ARBOR NETWORKS

Listening to the Network: Leveraging Network Flow Telemetry for Security Applications

Darren Anstee
EMEA Solutions Architect

Introduction

- **Security has an increased focus from ALL businesses, whether they are an enterprise, ISP, IDC or OTT application service provider.**
  - Better awareness of issues & tighter regulation
  - Main-stream press coverage = senior management focus
  - Huge financial / brand costs when something goes wrong

- **So, why is 'Flow relevant to security?**
  - Flow leverages our investment in the routers / switches within our infrastructure to identify threats to our networks and services
  - Flow is generated regardless of traffic symmetry
  - Flow can be used to detect malware infected hosts, zero-day exploits, attacks, inside misuse / abuse, DDoS etc..
  - Flow can provide a network wide picture of what is actually going on (context)

ARBOR NETWORKS

How can ‘Flow Help us?

- **Flow can help us to understand how our networks are used:**
  - We can use flow to build a model of who uses what, when, how often and how much. This can give us a baseline for normal network activity
  - And, we can detect abnormal / malicious / unusual traffic on our networks.
  - We can classify what is going on, in context, to establish our risk.
  - And, we get valuable forensic data.

- **Flow should be one of the key mechanisms we have for monitoring our network, service and data security.**

ARBOR NETWORKS

Agenda

- Introduction
- **What is 'Flow?**
- How can we use 'Flow for Security Applications
- Flow Security Use Cases
  - Network / Data Integrity - Bot Detection
  - Service Availability - DDoS Detection

ARBOR NETWORKS

‘Flow, the Voice of the Network

- **Why ‘Flow?**
  - Netflow v5/v7/v8/v9, sFlow v4/v5, Jflow, cflow, Netstream v5/v9, IPFix, Flexible Netflow
  - Routers and switches support different versions / types.
    - Cisco, Juniper, Alcatel, Huawei, Foundry, HP, Brocade

- **‘Flow maintains traffic data in Flow Records in a flow cache, and optionally exports that flow data to a collection/analysis system.**
- **Flow Records represent a form of network telemetry which can describe the traffic streams headed to / passing through a router**
  - Flow Record = uni-directional traffic flow
  - Bi-directional conversations will be represented by at least two Flow Records (and maybe more).

ARBOR NETWORKS

Flow Records, Key and Non-Key Fields

- **Using Netflow v5 Record (still most common).**

Key Fields	Non-Key Fields / Counters
• Source IP Address • Destination IP Address • Source TCP/UDP Port • Destination TCP/UDP Port • Input IfIndex • Protocol • Type of Service	• Packet Count • Byte Count • First Packet Time • Last Packet Time • Output ifIndex • TCP Flags • Next Hop Address • Source AS Number • Dest. AS Number • Source Prefix Mask • Dest. Prefix Mask

ARBOR NETWORKS

Flow Record Export

1.Create and update flows in NetFlow cache

Key fields in yellow
Non-key fields white

SrcIf	SrcIPadd	DstIf	DstIPadd	Protocol	TOS	Flgs	Pkts	Src Port	Src Msk	Src AS	Dst Port	Dst Msk	Dst AS	NextHop	Bytes/ Pkt	Active	Idle
Fa1/0	173.100.21.2	Fa0/0	10.0.227.12	11	80	10	11000	00A2	/24	5	00A2	/24	15	10.0.23.2	1528	1745	4
Fa1/0	173.100.3.2	Fa0/0	10.0.227.12	6	40	0	2491	15	/26	196	15	/24	15	10.0.23.2	740	41.5	1
Fa1/0	173.100.20.2	Fa0/0	10.0.227.12	11	80	10	10000	00A1	/24	180	00A1	/24	15	10.0.23.2	1428	1145.5	3
Fa1/0	173.100.6.2	Fa0/0	10.0.227.12	6	40	0	2210	19	/30	180	19	/24	15	10.0.23.2	1040	24.5	14

2.Expiration

- Inactive timer expired (15 sec is default)
- Active timer expired (30 min (1800 sec) is default)

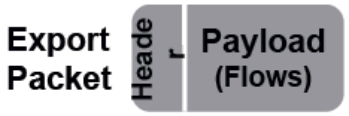
SrcIf	SrcIPadd	DstIf	DstIPadd	Protocol	TOS	Flgs	Pkts	Src Port	Src Msk	Src AS	Dst Port	Dst Msk	Dst AS	NextHop	Bytes/ Pkt	Active	Idle
Fa1/0	173.100.21.2	Fa0/0	10.0.227.12	11	80	10	11000	00A2	/24	5	00A2	/24	15	10.0.23.2	1528	1800	4

4.Export version

Non-Aggregated Flows—Export Version 5 or 9

5.Transport protocol

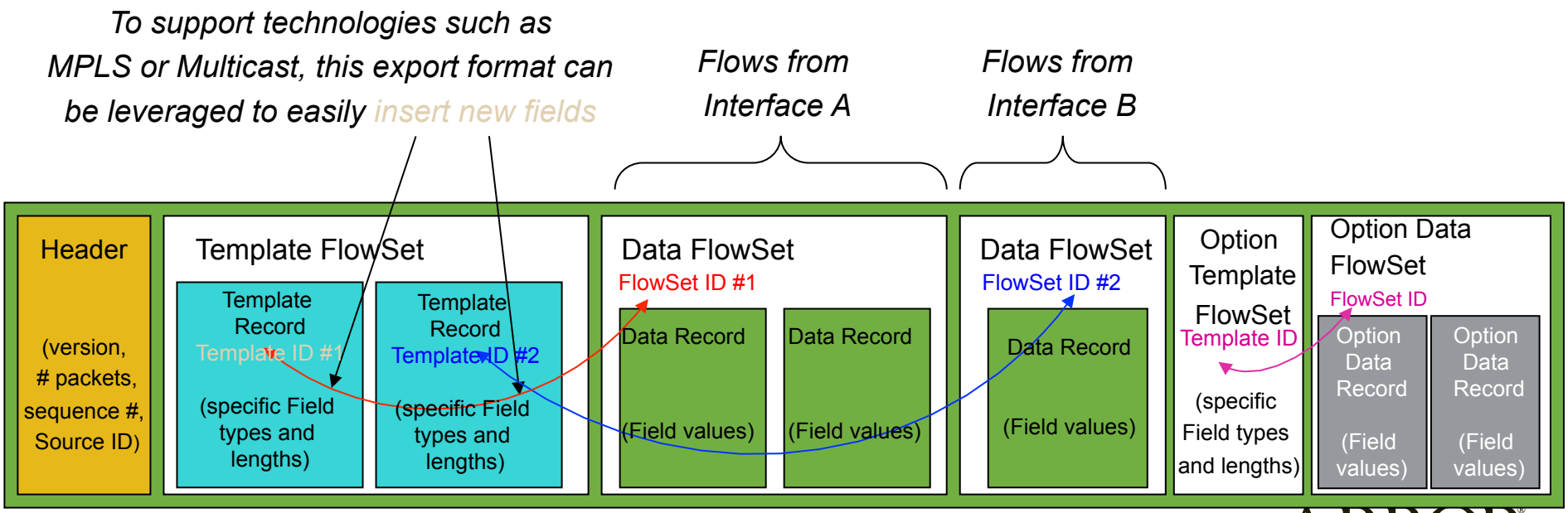
30 Flows per 1500 byte export packet



ARBOR NETWORKS

Extensible Flow : Netflow v9

- **Created to provide flexibility**
  - Additional ‘fields’ can be added to Netflow records.
- **Supported by Cisco, Juniper, Alcatel, Huawei etc…**
- **Required for routers to export Flow Records for MPLS, Multicast and IPv6 traffic.**



Extensible Flow : Flexible Netflow / IPFix

- **Flexible Netflow (Cisco)**
  - Allows user configurable Netflow Templates
    - Key, non-key, counter, time-stamp fields
  - Customised Netflow cache(s) for specific applications
  - Can reduce overhead:
    - Only ‘relevant’ information is sampled
    - Only ‘specified’ fields are stored
  - Introduces many new key / non-key fields
    - Can include NBAR and header / payload extracts.
  - Uses Netflow v9 format for export.
- **IPFix**
  - Standardised - RFC 5101, 5102
  - Similar export format to Netflow v9 but not identical
    - Version 10, sequence number counting etc..
    - Variable length fields etc..

ARBOR NETWORKS

Netflow Considerations

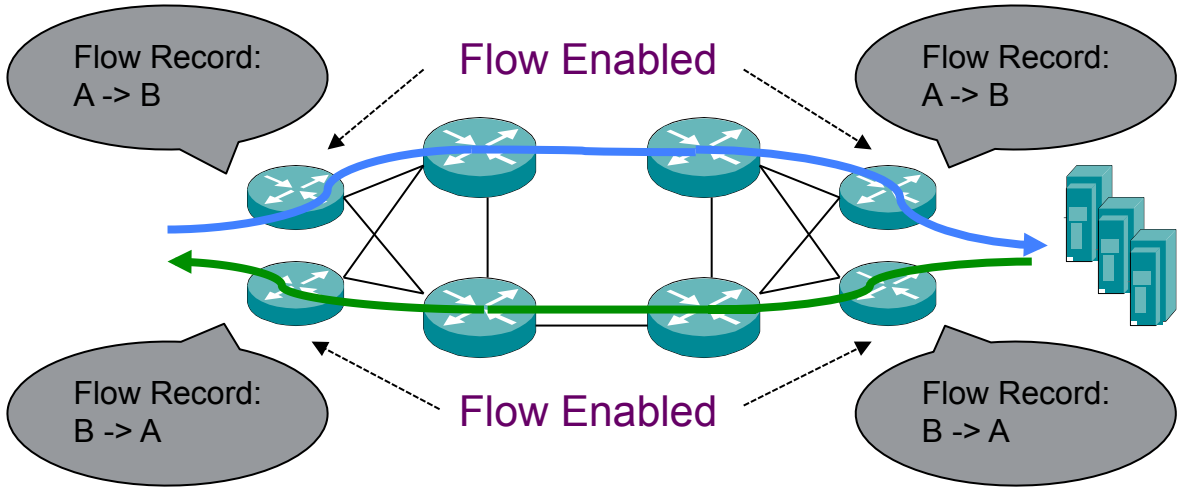
- **Sampled or Un-Sampled ‘Flow?**
  - Un-sampled ‘Flow is useful for troubleshooting, forensics, traffic analysis, and behavioral/relational anomaly-detection
  - Sampled ‘Flow is useful for traffic analysis and behavioral/relational anomaly-detection.
  - The choice comes down to router support / monitored and traffic volume / collection capabilities.

- **Monitoring with ‘Flow can scale for very large amounts of traffic**
  - Phone bill v’s wire-tap = scalability
    - Who’s talking to whom, over what protocols and ports, for how long, at what speed, for what duration, etc.
  - ‘Flow allows the routers / switches within the network infrastructure to be used as probes

ARBOR NETWORKS

Netflow Considerations, Where to Listen?

- **At network entry and exit points, in front of critical infrastructure to e.g. data-centre, extranet connection, internet gateway, peering edge, wherever we want visibility etc..**
- **Ingress 'Flow generation should typically be enabled on all router interfaces.**
  - Egress 'Flow generation in certain situations.
- **If traffic crosses multiple Flow enabled routers, multiple Flow Records may be generated representing the same traffic.**



ARBOR NETWORKS

Agenda

- Introduction
- What is ‘Flow?
- **How can we use ‘Flow for Security Applications**
- Flow Security Use Cases
  - Network / Data Integrity - Bot Detection
  - Service Availability - DDoS Detection

ARBOR NETWORKS

How can 'Flow Help us with our Security Posture?

- **As I said earlier….**

- **Flow can help us to understand how our networks are used.**
  - We can use flow to build a model of who uses what, when, how often and how much. This can give us a baseline for normal network activity
  - And, we can detect abnormal / malicious / unusual traffic on our networks.
  - We can classify what is going on in context, to establish our risk.
  - And, we can get valuable forensic data.
  - We can discover which customers / services share which infrastructure. This helps us to ensure availability

Behavioral Detection of Malware Infected Devices And DDoS Attacks

ARBOR NETWORKS

How can we use Flow?

- **We can look at the flow cache on each router. But….**
- **When Flow is enabled on router / switch infrastructure we can use a dedicated analysis systems to collect, detect, report on, and correlate observed activity.**
- **We can:**
  - See collated data across multiple devices.
  - Contrast current / historic traffic levels and patterns.
  - Detect bots / DDoS / insider misuse more easily.
  - Mine historical flow logs for forensic information.
- **Open source and commercial collection / analysis tools are available which greatly enhance the utility of Flow.**

ARBOR NETWORKS

How can we use Flow?

- **Multiple open source tools available:**
  - Nfdump / Nfsen
    - http://nfdump.sourceforge.net/
    - http://nfsen.sourceforge.net/
  - Stager
    - http://software.uninett.no/stager/
  - WebView Netflow Reporter
    - http://wvnetflow.sourceforge.net/
  - FlowViewer
    - http://ensight.eos.nasa.gov/FlowViewer/
  - Argus
    - http://www.qosient.com/argus/downloads.shtml
  - Others :
    - http://www.switch.ch/network/projects/completed/TF-NGN/floma/software.html
- **Commercial Tools**
  - More flexible, easier to configure, more scalable and supported.

