Detecting malicious hosts with Darknets

Security Camp
Boston University
10 March 2006

Christopher Misra

UMass statistics

- 25,000 active hosts
- 142 buildings
- All 42 Residential buildings networked
- 12000+ Residence hall connections (port-per-pillow)
- 8000+ Academic building connections
- 1200+ Cisco 24 port Switches
- 8 Cisco core routers
- 600 Off-campus dial-in modem lines
- 500Mb/s - Commodity Internet connections
- 155Mb/s - Internet2 connection
  - Each over leased GigE private fiber

Our Environment

- 2 Enterprises in one: Academic & Residential
- Residential network
  - 12,000 Residential hosts
  - We operate functionally as an ISP
  - However, they are still our students, we need to teach them
- Academic network
  - Traditionally open network
  - Academic freedom
  - Minimal content inspection

Higher education challenges

- Students exploring the boundaries of policy
- Many systems within our enterprise boundary, but not centrally managed.
- Sense of academic freedom occasionally leads students to make poor choices
  - Copyright
  - Hacking/Cracking
  - Wireless everywhere

Boundary Conditions

- Require robust security with limited staffing and budgets
  - More to do than we could hope to accomplish
- Parts of the enterprise are highly de-centralized.
  - Self-supporting
  - Many unmanaged systems

Incident Response

- Given the nature of our network and user base, we have a high proportion of incidents
  - DMCA, Spam, peer-to-peer
- Open network, complex challenges
  - allow most, deny bad
  - botnet hosts and C&C
  - limited edge filtering
  - Large pipes

Incident Response

- Since we don't manage many of the hosts on our campus, we need to find solutions that meet this need.
- Robust help desk to meet the needs of campus users
- How can we automate detection of activity?
  - And leverage help desk to improve security

Detection

- We don’t use any agent based detection on endpoint systems
- Heterogeneous environment
- Majority of these devices are not centrally managed

Netflow

- Originally developed by Cisco
- Versions
  - 5: most frequently in use (what we use)
  - 7,8: Cisco specific
  - 9 : Basis for IETF IPFIX (IP Flow Information Export)
- IPFIX
  - An effort to standardize flow export
  - http://ipfix.doit.wisc.edu/

IPFIX

- Architecture for IP Flow Information Export
  - Architecture for the selective monitoring of IP flows, and for the export of measured IP flow information from an IPFIX device to a collector

http://www.ietf.org/internet-drafts/draft-ietf-ipfix-architecture-09.txt

- Information Model for IP Flow Information Export
  - Protocol for transmitting information related to measured IP traffic over the Internet

http://www.ietf.org/internet-drafts/draft-ietf-ipfix-info-11.txt

IPFIX: Security Analysis and Intrusion Detection

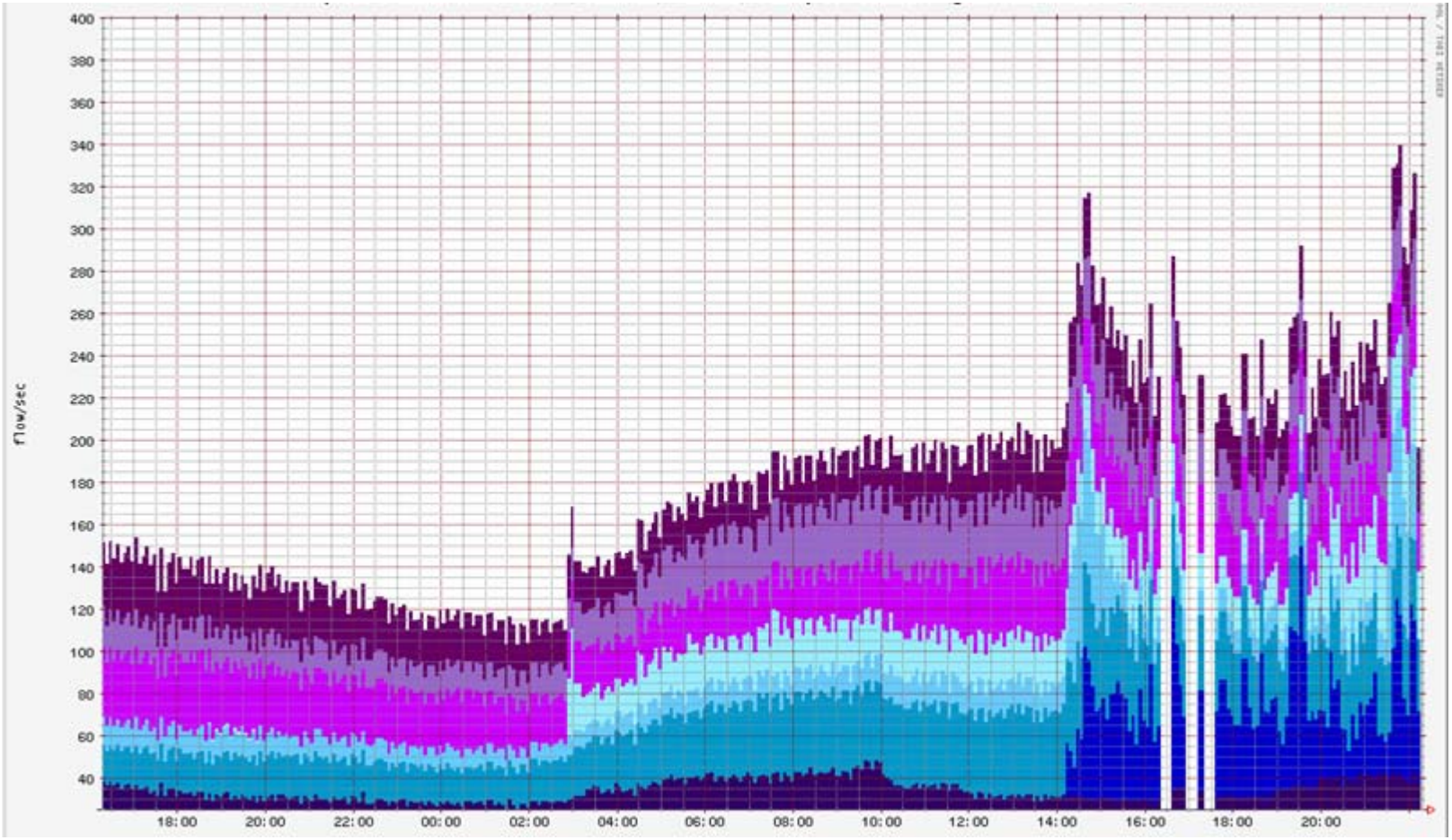
- IPFIX provides information about the traffic in a network. Therefore, it is very well suited to take a key role in the detection of network threats
  - intrusions
  - propagation of viruses and worms
  - port scanning
- Detection
  - visually monitor events (with a console)
  - collect data from sensors (with one or more event collectors)
  - store data from sensors (in a database)

