CONTINUOUS MONITORING: A BIG DATA CHALLENGE

RANDY MARCHANY
VA TECH IT SECURITY OFFICE

(C) MARCHANY 2011

TEAMWORK

Share Victory. Share Defeat.

VT CYBER SECURITY STRATEGY

University has 3 main business processes

- Academic, Administrative, Research

Academic

- Open access needed – THE ISP MODEL

Administrative

- Traditional corporate security model

Research

- Hybrid
  - Open access
  - Restricted research, e.g. ITAR

Must design a strategy that covers all 3 areas

VA TECH IT SECURITY STRATEGY

Based on ISO 27002, NIST 800-53 Standards

Implementing the 20 Critical Controls

Continuous Monitoring based on NIST 800-137

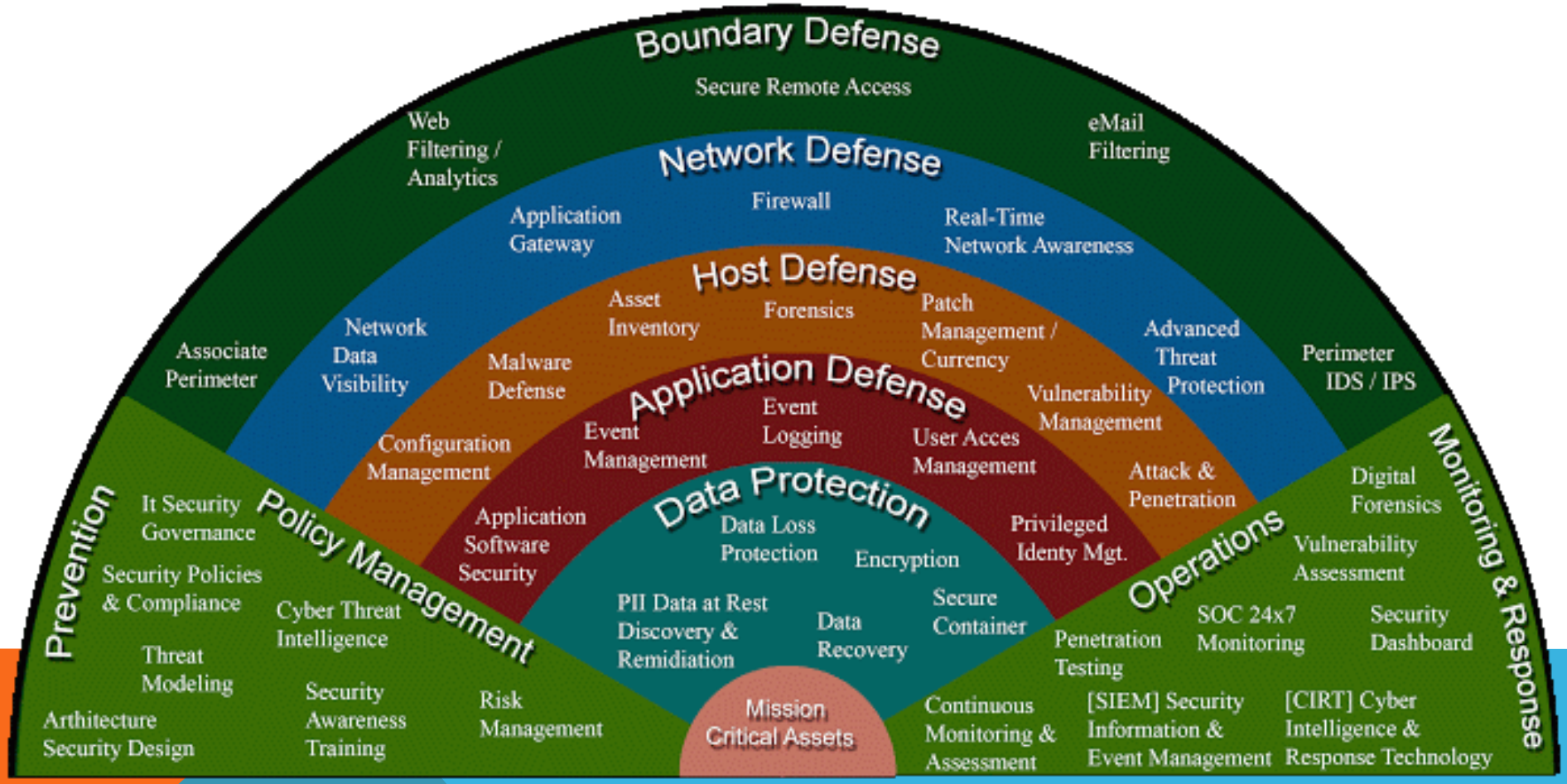
BYOD

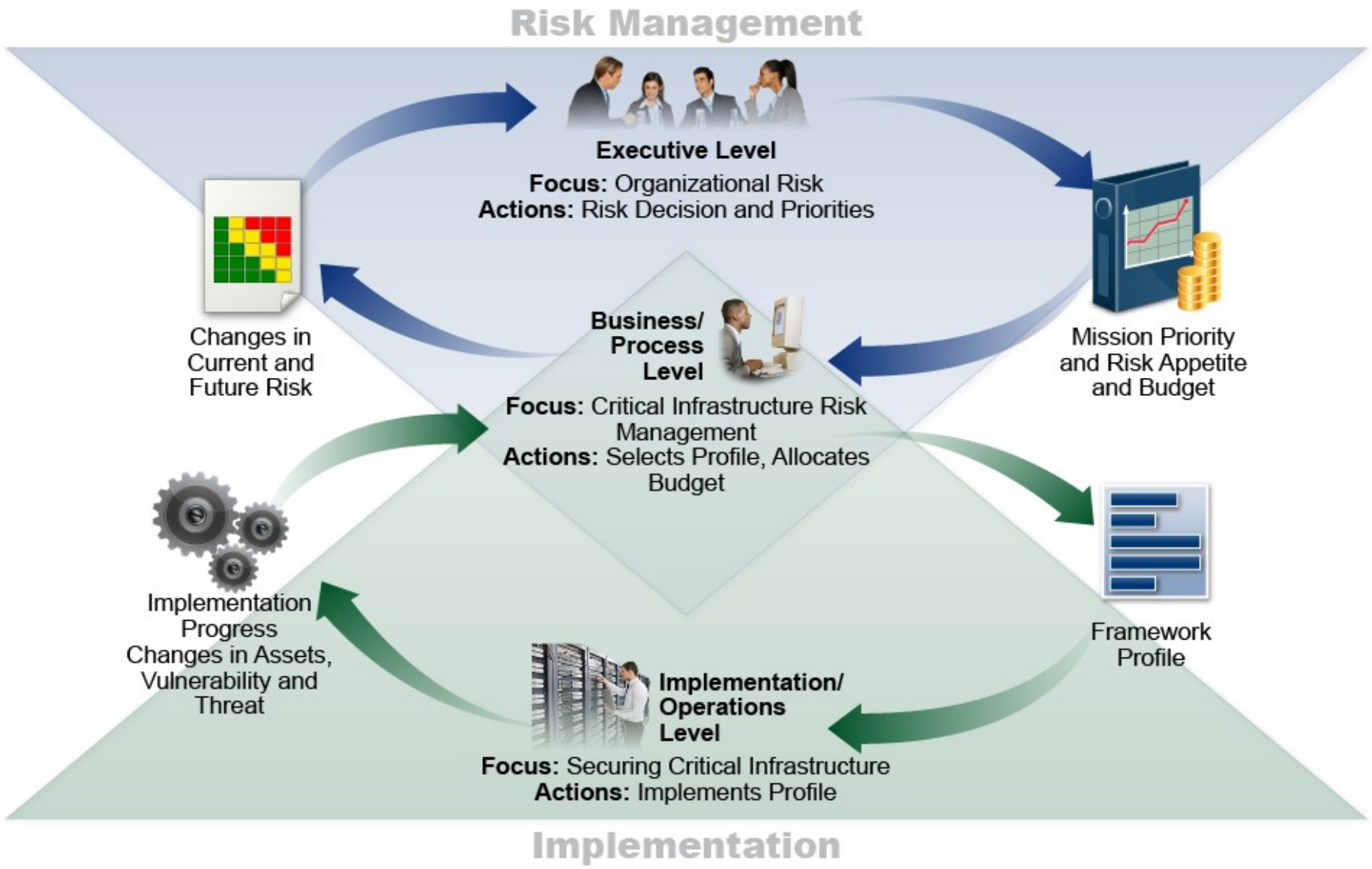
- All students required to purchase their own computers, bring their own smartphones. We've been doing this since 1984

Protect sensitive data regardless of location

Don't care what comes in the net. Worry about what leaves the net.

