

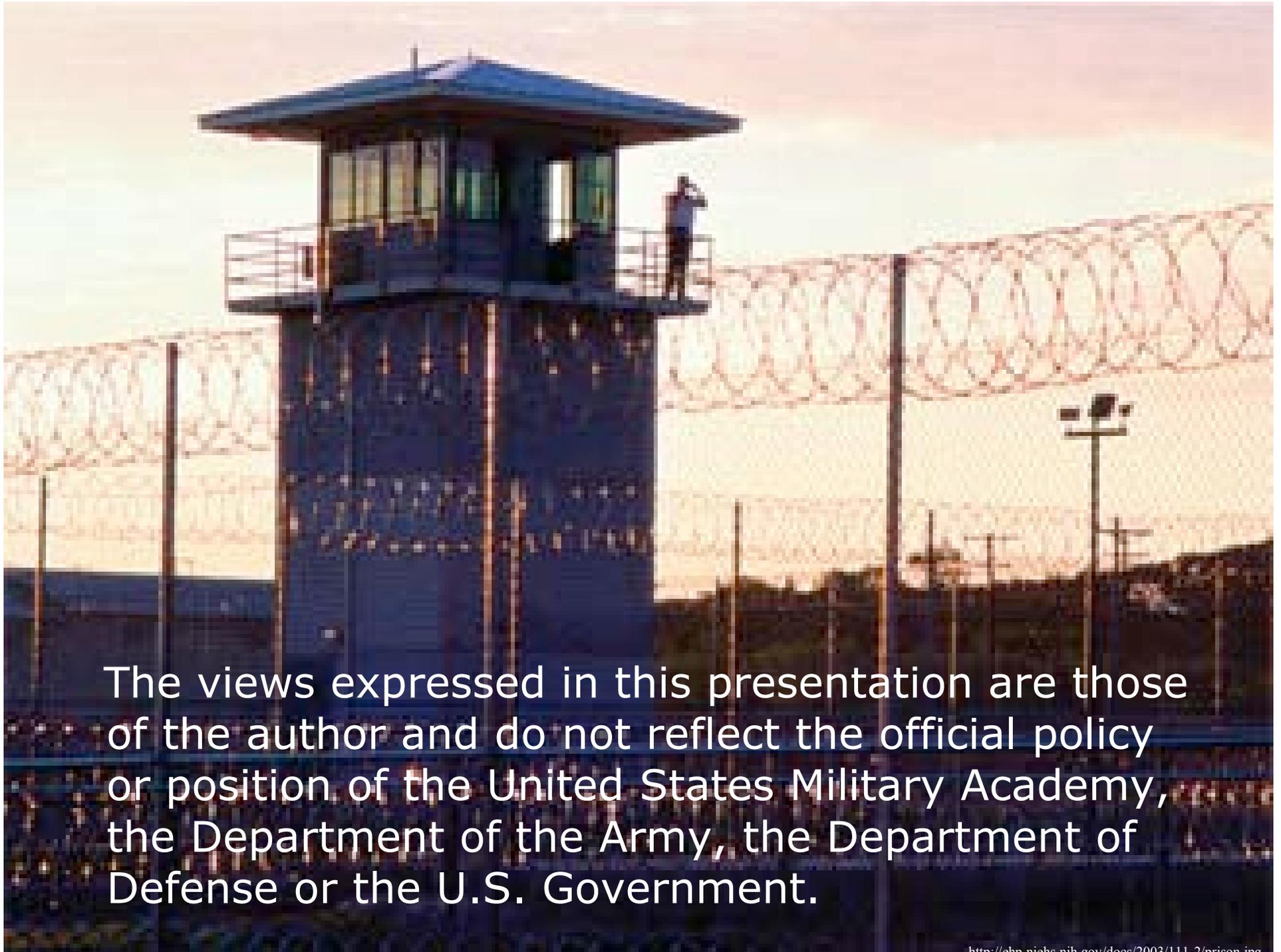
Malware Cinema

*A Picture is Worth a
Thousand Packets*

Gregory Conti

www.cc.gatech.edu/~conti

conti@cc.gatech.edu



The views expressed in this presentation are those of the author and do not reflect the official policy or position of the United States Military Academy, the Department of the Army, the Department of Defense or the U.S. Government.

information visualization is the use of interactive, sensory representations, typically visual, of abstract data to reinforce cognition.

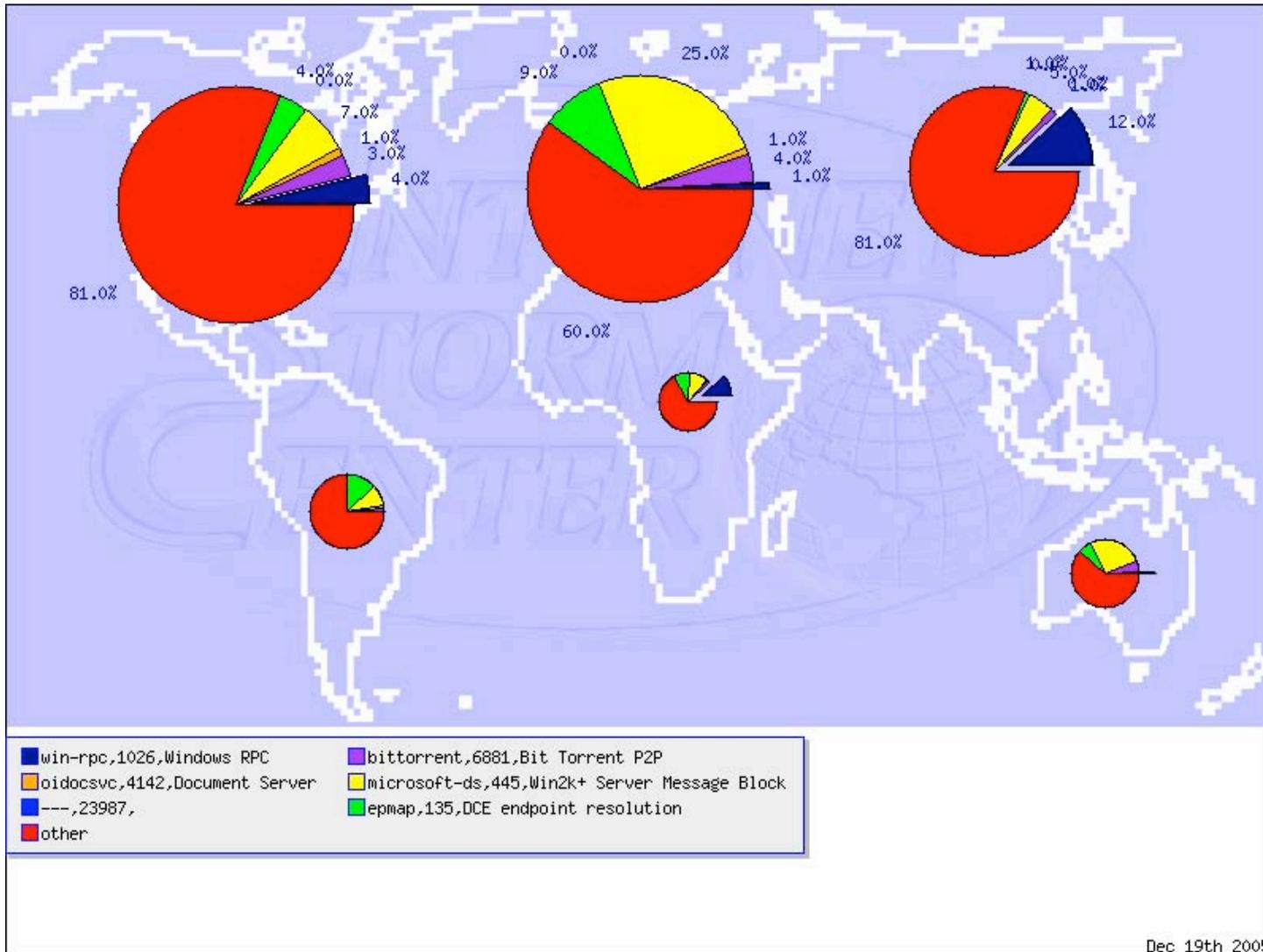
Gartner's Hype Cycle



-Gartner Group

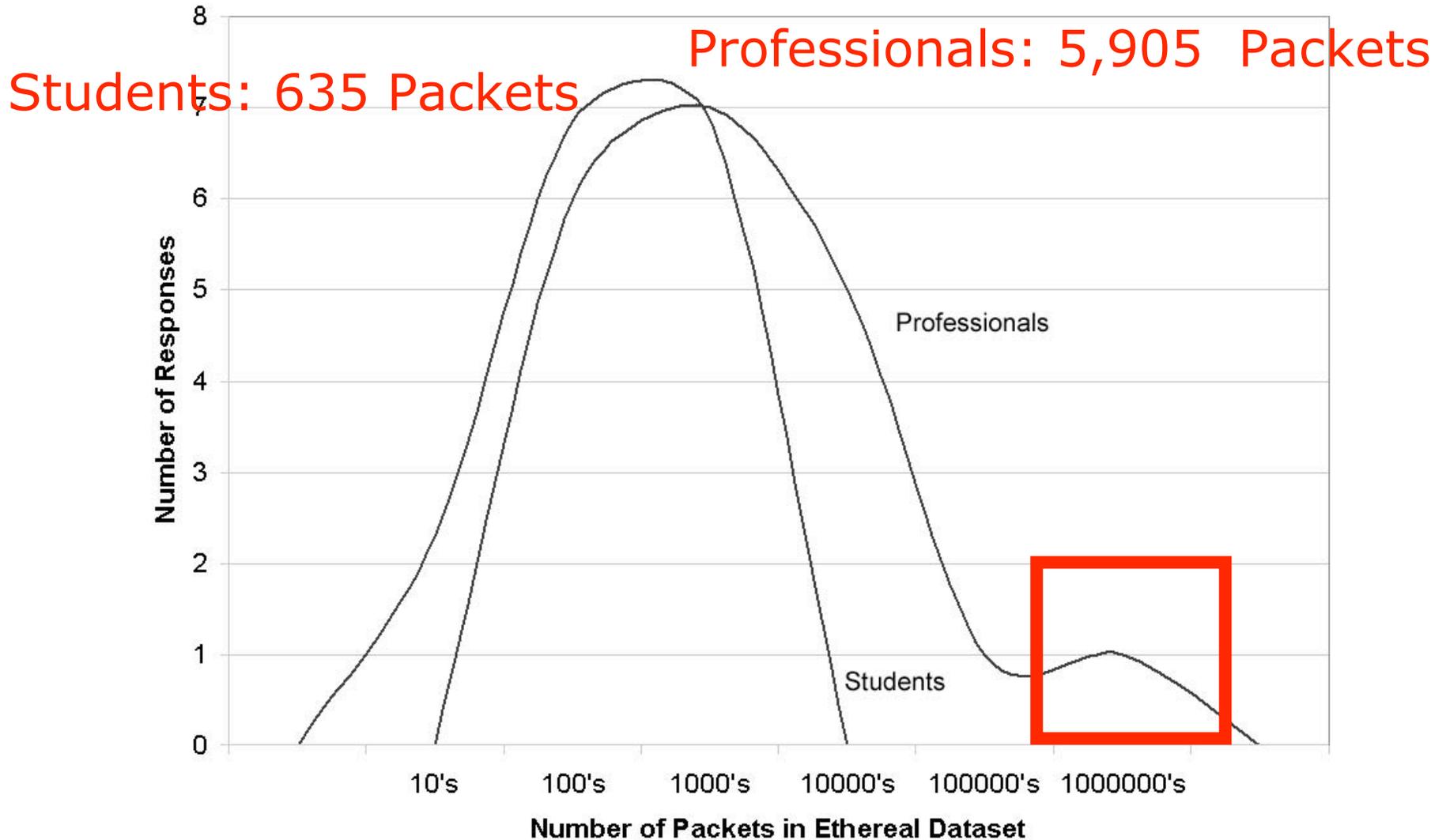
Thanks go to Kirsten Whitely for the Gartner curve idea
<http://java.sun.com/features/1998/03/images/year3/original/gartner.curve.jpg>

SANS Internet Storm Center



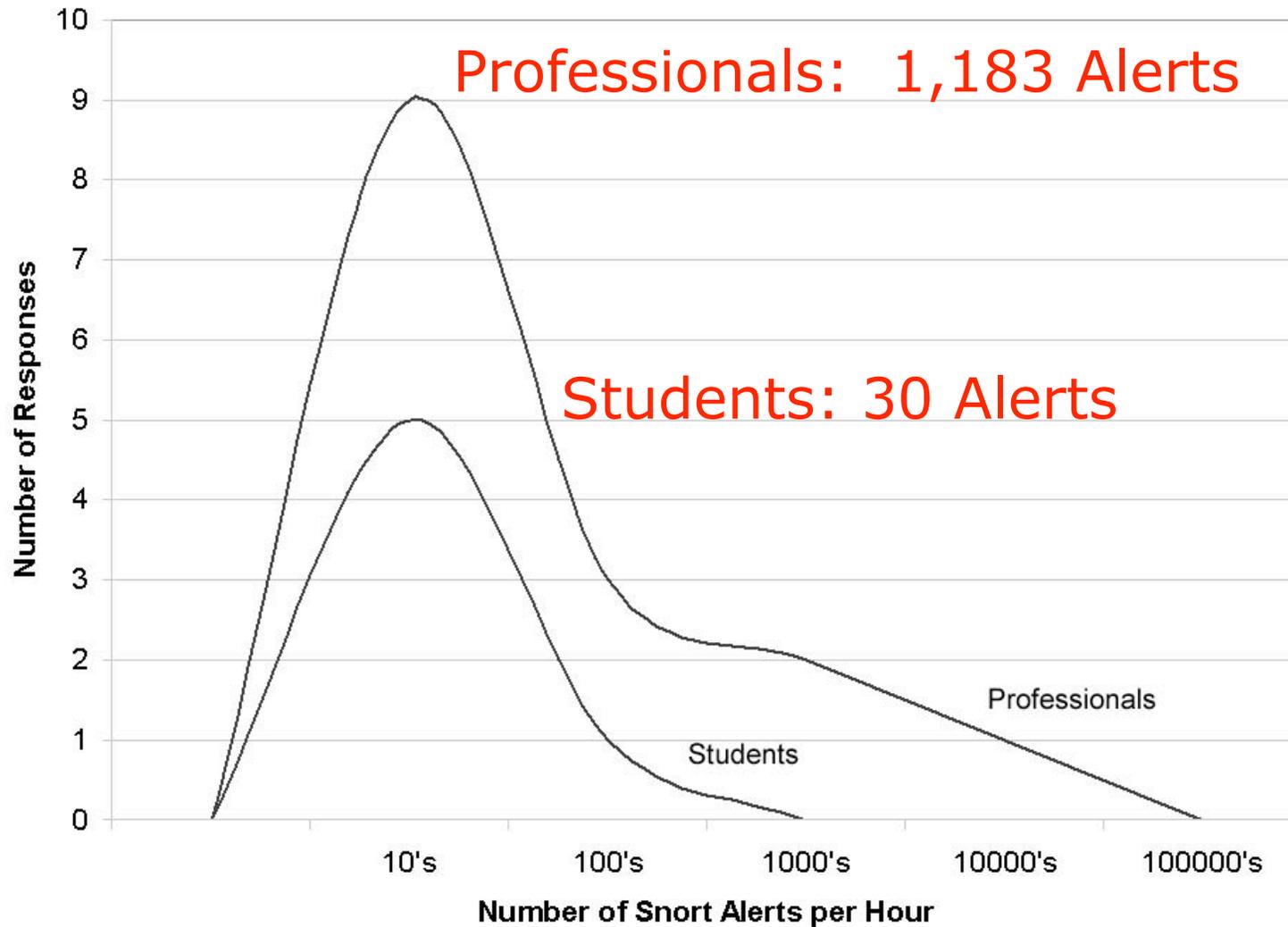
Ethereal's Tipping Point

(for the human)

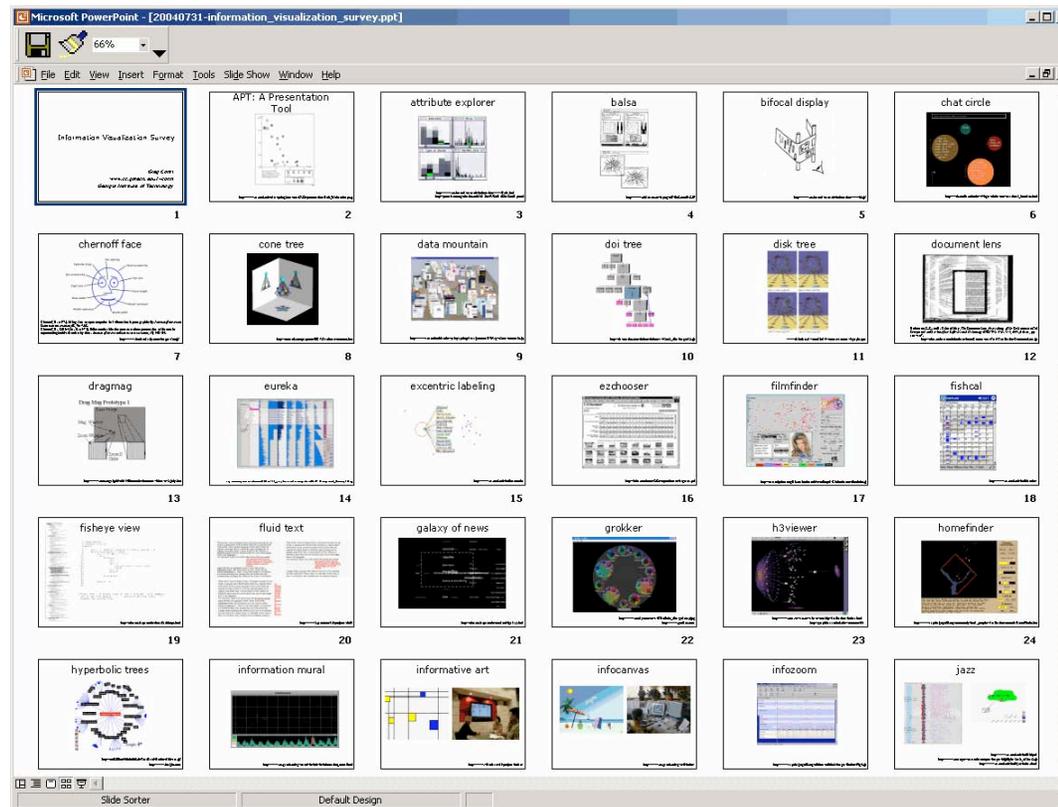


Snort's Tipping Point

(for the humans)



General InfoVis Research...



powerpoint of classic systems is here...

http://www.rumint.org/gregconti/publications/20040731-information_visualization_survey.ppt

see InfoVis proceedings for more recent work...

