Clarified Recorder and Analyzer for Visual Drill Down Network Analysis

Jani Kenttälä[1], Joachim Viide[1], Timo Ojala[2], Pekka Pietikäinen[3], Mikko Hiltunen[1], Jyrki Huhta[1], Mikko Kenttälä[2], Ossi Salmi[2], and Toni Hakanen[2]

[1] Clarified Networks
Hallituskatu 9 A 21, 90100 Oulu, Finland
{jani,jviide,mikko,jyrki}@clarifiednetworks.com
[2] MediaTeam Oulu, University of Oulu
P.O. Box 4500, 90014 University of Oulu, Finland
{firstname.lastname}@ee.oulu.fi
[3] Oulu University Secure Programming Group, University of Oulu
P.O. Box 4500, 90014 University of Oulu, Finland
{pekka.pietikainen}@ee.oulu.fi

Abstract. This paper presents the Clarified system for passive network analysis. It is based on capturing complete packet history and abstracting it in form of different interactive high-level visual presentations. They allow for drilling from the high-level abstractions all the way down to individual packets and vice versa. The applicability of the system is demonstrated with the daily management of a large municipal wireless network.

Keywords: Iterative interactive network traffic visualization.

1 Introduction

We present the Clarified Recorder and Analyzer system for passive network analysis. It is a further developed and commercialized version of the freely available HowNetWorks tool that won the first price in VMware Ultimate Virtual Appliance Challenge in 2006 [5]. The system is based on capturing complete packet history, which is abstracted by various visual presentations. They represent different aspects of the packet data, for example individual events, identities, flows between identities, or causal relationships of flows. The visualizations facilitate visual drilling down from high-level visual abstractions to the level of individual packets and back. This facilitates high-level visual analysis of complicated network problems without tedious and time consuming detailed study of large amounts of packet data. Further, the availability of the captured packet data allows reactive assessment of security threats based on real traffic in the network. We describe the design and implementation of the proposed system and report its usage in the daily management of a large municipal wireless network.

2 System Design and Implementation

The architecture of the Clarified system is illustrated in Fig. 1. The Recorder captures traffic from one of more network interfaces creating a track, which is imported by the

S.B. Moon et al. (Eds.): PAM 2009, LNCS 5448, pp. 122–125, 2009.

Analyzer for iterative visual drill drown analysis. The Recorder uses an approach similar to the Time Machine [1], so that data collection agents (taps) are placed throughout the network. The Recorder collects packets from one or multiple taps, possibly filters the packet stream with a set of predefined filters, and stores the packets into multiple pcap files in a ring buffer fashion, together with a flow index computed on the fly [4]. The Recorder is able to simultaneously record and export the captured data to the Analyzer. The main design goals for the Recorder have been protocol independence and capturing of complete network traffic. The software components can be deployed on a standard off-the-shelf PC. The hardware dictates the scalability of the Recorder in terms of the effective packet rate that can be recorded without packet loss. The performance of the tap implemented in native C is limited by libpcap and the speed of the disk. With a high-performance RAID5 array and an optimized version of libpcap [6] multiple Gigabit Ethernet streams can be stored to disk without loss of data. The current bottleneck in the Recorder is the indexer written in Python, which is able to process about 40000 pps (~250 Mbps) on a high-end quad-core 2.6 GHz Xeon machine.

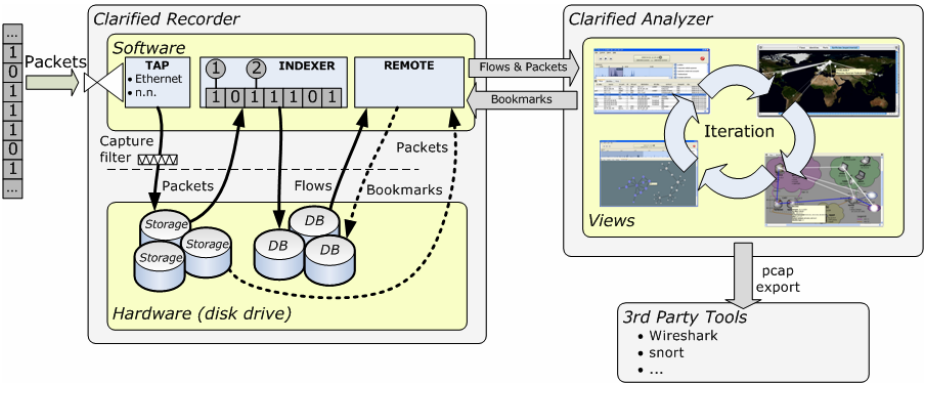

Fig. 1. The architecture of the Clarified system

The Analyzer imports a track (i.e. a flow index) created by a Recorder. The track is represented and manipulated on a timeline, which allows 'time travel' within the track. The Analyzer can conduct real-time analysis by importing a 'live' track that is currently being recorded by a Recorder. The Analyzer illustrates causal relationships [4] between different events in the track and other aspects with different types of visualizations, including *connection graph* (flows between identities), *layer graph* (relationships between entities), *association graph* (combination of connection graph and layer graph), *topology view* (flows rendered on a network diagram), *earth view* (data on a world map) and *DNS timeline* (TTL values of DNS resource records). The so-called monitors provide basic textual views of the packet data, including *flow monitor* (flows), *identity monitor* (identities) and *port monitor* (port statistics).

