



Incorporating dynamic list structures into YAF

**Software Engineering Institute/CERT
Network Situational Awareness**

**Dan Ruef - SEI
Emily Sarneso - SEI**



Agenda

IPFIX limitations

IPFIX extensions

List Structure Details

New in YAF

Mediators

yInspector

Limitations & Future Work

IPFIX Limitations

Fixed structured templates

- Templates contain a fixed set of information elements
- Unable to change elements depending on the data
- Unable to handle multiple occurrences of the same element
- Difficult to maintain relationships of hierarchical data
- Creating “single-use” templates is inefficient

Weak capabilities for lists

- Lists could be embedded in a variable length field
- Collector needs a priori knowledge to parse

New Requirements

Full Packet Capture

Maintain/Analyze Relationships

Security

Monitoring

Maintenance

Why IPFIX?

Template Mechanism

- As long as the Information Element is defined in the Information Model with a TLV {type, length, value}, it can be encoded

New IPFIX Capabilities

Basic List

- List of zero or more instances of an Information Element

Sub Template List

- List of zero or more instances of a structured data type defined by a template

Sub Template Multi List

- List of zero or more instances of a structured data type defined by different template definitions

Templates

Templates are sent before data is exported

When templates are defined, there is no concept of nested templates

- IPFIX Collector does not know what you intend to transport in lists

They are sent across the wire as equals

A template can contain a BL, STL, and/or STML

- Lists can be nested – necessary for maintaining relationships
- Some nested hierarchies are better than others
 - STL of 1 element = BL

Data Variability

The structure of the listed data is not chosen until the data is encoded and transmitted

How does this help?

- Data Specific Templates
- Variable Length Lists
- Model Hierarchical Relationships
- Nest Lists within Lists
- Multiple Occurrences of Data Types

YAF uses this flexibility to create data records that only contain elements it has data for

- Reduces null elements
- Relieves template management problem

New YAF Features

Deep Packet Inspection

SSL Certificate Capture

p0f

Tunneling Protocols

DNS

YAF Application Labeling & DPI

Application Labeling

- HTTP, SSH, SMTP, Gnutella, YMSG, DNS, FTP, SSL/TLS, SLP, IMAP, IRC, RTSP, SIP, RSYNC, PPTP, NNTP, TFTP, Teredo, POP3, DHCP, SMB, SNMP, AIM, SOCKS
- Compare flow's payload against configurable regular expressions and protocol decoding plug-ins
- Label 80 regex HTTP/d/.d/b

Deep Packet Inspection

- Based on Application Labeling
- If labeling succeeds, dive in further and pull out interesting strings

YAF IPFIX Templates

Before

0	1 - 15	16 - 31
Set ID = 2		Length = FFF
Template ID		Field Count
0	flowStartMilliseconds 152	Field Length = 8
0	flowEndMilliseconds 153	Field Length = 8
0	octetTotalCount 85	Field Length = 8
1	octetTotalCount 85	Field Length = 8
Reverse PEN		29305
0	packetTotalCount 86	Field Length = 8
1	packetTotalCount 86	Field Length = 8
Reverse PEN		29305
0	sourceIPv4Address 8	Field Length = 4
0	destinationIPv4Address 12	Field Length = 4
0	sourceTransportPort 7	Field Length = 2
0	destinationTransportPort 11	Field Length = 2
0	protocolIdentifier 4	Field Length = 1
0	flowEndReason 136	Field Length = 1
1	silkAppLabel 33	Field Length = 2
CERT PEN		6817
0	tcpSequenceNumber 184	Field Length = 4
1	tcpSequenceNumber 184	Field Length = 4
Reverse PEN		29305
1	initialTCPFlags 14	Field Length = 1
CERT PEN		6817
1	unionTCPFlags 15	Field Length = 1
CERT PEN		6817
1	reverseInitialTCPFlags 16398	Field Length = 1
CERT PEN		6817
1	reverseUnionTCPFlags 16399	Field Length = 1
CERT PEN		6817
0	vlanId 58	Field Length = 2
1	payload 18	Variable Length
CERT PEN		6817
1	reversePayload	Variable Length
CERT PEN		6817

