BRO @ LBL

Aashish Sharma and Vincent Stoffer
Lawrence Berkeley National Lab

Bro Exchange Aug 7-8, 2012, NCAR, CO.

Outline

1. Overview
2. Tapping
3. Catch and release
4. Software.bro
5. Finding spear phishing
6. Input Framework
7. Deep Bro
8. IPv6
9. Time Machine
10. Syslog2bro
11. Fast searching

Quick overview

Lawrence Berkeley National Laboratory

- Located in Berkeley, CA
- DoE research facility run by UC
- Focus on scientific research in many fields
- 13 Nobel prize winners

Computing overview

- ~5000 users ~10,000 hosts
- Distributed computing resources
- Many guests and visitors
- IT security responsible for everything
- Aashish & Vincent

History with Bro

- Long history with the Berkeley Lab
  - Using Bro since 1995
  - Vern Paxson created Bro at LBL
  - Partner with ICSI and Bro developers
- Many Bros throughout our network for different tasks
- Monitoring at 10G for many years
- Multiple clusters running
- Bro is our primary tool for active response, monitoring, alerting, forensics, incident response and everything else!

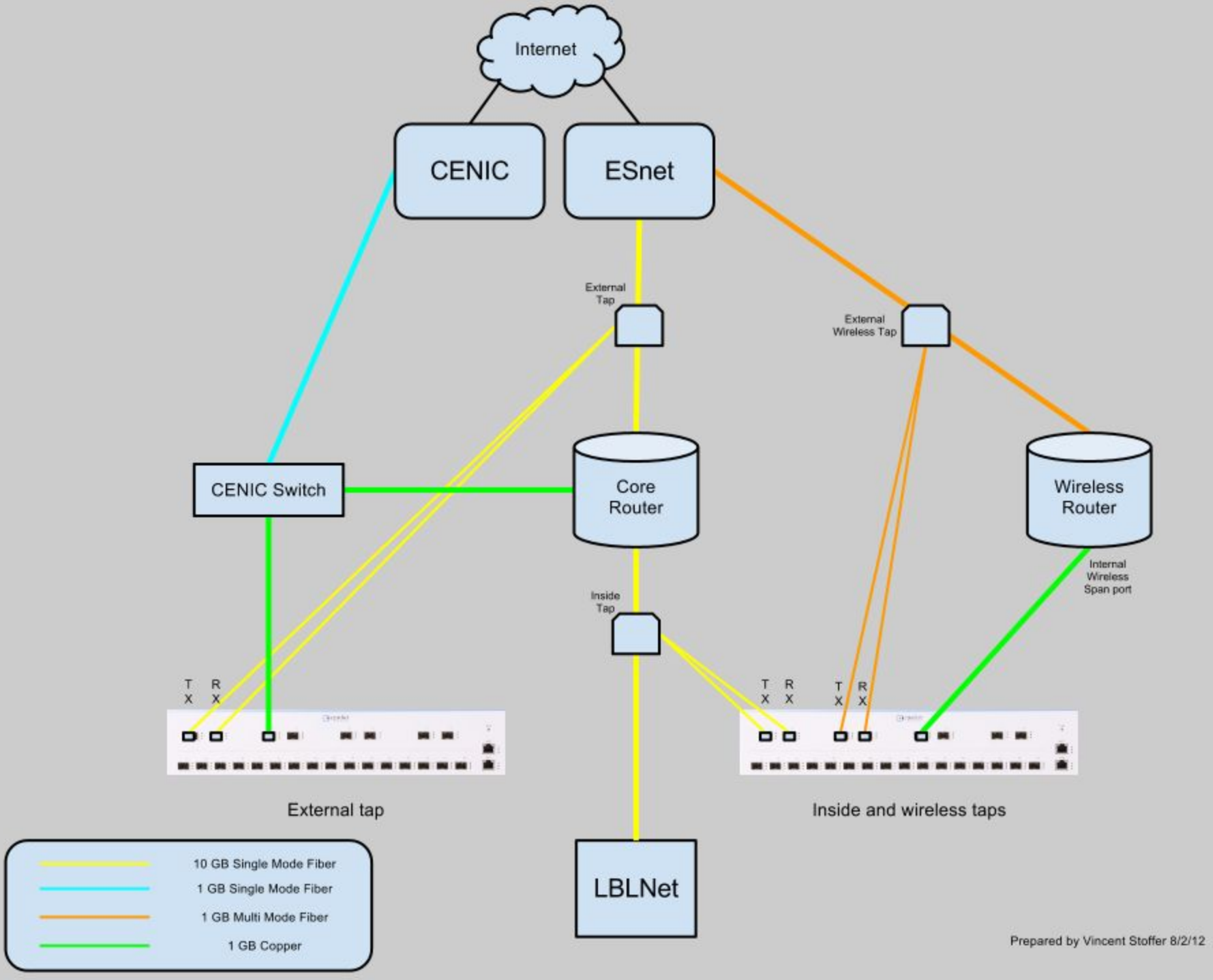
Tapping infrastructure

Need to tap 10G at our border and distribute to many bros:

