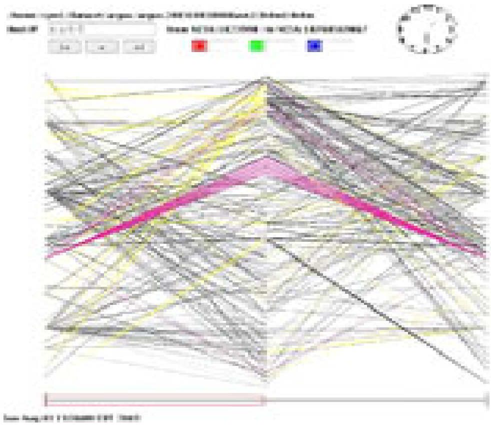


VisFlowConnect-IP:

A Link-Based Visualization of Netflows for Security Monitoring

William Yurcik

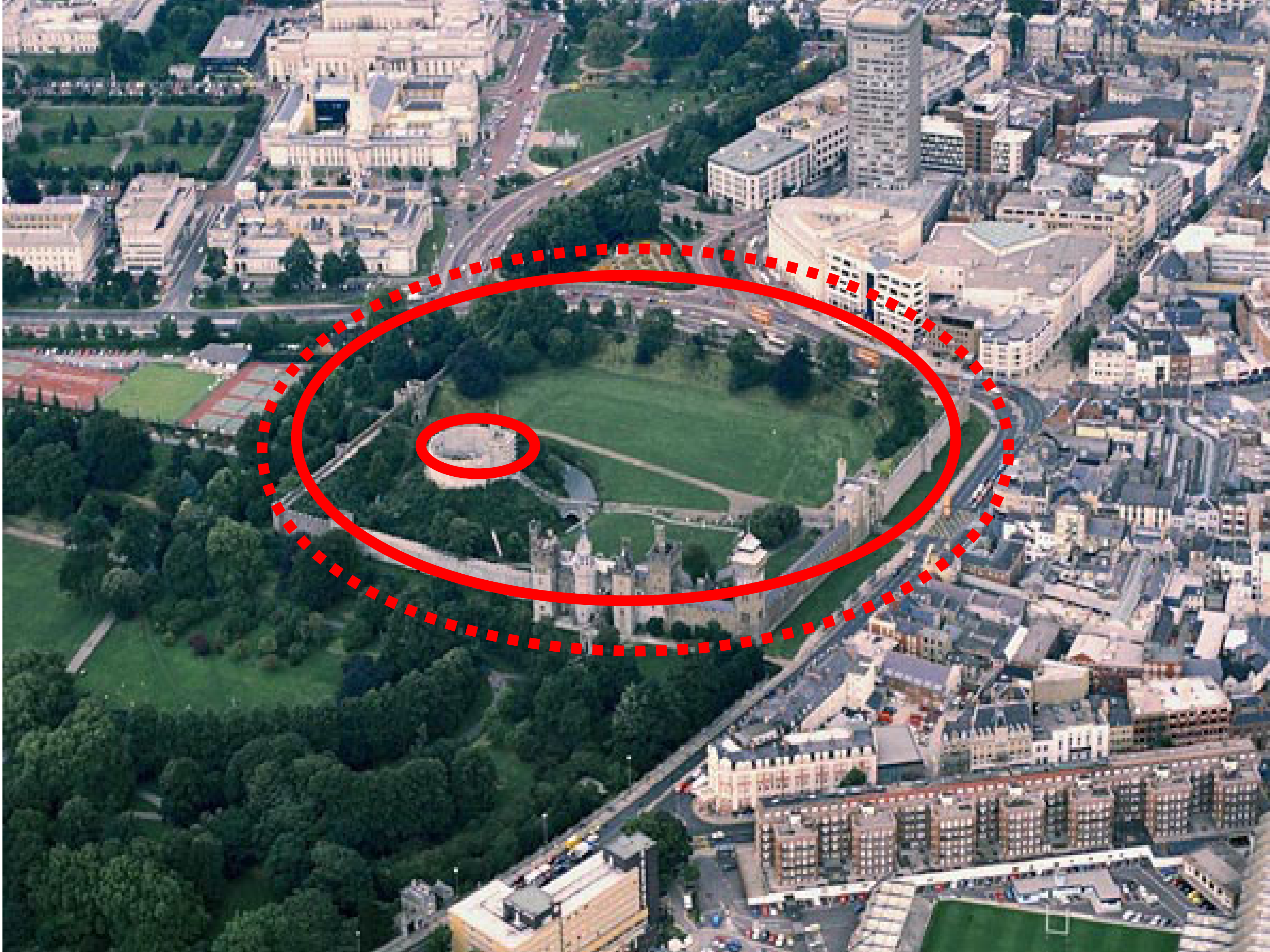
<byurcik@ncsa.uiuc.edu>

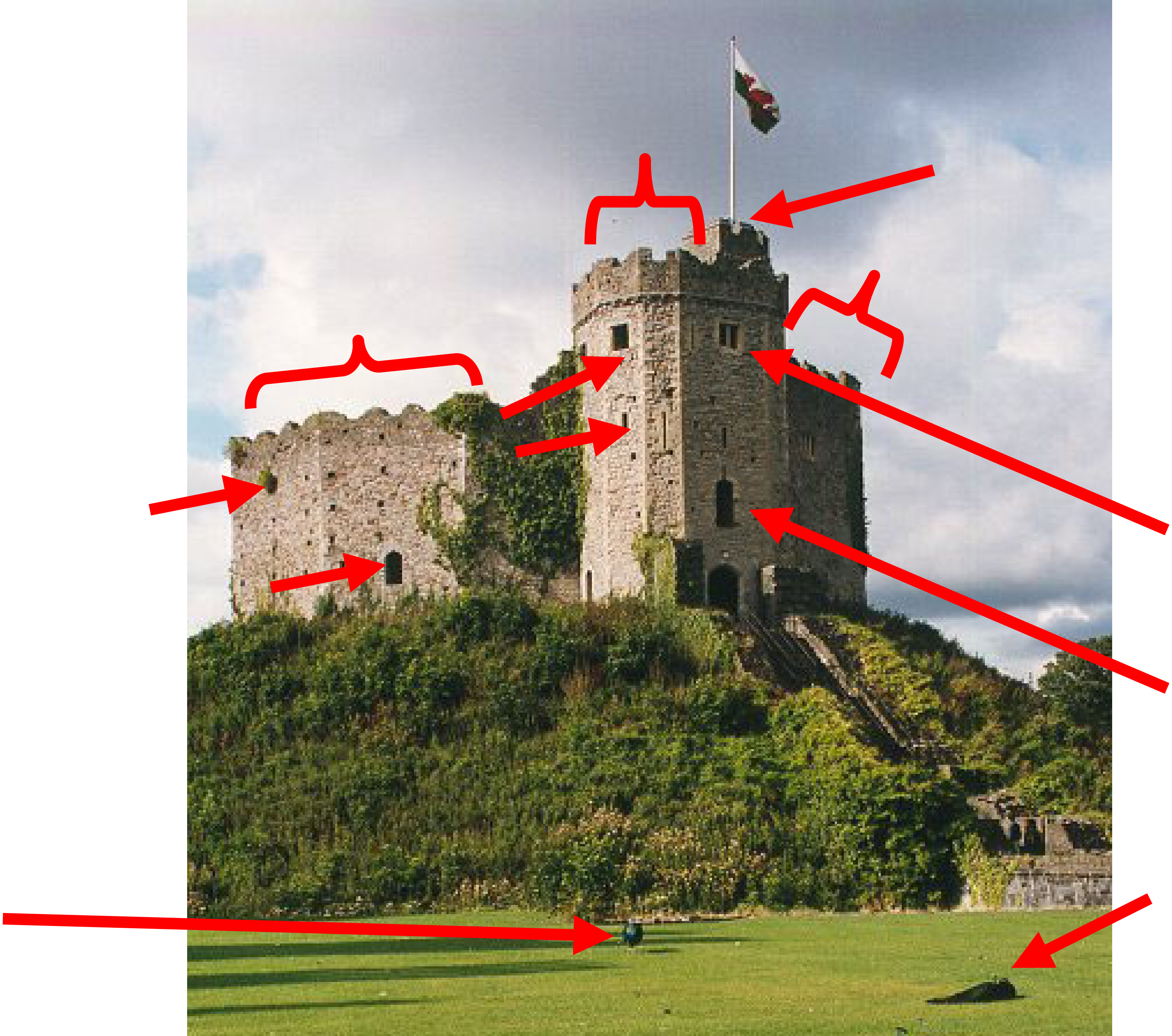
National Center for Supercomputing Applications (NCSA)
University of Illinois at Urbana-Champaign

FIRST'06 Baltimore Maryland USA

- Motivation
- Network Visualization for Security
- Our Approach: VisFlowConnect-IP
- Use Examples
- Future Work: Link-Based Clustering
- Summary

- **Motivation**
- Network Visualization for Security
- Our Approach: VisFlowConnect-IP
- Use Examples
- Future Work: Link-Based Clustering
- Summary



