Logging & free software

2013. RMLL

Peter Czanik / BalaBit

About me

- Peter Czanik from Hungary
- Community manager at BalaBit: syslog-ng upstream
- BalaBit is an IT security company with HQ in Budapest, Hungary
- 150 employees
  - almost 80% can configure a Linux kernel :)

About me

- Peter Czanik from Hungary
- Community manager at BalaBit: syslog-ng upstream
- BalaBit is an IT security company with HQ in Budapest, Hungary
- 150 employees
  - almost 80% can configure a Linux kernel :)

Topics

- What is syslog? And syslog-ng?
- What to log and how to log?
- Free-form messages versus name-value pairs
- The new buzzword: journal
- Standardization efforts
- Name-value pairs at work

What is syslog?

- Logging: recording events
- Syslog:
  - Application: collecting events
  - Data: the actual log messages
  - Protocol: forwarding events
- History:
  - Originally developed as a logging tool for sendmail
- Format:
  - <38>2013-02-13T11:43:58 localhost sshd[1234]: Accepted password for root from 192.168.101.1 port 38420 ssh2

What is syslog-ng

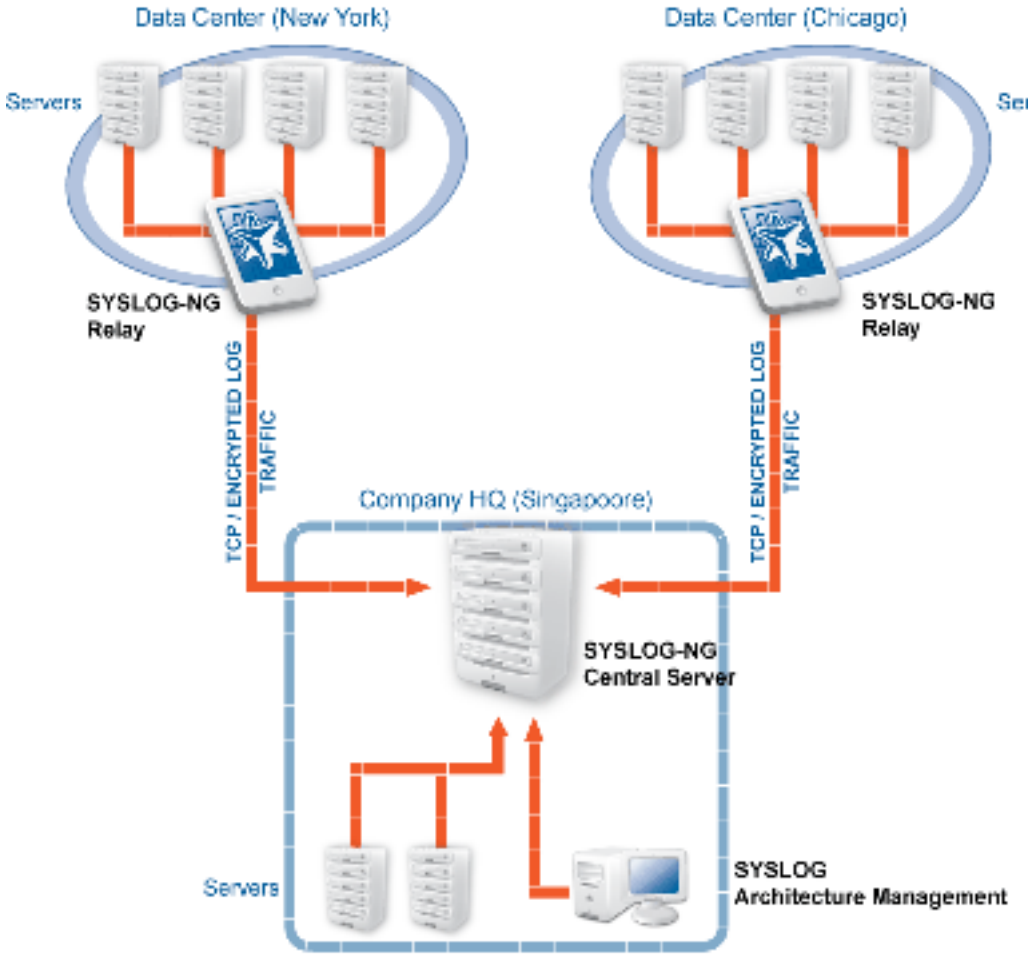
- syslog-ng: “Next Generation” syslog, since 1997
- Focus on central log collection
- “Swiss army knife” of logging
  - High performance
  - More input sources (files, programs, and so on)
  - More destinations (databases, encrypted net, and so on)
  - Better filtering (not only priority, facility)
  - Processing (rewrite, parse, correlate, and so on)
- OSE vs. PE
- rsyslog, logstash, graylog2, etc.