CONTINUOUS MONITORING – THE DREAM




Risk Management
Executive Level
Focus: Organizational Risk
Actions: Risk Decision and Priorities
Changes in Current and Future Risk
Mission Priority and Risk Appetite and Budget
Business/ Process Level
Focus: Critical Infrastructure Risk Management
Actions: Selects Profile, Allocates Budget
Implementation Progress Changes in Assets, Vulnerability and Threat
Framework Profile
Implementation/ Operations Level
Focus: Securing Critical Infrastructure
Actions: Implements Profile
Implementation

Nobody ever
listens to me....

THE AWFUL TRUTH

Keeping someone from getting inside has failed miserably

Firewalls are not effective PROTECTION devices.

They are effective DETECTION devices

Change the strategy

- Assume they are in so go hunt for the compromised hosts
- Monitor outbound traffic
- Prevent their command and control communication
- Inbound monitors server side attacks; outbound monitors client side attacks

A HORSE BY ANY OTHER NAME

Network Security Monitoring aka Extrusion Detection aka Continuous Monitoring

Use your most powerful defense tool

The Network

Collection

Detection

Analysis

DEVICE DRIVEN VS. DATA DRIVEN

Protecting the device is useless

Protecting the data is key.

It's not a breach until it leaves your network ☺

There are no device breach notification laws ☺

Protect the data so who cares where it's stored

ATTACKER STRATEGY – GET IN, STAY IN, DAMAGE

Initial compromise

- Phishing
- Software vulnerability
- DEFENSE: harden endpoints, patch systems, user awareness

Maintain communication with the mother ship

- command and control of compromised machines
- DEFENSE: Protect root/admin access, identify access patterns

Cause damage

- DEFENSE: interrupt communication channels to bad sites

COLLECTION

INTRUSION VS. EXTRUSION

Intrusion detection is the process of identifying unauthorized activity by inspecting inbound network traffic

Extrusion detection is the process of identifying unauthorized activity by inspecting outbound network traffic

Network forensics is the art of collecting, protecting, analyzing and presenting network traffic to support remediation or prosecution

CM SECURITY PRINCIPLES

Some intruders are smarter than you

Many intruders are unpredictable

Prevention eventually fails

Defensible networks can be watched; they are monitored

Defensible networks limit an intruder's freedom to maneuver; they are controlled

Defensible networks offer a minimum number of services and client-side applications; they are minimized

Defensible networks can be kept current

- source: Extrusion Detection, R. Betjlich

NSM PRINCIPLES

Intruders who communicate with victims can be detected

Detection through sampling is better than no detection at all

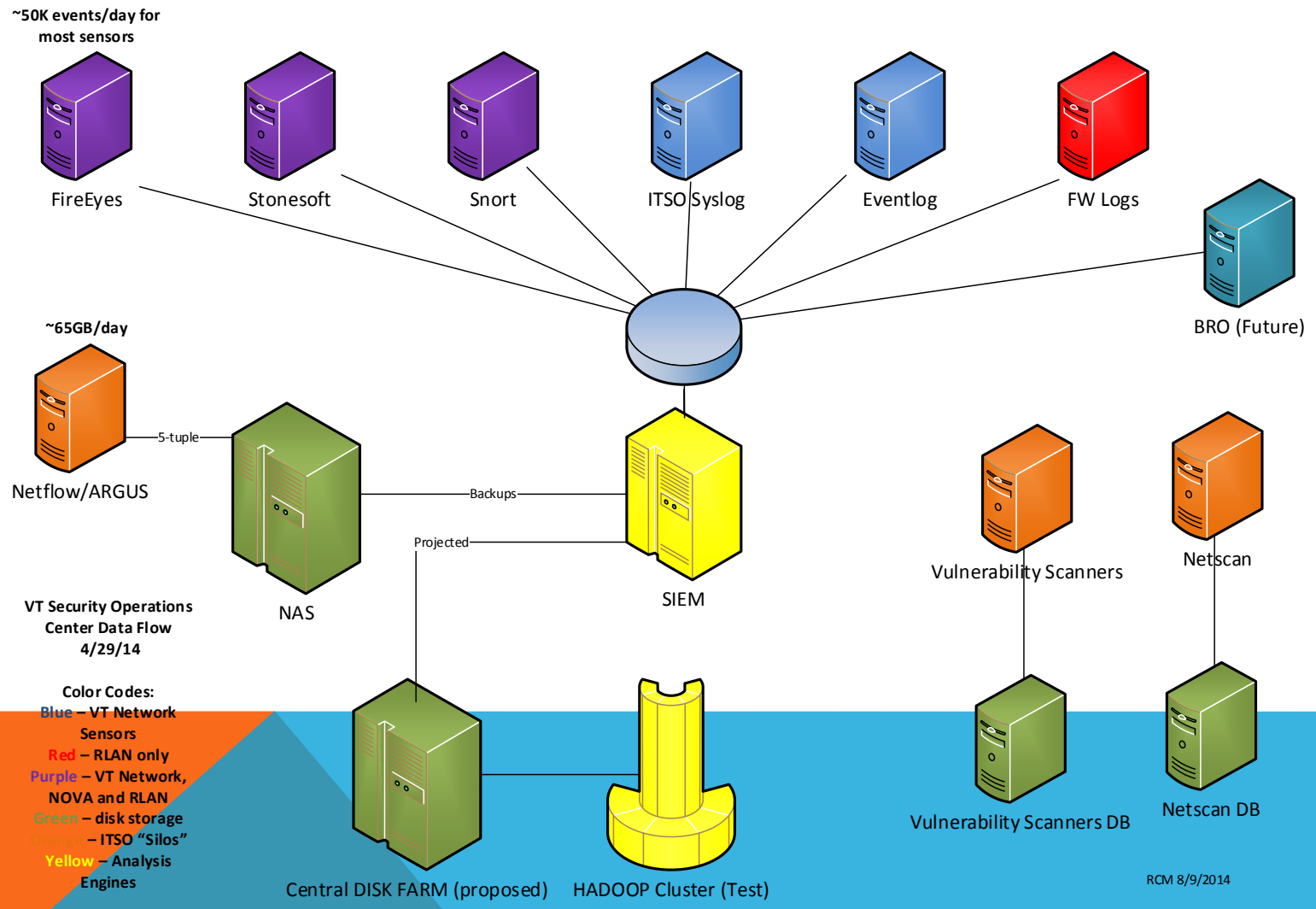
Traffic analysis determines who's talking, for how long and when