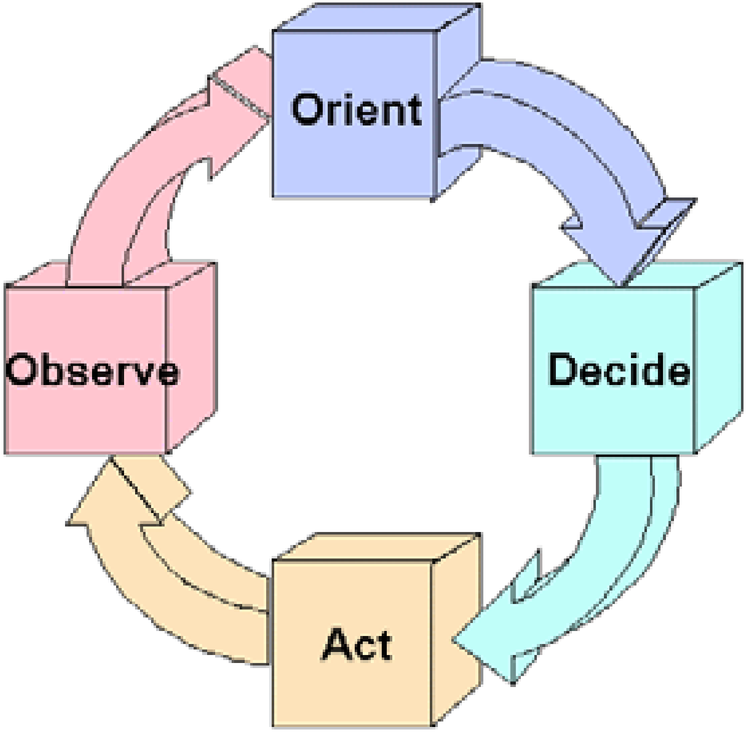


More Lessons Learned from Castles

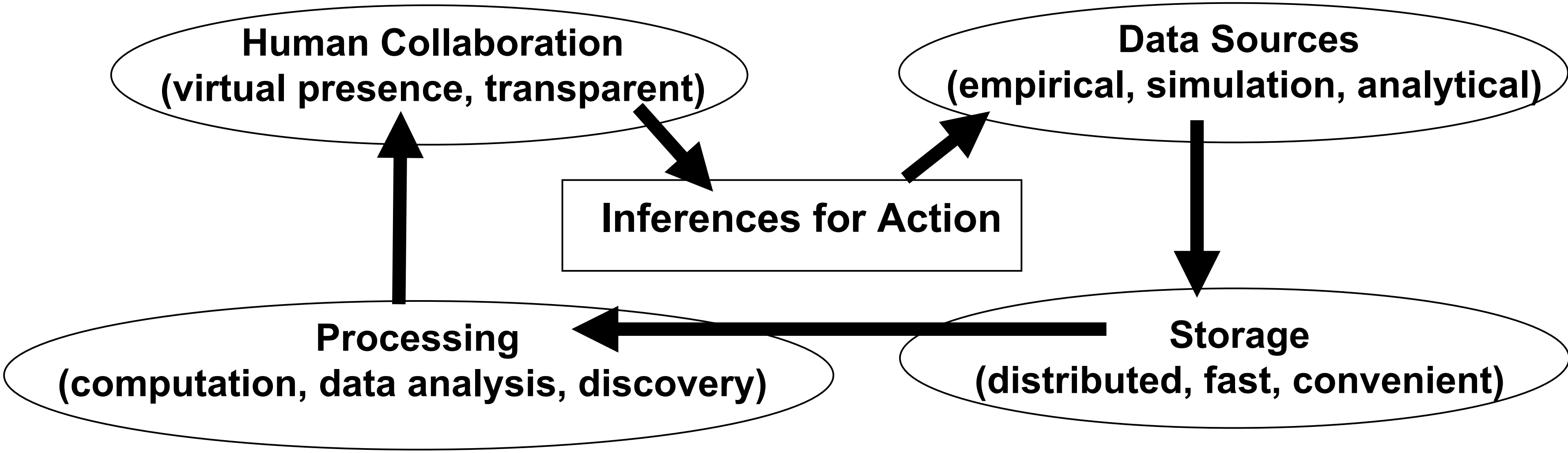
- **Even medieval castles have monitoring systems for their innermost keeps**
- **Internet security should be designed like a castle, with multiple layers of defenses for an attacker to avoid detection**
  - Reduces the space of actions that an attacker can take and remain undetected
  - Components of a security monitoring framework can monitor each other
- **Have clear observation points**
  - **Internet analogy are data source and process**

Fort McHenry

OODA Loop



OODA Loop for Internet Security



Human Collaboration
(virtual presence, transparent)

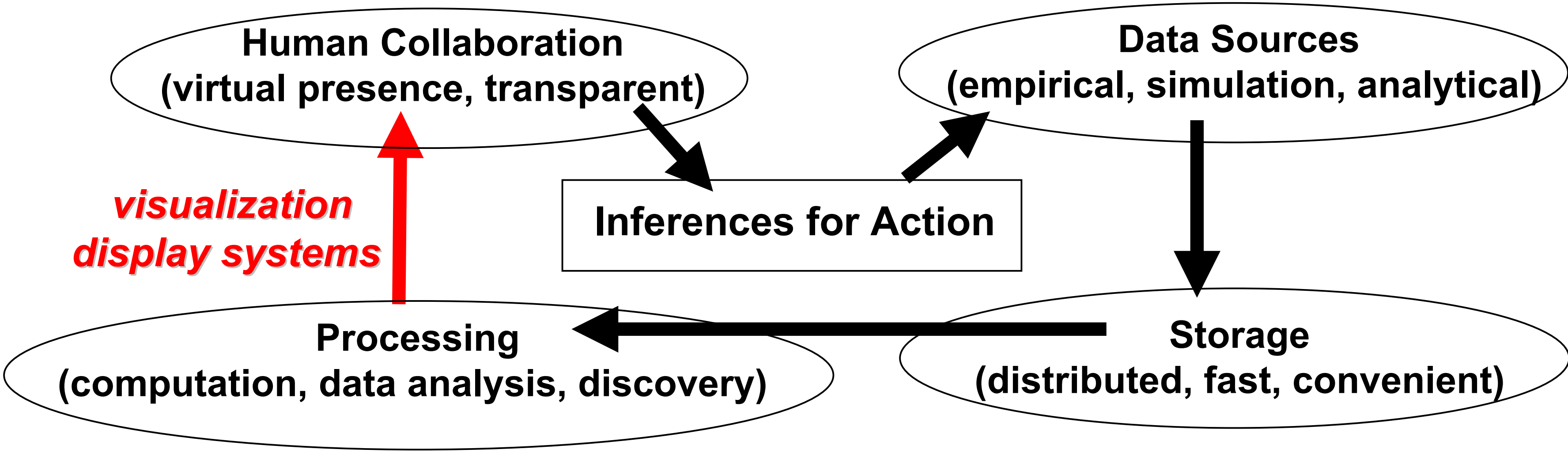
Data Sources
(empirical, simulation, analytical)

Inferences for Action

Processing
(computation, data analysis, discovery)

Storage
(distributed, fast, convenient)

Visualization in OODA Loop



Human Collaboration
(virtual presence, transparent)

Data Sources
(empirical, simulation, analytical)

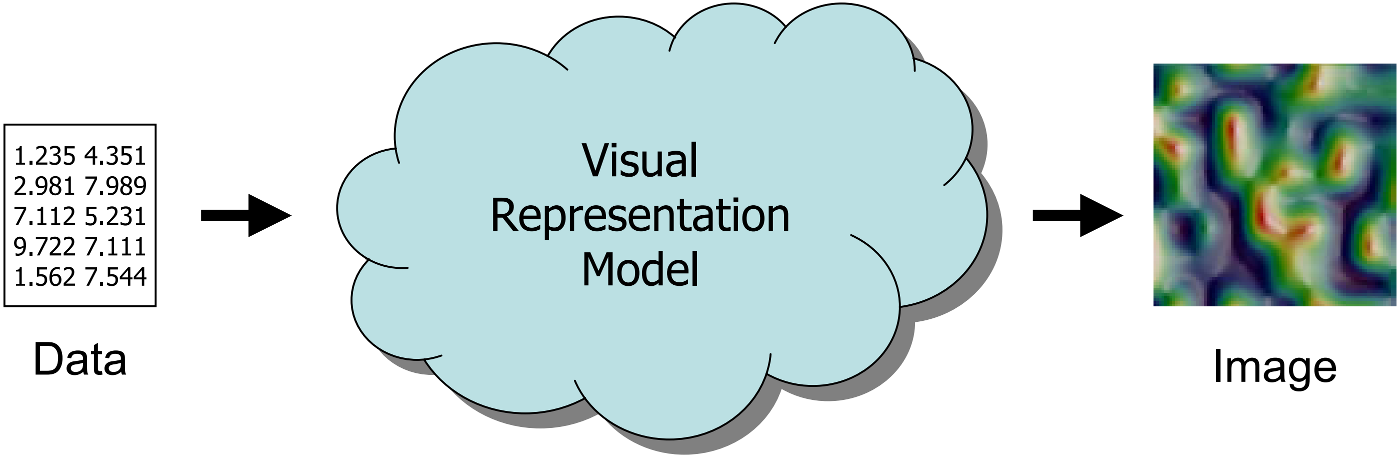
Inferences for Action

visualization display systems

Processing
(computation, data analysis, discovery)

Storage
(distributed, fast, convenient)

What is Visualization?



