Fixing Today's Internet!

Problem

In solving old problems and making new discoveries, scientific communities such as Astronomy, Earth Science, and High Energy Physics rely on globally-distributed high-performance computing environments. These environments have helped advance collaborative scientific projects, but not nearly at levels commensurate with their potential. For example, local area Ethernet networks (FastEthernet and Gigabit Ethernet) could readily provide 100 Mbps and 1000 Mbps data transfers, respectively, in campus networks. Yet 90% of end-users' transactions achieve less than 10 Mbps data rates; and only 1% receive anything close to 100 Mbps![1] Two problems inhibit a more advantageous use of computing environments for greater advances in scientific projects:

1. User expectations for shared infrastructure are too low. With 10 Mbps rates as the perceived norm, many flaws that could be fixed go unreported and unresolved, slowly degrading the system. Resetting expectation levels of scientists and support staff is a major challenge that we propose to address.
2. Practices for resolving computing problems often take longer than necessary. Delays are caused, in part, by an opaque network that presents the same symptom of slow performance to end-users regardless of the underlying cause. The challenge involved in addressing this problem is developing procedures and guidelines that truly connect with users' needs and knowledge and direct them in identifying and accurately defining reported problems so that they are efficiently resolved by the appropriate people.

Proposal and Significance

In the NSF-funded *Bridging the Gap Workshop* (BTG) in August 2005, we brought together 40 scientists, network operators, application developers, and network wizards from across the country. From these diverse perspectives, the group examined the causes of and potential resolutions to the two fundamental problems mentioned above. Based on Workshop findings, we have developed an approach for reducing the impact of these problems that we propose to implement in a pilot site.

This approach uses a combination of automatic tools to aid in the detection of multiple problems coupled with self-guided troubleshooting documentation tailored to the needs, priorities, and levels of understanding of stakeholders from each of the four groups. Our proposed pilot project will assure the availability of tools and will develop, implement, and test this self-guided documentation with a scientific community.

This proposed project rests on scaleable methods. If our pilot succeeds, results will be expanded across scientific communities so that end-users in any community can easily find and use diagnostic resources to promote progressive escalation through multiple levels of technical support. These fundamental improvements in communication across

---

[1] Weekly Abilene data is available from http://netflow.internet2.edu/weekly/

functional areas will enable engineers and scientists to collaboratively realize network performance potential, ushering in dramatic advances in research.

Causes of the Problems

1) Low Expectations.
Computing and network power and its potential for end-users' scientific activities is unprecedented. Scientists today use local computing resources that easily surpass the CPU, memory, and storage capacity of the 80's generation supercomputer. A decade of performance improvement (e.g., doubling the CPU clock speed every 18 months) in these computer resources leads to conclusion that faster computers reduce the time it takes to run a job, or more jobs can be run in a fixed amount of time (e.g., applications will continue to run faster on newer CPUs).

Yet end-users' expectations for network performance are unrealistically low, shaped by (a) previous experience, (b) poor application design/implementation, (c) faulty mental models of the end-to-end system, and (d) infrastructure limitations (users have been conditioned to accept poor performance from Internet-based applications). The effect of these low expectations is that most scientific communities are regularly operating at less than 0.1% of the backbone network capacity. In fact, most scientific environments could achieve 100 Mbps, or a significant fraction of that capacity on a regular basis. Failure to achieve these results indicates that real problems exists, and those problems should be eliminated not ignored.

In the local area, Ethernet networks (FastEthernet at 100 Mbps and Gigabit Ethernet at 1000 Mbps) have become the de facto standard for campus networks. Scientists also hear about massive DWDM-based wide area networks that promise terabit/sec networks spanning the globe. These two factors should be driving expectation up, but actual experience shows that networks do not behave like computers. Data collected on the Abilene network shows that while bulk data transfer programs send about 25% of all the packets seen on the backbone, over 90% of those transfers achieve less than 10 Mbps data rates. Even worse, less than 1% of those transfers is even close to the 100 Mbps range. Operational responses such as "the NOC doesn't see any problems" also drive home the impression that these poor performance numbers are to be expected. Thus, scientists are conditioned to expect poor performance when applications use today's shared network infrastructure.

2) Lengthy Network Problem-Solving Practices
Flaws can exist in multiple places and at multiple levels of the OSI stack. For example, application performance can be degraded by (a) infrastructure problems, (b) improper Operating System (OS) settings, and (c) poor application design. Yet, no matter what the problem is, or how many problems exist, the user sees one symptom – the application doesn't perform well.