ARBOR NETWORKS

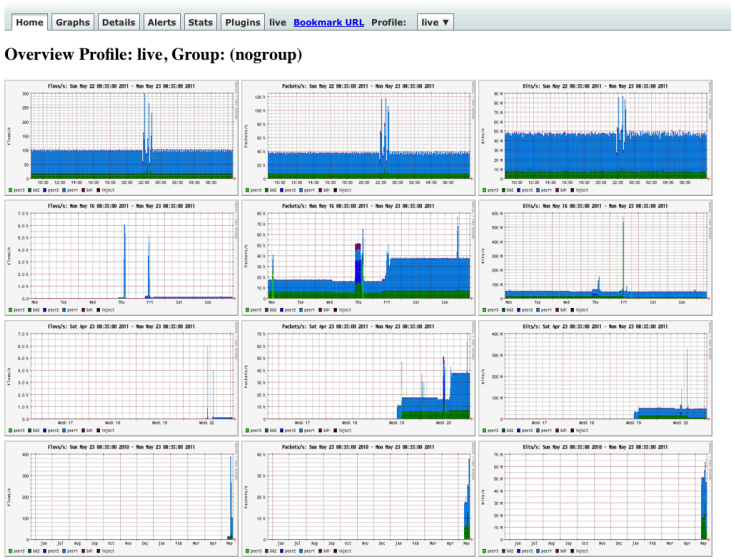
Flow Security Applications

- **Flow can help us to ensure network and data integrity and confidentiality + service availability.**
- **Numerous papers on the use of Flow for security applications:**
  - http://www.first.org/global/practices/Netflow.pdf
  - http://www.cert.org/flocon/2011/presentations/Krmicek_Detecting.pdf
  - http://www.ietf.org/proceedings/78/slides/NMRG-9.pdf
  - http://www.math.bme.hu/~slovi/temalabor3.pdf
  - Using machine learning techniques to identify botnet traffic. Livadas C., Walsh, R., Lapsley, D., Strayer, T. In: Proceedings of the 31st IEEE Conference on Local Computer Networks, 2006
  - Traffic aggregation for malware detection. Yen, T.-F., Reiter, M. K. . In: Proceedings of the 5th international conference on Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA '08), 2008
  - These are just a sample
- **Going to look at some (simple) examples**
  - Much more complex mechanisms available, see papers above

ARBOR NETWORKS

How can we use Flow?

- **Using :**
  - Nfdump / Nfsen, as an example
- **Why Nfdump / Nfsen?**
  - Flexible data-collection
    - Netflow v5 / v9, Sflow
  - Collated view of flow data
  - Good performance and scalability
  - Flexible, ad-hoc filtering of data
    - Good for investigating what is going on
  - Relatively easy to install / configure
    - Can be ‘working’ in less than a day



- **Why not send flow straight to a database?**
  - Scale, performance, scale, performance…….
- **Why not send flow to an event correlation system (splunk)?**
  - Flow is not refined enough
  - Use splunk for correlation of infection indicators from flow.

ARBOR NETWORKS

Agenda

- Introduction
- What is ‘Flow?
- How can we use ‘Flow for Security Applications
- **Flow Security Use Cases**
  - Network / Data Integrity – Indications of Malware Infection
  - Service Availability - DDoS Detection

ARBOR NETWORKS

Using Flow for Bot Identification

- Malware
  - Short for *malicious software*. Programming (code, scripts, active content, and other software) designed to disrupt or deny operation, gather information that leads to loss of privacy or exploitation, gain unauthorized access to system resources, and other abusive behavior. Source : US-CERT
- Botnet
  - In malware, a botnet is a collection of infected computers or bots that have been taken over by hackers (also known as bot herders) and are used to perform malicious tasks or functions Source : Wikipedia
- **Flow can help us ensure data and network integrity by providing cross-network visibility of malware infected devices:**
  - **Based on behavioral analysis / anomalies – zero day**
  - **Based on a match to 'known' behavior - CnC server**

ARBOR NETWORKS

Using Flow for Bot Identification

- **Detection via (simple) behavioral analysis / anomalies**
  - Allows us to detect zero day infections (no signature)
  - Utilises a match, or matches, on unusual host behavior:
    - Unusual outbound SMTP (Spam generation)
    - Off-net DNS queries
    - Scan detection
    - Based on outbound (DDoS) behavior
    - Other indicators – long-lived flows, unusual high volume transfers to external hosts etc..
  - Match more than one behavior, the likelihood of compromise grows

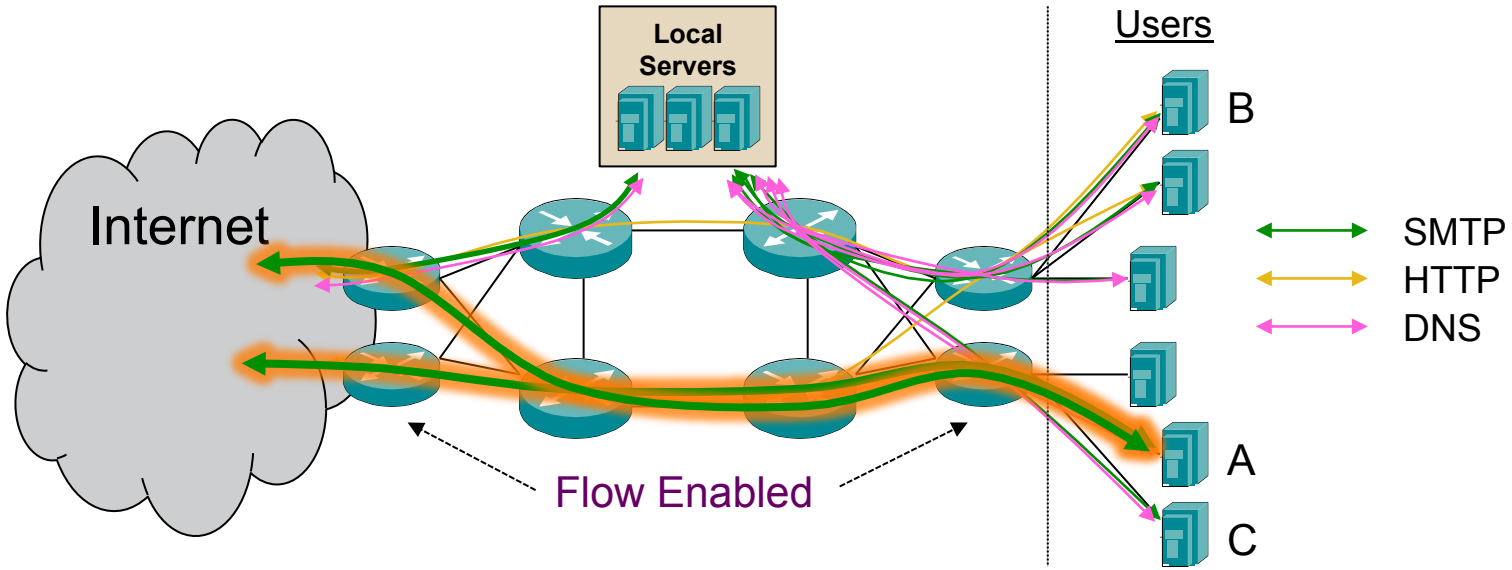
ARBOR NETWORKS

Using Flow for Bot Identification

- **Using a test network for examples, with real malware samples.**
  - Users on the 10.2.24.0/24 subnet
- **NOTE: Even if firewalls block traffic, routers / switches will still generate flow.**
- **NOTE: Even if routers / switches block traffic, they will still generate flow.**

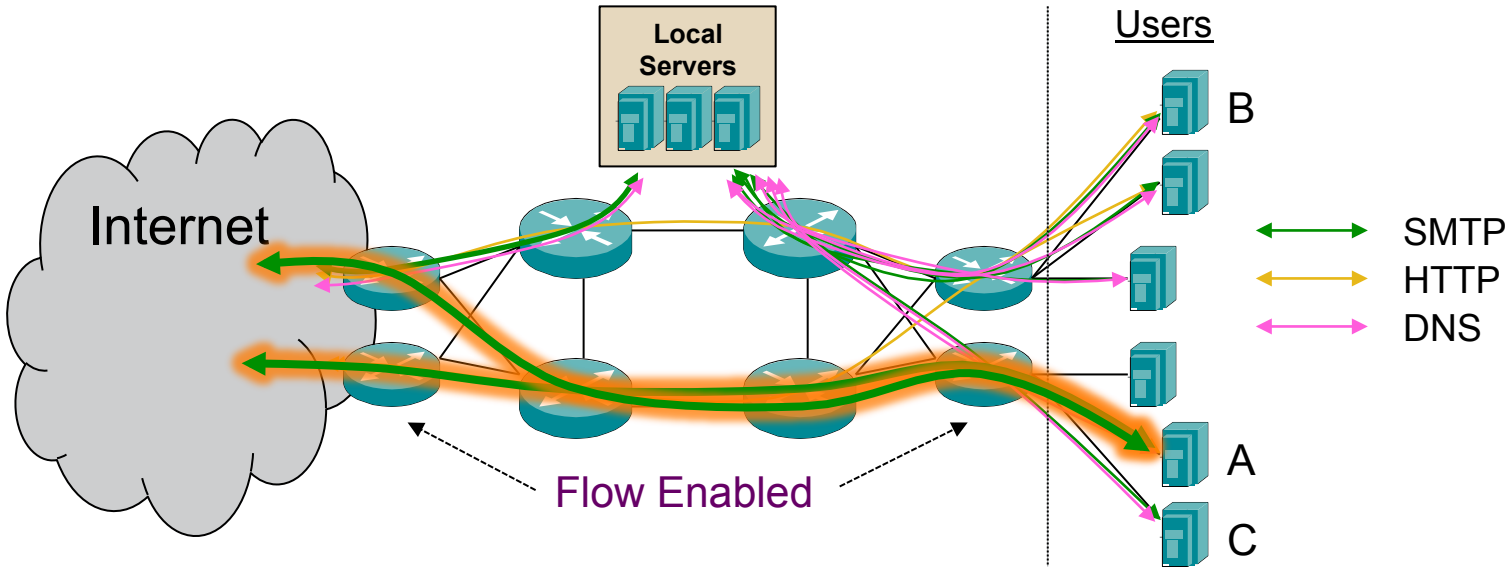
ARBOR NETWORKS

Using Flow for Bot Identification : Outbound SMTP



Host	Outbound SMTP	Off-Net DNS	Scanning	Outbound DDoS	Long Lived	High Volume	Possible Compromise
A							?
B							?

Using Flow for Bot Identification : Outbound SMTP



Host	Outbound SMTP	Off-Net DNS	Scanning	Outbound DDoS	Long Lived	High Volume	Possible Compromise
A	✔						?
B							?

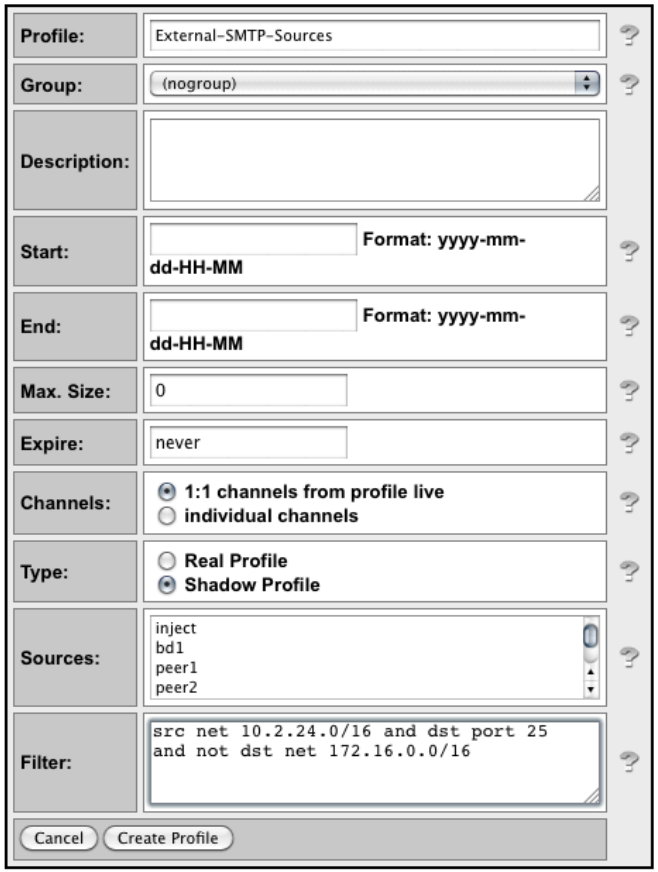
Using Flow for Bot Identification : Outbound SMTP

- **Bots can be used for Spam generation.**
  - Users do not normally use multiple external SMTP servers / send very large volumes of email. We can look for this behavior.
- **We can use nfdump to generate a list of sources, ranked by number of packets (we could use flows, bytes etc..)**
  - Traffic destined to port 25
  - Not going to local servers (172.16.0.0/16 in this case)
  - Constrain source based on desktop / customer address space (10.2.24.0/24 in this case)