Why to log?

- It helps the everyday work of IT specialists:
  - Developers: debug logs
  - Admins/operators: system logs about health and usage
  - Security: investigation and incident response
- Only if logs are managed!

How to log?

- Short answer: centrally
- Long: centrally, because:
  - Ease of use: one place to check instead of many
  - Availability: even if the sender machine is down
  - Security: logs are available even if sender machine is compromised

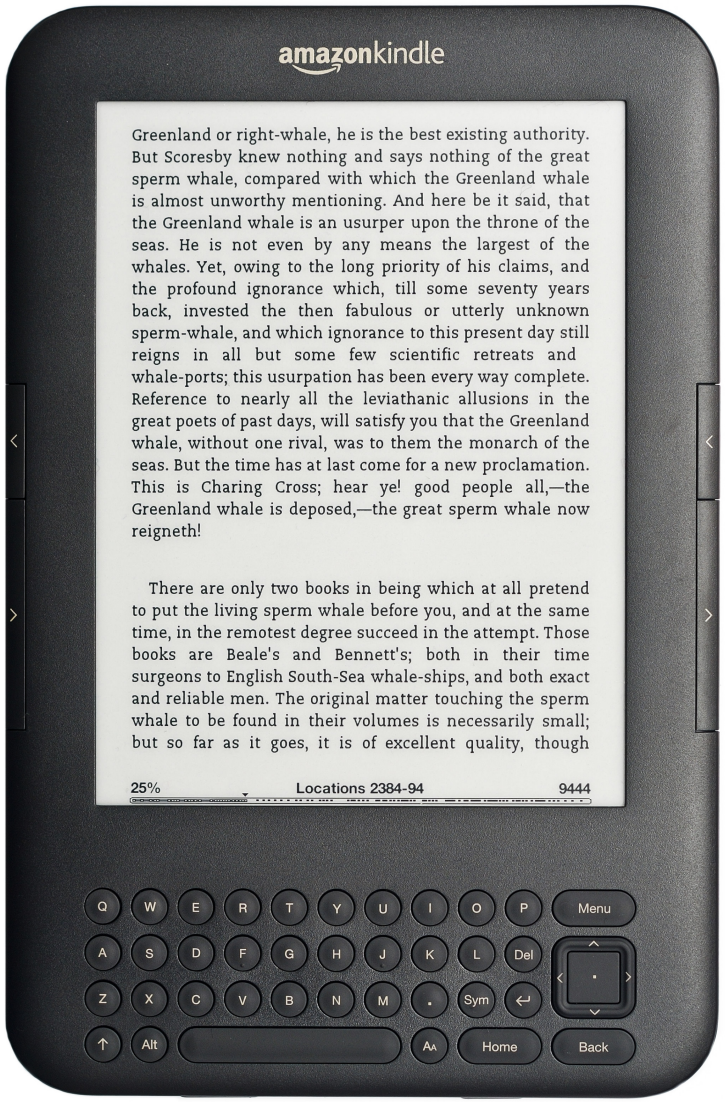


Which syslog-ng version is the most popular?