http://www.ietf.org/internet-drafts/draft-ietf-ipfix-as-06.txt

Detection: Netflow

- Visual analysis is often the most efficient detector for us
- Great tool for detecting Denial of Service attacks
  - Prone to data loss under abnormal load
- Useful tool for post-incident analysis
  - Provided the data has not been cycled off the system
- As links become faster, many flow exports are sampled
- Useful for Capacity planning and DoS detection, but of limited use for forensic purposes

Detection: Netflow



Detection: Netflow data

srcIP	dstIP	prot	srcPort	dstPort	octets	packets
80.116.163.85	xxx.yyy.131.204	17	3111	1434	404	1
81.3.162.10	xxx.yyy.131.182	17	1514	1434	404	1
200.74.27.228	xxx.yyy.131.246	6	447	8080	40	1
200.74.27.228	xxx.yyy.131.246	6	64068	80	40	1
200.74.27.228	xxx.yyy.131.246	6	50265	3128	40	1
142.179.169.213	xxx.yyy.131.178	17	1126	1434	404	1
213.60.21.96	xxx.yyy.131.171	17	1923	1434	404	1
212.180.2.68	xxx.yyy.131.114	6	63559	41544	40	1
200.29.164.162	xxx.yyy.131.233	17	1051	1434	404	1
202.103.13.62	xxx.yyy.131.35	6	9001	30185	40	1
213.119.233.63	xxx.yyy.131.7	17	1246	1434	404	1
216.51.150.219	xxx.yyy.131.7	17	1157	1434	404	1
24.112.24.160	xxx.yyy.131.122	17	1129	1434	404	1

Detection: Darknets

- We use NetFlow heavily for diagnostics and capacity planning
- Large campus network requires campus IGP
- Mix together and what do you get?
  - Pointing unused (dark) network space at a flow sensor provide good data with low overhead
  - Non-intrusive sensor

Darknets

- A darknet collector listens to one or more blocks of routed, allocated, but unused IP address space.
- Because the IP space is unused (hence "dark") there should be very little if any legitimate traffic entering the darknet
- ***Team Cymru Darknet Project***
  - http://www.cymru.com/Darknet/index.html

Darknets: Campus IGP

- Large, complex campus network necessitates an IGP
  - Dynamic network provisioning
  - Network retirement
  - Many hands
- We use hold-down (nailed-up) routes anyways
  - Static route at the border to minimize route flapping
  - Pointing our address space to Null0 with a high metric
  - Fail safe

Darknets: Campus IGP

- Originally we static routed unused address space to our darknet
  - Not scalable with many hands in the network
  - Relatively narrow aperture sensor

- Why not inject hold-down routes for unused space to a stub router?
  - And generate these netflow records in one place
  - Doesn't need a lot of horsepower
  - Unused space dynamically falls in to the Darknet

Darknets: Bogons, Martians, and friends

- In addition to aggressive ingress and egress filtering we filter bogons, martians, etc.
- A bogon prefix is a route that should never appear in the Internet routing table.
  - No packet routed over the public Internet should never have a bogon source address
  - Commonly found as the source addresses of DDoS attacks.

http://www.cymru.com/Bogons/index.html

Darknets: Bogons, Martians, and friends

- A martian prefix is a reserved and special use IPv4 prefix
- RFC 1918, RFC 3330
- Filtered by campus ingress/egress filters
  - Also interesting to catch in the darknet.
  - Often result of misconfigurations.

