# INCITE – Edge-based Traffic Processing and Inference
for High-Performance Networks

Richard Baraniuk, Edward Knightly, Robert Nowak, Rolf Riedi, Rice University
Wu-chun Feng, LANL
Les Cottrell, SLAC

Summary

The explosive growth of high-speed computer networks, combined with rapid and unpredictable developments in applications and workloads, has rendered network modeling, control, and performance prediction increasingly demanding tasks. High-end applications critical to the DOE mission, including distributed computation, remote visualization, and high-capacity data transfers, routinely fail to meet end-to-end performance expectations when deployed on high-speed networks. The INCITE (InterNet Control and Inference Tools at the Edge) Project aims to transform modern high-speed inter-networks into manageable and predictable systems to enable these critical applications. Our interdisciplinary team is developing new theory and methods for network monitoring, probing, and analysis based solely on edge-based measurement at hosts and/or edge routers.
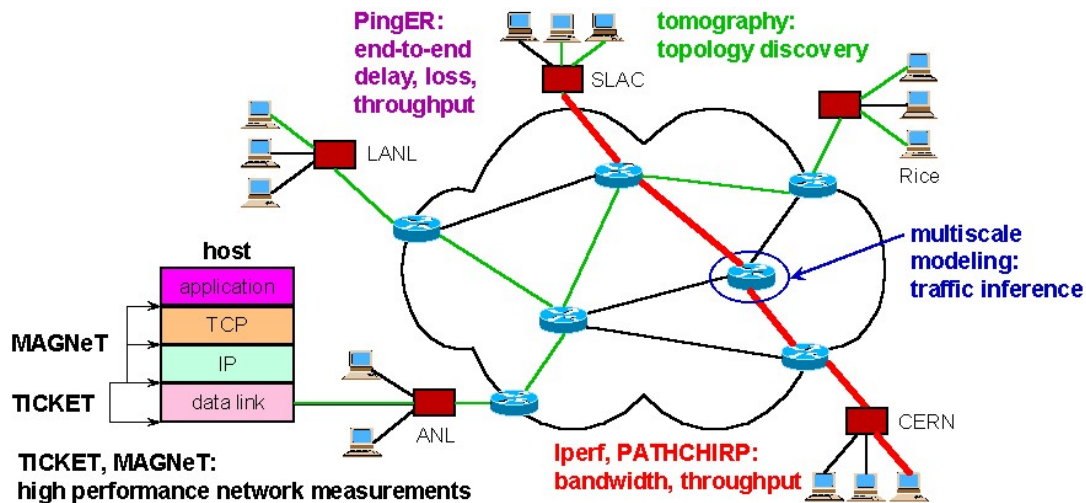
Distributed applications running on clusters and computational grids are complex and difficult to analyze. Moreover, optimizing their performance requires that end-systems have knowledge of the internal network traffic conditions and services. Without special-purpose network support (at every router), the only alternative is to indirectly infer dynamic network characteristics from edge-based network measurements. There is a great need for theories and methods to understand the complexities of distributed applications and network environments with the ability to choose the level of detail to fit the task, be it debugging, tuning, monitoring, or control.

The INCITE Project is developing on-line tools to characterize and map host and network performance as a function of space, time, application, protocol, and service. In addition to their utility for *trouble shooting* problems, these tools will enable a new breed of applications and operating systems that are *network aware* and *resource aware*.

Monitoring tools

INCITE's monitoring tools include *MAGNET* (Monitoring Apparatus for General kerNel-Event Tracing), *MUSE* (MAGNET User-Space Environment), and *TICKET* (Traffic Information-Collecting Kernel with Exact Timing). Together they act as a kind of "network oscilloscope" that can measure (capture packets) at different points in a host, cluster, network, or Grid, from the application to data link layer.

MAGNET and MUSE permit applications and developers to obtain detailed information about the environment on a host and enable new *resource aware* applications that adapt to changes in their environment (load balancing when needed, sensing when a node's resources are scarce or are bottle-necked, and so on). MUSE monitors without requiring modification or re-linking of applications; TICKET serves as a high-speed "tcpdump" replacement.

Edge-based probing tools

INCITE's probing tools include *PathChirp*, ABwE, *NetTomo*, and *NetTopo*. Much as x-rays probe our bodies to find cracks and breaks in bones, these tools inject probe packets into a network to determine its conditions and characteristics. pathChirp and ABwE quickly (<1s) and with little impact estimate the available bandwidth, delay distribution, and tight link along end-to-end paths using an exponentially spaced probe train (pathChirp) and closely spaced packets (ABwE). And much as x-ray tomography reconstructs a 3-d internal view of a person, NetTomo localizes delays and losses on individual network links by injecting probes along multiple network paths. NetTopo allows a user to discover the internal topology of a network through edge-based probing. NetTopo and ABwE have been incorporated into the IEPM traceroute/bandwidth visualization toolkit that assist in identifying performance

problems for Grid, Esnet, and HENP sites. Edge-based probing enables applications to become *network aware*. ABwE is now deployed at about 30 major Grid and HENP sites and 20 PlanetLab sites; we are also working to deploy it in MonALISA. ABwE data has been made available by Grid Services following the GGF NMWG schema definitions.
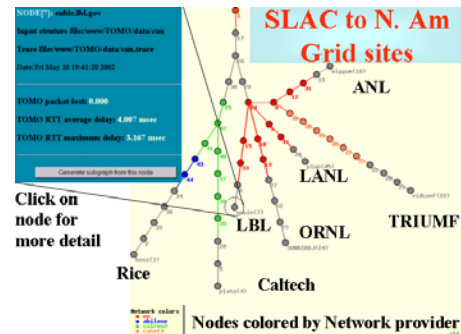
Topology from SLAC to various North American sites, with color indicating service provider.

We are tackling the vexing problems of automating performance change detection and gathering and reporting information to assist in resolution. We have developed and deployed sophisticated multiscale traffic models and analysis software based on wavelets and multifractals. Finally, we have developed TCP-LP (Low Priority) that utilizes only the excess network bandwidth to simplify background transfers of large files across networks.

INCITE users include: Globus, Particle Physics Data Grid Collaboratory Pilot, Scientific Workspaces of the Future, SciDAC Center for Supernova Research, TeraGrid, Transpac at Indiana U., San Diego Supercomputing Center, ns-2, Telecordia, CAIDA, Autopilot, TAU, and the European GridLab project. For more information, see the INCITE website at *incite.rice.edu.*