Research & Education Community Investment in National-footprint Network Services (RECINNS): Requirements gathering summary

Final - June 2017

Introduction

Internet2 began discussions with the community in 2016 for guidance on a strategy for the next community investment in national-footprint infrastructure. A combination of a rapidly changing environment, growth in demand, and hitting the 20-year anniversary for the Internet2 network combined to make it vital to engage the community deeply in the process.

A call for papers at the Technology Exchange in the Fall of 2016 resulted in exceptional input from a broad set of community participants - more than 30 papers in all. Those who submitted were invited to a workshop in Tempe in January 2017, which was well-attended and engaging. At that workshop, a team was formed to continue a requirements development effort.

This document summarizes the discussions to date, and includes input from the papers, workshop and subsequent team meetings.

Working groups, principles and themes

From the meeting in Tempe, two working groups were formed: Supporting the Academic Enterprise and Supporting Data-intensive Science and Network Research. The team that emerged from the workshop felt it was important to make a distinction between the two, at least at the requirements definition stage, since each has unique requirements and also some that are contradictory. For example, enterprise users are concerned about making maximum use of available bandwidth, where research users need significant headroom to support bursty traffic.

Two key guiding principles emerged from the conversation. They are:

Ecosystem-based collaboration

A user's perception of the quality of a service depends on how well the service can be delivered end-to-end. The R&E environment has many players (regional networks, campuses, NRENs, international service providers, cloud providers, etc.) whose infrastructure may participate in that end-to-end delivery. The service commitment for each of those players ends at the edge of their

infrastructure. Collaborative end-to-end Service Level Agreements (SLA) that cross service provider organizational boundaries usually do not exist.

Multi-institutional research collaborations stress this model, especially ones that include international partners. When the end-to-end service doesn't perform as expected, ad hoc efforts to assemble a team of support personnel from the various service providers along the path between the collaborators ensues. These efforts are time consuming, resource intensive and often lead to less than optimal outcomes for the end users.

The community has responded over the years to address this end-to-end challenge. The perfSONAR tool[1] is one example: by placing active performance measuring devices at key locations one can isolate where performance is degraded along a given path, which is a useful tool for debugging. Another is the idea of a *facilitator*, a person skilled in various aspects of cyberinfrastructure who can help the end user maximize their use of the infrastructure. Often engagements with a CI engineer (or team of specialists) results in a good outcome, but not without a lot of work.

The team has recognized that a more comprehensive approach would be to build a closer collaboration between community service providers and offer ecosystem-wide end-to-end services that are supported more robustly than is possible today. The challenges of doing so are many - including potentially challenging current business models - but the initial approach to addressing the problem lies in experimenting, which brings us to the second key guiding principle.

Experimentation and continuous development

In a rapidly changing environment, prognostication about future needs is difficult - most of us are generally constrained by the immediate problems before us and even if we see trends emerging, articulation of specific solutions that anticipate those trends remains elusive.

Today's R&E environment *is* changing rapidly. Traffic demand is growing, and fulfilling that demand with current technology platforms is becoming increasingly difficult to accomplish affordably. Traffic patterns are also changing - research is more collaborative, enormous amounts of data are being generated that need to be moved and processed, community members are relying more heavily on cloud services.

The team has proposed that an appropriate answer to this rapid evolution is to make small investments in experimenting with new approaches - rather than one large "rip and replace" upgrade. One advantage is that new ideas can be tried and, with the appropriate planning and expectation setting, scaled up or abandoned depending on outcome. This "fail fast" approach is more resilient to changes in demand and technology market forces.

[1] http://www.perfsonar.net/

Another advantage to the incremental approach is that it allows for new collaborations and partnerships to emerge - which ties in nicely with the first guiding principle (ecosystem). Not all of the community investment needed will be for technology - experiments and pilots that require new collaborations may show us that investments in people and skills development are needed to properly support bespoke service needs.

Taken together, these two key principles point to the need to embrace new ideas that solve emerging problems, build collaborations and community-wide processes that can pilot and evaluate those solutions efficiently, and work to scale up the solutions found to be effective and thus drive larger scale investments in technology, people and partnerships.