It is NOT security event mgt, network based forensics or intrusion prevention

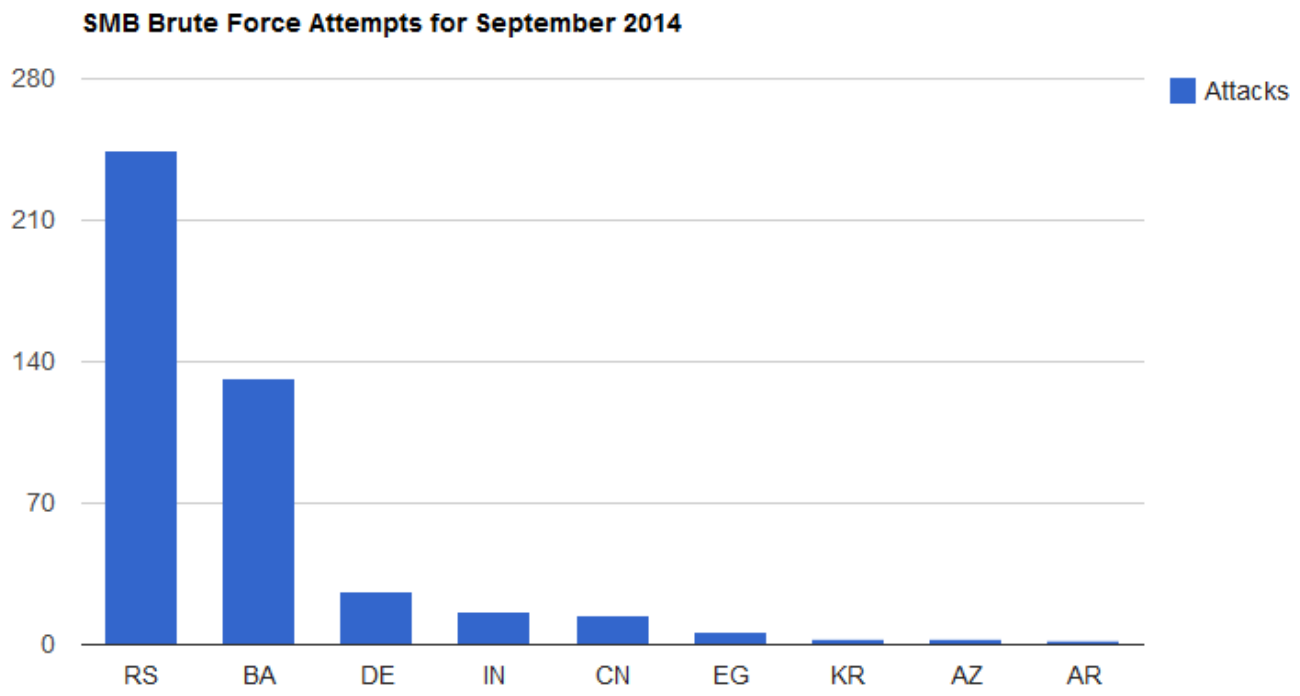
NSM data forms

- Full content data – header and application data
- Session data – aka flows, streams, conversations
- Statistical data – net activity highlighting deviation from norms
- Alert data – judgment by device about an event

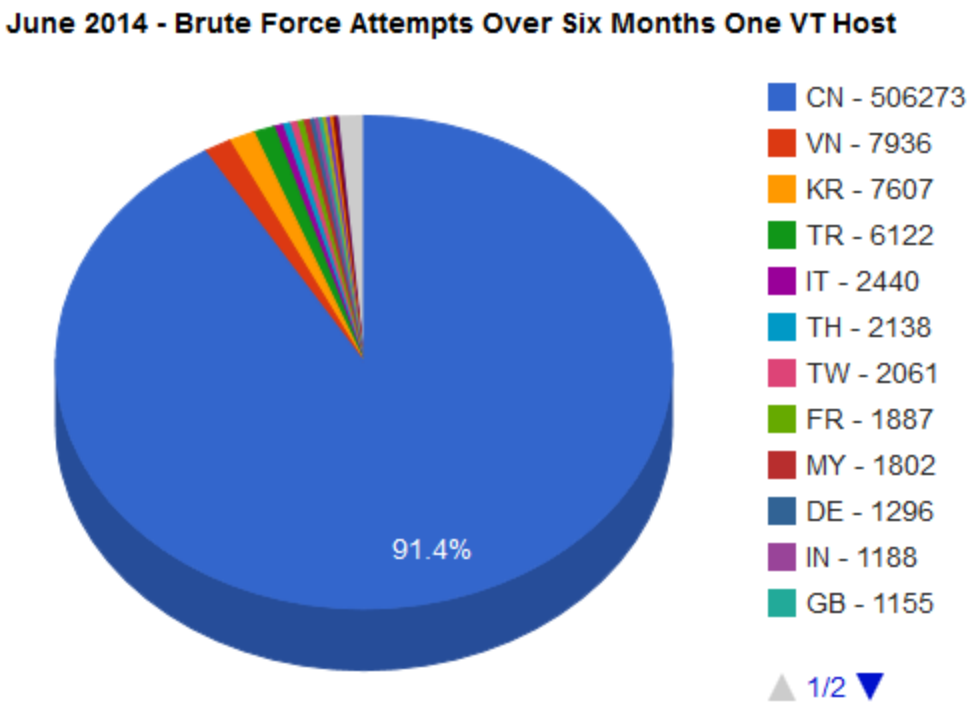
PROPOSED IMPLEMENTATION



DETECTION



US Count: 173705


June 2014 - Brute Force Attempts Over Six Months One VT Host
91.4%
CN - 506273
VN - 7936
KR - 7607
TR - 6122
IT - 2440
TH - 2138
TW - 2061
FR - 1887
MY - 1802
DE - 1296
IN - 1188
GB - 1155
1/2

RESPONSE STEPS TO MALWARE INFECTION

FireEye monitors for suspicious callbacks (COLLECTION)

malware callback detected (DETECTION)

isolate compromised host

PII on the host? Encrypted? Yes – wipe, reinstall. No – potential PII breach

ANALYSIS

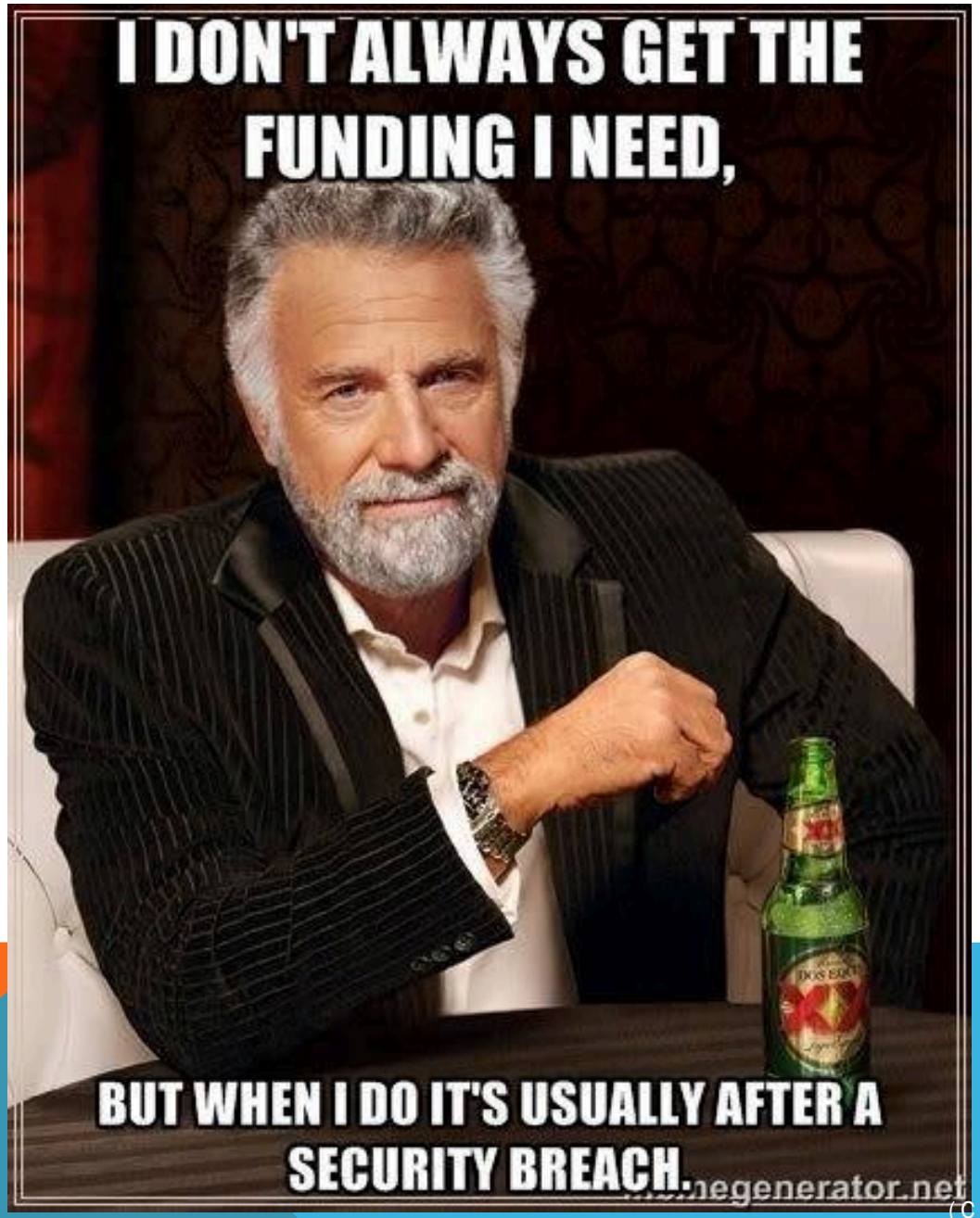
query: "when was the earliest comm between bad domain & compromised host?"

