

Déjà Vu

Describing recent breakthrough achievements demonstrating potential new economies in access to high-performance communications and computing, a colleague used the Yogi-ism, “It’s *déjà vu* all over again.” The referenced memory was the convergence, during the 1980s, of technology innovations in distributed communications systems and in the commoditization of computing—a convergence that led to the Internet. Like the first-generation Internet, these latest innovations may take the array of capabilities and tools that were once the exclusive province of a few federally sponsored, “big science” researchers and computer scientists and open them to all faculty and students, to scholarly research, and to learning and teaching.

Affordable Terascale Communications “An *ideal* electrical communications system can be defined as one that permits any person or machine to reliably and instantaneously communicate with any combination of other people or machines, anywhere, anytime, and at zero costs.”¹ With these historic words in 1964 Paul Baran, of the RAND Corporation, established the Holy Grail of twenty-first-century communications. Baran’s ideal was viewed by the telecommunications and computing industry as radical. Baran’s solace, perhaps even amusement, was that he foresaw both that the demands generated by computers would make the 1960s communications system obsolete and that computer technology would be a critical element in the replacement of that 1960s system. The emergence of ARPANET in the late 1960s, the

development of the internetworking protocol in the 1970s, and the establishment of NSFNET in the mid-1980s culminated in the commodity Internet of today—and in the first phase of the pursuit of Baran’s ideal communications system.

The second phase began in September 2003, when a consortium of U.S. research universities, Internet2, Cisco, and Level(3) Communications announced the creation of the National LambdaRail, a nonprofit initiative known as NLR (<http://www.nationallambdarail.org/>), to build a national optical research network reflecting the full potential of new and emerging communications technologies. This national infrastructure will enable terabit-per-second networks reflecting two to three orders of magnitude improvement in price per unit of bandwidth. It has been described most succinctly by Peter O’Neil, of the University Corporation for Atmospheric Research: “The fundamental and overriding goal of NLR is to provide an enabling experimental infrastructure for new forms and methods of science and engineering.”²

NLR leverages several new ideas as the basis for a very different communications infrastructure model. First, a reasonable assumption in today’s market is that with large-scale procurements, consortia of communications users can acquire optical fiber at a price at or near cost. In conjunction with Internet2’s FiberCo initiative, NLR has access to or has options to acquire thousands of miles of fiber interconnecting most of the nation’s largest cities.³

Second, NLR is lighting the fiber with Dense Wavelength Division Multiplexing (DWDM) optical technology. The initial deployment will be capable of supporting

on each available fiber pair up to 40 simultaneous light wavelengths, or lambdas, and data rates per lambda of 10 gigabits per second (Gb/s). DWDM technology is theoretically capable of supporting hundreds of lambdas per fiber pair and data rates per lambda of over 100 Gb/s.

What NLR is not doing, except on a special case-by-case basis, is deploying and relying on the established telecommunications industries’ significant investments in SONET technology, used to light fiber and to provide redundancy. Some telecommunications industry analysts estimate that SONET technology adds a tenfold premium to the cost of network services based on large-scale optical infrastructure.⁴ To provide redundant routing capabilities, NLR will develop the national backbone as a series of interconnected loops to provide alternative paths, deploy management and control capabilities within the DWDM optical technology domain, and utilize rerouting capabilities at the network services level.

At the national and regional levels, higher education institutions generally pay tens of dollars per megabits per second (Mb/s) per mile per year for dedicated communications links. An initial analysis of the NLR cost structure suggests the possibility of dedicated inter-regional lambdas costing less than ten cents per Mb/s per mile per year.

The NSFNET expansion based on the original ARPANET concept stimulated significant regional and statewide network efforts in the late 1980s, bringing most of the U.S. colleges and universities into the Internet world. Likewise, this national optical infrastructure effort today is already resulting in regional infrastructure development

efforts that reflect the architecture and technology being deployed by NLR. Higher education institutions in most regions of the United States have either built,⁵ are building,⁶ or are organizing to build⁷ the regional infrastructure to interconnect with the NLR infrastructure.

Commodity Supercomputing

The National Science Foundation (NSF), the Department of Energy (DOE), and other agencies are funding multimillion-dollar projects to enable the petascale computation necessary to model and solve the most complex science and engineering problems. This will require computing capable of operating at trillions of floating-point operations per second (TFlop/s) in a scalable distributed computer grid enabled by optical network technologies of the sort being deployed by NLR. Given the cost for the requisite computational facilities, up to \$10 per million floating-point operations per second (MFlop/s), there are very few higher education institutions in this arena. For instance, the top-rated supercomputer—the 35.86-TFlop/s Earth Simulator Center at Yokohama, Japan—is estimated to have cost \$250 million to \$350 million, or between \$7 and \$10 per MFlop/s.⁸