<http://www.infovis.org/symposia.php>

Potential DataStreams

Traditional

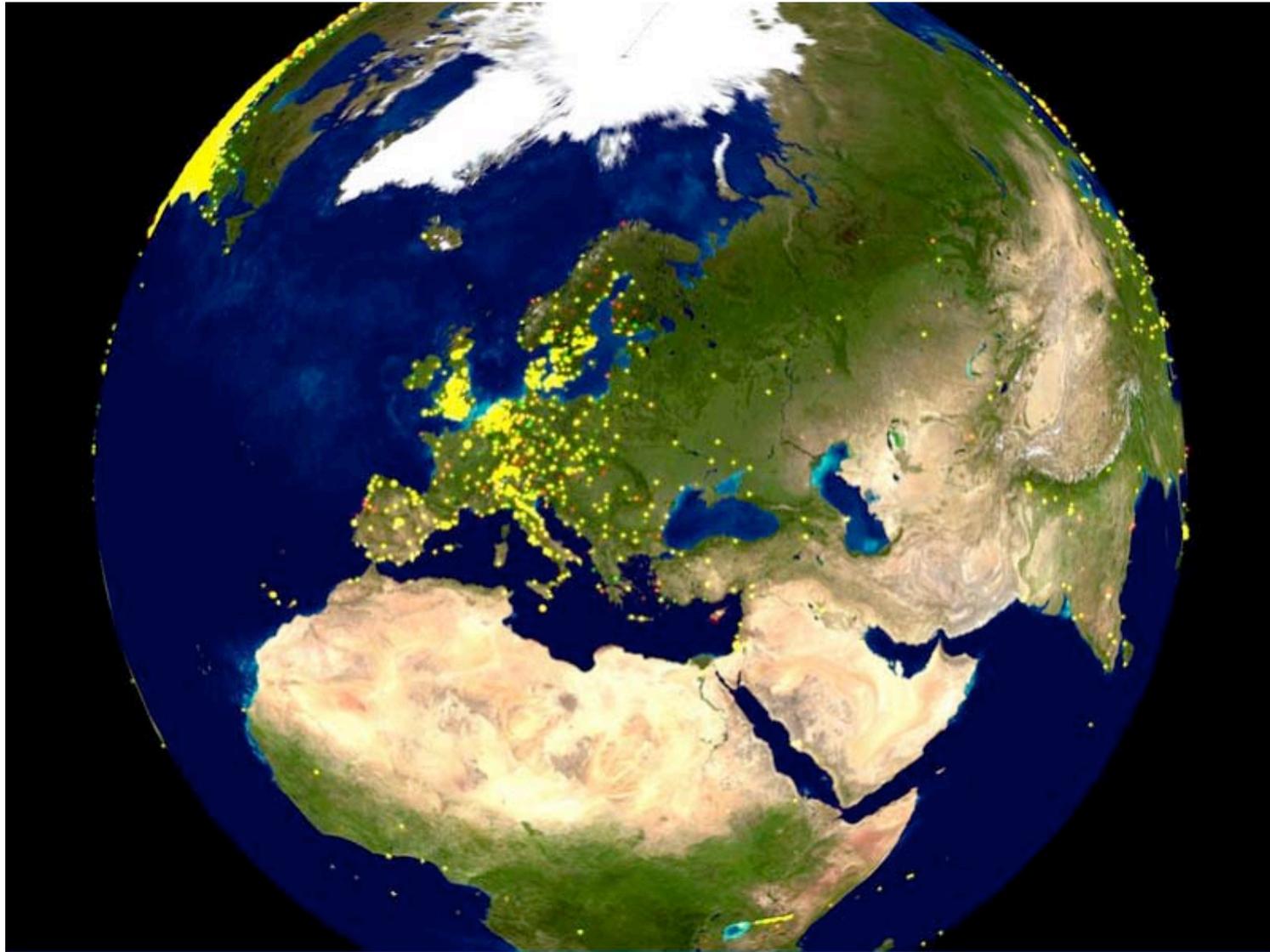
- packet capture
- IDS/IPS logs
- syslog
- firewall logs
- anti-virus
- net flows
- host processes
- honeynets
- network appliances

Less traditional

- p0f
- IANA data (illegal IP's)
- DNS
- application level
- extrusion detection systems
- local semantic data (unassigned local IPs)
- inverted IDS
- geolocation (MaxMind?)
- vulnerability assessment
 - nessus, nmap ...
- system files

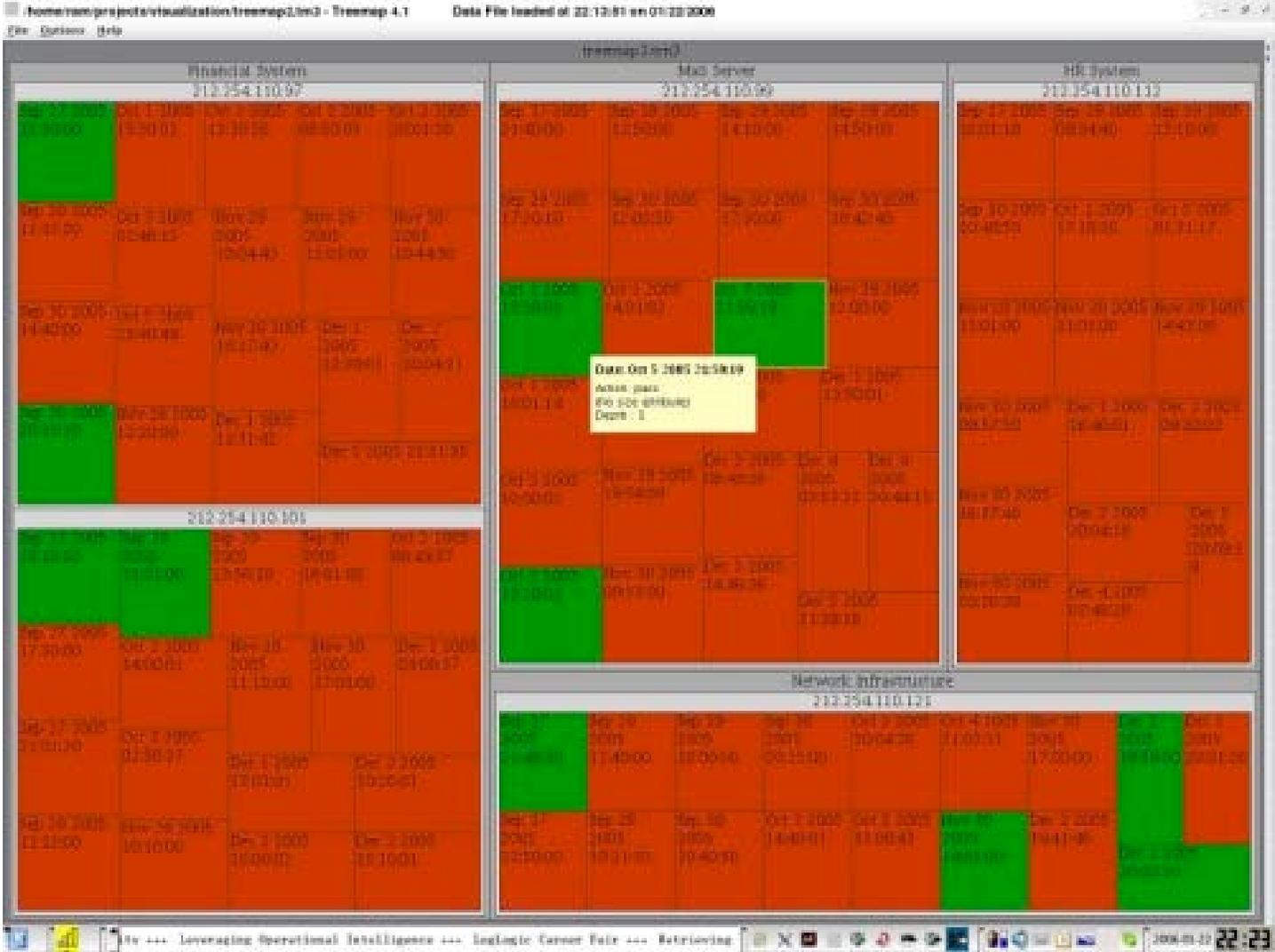
Rootkit Propagation

(Dan Kaminsky)



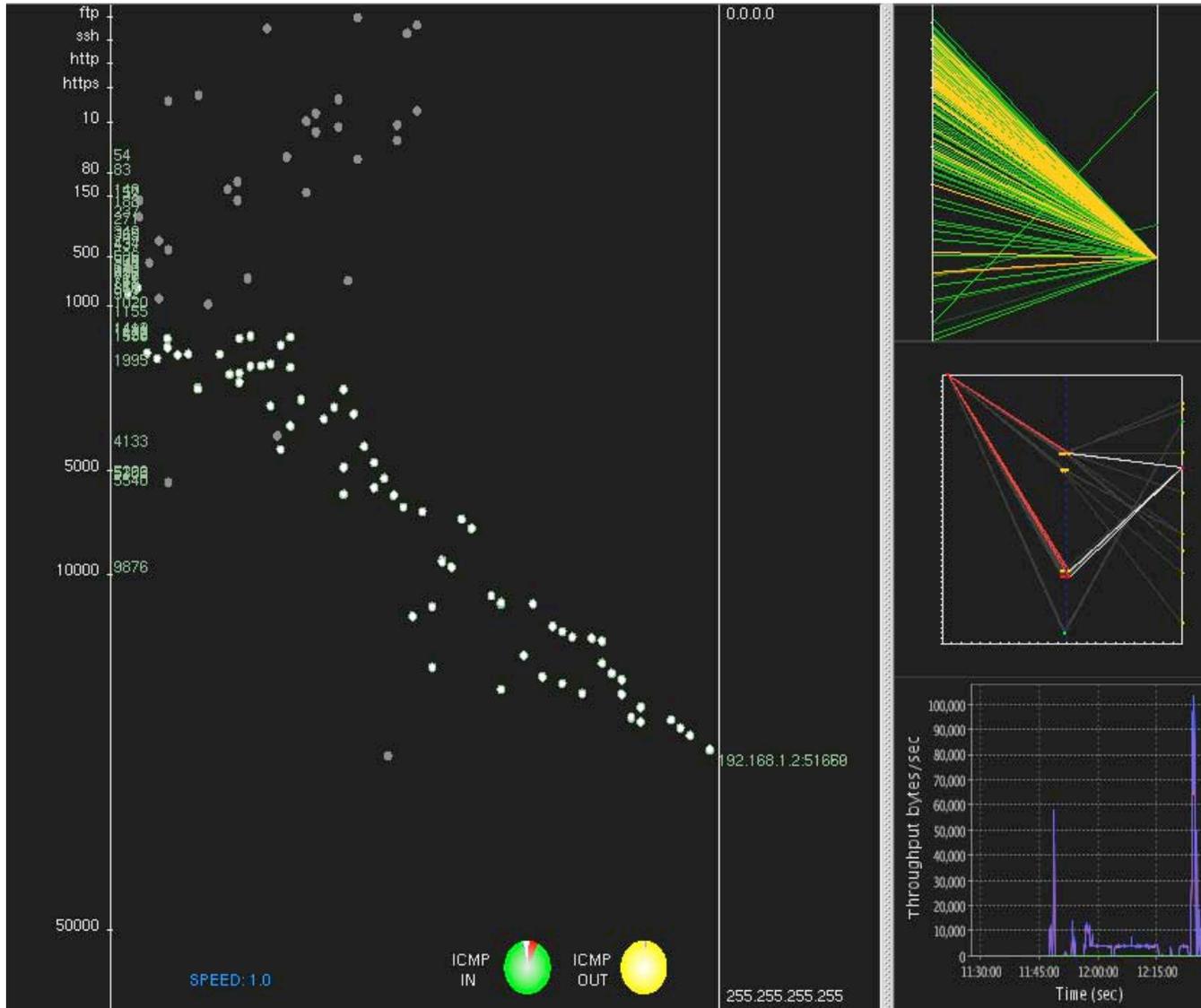
<http://www.doxpara.com/>

Firewall Data (Raffy Marty)



Firewall Data

(Chris Lee)

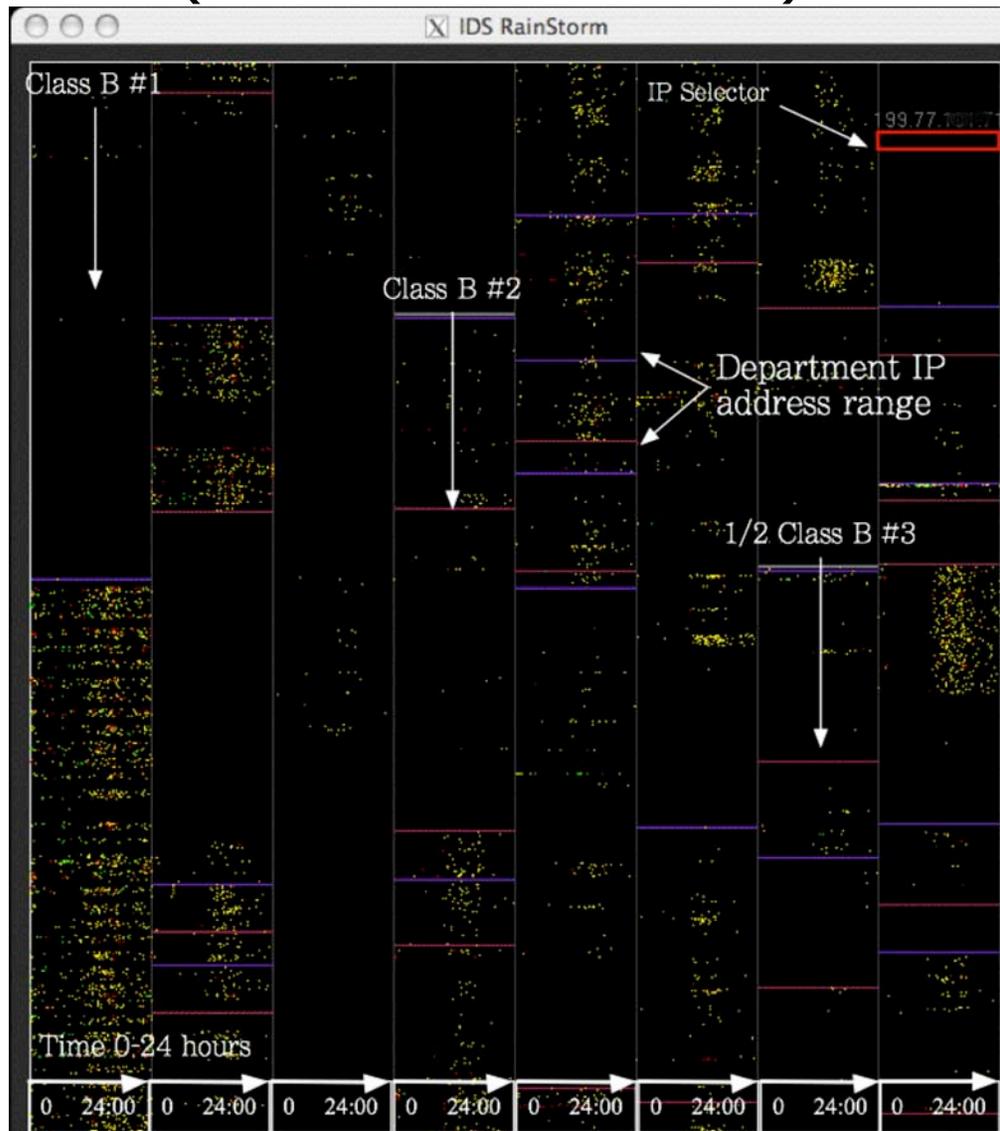


"Visual Firewall: Real-time Network Security Monitor"

Chris P. Lee, Jason Trost, Nicholas Gibbs, Raheem Beyah, John A. Copeland (Georgia Tech)

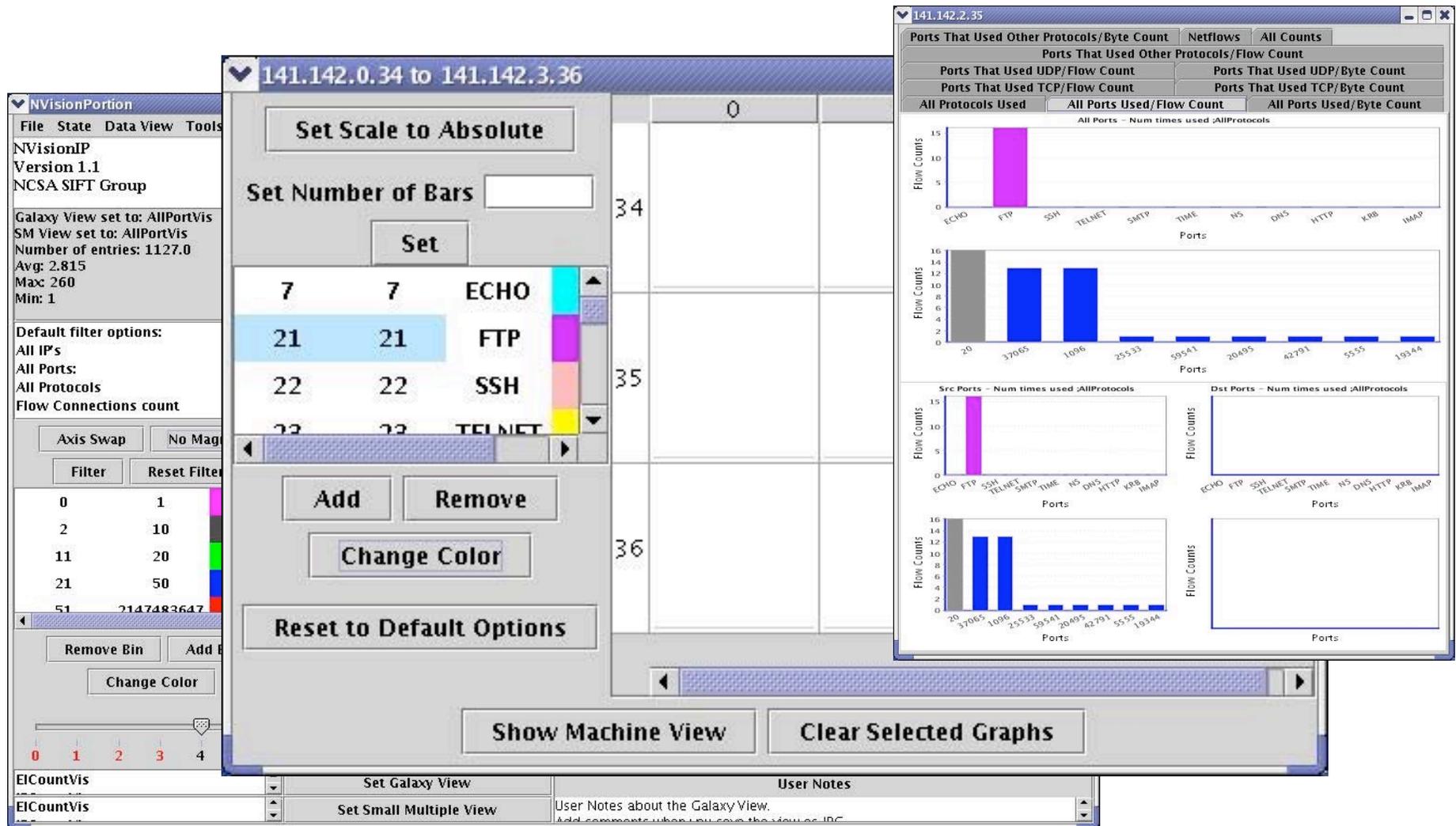
IDS Alerts

(Kulsoom Abdullah)