1.235	4.351
2.981	7.989
7.112	5.231
9.722	7.111
1.562	7.544

Data

Visual Representation Model

Image

Visualization Can Help

Empirical Data:

Visual vs Numerical (Visual Wins!)*

Visual vs Auditory (Visual Wins)*

Visual vs Tactile (Visual Wins)*

Visual Spatial vs Visual Color (Visual Spatial Wins!)*

[Chris Wickens, National Academy of Sciences Workshop on Visualizing Uncertainty, March 3, 2005]

Visualization Can Help

Empirical Data:

Visual vs Numerical (Visual Wins!)*

Visual vs Auditory (Visual Wins)*

Visual vs Tactile (Visual Wins)*

Visual Spatial vs Visual Color (Visual Spatial Wins!)*

[Chris Wickens, National Academy of Sciences Workshop on Visualizing Uncertainty, March 3, 2005]

How?

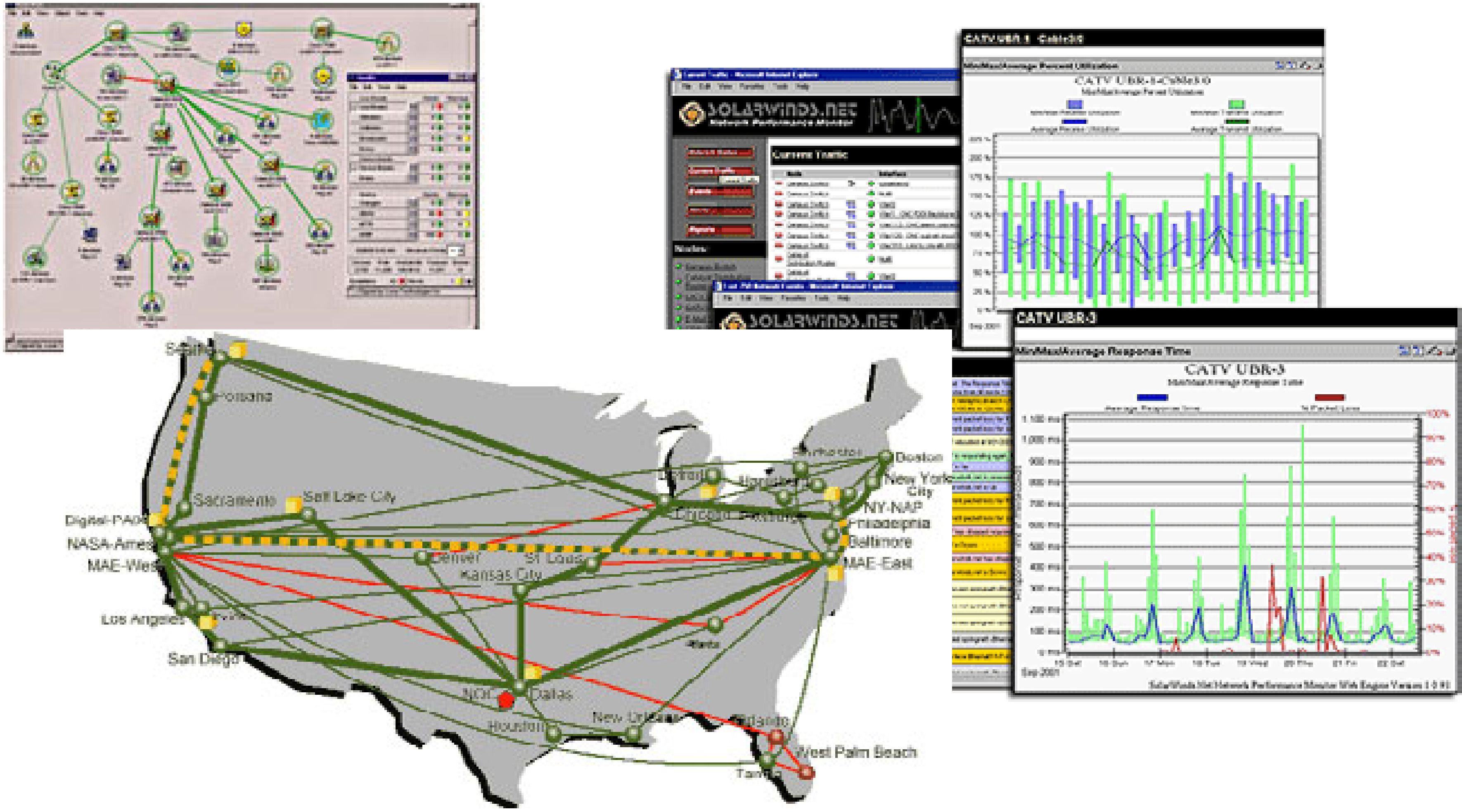
1) See Previously Obscured Things

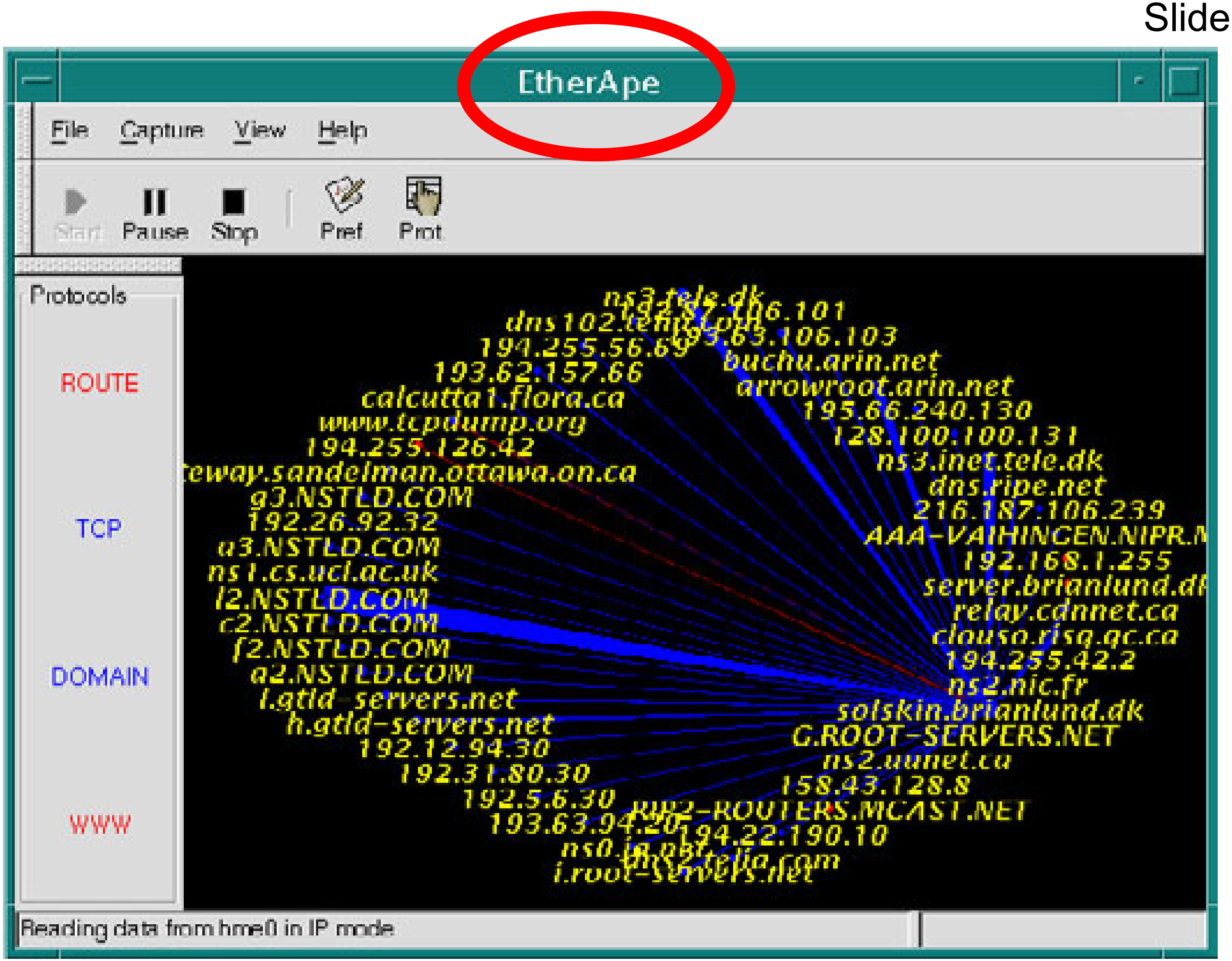
2) See New Things Faster (I never saw that before)

3) Share Insights (Do you see what I mean?)