- Had been using older Apcon 10G taps
- My first project at the lab was to update the tapping infrastructure
- Moved to cPacket cVu 240g devices which can handle aggregation, load-balancing and distribution
- Very flexible I/O routing but expensive

Tapping continued

- Aggregation of provider circuits
- Simple distribution of external tap to bros for many purposes
- Internal and wireless collected on another cVu and distributed
- Clusters are fed a single 10G aggregated link
  - another cPacket device does MAC rewriting for worker load balancing
  - Feeds each worker node 1G


Internet
CENIC
ESnet
External Tap
External Wireless Tap
CENIC Switch
Core Router
Wireless Router
Internal Wireless Span port
Inside Tap
T X
R X
T X
R X
T X
R X
External tap
Inside and wireless taps
LBLNet
10 GB Single Mode Fiber
1 GB Single Mode Fiber
1 GB Multi Mode Fiber
1 GB Copper
Prepared by Vincent Stoffer 8/2/12

Tapping part 3

- Goal to collect and aggregate as many different data sources as possible:
  - SMTP
  - DNS
  - "Deep Bro" internal network tapping
  - SYSLOG - Syslog2bro
- Always a need for more taps and aggregation. Being able to aggregate multiple data sources to the same Bro is very powerful

Tapping part 4

- Scaling well so far, drop rate is low on clusters and time machines, but...
- We are moving towards 100G at the border and pushing enough traffic that aggregating 10G in/out is starting to over-subscribe our 10G distribution links
- Working with vendors for 100G tapping/distribution solutions and considering our options for how to configure the cluster

Active response

- Dynamic firewall - Bro border protection
- Scanning policies
- Blocking via several mechanisms:
  - IP blocking - router ACLs and blocking appliance
  - Null zero routing - when blocking isn't enough
  - DHCP jailing - internal hosts
    - Malware infected
    - Out of compliance with software versions
    - Incident response
- Whitelisting and blacklisting capabilities

1: Catch-n-Release

- Deployed with Scan detection
- Mechanism to use limited amount of ACL's (or null zero routes) in an effective manner
- Once a system is released after a certain duration of time, a hair trigger response for the next scanning activity.
- This is what makes our dynamic firewall

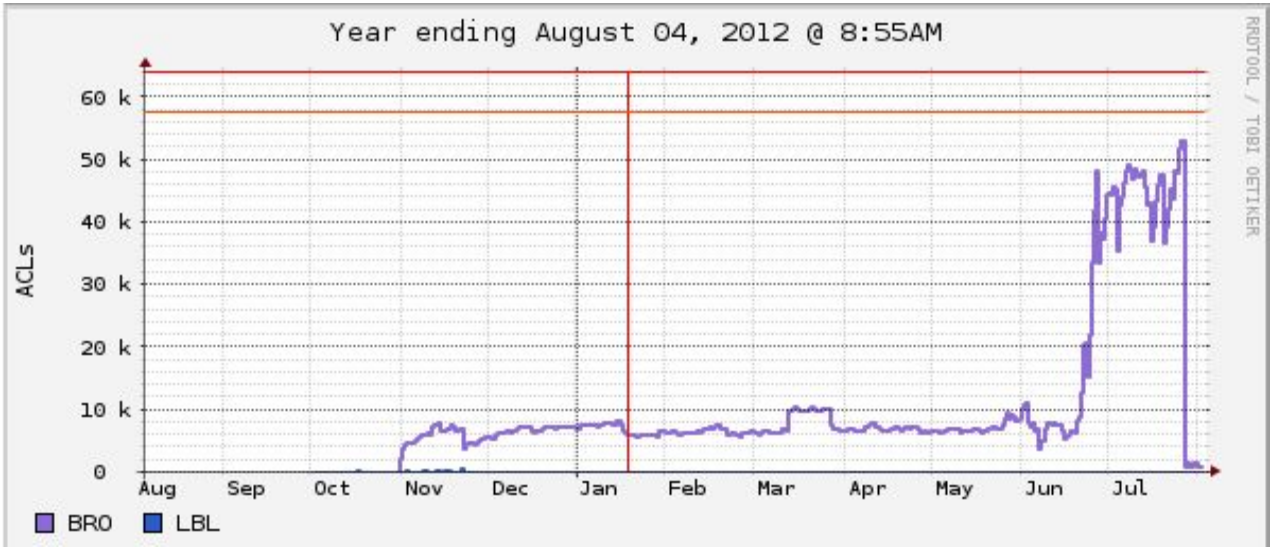
A side effect of catch-n-release - Discovering new unknown worms and botnets