Supporting principles

The team also derived a few important supporting principles:

- For the academic enterprise: community investments in infrastructure should provide production services with unique value propositions that are distinct from those available from commercial service providers.
- For discipline research: the community should invest in a flexible platform that allows for facilitated customization of services that support specific science workflows - including the people and processes required to make that work.
- For infrastructure research: the community should invest in a flexible-scale, non-disruptive testbed platform that allows for exploration of new infrastructure technologies and, where appropriate, allows for close partnerships with industry to effect a rapid transition to production products and services.
- Community-wide skills and processes - the community should work to adopt DevOps principles and processes and ensure that the current and future technical workforce is trained and equipped to operate and thrive an environment of deep customization and programmability.

Themes

The following major areas of discussion emerged during the group meetings, and are summarized thematically with suggested further work or liaison activity described as appropriate.

Cloud integration

Internet2 can play a role today in helping support the academic and research use of commercial cloud providers. The main value proposition is straightforward - Internet2 should aggregate the community's demand and provide a premium access service for the community via existing enterprise connectivity programs (such as Microsoft Azure Express Route, AWS Direct Connect, etc.) and the national backbone.

For the longer term, Internet2 should represent the community directly with commercial cloud providers to define premium access services more tailored to the community's need to support specialized research use of cloud resources. The community needs the flexibility to optimize use between existing community HPC resources and commercial clouds which will require orchestration tools and capabilities that are not generally available today. Since this is an active area of discussion and research, Internet2 should seek to leverage existing community-based efforts for guidance, and assist the community in partnering with and influencing large cloud providers.

This is an area particularly ripe for experimentation. The group discussed a proposal from RENCI and MCNC for a hybrid cloud pilot that was recently sent to Internet2. Given that there is a clear near-term need, additional proposals from the community (perhaps including industry members) should be encouraged and pursued as soon as possible.

Mobility and IOT

The working groups discussed the emergence of Internet of Things (IOT) traffic on campuses, and expressed concerns about the impact on campus architecture and capacity. It was pointed out that traffic patterns are changing, sometimes in unanticipated ways, and that overall bandwidth needs are growing. The impact this will have on regional networks or the national backbone, however, is unclear.

The modern academic campus is like a small city - and some of the thought leadership being applied to smart city projects might prove useful in planning for a more IOT-influenced future. There are a number of smart city initiatives going on in the community - including national efforts such as US Ignite - that have an interest in providing a "smart edge". Internet2 should reach out to those communities to seek to understand what kinds of experiments and pilots might be useful for planning and evaluation, while keeping in mind that it will be necessary to be tightly in sync with regional network and campus partners since discerning where "the edge" should be will be part of the consideration for future ecosystem architecture development.

Programmability and Orchestration

The current programmability offered in AL2S has been useful for the network and cloud research community. Since the current AL2S API (the GENI Aggregate Manager) has been used fairly extensively, stability of that API is a key requirement. However, there is also a desire for increased capability and deeper levels of programmability - researchers would like to have direct access to infrastructure slices with as thin a stack in the way as possible.

Also, there is an emerging need for applications or brokers to request infrastructure services with specific characteristics - for example a network path with consistent latency or one with specific security/compliance characteristics. These bespoke services need to be provided end-to-end across the R&E ecosystem, and thus require orchestration and operational support across multiple service provider entities. Coordination will be required with other community

infrastructure operators (regionals, other NRENs, campuses, etc.), those in the research community who are doing work in this area, and those in the vendor community who are addressing this need in order to define experiments that point the way towards common service characteristics and metrics that can be supported ecosystem-wide.

Security and Privacy

Beyond current day needs to protect users and infrastructure from attacks by bad actors, it will become increasingly important to consider the implications of IoT - the projected growth in this sector and the plethora of devices with little or no security built-in combine to form a concerning trend.