- Motivation
- **Network Visualization for Security**
- Our Approach: VisFlowConnect-IP
- Use Examples
- Future Work: Link-Based Clustering
- Summary

Current Net Vis Security Ops Tools

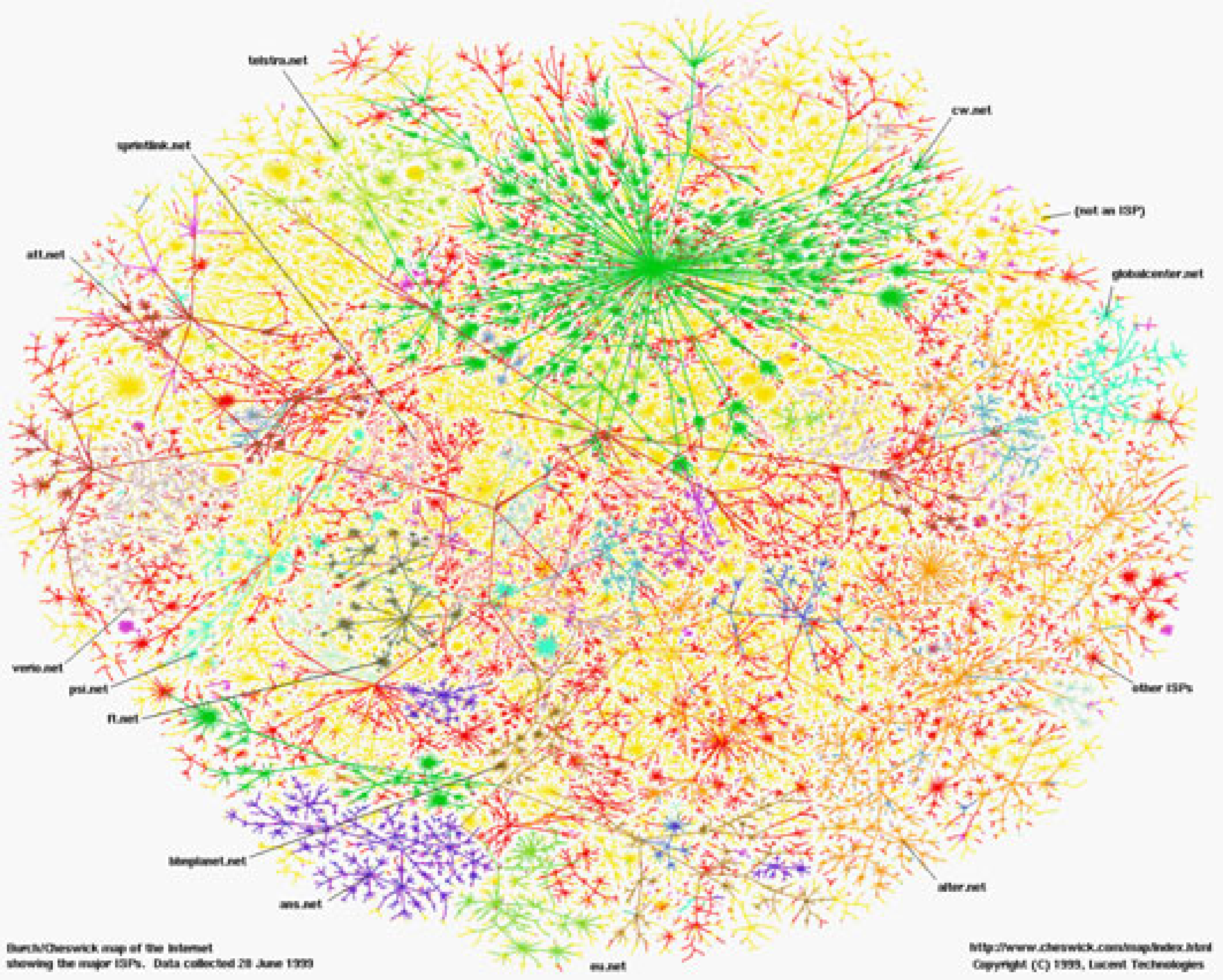




Reading data from hme0 in IP mode

Etherape by Juan Toledo can be found at
http://etherape.sourceforge.net/
screenshot: http://www.solaris4you.dk/sniffersSS.html

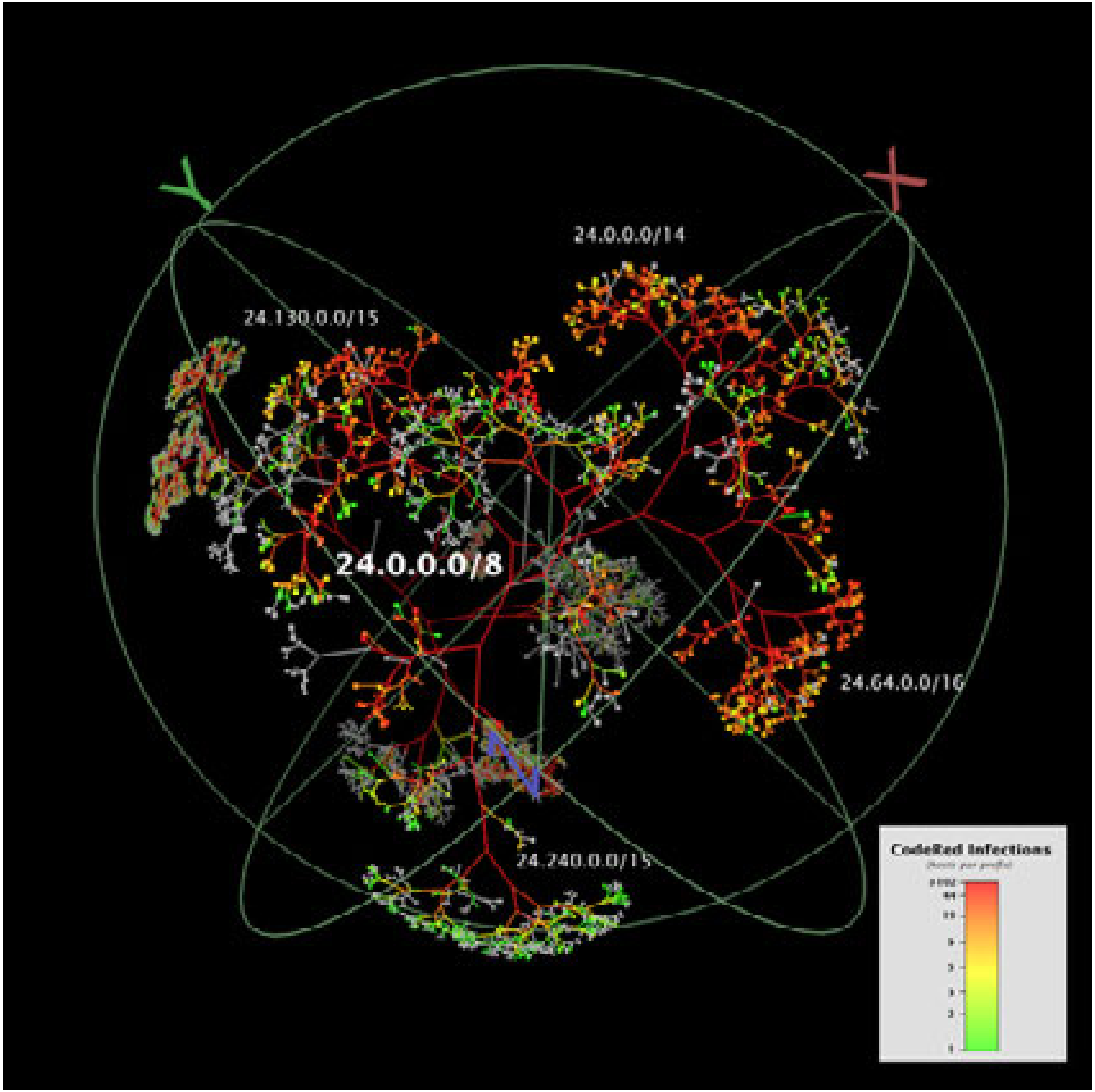
Lumeta's Peacock Diagrams



Burch/Cheswick map of the Internet
showing the major ISPs. Data collected 29 June 1999

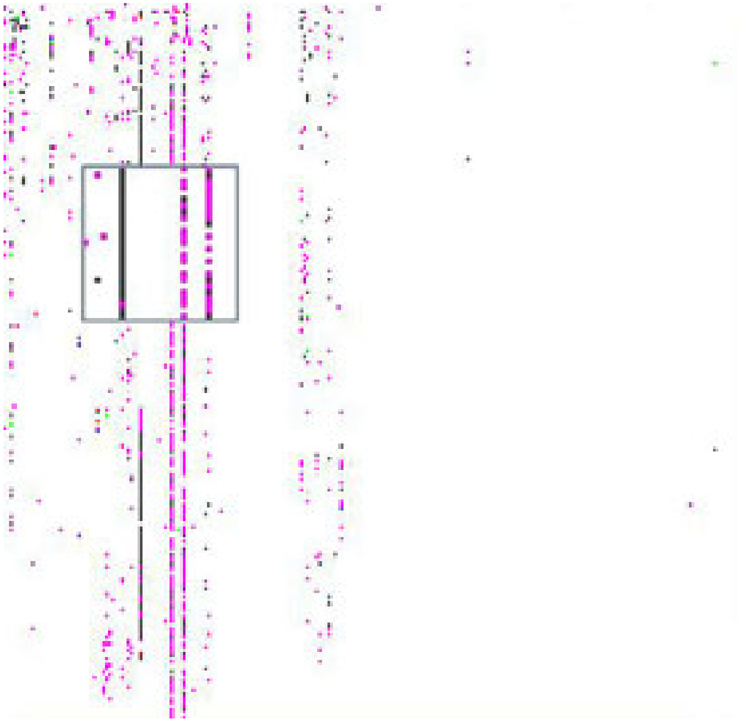
http://www.cheswick.com/map/index.html
Copyright (C) 1999, Lucent Technologies

