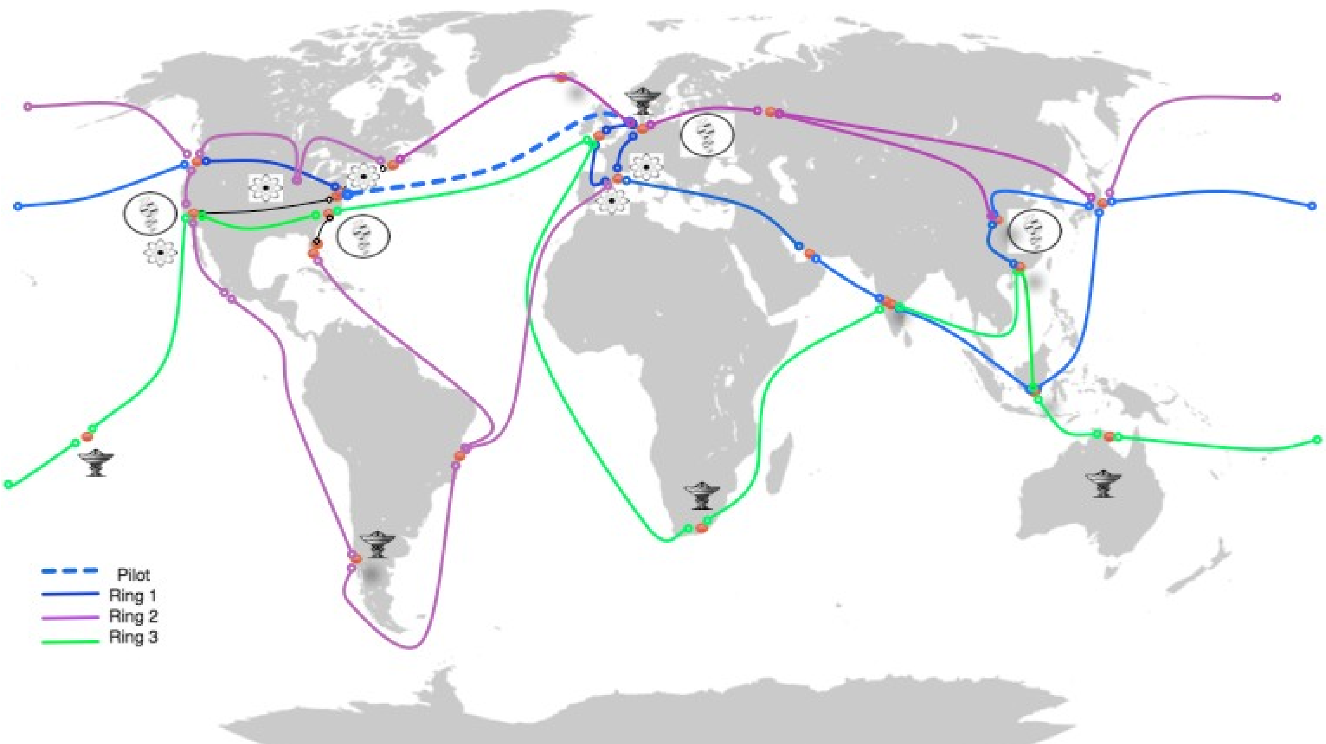

Figure - Example of a global optical rings with open lightpath exchanges

Resolutely Open Access

To support the network's purpose to be *the* advanced network substrate for global collaboration, it is essential that no single point in the network become a bottleneck, either capacity or policy, to another global user wishing to collaborate with any given region or any given region that sits behind an individual node on the network. The only way to establish such assured capability is to set firm policy to always require the most inclusive possible standard of acceptable use across the entire network. A resolution to maintain only the most open access at all points on the network is the only way to assure no differential encroachment on the global capability set.

All components of the infrastructure must allow a combination of pure research, academic and educational purposes to be fulfilled. This may include commercial partnerships and non-traditional uses that may be outside the mission of an individual contributor. The infrastructure must be implemented in a way where usage policy is open and conditions of use are broadly construed.

Accessing Open Capacity: Open Lightpath Exchanges

The new network architecture should include open light path exchanges that are coordinated at the intersection of major global capacity paths to allow each major region of the world to attach to at least two exchange points. They should be operated with a common set of inclusive principles that assure the highest standard of free and open exchange of information among all participating parties in the global collaboration. No ruling to restrict traffic or utilization at any Open Exchange point should be allowed that unequally restricts any party in the global collaboration from working with any other party.