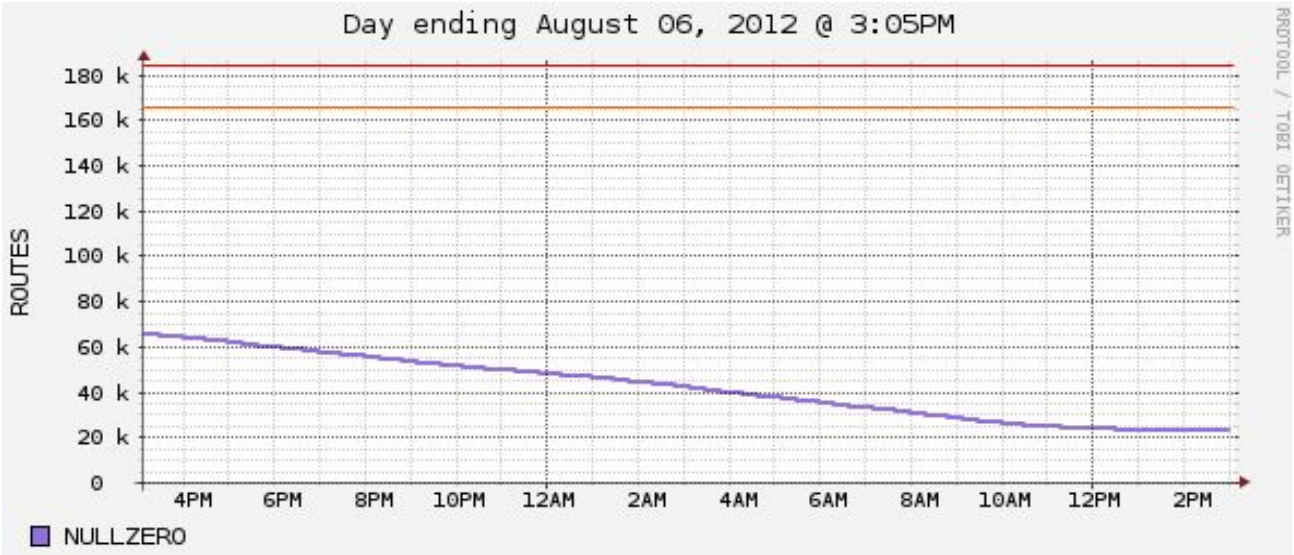
- Morto (Jun 2011) Windows RDP worm
  - Scanning activity is still ongoing at same rate - Morto-II
- CVE-2012-2122-mysql-authentication-bypass
- alien-worm (Jun 2012)
  - Worm made up solely of Video Cameras and DVR's
  - An embedded linux kernel possibly from a single OEM manufacturer
  - Popular cameras - buy on amazon for $60-$900
  - Current conservative estimates are at about 4M infected devices.
  - So what - Big deal?

Some examples of confirm compromises

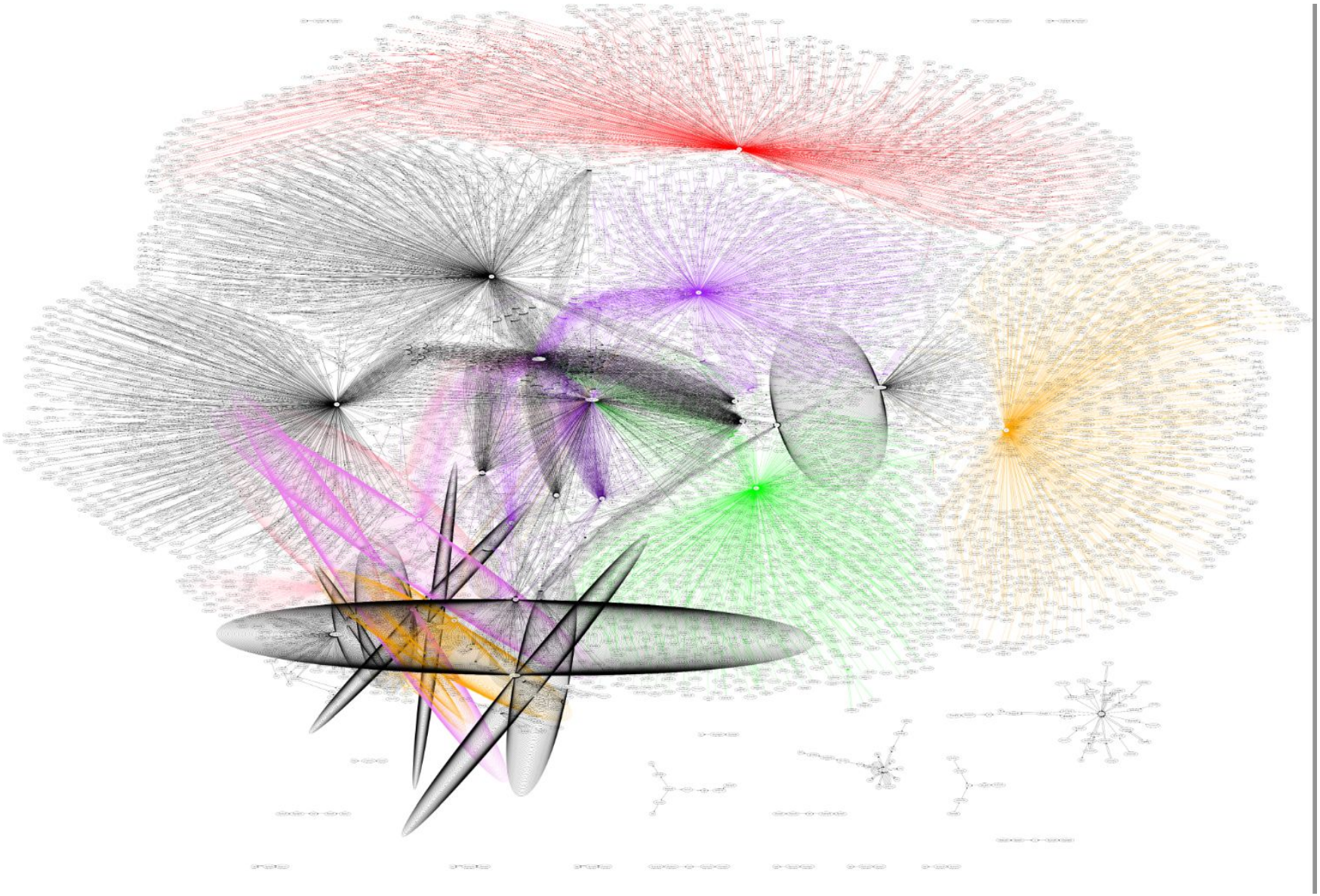
Embedded linux kernel versions: 3.x, 2.6.x, 2.4.x-2.6.x, 2.4.x, 2.2.x-3.x

