Near Real-Time Multi-Source Flow Data Correlation

Carter Bullard
QoSient, LLC
carter@qosient.com

FloCon 2013
Albuquerque, New Mexico
Jan 7-10, 2013
Problem Statement

• Cyber incident attribution and forensics, is a complex process.

• To assist in security incident response, recognizable hostile activity needs to be associated with other information system behavior in order to understand the complete cyber security incident life cycle
  
  • SSH flow is used to log in as root, and run a rogue program to exfiltrate data from the enterprise.

• Flow status information provides great transactional network traffic audit data

• How to correlate flow data to provide extended situational awareness to address complex cyber security events.
Data Correlation

• Semantic Enhancement
  • Classification Information
    • Geospatial - Geolocation information
    • Netspatial - Virtual positioning information
    • Community of Interest
    • Application
  • Origination information
    • User / Machine

• State based alerting / alarming
  • Intrusion detection
Data Correlation Strategies

- Flow Attribute Matching
  - Flow Identifiers
  - Protocol specific identifiers
  - Packet Dynamics
    - Inter-packet arrival times
    - Packet Size
  - Transactional Dynamics
    - Duration

- Non-flow Attribute Matching
  - Time
  - Observation Domain
  - Cross Domain Transactional Keys
  - Flow Identifiers
Distributed Situational Awareness

Attack Scenarios - External Threats

Data Plane
Command and Control
Attack Traffic

White/Visible Node
Black/Non-Visible Node
Distributed Situational Awareness

Attack Scenarios - Interior Exterior Spoofing
Spoof Correlation

• Simple multi-domain flow correlation

• However, with NAT, encryption, tunneling, traditional flow correlation is not possible.
  • No applicable flow identifiers for matching
  • Flow granularity mismatch

• Need flow metadata to make assessment
  • Content
  • Time
  • Packet dynamics (PD).

• Absence of correlation is the key
  • Statistical systems are unusable
End-to-End Situational Awareness
Network Optimization - Black Core Mesh

System Layer 2-7 Flow Data
- System Communication Efficiency
- Connectivity / Availability
- Offered Load / Loss / Jitter
- One-Way Delay (GPS synchronization)
- Round Trip Delay

Comprehensive Layer 2-7 Flow Data
- Site Communication Efficiency
- Enterprise Communication Efficiency
- Site Offered Load / Loss / Jitter
- Network Transit Times

Comprehensive Layer 2-4 Flow Data
- Network Path Assurance / Status
- Reachability / Availability Assessment
- One-Way Delay (GPS Synchronization)

Comprehensive Communication Efficiency
- End-to-End Communication Efficiency
- Reachability / Connectivity
- Received Load / Loss / Jitter
- Network Transit Times

ISP Communication Efficiency
- Ingress Available Capacity / Loss / Jitter
- One-Way Delay (GPS Synchronization)
- Network Path Status

Interface Status / Transitional Events
- Bulk Link Statistics

SNMP RMOM Element Statistics/Traps
- Interface Status / Transitional Events
- Bulk Link Statistics

Information System Repository
- Comprehensive Flow Monitor
- SNMP RMOM Style Monitor
- Core Service Provider Management Domain
- Enterprise Management Domain
- Core Service Provider Management Domain

ARGUS
Radium
Data Flow Machine Architectures
Radium
Data Flow Design

![Diagram of Radium Data Flow Design](image)
Radium
Data Flow Design

RADIUM
Streaming Analytics Correlation
Enterprise Border Awareness
Outside Inside / Them vs Us

Comprehensive Flow IS
Black/Non-Visible Node
White/Visible Node
Argus Sensor
Enterprise Border Awareness
Core Network
Enterprise Domain

Situational Awareness Data
Data Plane
Unobservable Traffic
Subnet Border Awareness
Local and Remote Strategies

- White/Visible Node
- Black/Non-Visible Node
- Comprehensive Flow IS
- Argus Sensor

- Data Plane
- Situational Awareness Data
- Unobservable Traffic
End System Awareness
Local and Remote Strategies

Comprehensive Flow IS
Black/Non-Visible Node
White/Visible Node
Argus Sensor
Situational Awareness Data
Data Plane
Unobservable Traffic
Local and Remote Strategies
Complex Comprehensive Awareness
Local and Remote Strategies

White/Visible Node
Black/Non-Visible Node
Comprehensive Flow IS
Argus Sensor

Data Plane
Situational Awareness Data
Flow - Non Flow Correlation

- Replay attack detection
  - Bi-Directional Protocol Time Uncoupling

- Stepping stone detection
  - Two completely independent flows, that share the same instantaneous burst behavior and packet size frequency distribution (shifted for encapsulations)

- Man vs Machine detection
  - Interactive vs Non-Interactive Session Detection
  - Packet, transaction and session jitter analysis

- Man-in-the-middle detection
  - Pass Thru - Detectable one-way latency, hop count, path resource modifications
  - Proxy - Connection setup time modifications, header attribute changes

- Performance as an Asset that needs Protection
  - Path Availability, Bandwidth, Latency, Jitter, MTU, ....
  - Continuous One-Way latency determinations
Supporting Slides