Caida’s Walrus



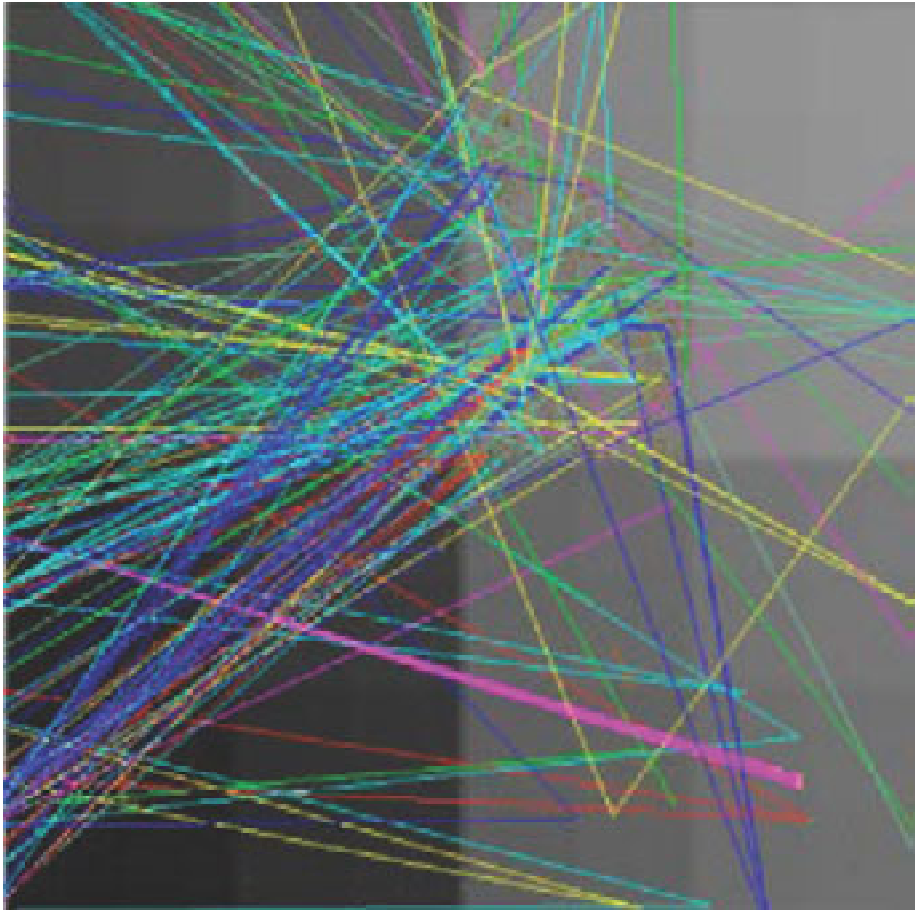
Research: Network Viz for Security

- **Host-based approaches**



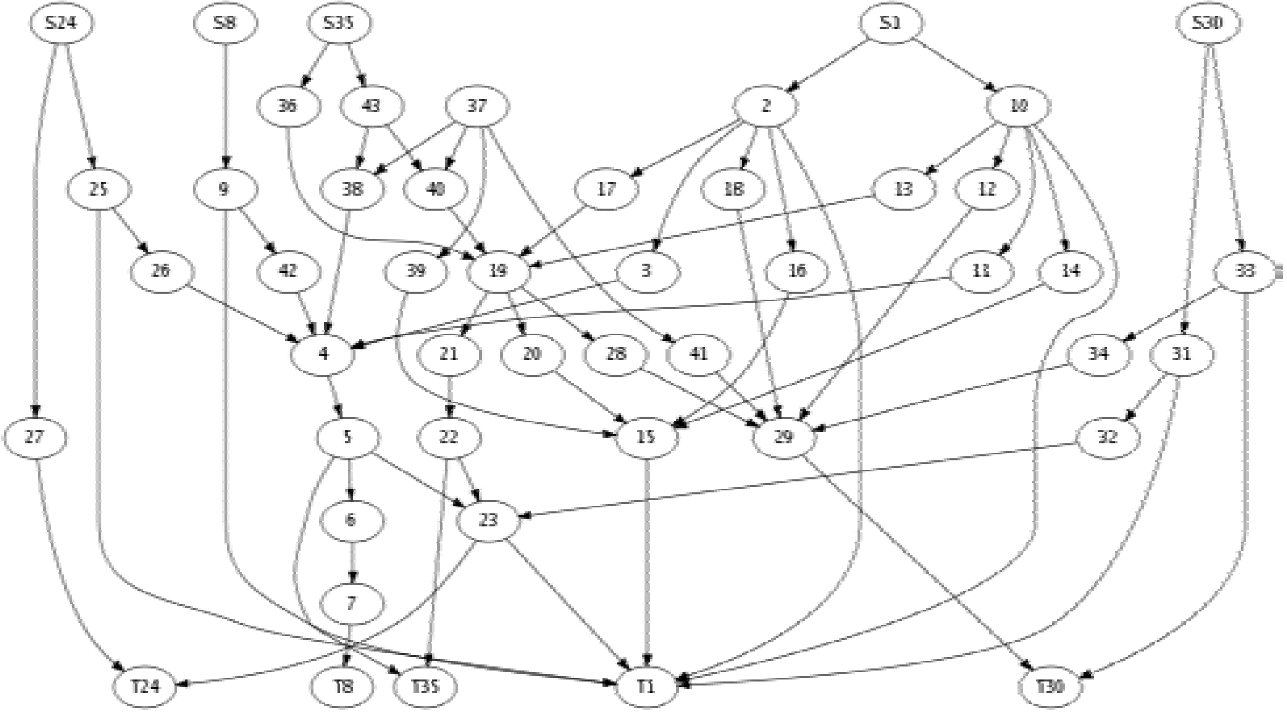
(NVisionIP- NCSA)

- **Link-based approaches**

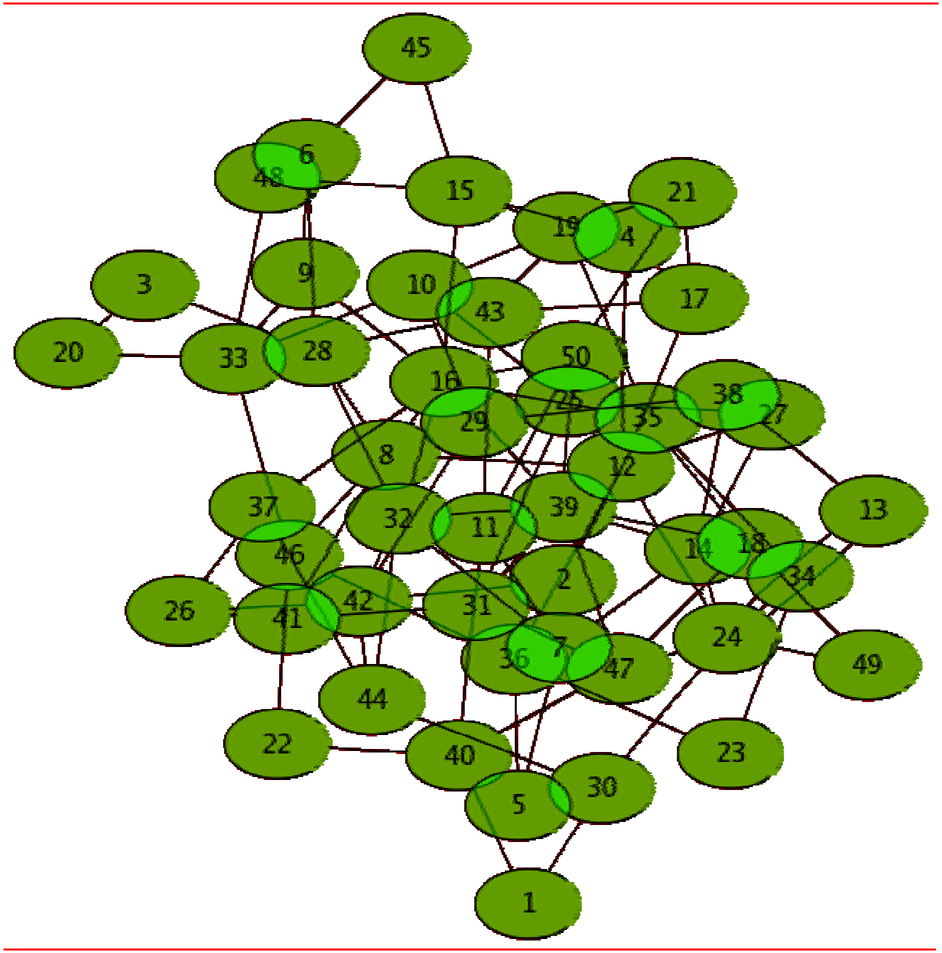


(Elisha-Teoh et al)