Netflows

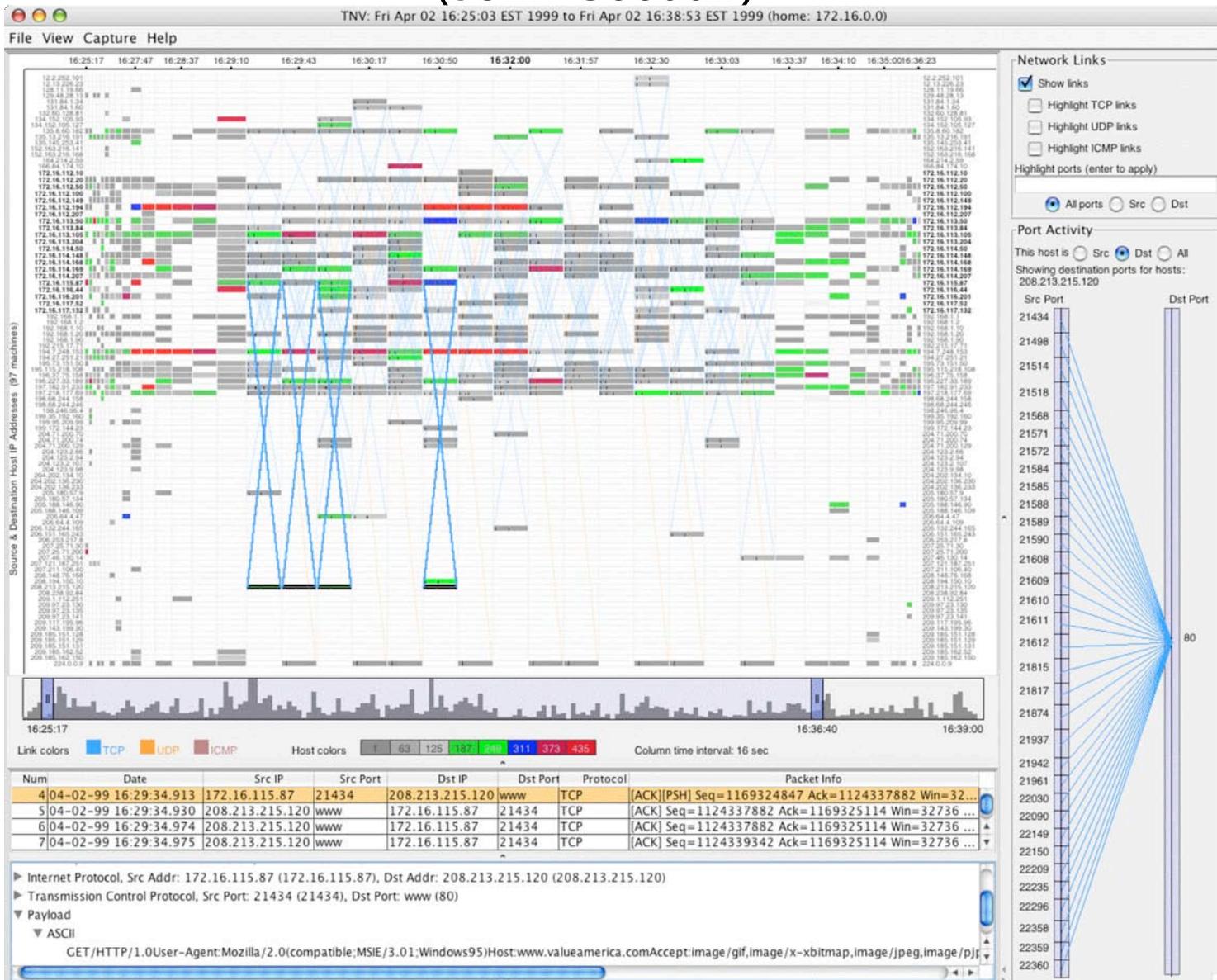
University of Illinois at Urbana-Champaign / Bill Yurcik



<http://security.ncsa.uiuc.edu/distribution/NVisionIPDownload.html>

Packet Level

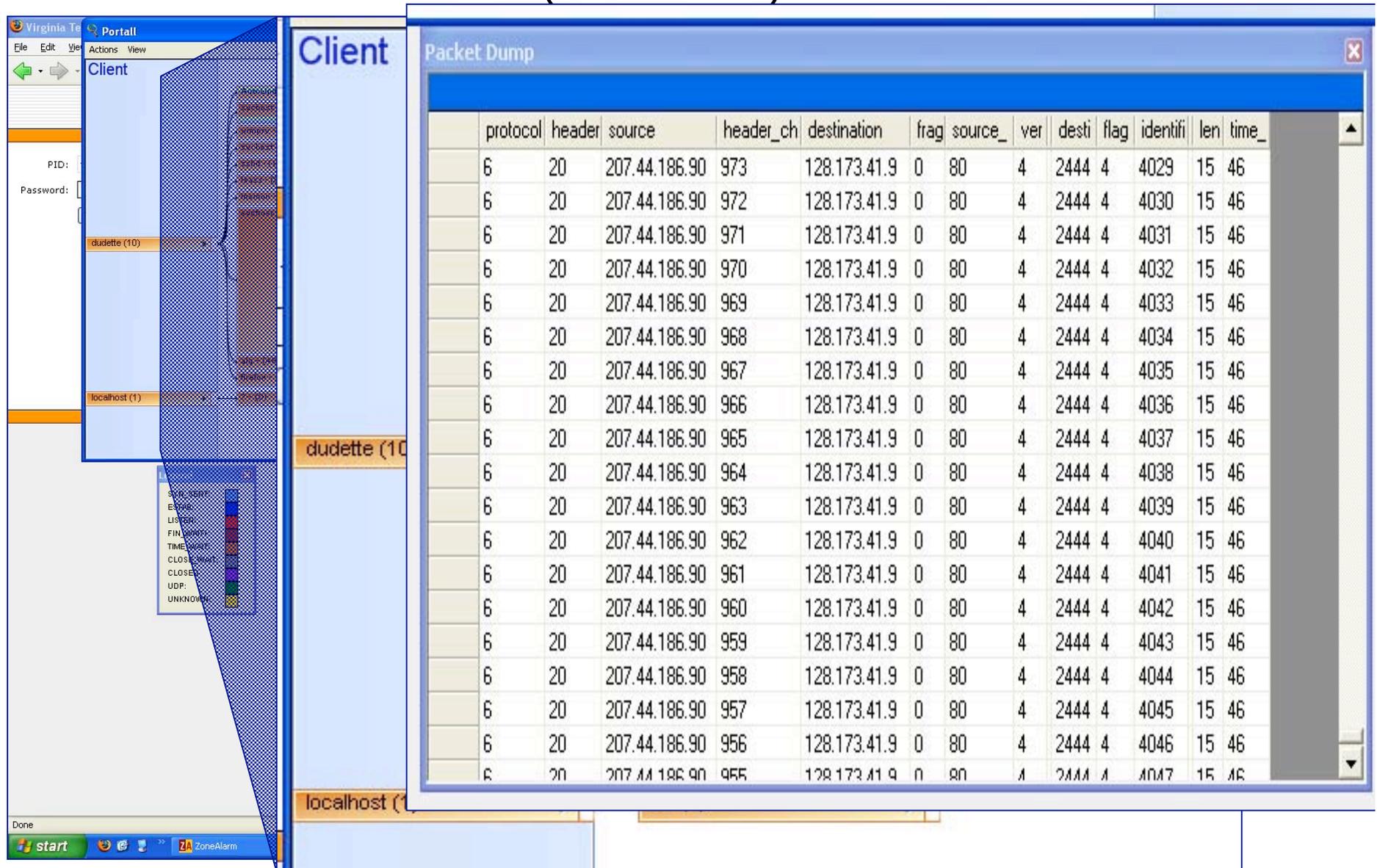
(John Goodall)



<http://userpages.umbc.edu/~jgood/research/tnv/>

Host Processes and Network Traffic

(Glenn Fink)



"Visual Correlation of Host Processes and Traffic" Glenn A. Fink, Paul Muessig, Chris North (Virginia Tech)

MD5

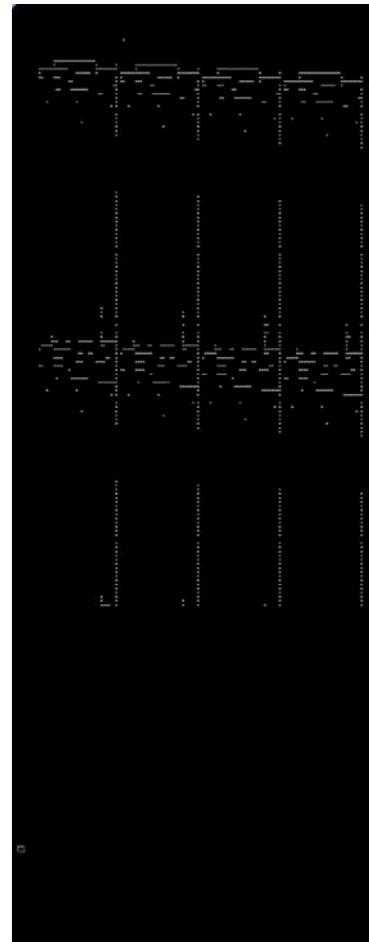
(Dan Kaminsky)



Hash 1



Hash 2



Diff



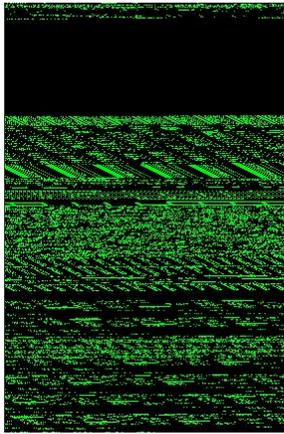
Animation

<http://www.doxpara.com/?q=node&from=10>

Comparing Executable Binaries

(Greg Conti)

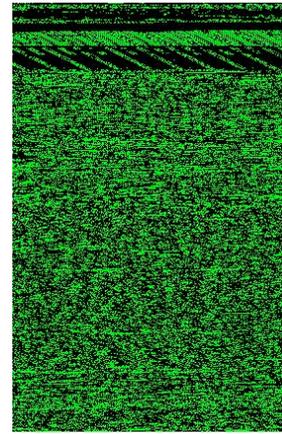
rumint.exe
(visual studio)



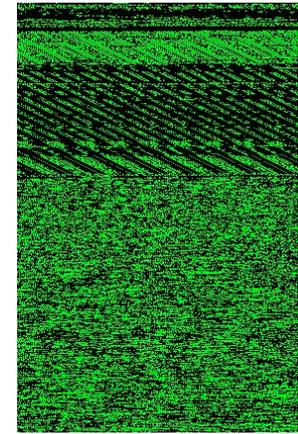
visualexplorer.exe
(visual studio)



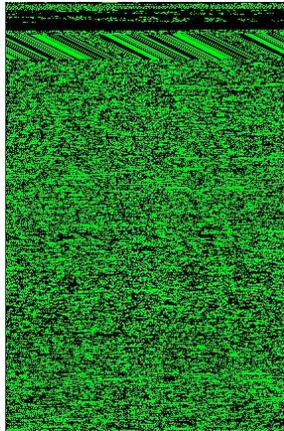
calc.exe
(unknown compiler)



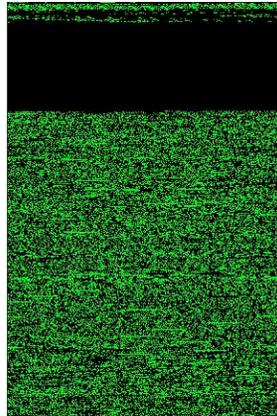
regedit.exe
(unknown compiler)



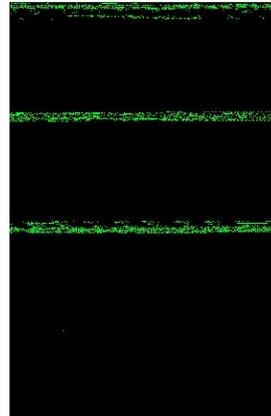
mozillafirebird.exe
(unknown compiler)



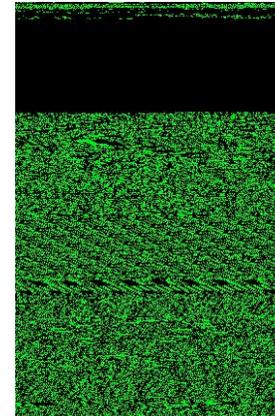
cdex.exe
(unknown compiler)

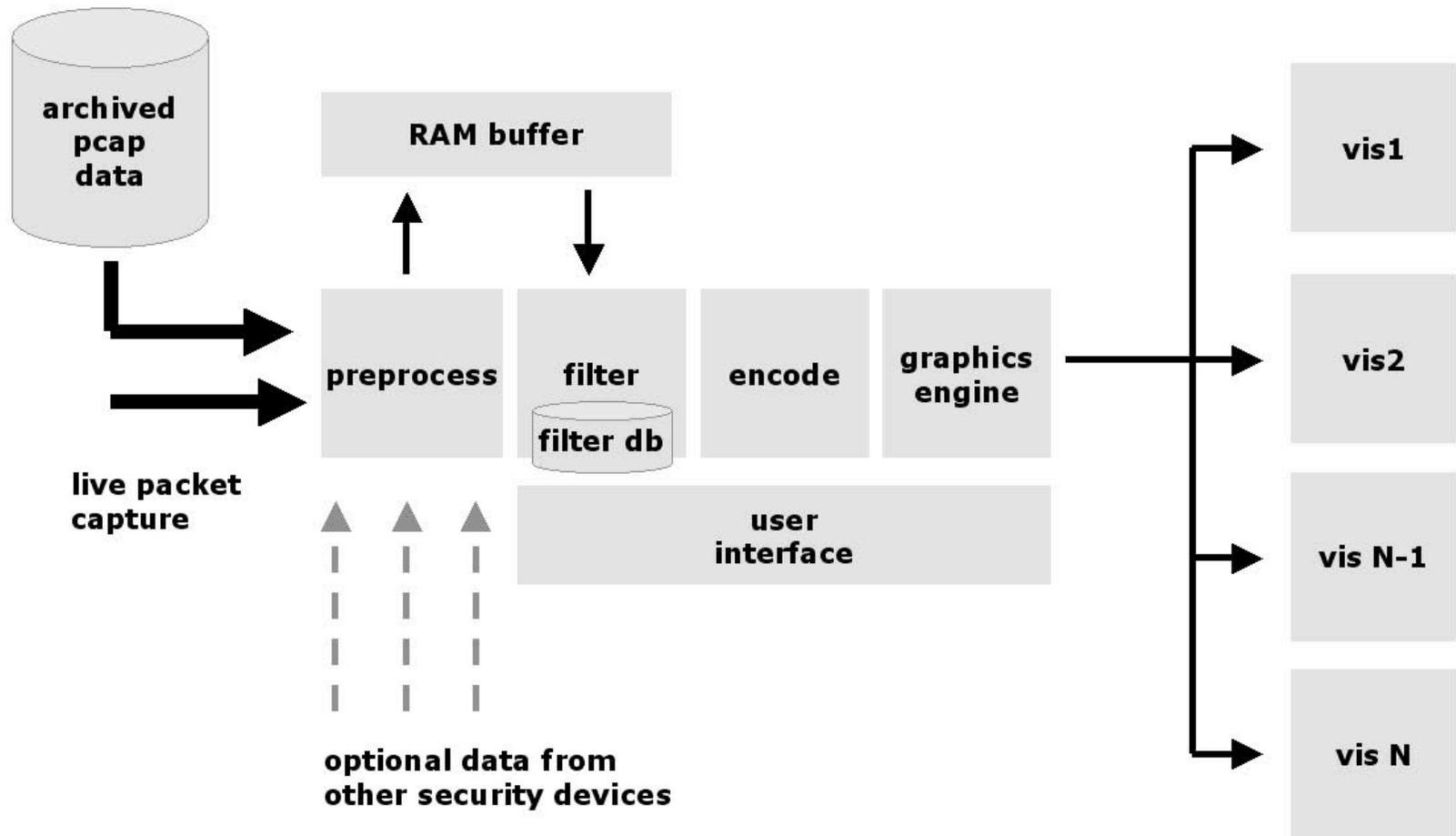


apache.exe
(unknown compiler)



ethereal.exe
(unknown compiler)





Ethereal Strengths

- Full view of all packet parameters
- Capture and display filters
- Dissect and analyze protocols

Snort Strengths

- Robust and configurable filtering
- High quality signature database
- Helps to focus human resources
- Flexibility
- Ability to access details of packets/alerts
- Open source

Ethereal Weaknesses

- Overwhelming detail / too much for human to process
- Impossible to properly visualize a large dataset without getting lost and confused
- GUI too cumbersome

Snort Weaknesses

- Too many false positives
- Reliance on known signatures
- Time and difficulty in selecting right set of signatures for a given network.
- Front end GUIs are poor

Ethereal

The screenshot displays the Ethereal network protocol analyzer interface. The main window is titled "(Untitled) - Ethereal" and features a menu bar (File, Edit, View, Go, Capture, Analyze, Statistics, Help) and a toolbar with various icons for file operations, capture, and analysis. Below the toolbar is a filter bar with the text "Filter: http" and buttons for "Expression...", "Clear", and "Apply".

The central pane shows a list of captured packets with the following columns: No., Time, Source, Destination, Protocol, and Info. The list includes several continuation packets and a series of HTTP requests and responses to mv7.pandora.nu. The selected packet (No. 224) is highlighted in blue.

