Using Vantage To Manage Complex Sensor Networks

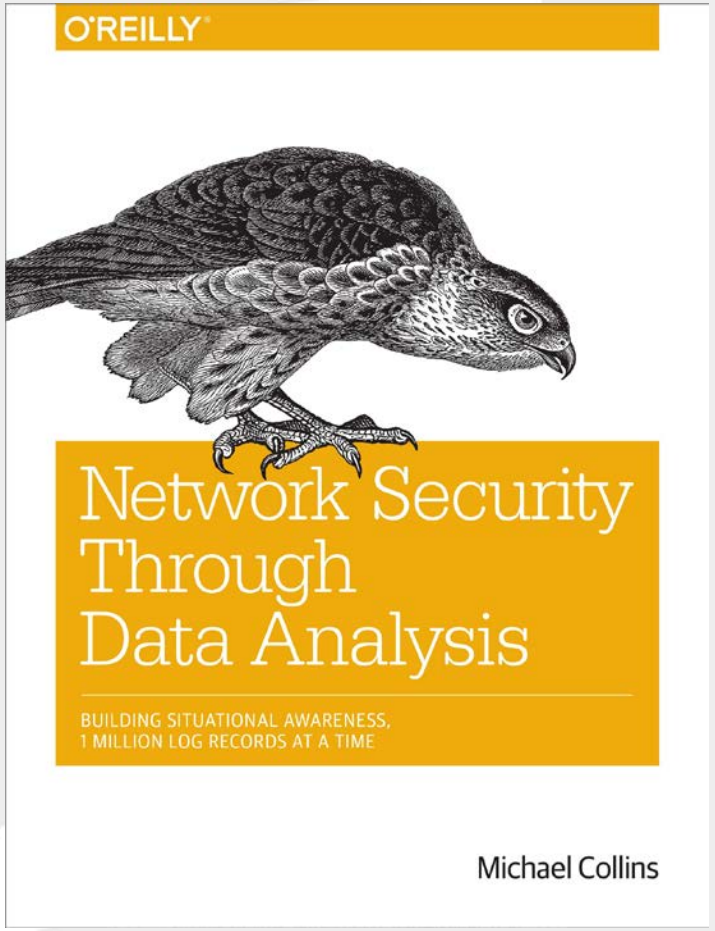
REDJACK

Flocon 2015

REDJACK

Biography

- Michael Collins,
  - Chief Scientist, RedJack
- Did a bunch of stuff at CERT
- Wrote a book on flow analysis



REDJACK

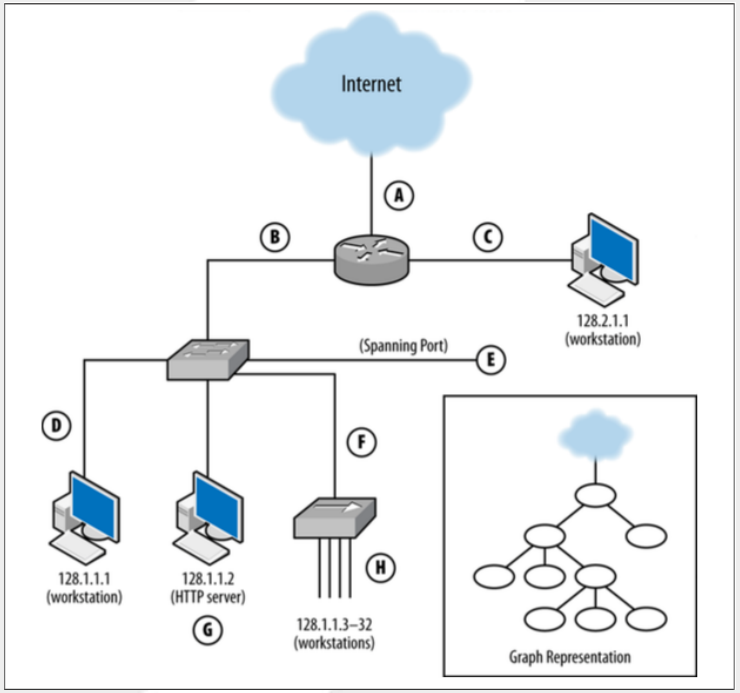
What is Vantage Analysis? (1)

- Study the network's structure to understand impact on monitoring
- Major concerns are blind spots and repetition
  - Blind spots: locations where traffic cannot be identified
  - Repetition: the same traffic recorded at multiple sensors

REDJACK

What is Vantage Analysis? (2)

- Vantage analysis consists of two major components
  - Classification of sensors by *vantage*, *domain*, and *action*
  - Mapping vantage options into a graph



REDJACK

Why Vantage Analysis?

