Argomenti Trattati

Due progetti open source (GPL) per l'analisi del traffico di rete da file PCAP:

- Xplico: un Network Forensic Analysis Tool -NFAT- (2007)
  - Cosa consente di estrarre/visualizzare
  - Come funzionano i NFAT e i problemi che devo affrontare
- CapAnalysis: strumento per l'analisi macroscopica di file pcap (2012)

Xplico: scopo del progetto

Riuscire a estrarre (e visualizzare) i dati applicativi trasportati dalla rete.

- Pagine web (HTTP)
- Email (SMTP, POP, IMAP, “WebMail”)
- VoIP (SIP, RTP, MEGACO, H323, …)
- Chat
- ...

I Protocolli

I principali problemi che devono affrontare i NFAT sono legati ai protocolli, alla loro varietà e alle loro evoluzioni nel tempo.

Protocolli:

- Standard/Aperti
- Proprietari/Chiusi

Esempi fra tutti, e fra loro distinti, sono:

- Gmail
- WhatsApp

I NFAT e la cifratura delle comunicazioni

Attualmente circa 78% (05/2019 Let's Encrypt Stats) del traffico HTTP è cifrato (SSL/TLS).

L’informazione utile è estraibile (acessibile) solo in due modi:

- Disponendo delle chiavi di cifratura (privata o di sessione)
- Acquisendo i dati attraverso un attacco Man-In-The-Middle (con chiavi con firma valida/certificata)

Trasporto dell'Informazione

L'informazione da estrarre la si può trovare trasportata in rete in varie forme e modalità:

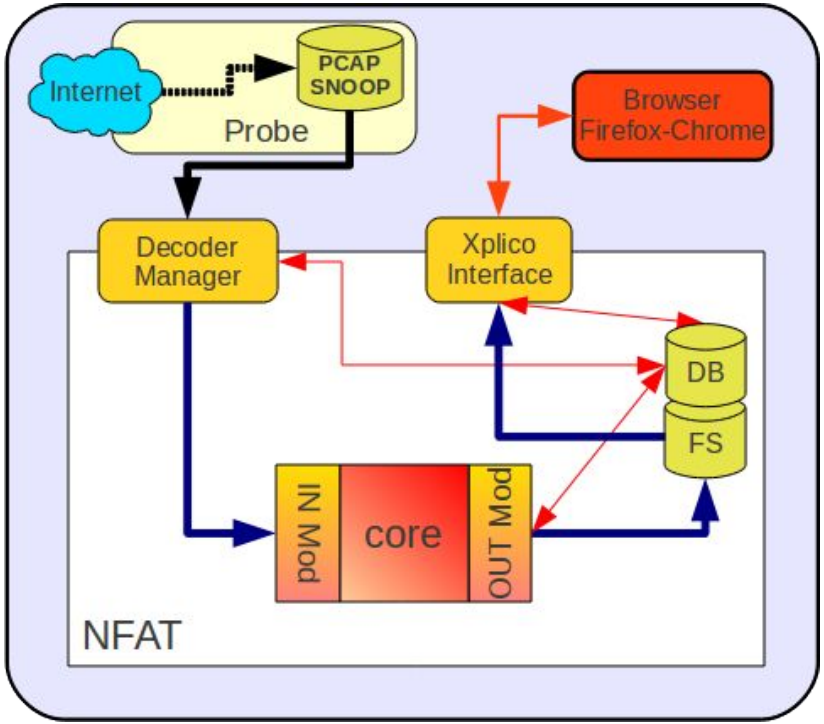
- HTTP, POP, IRC, ... : con un unico flusso TCP
- FTP, VoIP (SIP, RTP, H323, MEGACO, ...): più “flussi” correlati fra loro ma limitati in numero
- FB Web chat, ..: trasportati dal HTTP in più messaggi e su più flussi, in un arco temporale che possiamo però considerare limitato
- P2P: molti “flussi” distribuiti anche in un ampio arco temporale

In alcuni casi la sola estrazione del dato non è sufficiente, un esempio sono le pagine Web.

