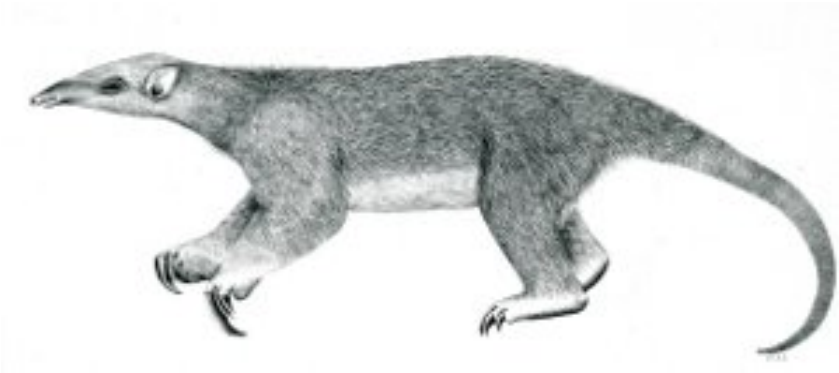
Tranalyzer2

txtSink

Text Output



Tranalyzer Development Team

Contents

1 txtSink

1.1 Description

The txtSink plugin provides human readable text output which can be saved in a file `PREFIX_flows.txt`, where `PREFIX` is provided via the `-w` option. The generated output contains a textual representation of all plugins results. Each line in the file represents one flow. The different output statistics of the plugins are separated by a tab character to provide better post-processing with command line scripts or statistical toolsets.

1.2 Dependencies

1.2.1 External Libraries

If gzip compression is activated (`GZ_COMPRESS=1`), then **zlib** must be installed.

Kali/Ubuntu: `sudo apt-get install zlib1g-dev`

Arch: `sudo pacman -S zlib`

Fedora/Red Hat: `sudo yum install zlib-devel`

Gentoo: `sudo emerge zlib`

OpenSUSE: `sudo zypper install zlib-devel`

Mac OS X: `brew install zlib`[1]

1.3 Configuration Flags

The configuration flags for the txtSink plugins are separated in two files.

1.3.1 txtSink.h

Name	Default	Description
`TFS_SPLIT`	1	Split the output file (Tranalyzer `-W` option)
`TFS_PRI_HDR`	1	Print a row with column names at the start of the flow file
`TFS_HDR_FILE`	1	Generate a separate header file (Section 1.4.1)
`TFS_PRI_HDR_FW`	0	Print header in every output fragment (Tranalyzer `-W` option)
`GZ_COMPRESS`	0	Compress the output (gzip)

The default suffix used for the flow file is `_flows.txt` and `_headers.txt` for the header file. Both suffix can be configured using `FLOWS_TXT_SUFFIX` and `HEADER_SUFFIX` respectively.

[1] Brew is a packet manager for Mac OS X that can be found here: https://brew.sh

1.3.2 bin2txt.h

bin2txt.h controls the conversion from internal binary format to standard text output.

Variable	Default	Description
HEX_CAPITAL	0	Hex number representation: 0: lower case, 1: upper case
IP4_NORMALIZE	0	IPv4 addresses representation: 0: normal, 1: normalized (padded with 0)
IP6_COMPRESS	1	IPv6 addresses representation: 1: compressed, 0: full 128 bit length
TFS_EXTENDED_HEADER	0	Whether or not to print an extended header in the flow file (number of rows, columns, columns type)
B2T_LOCALTIME	0	Time representation: 0: UTC, 1: localtime
B2T_TIME_IN_MICRO_SECS	1	Time precision: 0: nanosecs, 1: microsecs
HDR_CHR	"%"	start character of comments in flow file
SEP_CHR	"\t"	character to use to separate the columns in the flow file

1.4 Additional Output

1.4.1 Header File

The header file PREFIX_headers.txt describes the columns of the flow file and provides some additional information, such as plugins loaded and PCAP file or interface used, as depicted below. The default suffix used for the header file is _headers.txt. This suffix can be configured using HEADER_SUFFIX.

```
# Header file for flow file: PREFIX_flows.txt
# Generated from: /home/test/file.pcap
#
# 666;03.03.2016_19:04:55;hostname;Linux;4.2.0-30-generic;#36-Ubuntu SMP Fri Feb 26 00:58:07
    UTC 2016;x86_64
#
# Plugins loaded:
# 00: protoStats, version 0.6.0
# 01: basicFlow, version 0.6.0
# 02: macRecorder, version 0.6.0
# 03: portClassifier, version 0.5.8
# 04: basicStats, version 0.6.1
# 05: tcpFlags, version 0.6.0
# 06: tcpStates, version 0.5.8
# 07: icmpDecode, version 0.6.0
# 08: connectionCounter, version 0.6.0
# 09: txtSink, version 0.5.8
#
# Col No.   Type        Name
1           24:N        Flow direction
2           10:N        Flow Index
3           15:N        Flow Status
4           25:N        System time of first packet
5           25:N        System time of last packet
6           25:N        Flow duration
7           8:R         Ether VlanID
8           28:N        Source IPv4 address
9           15:N        Subnet number of source IPv4
10          8:N         Source port
11          28:N        Destination IP4 address
12          15:N        Subnet number of destination IP
13          8:N         Destination port
```

```
14          7:N         Layer 4 protocol
15          9:N         Number of distinct Source/Destination MAC addresses pairs
16          27_27_10:R  Source MAC address, destination MAC address, number of packets of MAC
    address combination
17          30_30:R Source MAC manufacturer, destination MAC manufacturer
...
```

The column number can be used, e.g., with `awk` to query a given column. For example, to extract all ICMP flows (layer 4 protocol equals 1) from a flow file:

```
awk -F'\t' '$14 == 1' PREFIX_flows.txt
```

The second column indicates the type of the column (see table below). If the value is repetitive, the type is postfixed with `:R`. Repetitive values can occur any number of times (from 0 to *N*). Each repetition is separated by a semicolon. The `'_'` indicates a compound, i.e., a value containing 2 or more subvalues.

#	Name	Description
1	I8	int8
2	I16	int16
3	I32	int32
4	I64	int64
5	I128	int128
6	I256	int256
7	U8	uint8
8	U16	uint16
9	U32	uint32
10	U64	uint64

#	Name	Description
11	U128	uint128
12	U256	uint256
13	H8	hex8
14	H16	hex16
15	H32	hex32
16	H64	hex64
17	H128	hex128
18	H256	hex256
19	F	float
20	D	double

#	Name	Description
21	LD	long double
22	C	char
23	S	string
24	C	flow direction[2]
25	TS	timestamp[3]
26	U64.U32	duration
27	MAC	mac address
29	IP4	IPv4 address
29	IP6	IPv6 address
30	IPX	IPv4 or 6 address
31	SC	string class[4]

[2] `A`: client→server, `B`: server→client
[3] U64.U32/S (See `B2T_TIMESTR` in `bin2txt.h`)
[4] string without quotes