Purpose: try to determine extent of breach

query: "What other VT hosts communicated with bad domain?"

Purpose: determine extent of breach (CONTAINMENT)

Speed is essential


I DON'T ALWAYS GET THE FUNDING I NEED,
BUT WHEN I DO IT'S USUALLY AFTER A SECURITY BREACH.
negenerator.net

ANALYSIS CHALLENGES

Data collection from various sensors

- 75GB netflow data daily
- 6 months of binary Argus data , ~25 fields
- 3-5 months CSV SQL data – structured DB, 7 fields
- 6 months CSV Hadoop – flat file 19 fields
- Snort sensor, IPS, IDS, host based firewall

Big data analysis techniques needed

Organizing database

Optimizing search techniques -

Vendor solutions aren't enough

ANALYSIS: THE DATABASES

Flat files

Traditional databases - Oracle, Postgres, SQL

- not designed to deal with log data
- Designed for data integrity not performance

No SQL (Hadoop)

The right data system

- A function of the volume of data stored, the type of data and who's analyzing it
- Analysis requires repetitive searches

ANALYSIS – RDBMS VS. LOG DATA

Ensures data integrity

Constant data queries so the system spends a lot of time tracking transactions

Users will asynchronously update existing contents

Log data doesn't change

An event occurs, the log data is never updated

Sensors are the only things that write to disk.

Users only read from disk.

EXAMPLE: FLOW ANALYSIS

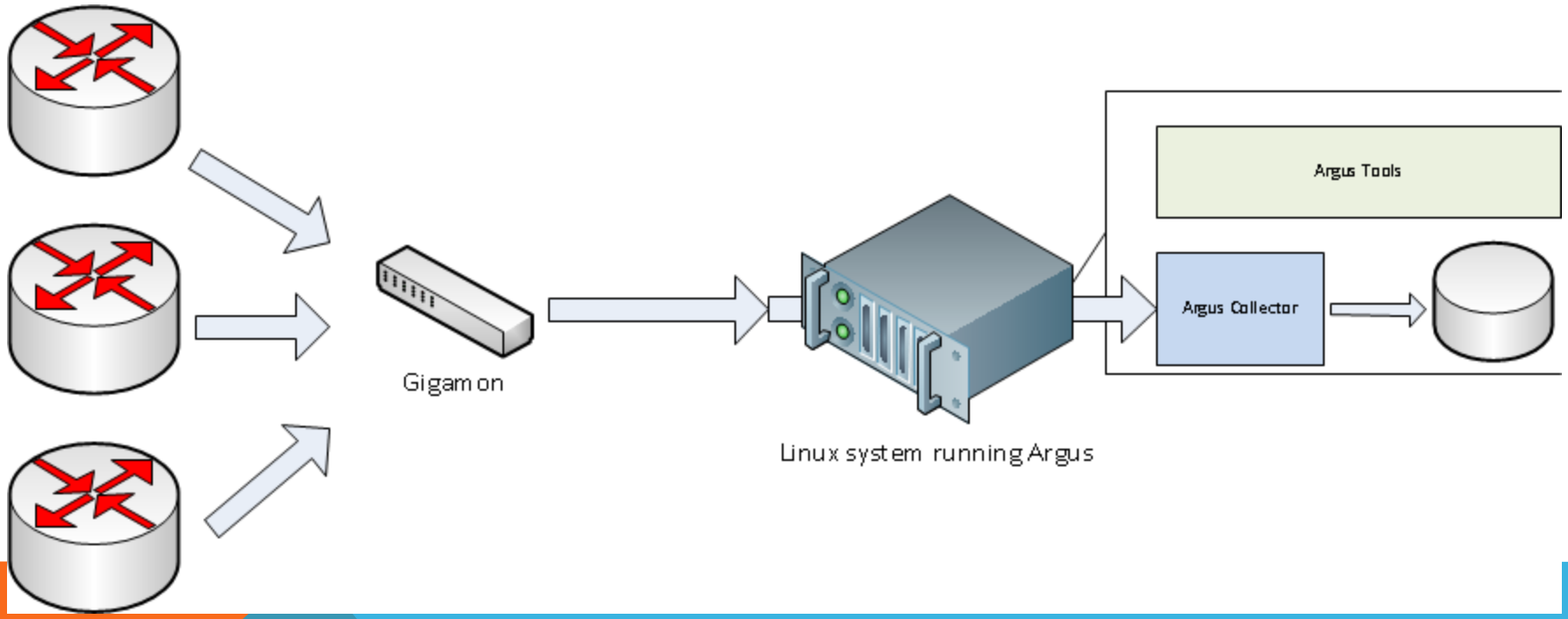
Network flows allow us to see a machine's traffic history.

At VT we use Argus.

- Pros:
  - Open source (FREE)
  - Well documented
  - Great toolset
- Cons
  - Runs off span port, not router flows*
  - Steep learning curve