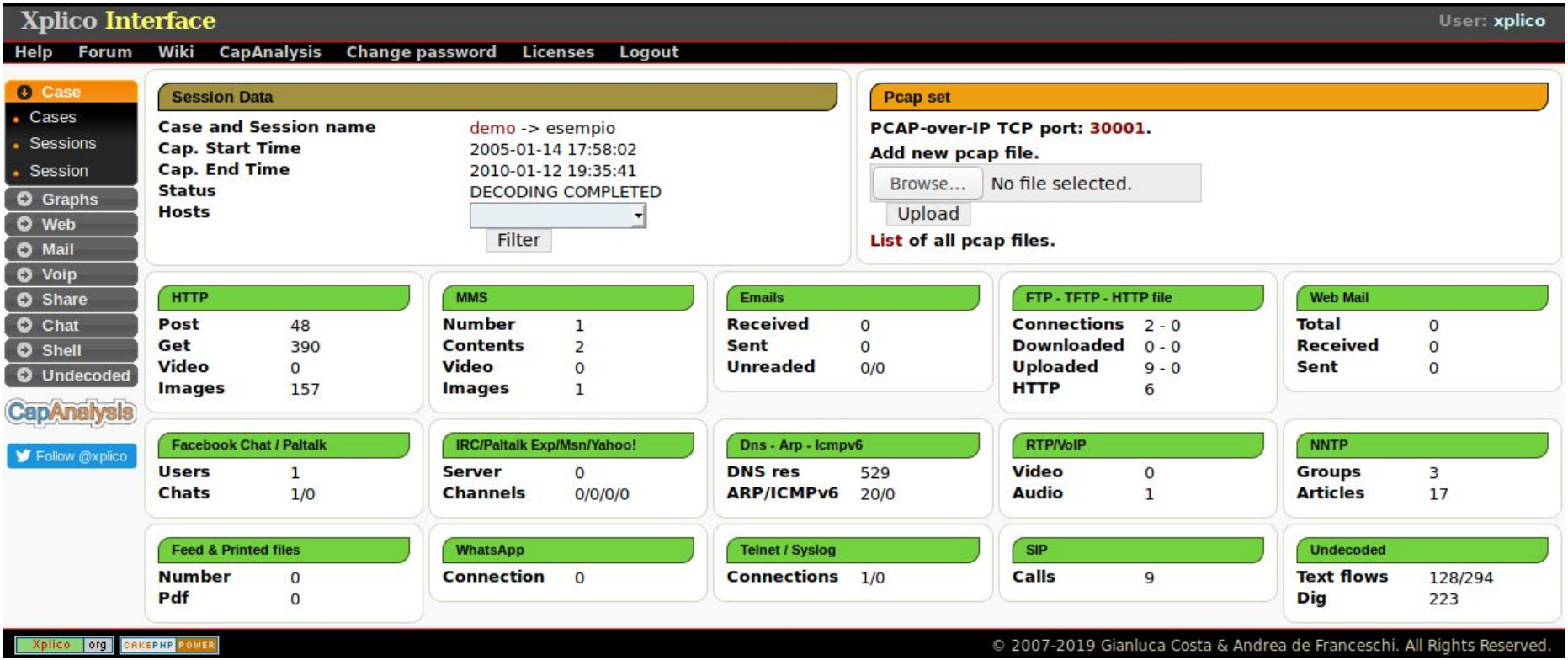
Architettura

Xplico come sistema di decodifica si compone di varie applicazioni:

- Un gestore delle decodifiche
- Xplico che ricostruisce il dato appl.
- Una serie di applicativi: aggregatori/manipolatori
- Un insieme di applicativi per la trascodifica (vlc, ffmpeg, …)
- Una interfaccia utente (Web User Interface)



Xplico


Xplico Interface
User: xplico
Help Forum Wiki CapAnalysis Change password Licenses Logout
Case
Cases
Sessions
Session
Graphs
Web
Mail
Voip
Share
Chat
Shell
Undecoded
CapAnalysis
Follow @xplico
Session Data
Case and Session name demo -> esempio
Cap. Start Time 2005-01-14 17:58:02
Cap. End Time 2010-01-12 19:35:41
Status DECODING COMPLETED
Hosts
Filter
Pcap set
PCAP-over-IP TCP port: 30001.
Add new pcap file.
Browse... No file selected.
Upload
List of all pcap files.
HTTP
Post 48
Get 390
Video 0
Images 157
MMS
Number 1
Contents 2
Video 0
Images 1
Emails
Received 0
Sent 0
Unreaded 0/0
FTP - TFTP - HTTP file
Connections 2 - 0
Downloaded 0 - 0
Uploaded 9 - 0
HTTP 6
Web Mail
Total 0
Received 0
Sent 0
Facebook Chat / Paltalk
Users 1
Chats 1/0
IRC/Paltalk Exp/Msn/Yahoo!
Server 0
Channels 0/0/0/0
Dns - Arp - Icmpv6
DNS res 529
ARP/ICMPv6 20/0
RTP/VoIP
Video 0
Audio 1
NNTP
Groups 3
Articles 17
Feed & Printed files
Number 0
Pdf 0
WhatsApp
Connection 0
Telnet / Syslog
Connections 1/0
SIP
Calls 9
Undecoded
Text flows 128/294
Dig 223
Xplico.org
CAKEPHP POWER
© 2007-2019 Gianluca Costa & Andrea de Franceschi. All Rights Reserved.