AT&T's Graphiz



Graphviz again

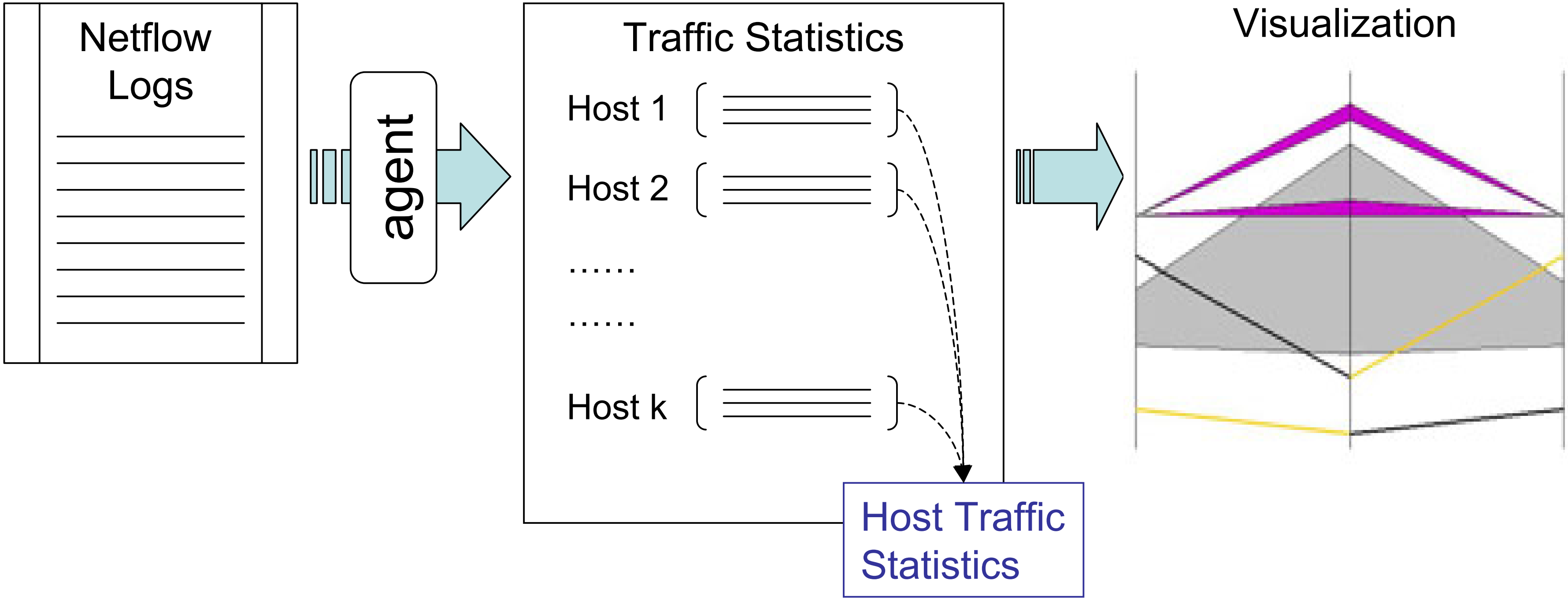


- Motivation
- Network Visualization for Security
- **Our Approach: VisFlowConnect-IP**
- Use Examples
- Future Work: Link-Based Clustering
- Summary

Our Design Goals

- Traffic dynamics over time
- Filtering
- Scalability
- Expose hidden structures & patterns for further investigation

System Architecture



Netflow Logs

agent

Traffic Statistics

Host 1

Host 2

……

……

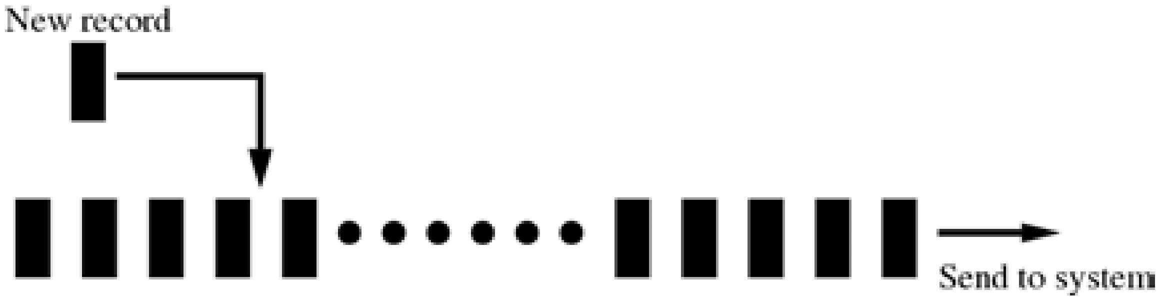
Host k

Host Traffic Statistics

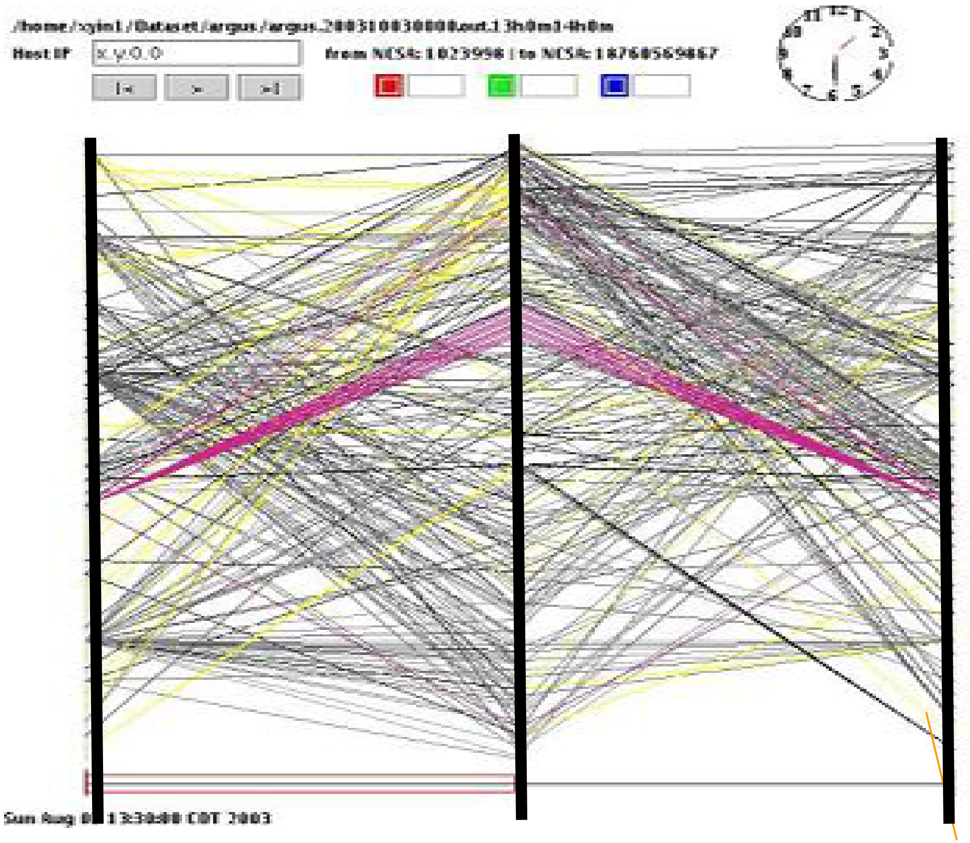
Visualization

Reading Netflow Logs

- An agent reads records log (or streaming)
  - send record to VisFlowConnect-IP when requested
- Reorder NetFlow records with record buffer
  - records are not strictly sorted by time stamps
  - use a record buffer



VisFlowConnect-IP



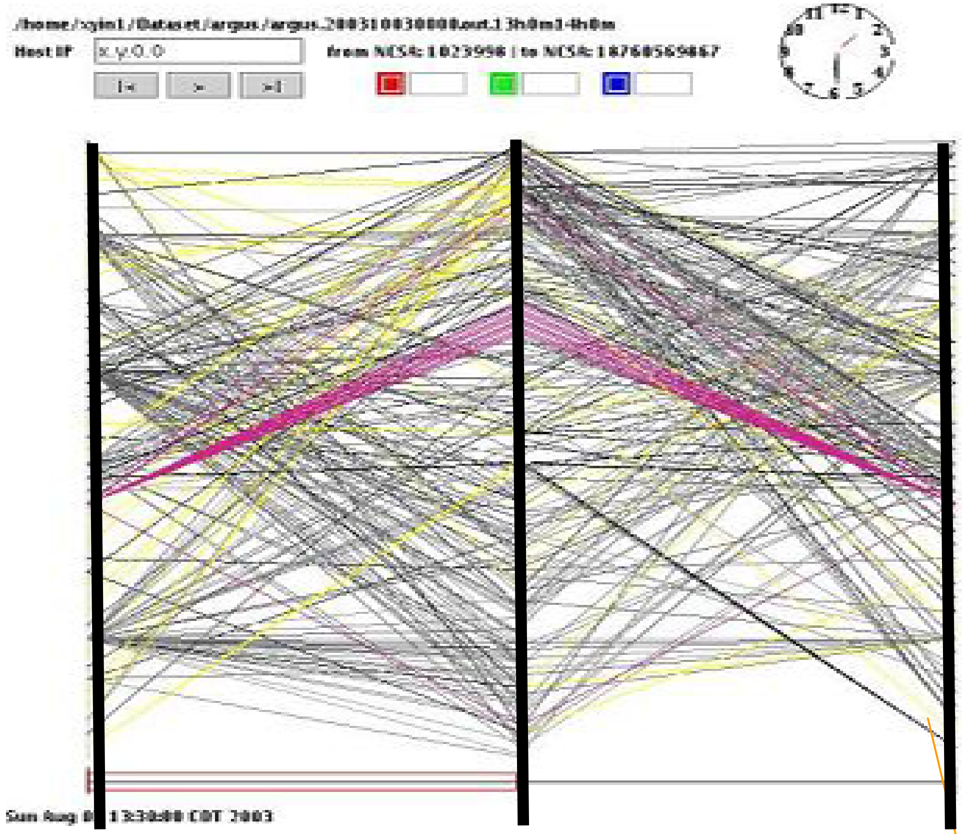
/home/xyin1/Dataset/argus/argus.2003100830000.out.13h0m14h0m