No.	Time	Source	Destination	Protocol	Info
187	20.101837	mv7.pandora.nu		HTTP	Continuation
191	20.104052	mv7.pandora.nu		HTTP	Continuation
192	20.108770	mv7.pandora.nu		HTTP	Continuation
193	20.110732	mv7.pandora.nu		HTTP	Continuation
194	20.112117	mv7.pandora.nu		HTTP	Continuation
199	23.466027	mv7.pandora.nu	mv7.pandora.nu	HTTP	GET /tempo-depot/notes/ HTTP/1.1
201	24.246294	mv7.pandora.nu	mv7.pandora.nu	HTTP	HTTP/1.1 304 Not Modified
203	30.183593	mv7.pandora.nu	mv7.pandora.nu	HTTP	GET /tempo-depot/notes/PC_side/Graphic...
205	31.583269	mv7.pandora.nu	mv7.pandora.nu	HTTP	HTTP/1.1 304 Not Modified
217	43.555787	mv7.pandora.nu	mv7.pandora.nu	HTTP	GET /tempo-depot/notes/ HTTP/1.1
219	44.483956	mv7.pandora.nu	mv7.pandora.nu	HTTP	HTTP/1.1 304 Not Modified
222	48.121447	mv7.pandora.nu	mv7.pandora.nu	HTTP	GET /tempo-depot/notes/PC_side/Graphic...
224	49.788289	mv7.pandora.nu	mv7.pandora.nu	HTTP	HTTP/1.1 304 Not Modified

The bottom pane shows a detailed view of the selected packet, displaying the protocol stack: PPP-over-Ethernet Session, Point-to-Point Protocol, Internet Protocol, Transmission Control Protocol, and Hypertext Transfer Protocol. The HTTP details section is expanded, showing the following fields:

- GET /tempo-depot/notes/ HTTP/1.1\r\n
- User-Agent: Opera/7.60 (Windows NT 5.1; U) [en] (IBM EVV/3.0/EAK01AG9/LE)\r\n
- Host: www.pandora.nu\r\n
- Accept: application/xhtml+xml;version=1.2, application/x-xml+voice+xml;version=1.2, text/...
- Accept-Language: ja;q=1.0,en;q=0.9\r\n
- Accept-Charset: shift_jis, utf-8, utf-16, iso-8859-1;q=0.6, *;q=0.1\r\n
- If-Modified-Since: Fri, 03 Sep 2004 14:56:50 GMT\r\n
- If-None-Match: "345bacaa032b8f9264ceed73ec6f64a7"\r\n
- Connection: Keep-Alive, TE\r\n
- TE: deflate, gzip, chunked, identity, trailers\r\n

The bottom-most pane shows the raw packet data in hexadecimal and ASCII format, with the ASCII portion corresponding to the HTTP headers shown in the pane above.

Ethereal can be found at <http://www.ethereal.com/>

http://www.pandora.nu/tempo-depot/notes/bloxxom/data/PC_side/Web_Browser/Bloxxom/ethereal.png

Potential DataStreams

Traditional

- **packet capture**
- IDS/IPS logs
- syslog
- firewall logs
- anti-virus
- net flows
- host processes
- honeynets
- network appliances

payload

byte frequency

packet length

ethertype

IP version

IP header length

IP differential services

IP total length

IP identification

IP flags

IP fragment

TTL

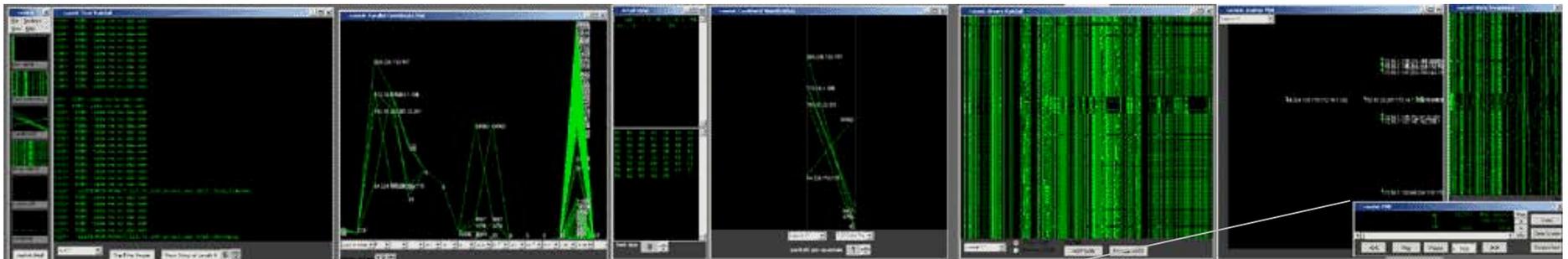
IP transport

IP header checksum

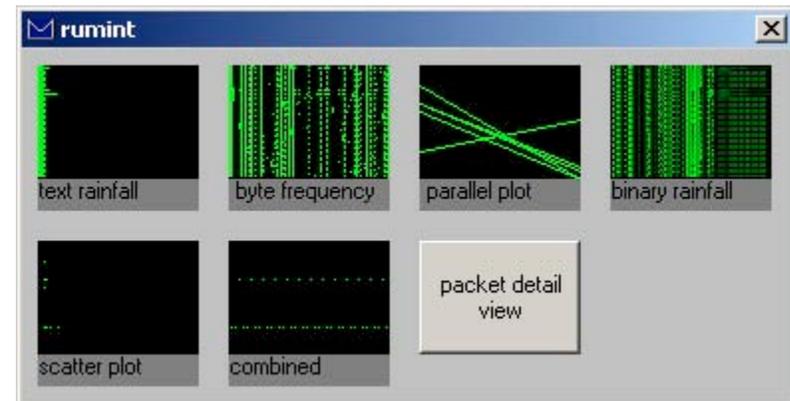
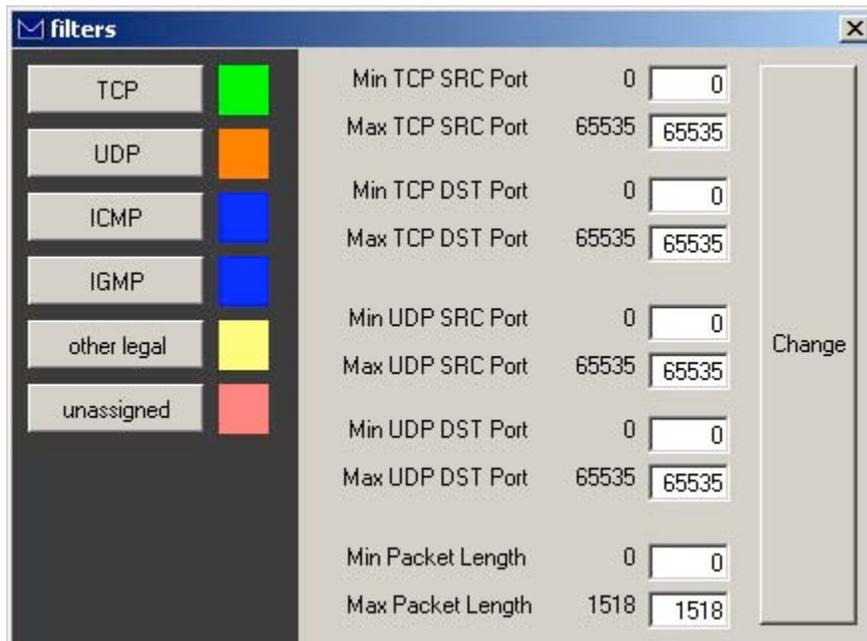
src/dst IP

src/dst TCP&UDP port

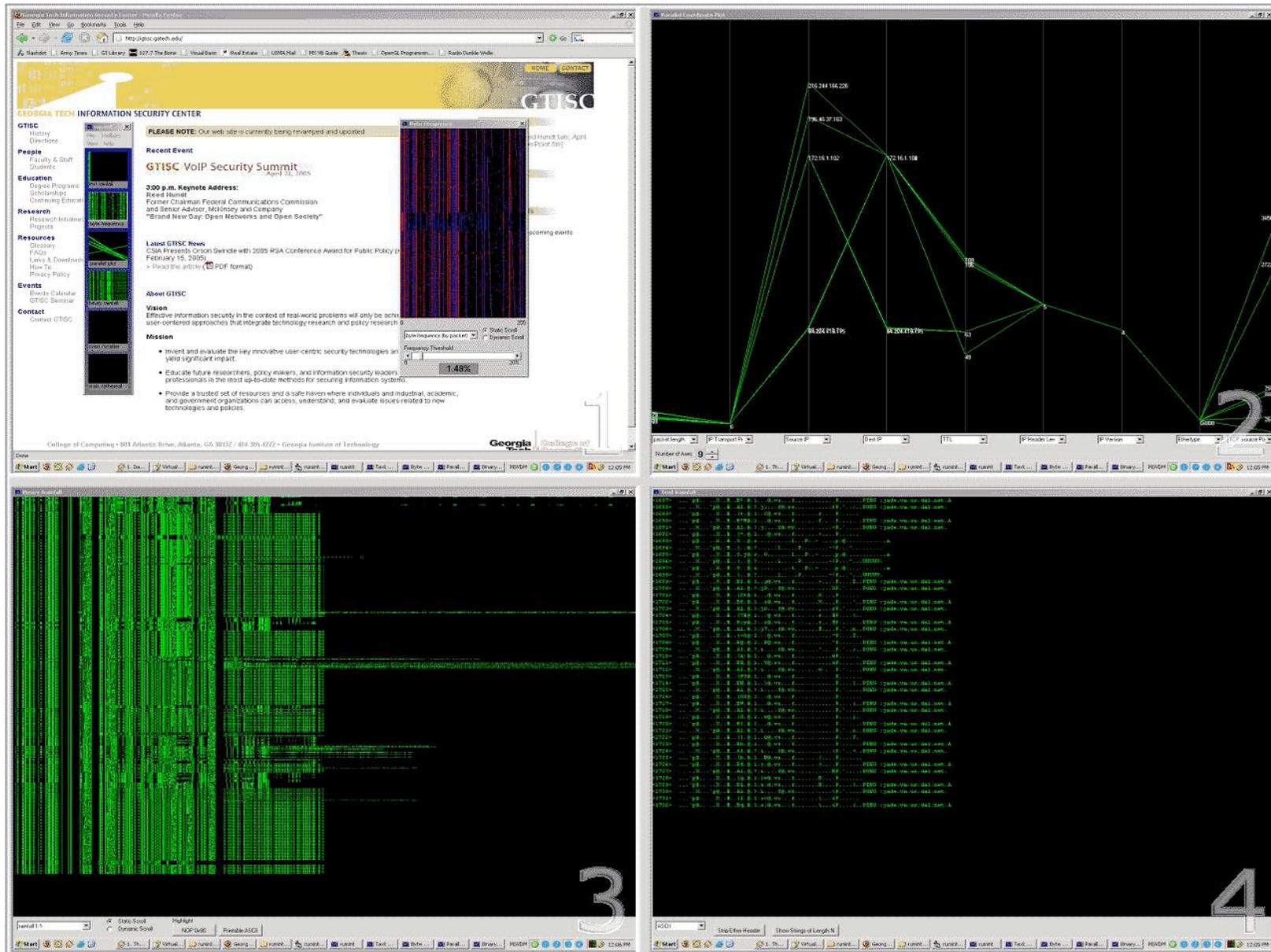
RUMINT



Filtering, Encoding & Interaction

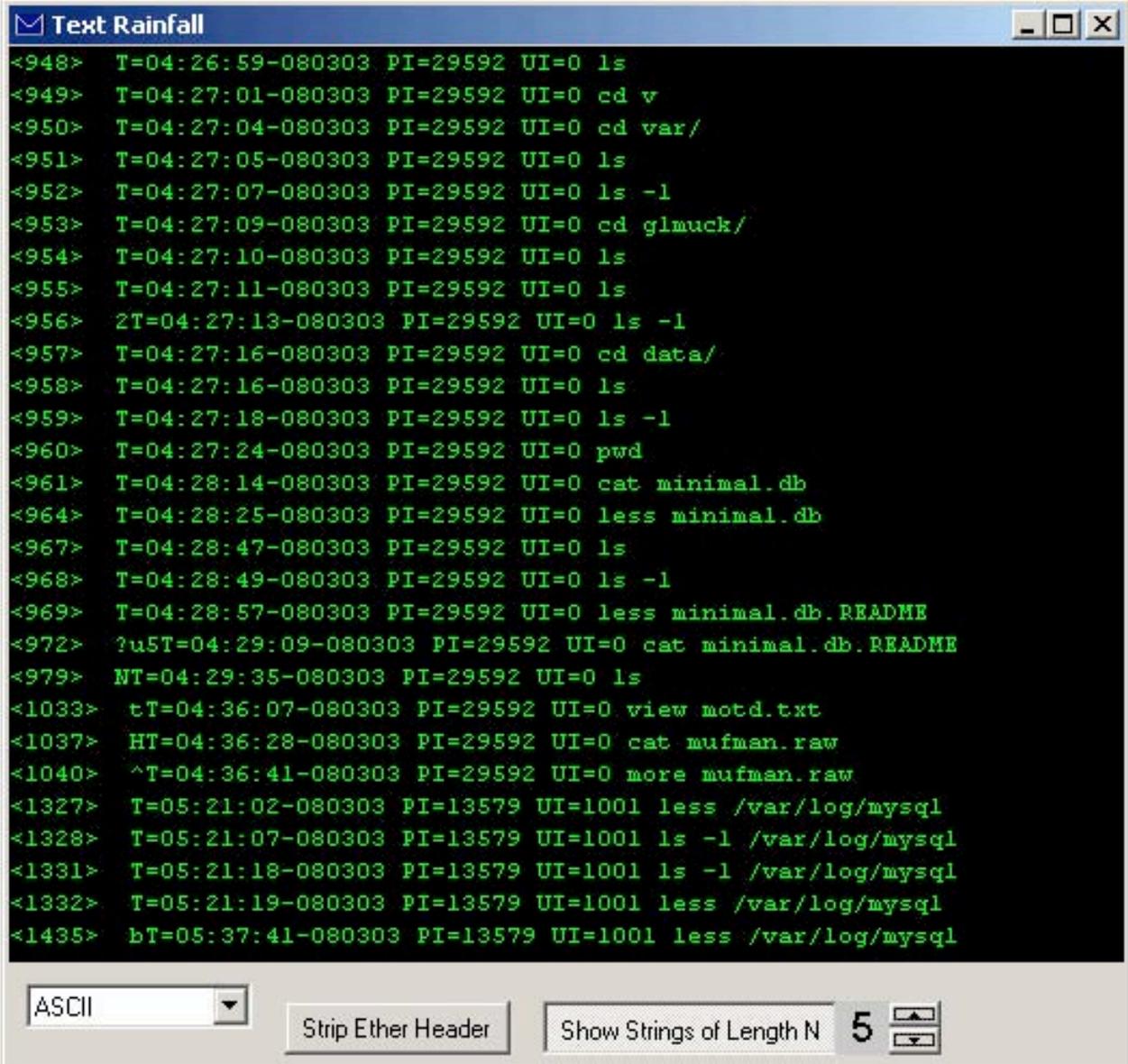


Multiple Coordinated Views...



Text

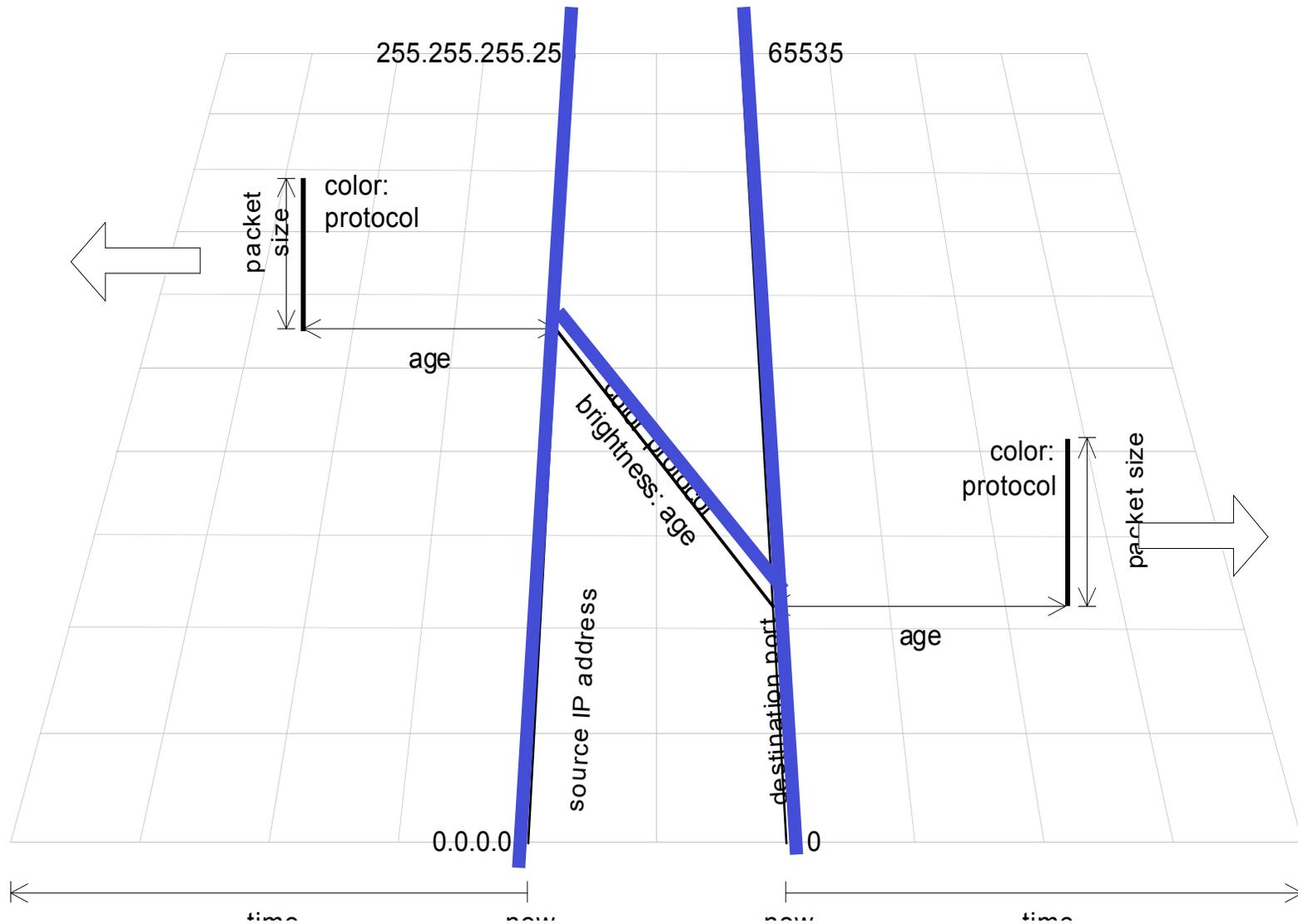
(on the fly strings)



