Flamingo: Visualizing Internet Traffic Data

Manish Karir, Jon Oberheide, Mike Goff
Networking R&D, Merit Network Inc.
• Introduction: What is Flamingo
• Visualizations
• The Flamingo Tool
  – Combining visualization with controls
• Case Studies
  – Traffic Anomaly
  – Network Scans
  – Worm traffic
  – P2P traffic
  – The Slashdot effect!
Introduction

• Flamingo is a unique software tool that enables 3D Internet traffic data exploration in real-time
• Provides a series of different visualization methods to illustrate different aspects of the data
• Based on information extracted from netflow records
• Includes additional tools/filters to allow people to easily extract “information” from raw netflow data
Introduction: Flamingo Architecture

• Client/Server Architecture
• A single server can support multiple clients
• A single server can act as collector for multiple netflow feeds
• Supports both aggregation as well as non-aggregation mode
Visualization Methods

• Based on Extended Quad-Tree Implementation
• Traffic Volume by src/dst IP prefix
• Traffic Volume by src/dst AS
• Traffic distribution across src/dst ports
• Traffic flows between src/dst IP prefixes
• Traffic flows between src/dst IP/ports
The Basic Quad-Tree
Traffic Volume by Src/Dst IP

• The 2D quad-tree map is used as the base of a visualization cube
• We plot prefixes from a BGP routing table onto the base of the cube, size of prefix determines size of representation on 2D base
• Longest prefix match is used to map netflow IP addresses onto BGP prefixes
• The z-axis/height is used to represent the volume of traffic
• Different color is used for each prefix
Traffic Flows by Aggregated Src/Dst IP

- Flows contain source and destination information, which might map to 2 different prefixes, so far we only have the ability to represent a single flow
- Solution: Use 2 inside surfaces of a cube, one for source, one for destination, represent a flow by a line between them
- Thickness of line represents relative traffic volume
Traffic Flows by Aggregated Src/Dst IP
Traffic Flows by Src/Dst IP and Port

- Flows contain source/destination port number information as well
- Solution:
  - Use base of cube to represent prefixes, both source and destination are on the same base
  - The z-axis is used to represent port numbers, source and destination
  - \((\text{srcIP}, \text{srcPort})\) \(\rightarrow\) \(\text{((x1,y1), z1)}\)
  - \((\text{dstIP}, \text{dstPort})\) \(\rightarrow\) \(\text{((x2,y2), z2)}\)
  - Line between these 2 points in 3D space represents a flow from \((\text{srcIP}, \text{srcPort})\) to \((\text{dstIP}, \text{dstPort})\)
  - Line thickness represents relative volume of traffic
  - Same color used for all flows with same source IP
Traffic Flows by Src/Dst IP and Port

(SrcIP, SrcPort)

(DstIP, DstPort)
Flamingo Visualization Tool
Flamingo Controls

Text Representation of Visualized Information

Slider Bar Controls

Address Range Configuration
Case Study: Traffic Anomaly Sunday- Oct 16, 2005

- Large burst of traffic visible outgoing from 141.213.x.x(x.x.umich.edu)
- Start time roughly – 12PM - End time roughly – 6PM
- Single srcIP/port – few(4) targetIP’s/multiple ports
- UDP flows
- Traffic pattern visible in the normal clutter
- We then proceed to examine the src (141.213.x.0/24) and target prefixes (216.74.128.0/18, 217.199.32.0/19) in more detail in the following sequence of images
Overall Traffic Pattern at Primary Router Sunday, Oct16 2005

Anomaly
Traffic Volume sourced from /24 subnet by individual hosts

141.213.x.x/32

Src IP Addresses 141.213.x.0/24

5 million flows

5 million flows
### Distribution of Target IP Addresses

**Few (4) Specific Target IP’s**

<table>
<thead>
<tr>
<th>Color</th>
<th>Volume</th>
<th>Src Addr -&gt; Dst Addr</th>
<th>Despription</th>
</tr>
</thead>
<tbody>
<tr>
<td>1410000</td>
<td>141.213</td>
<td>/32 -&gt; 217.199.143/32</td>
<td>N/A</td>
</tr>
<tr>
<td>741000</td>
<td>141.213</td>
<td>/32 -&gt; 217.74.113/32</td>
<td>N/A</td>
</tr>
<tr>
<td>659000</td>
<td>141.213</td>
<td>/32 -&gt; 216.74.143/32</td>
<td>N/A</td>
</tr>
<tr>
<td>79000</td>
<td>141.213</td>
<td>/32 -&gt; 216.74.120/32</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Src IP Addresses**

141.213.x.0/24

**Dst IP Addresses**

217.74.120.0
Distribution of flows in terms of src/dst ports/protocol

One or two Src Ports

Dst ports Vary 0-65K
Case Study: Worm Traffic/Port 42 Scans

Scan:
Dst: 35/8 Port:42
Src: 219.188.209/24 Port: various
Case Study: /24 Network Scan
ssh scans
ssh scan
Case Study: Slashdot Event Oct 31, 2004

Traffic Volume

Flow Volume
Zotob Worm Infection
P2P Traffic
Darkspace Traffic Visualization
Conclusion

• The Flamingo Visualization Tool provides users with the ability to easily explore and extract meaning information regarding traffic flows in their network

• More details can be found at:
  – http://flamingo.merit.edu