The main design goal for the Analyzer has been high-level analysis using 'living views'. It refers to the functional relationship between the visualizations provided by the Analyzer and the underlying packet data or network documentation. For example,

given a particular visualization the user can access the corresponding packet data in real time, which allows for drilling down from a high-level overview down to the underlying flows and the details of individual packets if needed be. Similarly, visualizations can be mapped to network documentation so that the flows contained in the packet data are rendered on the network diagram, allowing visual inspection and understanding of the network traffic. Further, visualizations can be used to filter the input data, filtering affecting other views as well. For example, the earth view can be used for painting the geographical region of interest, on which further analysis will be limited. The packet data of interest can also be exported to 3rd party tools in pcap format. The AAA capabilities include centralized cookie-based user authentication, track-specific authorization, and accounting of Analyzer use and Recorder logins. The Analyzer is implemented in Python.

3 Usage in a Daily Management of a Municipal Wireless Network

panOULU is a municipal wireless network in the City of Oulu in northern Finland [3]. As of now the panOULU network has about 1050 WiFi APs, which provide open (no login, authentication or registration) and free (no payment) wireless internet access to the general public with a WiFi-equipped device. panOULU network has been operational since October 2003. The network usage has been growing constantly so that in September 2008 15127 unique WiFi devices used the panOULU network, totaling ~370000 sessions and ~13.9 million minutes of online time. For a comprehensive description of the panOULU network see [2].

The Recorder and Analyzer have been used in the daily management of the panOULU network since 2006. Three Recorder instances are deployed on a Dell PowerEdge 2900 with 4 CPUs and 8 GB of RAM. The machine has 3 TB of local disk, of which 2 TB is reserved for storing packet data. One Recorder instance is recording outside the NAT router, and two instances inside the NAT router. Capturing packets on both sides of the NAT allows inspecting possible NAT traversal problems, and deploying multiple Recorder instances with different filter configurations provides better scalability. Each instance is configured to record the first 100 bytes of each packet for the purpose of avoiding recording of private data possibly contained in the payloads. Since the average traffic is ~20000 pps, the 2 TB of disk is sufficient for recording ~11.5 days of traffic. So far peak rates have been ~71000 pps, which the multiple Recorder deployment has been able to capture without any packet loss.

The ‘time travel’ functionality is one of the main benefits of the Clarified system in the daily network management. If and when something abnormal happens, we can always easily return to the incident, as long as the data is still available on disk. A typical case is a rogue IPv6 router. Windows Vista operating system has an interesting ‘feature’: if a Vista machine is configured to use IPv6 routing (which is the default) and Internet connection sharing is enabled, the machine starts sending faulty IPv6 router advertisements, which mess up IPv6 routing in the network. As Vista is becoming more popular, so are rogue IPv6 routers in the panOULU network. To counter the problem we have implemented RogueIPv6Alerter as a Recorder helper application. Upon detecting a faulty IPv6 router advertisement it creates an alert bookmark identifying the event on the timeline. If we are able to locate the machine,

we inform the user of the problem. If not, the machine is disconnected and placed on the list of blocked machines. The next time the machine tries to connect to the network it is automatically redirected to a web page explaining the problem and providing instructions for fixing the problem. Another typical case is an infected machine. We monitor the traffic for patterns involving large amounts of ARP traffic, which is typical for viruses. If and when such patterns are detected, we engage Analyzer to identify the infected machine for corrective action as with rogue IPv6 routers.

4 Conclusion

We presented the Clarified Recorder and Analyzer for visual drill down of network problems using different interactive visual abstractions of the captured packet data. By capturing complete packet history we guarantee that we can always drill down to individual packets when necessary. If we would sample the packet stream, there would be no chance to obtain the original flow information later. However, the trade-off is that the system does not scale up for high speed core networks. We are currently expanding the system towards wiki-based collaborative analysis.

Note. Videos illustrating the usage of the Analyzer are available online at http://www.clarifiednetworks.com/Videos.

References

1. Kornexl, S., Paxson, V., Dreger, H., Feldmann, A., Sommer, R.: Building A Time Machine for Efficient Recording and Retrieval of High-Volume Network Traffic. In: Internet Measurement Conference 2005, pp. 267–272 (2005)
2. Ojala, T., Hakanen, T., Salmi, O., Kenttälä, M., Tiensyrjä, J.: Supporting Session and AP Mobility in a Large Multi-provider Multi-vendor Municipal WiFi Network. In: Third International Conference on Access Networks (2008)
3. panOULU network, `http://www.panoulu.net`
4. Pietikäinen, P., Viide, J., Röning, J.: Exploiting Causality and Communication Patterns in Network Data Analysis. In: 16th IEEE Workshop on Local and Metropolitan Area Networks, pp. 114–119 (2008)
5. VMware Ultimate Virtual Appliance Challenge, `http://www.vmware.com/company/news/releases/uvac_winners.html`
6. Wood, P.: A libpcap Version which Supports MMAP Mode on Linux Kernels 2.[46].x., `http://public.lanl.gov/cpw/`