- Help:
  - Current version is v3.4 (5 months old)
  - SLES: v2.0
  - Gentoo: v3.2
  - Debian: v3.3
  - Fedora: v3.4
  - OpenSUSE: v3.4

Version 1.6

- V1.6
- :-)



Free-form log messages

- Most log messages are: date + hostname + text

Mar 11 13:37:56 linux-6965 sshd[4547]: Accepted keyboard-interactive/pam for root from 127.0.0.1 port 46048 ssh2

- Text = English sentence with some variable parts
- Easy to read by a human

Why does not it scale?

- Information is presented differently by each application
- Few logs (workstation) → easy to find information
- Many logs (server) → difficult to find information
- Difficult to process them with scripts

- Answer: structured logging
  - Events represented as name value pairs

PatternDB

- syslog-ng: uses name-value pairs internally
  - Date, facility, priority, program name, pid, etc.
- PatternDB message parser:
  - Can extract useful information into name-value pairs
  - Add status fields based on message text
  - Message classification (like LogCheck)
- Example: an ssh login failure:
  - user=root, action=login, status=failure
  - classified as “violation”

Related features

- Correlation
- JSON formatting and parser
- MongoDB destination
- AMQP destination

Journal

- The logging component of systemd
- Uses name-value pairs internally:
  - Message
  - Trusted properties
  - Any additional name-value pairs
- Native support for name-value pair storage
- Persistent log storage can be disabled
- Logs can be forwarded to syslog-ng through a socket
- syslog-ng can filter, process logs and forward them to central log server

Journal: the enemy?

- FAQ: Q: is journal the enemy? A: No!
- Journal is local only (syslog-ng: client – server)
- Journal does not filter or process log messages
- Journal is limited to Linux/systemd (syslog-ng: all Linux/BSD/UNIX)

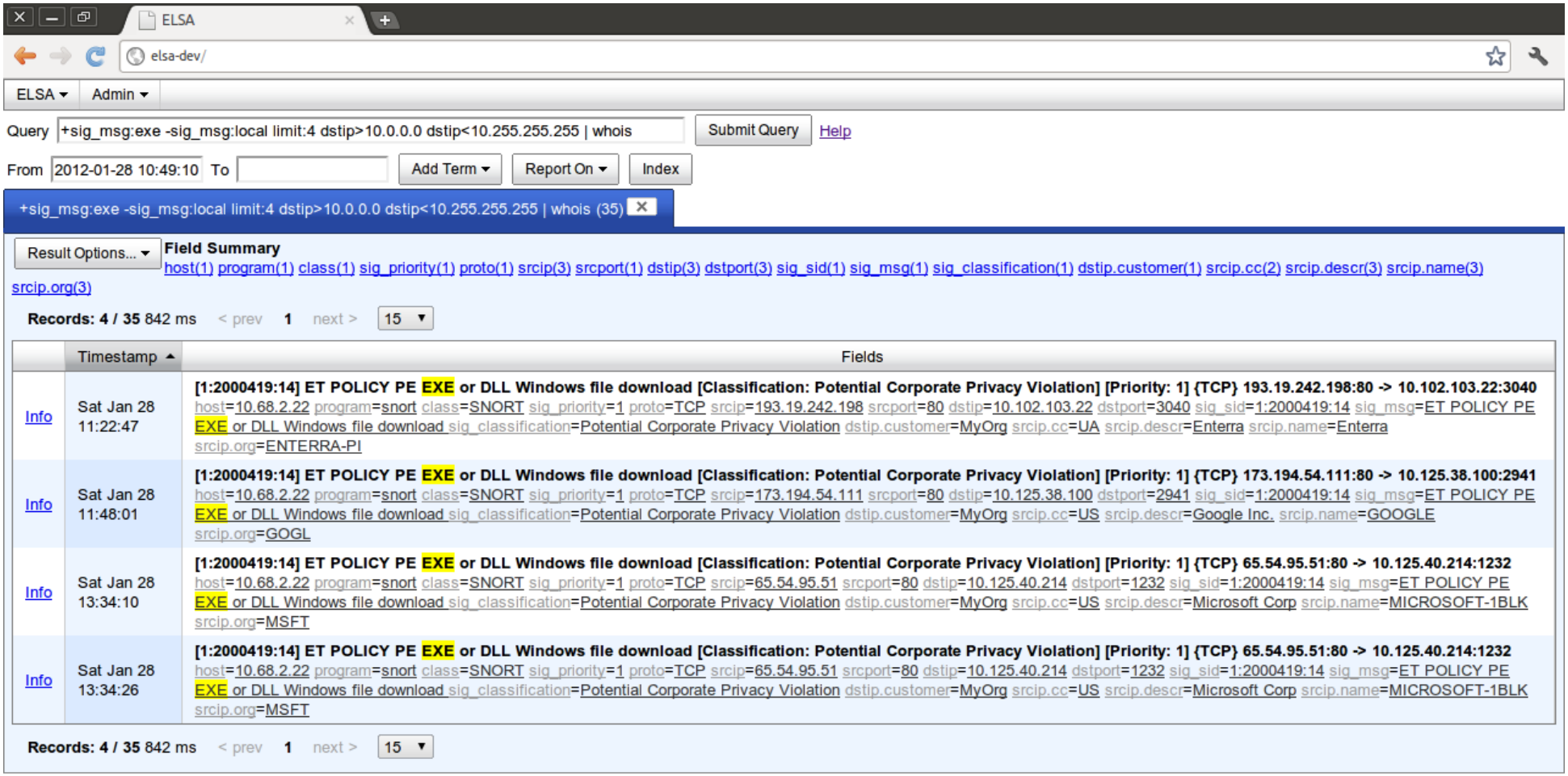
Common Event Expression (CEE)

