Scalable visual traffic analysis

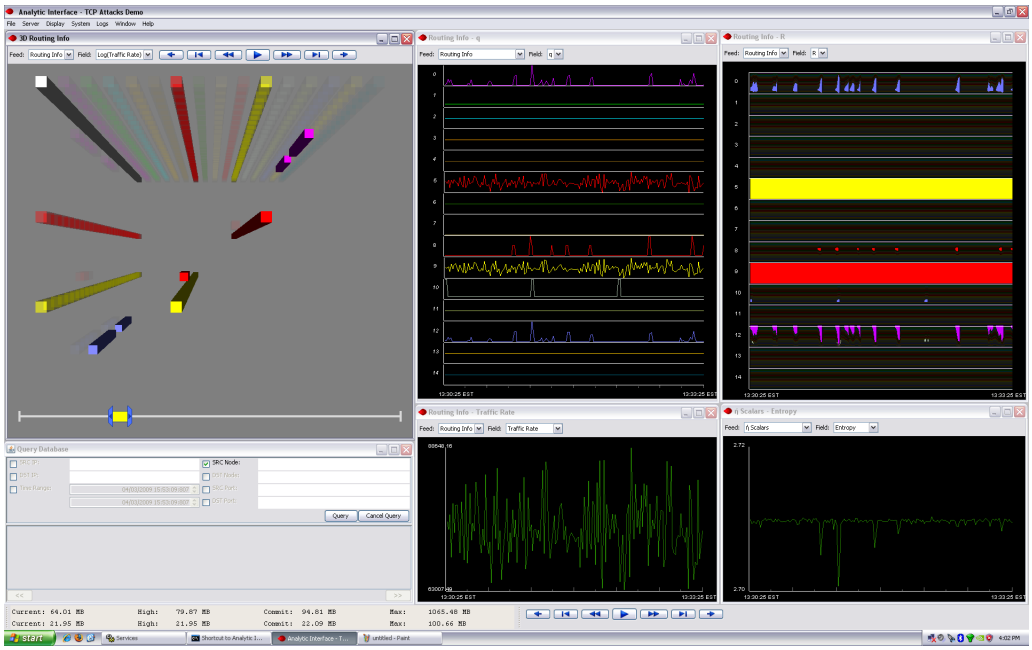
Steve Huntsman, Chris Covington, and John Franklin

Equilibrium Networks' scalable visual traffic analysis system combines visualization of matrix time series with other complementary data reduction methods inspired by information theory and statistical physics to provide deep insight into network traffic. A single classification scheme is used for both source and destination attributes based on layer 3 and 4 packet headers in order to capture bulk and policy-relevant features of network traffic in a normalized format. The system leverages Snort for sensing (though the feasibility of higher-throughput custom sensing architectures has been demonstrated for serial link speeds of 10+ Gbps) and can be adapted to incorporate special packet capture libraries or hardware such as DAGs for high-performance timestamping and data storage, as well as commodity hardware and software for data acquisition.

Using a small number of source/destination attributes enables visualization strategies that are essentially independent of the monitored link bandwidth, since only the nonzero elements of a matrix of fixed size must be displayed for any given time slice. In practice a small number of such attributes suffices because only coarse-grained levels of description of network addresses, TCP services, flow status and the like are necessary to provide a summary of the network traffic appropriate for traffic analysis. Both 2D and 3D displays are provided for the same data that complement each other and are designed to avoid common pitfalls such as occlusion and resolution limitations.

By classifying traffic patterns in an extremely compact and mathematically convenient way, the display intervals can be dynamically resized and shifted, allowing the review of a day's worth of traffic in minutes. Moreover, the classification protocol allows fusion of multiple sensor streams in a conceptually straightforward way and with minimal overhead, so that parallelization can be properly leveraged to accommodate increasing network link speeds. By sampling traffic if and as necessary in a way consistent with the normalization protocol, and using technologies such as MySQL Proxy, the data storage can accommodate a wide range of link speeds and hardware capabilities.

The system's capabilities raise the prospect of significantly reduced time and effort spent on dealing with log files, combined with the ability to view the entirety of traffic, and not just alerts. Our system provides an extensible and scalable infrastructure for security visualization and mathematical characterization of network traffic. It can be used for intrusion detection as well as network behavior analysis, and its basic elements are planned for release under a free/open-source license.



Screenshot of benign TCP traffic from a gigabit network testbed link near saturation.