Since the 1980s, scientists like Eugene Brooks, at the Lawrence Livermore National Laboratory, and Thomas Sterling, at the Jet Propulsion Laboratory, have been advocates for the potentially extraordinary power and economies in large-scale parallel computer architectures. Noting that the future of high-performance computing was dependent on utilizing architectures aligned with, and leveraging, the cost-performance gains being realized in processor technology and consistent with Moore's law, Brooks gave substance to the refrain, "No one will survive the attack of the killer micros."⁹ Sterling is one of the developers of the Beowulf model for high-performance computer systems. Beowulfs are parallel systems built with commodity hardware and open-source software. The promise of this model is to deliver a price-to-performance ratio that puts high-performance computing into the hands of individual researchers and into small laboratories and that opens this alternative to a wide array of applications.¹⁰

The top 100 high-performance computer systems on the "TOP500 Supercomputer Sites" list (<http://www.top500.org>), which ranks the sites according to the measured performance of the largest problem run on each system, all have a performance capability greater than one TFlop/s. Assuming that the upper budgetary bounds of most laboratories or academic departments for computing and network facilities range from a few hundred thousand dollars to a million dollars, the goal of broad deployment of terascale computing facilities suggests that the maximum cost must be well under \$1 per MFlop/s.

A supercomputer assembled by Virginia Tech has met that goal. In November 2003, a TOP500 spokesperson announced: "Virginia Tech's System X was only the third system to exceed the 10 TFlop/s benchmark, at 10.28 TFlop/s." This system has a theoretical peak-performance capability of 17.6 TFlop/s. The total cost of the computer and communications hardware configuration was \$5.2 million, for a cost per theoretical peak performance of 30¢ per MFlop/s. The Virginia Tech project and system name, "System X," was chosen because its original goal was to build a ten trillion Flop/s (10 TFlop/s) system utilizing commodity components—1,100 Apple Power Mac G5s based on dual IBM 64-bit PowerPC 970 processors—and because it is based on Apple's Mac OS X operating system.¹¹

Notable problems need to be solved to increase the utility of large-scale parallel computer systems. However, with this milestone in building a large-scale cluster supercomputer with commodity components, we have compelling evidence that a broad set of the nation's scholars and researchers may be able to participate in the supercomputing-dependent data assembly and analyses, simulations, and computations required to solve our world's most complex science and engineering questions.

A Universal Cyberinfrastructure

With these demonstrated new economies in terascale computing and communications, the U.S. cyberinfrastructure will include, within this decade, thousands of interconnected TFlop/s computing systems accessible to the majority of the higher education research community.¹²

These accomplishments, representing the progress in distributed communications infrastructure and in high-performance computing, have profound implications for academia, for the United States, and for the world. Timely solutions to the complex and serious challenges in today's world require broad collaborations tapping every possible contributor and require open, frictionless access to the most powerful computation and data-assembly facilities. The dominant view of today's Internet is of a huge consumer network. But its greatest potential lies in its transformation to a global producer network. Higher education can lead this transformation: every member of the community is a potential producer and contributor of research and scholarship in this new world, and every member can help build the infrastructure and the literacy on which it depends.

Notes

1. Paul Baran, *On Distributed Communications* (Santa Monica, Calif.: RAND Corporation, 1964), <<http://www.rand.org/publications/RM/RM3767/RM3767.chapter1.html>> (accessed March 23, 2004).
2. Peter O'Neil, e-mail to NLR principals, December 1, 2002.
3. See <<http://www.internet2.edu/pubs/200310-WIS-FC.pdf>> (accessed March 23, 2004).
4. Roxane Googin, "How Networking Advances Screwed Up the Economy," *isen.com Smart Letter #64* (December 16, 2001), <<http://www.isen.com/archives/011216.html>> (accessed March 23, 2004).
5. For example: California, <<http://www.cenic.org/>> and Illinois, <<http://www.iwire.org/>>.
6. For example: Ohio, <<http://www.tfn.oar.net>>.
7. For example: Florida, <<http://www.flrnet.org/>> and the mid-Atlantic region, <<http://www.midatlantic-terascale.org/>>.
8. See the "TOP500 Supercomputer Sites" list: <<http://www.top500.org/>>.
9. On Brooks, see "Killer Micros Change Computing," *Up Close on LDRD*, June 2003, p. 2, <www.llnl.gov/llnl/06news/Employee/articles/2003/06-27-03-newsline-ldrd.pdf> (accessed March 23, 2004).
10. On Sterling, see <www.cacr.caltech.edu/~tron/>.
11. For more on this project, see <http://computing.vt.edu/research_computing/terascale/>.
12. For more on this, see *Revolutionizing Science and Engineering through Cyberinfrastructure*, Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure, Daniel E. Atkins, chair, January 2003, <www.cise.nsf.gov/sci/reports/atkins.pdf> (accessed March 23, 2004).

Erv Blythe is Vice President for Information Technology at Virginia Tech. With this column, he begins a two-year term as Editor of the New Horizons department in *EDUCAUSE Review*.