- Journal, syslog-ng, Windows eventlog, rsyslog, auditd, and so on are based on name-value pairs
- All use different field names
- Standardization is a must: CEE → Common Event Expression
- Events: name-value pairs instead of free-form text
  - Taxonomy: name-value pairs to describe events (example: action, like “login”)
  - Dictionary: name-value pairs for event parameters (example: “user”)
- PatternDB can turn free-form messages into CEE
- “Holy Grail”, but too many legacy log messages

Name-value pairs in action: ELSA

- ELSA: Enterprise Log Search and Archive
- Based on syslog-ng, PatternDB and MySQL
- Simple and powerful web GUI
- Extreme scalability
- Patterns focused on network security:
  - Firewalls: Cisco, iptables
  - IDS: Snort, Suricata, Bro
  - HTTP, Windows logs, etc.

Search

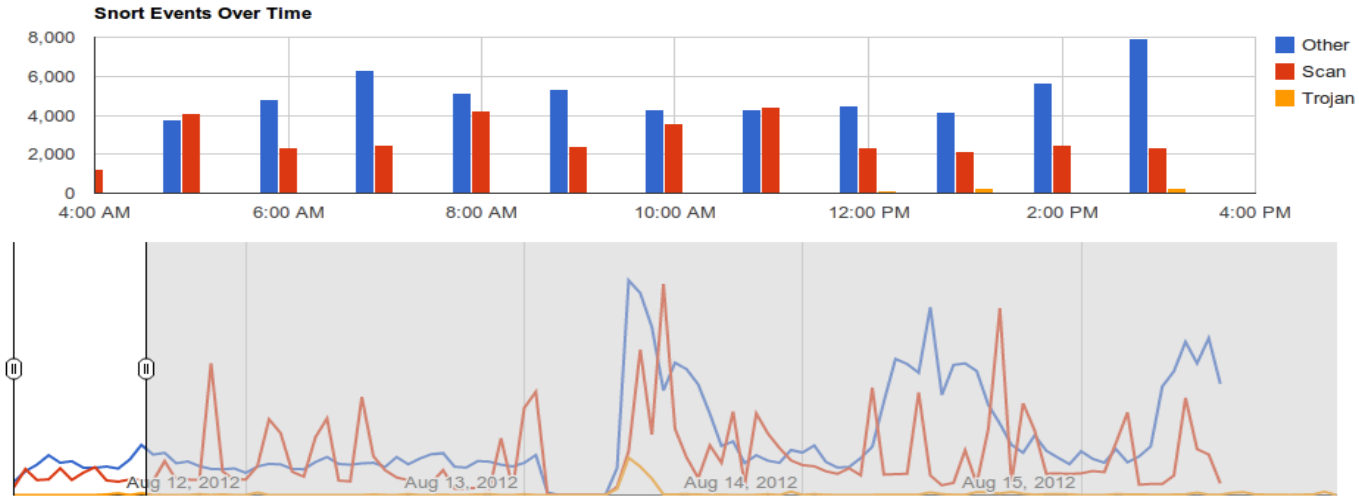

ELSA
elsa-dev/
ELSA
Admin
Query +sig_msg:exe -sig_msg:local limit:4 dstip>10.0.0.0 dstip<10.255.255.255 | whois
Submit Query
Help
From 2012-01-28 10:49:10 To
Add Term
Report On
Index
+sig_msg:exe -sig_msg:local limit:4 dstip>10.0.0.0 dstip<10.255.255.255 | whois (35)
Result Options...
Field Summary
host(1) program(1) class(1) sig_priority(1) proto(1) srcip(3) srcport(1) dstip(3) dstport(3) sig_sid(1) sig_msg(1) sig_classification(1) dstip.customer(1) srcip.cc(2) srcip.descr(3) srcip.name(3) srcip.org(3)
Records: 4 / 35 842 ms
< prev
1
next >
15
Timestamp
Fields
Info
Sat Jan 28 11:22:47
[1:2000419:14] ET POLICY PE EXE or DLL Windows file download [Classification: Potential Corporate Privacy Violation] [Priority: 1] {TCP} 193.19.242.198:80 -> 10.102.103.22:3040
host=10.68.2.22 program=snort class=SNORT sig_priority=1 proto=TCP srcip=193.19.242.198 srcport=80 dstip=10.102.103.22 dstport=3040 sig_sid=1:2000419:14 sig_msg=ET POLICY PE EXE or DLL Windows file download sig_classification=Potential Corporate Privacy Violation dstip.customer=MyOrg srcip.cc=UA srcip.descr=Enterra srcip.name=Enterra srcip.org=ENTERRA-PI