argus

ARGUS


Argus Tools
Argus Collector
Gigamon
Linux system running Argus

FLOW ANALYSIS – ALL FLOWS

How many phone home attempts were there?

```
ll@wheaties:/opt/incidents/RICH-TEST-INCIDENT$ ra -r 198.82.[redacted]-2011-09-27.ra -L0 - host 84.247.61.32
       StartTime      Flgs  Proto            SrcAddr  Sport   Dir            DstAddr  Dport  TotPkts   TotBytes State
 12:40:23.517995  e           tcp     198.82.[redacted].3551      ->       84.247.61.32.www          4        651   CON
 12:40:45.114974  e           tcp     198.82.[redacted].3614      ?>       84.247.61.32.www          1        391   CON
 12:40:55.177561  e           tcp     198.82.[redacted].3614     <?>       84.247.61.32.www         14      11782   CON
 12:41:03.943470  e r         tcp     198.82.[redacted].3614     <?>       84.247.61.32.www         10      10168   CON
 12:41:10.215085  e r         tcp     198.82.[redacted].3614     <?>       84.247.61.32.www          6       7190   CON
 12:41:16.112838  e d         tcp     198.82.[redacted].3614     <?>       84.247.61.32.www         31      27282   CON
 12:41:22.850172  e d         tcp     198.82.[redacted].3614     <?>       84.247.61.32.www         27      20152   CON
 12:41:30.016314  e d         tcp     198.82.[redacted].3614     <?>       84.247.61.32.www          6       5836   CON
 12:41:35.286870  e d         tcp     198.82.[redacted].3614     <?>       84.247.61.32.www         32      29130   CON
 12:41:40.304984  e           tcp     198.82.[redacted].3614     <?>       84.247.61.32.www          4       1606   CON
 12:41:46.791524  e d         tcp     198.82.[redacted].3614     <?>       84.247.61.32.www         17       9246   CON
 12:41:53.834717  e d         tcp     198.82.[redacted].3614     <?>       84.247.61.32.www         24      20986   CON
 12:41:59.215328  e           tcp     198.82.[redacted].3614     <?>       84.247.61.32.www         23      20518   CON
 12:42:05.124440  e d         tcp     198.82.[redacted].3614     <?>       84.247.61.32.www         23      20522   CON
 12:42:10.293157  e d         tcp     198.82.[redacted].3614     <?>       84.247.61.32.www         17      13326   CON
 12:42:18.084731  e d         tcp     198.82.[redacted].3614     <?        84.247.61.32.www          4       5704   CON
 12:42:32.223575  e d         tcp     198.82.[redacted].3614     <?>       84.247.61.32.www         41      32560   CON
 12:42:37.362093  e d         tcp     198.82.[redacted].3614     <?>       84.247.61.32.www         21      14950   CON
 12:42:42.917966  e d         tcp     198.82.[redacted].3614     <?>       84.247.61.32.www         19      16190   CON
 12:43:04.632456  e d         tcp     198.82.[redacted].3614     <?>       84.247.61.32.www          7       5896   CON
 12:43:25.097600  e           tcp     198.82.[redacted].3614      ?>       84.247.61.32.www          1         66   CON
 12:44:04.998[redacted]7  e           tcp     198.82.[redacted].3614     <?        84.247.61.32.www          1       1426   CON
 12:45:24.[redacted]186  e           tcp       84.247.61.32.www       ?>     198.82.[redacted].3614          1       1426   CON
 12:47:2[redacted]438346  e s         tcp       84.247.61.32.www      <?>     198.82.[redacted].3614         36      33622   CON
 12:47:31.738497  e i         tcp       84.247.61.32.www      <?>     198.82.[redacted].3614          7       7256   CON
 12:47:41.918354  e s         tcp       84.247.61.32.www      <?>     198.82.[redacted].3614         23      20542   CON
```

VA SCAN 2014 - Marchany

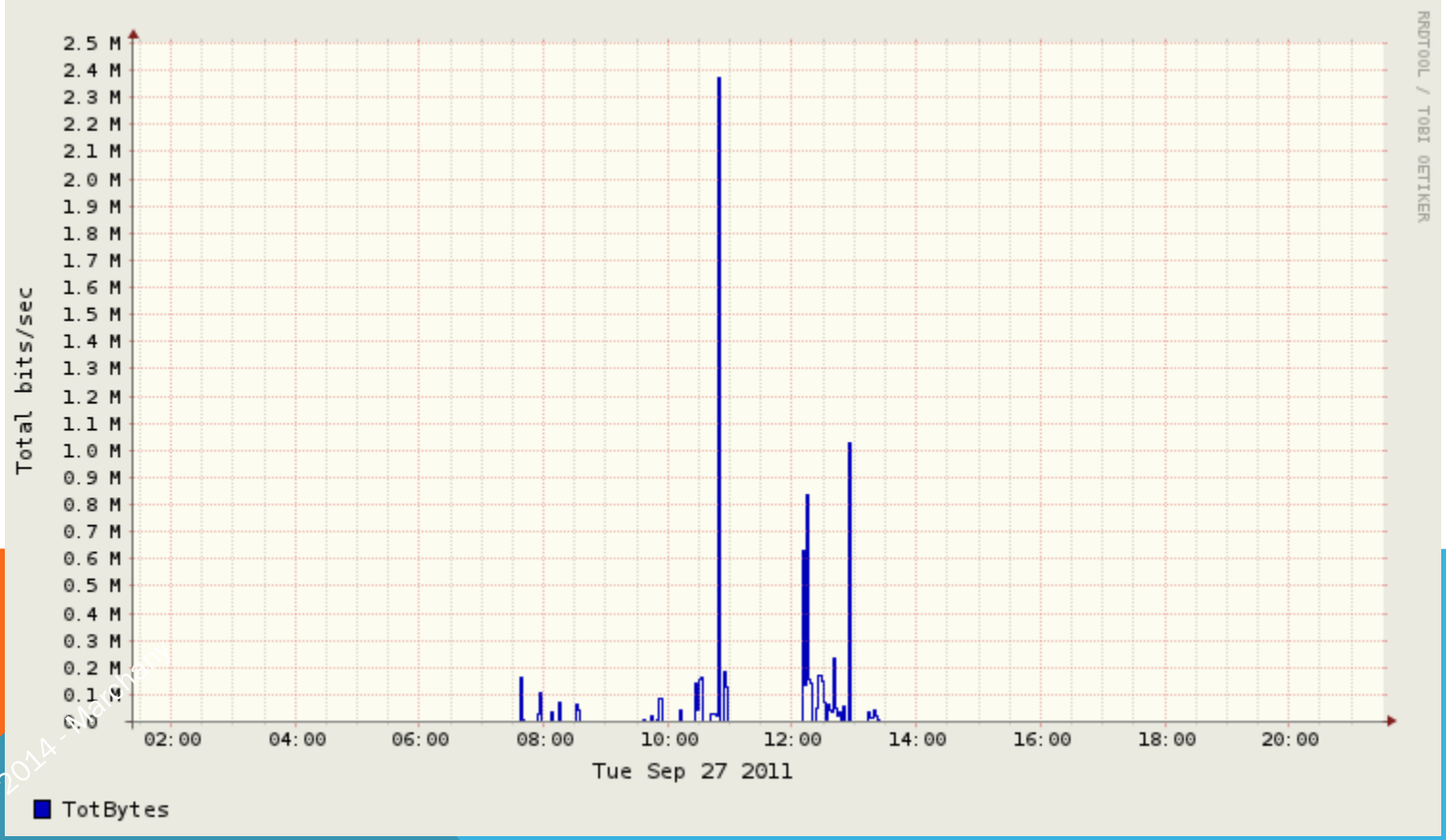
FLOW ANALYSIS – TOP TALKERS

Who were the top talkers?

```
will@wheaties:/opt/incidents/RICH-TEST-INCIDENT$ /opt/scripts/reporting/toptalkers.report 198.82.1[illegible]-2011-09-27.ra
                                                                  43177556    42329
                                          cds82.sea9.msecn.net   17925333    15960
                                          cds45.sea9.msecn.net   14844152    13383
                         168-215-117-34.static.twtelecom.net    4333900     4545
                                         cds171.sea9.msecn.net    3060458     2859
                a72-247-242-237.deploy.akamaitechnologies.com    2864955     3232
                                              crl2.entrust.net    2019862     2272
                                        solutions.sciquest.com    1713995     1828
                                                205.128.76.126    1512386     1964
                                                 cip.vtcrc.com    1395273     1387
                                          cds18.sea9.msecn.net    1111249     1018
                                         cds117.sea9.msecn.net    1061034      986
                              u15311009.onlinehome-server.com     875976      840
                                             bravo.gendns1.com     852143      813
                a72-247-242-238.deploy.akamaitechnologies.com     810180      958
                                                   72.21.91.41     757878      691
                               cdn-208-111-160-6.iad.llnw.net     682859      719
                                   iad04s01-in-f148.1e100.net     682820      814
                             cdn-208-111-161-252.iad.llnw.net     660856      579
                                                 208.71.123.86     594682      667
                   a23-2-12-132.deploy.akamaitechnologies.com     593180      736
                                                74.208.123.127     592815      550
                                          cds12.sea9.msecn.net     505670      508
                                                93.184.216.169     505279      501
                  a72-246-30-18.deploy.akamaitechnologies.com     479920      826
                                                 208.254.9.227     473130      629
                                           tiger.host-care.com     455784      587
                                          fcds141.iad.llnw.net     419320      432
                             dedicated515.inmotionhosting.com     396554      400
                                   iad04s01-in-f149.1e100.net     379349      499
                                         cds163.sea9.msecn.net     369654      349
```