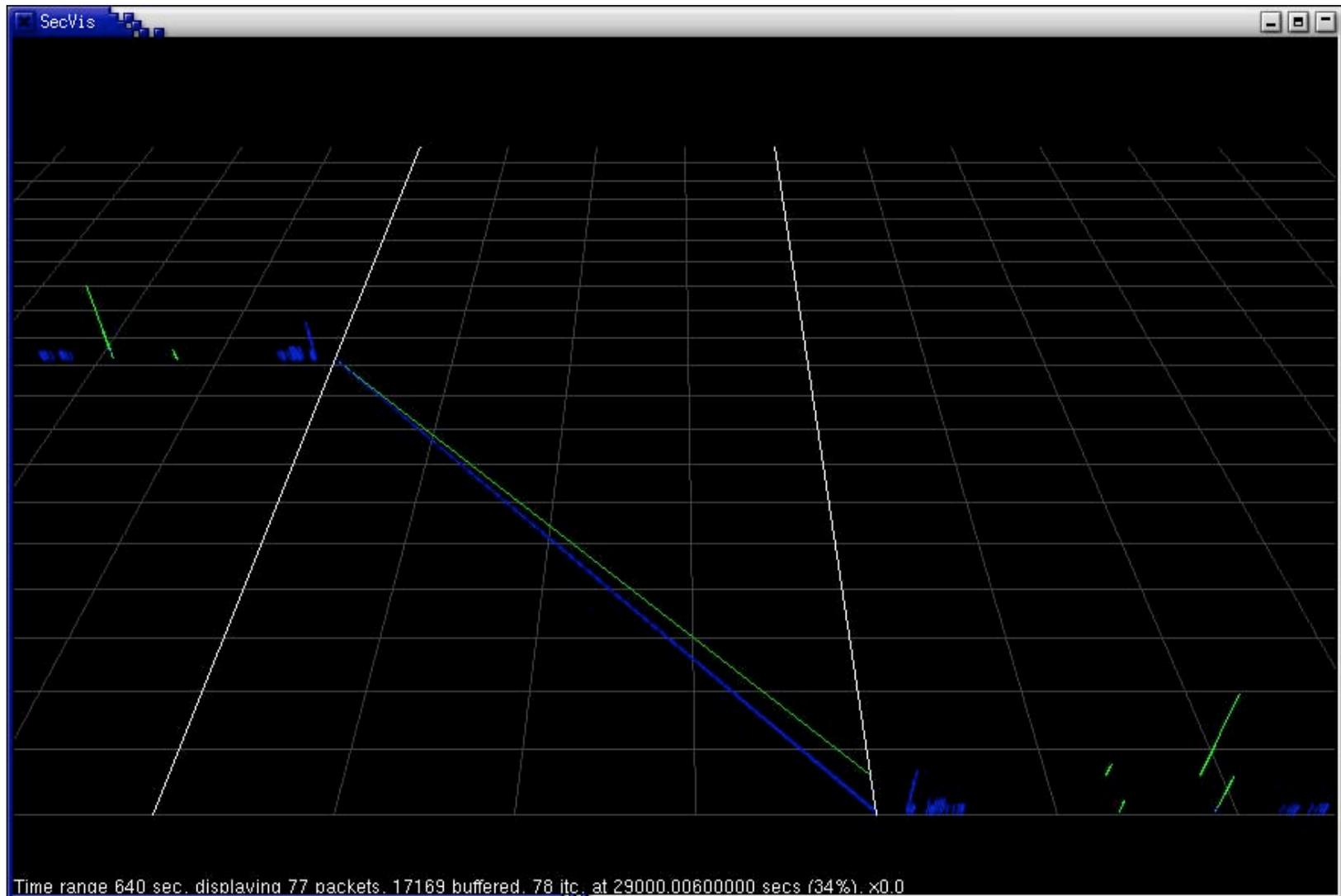
```
Text Rainfall
<948> T=04:26:59-080303 PI=29592 UI=0 ls
<949> T=04:27:01-080303 PI=29592 UI=0 cd v
<950> T=04:27:04-080303 PI=29592 UI=0 cd var/
<951> T=04:27:05-080303 PI=29592 UI=0 ls
<952> T=04:27:07-080303 PI=29592 UI=0 ls -l
<953> T=04:27:09-080303 PI=29592 UI=0 cd glmuck/
<954> T=04:27:10-080303 PI=29592 UI=0 ls
<955> T=04:27:11-080303 PI=29592 UI=0 ls
<956> T=04:27:13-080303 PI=29592 UI=0 ls -l
<957> T=04:27:16-080303 PI=29592 UI=0 cd data/
<958> T=04:27:16-080303 PI=29592 UI=0 ls
<959> T=04:27:18-080303 PI=29592 UI=0 ls -l
<960> T=04:27:24-080303 PI=29592 UI=0 pwd
<961> T=04:28:14-080303 PI=29592 UI=0 cat minimal.db
<964> T=04:28:25-080303 PI=29592 UI=0 less minimal.db
<967> T=04:28:47-080303 PI=29592 UI=0 ls
<968> T=04:28:49-080303 PI=29592 UI=0 ls -l
<969> T=04:28:57-080303 PI=29592 UI=0 less minimal.db.README
<972> ?u5T=04:29:09-080303 PI=29592 UI=0 cat minimal.db.README
<979> NT=04:29:35-080303 PI=29592 UI=0 ls
<1033> tT=04:36:07-080303 PI=29592 UI=0 view motd.txt
<1037> HT=04:36:28-080303 PI=29592 UI=0 cat mufman.raw
<1040> ^T=04:36:41-080303 PI=29592 UI=0 more mufman.raw
<1327> T=05:21:02-080303 PI=13579 UI=1001 less /var/log/mysql
<1328> T=05:21:07-080303 PI=13579 UI=1001 ls -l /var/log/mysql
<1331> T=05:21:18-080303 PI=13579 UI=1001 ls -l /var/log/mysql
<1332> T=05:21:19-080303 PI=13579 UI=1001 less /var/log/mysql
<1435> bT=05:37:41-080303 PI=13579 UI=1001 less /var/log/mysql
```

ASCII [v] Strip Ether Header Show Strings of Length N 5 [v]

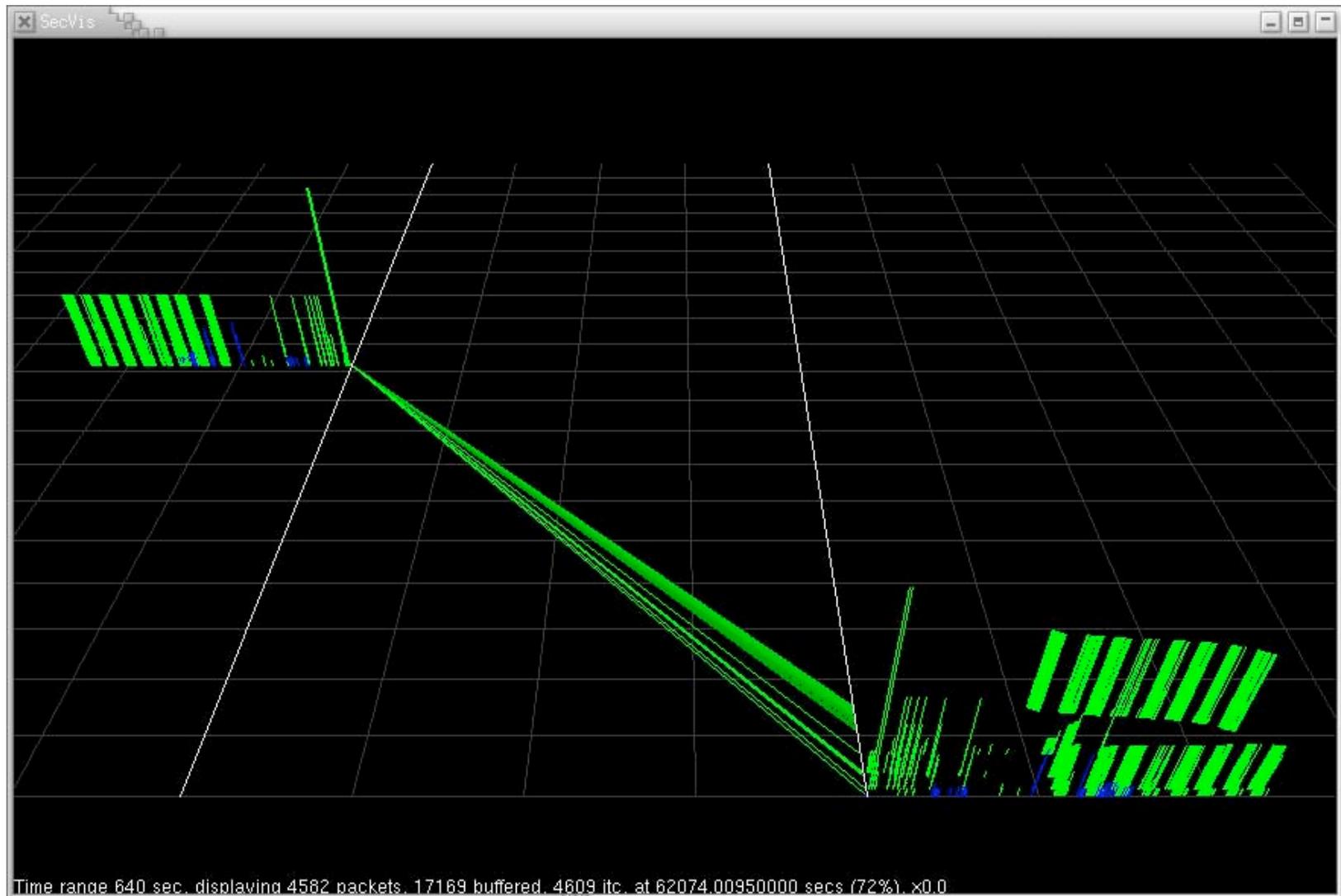
Krasser Visualization



Routine Honeynet Traffic (baseline)



Compromised Honeypot



Binary Rainfall Visualization

(single packet)

Bits on wire...

0	1	1	0	1	1	1	0	0	1	0	1	0	1	0	0	1	0	1	1	1	1	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Binary Rainfall Visualization

(single packet)

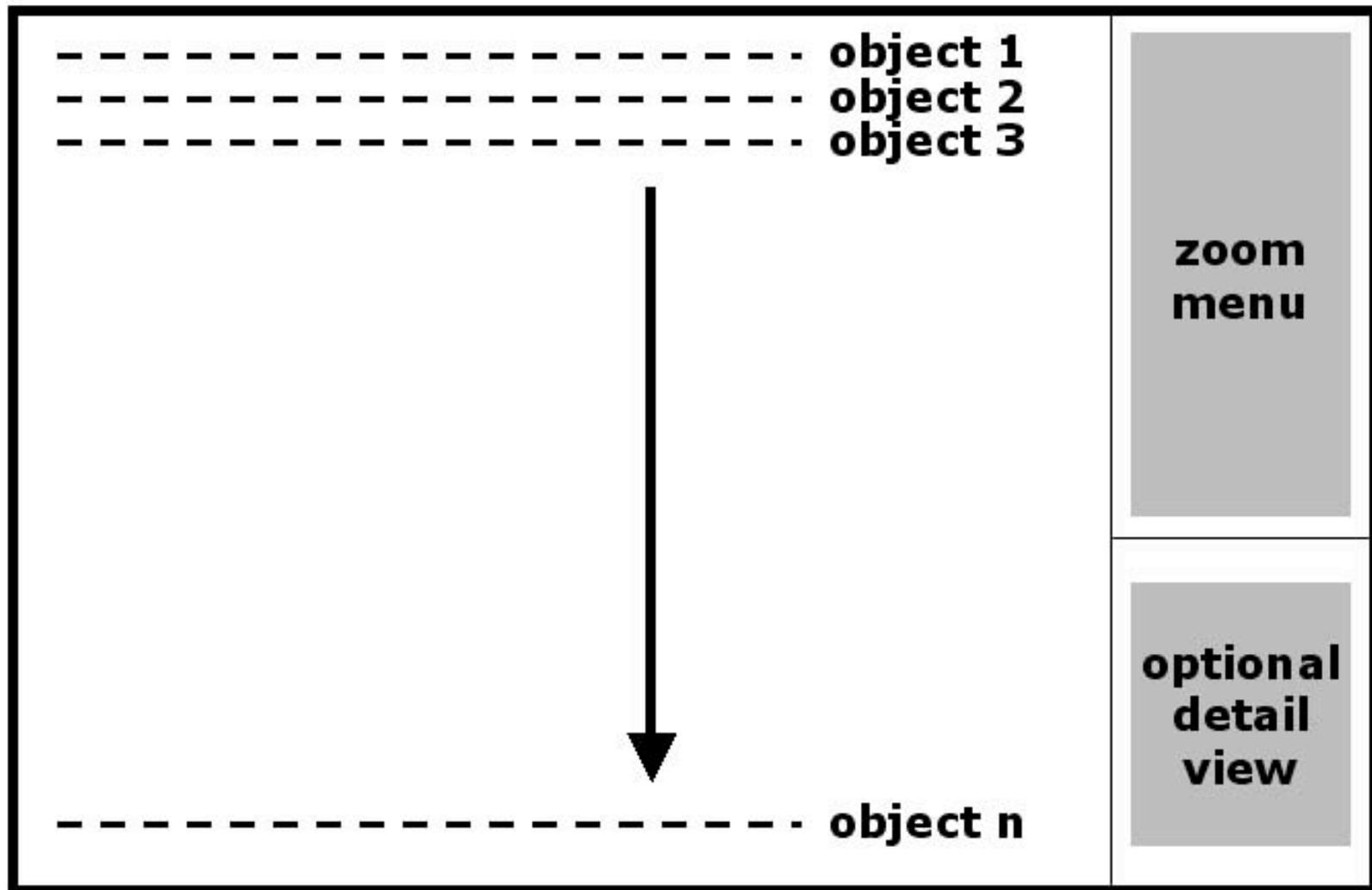
Bits on wire...

0	1	1	0	1	1	1	0	0	1	0	1	0	1	0	0	1	0	1	1	1	1	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

View as a 1:1 relationship (1 bit per pixel)...

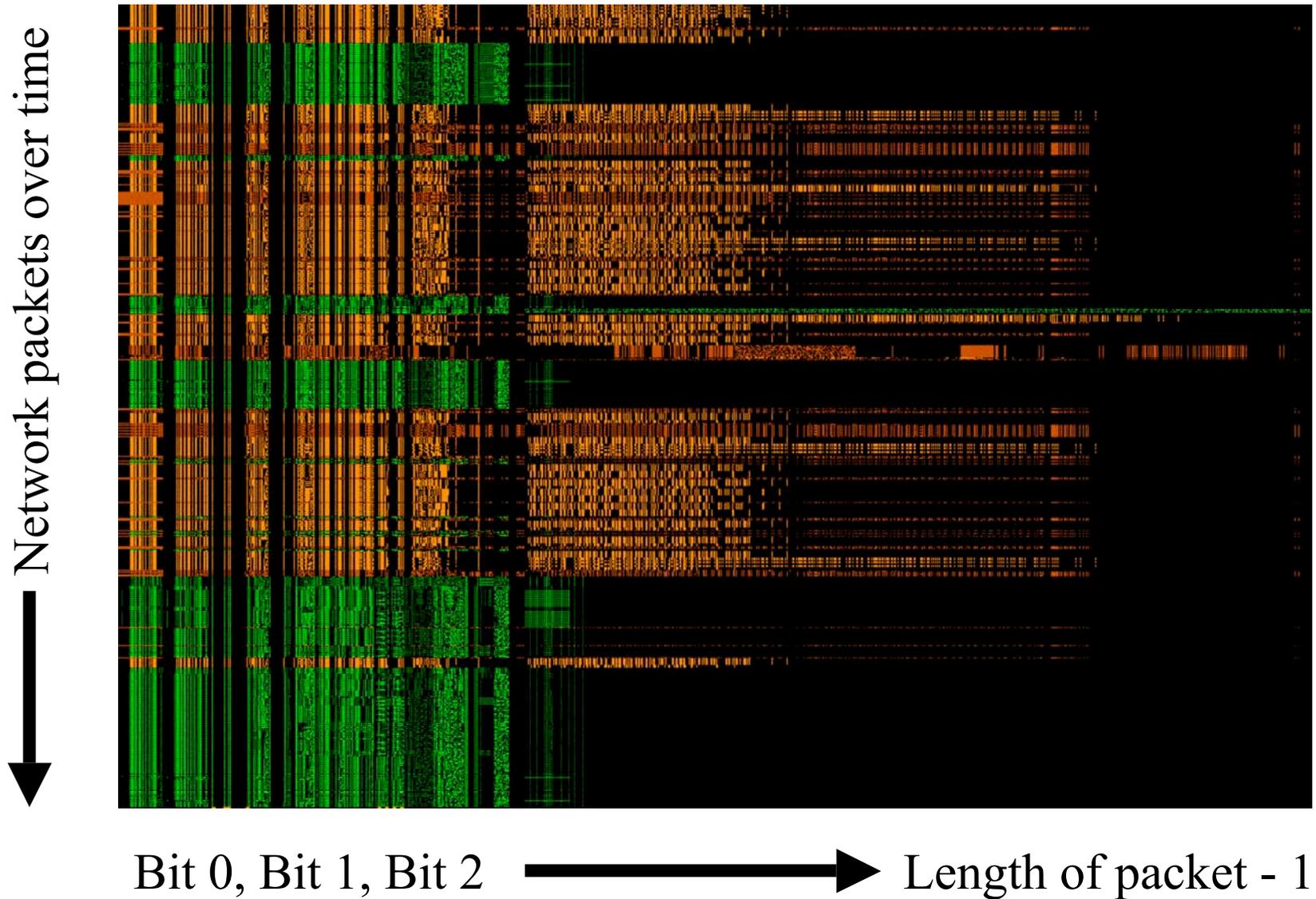
0	1	1	0	1	1	1	0	0	1	0	1	0	1	0	0	1	0	1	1	1	1	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

24 Pixels

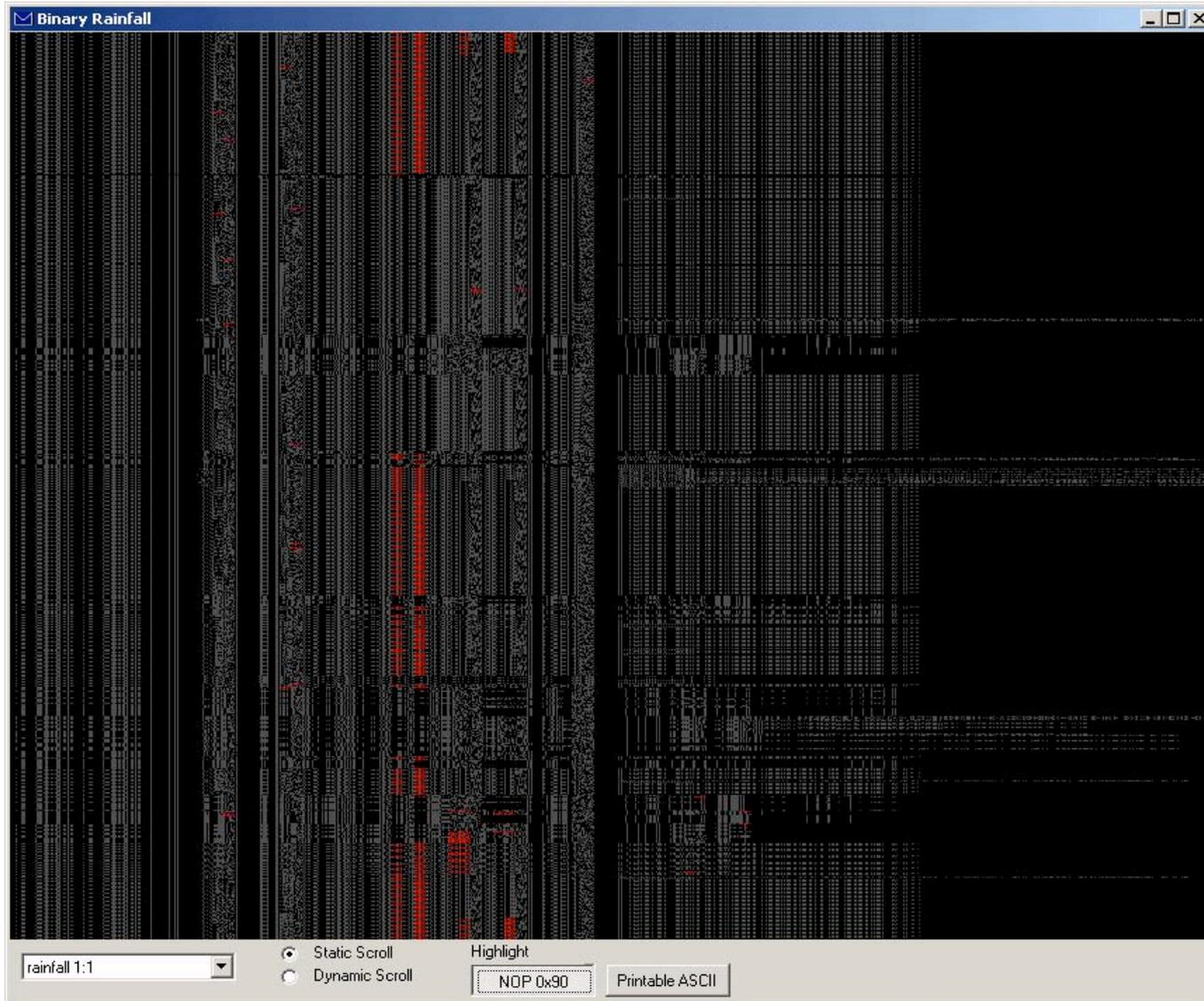


bit 1 bit n

Encode by Protocol



On the fly *disassembly*?



