Western Hemisphere Research and Education Networks - Links Interconnecting Latin America (WHREN-LILA) Research Experiences for Undergraduates (REU) Project

The WHREN (Western Hemisphere Research and Education Networks) project addresses the existing and future needs for improved North American (especially the U.S.)–South American network connectivity. Activities focus on the need for improved connectivity through new network links: LILA (Links Interconnecting Latin America). WHREN formed a consortium of organizations from across the Western Hemisphere to participate in developing and operating a next-generation model for international networking that is now fostering collaborative research and advance education throughout the Western Hemisphere and other world regions. WHREN also serves to increase the rate of discovery both in the U.S. and across the Western Hemisphere. U.S. researchers are part of communities of scientists undertaking experiments that require increased and improved network resources throughout the Americas.

The October 2006 NSF International Research Network Connections (IRNC) program review of the Western-Hemisphere Research and Education Networks – Links Interconnecting Latin America (WHREN-LILA) is a project, award #0441095, recommended to investigate the possibility of collecting netflow data on the IRNC funded links. "If possible, netflow data should be collected and analyzed on the IRNC links to assure appropriate use of the assets." The NSF awarded FIU's proposal for a Research Experience for Undergraduates (REU) program to research and implement these recommendations. Funding this REU request will have a direct and positive effect on helping the WHREN-LILA team to respond to recommendations of the IRNC program review panel.

Research on Network Monitoring and Measurement Improvements for LILA links

In a previous REU project, under STI award #0231844, FIU undergraduate students augmented the functional capabilities of the Monitoring Agents in A Large Integrated Services Architecture (MonALISA[1]) with FlowTools[2], via an UDP-listening agent, ApMon[3]. The resulting tool enabled the creation of dynamic and real time views to improve understanding of the traffic flow characteristics for the networks that connect to the AMPATH international exchange point in Miami, with traffic flows to Internet2's Abilene network.

For this REU project, we plan to build upon the software base of the first REU project to extend functionality beyond a set of independent routers. The proposed research is to improve understanding of traffic patterns and anomalies across the LILA links. The proposed REU students will be included in planning discussions, development, documentation and execution of the proposed research activity. They will be involved in

[1] http://monalisa.caltech.edu/
[2] FlowTools: http://www.splintered.net/sw/flow-tools/
[3] ApMon: http://monalisa.caltech.edu

the experiment design to collect netflow data using the existing flow tools developed from the previous REU award to integrate Cisco NetFlow technology and MonALISA. The REU students will assist in the interpretation of the netflow data, which will then lead to the extension of the current tool set.

The focus of this REU research project has two perspectives. The first perspective is to increase understanding of traffic patterns from a coordinated perspective. A coordinated perspective would permit viewing flow traffic, from multiple sources, as a correlated system. The second perspective is to increase awareness of network anomalies by measuring Round Trip Time (RTT) variations.

Work Plan Summary

Network flow information is available from a NetFlow services device. NetFlow (Cisco) is a widely deployed router-based traffic monitoring mechanism. FlowTools (FlowTools) is an open-source NetFlow analysis toolset underlying the data gathering and analysis infrastructure of our project. A Netflow device is a Cisco router or switch that supports Netflow Services and which exports NetFlow records (Fullmer & Romig, 2000). Netflow services can be used to interpret flows across network links to answer "Who", "When", "How" and "How much" types of questions, which can be helpful to identify source of intrusions and other security concerns.

The challenge is that there is a lack of understanding of traffic pattern from a coordinated perspective. There are tools to interpret traffic patterns as a set of independent routers. What is missing is a coordinated understanding of the traffic patterns in a network, not viewed as a set of independent routers, but correlated. A coordinated perspective would permit viewing flow traffic, from multiple sources, as a correlated system, and provide data to answer questions of the following kind: what are the differences in traffic flows at both ends of the link? Are the links congested? Who are the top talkers at each end of the link? How reliable is the sampled netflow data that is being collected at both ends? For example, do the flows match, or does the sampling correlate at both ends? What real-time data can we gather on a multi-domain NetFlow architecture and what are the challenges associated with this effort?

A sensible approach to resolving these questions would rely on previous REU work (STI award #0231844, that integrated Cisco NetFlow technology and MonALISA) , extending the current infrastructure to incorporate data collection from various routing domains (autonomous systems). The external domains would reside on the far end of long international research links (namely WHREN-LILA), NetFlow data should be collected at a repository close to the originating NetFlow device; as is well documented in the industry's best practice. A NetFlow collector at AMPATH would serve as a central repository for the data and provide sufficient storage for historical data collection for a period to be determined by the requirements of this research project.

A second component to our proposal is to include a function for round trip time (RTT) sensitivity and incorporate router statistics (CPU/memory utilization). On very long inter-continental circuits, measuring RTT in real time could be very useful in detecting anomalies, such as a failure that causes traffic to reroute over alternate, longer latency paths. Higher RTT can also be caused by a spike in router CPU utilization due to the setup or teardown of virtual circuits, MPLS circuits or other intensive activity (desired or undesired). This is of practical significance for the LILA east circuit, from Miami to Sao Paulo. In the event of circuit failure, it is possible that traffic destined to the U.S. could be rerouted over to Europe, taking a much longer route. The data collected from an RTT anomaly sensitivity tool could allow detection quickly and a trace of when it happened and possible cause. A novel approach to this solution is to integrate the RTT sensitivity data with the Netflow correlated data. The goal is to detect anomalies using RTT / router performance sensitivity to then identify the source using the NetFlow correlated data.