Most specialized disciplines have a unique domain-specific language that is used by the community for internal discussions. That is, network operators, application developers, domain scientists, and network wizards, all have unique languages. Yet, there are times when these groups must “speak the same language” to deal with technical problems whose causes reach across many network domains. Many diagnostic tools have been developed and made available to help define and locate these underlying problems that impact application performance. As the BTG report shows, however, simply developing or widely deploying tools will not solve some of the underlying social problems. Those problems must be solved in a different manner. When end-to-end performance problems require combined stakeholder involvement in finding and resolving their source, all these communities must communicate effectively

The BTG Workshop began this new approach by including experts in the field of social communications in the discussions between the four communities listed above. As that report shows, developing a common understanding of what language each technical community uses to describe problems and their solutions is an important pre-requisite to real progress in this area.

Proposed Methods and Objectives

1) Resetting Expectations.
Our methods aim to develop a greater awareness and understanding among scientists about what performance levels are routinely achievable. After implementing our pilot approach to resolving the “low expectation problem,” scientists will be able to:

- Communicate effectively with the campus network operations staff to ensure that they know how their labs and offices are connected.
- Be aware of how upgrades are performed and how to plan/finance them.
- Use tools with intuitive user interfaces that can assist in performing basic trouble shooting tasks.

Specifically, we will achieve these objectives through the following approaches:

- Outreach activities (i.e., workshop presentations, web sites) that highlight how some users have managed to achieve over 100 Mbps data transfer rates.
- Highlight performance activities (i.e., contests, science community results, record setting events) that demonstrate the network’s potential.
- Support wide-spread deployment of automatic tools that can be used to verify that 100 Mbps data transfer rates are achievable.

2) Introducing Effective Problem Resolutions Practices
While the BTG report helps define the ultimate goal of combined improvements in cross-functional communication and tool use, it does not chart an explicit path that will get us to this goal. Work needs to be done to develop, test, evaluate, and refine explicit tasks that can scale to meet the needs of many different scientific communities. We propose a multidisciplinary approach that tightly couples deploying tools and creating self-guided documentation to assist numerous scientific communities. This documentation would include presentations (language, content and forms) and troubleshooting guides that are directly relevant and meaningful to scientists’ “worlds of work” and that take time and resource constraints into account.

After working with the scientific community, our objective is to provide documentation and tools that assure:

- Improved initial trouble reporting process by both pre-deploying automatic tools (NPAD, NDT, etc) and generating interactive, context-specific documentation that directs debugging and repair activities.
- Tool-generated reports and instructions that are clear and specific enough where the true end-user can accurately identify the failing subsystems (e.g. application, end-system, network path), sufficiently to identify someone with the right expertise for the first (or next) level of technical support.
- Recognition that the true end-user is the only participant with a vested interested in the ultimate outcome. The end user will be able to periodically refresh the diagnostic reports to confirm any reported status changes and to reassert problem reports to technical support people.
- Diagnostic tools and reports will present identical views of the network to the end-user and to multiple levels of technical support. This is likely to include methods to hide or expose excess measurement detail.
- At any given audience level, the reports will include navigational aids to steer people to the proper level of measurement detail.
- At any given audience level, the reports will include tutorial information, to help support staff to stretch to the next skill level or to enable aggressive end-users to take action on their own.
- The diagnostic resources (automatic tools and self-guided documentation) will have multiple starting points for end-users (or support staff) having different backgrounds or points of view. Each starting point will lead the user (or support person) down a debugging path that is sensible from their own point of view.
- The starting point should be easily located via public means, such as Google.
- Ultimately, an end-user should be able to easily motivate a "normal" campus support staff to debug and correct all common networking problems, without involving people at any higher level on expertise than absolutely needed. And, in the process, better educate themselves to be more prepared to correct the next problem.

Conclusion

Two major issues – low user expectations and application level performance identification – have a negative impact on the future growth of network usage by scientific communities. A team of researchers from the University of Michigan, Pittsburgh Supercomputer Center, and Internet2 are defining a scaleable set of tools and self-guided documentation that can be used by scientists and their support staff to resolve these issues. This team is proposing to work with the Sloan Digital Sky Survey (SDSS) program on a pilot project to develop, evaluate, and revise these tools and documents will scale to meet the needs of numerous scientific communities.