Binary Rainfall Visualization

(single packet)

Bits on wire...

0	1	1	0	1	1	1	0	0	1	0	1	0	1	0	0	1	0	1	1	1	1	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

View as a 1:1 relationship (1 bit per pixel)...

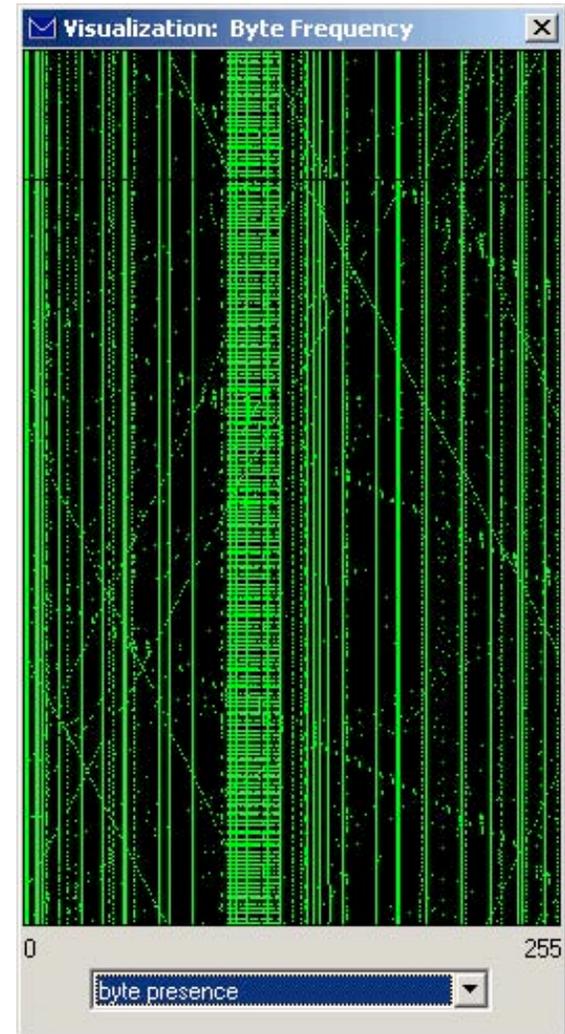
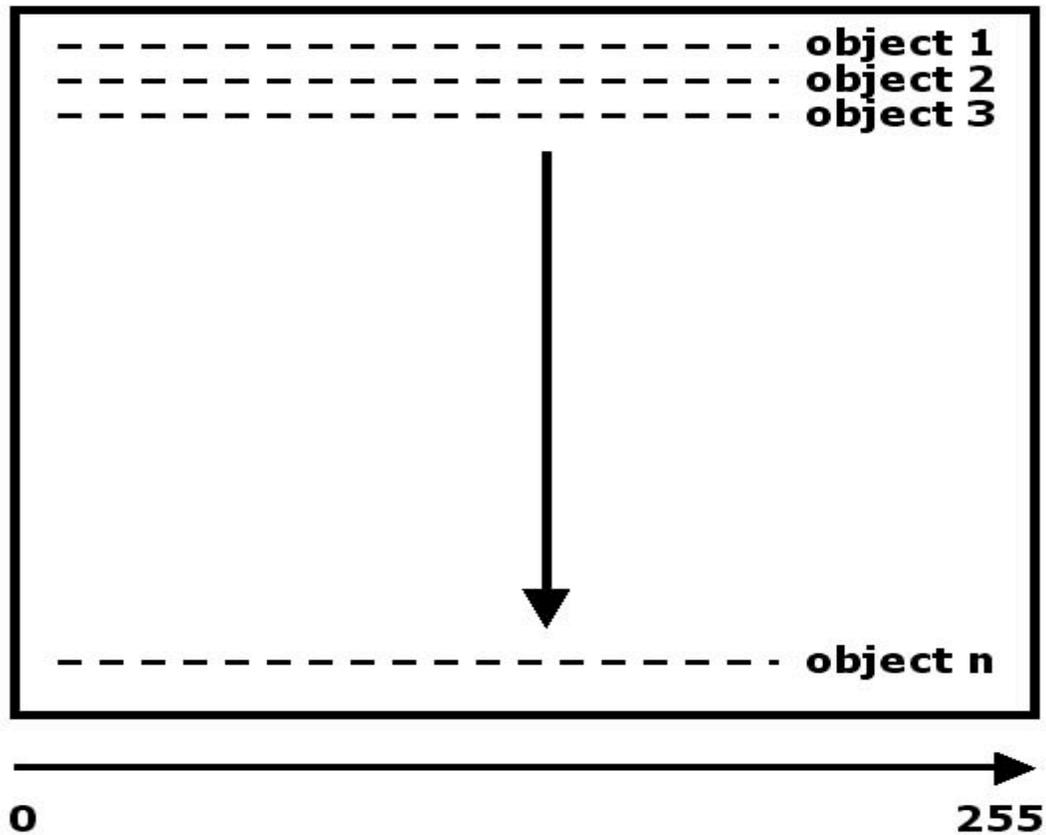
0	1	1	0	1	1	1	0	0	1	0	1	0	1	0	0	1	0	1	1	1	1	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

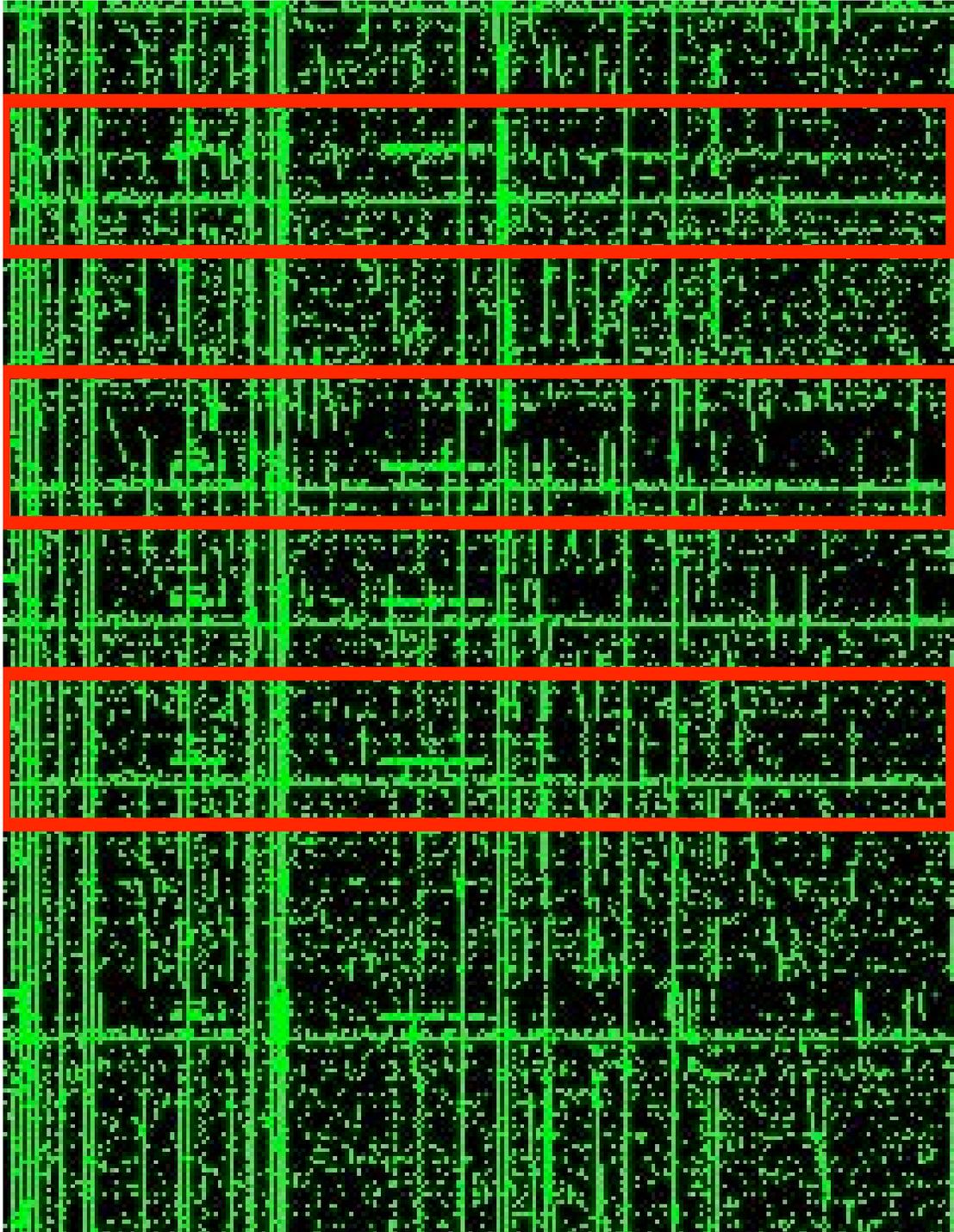
View as a 8:1 relationship (1 byte per pixel)...

0	1	1	0	1	1	1	0	0	1	0	1	0	1	0	0	1	0	1	1	1	1	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