Spikes in ACL counts

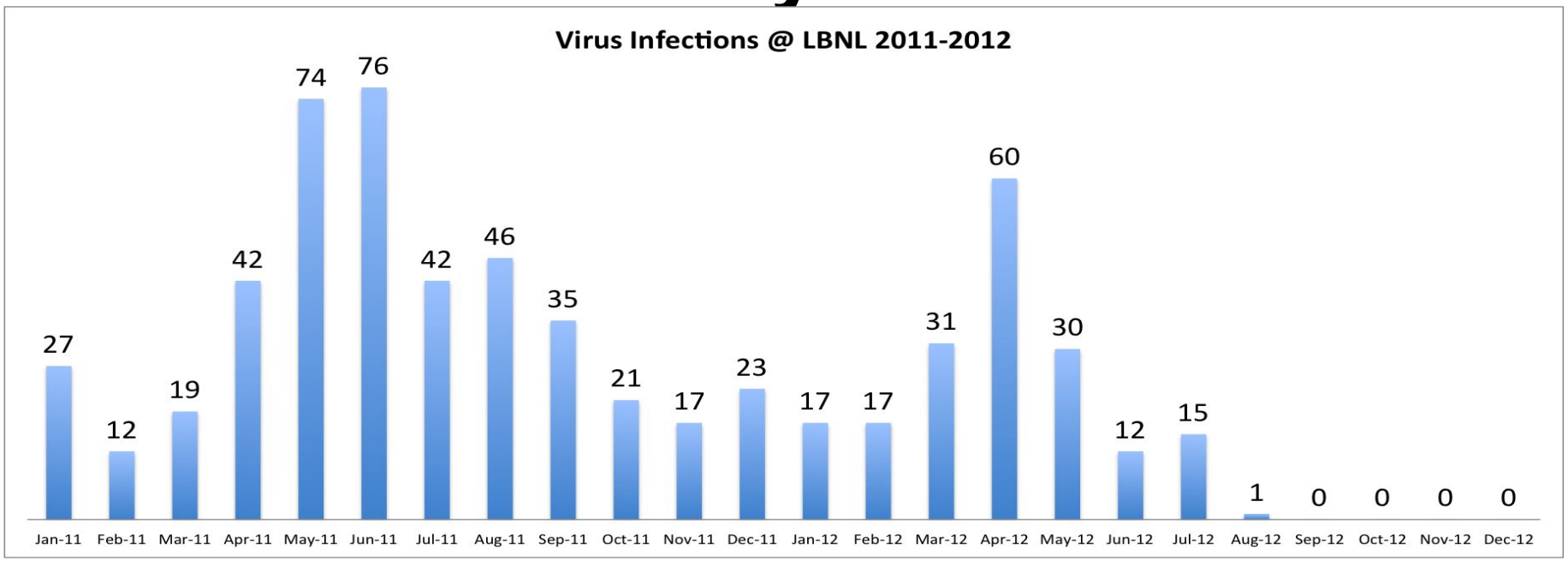




Alien-worm - Distributed coordinated Scanning



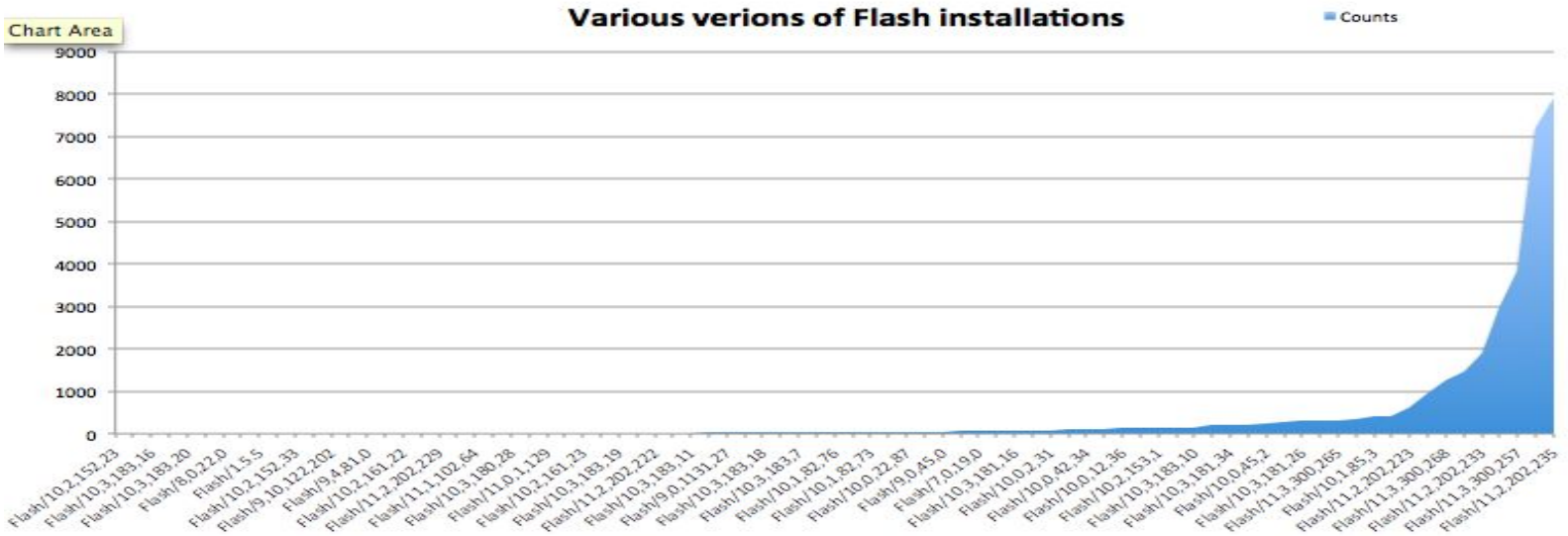