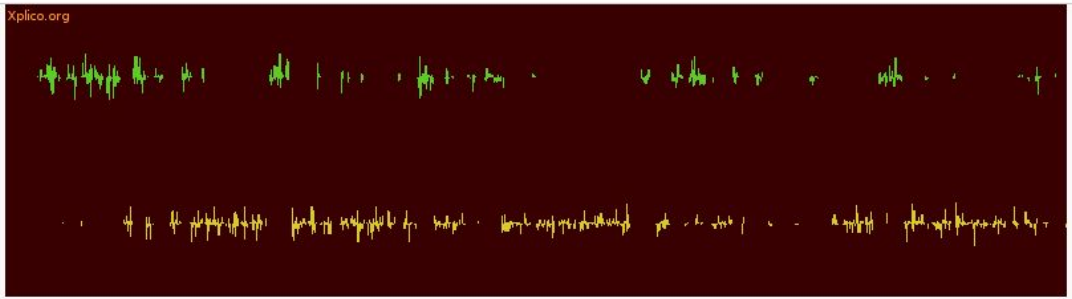
Presentazione contenuti HTTP

Date	Url	Size	Method	Info
2009-12-14 21:02:07	bs.serving-sys.com/BurstingPipe/adServer.bs?cn=sb&c=17&pli=1167587&PluID=0&w=160&h=600&ord=1260824525	986	GET	info.xml
2009-12-14 21:02:06	ad.yieldmanager.com/iframe3?VaUDAB93CQCE.jgAAAAAAN02DwAAAAAAAAApHAoAAAAAAAoAAQACF4owAgAAAAAAe	1280	GET	info.xml
2009-12-14 21:02:05	es.mg41.mail.yahoo.com/fc/fc?p=mail_candygram&bg=transparent&l=SKY&f=150557781&id=3&cbk=fcLoaded&tgt=_	3345	GET	info.xml
2009-12-14 21:02:01	es.mg41.mail.yahoo.com/dc/blank.html?bn=240.3&.intl=es&.lang=es-ES&nocache=1260824520977	460	GET	info.xml
2009-12-14 21:02:00	es.mg41.mail.yahoo.com/fc/fc?p=mail_candygram&bg=transparent&l=SKY&f=150557152&id=2&cbk=fcLoaded&tgt=_	3342	GET	info.xml
2009-12-14 21:01:19	es.mg41.mail.yahoo.com/dc/rs?log=LaunchSequenceCompleted;SuccessfulLaunch;PostLaunchSequenceCompleted&&.	110	GET	info.xml
2009-12-14 21:01:14	es.mg41.mail.yahoo.com/fc/fc?p=mail_candygram&bg=transparent&l=SKY&f=150557782&id=1&cbk=fcLoaded&tgt=_	3340	GET	info.xml
2009-12-14 21:01:10	es.mg41.mail.yahoo.com/dc/launch?.gx=0&.rand=1718390485&action=showLetter&umid=1_1545_AN9uUtQAAMlWSy	34378	GET	info.xml
2009-12-14 21:01:09	m.es.yahoo.com/_ylt=AtTaZM5Fg2s7Ze5.J5SBf4Bdoq5_/SIG=12n6h3s3e/**http%3A//mrd.mail.yahoo.com/msg%3Fmid=	201	GET	info.xml
2009-12-14 21:01:09	mrd.mail.yahoo.com/msg?mid=1_1545_AN9uUtQAAMlWSyaktwdveFKRZrA&fid=Inbox	88	GET	info.xml
2009-12-14 21:01:04	m.es.yahoo.com/sda2?intl=es&f=2142258060&p=yahoo&l=WDPA1&pos=wdpa1&bkt=bkt705&cbg=%235c758e&ctxt=	1800	GET	info.xml
2009-12-14 21:01:04	ad-emea.doubleclick.net/imp;v1;f;219952527;0-0;0;43387308;1|1;34415411|34433289|1;;cs=k%3fhttp://ad-emea.dou	0	GET	info.xml
2009-12-14 21:01:03	m.es.yahoo.com/_ylt=AoSvX4t7A8KmmlKh_Wd7E7tdoq5_/SIG=11t9fited/**https%3A//login.yahoo.com/config/mail%3F.	175	GET	info.xml
2009-12-14 21:00:50	ad-emea.doubleclick.net/imp;v1;f;219952527;0-0;0;43387308;1|1;34415411|34433289|1;;cs=k%3fhttp://ad-emea.dou	0	GET	info.xml
2009-12-14 21:00:49	ads.yimg.com/ev/eu/any/http://ads.yimg.com/ev/eu/any/esbckup.gif	91	GET	info.xml
2009-12-14 21:00:48	m.es.yahoo.com/	37832	GET	info.xml