3 Pixels

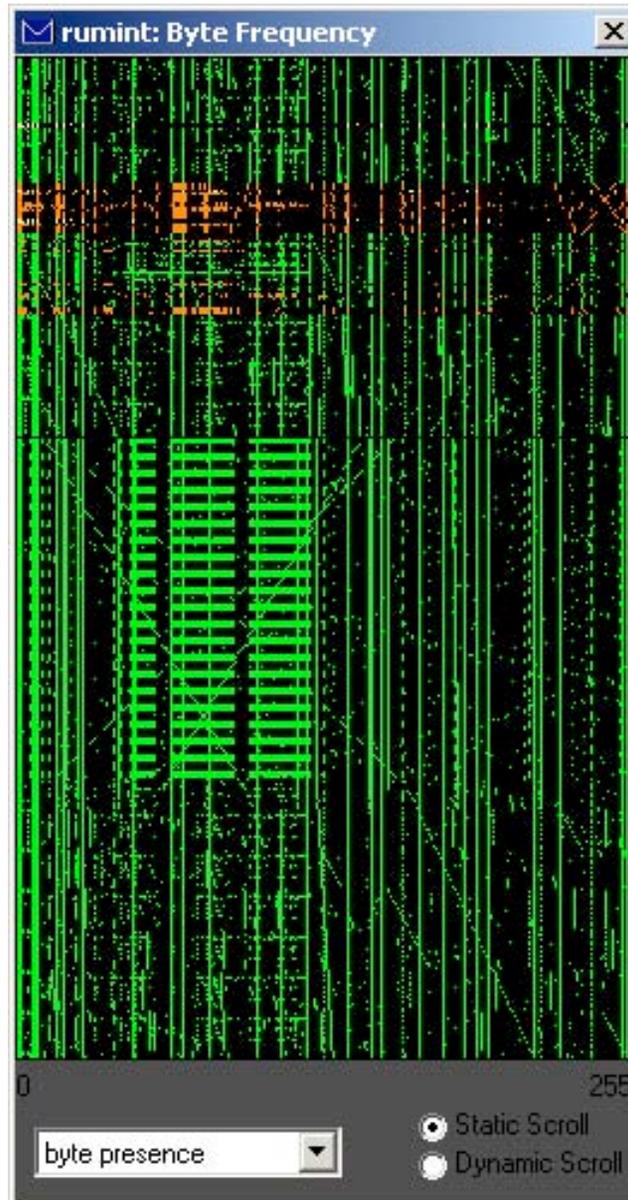
Byte Visualization





Open SSH Diffie- Hellman Key Exchange

Zipped Email Attachment



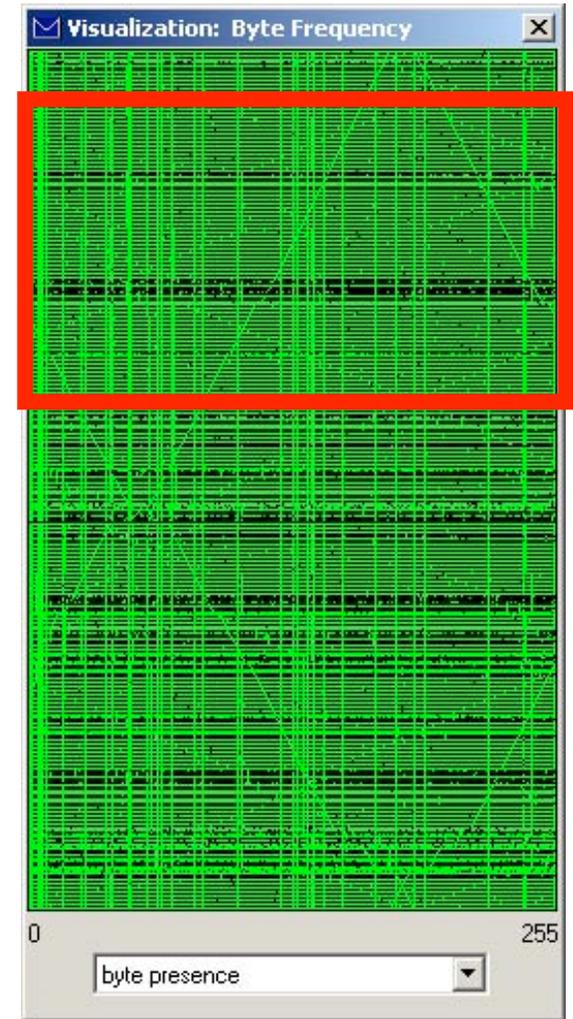
Byte Presence



dictionary file via HTTP

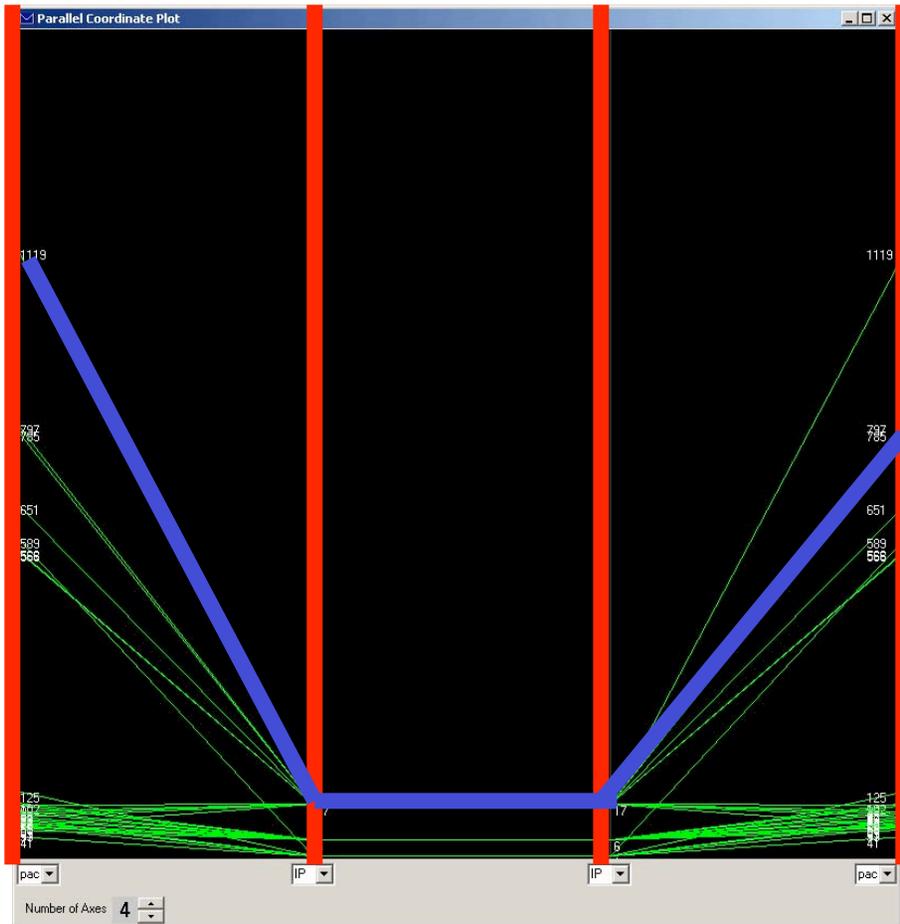


ssh



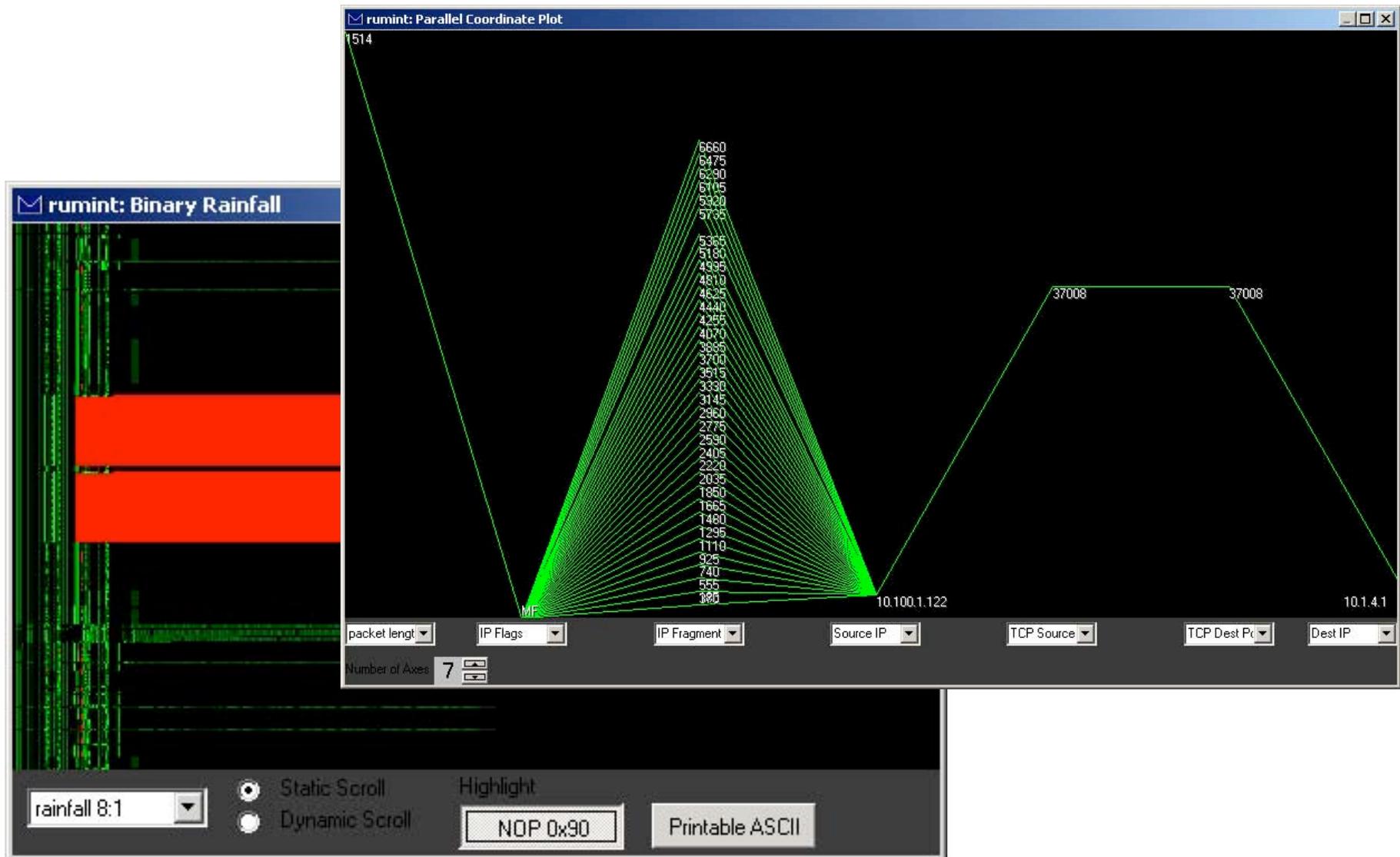
SSL

Parallel Coordinates

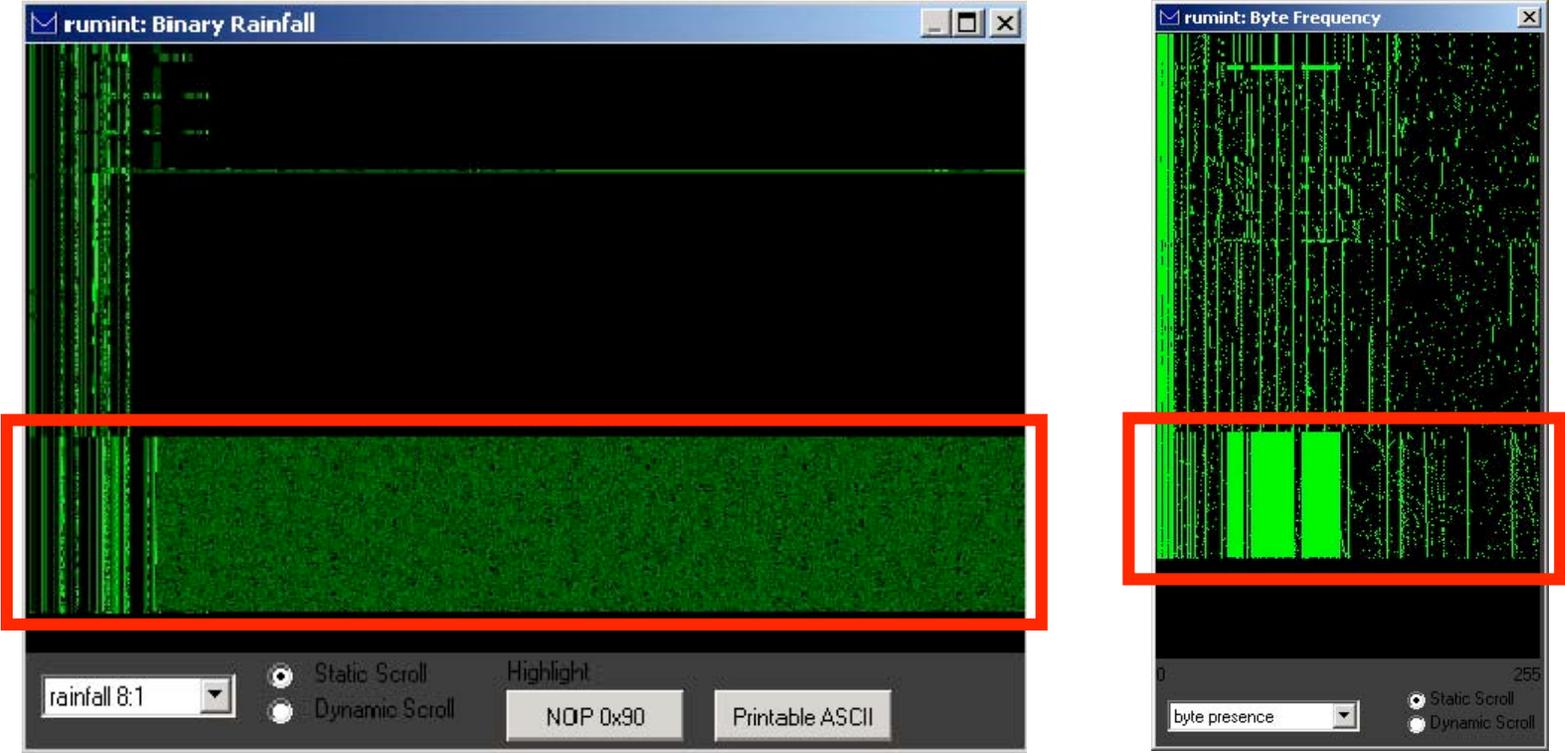


- goal: plot any data fields
- dynamic columns
- change order for different insight
- intelligent lookup and translation of fields
 - e.g. IP transport protocol

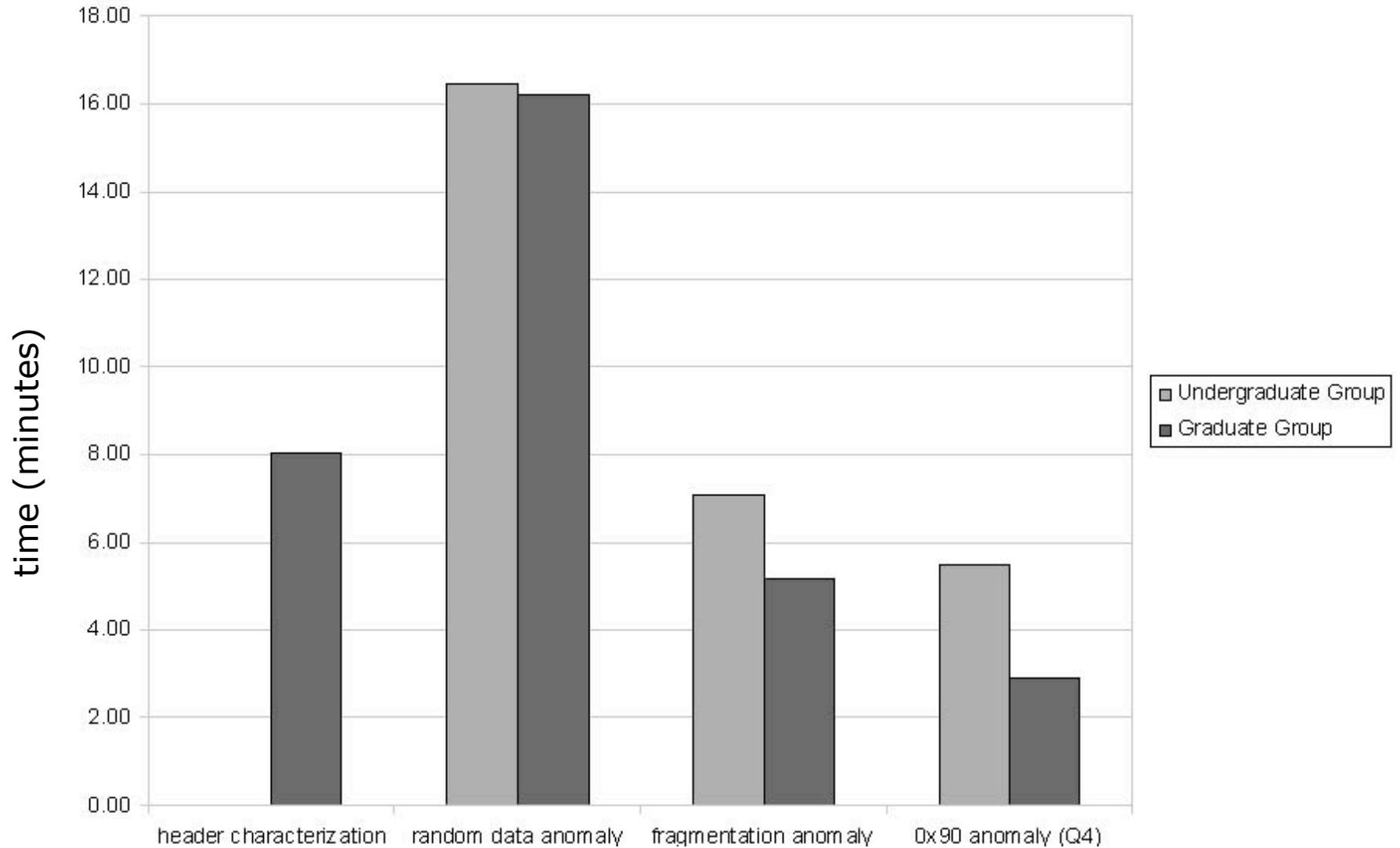
Identify and Precisely Locate x90 Anomaly



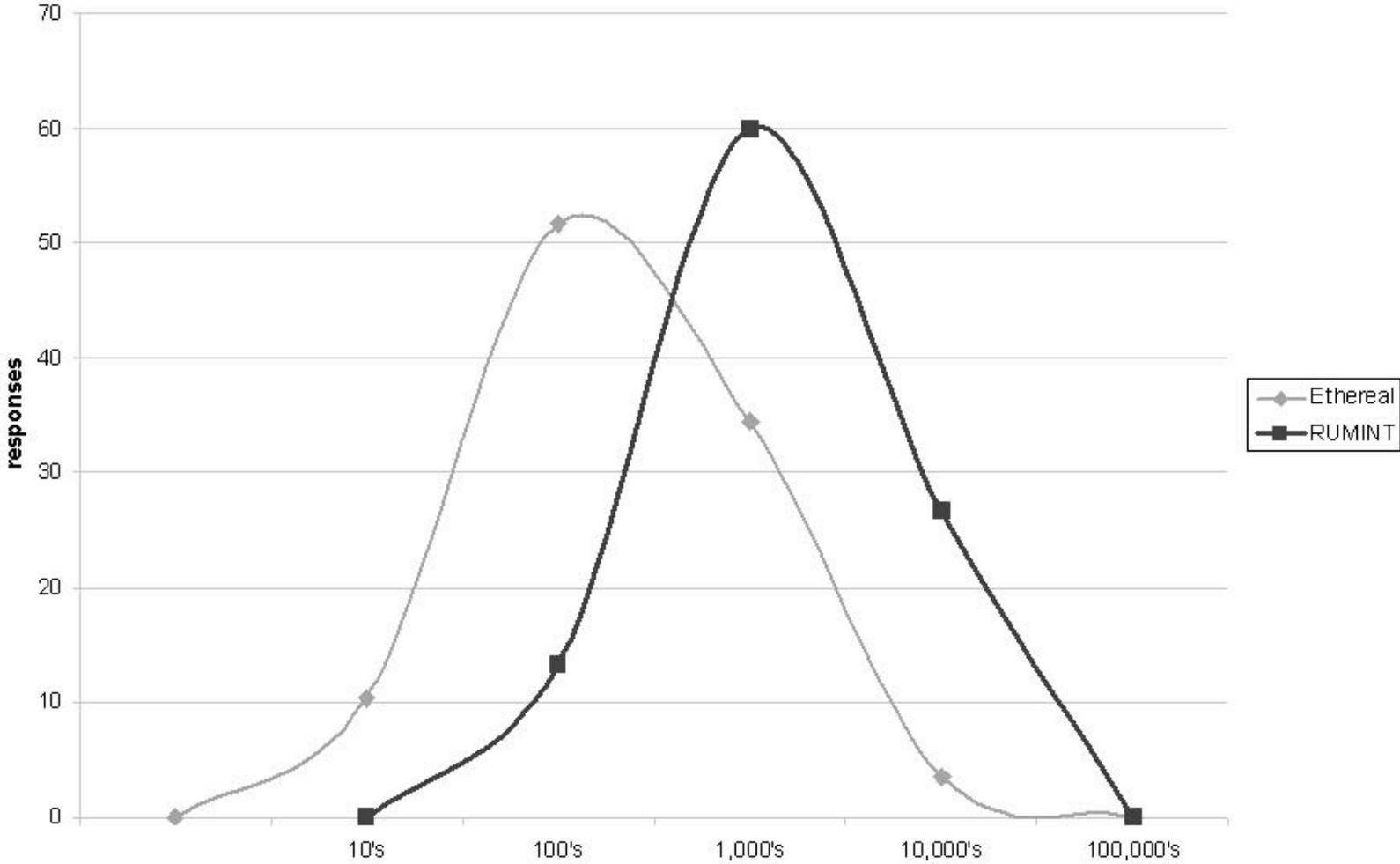
Identify and Precisely Locate Possible Random Payload Anomaly



Task Completion Time



RUMINT Tipping Point



System Requirements

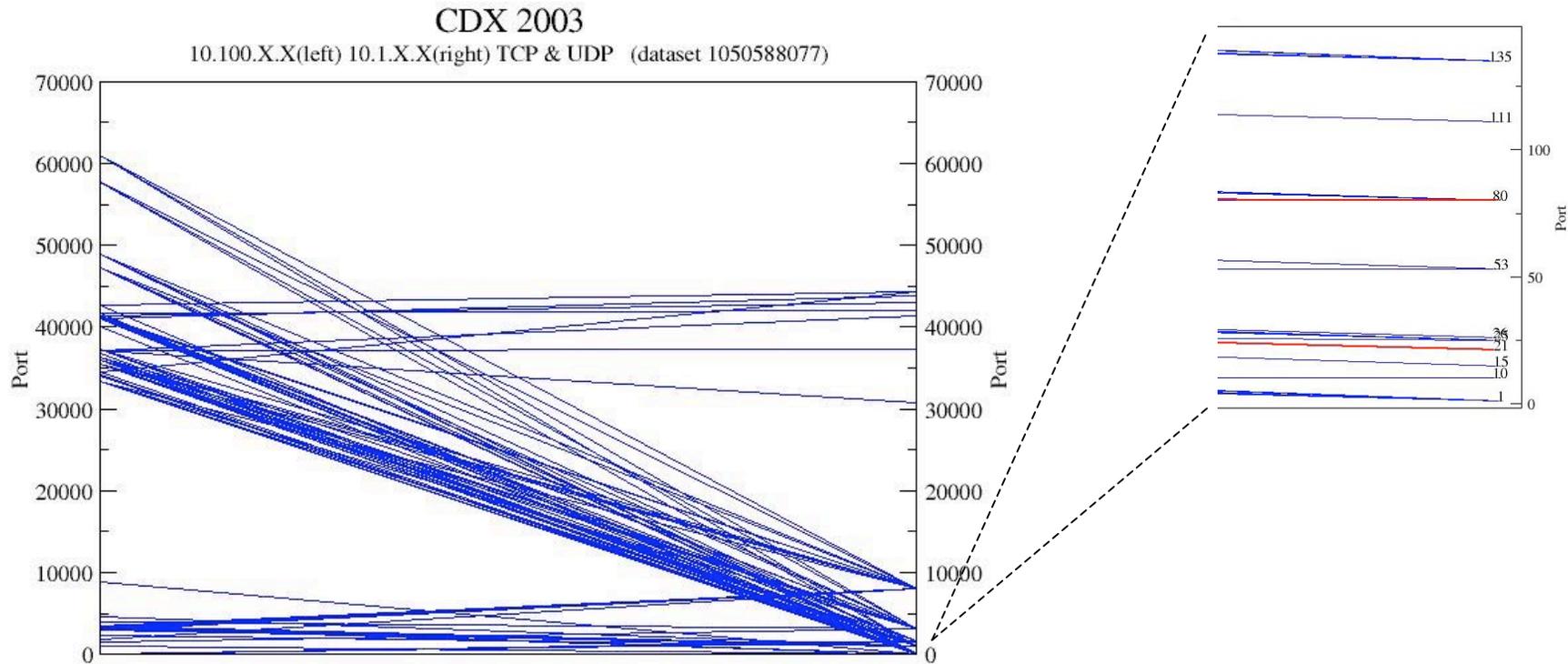
- IP over Ethernet
- Tested on Windows XP
- ~256+ MB Ram
- Processor 300MHZ (minimum)
- The more screen real estate the better
- Latest winpcap

- Development
 - Visual Studio 6
 - port to GCC and Open GL
 - PacketX for now
 - Go direct to (win)pcap