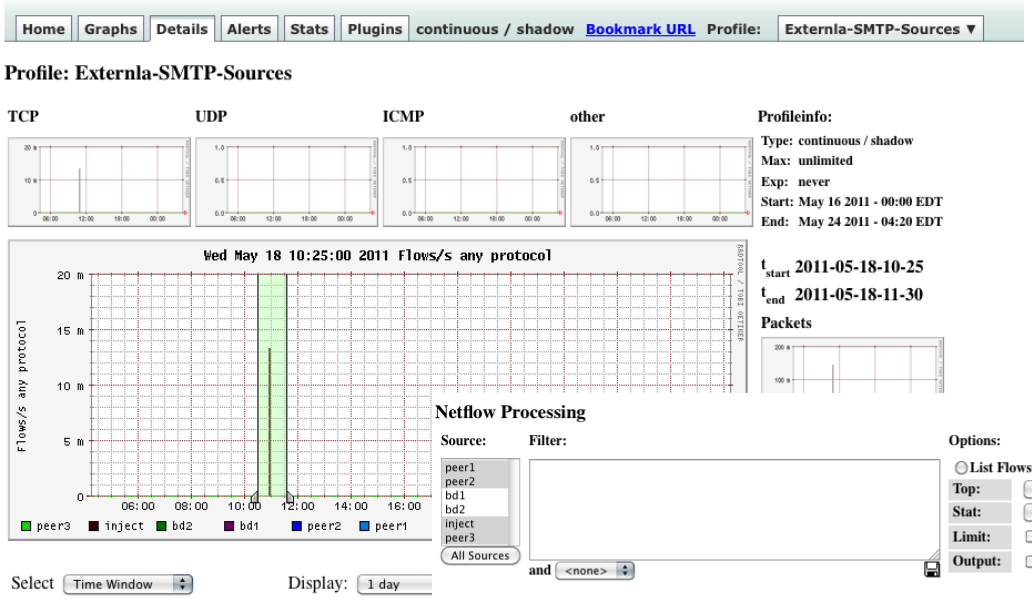
```
nfdump -R . -t 2011/05/02.00:00:00-2011/05/09 -s srcip/packets 'src net 10.2.24.0/24 and dst port 25 and
not dst net 172.16.0.0/16'
Top 10 Src IP Addr ordered by packets:
Date first seen          Duration Proto      Src IP Addr    Flows(%)     Packets(%)     Bytes(%)       pps      bps
bpp
2011-05-03 00:05:50.647     2.366 any        10.2.24.30      4(100.0)      43(100.0)    1752(100.0)       18
5923    40

Summary: total flows: 4, total bytes: 1752, total packets: 43, avg bps: 5923, avg pps: 18, avg bpp: 40
Time window: 2011-05-03 00:05:50 - 2011-05-03 00:05:53
Total flows processed: 3603568, Blocks skipped: 0, Bytes read: 270680004
Sys: 0.483s flows/second: 7456156.7  Wall: 0.476s flows/second: 7567929.5
```

Using Flow for Bot Identification : Outbound SMTP

- **Can see this visually in nfsen**
  - Need the correct profile configured to simplify investigation



Using Flow for Bot Identification : Outbound SMTP



Statistics timeslot May 18 2011 - 10:25 - May 18 2011 -

Channel:	Flows: all:	tcp:	udp:	icmp:	other:	Packe all:	tcp:	udp:
peer1	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s
peer2	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s
bd1	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s
bd2	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s
inject	0.0 /s	0.0 /s	0 /s	0 /s	0 /s	0.0 /s	0.0 /s	0 /s
peer3	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s

All None Display: Sum Rate

```
** nfdump -M /data/nfsen/profiles-data/live/peer3:inject:peer2:peer1 -T  -R 2011/05/18/nfcapd.201105181025:2011/05/18/nfcapd.201105181130 -n 10 -s srcip/packets
nfdump filter:
(( ident peer1) and (
src net 10.2.24.0/24 and dst port 25 and not dst net 172.16.0.0/16
)
or
( ident peer2) and (
src net 10.2.24.0/24 and dst port 25 and not dst net 172.16.0.0/16
)
or
( ident inject) and (
src net 10.2.24.0/24 and dst port 25 and not dst net 172.16.0.0/16
)
or
( ident peer3) and (
src net 10.2.24.0/24 and dst port 25 and not dst net 172.16.0.0/16
))
Top 10 Src IP Addr ordered by packets:
Date first seen          Duration Proto       Src IP Addr    Flows(%)     Packets(%)       Bytes(%)         pps      bps   bpp
2011-05-03 00:05:50.647     2.366 any          10.2.24.30        4(100.0)       43(100.0)     1752(100.0)       18     5923    40

Summary: total flows: 4, total bytes: 1752, total packets: 43, avg bps: 5923, avg pps: 18, avg bpp: 40
Time window: 2011-05-03 00:05:50 - 2011-05-03 00:05:53
Total flows processed: 19150, Blocks skipped: 0, Bytes read: 1013704
Sys: 0.009s flows/second: 1915191.5  Wall: 0.005s flows/second: 3310285.2
```

Using Flow for Bot Identification : Outbound SMTP

- **Use nfdump to drill down**
  - Which SMTP servers 10.2.24.30 tried to connect to

nfdump -R . -t 2011/05/02.00:00:00-2011/05/09.00:00:00 '**src host 10.2.24.30 and dst port 25**'

```
Date flow start          Duration Proto      Src IP Addr:Port          Dst IP Addr:Port    Packets
Bytes Flows

2011-05-03 00:05:50.647    0.516 TCP        10.2.24.30:1049  ->    94.100.176.20:25          4
168     1
2011-05-03 00:05:51.009    0.093 TCP        10.2.24.30:1051  ->     74.125.95.27:25          4
168     1
2011-05-03 00:05:51.080    1.191 TCP        10.2.24.30:1052  ->     66.111.4.73:25           4
168     1
2011-05-03 00:05:52.235    0.778 TCP        10.2.24.30:1053  ->   216.157.130.15:25          31
1248     1
Summary: total flows: 4, total bytes: 1752, total packets: 43, avg bps: 5923, avg pps: 18, avg
bpp: 40
Time window: 2011-05-03 00:05:50 - 2011-05-03 00:05:53
Total flows processed: 3603568, Blocks skipped: 0, Bytes read: 270679808
Sys: 0.499s flows/second: 7207626.1  Wall: 0.495s flows/second: 7274923.7
```

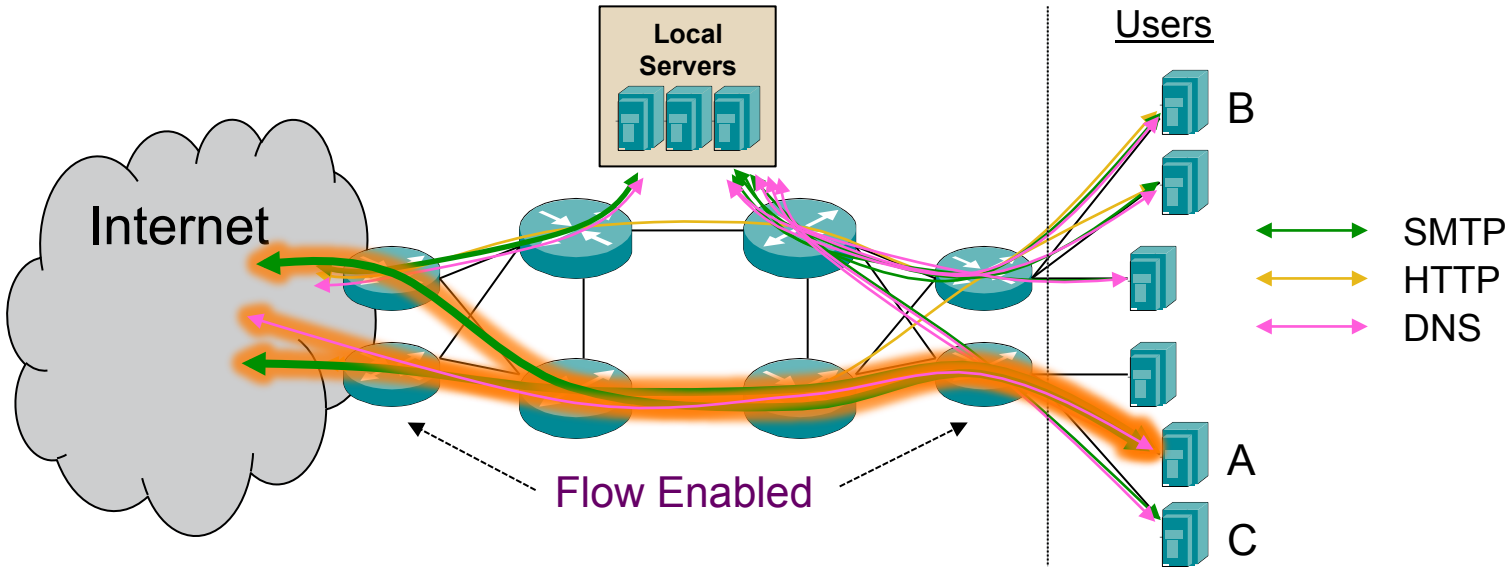
ARBOR NETWORKS

Using Flow for Bot Identification : Outbound SMTP

- **Host attempted to use four different SMTP servers**
  - Succeeded in utilizing one of them – unusual behavior
- **Also we can resolve the IP addresses of the servers to see if they look unusual, in this case:**
  - mxs.mail.ru
  - mx4.messagingengine.com
  - mail7.hsphere.cc
- **This may not be normal (dependent on your users), but now we know what question to ask.**
- **However, this is just one indicator.**
  - We can correlate the results of multiple indicators
  - Develop a higher confidence that a host is compromised.
- **NOTE: Script and cron for periodic, automated reports.**

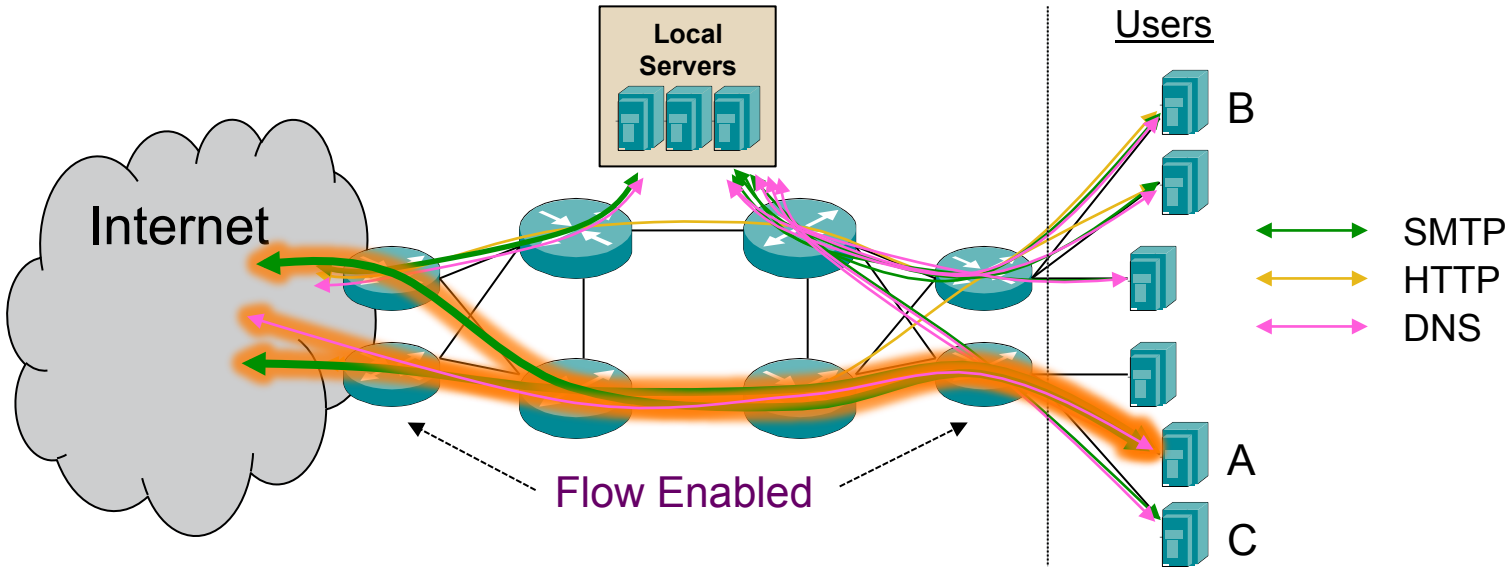
ARBOR NETWORKS

Using Flow for Bot Identification : Non-Local DNS



Host	Outbound SMTP	Off-Net DNS	Scanning	Outbound DDoS	Long Lived	High Volume	Possible Compromise
A	✔						?
B							?

Using Flow for Bot Identification : Non-Local DNS



Host	Outbound SMTP	Off-Net DNS	Scanning	Outbound DDoS	Long Lived	High Volume	Possible Compromise
A	✔	✔					✔
B							?