Presentazione chiamate VoIP

Date	From	To	Duration	Info
2007-10-31 12:14:23	"FreeSwitch" <sip:5555551212@192.168.1.	<sip:6580@192.168.1.12>	0:0:0	info.xml
2007-10-31 12:14:23	"FreeSwitch" <sip:5555551212@192.168.1.	<sip:6580@192.168.1.12>	0:0:19	info.xml
2007-10-31 12:01:39	"FreeSwitch" <sip:5555551212@192.168.1.	<sip:6580@192.168.1.12>	0:0:18	info.xml
2007-10-31 12:01:38	"FreeSwitch" <sip:5555551212@192.168.1.	<sip:6580@192.168.1.12	0:0:0	info.xml
2005-01-14 17:58:27	"Ivan" <sip:Ivan@Verso.com>	"Ivan" <sip:Ivan@Verso.com>	0:0:0	info.xml
2005-01-14 17:58:27	"Ivan Alizade" <sip:5514540002@200.57.7.	"francisco@bestel.com" <sip:francisco@bes	0:0:0	info.xml
2005-01-14 17:58:20	"america" <sip:francisco@bestel.com>	"america" <sip:francisco@bestel.com>	0:0:0	info.xml
2005-01-14 17:58:07	"Ivan" <sip:Ivan@Verso.com>	"Ivan" <sip:Ivan@Verso.com>	0:0:0	info.xml
2005-01-14 17:58:02	<sip:200.57.7.195:55061	"francisco@bestel.com" <sip:francisco@bes	0:0:24	info.xml

Date:	2010-03-07 11:58:31	
From:	192.168.0.121	play
To:	62.94.199.34	play
Duration:	0:3:48	
Info	info.xml	



Dato estratto e flussi TCP/UDP

In Xplico ad ogni dato estratto è associato file XML che elenca quali flussi sono stati elaborati per estrarre l'informazione. Un esempio è il seguente e riferibile a una chiamata VoIP (SIP).

```
--- Decoding info: stream 0 ---
udp
    udp.srcport 5060
    udp.dstport 5061
ip
    ip.proto 17
    ip.src 200.57.7.195
    ip.dst 200.57.7.204
    ip.offset 14
eth
    eth.src 00:03:ba:94:63:3e
    eth.type 2048
pol
    pol.layer1 1
    pol.count 1
    pol.offset 24
    pol.file /opt/xplico/pol_1/sol_1/decode/xplico.org_sample_
    pol.sesid 1
    pol.polid 1
```