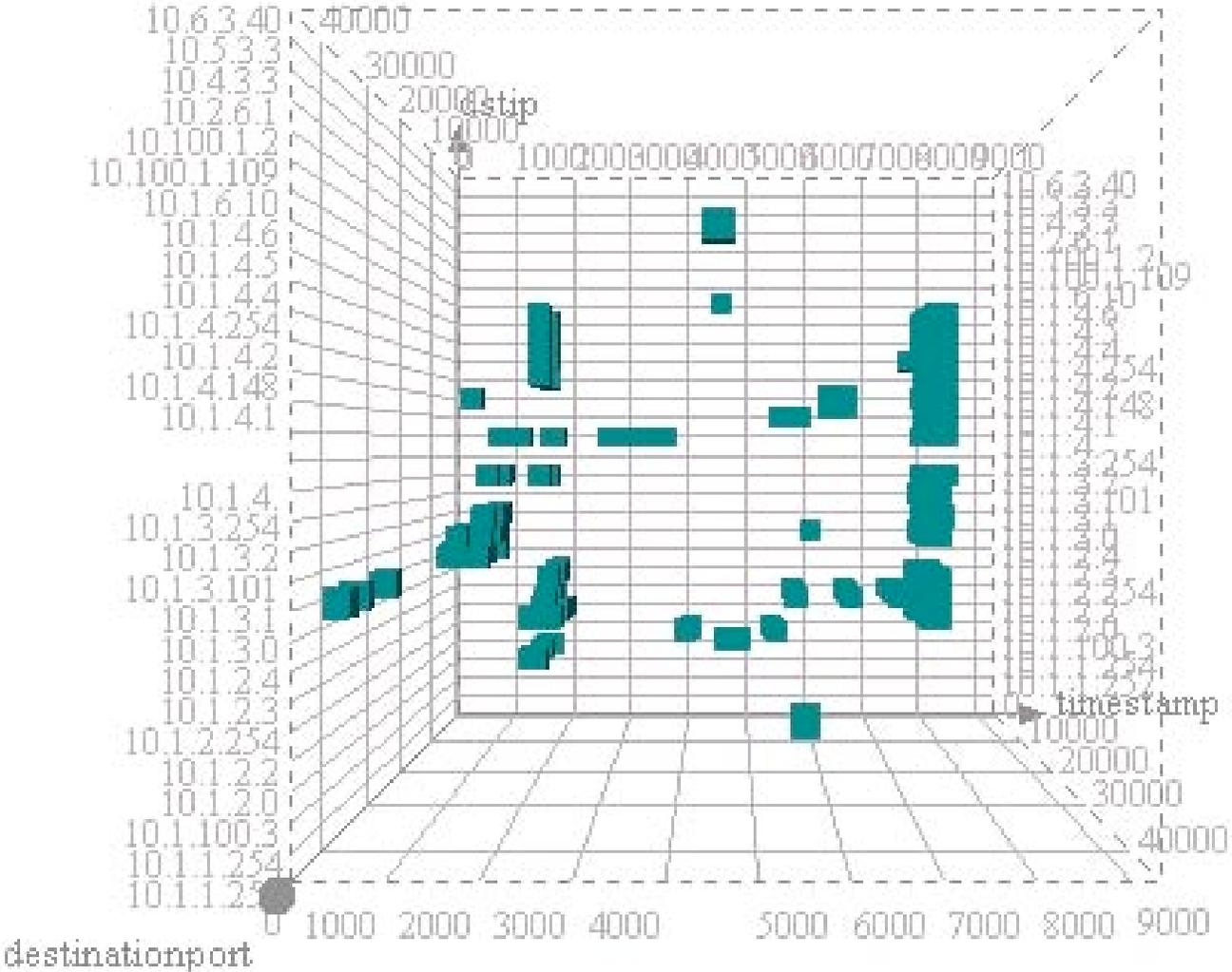
Demo

Attacking the Analyst

AutoScale Attack/Force User to Zoom

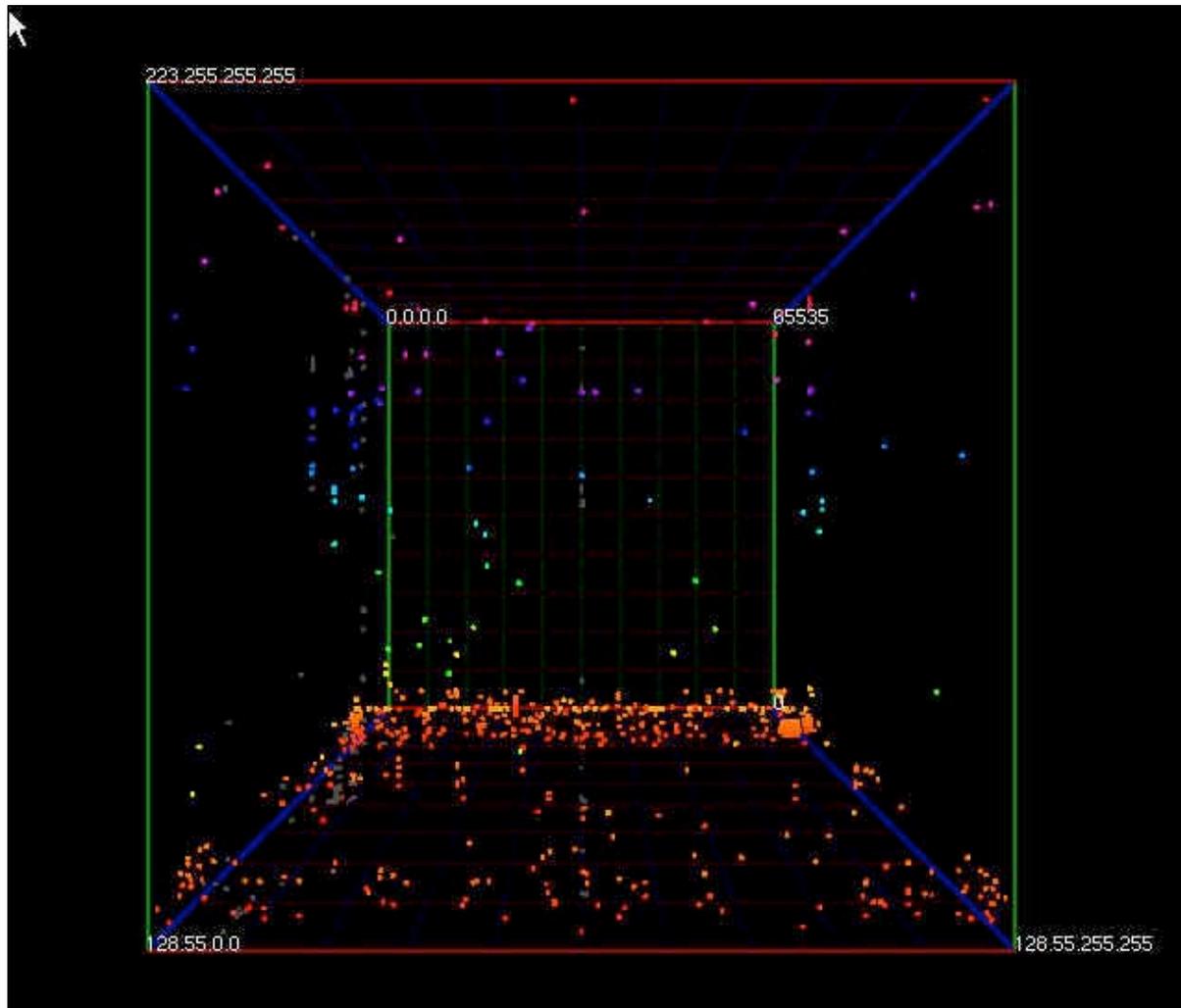


Labeling Attack



CDX 2003 Dataset
X = Time
Y = Destination IP
Z = Destination Port

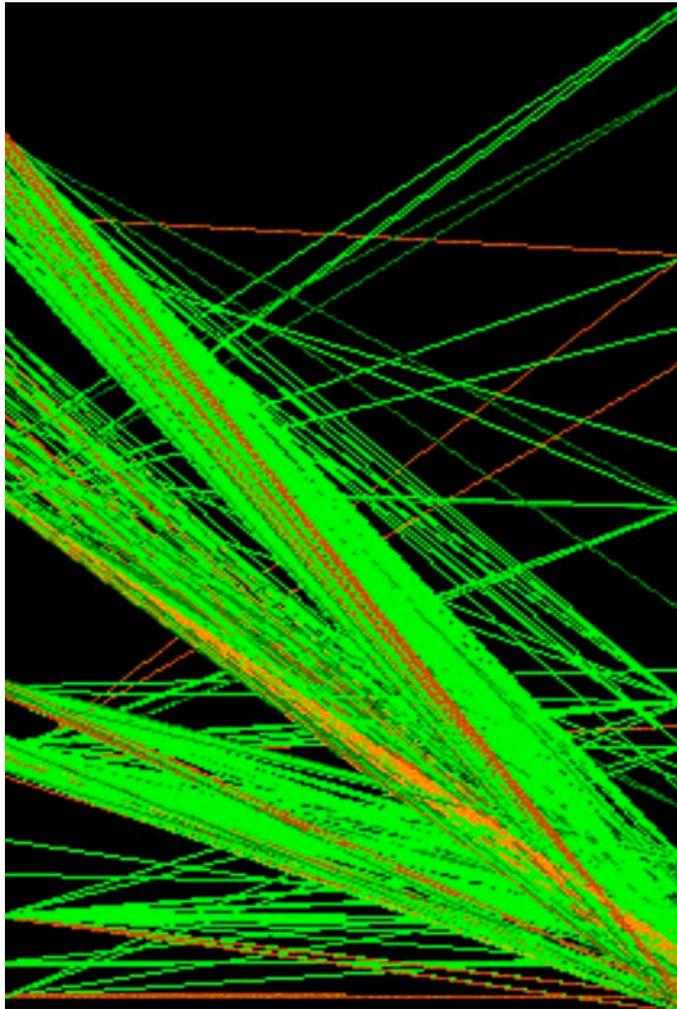
Precision Attack



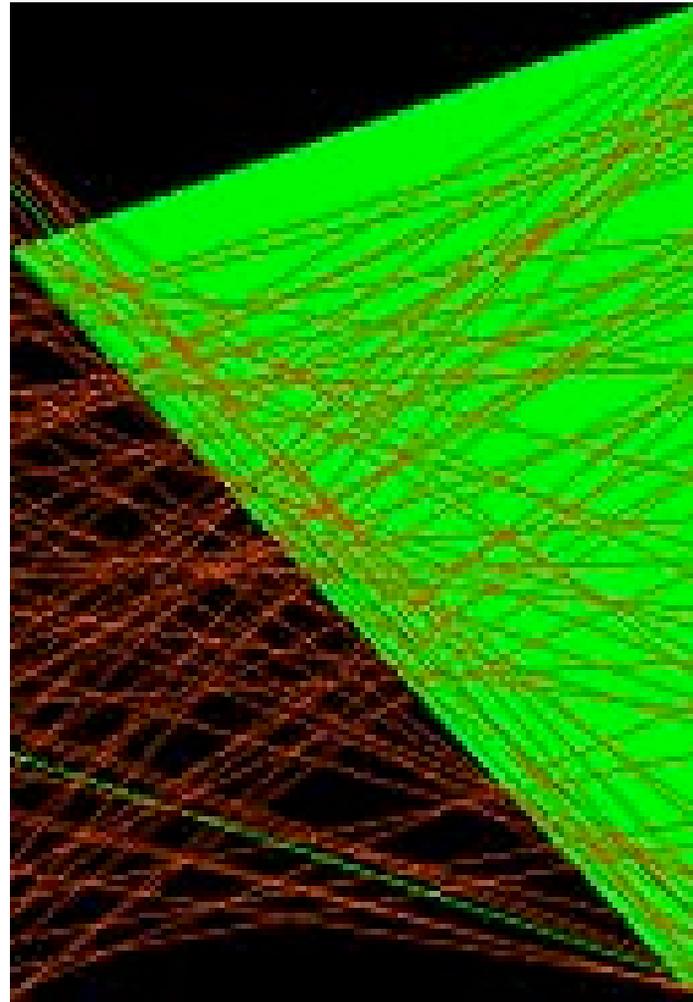
<http://www.nersc.gov/nusers/security/Cube.jpg>

<http://developers.slashdot.org/article.pl?sid=04/06/01/1747223&mode=thread&tid=126&tid=172>

Occlusion



Jamming



Attack Demo

Attacking the Analyst...

G. Conti, M. Ahamad and J. Stasko;
"Attacking Information Visualization System Usability: Overloading and Deceiving the Human;" *Symposium on Usable Privacy and Security (SOUPS)*; July 2005. [On the CD...](#)

G. Conti and M. Ahamad; "A Taxonomy and Framework for Countering Denial of Information Attacks;" *IEEE Security and Privacy*. (accepted, to be published) [Website...](#)

**Attacking Information Visualization System Usability
Overloading and Deceiving the Human**

Gregory Conti
College of Computing
Georgia Institute of Technology

Mustafa Ahamad
College of Computing
Georgia Institute of Technology

John Stasko
College of Computing
Georgia Institute of Technology

Abstract
Information visualization is an effective way to easily comprehend large amounts of data. For such systems to be truly effective, the information visualization designer must be aware of the ways in which their system may be manipulated and protect their users from attack. In addition, users should be aware of potential attacks in order to minimize or negate their effect. These attacks target the information visualization system as well as the perceptual, cognitive and motor capabilities of human end users. To identify and help counter these attacks we present a framework for information visualization system security analysis, a taxonomy of visualization attacks and technology independent principles for countering malicious visualizations. These themes are illustrated with case studies and working examples from the network security visualization domain, but are widely applicable to virtually any information visualization system.

CR Categories: [I.5.2 Information Systems]: Information Interfaces and Presentation - User Interfaces; [C.1.3 Computer-Communication Networks]: Network Operations: Network Monitoring; [C.2.0 Computer-Communication Networks]: General - Security and Protection

Keywords: malicious visualizations, usability attacks, denial of information, secure visualization, information visualization

1 Introduction
All but the most trivial visualization systems must be designed with security in mind. Information visualization systems are constantly exposed to attack, either from malicious entities attempting to overwhelm, mislead or distract the human viewer or from non-malicious entities that accomplish the same result by accident. Some might believe that today's systems are not potential targets for attack. Clearly there are many domains where security is of minimal importance, but increasingly information visualization systems are being used to support critical decision making. For example, intelligence analysis, law enforcement, network security and business decision-support systems exist in an adversarial environment where it is likely that malicious entities are actively attempting to manipulate human end users. We believe that there is a clear threat today and there will be a growing problem into the foreseeable future. For information visualization systems to maintain relevance security must be considered. Information visualization systems inherently have the human tightly coupled in the system loop. In most cases, the human is the decision maker who will act upon our not act.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Symposium on Usable Privacy and Security '05, July 6-8, 2005, Pittsburgh, PA, USA.
Copyright 2005 ACM 1-58113-810-0/05/000004...\$5.00.

upon the information presented and, as a result, the human is a high-target and likely target. Data points in the information visualization system may be attacked, from data collection to processing to final visualization, in order to impact human interpretation. A "minor" compromise of a single bit may have significant impact on the human; consider a change in the foreground color of a scatter plot to the background color. Major compromises may have far greater impact. Our primary goal is to identify these threats and vulnerabilities, as well as develop principles to counter or mitigate these attacks. By identifying the threats and weaknesses of their system, designers can make appropriate decisions to mitigate these vulnerabilities.

To see a simple attack in action, consider a visual intrusion detection system designed to supplement classical anomaly-based and signature-based intrusion detection systems. Such systems are typically co-located with a firewall at the border between the internal institutional network and the public Internet. This vantage point allows the system to observe and collect selected data from network traffic at entry and egress from the internal network. Our example system collects header data from network traffic and visualizes it in real-time. In particular, it captures the source and destination addresses of communicating network nodes, network protocols in use, source and destination ports (used for process to process communication across an Internet Protocol (IP) network, e.g. port 80 for a web server) as well as calculates a checksum for each record. An adversary may easily inject arbitrary data into the visualization system, intermingled with legitimate users' traffic, due to weaknesses in current networking protocols. In our example, the adversary knows the system operator on the night shift is red-green colorblind. They also know that the default settings on the visualization system map the very common 09A+% of traffic Transmission Control Protocol (TCP) to green, the User Datagram Protocol (UDP) to blue and the Internet Control Management (ICMP) protocol to red. In addition, the attacker knows that the target node has serious ICMP and UDP vulnerabilities. The attacker waits until late in the operator's shift and launches an ICMP based attack. The already tired operator does not notice the red packet amidst the much greater noise of green packets. In this case, the attacker took advantage of the visualization system's color mapping to target a specific user, but many other techniques could have been used. We will describe and illustrate these attacks in later sections.

To help counter usability attacks against visualization systems this work includes several novel contributions: a framework for information visualization system security analysis, a taxonomy of malicious attacks as well as technology independent principles for designing information visualization systems that will resist attack. We illustrate and validate these contributions with results from the design, implementation and real-world use of a visual network intrusion detection system [1].

Future Vision



Directions for the Future...

We are only scratching the surface of the possibilities

- **attack specific community needs**
- **plug-ins**
- launch network packets?
- **protocol specific visualizations**
 - including application layer (e.g. VoIP, HTTP)
- Open GL
- graph visualization+
- screensaver/wallpaper snapshot?
- work out GUI issues
- database of filters / smart books
- stress testing
- evaluate effectiveness

For more information...

- G. Conti, K. Abdullah, J. Grizzard, J. Stasko, J. Copeland, M. Ahamad, H. Owen and C. Lee; "Countering Security Analyst and Network Administrator Overload Through Alert and Packet Visualization;" IEEE Computer Graphics and Applications (CG&A), March 2006.
- G. Conti, J. Grizzard, M. Ahamad and H. Owen; "Visual Exploration of Malicious Network Objects Using Semantic Zoom, Interactive Encoding and Dynamic Queries;" IEEE Symposium on Information Visualization's Workshop on Visualization for Computer Security (VizSEC); October 2005.
- G. Conti; "Beyond Ethereal: Crafting A Tivo for Security Datastreams;" Black Hat USA; July 2005.
- G. Conti, M. Ahamad and J. Stasko; "Attacking Information Visualization System Usability: Overloading and Deceiving the Human;" Symposium on Usable Privacy and Security (SOUPS); July 2005.
- S. Krasser, G. Conti, J. Grizzard, J. Gribschaw and H. Owen; "Real-Time and Forensic Network Data Analysis Using Animated and Coordinated Visualization;" IEEE Information Assurance Workshop (IAW); June 2005.
- G. Conti; "Countering Denial of Information Attacks with Information Visualization;" InterzOne 4; March 2005.
- G. Conti and K. Abdullah; " Passive Visual Fingerprinting of Network Attack Tools;" ACM Conference on Computer and Communications Security's Workshop on Visualization and Data Mining for Computer Security (VizSEC); October 2004.
- G. Conti; "Network Attack Visualization;" DEFCON 12; August 2004.
- G. Conti; "Network Security Data Visualization;" InterzOne3; April 2004.

www.cc.gatech.edu/~conti
www.rumint.org

On the CD...

- Talk slides
- Code
 - rumint
- Papers
 - SOUPS Malicious Visualization paper
 - Hacker conventions article
 - Ethereal / Snort Survey



See also: www.cc.gatech.edu/~conti and www.rumint.org

Feedback Requested...

- Tasks
- Usage
 - provide feedback on GUI
 - needed improvements
 - multiple monitor machines
 - performance under stress
 - bug reports
- Data
 - interesting packet traces
 - screenshots
 - with supporting .rum and .pcap files, if possible
- Pointers to interesting related tools (viz or not)
- New viz and other analysis ideas

Acknowledgements

404.se2600, Kulsoom Abdullah, Sandip Agarwala, Mustaque Ahamad, Bill Cheswick, Chad, Clint, Tom Cross, David Dagon, DEFCON, Ron Dodge, EliO, Emma, Mr. Fuzzy, Jeff Gribshaw, Julian Grizzard, GTISC, Hacker Japan, Mike Hamelin, Hendrick, HoneyNet Project, InterzOne, Jinsuk Jun, Kenshoto, Oleg Kolesnikov, Sven Krasser, Chris Lee, Wenke Lee, John Levine, Michael Lynn, David Maynor, Neel Mehta, Jeff Moss, NETI@home, Henry Owen, Dan Ragsdale, Rockit, Byung-Uk Roho, Charles Robert Simpson, Ashish Soni, SOUPS, Jason Spence, John Stasko, Strick, Susan, USMA ITOC, IEEE IAW, VizSEC 2004, Grant Wagner and the Yak.



Greg Conti

conti@cc.gatech.edu

www.cc.gatech.edu/~conti

www.rumint.org