Host IP x.y.0.0

from NCSA:1023998 | to NCSA:18768569887

Sun Aug 0[illegible] 13:30:00 CDT 2003

VisFlowConnect-IP Main View



/home/xyin1/Dataset/argus/argus.20031008000.out.13h0m1-4h0m

Host IP x.y.0.0

from NCSA:1023998 | to NCSA:18760569867

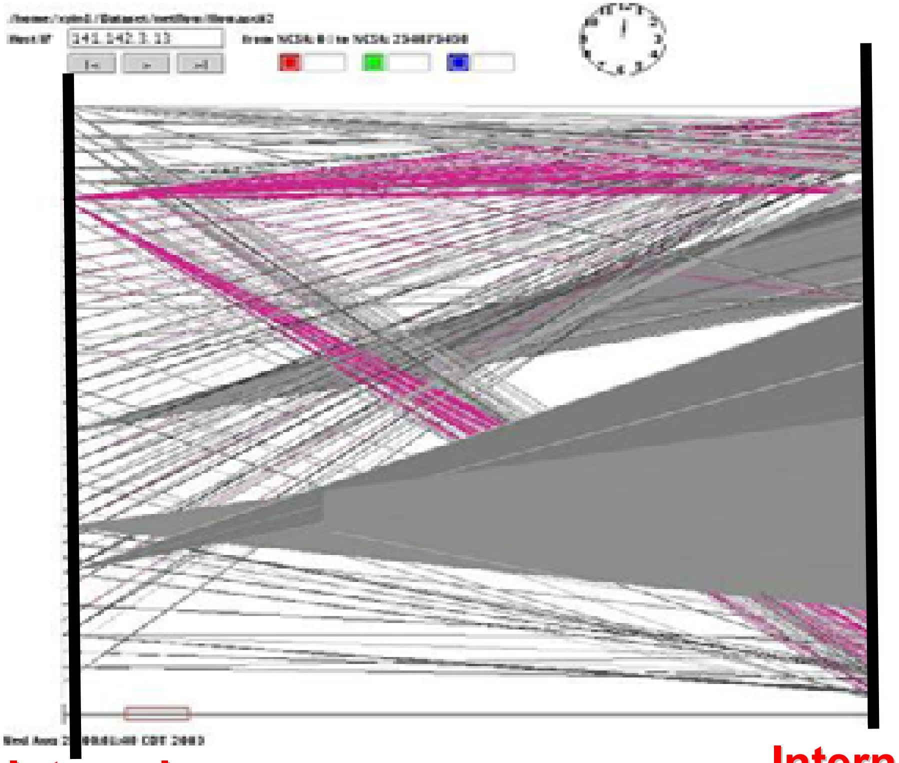
Sun Aug 0 13:30:00 CDT 2003

outside domains axis

inside hosts axis

outside domains axis

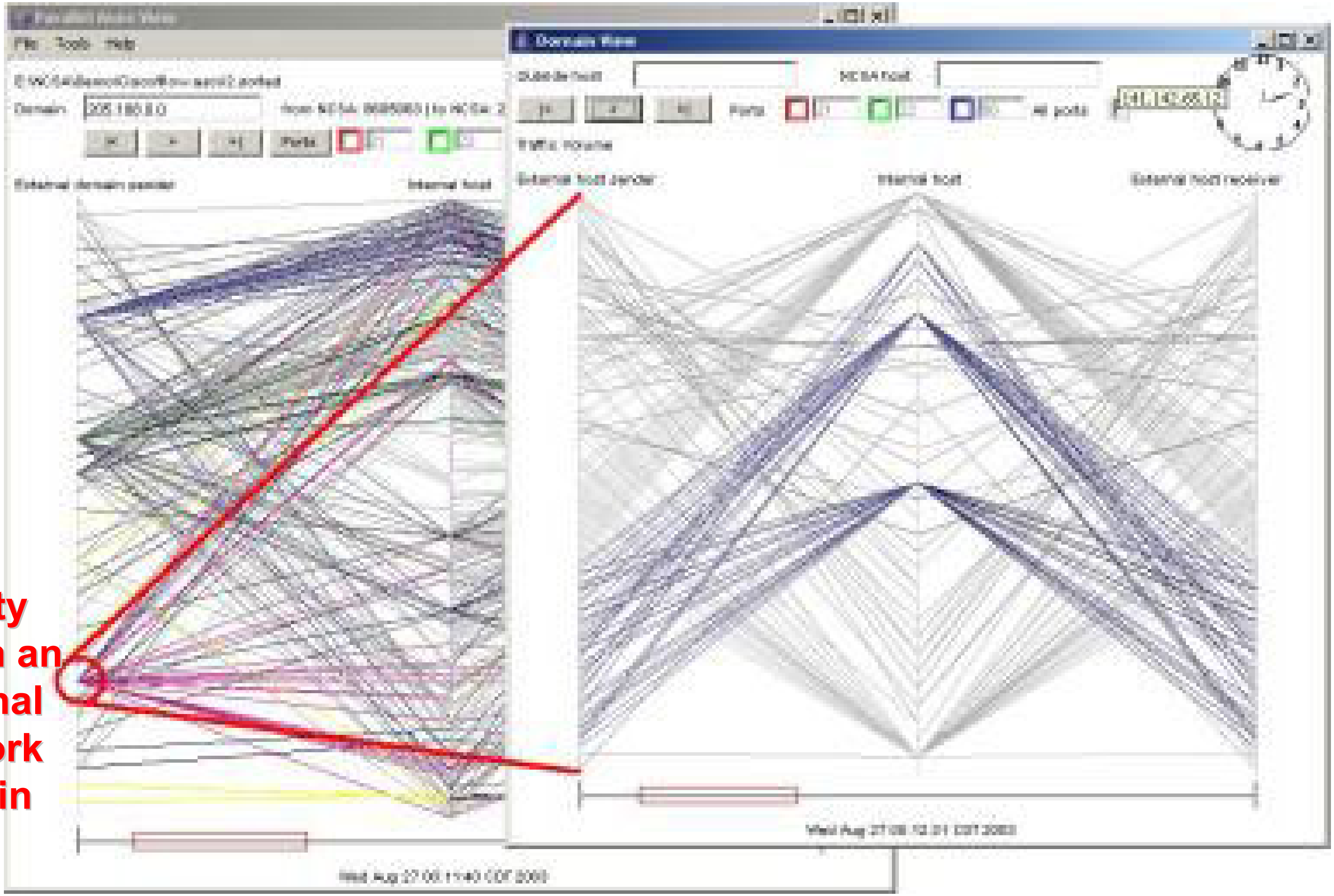
VisFlowConnect-IP Internal View



Internal
network
sources

Internal
network
receivers

VisFlowConnect-IP Domain View



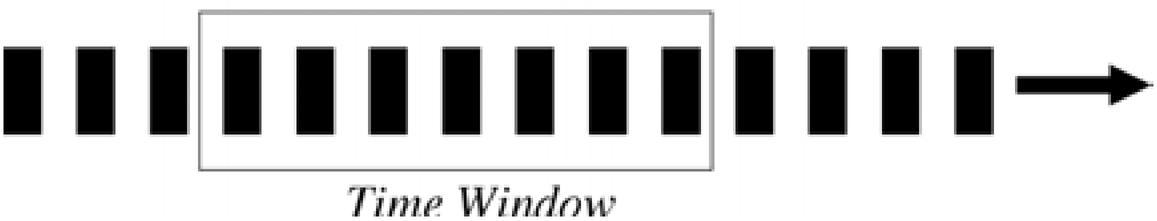
see activity within an external network domain

Creating Dynamic Animation

- Visualizing traffic statistics with time
  - update visualization after each time unit
- How to arrange domains/hosts?
  - 100s of domains/hosts; added/removed in time
  - fairly stable positioning
- Solution: sort by IP
  - domain/hosts move up or down