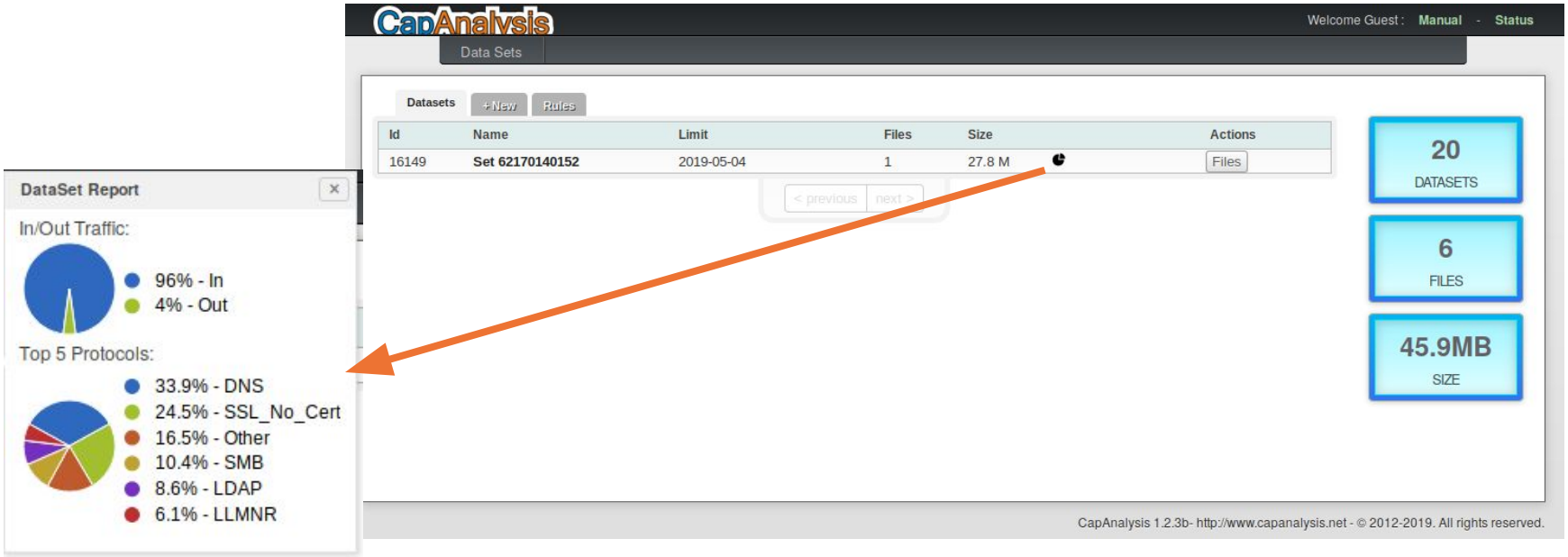
```
--- Decoding info: stream 1 ---
udp
    udp.srcport 8000
    udp.dstport 40376
ip
    ip.proto 17
    ip.src 200.57.7.204
    ip.dst 200.57.7.196
    ip.offset 14
eth
    eth.src 00:00:00:60:dd:19
    eth.type 2048
pol
    pol.layer1 1
    pol.count 499
    pol.offset 106426
    pol.file /opt/xplico/pol_1/sol_1/decode/xplico.org_sample_
    pol.sesid 1
```

CapAnalysis: scopo del progetto

- Visualizzare e classificare i flussi di un dataset (di file pcap) sulla base:
  - Tipo di protocollo (utilizzando nDPI di nTOP)
  - Numero di pacchetti inviati/ricevuti
  - Byte inviati/ricevuti
  - Porte di comunicazione utilizzate
  - Data-ora
  - Byte persi (per i flussi TCP) per la singola direzione
  - Durata del flusso
  - Collocazione geografica nella quale si trova il servizio remoto
- Poter filtrare i flussi in base:
  - All'IP o all'host name
  - Alla porta (TCP/UDP) o al protocollo
  - Alla collocazione geografica del Servizio
  - Alla dimensione dei dati scambiati

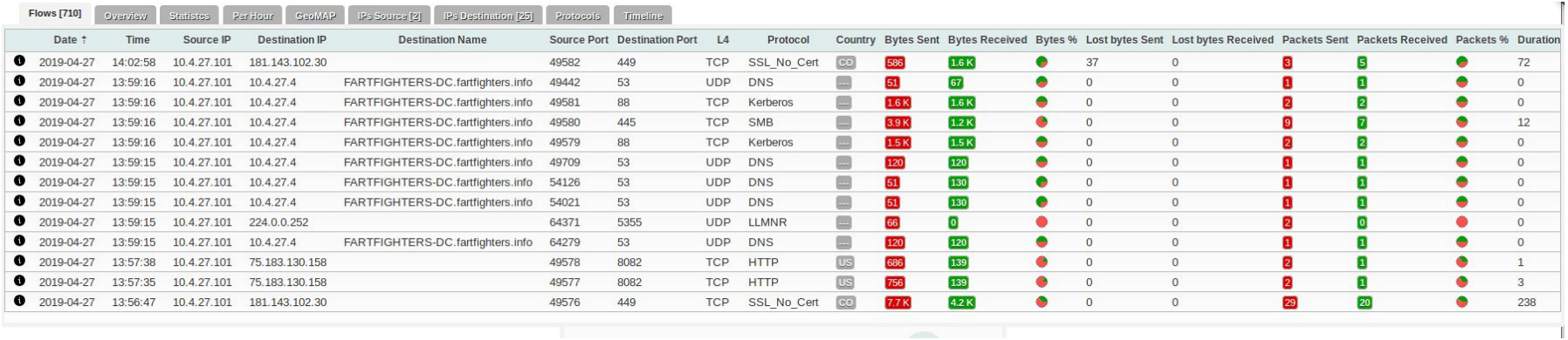
I Dataset

Un dataset è una collezione di file PCAP



Dati visualizzabili dal Dataset (1)