After

0	1 - 15	16 - 31
Set ID = 2		Length = FFF
Template ID		Field Count
0	flowStartMilliseconds 152	Field Length = 8
0	flowEndMilliseconds 153	Field Length = 8
0	octetTotalCount 85	Field Length = 8
1	octetTotalCount 85	Field Length = 8
Reverse PEN		29305
0	packetTotalCount 86	Field Length = 8
1	packetTotalCount 86	Field Length = 8
Reverse PEN		29305
0	sourceIPv4Address	Field Length = 4
0	destinationIPv4Address 12	Field Length = 4
0	sourceTransportPort 7	Field Length = 2
0	destinationTransportPort 11	Field Length = 2
0	protocolIdentifier 4	Field Length = 1
0	flowEndReason 136	Field Length = 1
1	silkAppLabel 33	Field Length = 2
CERT PEN		6817
0	vlanId 58	Field Length = 2
0	subTemplateMultiList	Variable Length

0	1 - 15	16 - 31
Set ID = 2		Length = 12
Template ID		Field Count
0	tcpSequenceNumber 184	Field Length = 4
1	tcpSequenceNumber 184	Field Length = 4
Reverse PEN		29305
1	initialTCPFlags 14	Field Length = 1
CERT PEN		6817
1	unionTCPFlags 15	Field Length = 1
CERT PEN		6817
1	reverseInitialTCPFlags 16398	Field Length = 1
CERT PEN		6817
1	reverseUnionTCPFlags 16399	Field Length = 1
CERT PEN		6817

0	1 - 15	16 - 31
Set ID = 2		Length = FFF
Template ID		Field Count
1	payload 18	Variable Length
CERT PEN		6817
1	reversePayload	Variable Length
CERT PEN		6817

Fixbuf API

```
fbSubTemplateMultiList_t *stml = NULL;

fbSubTemplateMultiListInit(&(rec.subTemplateMultiList), 0, 2);

stml = fbSubTemplateMultiListGetNextEntry(&(rec.subTemplateMultiList), stml);

fbSubTemplateMultiListEntryInit(stml, YAF_TCP_FLOW_TID, tcpTemplate, 1);

/* Fill with data*/

stml = fbSubTemplateMultiListGetNextEntry(&(rec.subTemplateMultiList), stml);

fbSubTemplateMultiListEntryInit(stml, YAF_PAYLOAD_TID, payloadTemplate, 1);

/* Fill with data*/
```

STML is initialized
Get first entry in STML
Initialize entry
Fill with data
Get Next Entry
Initialize Entry
Fill with data
...

Protocol Specific Templates

YAF DNS Example

0	1 - 15	16 - 31
Set ID = 2		Length = 64
Template ID		Field Count
1	subTemplateList	Variable Length

YAF DNS Template

Resource Record Template

0	1 - 15	16 - 31
Set ID = 2		Length = FFF
Template ID		Field Count
0	subTemplateList	Variable Length
1	dnsTTL	Field Length = 4
CERT PEN		6817
1	dnsQueryType	Field Length = 2
CERT PEN		6817
1	dnsQueryResponse	Field Length = 1
CERT PEN		6817
1	dnsAuthoritative	Field Length = 1
CERT PEN		6817
1	dnsNXDomain	Field Length = 1
CERT PEN		6817
1	dnsRRSection	Field Length = 1
CERT PEN		6817
1	dnsQueryName	Variable Length
CERT PEN		6817

A Record

0	1 - 15	16 - 31
Set ID = 2		Length = 4
Template ID		Field Count
0	sourceIPv4Address	Field Length = 4