Using Flow for Bot Identification : Non-Local DNS

- **Most network hosts will utilise the local DNS servers**
  - There will be legitimate exceptions
- **As with SMTP we can query our flow data:**
  - My local DNS server is 10.2.0.25
  - Constraining the src addresses to be within my ‘user’ space.

nfdump -R . -t 2011/05/02.00:00:00-2011/05/09.00:00:00 -s srcip/packets **'src net 10.2.24.0/24 and dst port 53 and not dst host 10.2.0.25'**
Top 10 Src IP Addr ordered by packets:
Date first seen Duration Proto Src IP Addr Flows(%) Packets(%) Bytes(%) pps bps bpp
2011-05-03 00:05:49.508 32.419 any **10.2.24.30** 5(100.0) 9(100.0) 555(100.0) 0 136 61

Summary: total flows: 5, total bytes: 555, total packets: 9, avg bps: 136, avg pps: 0, avg bpp: 61
Time window: 2011-05-03 00:05:49 - 2011-05-03 00:06:21
Total flows processed: 3603568, Blocks skipped: 0, Bytes read: 270685604
Sys: 0.523s flows/second: 6886263.7 Wall: 0.491s flows/second: 7327244.2

ARBOR NETWORKS

Using Flow for Bot Identification : Non-Local DNS

- **We can see which DNS servers 10.2.24.30 tried to use by drilling down into our forensic flow information:**

nfdump -R . -t 2011/05/02.00:00:00-2011/05/09.00:00:00 -o long **'src host 10.2.24.30 and dst port 53 and not dst host 10.2.0.25'**

```
Date flow start          Duration Proto      Src IP Addr:Port          Dst IP Addr:Port   Flags Tos  Packets
Bytes Flows
2011-05-03 00:05:49.645     0.039 UDP        10.2.24.30:1044  ->      128.8.10.90:53    ......  0        1
49     1
2011-05-03 00:05:49.508     0.944 UDP        10.2.24.30:1025  ->      172.24.50.1:53    ......  0        5
338     1
2011-05-03 00:05:49.724     0.139 UDP        10.2.24.30:1046  ->    202.12.27.33:53     ......  0        1
49     1
2011-05-03 00:05:49.897     0.087 UDP        10.2.24.30:1047  ->     198.41.0.4:53      ......  0        1
48     1
2011-05-03 00:06:21.852     0.075 UDP        10.2.24.30:1055  ->      172.24.50.1:53    ......  0        1
71     1
Summary: total flows: 5, total bytes: 555, total packets: 9, avg bps: 136, avg pps: 0, avg bpp: 61
Time window: 2011-05-03 00:05:49 - 2011-05-03 00:06:21
Total flows processed: 3603568, Blocks skipped: 0, Bytes read: 270685632
Sys: 0.509s flows/second: 7066304.6  Wall: 0.503s flows/second: 7158714.5
```

ARBOR NETWORKS

Using Flow for Bot Identification : Non-Local DNS

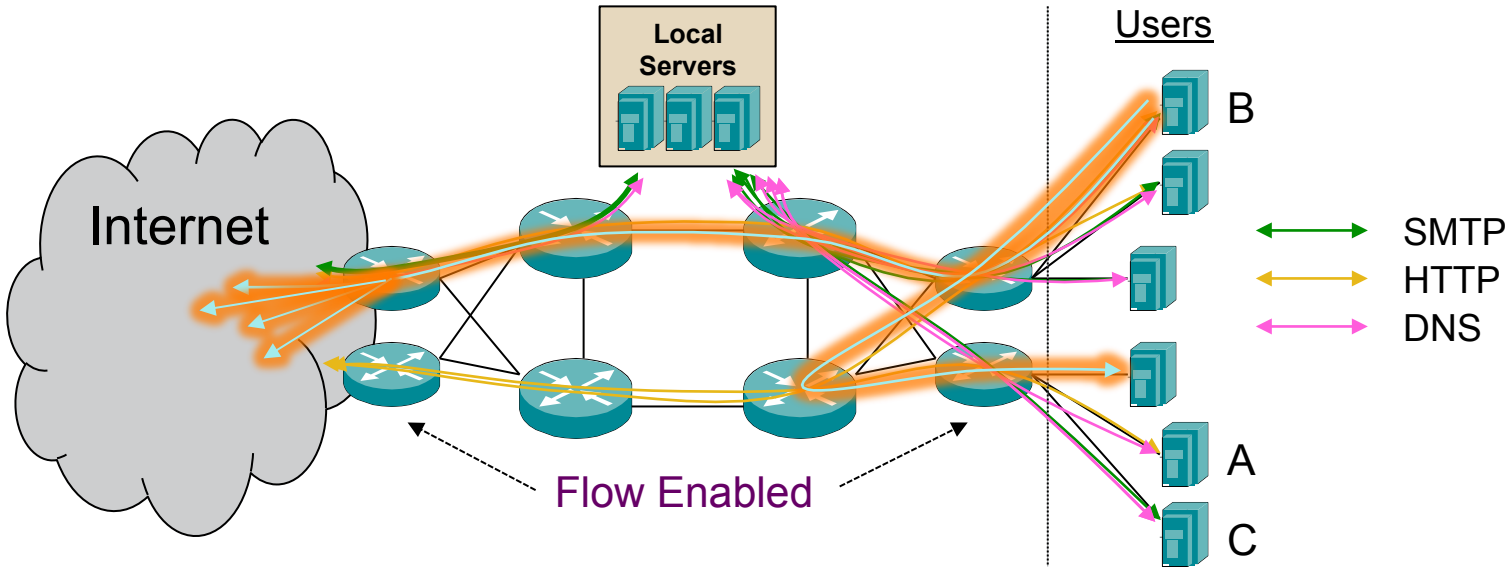
- **If we resolve the DNS servers we can see that they were a,d and m root server instances.**
  - Unusual for a user host.
- **And, this is the same user IP as before :**
  - Multiple indicators for the same IP
  - So, probably worth investigating this machine further.
  - Or, we can look for other indicators….

Host	Outbound SMTP	Off-Net DNS	Scanning	Outbound DDoS	Long Lived	High Volume	Possible Compromise
A	✔	✔					✔
B							?
C							?

- **NOTE: Can use a database (MySQL, for example) or splunk to correlate the results of indicators.**

ARBOR NETWORKS

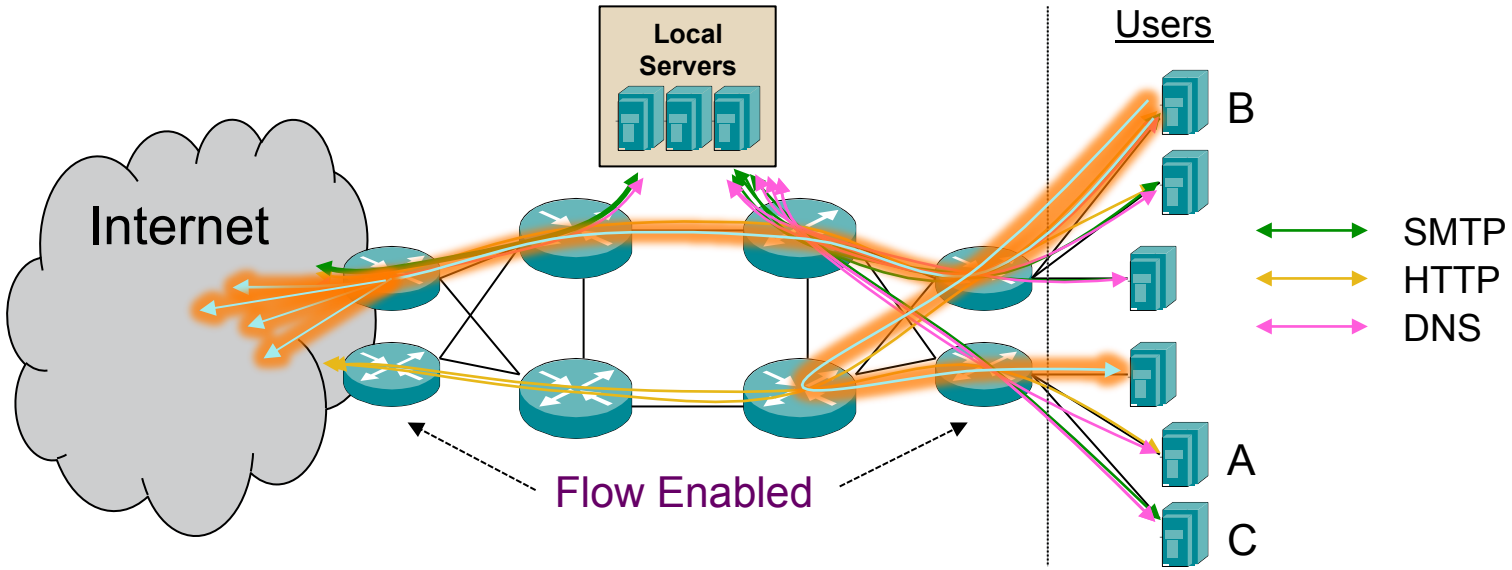
Using Flow for Bot Identification : Scanning



Host	Outbound SMTP	Off-Net DNS	Scanning	Outbound DDoS	Long Lived	High Volume	Possible Compromise
A							?
B							?

ARBOR NETWORKS

Using Flow for Bot Identification : Scanning



Host	Outbound SMTP	Off-Net DNS	Scanning	Outbound DDoS	Long Lived	High Volume	Possible Compromise
A							?
B			✔				?

Using Flow for Bot Identification : Scanning

- **Scans from a host are another possible indicator**
  - Can also be due to mis-configuration, NMS applications, Windows Browser / SMB traffic
- **As before we can search for scans in our flow data:**

nfdump -R . -t 2011/05/16.00:00:00-2011/05/23.00:00:00 -s srcip/flows -s dstport/flows '**src net 10.2.24.0/24 and proto tcp and ((flags S and not flags FRAUP) or (flags SR and not flags FAUP))**'
Top 10 Src IP Addr ordered by flows:

Date first seen	Duration	Proto	Src IP Addr	Flows(%)	Packets(%)	Bytes(%)	pps	bps	bpp
2011-05-17 11:46:34.368	17789.053	any	**10.2.24.32**	274(98.2)	1056(45.2)	134528(63.7)	0	60	127
2011-05-17 13:27:07.943	365.841	any	10.2.24.6	4(1.4)	1024(43.8)	61440(29.1)	2	1343	60
2011-05-17 16:23:33.212	0.000	any	10.2.24.33	1(0.4)	256(11.0)	15360(7.3)	0	0	60

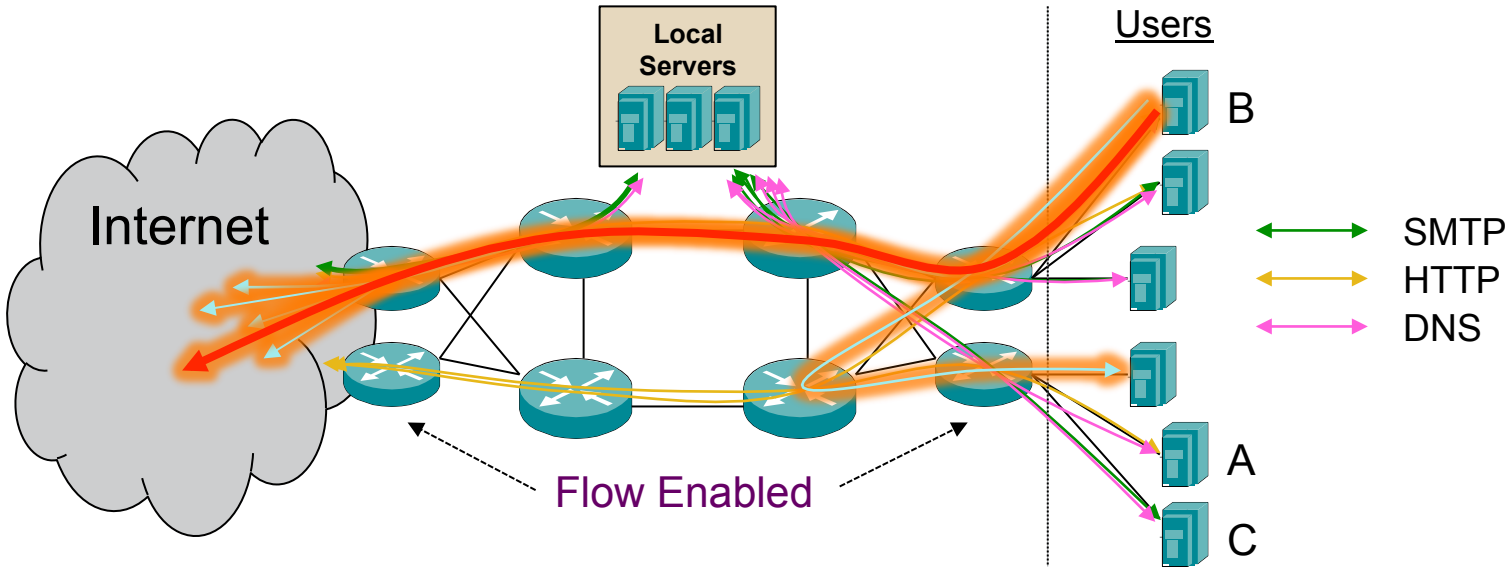
Top 10 Dst Port ordered by flows:

Date first seen	Duration	Proto	Dst Port	Flows(%)	Packets(%)	Bytes(%)	pps	bps	bpp
2011-05-17 11:46:34.368	0.859	any	**27031**	272(97.5)	544(23.3)	23936(11.3)	633	222919	44
2011-05-17 13:28:05.839	10527.373	any	22	4(1.4)	1024(43.8)	61440(29.1)	0	46	60
2011-05-17 14:00:03.569	9779.852	any	80	2(0.7)	512(21.9)	110592(52.3)	0	90	216
2011-05-17 13:27:07.943	0.000	any	23	1(0.4)	256(11.0)	15360(7.3)	0	0	60

Summary: total flows: 279, total bytes: 211328, total packets: 2336, avg bps: 95, avg pps: 0, avg bpp: 90
Time window: 2011-05-17 11:46:34 - 2011-05-17 16:43:03
Total flows processed: 4478280, Blocks skipped: 0, Bytes read: 317012836
Sys: 0.689s flows/second: 6490693.6 Wall: 0.655s flows/second: 6827801.7

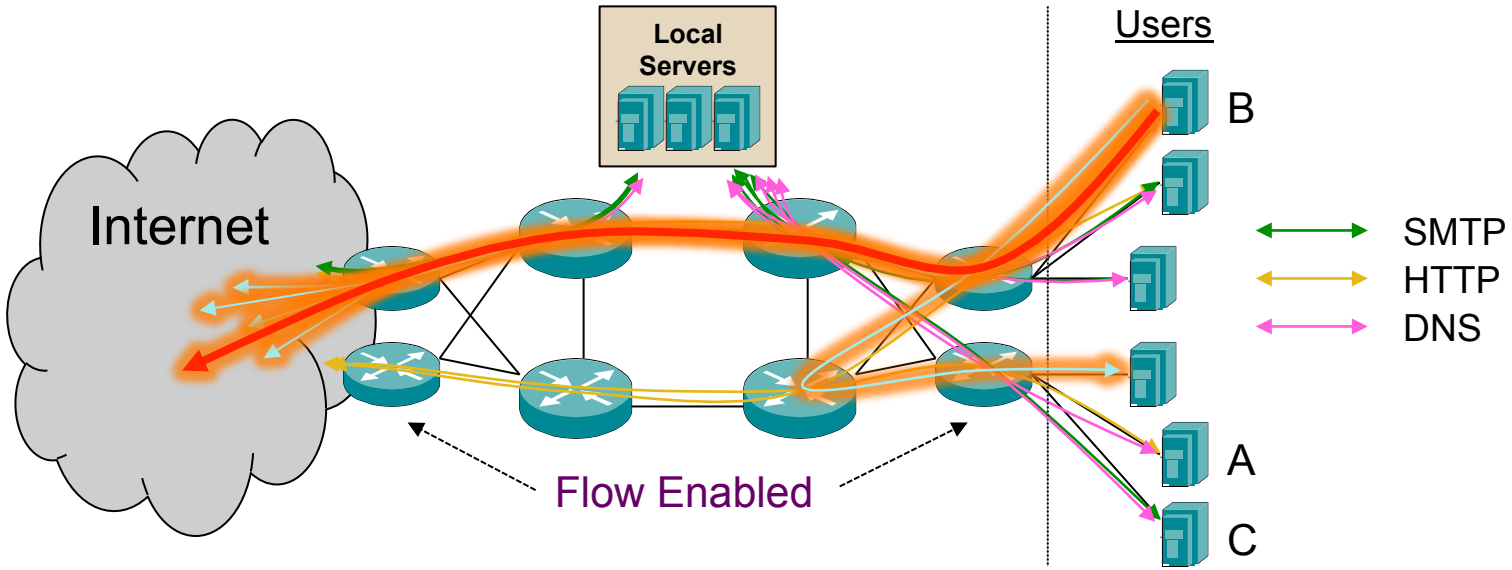
ARBOR NETWORKS

Using Flow for Bot Identification : Outbound DDoS



Host	Outbound SMTP	Off-Net DNS	Scanning	Outbound DDoS	Long Lived	High Volume	Possible Compromise
A							?
B			✔				?

Using Flow for Bot Identification : Outbound DDoS



Host	Outbound SMTP	Off-Net DNS	Scanning	Outbound DDoS	Long Lived	High Volume	Possible Compromise
A							?
B			✔	✔			✔

ARBOR NETWORKS

Using Flow for Bot Identification : Outbound DDoS

- **Outbound DDoS traffic is another (strong) indicator**
  - Even if the traffic doesn't make it out of the network Flow will still be generated.
- **Look for common attacks types:**
  - SYN Flood, RST Flood, UDP Flood, ICMP Flood etc..
- **Implement detection 'thresholds' by using a combination of 'pps' and 'packets' filters when searching for flows.**

nfdump -R . -t 2011/05/02.00:00:00-2011/05/09.00:00:00 **'src net 144.0.0.0/8 and proto icmp and pps > 100 and packets > 3000 and duration > 30000'**

```
…..snip
2011-05-03 18:35:55.973    59.967 ICMP      10.2.24.32:0    ->  XXX.255.182.167:8.0      6716    402978    1
…..snip
```

- **ICMP Flows with pps rate > 100 and with more than 3K packets counted and duration of more than 30 seconds.**
  - NOTE: nfdump cannot filter on a duration longer than your active flow expiry timer.

ARBOR NETWORKS

Using Flow for Bot Identification : Outbound DDoS

- If we detect a host generating any unusual, malware related, behavior use the flow log as a forensic tool to try establish potential CnC server addresses

```
nfdump -R . -t 2011/05/03.18:30:00-2011/05/03.18:40:00 'src host 10.2.24.32'
…..snip
2011-05-03 18:35:43.389    0.000 UDP      10.2.24.32:138   ->   172.24.50.103:138      10    2080    1
2011-05-03 18:35:54.461    0.469 UDP      10.2.24.32:1025  ->    172.24.50.1:53         1     61     1
2011-05-03 18:35:54.952   56.409 TCP      10.2.24.32:1048  ->  XXX.186.38.173:5050      5    441     1
2011-05-03 18:35:43.389    0.000 UDP      10.2.24.32:138   ->   172.24.50.103:138      10    2080    1
2011-05-03 18:35:55.973   59.967 ICMP     10.2.24.32:0     ->  XXX.255.182.167:8.0     6716   402978    1
…..snip
```

- Outbound connection just before the attack flow.
  - This might be perfectly valid
  - 5050 is one of the yahoo messenger ports
  - the destination IP resolves to a .cn domain
- But, Flow has given us the ability to investigate.
  - Now we can ask **the right questions** etc..
- We can then search our flowlog to see if any other hosts connect to our potential CnC address – as they will also need investigation / clean-up

ARBOR NETWORKS

Using Flow for Bot Identification : Other Indicators

- Other potential indicators of security issues using flow
  - Large volumes of traffic leaving our network unexpectedly
    - Indicative of file transfers / streaming / p2p etc..

```
nfdump -R . -a -L +20M -t 2011/05/16.00:00:00-2011.05/23.00:00 'src net 10.2.24.0/24'
Byte limit: > 20000000 bytes
Date flow start          Duration Proto      Src IP Addr:Port          Dst IP Addr:Port   Packets    Bytes Flows
2011-05-16 04:49:21.887  7041.709 TCP        10.2.24.27:22    ->       10.1.15.16:61734    1.1 M  234.2 M  4205
Summary: total flows: 16009, total bytes: 527.9 M, total packets: 2.5 M, avg bps: 3314, avg pps: 1, avg bpp: 210
Time window: 2011-05-03 00:05:49 - 2011-05-17 18:01:12
Total flows processed: 77661319, Blocks skipped: 0, Bytes read: 4126141420
Sys: 8.529s flows/second: 9105086.8  Wall: 8.532s flows/second: 9101761.7
```

  - Long lived flows to external hosts
    - Key logging, CnC Connections etc..
    - Remember that we cannot search directly (using nfdump) for durations longer than our active flow expiry so must post process.

```
nfdump -R . -a -t 2011/05/09.00:00:00-2011/05/16.00:00 'src net 10.2.24.0/24' | awk '{if ($3 > 86400) {print $0;};}'
Date flow start          Duration Proto      Src IP Addr:Port          Dst IP Addr:Port   Packets    Bytes Flows
2011-05-13 10:09:22.432 357995.390 ICMP      10.2.24.6:8      ->      10.2.24.27:0.0      3584   595968    14
```

ARBOR NETWORKS

Using Flow for Bot Identification : Known CnC

- **As well as behavioral anomalies, we can also look for traffic towards 'known' CnC servers.**
  - Need a list of known CnC IPs.
  - These lists can be LARGE.
  - Lists can be obtained from a variety of sources e.g.
    - http://www.emergingthreats.net/index.php/rules-mainmenu-38.html
    - http://www.sunbeltsoftware.com/Malware-Research-Analysis-Tools/ThreatTrack/
  - **Search our flow logs to establish if any connections match our list of CnC IPs**
    - Using an Arbor list here

proto tcp AND ((port 5276 AND (host 210.166.220.222)) OR (port 6660 AND (host 84.208.29.17 OR host 69.61.21.115 OR host 67.198.195.194 OR host 194.14.236.50 OR host 217.174.199.222 OR host 195.13.58.57 OR host 64.32.20.108)) OR (port 6661 AND (host 202.156.1.18)) OR (port 6662 AND (host 84.27.119.230)).....VERY LONG

ARBOR NETWORKS

Using Flow for Bot Identification : Known CnC

```
nfdump -R . -f /root/cnc_list.txt -t 2011/05/02.00:00:00-2011/05/09.00:00
Date flow start          Duration Proto      Src IP Addr:Port          Dst IP Addr:Port   Packets    Bytes
Flows
2011-05-03 18:35:42.016    15.956 TCP         10.2.24.32:1046  ->     64.74.223.46:80         73
5347     1
2011-05-03 18:35:42.016    15.956 TCP       64.74.223.46:80    ->       10.2.24.32:1046       80
101855     1
Summary: total flows: 2, total bytes: 107202, total packets: 153, avg bps: 53748, avg pps: 9, avg
bpp: 700
Time window: 2011-05-03 18:35:42 - 2011-05-03 18:35:57
Total flows processed: 77340971, Blocks skipped: 0, Bytes read: 4109552773
Sys: 462.223s flows/second: 167323.9   Wall: 467.673s flows/second: 165374.0
```

- **We can clearly see a host within our user / customer address range.**

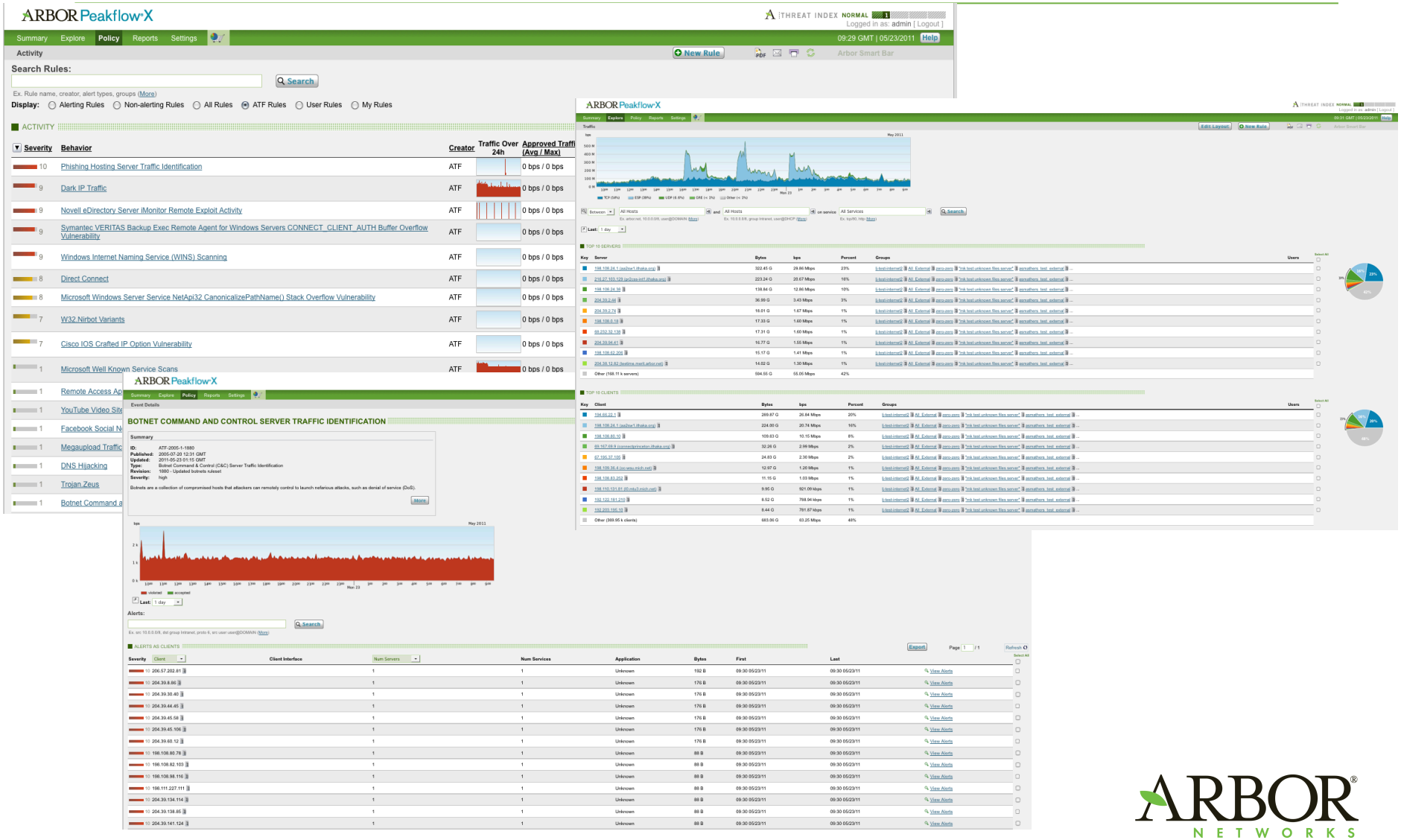
ARBOR NETWORKS

Using Flow for Bot Identification

- **Flow is a cost-effective and scalable way of detecting malware infected hosts.**
  - Leverages the functionality available within routers / switches
  - We can see 'inside' the network
- **Not reliant on signatures (zero-day)**
- **Provides multiple 'indicators' that a host may be infected**
  - The more indicators, the more likely the host is compromised
- **Detailed forensic data to establish exposure.**
- **Why use flow over firewall logs?**
  - Pervasive visibility, context, scalability, standardized record formats, easy to use open-source tools.
- **Flow can help us ensure the integrity of our networks / data.**