Elenco dei flussi di dati presenti nel Dataset



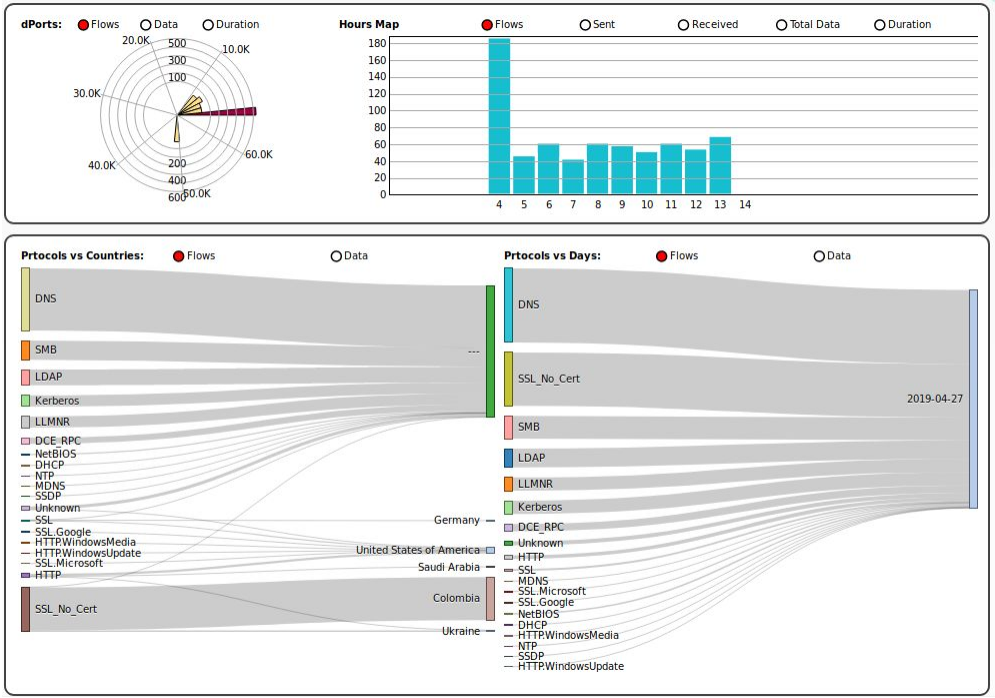
Flows [710] | Overview | Statistics | Per Hour | GeoMAP | IPs Source [2] | IPs Destination [25] | Protocols | Timeline

Date	Time	Source IP	Destination IP	Destination Name	Source Port	Destination Port	L4	Protocol	Country	Bytes Sent	Bytes Received	Bytes %	Lost bytes Sent	Lost bytes Received	Packets Sent	Packets Received	Packets %	Duration
2019-04-27	14:02:58	10.4.27.101	181.143.102.30		49582	449	TCP	SSL_No_Cert	CO	586	1.6 K		37	0	3	5		72
2019-04-27	13:59:16	10.4.27.101	10.4.27.4	FARTFIGHTERS-DC.fartfighters.info	49442	53	UDP	DNS	---	51	67		0	0	1	1		0
2019-04-27	13:59:16	10.4.27.101	10.4.27.4	FARTFIGHTERS-DC.fartfighters.info	49581	88	TCP	Kerberos	---	1.6 K	1.6 K		0	0	2	2		0
2019-04-27	13:59:16	10.4.27.101	10.4.27.4	FARTFIGHTERS-DC.fartfighters.info	49580	445	TCP	SMB	---	3.9 K	1.2 K		0	0	9	7		12
2019-04-27	13:59:16	10.4.27.101	10.4.27.4	FARTFIGHTERS-DC.fartfighters.info	49579	88	TCP	Kerberos	---	1.5 K	1.5 K		0	0	2	2		0
2019-04-27	13:59:15	10.4.27.101	10.4.27.4	FARTFIGHTERS-DC.fartfighters.info	49709	53	UDP	DNS	---	120	120		0	0	1	1		0
2019-04-27	13:59:15	10.4.27.101	10.4.27.4	FARTFIGHTERS-DC.fartfighters.info	54126	53	UDP	DNS	---	51	130		0	0	1	1		0
2019-04-27	13:59:15	10.4.27.101	10.4.27.4	FARTFIGHTERS-DC.fartfighters.info	54021	53	UDP	DNS	---	51	130		0	0	1	1		0
2019-04-27	13:59:15	10.4.27.101	224.0.0.252		64371	5355	UDP	LLMNR	---	66	0		0	0	2	0		0
2019-04-27	13:59:15	10.4.27.101	10.4.27.4	FARTFIGHTERS-DC.fartfighters.info	64279	53	UDP	DNS	---	120	120		0	0	1	1		0
2019-04-27	13:57:38	10.4.27.101	75.183.130.158		49578	8082	TCP	HTTP	US	686	139		0	0	2	1		1
2019-04-27	13:57:35	10.4.27.101	75.183.130.158		49577	8082	TCP	HTTP	US	756	139		0	0	2	1		3
2019-04-27	13:56:47	10.4.27.101	181.143.102.30		49576	449	TCP	SSL_No_Cert	CO	7.7 K	4.2 K		0	0	29	20		238