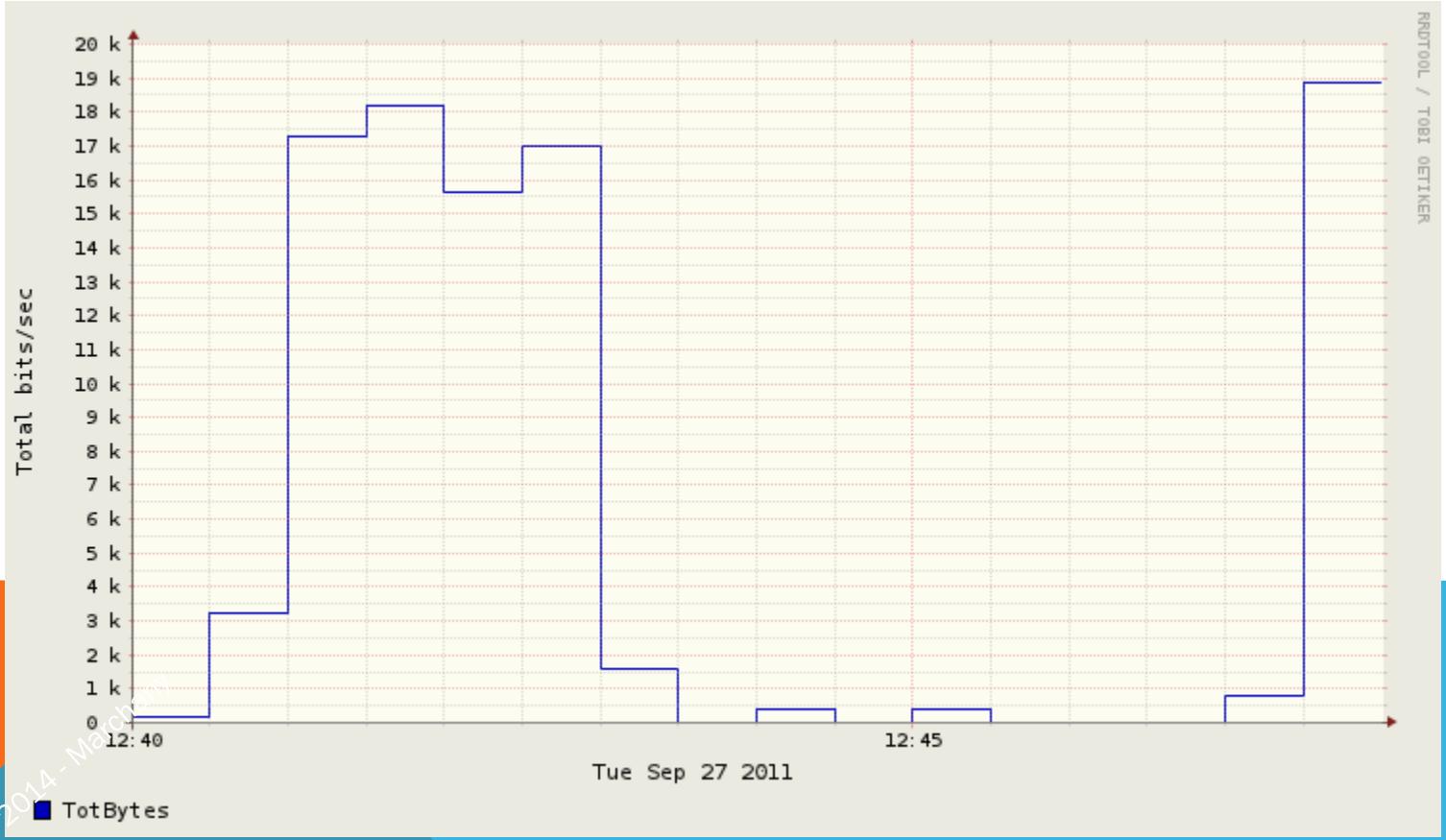
FLOW ANALYSIS – TRAFFIC GRAPHS

Are there any traffic patterns?



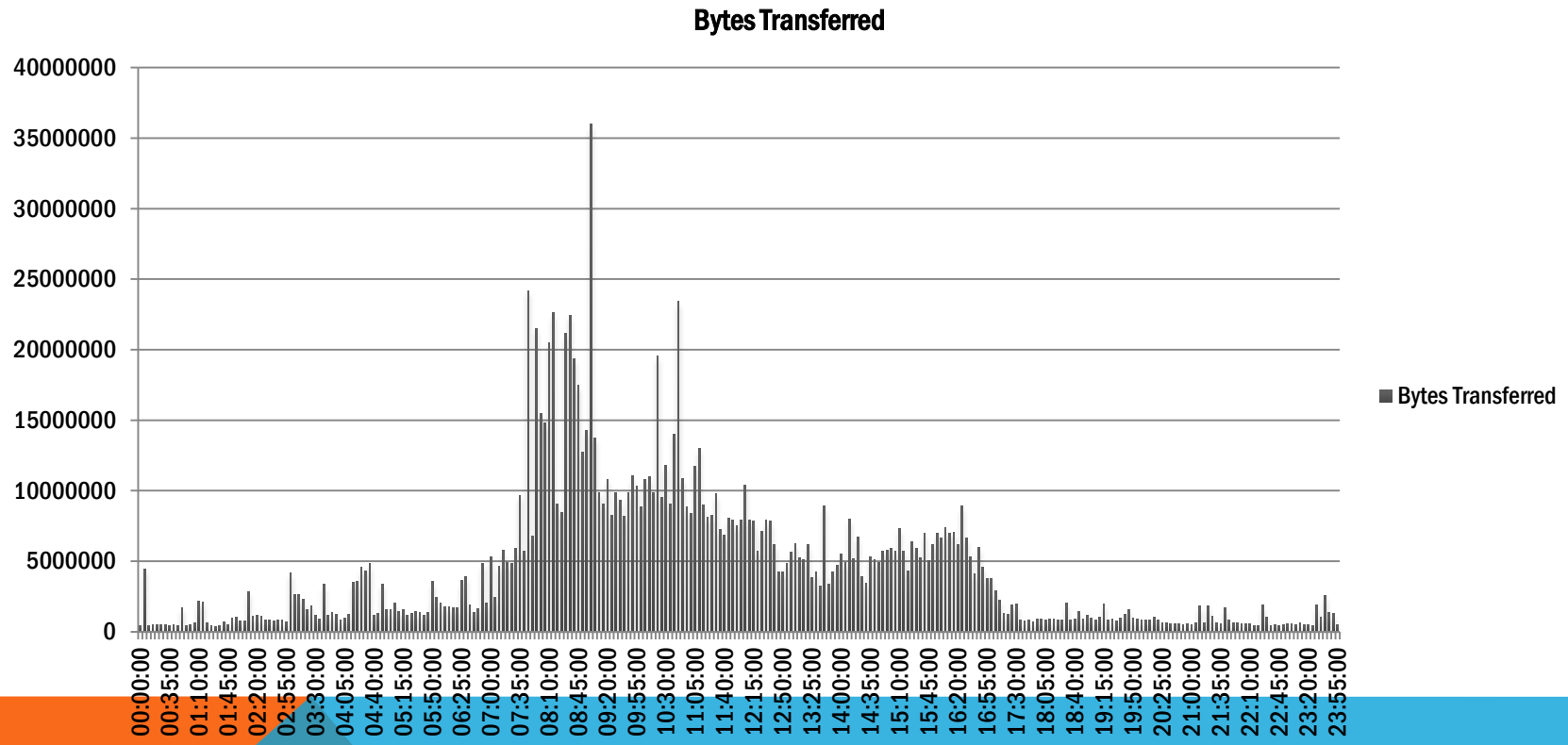
FLOW ANALYSIS – C&C GRAPHS

How much data was transferred to the C&C server?