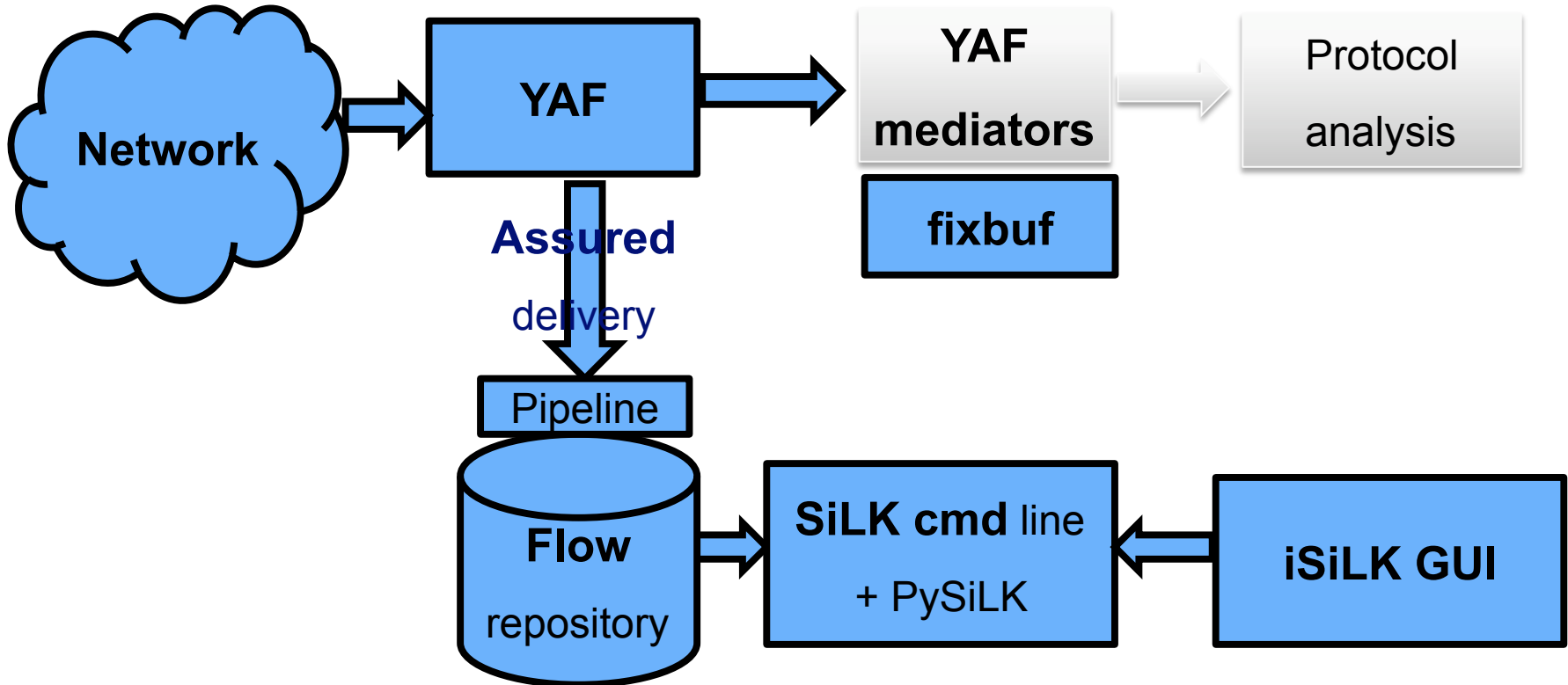
MX Record

0	1 - 15	16 - 31
Set ID = 2		Length = FFF
Template ID		Field Count
1	dnsMXPreference	Field Length = 2
CERT PEN		6817
1	dnsMXExchange	Variable Length
CERT PEN		6817

NS Record

0	1 - 15	16 - 31
Set ID = 2		Length = FFF
Template ID		Field Count
1	dnsNSdname	Variable Length
CERT PEN		6817

YAF Mediators



Spread Mediators

What is Spread?

- Spread is an open source toolkit that provides a publish/subscribe messaging service

Templates are managed per group

Messages can be multicast or sent to 1 or more subscribed groups

Collectors can subscribe to 1 or more groups

Spread groups can be leveraged to collect data specific records from YAF

YAF MySQL Mediator

a.k.a. yInspector

Listens for connections from YAF via the network

Parses Flow and DPI Data and inserts into a MySQL Database

A web front end was created to query the database

yInspector

yInspector
DPI - you know you want to look

Home Query Top 10

Select Options

- Source IP Address
- Source Port
- Flow Start Time
- Vlan
- Packet Count
- Octet Count
- flowEndReason
- Initial TCP Flags
- Destination IP Address
- Destination Port
- Flow End Time
- silkAppLabel
- Reverse Packet Count
- Reverse Octet Count
- Protocol
- Union TCP Flags

Where Options

Source IPv4 Address:

Destination IPv4 Address:

Source Port:

Destination Port:

Protocol: ALL TCP UDP

Vlan:

flowStartTime: 6/27/2010 0h

flowEndTime: 12/28/2010 0h

silkAppLabel:

Protocol Specific Options

User Agent:

HTTP Get:

HTTP Server String:

Protocol Specific Field Names

- FTP
- IMAP
- RTSP
- SIP
- SMTP
- SSH

Links

- SEI
- CERT
- NETSA
- NETSA TOOLS
- YAF

yInspector
DPI - you know you want to look

Home Query Top 10

Results Table

Double click any cell in the row to reveal all DPI and flow data for the flow

Query Results Total: 301 Records

	srcip4	srcport	dstport	packetTotalCount
+ 10.20.128.48	6610	80	4	
+ 10.45.14.186	9146	80	4	
+ 10.45.249.27	24429	80	4	
+ 10.168.5.220	31281	80	4	
+ 10.20.180.33	5553	80	4	
+ 10.20.5.200	5173	80	4	
+ 10.20.140.170	9444	80	4	
+ 10.20.171.171	32033	80	4	
+ 10.168.5.224	30471	80	4	
+ 10.168.5.252	47920	80	4	
+ 10.20.64.243	57123	80	4	
+ 10.45.112.60	27553	80	4	
+ 10.20.18.175	56973	80	4	
+ 10.20.92.203	24140	80	4	
+ 10.45.11.30	48821	80	4	
+ 10.20.60.229	37478	80	4	
+ 10.45.101.67	38210	80	4	
+ 10.168.5.252	47919	80	4	