Per ogni flusso: Data, IP sorgente e destinazione, porte, protocollo applicativo (nDPI), byte inviati e ricevuti, byte persi, durata, ...

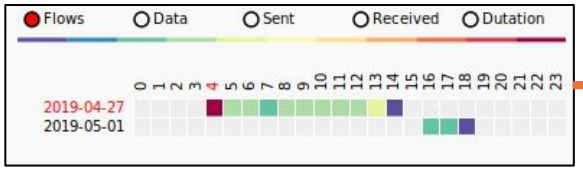
Dati visualizzabili dal Dataset (2)

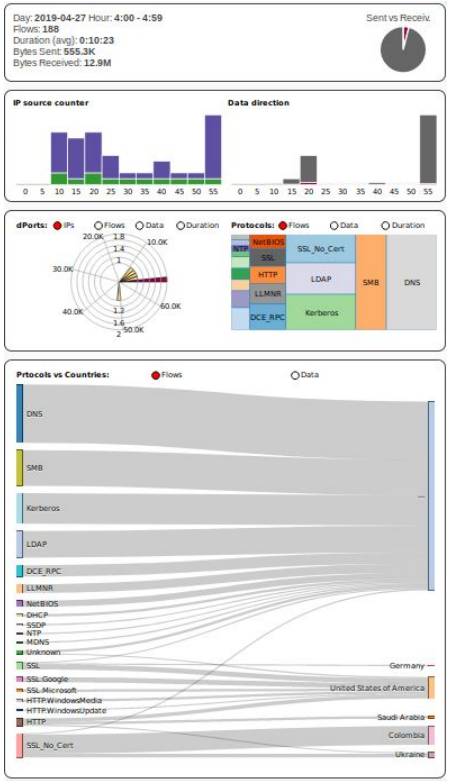
- Mappa delle porte destinazione utilizzate in base al loro uso o alla relativa mole di dati.
- Mappa temporale nelle 24h riferita al numero di flussi o alla quantità di dati scambiati.
- Mappa dei protocolli applicativi in riferimento all'area geografica o al loro volume (flussi o dati).



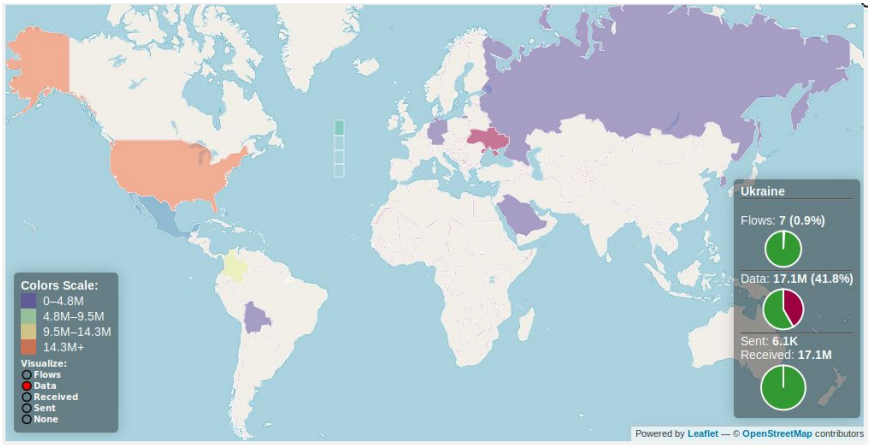
Dati visualizzabili dal Dataset (3)

Le informazioni sui dati aggregati sono consultabili In base al giorno e all'orario.