A key value that the R&E community infrastructure can bring is to support innovative security pilot projects and research efforts at scale. A variety of network measurements and other information about the network should be made available to researchers in an automated and secure manner that preserves individual privacy. Additionally, research that requires access to real traffic at scale should be supported in a manner that doesn't compromise the integrity of the underlying network service. Fulfilling these needs will require coordination with security researchers, organizations engaged in security standards such as NIST, and funding agencies such as the NSF.

Peering

The current TR/CPS peering service has been valuable for the community, but with changing traffic patterns and overall growth, the time is right to consider adjustments. Most regionals and some campuses use TR/CPS as a part of their peering strategy, which often include other open or direct peerings with key content providers. Those strategies often end up optimizing peering locally (i.e. on a regional network basis). The community should investigate if a broader, more nationally optimized strategy would provide better service for all or for some community members.

A key recommendation is to coordinate with the Quilt "blender" group that has been discussing some of these alternative optimization strategies for the collection of regional networks. Other options (such as a liaison with the Internet2 Network Technical Advisory Committee - NTAC) should be considered to discuss strategies and opportunities to pilot new ideas in this area.

Testbeds for Network and Distributed Systems Research

There was significant interest and discussion in the group meetings about the need to support testbeds for the network and distributed systems research community. The next generation network should continue to deliver a flexible platform upon which this community can build experiments, with key characteristics such as isolation (between experiments, and between any experiment and production services), service guarantees, the ability for applications to choose

optimal paths, and enough control and predictability so that experimental results can be repeated reliably.

Current examples such as the NSF-funded Global Environment for Network Innovation (GENI) and the Information Centric Networking testbed (ICN, sponsored by Cisco) rely on hardware/software nodes built by the research community that reside on campuses. As those projects evolve in the future, functions in those nodes that are generally useful and that become stable and lean towards production should be incorporated into the ecosystem-wide infrastructure that is supporting the testbed. This will require that what is now a transport network handoff become more of a flexible edge node that can support a full range of Network Functional Virtualization capabilities with the ability to reliably isolate each experiment.

As of this writing, the NSF has a let a solicitation called “Tomorrow’s Internet Project Office” (TIPOFF). Once the award is made, it will be key to liaise with the winner and coordinate requirements.

Telemetry and Analytics

A strong push into streaming telemetry and analytics platforms will bring great benefits operationally, but also provide opportunities to help the community innovate in this area, at a national scale. Insufficient transparency about how the network is performing limits the ability of application developers to improve Quality of Experience (QoE) for end users. This is an area where the R&E community should be at the forefront - providing the kind of transparency and availability of data for research purposes that commercial service providers are reluctant to make available.

There are a number of active research projects in this area that can be solicited for requirements and perhaps active collaborations.

Next steps

This summary was intended to capture key points in the papers submitted by the community and subsequent discussions in the workshop, working group meetings and elsewhere. During the next phase of this effort, we will continue to evolve conversations about requirements, extend those conversations to a broader stakeholder base, and continue to mine the rich content in the papers. In addition, Internet2 will continue to support community-led efforts that can provide important guidance for future requirements, such as the nascent National Research Platform[2], the upcoming NSF award for the TIPOFF solicitation (Tomorrow’s Internet Project Office)[3], the Information-Centric Network (ICN) testbed[4], and others.

---

[2] http://prp.ucsd.edu/events/the-first-national-research-platform-workshop
[3] https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505384
[4] https://goo.gl/FP4H9Q

Internet2 staff will now turn attention towards organizing a set of proof of concept (PoC) projects that will explore technology, operational and organizational opportunities that require more exploration - using as guideposts the two main principles of ecosystem-wide development and experimentation. We will seek out community members who are willing to engage and provide resources to partner in these efforts, and find alignment with the overall strategy of investing effort now to begin shaping our future infrastructure. The intent of these individual PoC efforts will vary, but the overall goal will be to provide the community with first-hand experience in the technologies and coordination strategies that will allow us to collectively support the global R&E community's needs for advanced infrastructure for teaching, learning and discovery in 2019 and beyond.