- Data collection is basically solved
  - Problem is what to do with data
  - Data collection can be self-defeating
- Network design is increasingly complex
  - Multiple hands on the levers
  - Increased mobility, ephemerality
- Information we need is scattered in multiple locations
  - Not sure you can do complete, non-repetitive monitoring

REDJACK

Implementing Vantage Analysis - Overview

- Vantage analysis consists of two major steps
  1. Classifying points by *vantage*, *domain*, *action*
  2. Identifying overlapping domains via graphs
- I'm going to go through each of these steps, and what's involved

REDJACK

Implementing Vantage Analysis - Vantage

- Vantage is expressed as pairs of IP ranges
  - Source range, destination range
  - Both directions recorded
- Special ranges:
  - Single addresses
  - Ports
  - "Internet" (everything that isn't your network)

Point	Source	Dest
G	128.1,2.1.1-32, Internet	128.1.1.2:tcp/80
	128.1.1.2:tcp/80	128.1.,2.1.1-32

REDJACK

Implementing Vantage Analysis - Domain

- *Domain* refers to the data that can be collected
  - *Network:* tcpdump, flow, &c. POV is from the wire.
  - *Host*: State information about the host (memory, logins, logouts)
  - *Service*: Specialized service logs (e.g., HTTP or FTP)
- Domain informs the fidelity of the data, and decisions about repetitive collection

REDJACK

Implementing Vantage Analysis - Action

- *Action* describes how a sensor reacts to or collects data
- Three basic actions
  - Report: passively describe what's seen
  - Event: decide whether or not to describe something
  - Control: alter traffic based on something