Info
Sat Jan 28 11:48:01
[1:2000419:14] ET POLICY PE EXE or DLL Windows file download [Classification: Potential Corporate Privacy Violation] [Priority: 1] {TCP} 173.194.54.111:80 -> 10.125.38.100:2941
host=10.68.2.22 program=snort class=SNORT sig_priority=1 proto=TCP srcip=173.194.54.111 srcport=80 dstip=10.125.38.100 dstport=2941 sig_sid=1:2000419:14 sig_msg=ET POLICY PE EXE or DLL Windows file download sig_classification=Potential Corporate Privacy Violation dstip.customer=MyOrg srcip.cc=US srcip.descr=Google Inc. srcip.name=GOOGLE srcip.org=GOGL
Info
Sat Jan 28 13:34:10
[1:2000419:14] ET POLICY PE EXE or DLL Windows file download [Classification: Potential Corporate Privacy Violation] [Priority: 1] {TCP} 65.54.95.51:80 -> 10.125.40.214:1232
host=10.68.2.22 program=snort class=SNORT sig_priority=1 proto=TCP srcip=65.54.95.51 srcport=80 dstip=10.125.40.214 dstport=1232 sig_sid=1:2000419:14 sig_msg=ET POLICY PE EXE or DLL Windows file download sig_classification=Potential Corporate Privacy Violation dstip.customer=MyOrg srcip.cc=US srcip.descr=Microsoft Corp srcip.name=MICROSOFT-1BLK srcip.org=MSFT
Info
Sat Jan 28 13:34:26
[1:2000419:14] ET POLICY PE EXE or DLL Windows file download [Classification: Potential Corporate Privacy Violation] [Priority: 1] {TCP} 65.54.95.51:80 -> 10.125.40.214:1232
host=10.68.2.22 program=snort class=SNORT sig_priority=1 proto=TCP srcip=65.54.95.51 srcport=80 dstip=10.125.40.214 dstport=1232 sig_sid=1:2000419:14 sig_msg=ET POLICY PE EXE or DLL Windows file download sig_classification=Potential Corporate Privacy Violation dstip.customer=MyOrg srcip.cc=US srcip.descr=Microsoft Corp srcip.name=MICROSOFT-1BLK srcip.org=MSFT
Records: 4 / 35 842 ms
< prev
1
next >
15

