The sys admin's daily grind: tcpflow and HugeURL

Double Feature

First the fun, then the pleasure: This month, we look at a TCP that administrators have to take seriously, followed by some URL fun. By Charly Kühnast

f you need to smooth the kinks out of your network services, very likely you will take Tcpdump and Wireshark out of your toolbox. But honestly, both are slightly less than intuitive, and you need to be a genuine expert to interpret the results. Tcpflow [1], on the other hand, is infinitely more intelligible. Instead of presenting the re-

sults packet by packet, it collects them as a data stream - thus, the "flow" in the tool's name.

Tcpflow organizes all of your open connections in source-target pairs and gives you a contiguous summary of the data traffic, with no need to worry about sequence numbers and packets out of order. The typical command looks some-

thing like:

tcpflow -i eth0 -c -e

The -i eth0 parameter selects the interface on which to listen; -c directs data on the console.

If you leave out the -c, the tool writes the content of each flow to a separate file. The file name is based on the source IP, target IP, and port numbers. The -e parameter, which does not seem to be available on some Linux distributions, color-codes the client/ server connections on the console. In the style of tcpdump, I can enable filters that restrict the display to

specific sources, targets, or ports. The following line

tcpflow -i eth0 -c -e port 143

only lists the IMAP connections. The screenshot in Figure 1 shows my browser's HTTP connection; I happen to be accessing the German version of Linux Magazine online.

This example shows how tcpflow also reveals the content of a cookie sent to me by the website.

The Longer, the Better

About a year ago, I wrote about the URL abbreviator Yourls (Your Own URLs) [2]. Yourls uses Apache's Mod_rewrite function to create practical short URLs from long ones. But, you have to consider whether these fast-food URLs are detrimental to our web culture. Thanks to the self-healing properties of our global village, the counterrevolution is already in full swing; irate administrators are fighting back with URLs of epic dimensions. To generate them, surf to HugeURL.com [3], a service that converts a simple web address (www.linux-magazin.de) into a monster (Figure 1).

Elie Edit View Şearch Terminal Help 010.0000.000.115.37895-080.237.227.140.000800: GET / HTTP/1.1 Host: www.linux.magazin.de User.Agent: Mozilla/5.0 (X1); U; Linux i686; en-U5; rv:1.9.2.14) Gecko/20110221 Ubuntu/10.10 (maverick) Firefox/3.6.14 Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8 Accept: fencoding: gzip.deflate Accept-Encoding: gzip.deflate Accept: Anterset: ISO 889-1, ut-1.8;q=0.7,s;q=0 Connection: keep-alive Cookie: __utma_22a=7292745027.1307663422.127 Marzabalbard.201.012728/LINUME.20040041 Usc298/LINUME.20140041 Usc298/LINUME.201400401 Usc298/LINUME.2014004001 Usc298/LINUME.201400401 Usc298/LINUME.2014004001 Usc298/LINUME.2014004001 Usc298/LINUME.2014004001 Usc298/LINUME.2014004001 Usc298/LINUME.20140040001 Usc298/LINUME.201400400001 Usc298/LINUME.201400400001 Usc298/LINUME.2014004000001 080.237.227.140.00080-010.000.000.115.37805: Date: Mon, 07 Mar 2011 11:38:16 GMT Server: Apache/2.2.8 (Ubuntu) PHP/5.2.4-2ubu (V3.0.0 mod ssl/2.2.8 OpenSSL/0.9.8g X-Powered-By: e2 Publish Expires: Mon, 26 Jul 1997 03:00:00 GMT Cache-Control: no-cache, must-revalidate Pragma: no-cache Last-Modified: Mon, 07 Mar 2011 11:38:16 GMT Served-by: www.linux-magazin.de Content-language: de-DE Keep-Alive: timeout=15, max=98 timeout=15, max=98 Keep-Alive: timeout=15 Connection: Keep-Alive Transfer-Encoding: chunked Content-Type: text/html; charset=utf-8 d307 <!-- pagelayout start -->... Figure 1: Tcpflow collates

packet dumps from the client (top left) and the server (bottom left) to create flows. The right-hand image shows one huge URL from HugeURL.com.

INFO

- [1] tcpflow: http://www.circlemud.org/~jelson/software/tcpflow/
- [2] "Short Story" by Charly Kühnast, Linux Magazine March 2010, pg. 59: http://www.linux-magazine.com/lssues/2010/112/SHORT-STORY
- [3] HugeURL.com: http://hugeurl.com

AUTHOR

Charly Kühnast is a Unix operating system administrator at the Data Center in Moers, Germany. His tasks include firewall and DMZ security and availability. He divides his leisure time into hot, wet, and eastern sectors, where he enjoys cooking, freshwater aquariums, and learning Japanese, respectively.