The concept and policy surrounding Open Lightpath Exchanges or Open Exchange Points will be both critical to the development of the GNA and difficult to create. Some issues surrounding the OPE are outlined in the NREN CEO Discussion Points document. The issues and policy surrounding OPEs deserve separate discussion and attention from the issues and policies surrounding the GNA network. This separate document will be created with the guidance supplied by the NREN CEOs as they focus on the OEP discussion points.

Inclusivity: Geographic reach and OPEs

The new global network would need to create at least two points of presence (for resiliency) that were reachable from each major region of the world. Such points of presence (OEPs) would need to be located at locations where current and future infrastructure was likely to be interconnected so as to encourage access from the greatest number of countries and localities.

The architecture should focus on consistent delivery of services across regions as well as what capabilities may exist in a given region. This means designing globally to the most exemplary requirements that more than a single party needs as opposed to seeking the lowest common denominator that all parties need or are capable of providing.

The new architecture must plan for geographies that need to be reached by the scientific world with R&E networks but are located within NRENs that may lack the resources or incentive to build to the rest of the world – Atacama desert, SKA, Antarctica. Any new architecture must solve this science access problem.

Services offered on the new infrastructure

While dozens of new capabilities may emerge on top of the new infrastructure over time, two basic services are essential to enable participation of the global NREN community in the project.

Note that these services depend on a number of unresolved issues such as funding, central or distributed control and further refinement of the definition of OEPs. These should be viewed as examples of important services that the NREN CEOs may wish to implement on the GNA.

Global Open Lightpath Exchange Service

The network must support a basic level of connectivity between all of its exchange points (OEPs), with the ability to offer unrestricted Vlan or VLAN-trunks between any endpoints on the network.

All participants (NRENs or NREN collaborations) will be expected to attach to an OEP and offer unrestricted connectivity without predatory or discriminatory pricing from the shared global exchange to end-users, including individual institutions, research groups and parties that other participants desire to reach within their service area.

A bandwidth management “headroom policy” will be in place that establishes how capacity and capabilities will be used to monitor, manage and augment capacity on the network to assure that advanced applications have ample and necessary capacity to achieve exemplary performance.

Global Wavelength Service

The network will allow individual participants to purchase, at average cost, additional dedicated capacity on the network, among, between and through any of the add-drop points established on the network. Wavelengths will be offered on a non-discriminatory and policy free basis to any participant in the service.

Such wavelengths will allow specialty use cases to contribute to the scale economies of operating the network, while also allowing specific science, research, or education initiatives to have dedicated capacity where necessary and ready interconnection to the global shared fabric as well.

Startup and Ongoing Operating Models

This infrastructure, as outlined above, is expected to cost between $1.5B and $2.0B 2013 US-dollars to acquire the 20-30 year IRU capabilities. An additional $100-150M/yr US-dollars will need to be generated to operate the network annually. Such funds will necessarily need to be derived through substantial investments by those nations who require these global capabilities. It is expected that large global economies like Europe, the United States and China will contribute on the level of $350M+ each, whereas smaller nations may contribute on the order of $50+M each. All contributors will benefit from equal access to the eventual acquired infrastructure regardless of their level of initial contribution.

Whenever funding is discussed questions of control arise. The relationship between funding and control is a fundamental issue that the NREN CEOs must discuss. Funding and control are also directly related to OEPs and the level of policy which must be implemented at the OEP. These are critical NREN CEO discussion points.

5. Next Steps

The next step in this process does not involve this document. Instead it involves the NREN CEO Summary document and the NREN CEO Discussion Points document. The ideas put forward and the answers to the questions raised in these two documents will guide the next revision of the main GNA document.

When input has been obtained from the Forum, the draft document will be revised and put forward for additional review and possible acceptance by the Forum as a guiding document in the development of the Forum’s network architecture. Ultimately, this document will represent the collective thinking of the Forum as it will be represented to various funding bodies and private sector partners.

When a Global Network Architecture and principles has been agreed upon, parties that are able and willing to contribute to developing this new network capability will be brought together, under the auspices of the Global R&E Network CEO Forum, to work on the various detailed projects to begin further refinement and ultimate implementation of this ambitious program. It is foreseen that parties will come from the networks of which the CEO is a member of the CEO Forum and others parties that are willing and able to contribute.

These projects and white papers include at least the following topics:

- Network Architecture
- Governance
- Sustainable funding
- Procurement Procedures
- Network Design
- Network Operations

Each of these documents will be reviewed by the NREN CEO Forum in the development process with the goal of Forum acceptance.