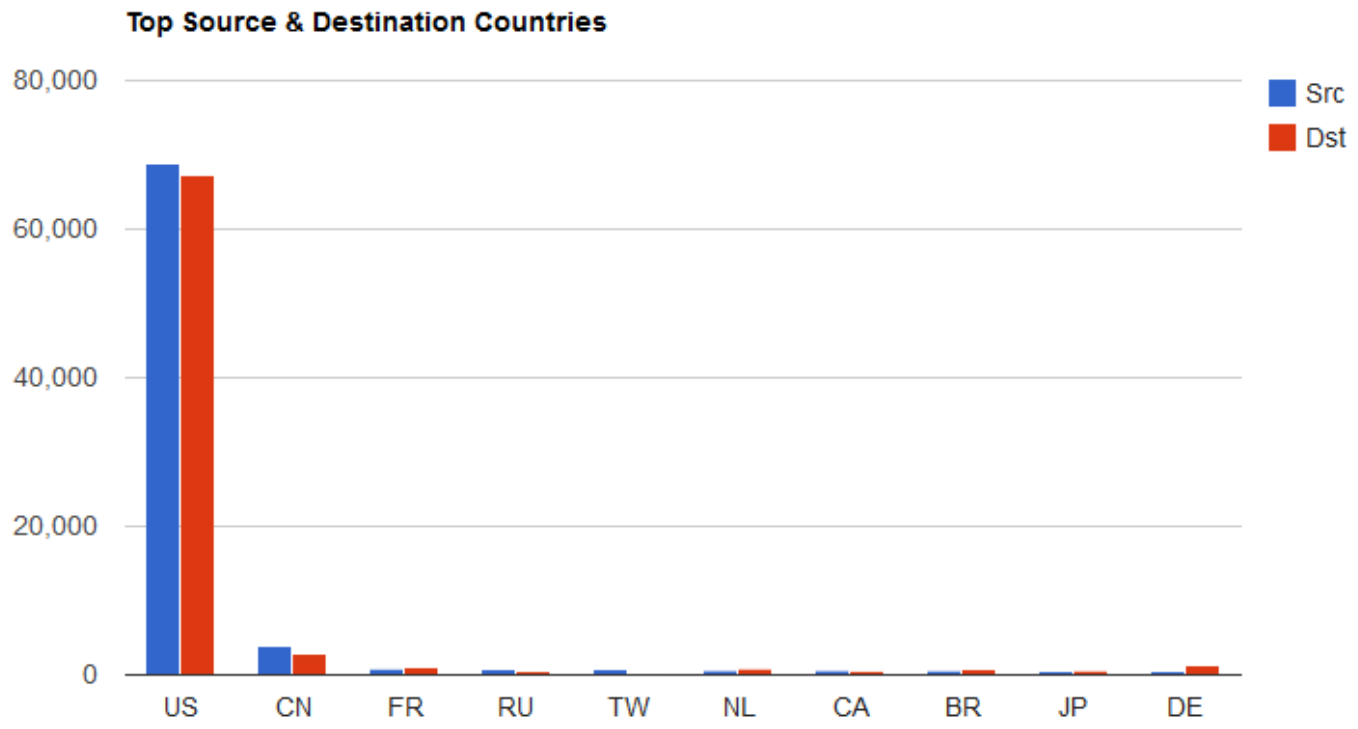


BANDWIDTH ANALYSIS

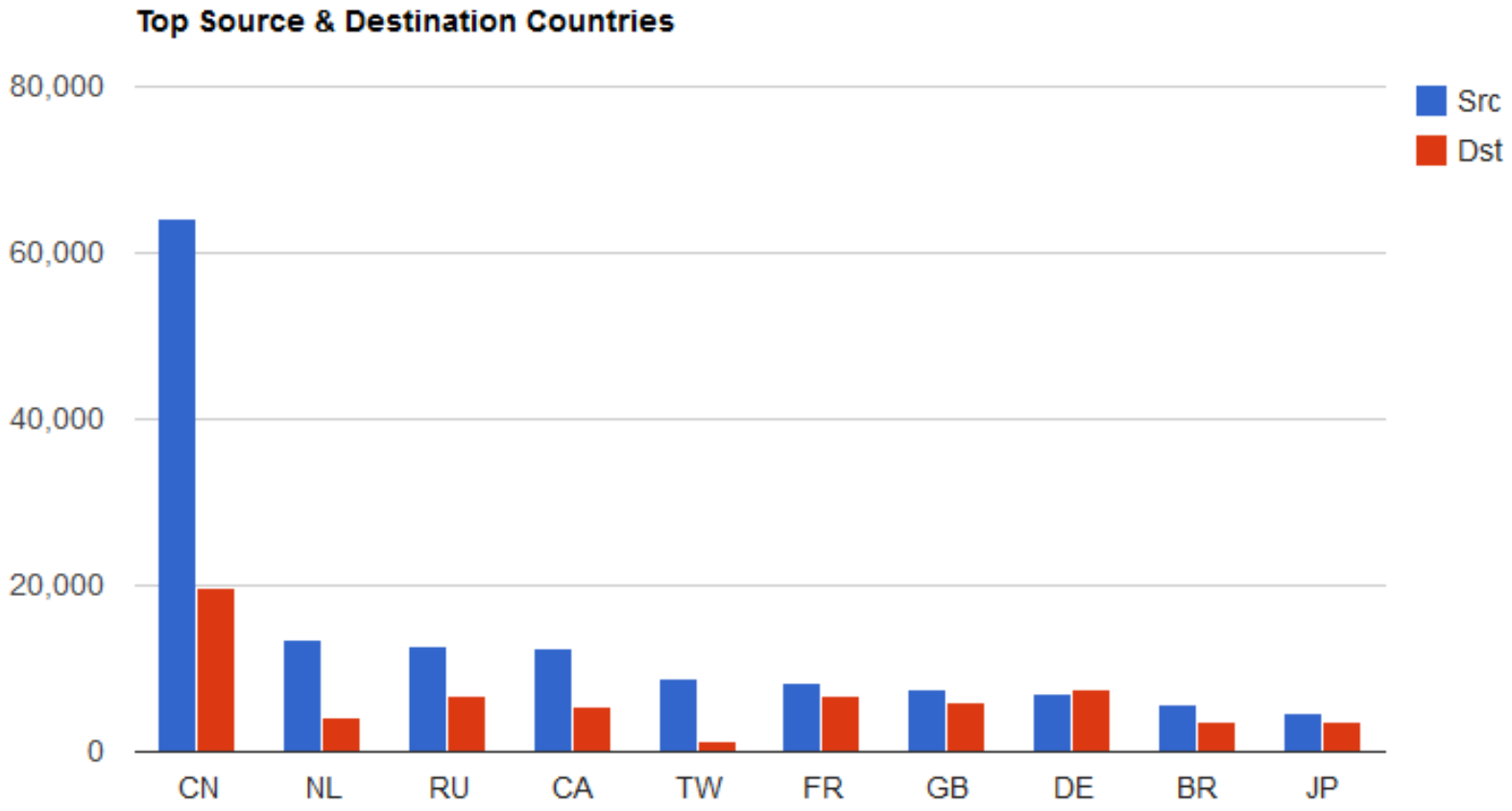
Argus observes and records traffic statistics for specific networks (departments)



EXAMPLE: WHERE DOES DATA GO?