Graph

Bro IDS | ELSA | Snort

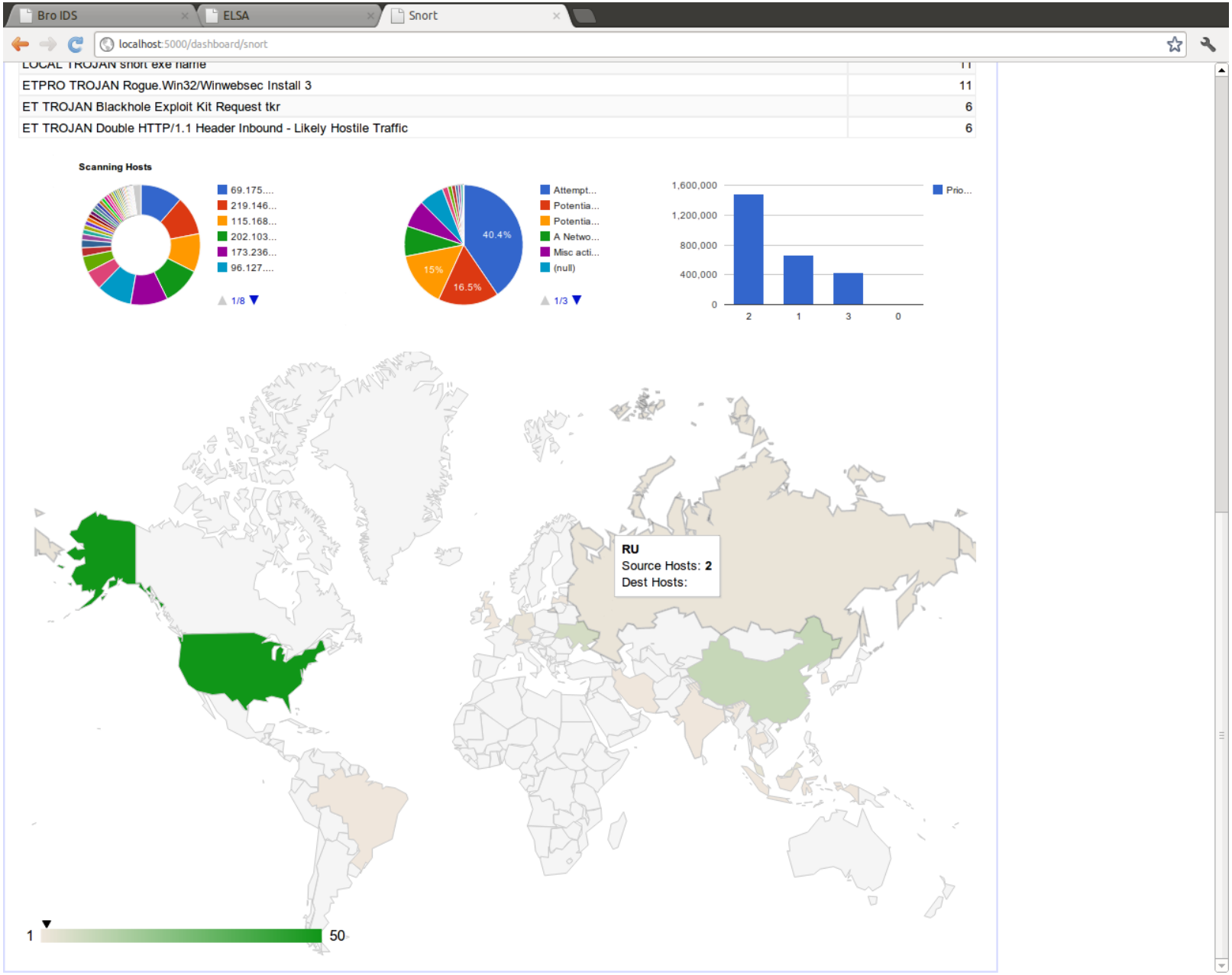
localhost:5000/dashboard/snort

Snort



snort.sig_msg	Count
ET TROJAN MS Terminal Server User A Login, possible Morto inbound	26680
ET TROJAN UPX compressed file download possible malware	859
ET TROJAN Suspicious User-Agent (C slash)	497
ET TROJAN ZeroAccess Outbound udp traffic detected	486
ET TROJAN Downloader User-Agent HTTPGET	195
ETPRO TROJAN Suspicious User-Agent (dnf)	170
ET TROJAN Trojan.Win32.Jorik.Totem.vg HTTP request	128
ET TROJAN Suspicious User-Agent (MSIE)	116
ET TROJAN IRC Nick change on non-standard port	114
ET TROJAN Potential DNS Command and Control via TXT queries	104
ET TROJAN Generic - 8Char.JAR Naming Algorithm	95
ET TROJAN Suspicious User-Agent (Presto)	68
ET TROJAN Clickfraud List Delivered To Client	63
ETPRO TROJAN Simda.C Checkin	31
ET TROJAN Java Archive sent when remote host claims to send an image	24
ET TROJAN Karagany/Kazy Obfuscated Payload Download	19
ET TROJAN Medfos/Midhos Checkin	15
LOCAL TROJAN short exe name	11
ETPRO TROJAN Rogue Win32/Winwebsec Install 3	11

Map


Bro IDS