2: Vulnerability detection using Bro - to address drive-by-downloads

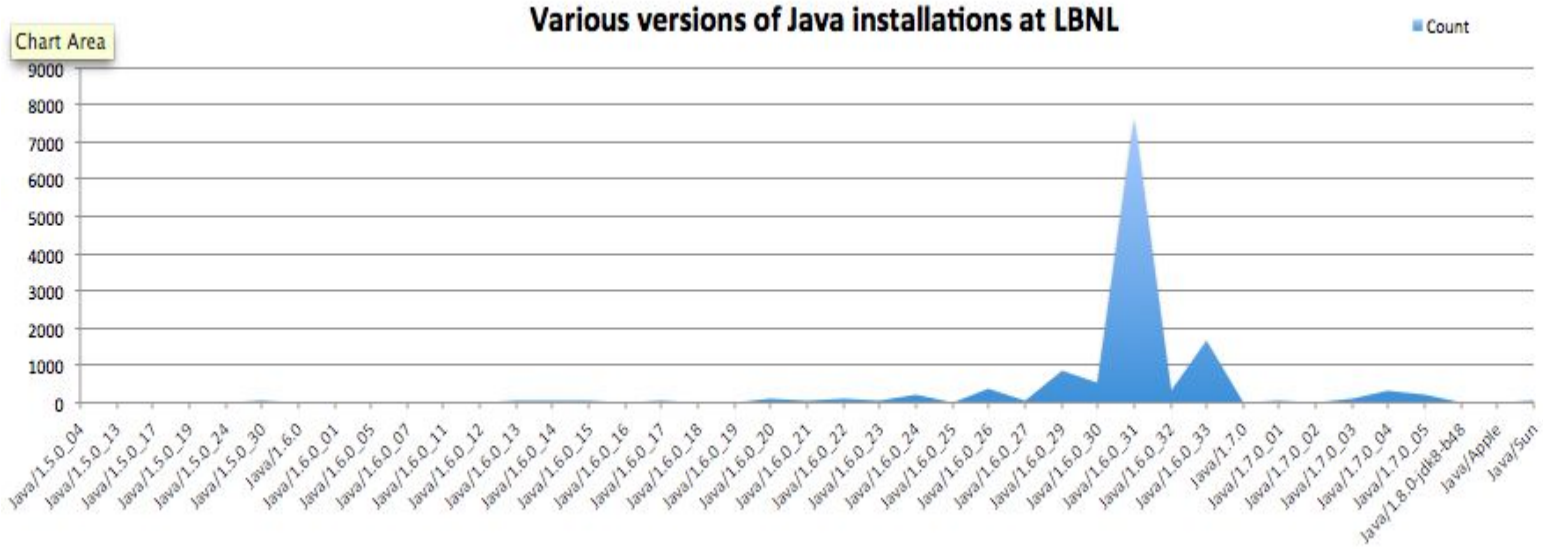


Aggressive patching and additional Influence of external factors:

- Google taking down various domains,
- New version of flash with auto-updates,
- Active blocking, RPZ of malicious domains etc

Java and Flash installations





IDS now allows you to

- Know what is running on your network - software.bro

but also

- Prioritize your patching and vulnerability management

3: Spear phishing: Looking for the very targeted stuff

- Came across: www.malware-tracker.com
- Ran some of the embedded pdf's and word documents collected over a period of time against it
- Needed some engineering to address volumes
  - python script - multiple buckets for clean, suspicious, infected
- At-least two new signatures issued by sophos after detection

Sample spear phish

Rating: EXPLOIT :1 MALWARE :1 SEVERITY :30 HITS :0 HAS_EXE :1

==========

Time: 1342007556.302290

From: <agulbra@nvg.unit.no>

To: <JOHNDOE@lbl.gov>

when_ts Wed, 11 Jul 2012 12:52:38 +0100

Subject: Re: Is that your document?

1342007557.718325 fPUHmL6ke86 212.58.56.90 64369 128.3.x.x 25 1 - 25 text/plain - - (empty)

1342007557.718325 fPUHmL6ke86 212.58.56.90 64369 128.3.x.x 25 1 part6.zip 29832 application/zip bb19060fde6e92bfaf5c585e56e3cb8e

/home/users/bro/extract/smtp-entity_212.58.56.90:64369-128.3.x.x:25_1.dat (empty)

1342007556.302290 fPUHmL6ke86 212.58.56.90 64369 128.3.x.x 25 1 lbl.gov <agulbra@nvg.unit.no> <johndoe@lbl.gov> Wed, 11 Jul 2012 12:52:38 +0100 agulbra@nvg.unit.no johndoe@lbl.gov -

- - Re: Is that your document? - - - 250 ok: Message 80405646 accepted 128.3.41.146,212.58.56.90

4: Input Framework - Leveraging the intelligence from the Community

Integrating Feeds from

- DoE CIRC, REN-ISAC (CIF), IID etc
  - IP address
  - host names
  - URLs
  - MD5

Example of Input-Framework -> Intel-Framework

Date: Mon, 6 Aug 2012 09:21:00 -0700 (PDT)

From: Big Brother <bro@.....lbl.gov>

To: ..... @lbl.gov

Subject: [IR (stomp) deep-bro] Intel::IID_SensitiveDNS_Lookup

Message: [hostname=speciallyregarding.com, property=malware, date_iso=20120706T143217Z] 2620:83:8000:140::3, 2001:400:613:18::74e

Sub-message: speciallyregarding.com

Connection: 2001:400:613:18::74e:63037 -> 2620:83:8000:140::3:53

Connection uid: IAlDTclj2Al

Email Extensions

---

orig/src hostname: <..... >

resp/dst hostname: nsx.lbl.gov

[Automatically generated]

5: "Deep-bro" - The real defense in depth

Q: Instead of standalone bro's why not make a cluster out of it?