Page 1 of 7 50 View 1 - 50 of 301

yInspector

yInspector
DPI - you know you want to look

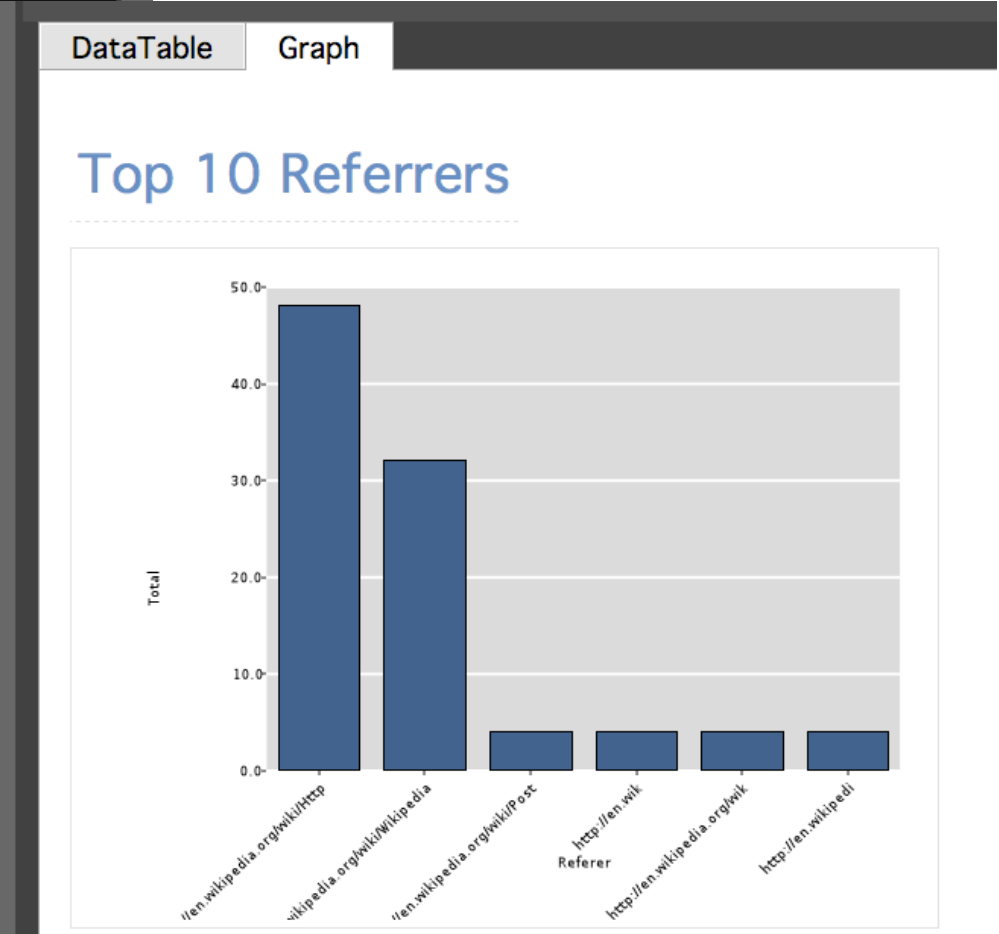
Home Query Top 10

DataTable Graph

Top 10 Referrers

Query Results Total: 6 Records

Referrer	Total
http://en.wikipedia.org/wiki/Http	48
http://en.wikipedia.org/wiki/Wikipedia	32
http://en.wikipedia.org/wiki/Post	4
http://en.wik	4
http://en.wikipedia.org/wik	4
http://en.wikipedi	4



Limitations

IPFIX Collectors still need to be aware of what is coming

Internal Templates are handled differently with lists

More responsibility on user to manage memory

Future Work

Deep Packet Inspection Enhancements

Machine Learning Capability for Protocol Recognition

Testing

Visualization Enhancements

Questions?

YAF available for download:

www.tools.netsa.cert.org

netsa-help@cert.org

Emily Sarneso

ecoff@cert.org