ELSA
Snort
localhost:5000/dashboard/snort
ETPRO TROJAN Rogue.Win32/Winwebsec Install 3 11
ET TROJAN Blackhole Exploit Kit Request tkr 6
ET TROJAN Double HTTP/1.1 Header Inbound - Likely Hostile Traffic 6
Scanning Hosts
69.175....
219.146...
115.168...
202.103...
173.236...
96.127....
1/8
40.4%
16.5%
15%
Attempt...
Potentia...
Potentia...
A Netwo...
Misc acti...
(null)
1/3
1,600,000
1,200,000
800,000
400,000
0
2
1
3
0
Prio...
RU
Source Hosts: 2
Dest Hosts:
1
50

So, why syslog-ng?

- 15 years of open source development
- High performance log management
- Flexible configuration
- Excellent documentation
- PatternDB message parsing

Questions? (and some answers)

- Questions?

- Some useful syslog-ng resources:
  - Syslog-ng: http://www.balabit.com/network-security/syslog-ng
  - Many books at http://oreilly.com/
  - ELSA (log analysis based on syslog-ng's patterndb): http://code.google.com/p/enterprise-log-search-and-archive/
  - Mailing list: https://lists.balabit.hu/pipermail/syslog-ng/
  - My blog: http://czanik.blogs.balabit.com/
  - My e-mail: czanik@balabit.hu

End

www.balabit.com