OAK RIDGE NATIONAL LABORATORY

FACT SHEET

MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

High-Precision, Layer-2 Network Monitoring at ORNL

There is considerable interest in optimizing network performance using various approaches to so-called network weather services. In general, these involve measurements at layer-3 based on PING, IPERF, PipeChar, and similar level-3 tools. Our interest in measurements at layer-2 can be summarized in an analogy. It doesn't make a lot of sense to correlating rainfall with lake level unless you first measure the position of the flood-control gates.

Objectives

The objectives of this project have been to develop layer-2 tools that would be applicable in a variety of wide area network environments. Prove their usefulness, and make them available to the DOE community.

Components and Test Procedure

The project is based on two major components. The first is a hardware-based system for precisely measuring the time-of-flight of layer-2 traffic in a wide area network. The second is the software system for analyzing the data. The hardware used for the data shown here was based on Duke ATM cell analyzers that could be programmed to generate a trigger pulse when they sensed the passage of a cell with a specified header and payload. That trigger in turn captured a highly accurate time-of-day reading in a GPS clock, causing it to store the time-of-day in an internal register stack. Cell transit times are measured to an accuracy of +/- 50 nanosec.

The software, running on Linux PCs, set up and supervised the testing. Each test consisted of 127 measurements of transit time (this was the height of the register stack in the GPS clock). The individual measurements were separated by 11 msec, with a “burst” of 127 such measurements requiring ~ 1 1/3 seconds. These bursts were repeated at 10-minute intervals. For end-to-end testing, the two Linux systems send data to each other, and exchange results by e-mail.

Results

To understand the results, it probably helps to look at the raw measurement data and follow it through the analysis process. Figure 1, to the right, shows six, randomly chosen burst measurements. The x-axis is time, the data points by separated by 11 msec. The y-axis is time-of-flight, measured to +/- 50 nanosec. The variation in the time-of-flight represents the fundamental “noise” in the various switch fabrics along the route.

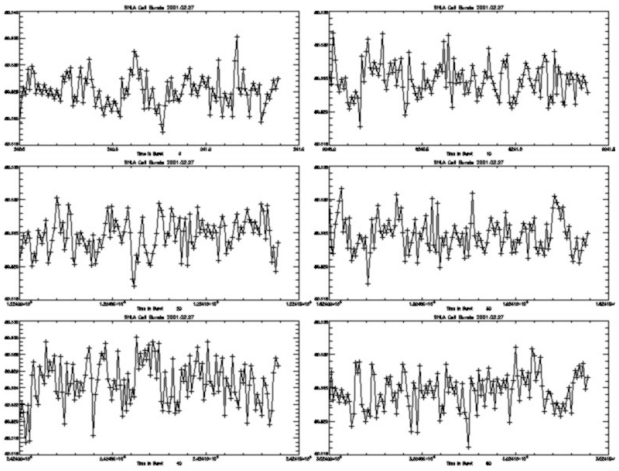

Fig. 1: Time of flight measurements

Remembering that these bursts measurements are repeated at 10-minute intervals, plotting the data from 24 hours yields the sort of image seen in the next figure. Here, all the data points from a burst appear as a single vertical line, and as can be seen, the network is relatively quiescent. Figures 3 and 4 illustrate two different departures from this quiescent state. Figure 3 illustrates a major SONET upset. The time of flight changes by 80 msec. On this scale, the “grass” in Figure 2 is less than the line width. Figure 4 shows an onset of congestion at the FNAL site. It also illustrates a fairly common feature in the data. The baseline time-of-flight changes by about 50 microsec at mid-day. This represents a change in the effective fiber path of roughly 10 miles, probably indicating a re-route from one POP to another for maintenance work.

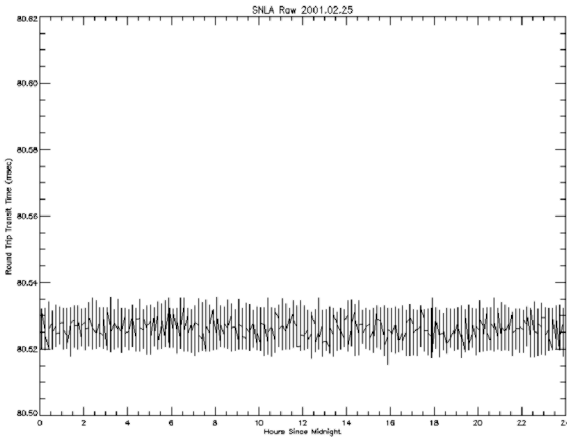

Figure 2: Quiescent data over 24 hours.