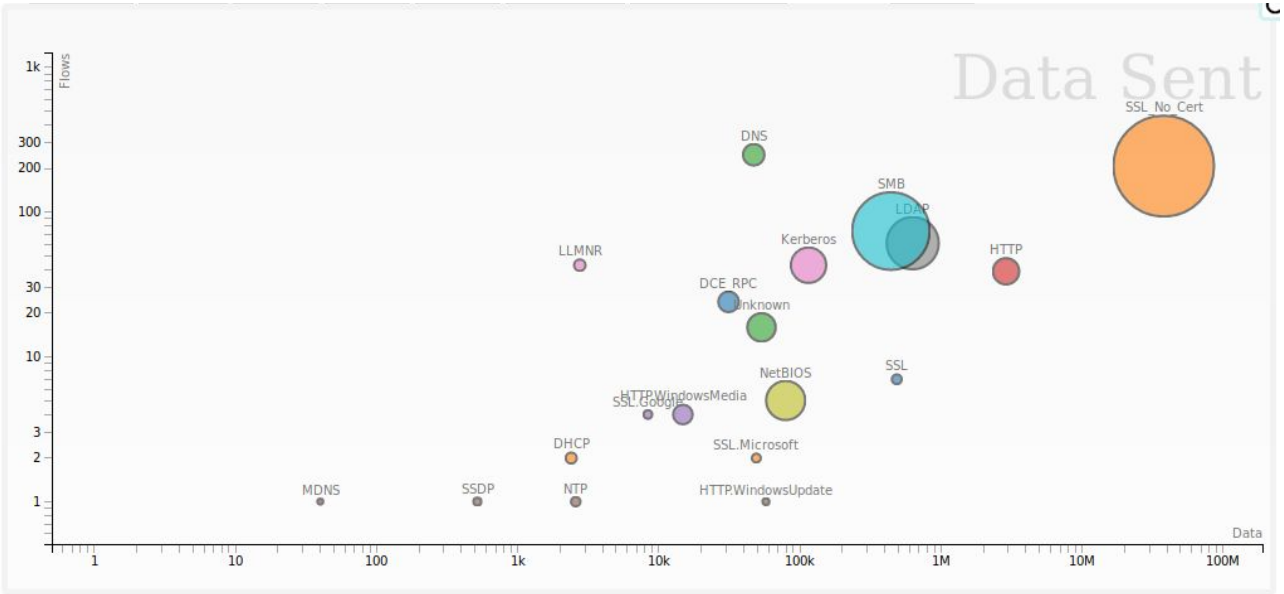


Mappa dei flussi o dei dati scambiati nelle varie aree geografiche.



Dati visualizzabili dal Dataset (4)

Grafico 3D che mette in relazione fra loro due due/tre tipi di dati riferibili al medesimo protocollo.



“Filtri” applicabili alla visualizzazione del Dataset

Alle informazioni descritte sono applicabili alcuni filtri il cui scopo è quello di permettere di individuare uno specifico sottoinsieme di flussi/dati all’interno del dataset.

I filtri disponibili consentono la selezione:

- di un sottoinsieme dei file PCAP del dataset
- dell’IP o della porta sorgente o dell’IP o della porta destinazione
- di uno o più protocolli applicativi
- delle aree geografiche
- dei volumi dei dati relativi al flusso (anche in base alla direzione)
- arco temporale (ad esempio una precisa ora del giorno)

Conclusioni

I NFAT sono strumenti complessi da realizzare e la loro complessità dipende dai protocolli e dalle informazioni che si desidera estrarre. Con la cifratura delle comunicazioni lo sviluppo di questi strumenti di analisi non si focalizza più unicamente all'estrazione dei dati applicativo, essendo disponibili, in tali condizioni, solamente metadati (informazioni molto parziali).

NetworkMiner: https://www.netresec.com/
Xplico: https://www.xplico.org/
PyFlag: http://www.pyflag.net/

Moloch: https://molo.ch/
Packetbeat: https://www.elastic.co