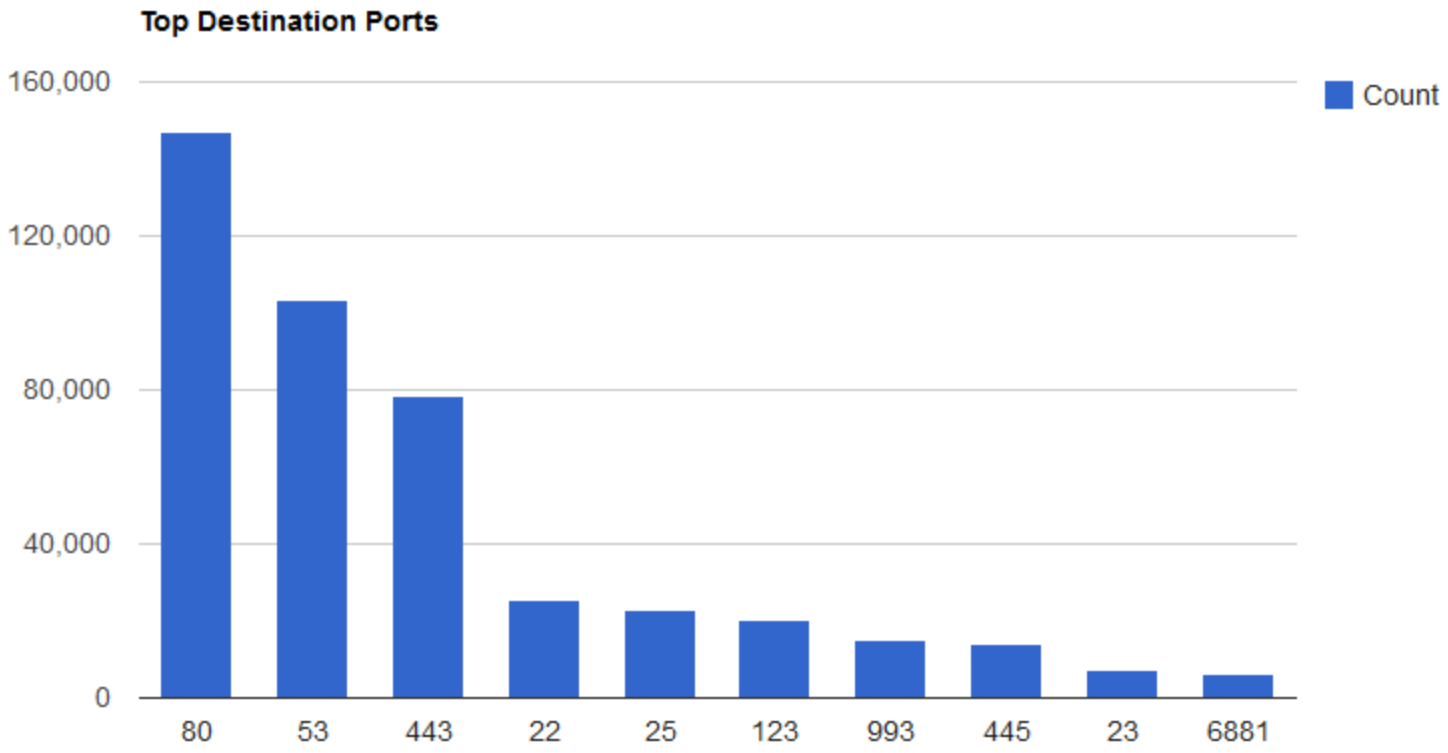


Souce - ITSO Argus Data September 2014



Souce - ITSO Argus Data Past 5 Months

EXAMPLE: TOP PORTS USED



Source ITSO Argus Data Past 5 Months

PARTNERS? SECURITY & NETWORK GROUPS

ITSO and Network mgt groups must closelycoordinate activities

CM "disruptors"

- MPLS
- High speed (>10GB) backbones
- SSL, application layer encryption

Span ports

- 10GB span on a 100GB backbone

These will blind CM sensors

Need to work within these constraints

THOUGHTS ON METRICS

Do you have this capability?

Forensic

- How long does something take to occur?
- How long has it been happening?
- How widespread is it?

Ratio

- What is leaving our net?
- Compromised hosts vs. total hosts

SECURITY METRICS

Executive Level

- Compliance with University Policies and Federal, State regulations
- Summary graphs

Operational

- Vulnerability scans, IDS/IPS data collection

Incident Handling

- Volume, type, severity

http://www.educause.edu/library/security-metrics

FUTURES? THE GOOD

Continuous monitoring increases the chances of successful detection and mitigation

Can tailor defenses based on critical assets

- find repeat offenders
- create targeted training

Provide justification for IT improvement

- Department and Enterprise staff, hardware, software, training

Enhance university research

- work with researchers by providing anonymized raw data

FUTURES – THE BAD

Edward Snowden effect

Must take significant precautions to safeguard collected data

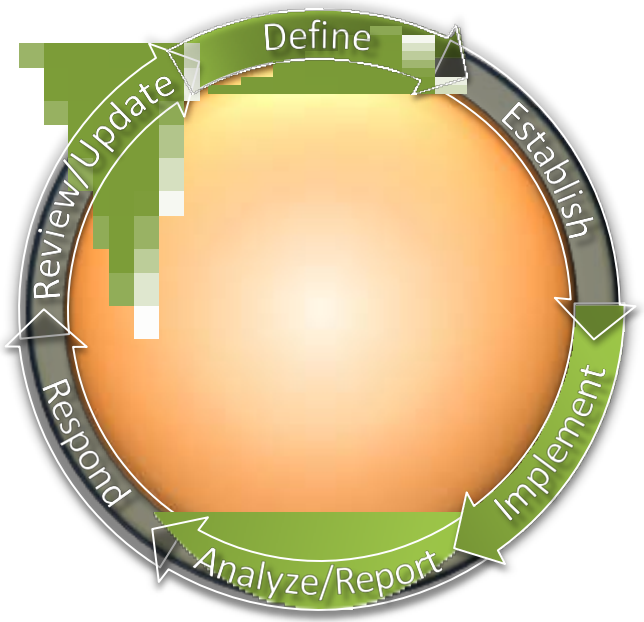
Anonymize local source information

Staffing

- need to hire data scientists

MORE GOOD NEWS

Figure 3-1. ISCM Process



Continuous Monitoring

- Maps to risk tolerance
- Adapts to ongoing needs
- Actively involves management

FINAL THOUGHTS

Continuous Monitoring focus is assurance and not compliance

Feed the right data to upper mgt for risk acceptance

Allows you to find compromised hosts before it becomes a serious problem

Requires a shift in security architecture design

Books:

Applied Network Security Monitoring, Sanders, & Smith, ISBN 978-0-12-417208-1

Extrusion Detection, Betjlich, ISBN 0-321—34996-2

Network Security Through Data Analysis , Collins, ISBN 978-1-449-35790

Network Forensics, Davidoff & Ham, 978-0-13-256471-7

QUESTIONS?

Contact Information

Randy Marchany

University IT Security Officer

VA Tech IT Security Office & Lab

1300 Torgersen Hall

Blacksburg, VA 24061

540-231-9523 (office) 540-231-1688(lab)

marchany@vt.edu

Twitter: @randymarchany

Blog: randymarchany.blogspot.com