ARBOR NETWORKS

Using Flow for Bot Identification



Agenda

- Introduction
- What is ‘Flow?
- How can we use ‘Flow for Security Applications
- **Flow Security Use Cases**
  - Bot Detection
  - DDoS Detection

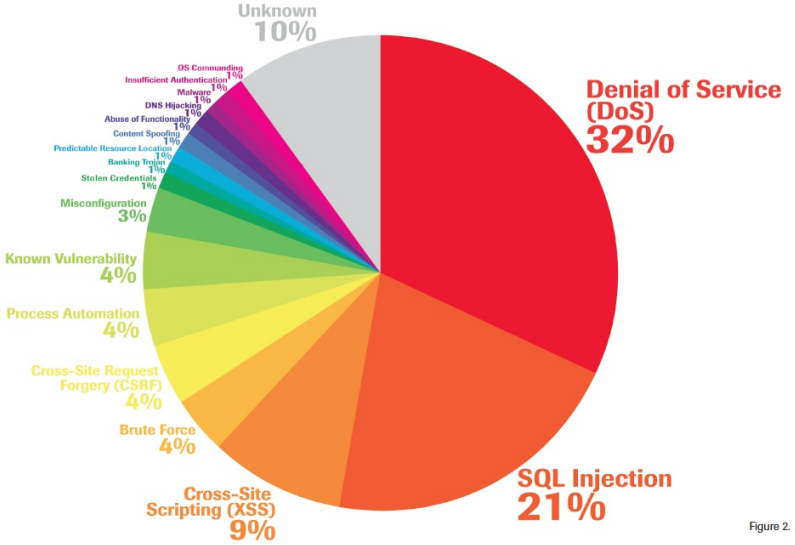
ARBOR NETWORKS

Using Flow for DDoS Detection : Primer

- **Flow can also help us to detect and classify DDoS attacks, a major threat to service availability.**

What is a Denial of Service attack?

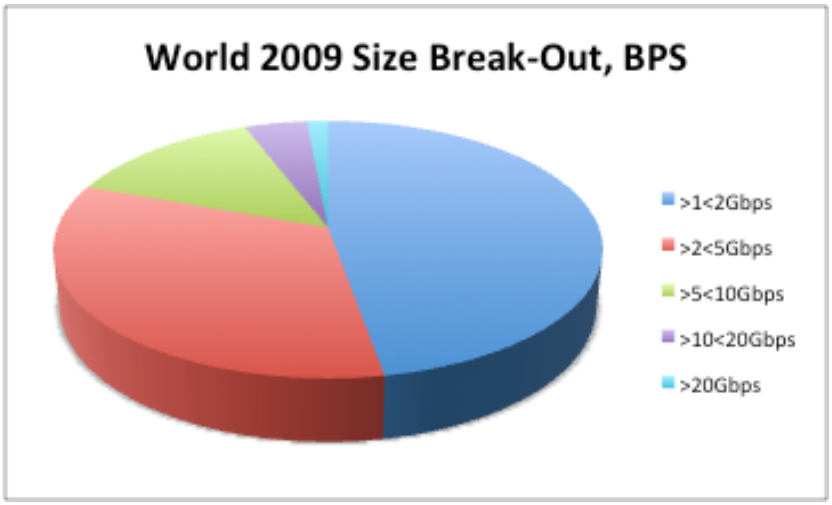
- **An attempt to consume finite resources, exploit weaknesses in software design or implementation, or exploit lack of infrastructure capacity**
- **Effects the availability and utility of computing and network resources**
- **Attacks can be *distributed* for even more significant effect**
- **The *collateral damage* caused by an attack can be as bad, if not worse, than the attack itself**

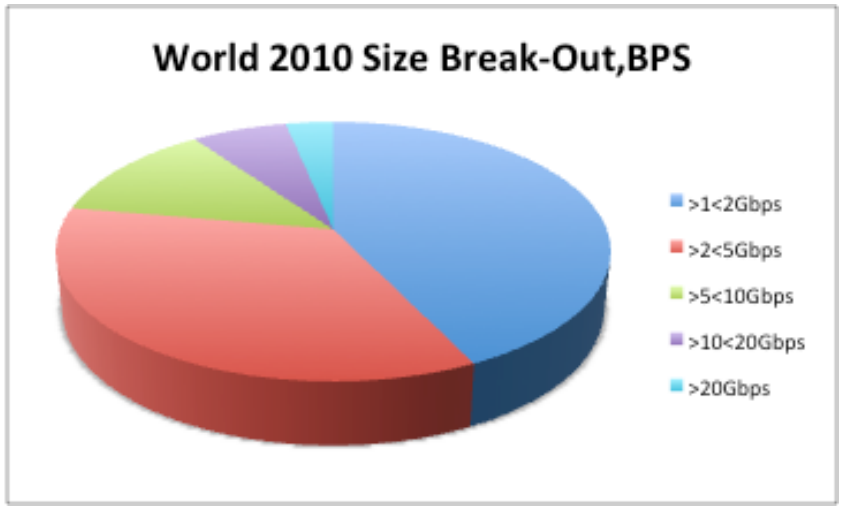


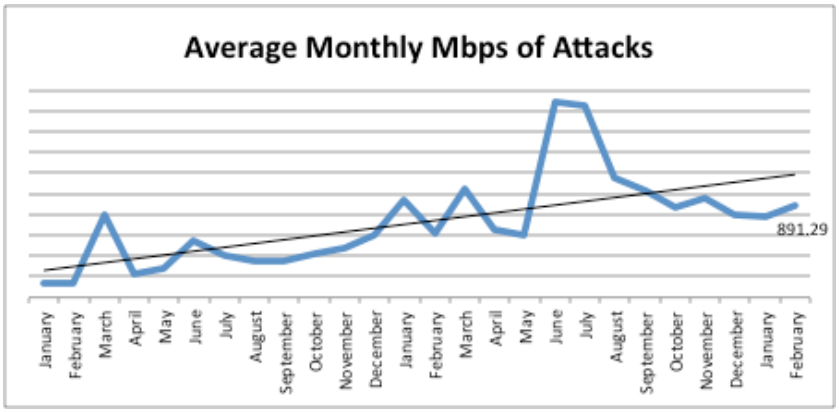
Source: Top Attack Methods, Trustwave WHID Report

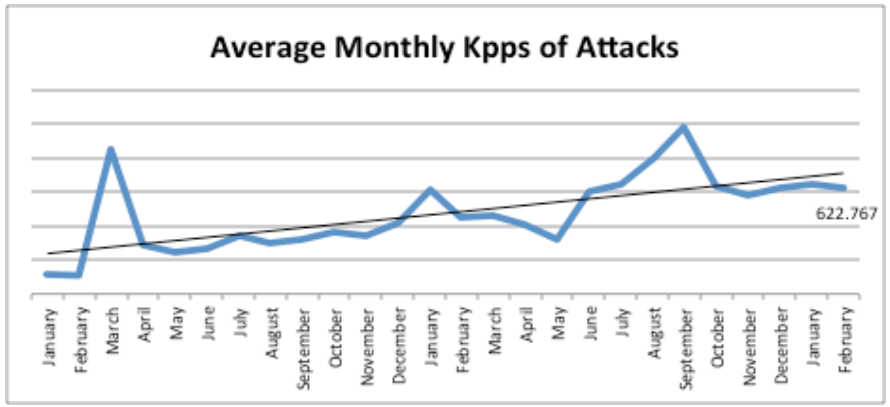
ARBOR NETWORKS

DDoS Data for 2010 – Arbor ATLAS Initiative









ARBOR NETWORKS

Flow Based Detection Techniques

- **Baseline Detection**
  - Detecting shifts in traffic above what is normally seen
  - Catches non standard application/protocol floods, multi-victim attacks, application attacks, changes in GeoIP traffic mix.

- **Misuse (Flood) Detection**
  - Detecting host traffic that exceeds normally accepted Internet behavior
  - Catches common attack vectors like SYN floods, ICMP floods, DNS floods

- **Fingerprint Detection**
  - Detecting known anomalous traffic behaviors indicative of a known threat. Malware detection, specific packet size attacks

ARBOR NETWORKS

Using Flow for DDoS Detection

- **We can 'classify' and 'trace-back' DDoS attacks (and other network events) using the Flow cache on our routers / switches.**
  - Difficult to do pro-active detection.
  - But, no need to export the flow, deploy collectors etc..

```
Demo-Peering-Rtr-1#show ip cache verbose flow
IP packet size distribution (6503M total packets):
   1-32   64   96  128  160  192  224  256  288  320  352  384  416  448  480
   .001 .037 .082 .000 .000 .000 .000 .000 .000 .000 .047 .160 .652 .000 .000

    512  544  576 1024 1536 2048 2560 3072 3584 4096 4608
   .015 .003 .000 .000 .000 .000 .000 .000 .000 .000 .000

IP Flow Switching Cache, 4456704 bytes
  421 active, 65115 inactive, 27650707 added
  305239226 ager polls, 0 flow alloc failures
  Active flows timeout in 1 minutes
  Inactive flows timeout in 15 seconds
IP Sub Flow Cache, 336520 bytes
  421 active, 15963 inactive, 27650707 added, 27650707 added to flow
  0 alloc failures, 0 force free
  1 chunk, 13 chunks added
  last clearing of statistics never
Protocol         Total    Flows   Packets Bytes  Packets Active(Sec) Idle(Sec)
--------         Flows     /Sec     /Flow  /Pkt     /Sec     /Flow     /Flow
TCP-FTP            79      0.0         1    60      0.0       0.0      19.6
TCP-WWW       23928188     11.6        21   287    251.2       3.3      37.9
TCP-SMTP           78      0.0         1    60      0.0       0.0      19.7
TCP-X               1      0.0       241    40      0.0       0.9      18.1
TCP-BGP        394923      0.1         1    48      0.2       2.7      13.6
TCP-Frag            5      0.0       142    40      0.0       0.3      18.5
TCP-other      373964      0.1       176    40     32.0       1.0      17.4
UDP-DNS        121704      0.0       411    61     24.3      55.0       2.5
UDP-NTP             2      0.0         1    76      0.0       0.0      15.3
UDP-other     1977456      0.9      2425   360   2326.5      59.7       0.6
ICMP             7675      0.0         1    77      0.0       0.1      15.8
IPv6INIP       843281      0.4      1263   377    516.7      60.7       0.1
IP-other         2930      0.0      2225    20      3.1       0.5      15.8
Total:       27650286     13.4       235   351   3154.3       9.3      33.3

SrcIf         SrcIPaddress    DstIf         DstIPaddress    Pr TOS Flgs  Pkts
Port Msk AS                   Port Msk AS   NextHop            B/Pk  Active
```

```
Gi3/0.1      172.254.201.154 Gi2/0        3.3.3.3        06 00 02     1
Gi3/0.1      199.100.65.192  Gi2/0        3.3.3.3        06 00 02     1
Gi3/0.1      196.29.143.198  Gi2/0        3.3.3.3        06 00 02  4189
Gi3/0.1      69.229.13.86    Gi2/0        3.3.3.3        06 00 02     1
Gi3/0.1      182.6.36.66     Gi2/0        3.3.3.3        06 00 02  3481
Gi3/0.1      228.101.203.81  Gi2/0        3.3.3.3        06 00 02  3482
Gi3/0.1      123.229.13.86   Gi2/0        3.3.3.3        06 00 02  2937
Gi3/0.1      166.59.246.210  Gi2/0        3.3.3.3        06 00 02  3481
Gi3/0.1      65.152.135.227  Gi2/0        3.3.3.3        06 00 02  1594
Gi3/0.1      13.188.213.198  Gi2/0        3.3.3.3        06 00 02     1
Gi3/0.1      137.245.160.226 Gi2/0        3.3.3.3        06 00 02  5219
Gi3/0.1      130.165.205.228 Gi2/0        3.3.3.3        06 00 02  3503
Gi3/0.1      216.67.56.246   Gi2/0        3.3.3.3        06 00 02     1
Gi3/0.1      166.234.39.115  Gi2/0        3.3.3.3        06 00 02  1741
Gi3/0.1      148.12.123.205  Gi2/0        3.3.3.3        06 00 02  3504
Gi3/0.1      215.81.78.225   Gi2/0        3.3.3.3        06 00 02     1
Gi3/0.1      168.89.28.106   Gi2/0        3.3.3.3        06 00 02  3480
Gi3/0.1      136.103.252.201 Gi2/0        3.3.3.3        06 00 02  4190
Gi3/0.1      209.193.14.202  Gi2/0        3.3.3.3        06 00 02   544
Gi3/0.1      32.247.103.10   Gi2/0        3.3.3.3        06 00 02   546
```