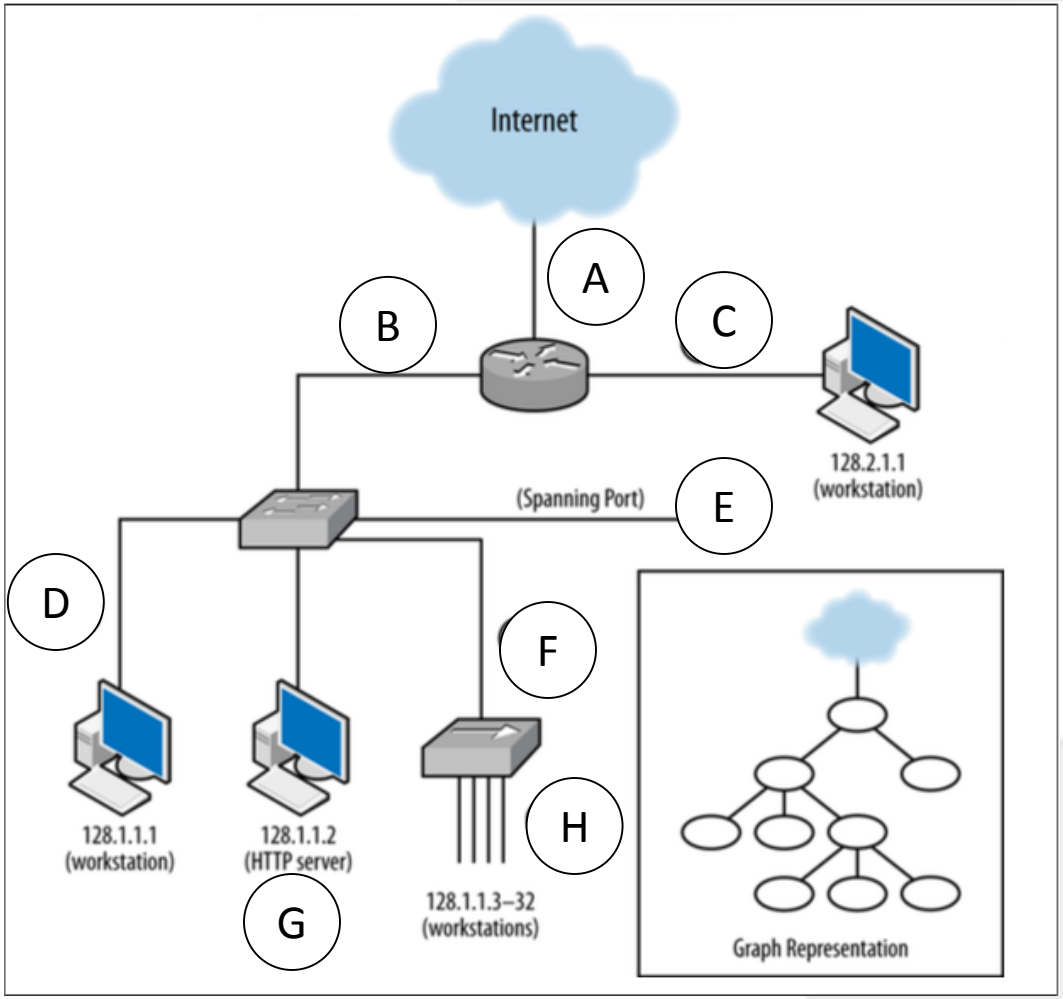
REDJACK

Implementing Vantage Analysis - Overlap

- Overlap happens when two sensors have similar vantage
- Identify by enumerating source/destination IP spaces and comparing each vantage point's set
- Best case scenario: same vantage, different domains
    - Then, pick the narrowest domain

REDJACK

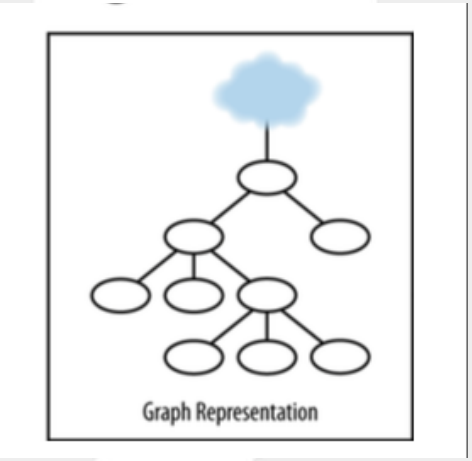
An Example Analysis



REDJACK

First: Figuring Out Instrumentation Points

- Promiscuous device (pcap):
  - Vantage: entire collision domain
- Switch (mirror port)
  - Vantage: all mirrored ports
- Router (monitoring port)
  - Vantage: VLANs
- Draw a graph showing ip sets at endpoints



REDJACK

Second: Create Table Showing Relations

Point	Source	Destination	Domain
A	I	128.1,2.1.1-32	Network
B	128.1,2.1.1-32	128.2.1.1, I	Network
C	128.2.1.1	128.1.1-32,I	Network
D	128.1.1.1	128.1.1.2-32,128.2.1.1,I	Network
E	128.1.1.1	128.1.1.2-32,128.2.1.1.,I	Network
	128.1.1.2	128.1.1.1,128.1.1.3-32,128.2.1.1,I	Network
	128.1.1.3-32	128.1.1.1-2,128.2.1.1,I	Network
F	128.1.1.3-32	128.1.1.1-2,128.2.1.1,I	Network
G	128.1,2.1.1-32,I	128.1.1.2:tcp/80	Service/HTTP
H	128.1.1.3-32	128.1.1.1-32,128.2.1.1,I	Network

REDJACK

Third, Evaluate Vantage Points

- Optimal calculation is likely to be NP-complete, suggest working greedy
  - Pick largest spaces, add in progressively finer spaces
- Identify overlapping vantages with different domains

REDJACK

Resources and Future Development

- This is a moving target
  - First version is in book
  - Updates at http://www.mpatrickcollins.com/
- Future moves:
  - Automating process
  - Errors in accountability