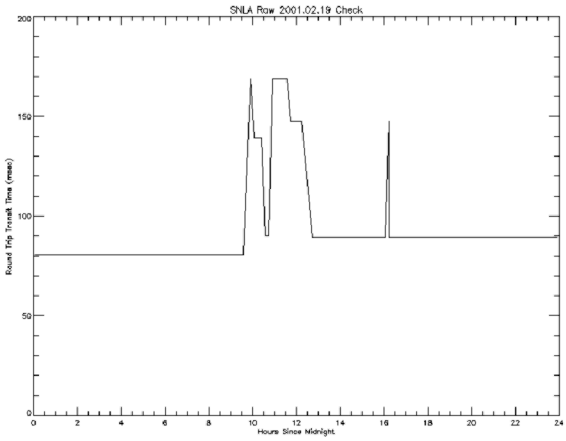

Figure 3: SONET Upset

Figure 5 illustrates a very different effect. We expected that the time-of-flight data would show essentially no correlation between measurements within the noise level represented by the "grass" seen in Figure 1 and Figure 2. However, when we calculate an auto-correlation function along the individual measurement points, they turn out to be strongly correlated, with periodicities that are persistent over times of many seconds. The periodicity turns out to be related to the "brass ring" effect as SONET packets traverse the cross-connects that bridge SONET rings in wide area networks. Figure 6 shows the same data, now analyzed in the frequency domain. There is a frequency peak associated with each SONET cross-connect. Interestingly enough, these peaks persist over many days as seen in Figure 7, changing when there is a major change in the underlying SONET routing or architecture.

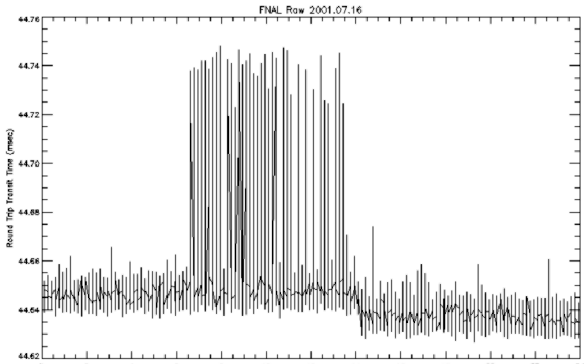

Figure 4: Congestion onset and minor reroute

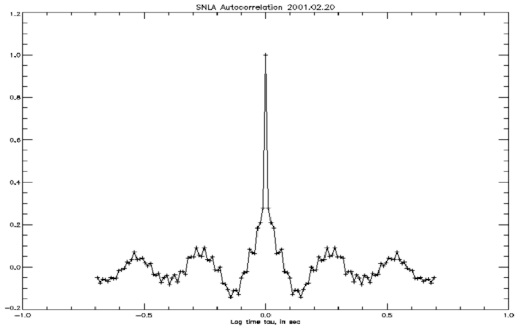

Figure 5: Long-period correlation (seconds) in time-of-flight data

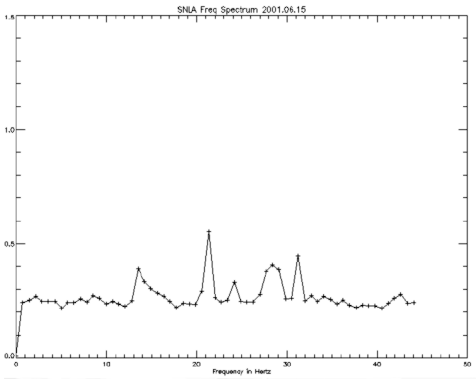

Figure 6: Same data, frequency domain

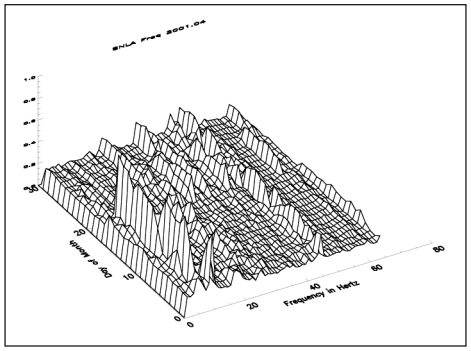

Figure 7: Frequency data, over one month