A: Yes

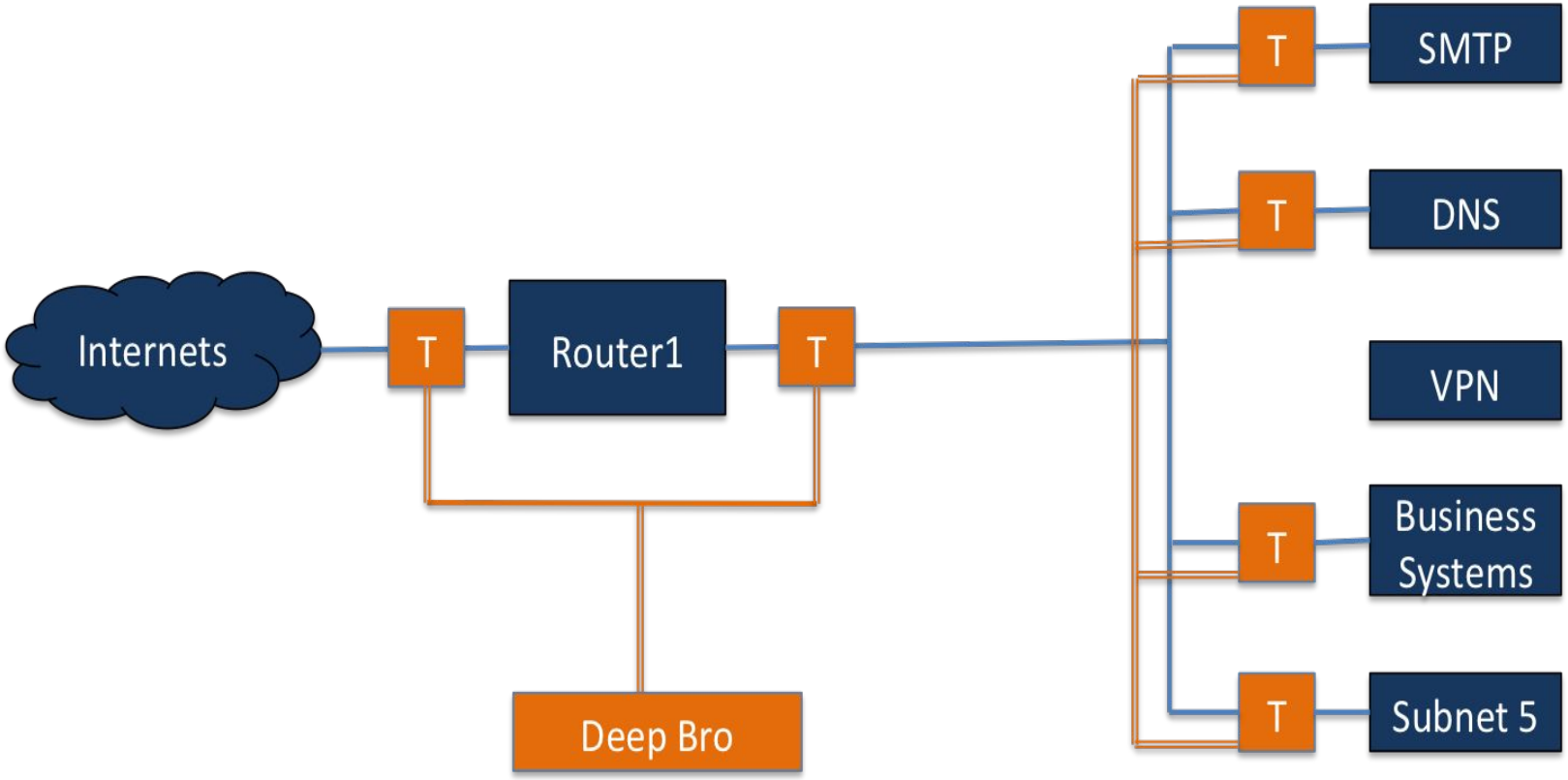
Q: Can we run bro on every subnet?

A: Ummm

Q: How about at-least the important subjects

A: Yes

deep-bro is name given to a bro cluster where worker nodes are spread throughout the internal subnets.



Deep-bro

For the starts, we converted the bro on subnets as part of cluster

- DNS
- SMTP
- VPN
- Business Systems

Plan is to expand to another 6-8 subnets asap

6: IPV6 - Trying to find unknowns

- LBL has Production wireless on IPv6
  - (LBLnet service, ACS, cluster on border)
- IP/mac address binding for dhcp jailing
  - ISC - yes, no, ummm, nah, not really, may be ....
- Tracker Ticket #833 extract the mac-address from the ICMPV6 using events
  - icmp_neighbor_advertisement
  - icmp_neighbor_solicitation
- Start alerting on attacks on ICMP protocol eg.
  - Rogue routers for fake router advertisements, build neighbor caches
  - Proactive response where a rogue RA results in another packet injected with lifetime of 0

Time-Machine

- Full-packet capture - 10TB a day
- Current Time-machine deployment
  - One Big time machine
  - Each cluster has a time-machine
  - Wireless Time machine
- Aim to keep full packets for as long as we can
- Hole in the marketplace - Solera etc aren't much useful to us
- trying various strategies
  - Priority deletions - encrypted, HTTP delete quickly while SMTP and class_all to hold on to for a long time, shrinking the capture sizes over period.

Time machine buckets

- Not truly a full packet capture
  - Various buckets with various different Cutoffs

Bucket	**Filters**	**Cutoff**
class "dns"	port 53	Xm
class "udp"	udp	Xm
class "icmp"	icmp	Xk
class "smtp"	port smtp or port 587	Xm
class "http"	port 80 or port 81 or port 631 or port 1080 or port 3128 or port 8000 or port 8080 or port 8888	Xm
class "encrypted"	port 22 or port 443 or port 993	Xk
class "all"	""	Xm

Instrumented SSH

- Keystrokes entered and responses sent with this version of SSHd is sent for analysis to Bro.
  - Sensitive information, such as passwords, is filtered out.

- Using various signatures, some complex and some fairly simple, Bro is able to alert us when an account appears compromised.
  - `unset HISTFILE` is still used by hackers to try and hide their tracks, but to us, its a smoking gun!

- Furthermore, once a compromise is confirmed, the logs from this version of SSH will help us determine the extent of the compromise and what, precisely, the intruder did.