- **TCP Flags field is logical OR of flags seen on all packets matching a flow.**
- **Just SYN indicates a problem.**
- **SYN Flood attack in this case.**

ARBOR NETWORKS

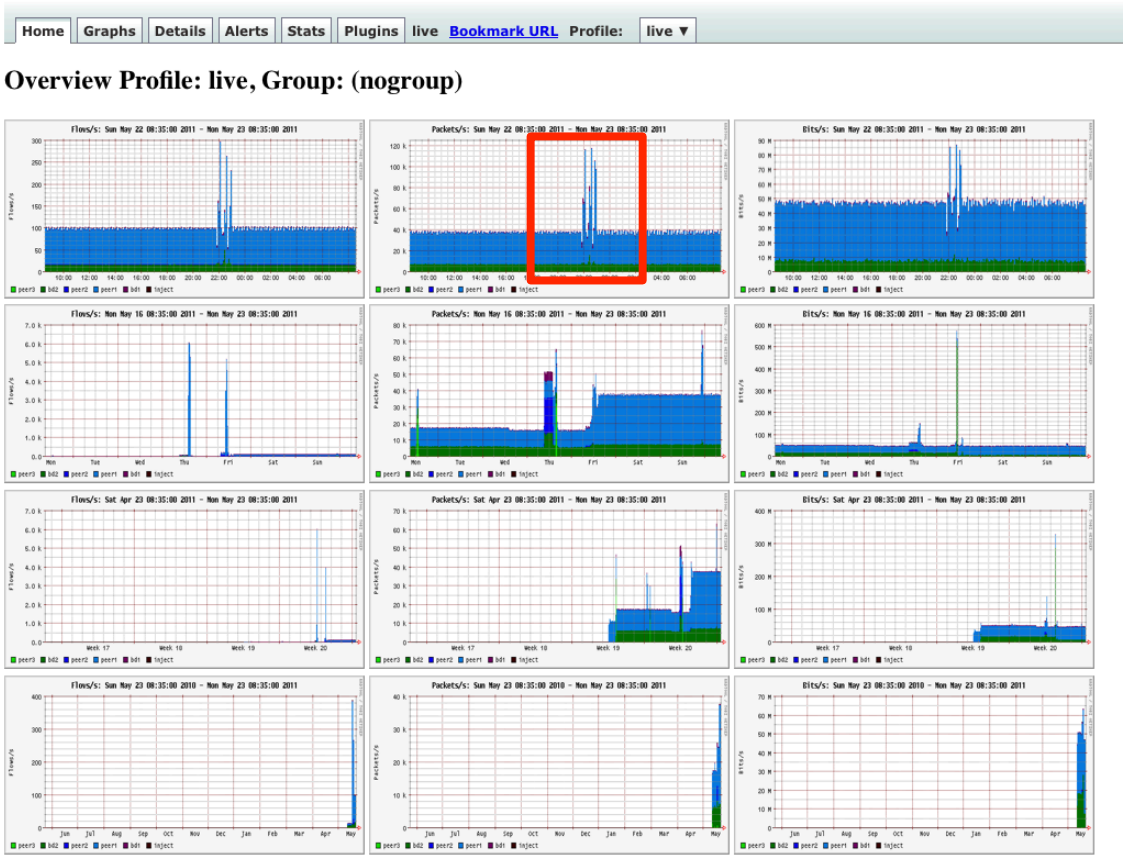
Using Flow for DDoS Detection

- **As with Bot detection we can use open-source tools for DDoS detection**
- **Nfsen can provide a graphical view of traffic.**
  - Many other tools available
  - Establish which routers / infrastructure carry traffic for which customer / service
  - Understand when / where there will be collateral damage.

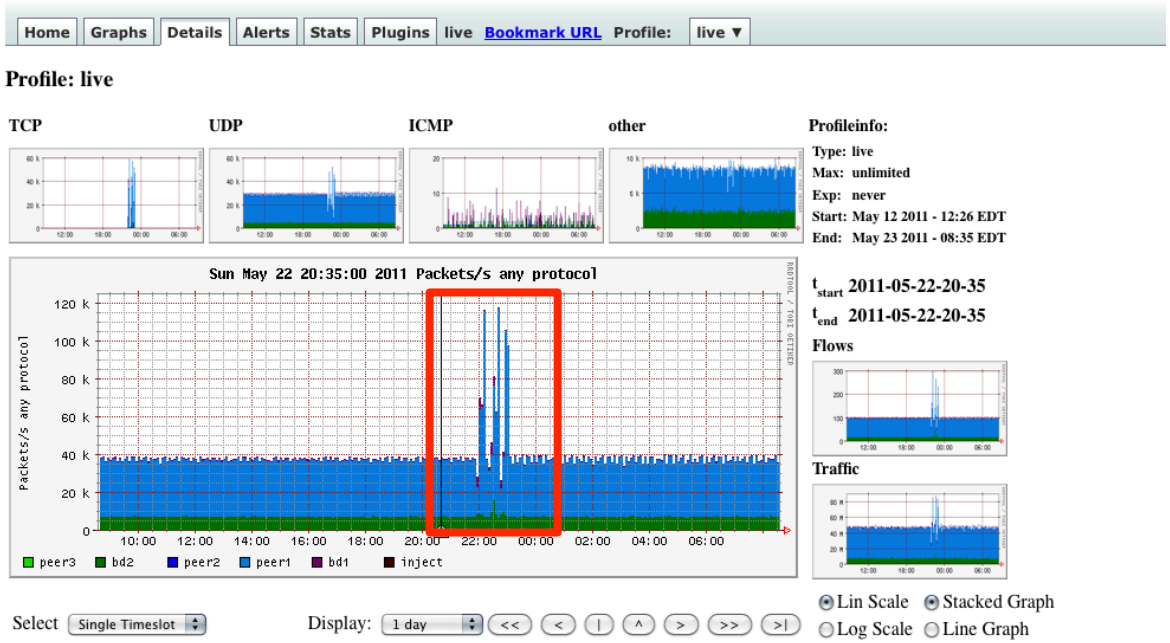
ARBOR NETWORKS

Using Flow for DDoS Detection

- Can use graphs to identify changes in traffic pattern.



Using Flow for DDoS Detection



Statistics timeslot May 22 2011 - 20:35

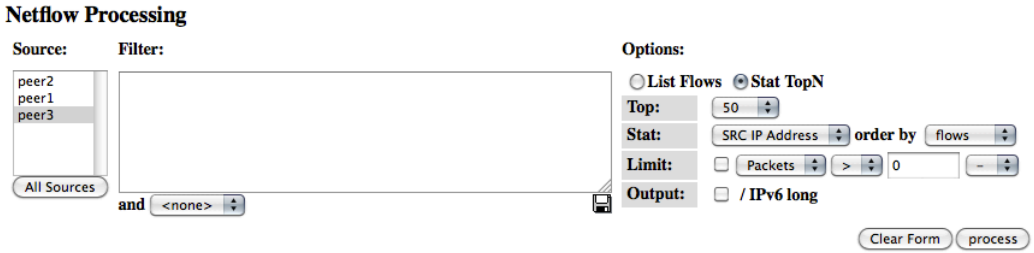
Channel:	Flows:					Packets:					Traffic:				
	all:	tcp:	udp:	icmp:	other:	all:	tcp:	udp:	icmp:	other:	all:	tcp:	udp:	icmp:	other:
inject	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 /s	0 b/s	0 b/s	0 b/s	0 b/s	0 b/s
bd1	0.0 /s	0.0 /s	0.0 /s	0 /s	0 /s	41.0 /s	34.1 /s	6.8 /s	0 /s	0 /s	140.4 kb/s	102.3 kb/s	38.1 kb/s	0 b/s	0 b/s
peer1	79.5 /s	0.4 /s	79.0 /s	0.0 /s	0.2 /s	29.1 k/s	0.6 /s	23.0 k/s	0.0 /s	6.1 k/s	38.5 Mb/s	246.1 b/s	21.3 Mb/s	4.1 b/s	17.1 Mb/s
peer2	0.6 /s	0.3 /s	0.2 /s	0.1 /s	0 /s	1.3 /s	0.5 /s	0.6 /s	0.2 /s	0 /s	835.5 b/s	196.7 b/s	487.5 b/s	151.3 b/s	0 b/s
bd2	17.2 /s	0.1 /s	14.7 /s	0.0 /s	2.4 /s	7.7 k/s	16.2 /s	5.4 k/s	1.7 /s	2.2 k/s	9.3 Mb/s	18.9 kb/s	7.0 Mb/s	559.8 b/s	2.3 Mb/s
peer3	0.0 /s	0.0 /s	0 /s	0 /s	0 /s	1.0 /s	1.0 /s	0 /s	0 /s	0 /s	453.3 b/s	453.3 b/s	0 b/s	0 b/s	0 b/s

All None Display: Sum Rate

ARBOR NETWORKS

Using Flow for DDoS Detection

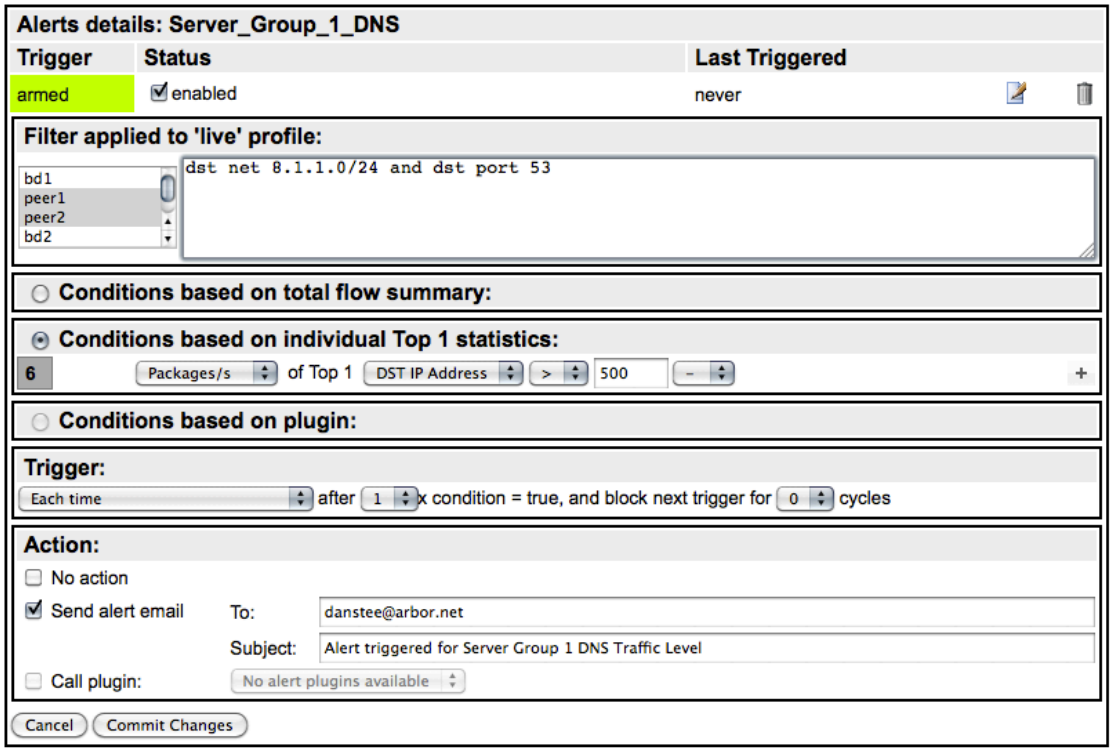
- **Can use nfdump, as before, to classify our traffic.**
- **Flow can provide both Detection and Classification information.**
  - Classification / Trace-Back data needed for mitigation




```
** nfdump -M /data/nfsen/profiles-data/live/peer3 -T  -r 2011/05/16/nfcapd.201105161255 -n 50 -s srcip/flows
nfdump filter:
(( ident peer3) and (
dst net 8.1.1.0/24 and dst port 53
))
Top 50 Src IP Addr ordered by flows:
Date first seen          Duration Proto       Src IP Addr     Flows(%)     Packets(%)       Bytes(%)         pps      bps   bpp
2011-05-16 13:15:40.825     1.071 any       39.144.11.56       2( 0.1)     3000( 0.0)    84000( 0.0)    2801   627450    28
2011-05-16 13:16:40.906     0.879 any      104.170.13.236      2( 0.1)     2500( 0.0)    70000( 0.0)    2844   637087    28
2011-05-16 13:16:40.914     0.875 any       65.42.231.27       2( 0.1)     1900( 0.0)    53200( 0.0)    2171   486400    28
2011-05-16 13:15:40.342     1.547 any      122.210.234.242     2( 0.1)     2800( 0.0)    78400( 0.0)    1809   405429    28
2011-05-16 13:16:40.910     0.877 any       32.139.0.121       2( 0.1)     1600( 0.0)    44800( 0.0)    1824   408665    28
2011-05-16 13:14:40.811     0.775 any      216.102.27.104      2( 0.1)     2200( 0.0)    61600( 0.0)    2838   635870    28
2011-05-16 13:14:40.627     0.634 any       42.182.32.232      2( 0.1)     1800( 0.0)    50400( 0.0)    2839   635962    28
2011-05-16 13:16:40.697     0.937 any       94.67.18.69        2( 0.1)     1500( 0.0)    42000( 0.0)    1600   358591    28
2011-05-16 13:16:40.909     0.878 any      171.171.171.171     2( 0.1)     2000( 0.0)    56000( 0.0)    2277   510250    28
2011-05-16 13:14:40.630     0.634 any       18.43.184.131      2( 0.1)     2200( 0.0)    61600( 0.0)    3470   777287    28
2011-05-16 13:14:40.627     0.634 any      209.59.196.209      2( 0.1)     2100( 0.0)    58800( 0.0)    3312   741955    28
2011-05-16 13:14:40.627     0.634 any       212.22.77.206      2( 0.1)     1800( 0.0)    50400( 0.0)    2839   635962    28
2011-05-16 13:16:40.697     1.091 any       36.0.121.53        2( 0.1)     2000( 0.0)    56000( 0.0)    1833   410632    28
2011-05-16 13:13:40.869     0.854 any      151.252.31.240      2( 0.1)     7100( 0.1)   198800( 0.1)    8313    1.9 M    28
2011-05-16 13:15:40.825     0.845 any      107.150.141.63      2( 0.1)     3100( 0.0)    86800( 0.0)    3668   821775    28
2011-05-16 13:14:40.811     0.933 any       40.220.143.39      2( 0.1)     2200( 0.0)    61600( 0.0)    2357   528188    28
2011-05-16 13:13:40.712     1.011 any      121.156.166.164     2( 0.1)     8000( 0.1)   224000( 0.1)    7912    1.8 M    28
2011-05-16 13:14:40.630     0.635 any      138.188.190.133     2( 0.1)     2400( 0.0)    67200( 0.0)    3779   846614    28
2011-05-16 13:14:40.627     0.634 any      133.70.200.223      2( 0.1)     3400( 0.0)    95200( 0.0)    5362    1.2 M    28
2011-05-16 13:16:40.694     0.940 any      219.115.157.130     2( 0.1)     2200( 0.0)    61600( 0.0)    2340   524255    28
2011-05-16 13:16:40.914     0.875 any       126.61.170.13      2( 0.1)     1200( 0.0)    33600( 0.0)    1371   307200    28
2011-05-16 13:15:40.346     1.547 any      113.173.209.116     2( 0.1)     3100( 0.0)    86800( 0.0)    2003   448868    28
2011-05-16 13:14:40.626     0.634 any      137.13.226.249      2( 0.1)     2100( 0.0)    58800( 0.0)    3312   741955    28
```