Darknets: Bogons, Martians, and friends

- Other fun ones include:
  - Unused 128.119.0/24 from our /16
  - Dumb viruses love scanning this space first
  - Unused 128.119.255/24
- Some people are fans of allocated but non-routed space filtering
  - Some DoD allocated space, etc.
- These are far more subject to local politics/peculiarities.
- Goal is to increase aperture of sensor without compromising network functionality.

Darknets: Bogons, Martians, and friends

- Static routes on Darknet router(s)
  - Scalability, etc
- Dynamic routes
  - Bogon Route Server Project
  - Peering is conducted over a multihop eBGP peering session.

http://www.cymru.com/BGP/bogon-rs.html

Darknets

- We still filter at our edge appropriately.
  - Best practices
- We also filter inappropriate sources at each subnet router interface
  - Unicast reverse-path forward filtering (Cisco)
- Makes our Darknet far more predictable
  - Signal-to-noise ratio much higher.

Detection: Darknets

- Non-content logging data source
  - Privacy retains high community value
- Non-signature based detection
- Catches malicious activity as well as anomalous traffic
  - Often due to misconfigured/faulty hosts

Detection: So what do we use it for

- Compromised host detection
  - This is our primary use
  - Many viruses still just scan campus subnets
    - *Though this seems to be changing*
- Network Mapping detection
  - Reconnaissance, etc
- Trends
  - There is a lot of garbage out there

Example Darknet Data

- 9 March 2006 20:00-20:59 EST

- 33,472 flow from Non-local sources to Darknet
- Top sources

IPaddr	flows	octets	packets
208.38.28.100	4539	375600	7825
202.97.170.135	3513	276908	6286
221.214.141.34	591	32652	800
202.105.233.25	581	45556	1085
221.1.220.30	508	24384	508

Darknet data: Flows by IP protocol

- 9 March 2006 20:00-20:59 EST

protocol	flows	octets	packets
6	3461	402916	8613
17	1676	354254	2167
1	301	42302	594
2	1	28	1

Darknet data: Flows by destination port

- 9 March 2006 20:00-20:59 EST

port	flows	octets	packets
80	7714	706378	14602
6346	3813	306939	5664
6348	3751	363570	6378
1026	1129	236651	1243
30592	723	43605	765
6349	623	51545	962

Detection

- User is identified based on initial registration information
- Username<-> MAC Address mapping
- MAC address to IP address mapping maintained in a database
  - With some historical information
- All data is considered confidential
  - Restricted access
  - Retained pursuant to policy

Device location

- To perform most isolation, detected device must be located

- Given an IP address we must be able to determine:
  - MAC Address at time of infraction
  - Associated username
  - Switch and port device is currently connected to

Isolation

- Isolation is enforced by changing network devices to limit the access of a non-compliant host.
- This protects the detected host from:
  - additional compromise
  - protects other hosts from it
  - provide a conduit for notifying the responsible individual(s)

Isolation

- Capable of isolating hosts at layer 2 and layer 3
  - Dependent on violation and severity of activity.
- Capable of disconnecting host from network completely
  - Shutting down interface on switch
- Isolation enforced at building edge and perimeter routers
  - Cisco ACLs

Notification

- In most cases, the user is notified by email
  - Even when isolated, users can access campus email services
- Help Desk is notified
  - Within trouble ticketing system
- Status is updated in real time
- In cases of managed computers, isolation may not be performed

Darknet Conclusions

- This data is useful
  - To the point of reconfiguring our stub router (located as a development box in our NOC) to have better availability
  - You don't realize how much you miss it until it is gone.
- Effective sensor with low overhead
  - We leveraged existing infrastructure
- Empirically as good a sensor as some commercial IDS sensors
  - Thought the vendors still dispute this

Automated Darknet Processing

- We currently have some automated processing tools to analyze the data
  - Hits per unit time
  - Number of targets
- Not yet brave enough to have these tools push the button
  - Still a person in the middle to approve isolation action
- Likely we will increase automation
  - Increased trust in validity of sensor

Sharing Darknets

- While we have drastically improved the sensor aperture, it is still scoped to campus address space
  - Difficult to see some trends
  - Early warning only to traffic that hits us
- There's no reason we couldn't look to coordinate our data with others
  - Policy issue not withstanding
- We have engaged in these partnership within closed, vetted, trusted communities of interest

Futures

- As we address these policy issues, can we organize this on a broader scale?
- Addressing these concerns in working group format
- Partnerships with groups like REN-ISAC
  - And with that…
- Questions: After Doug Pearson's presentation…