- Code available at: `http://code.google.com/p/auditing-sshd/`

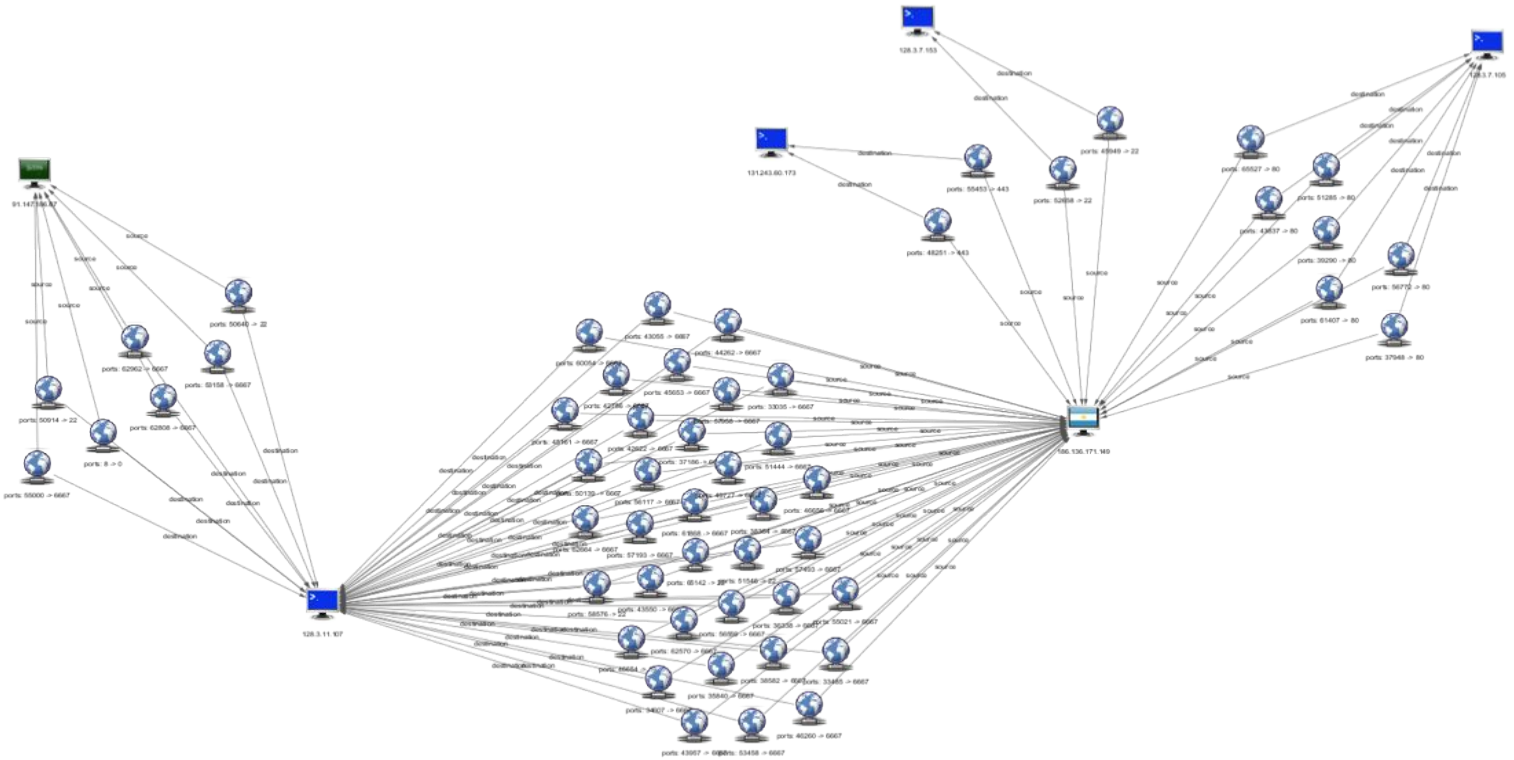
syslog2bro

- Block bruteforce SSH scanner on border
- Slightly more sophisticated attackers use a botnet where they would try one account per host/ip
- feed syslog data to bro - let it make correlations and initial drops
- Bro-2.0 has limited (udp) capability to sniff syslog on the wire - eliminating the need to feed it syslog

Got data, now what ?

- All the deployments and collection pose problem for data handling
- Frequent Searches
- Understand the data

Lynxeon - Graphical interface for Bro



Built-in Analytics +
Timeline +
Connection graphs +
Scriptable - similar to the bro policy + -
Bro policies work on the wire, While lynxeon provides a capability to run the policy on the logs

Searching logs: grep isn't enough...

- fgrep
  - 30+ mins - 2 hours
- GNU parallel - try it +++
  - 2-10 mins
- Hadoop - Fail
- Oracle database - Fail
- Biggest fastest disk array - Fail
- SSD - Fail

Mining bro logs

- Google BigQuery
  - < 10s
  - Presently 2 billion conn logs search results in < 10 seconds
  - *subject to the size of results too.
- Problems
  - new columns in the table ? Reindex ?
  - How do I account for what is not going into the Big Query?
  - Multiple-Columns, inner-join/outer-join ?
- Pricing
  - woha! I just ran a $22,143.99 query ?

100 Gb Roadmap

- Using bro cluster approach - solved problem
- Break 100Gb into multiple 10Gb feeds
- Exploring tapping infrastructure capabilities

- Various questions
  - Time machine ?
  - Selective logging/monitoring ?

Questions

Aashish Sharma - asharma@lbl.gov

Vincent Stoffer - vstoffer@lbl.gov