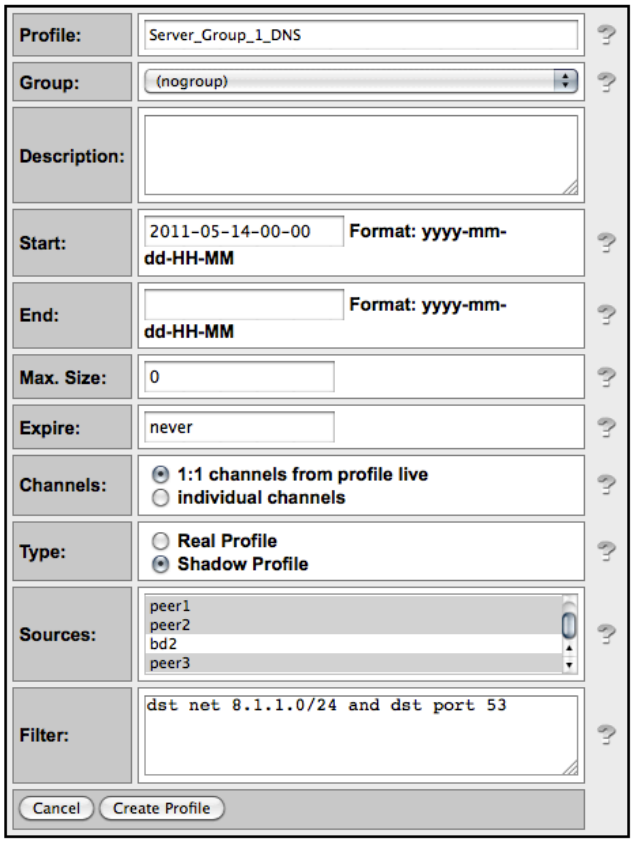
Using Flow for DDoS Detection

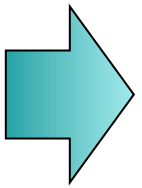
- **Can also use Alerts of pro-active detection of specific traffic**
  - SYN floods, UDP floods etc..
- **Can also use plugins (freely available) which extend this functionality**



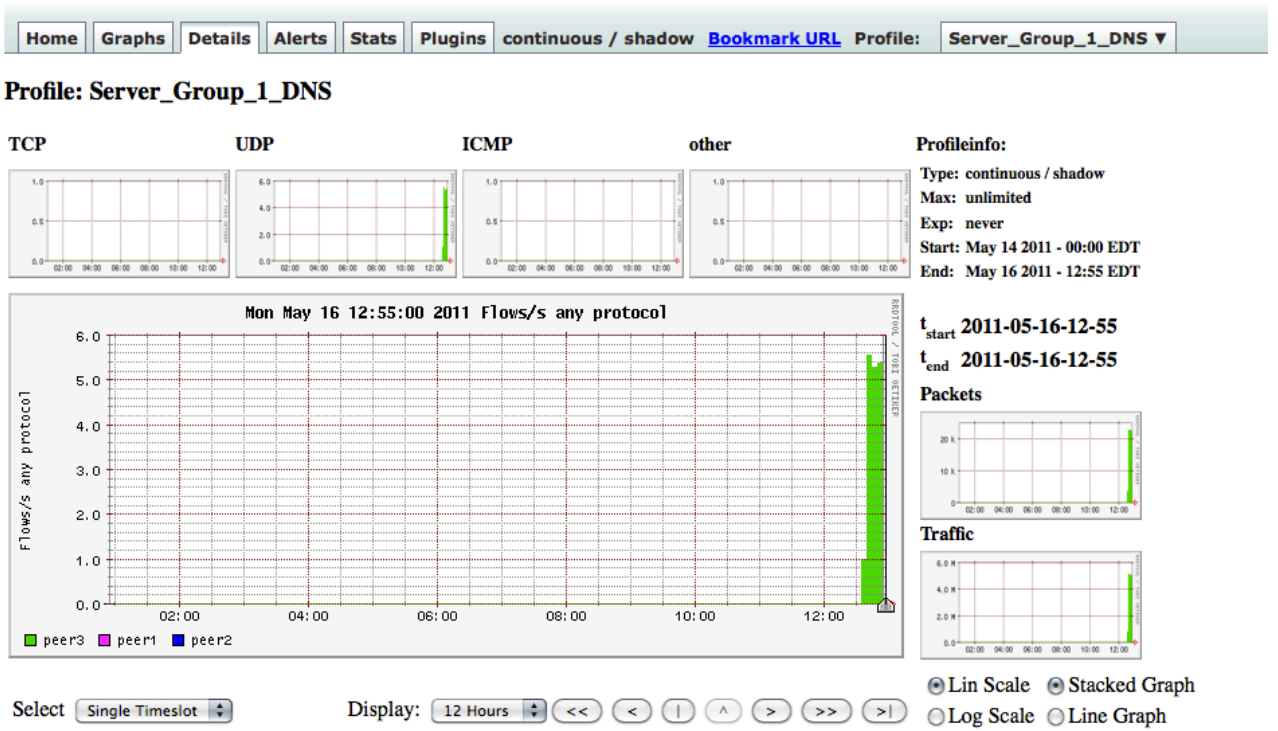
Using Flow for DDoS Detection

- **We can create Profile(s) (retrospectively) to more easily visualize traffic changes.**
  - Can include filters in Profile to zoom in
  - Can help us to trace traffic across the network, visualise which routers are reporting the change.

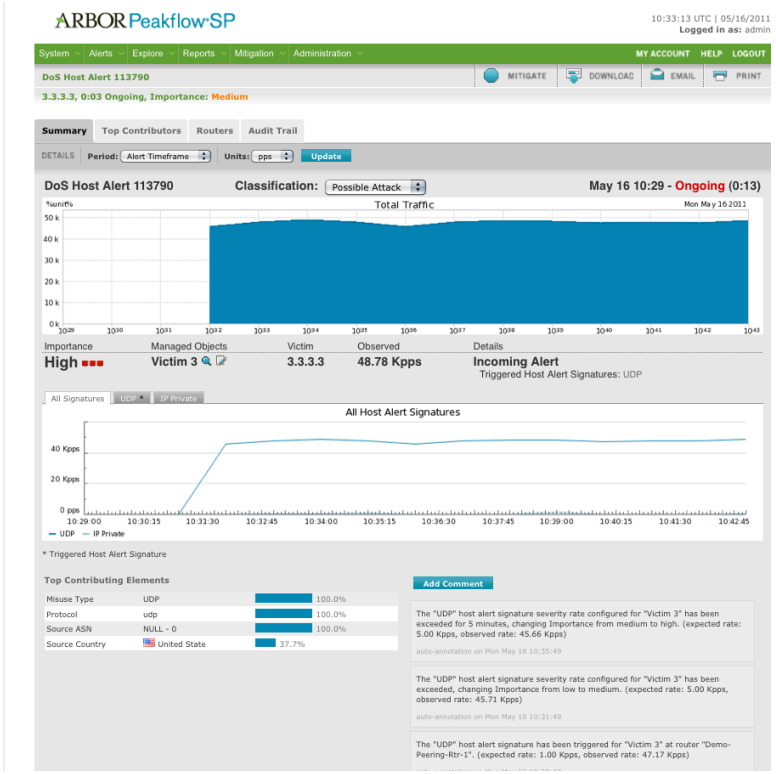


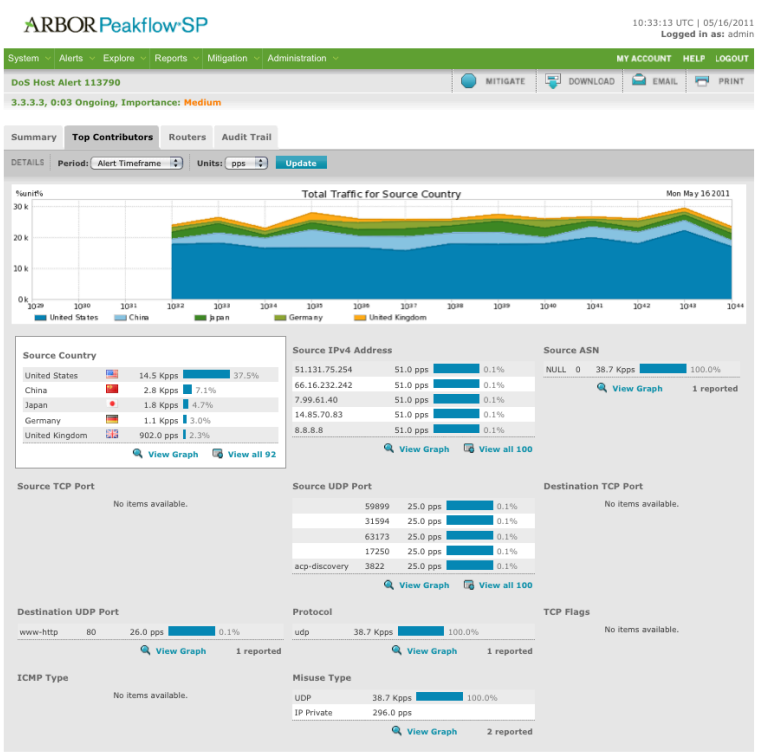


Using Flow for DDoS Detection



Using Flow for DDoS Detection





Using Flow for DDoS Detection

- **DDoS poses a growing service availability risk**
- **Cost-effective and scalable way of detecting and classifying DDoS attacks.**
  - Leverages the functionality available within the routers / switches
  - Can monitor very large traffic volumes, across multiple routers, over an unlimited geographic area.
  - Collection can be centralized or distributed, dependent on scale / processing requirements.
  - Provides pro-active detection, classification and trace-back of events.
- **Not reliant on signatures (zero-day)**
- **Does not introduce additional state into the network**
  - which increases the attack surface.
- **Helps us ensure the availability of our services.**

ARBOR NETWORKS

ARBOR NETWORKS

Thank You

Darren Anstee
darren@arbor.net