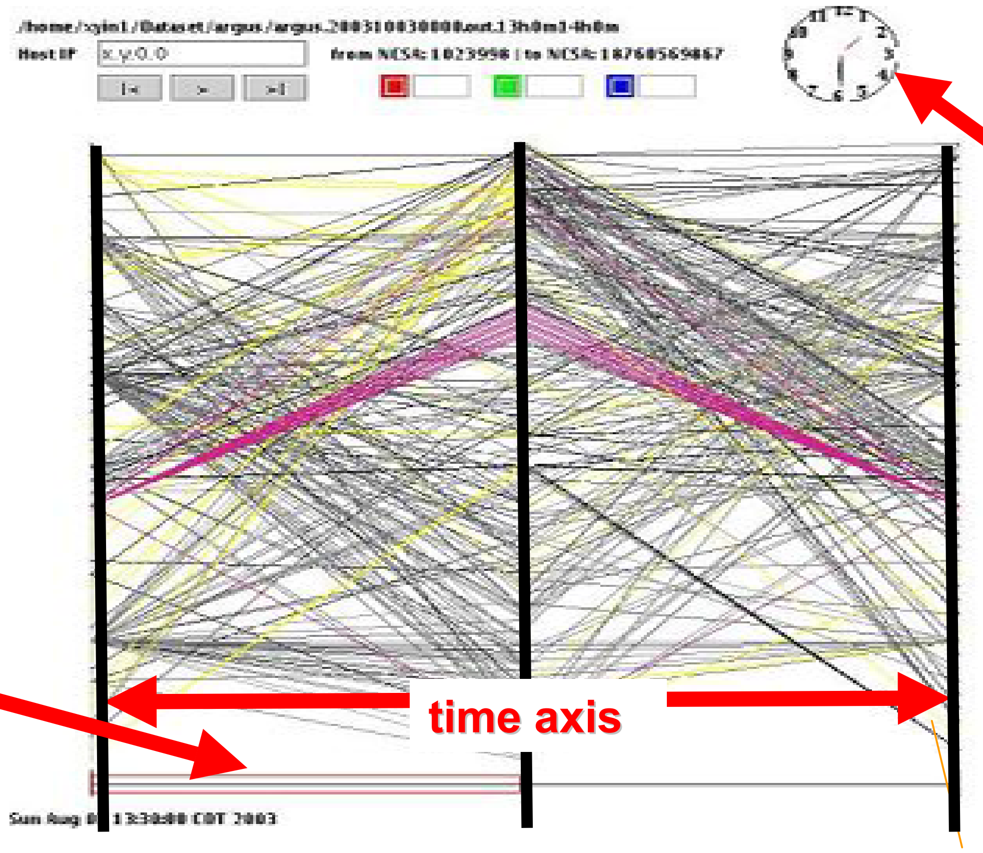
Time Window

- User is usually interested in most recent traffic (e.g., in last minute or last hour)
- VisFlowConnect-IP only visualizes traffic in a user adjustable time window



  - Update traffic statistics when
    - A record comes into time window
    - A record goes out of time window

Time Dynamics



/home/xyin1/Dataset/argus/argus.2003100830000.out.13h0m14h0m

Host IP x.y.0.0 from NCSA:1023998 | to NCSA:18768569867

Sun Aug 0 13:30:00 CDT 2003

analog clock

time window

time axis

timestamp

Filtering/Highlighting Capability

- Approach
  - – Filter out “good” traffic
    - User specifies a list of filters:

+: (SrcIP=141.142.0.0−141.142.255.255), (SrcPort=1−1000)

//keep all records from domain 141.142.x.x, from port 1 – 1000

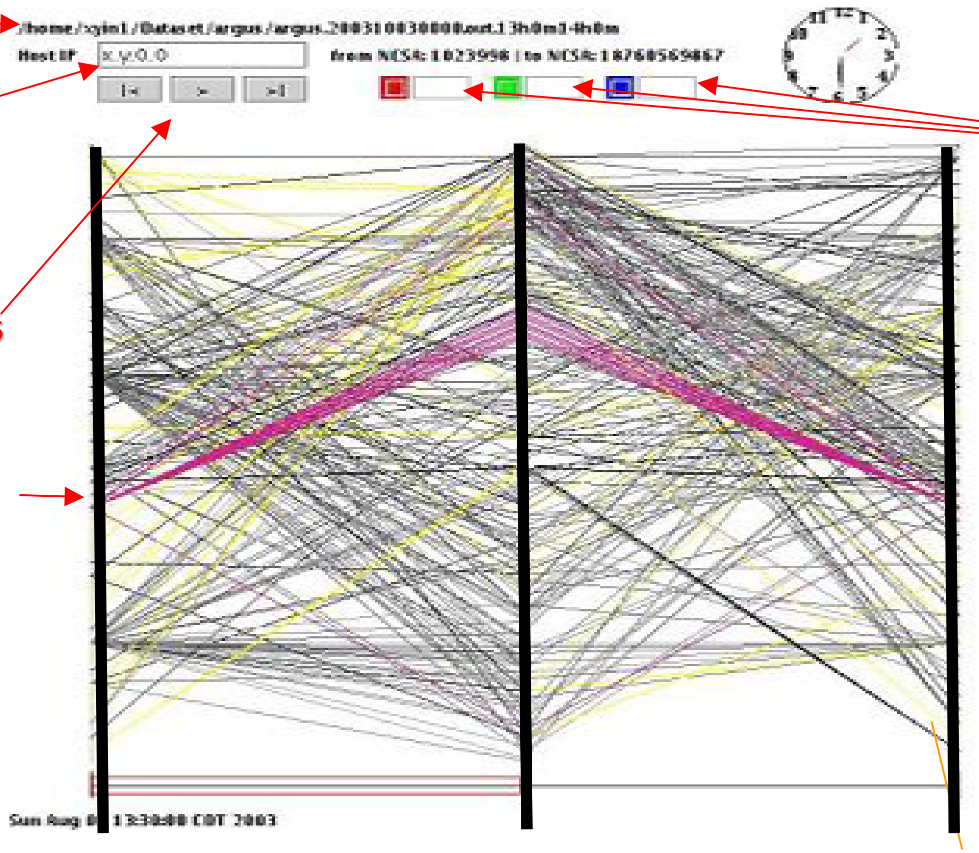
−: (SrcPort=80)

−: (DstPort=80)

//discard records of http traffic

- – Highlight “traffic of interest”
  - traffic colored by port

Highlighting “Traffic of Interest”


/home/xyin1/Dataset/argus/argus.2003100300008out.13h0m14h0m

Host IP x.y.0.0

from NCSA: 1023998 | to NCSA: 18760569867

Sun Aug 0 13:30:00 CDT 2003

File I/O

Net Domain

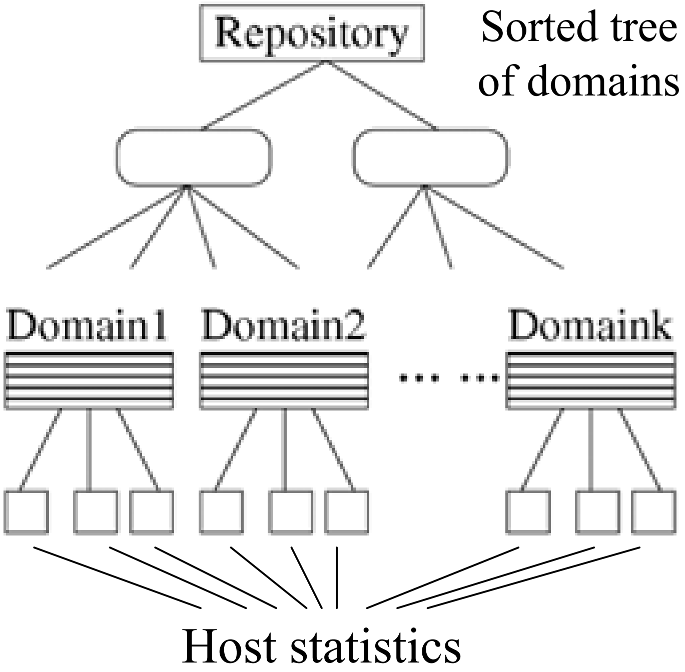
VCR controls

highlighted flow

highlighted ports

Storing Traffic Statistics

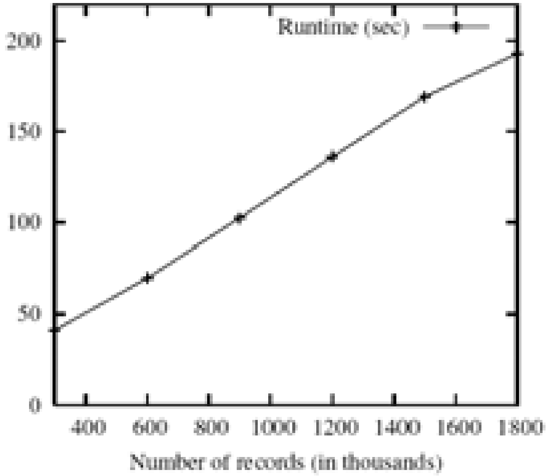
- Store traffic statistics involving each domain by a sorted tree
  - only necessary information for visualization is stored
  - statistics for every domain or host can be updated efficiently

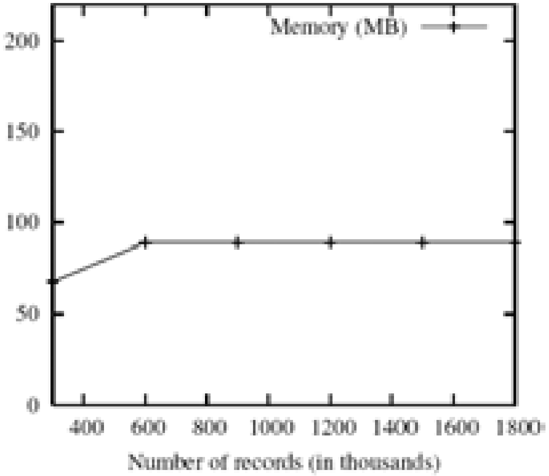


Scalability Experiments

Runtime & Memory

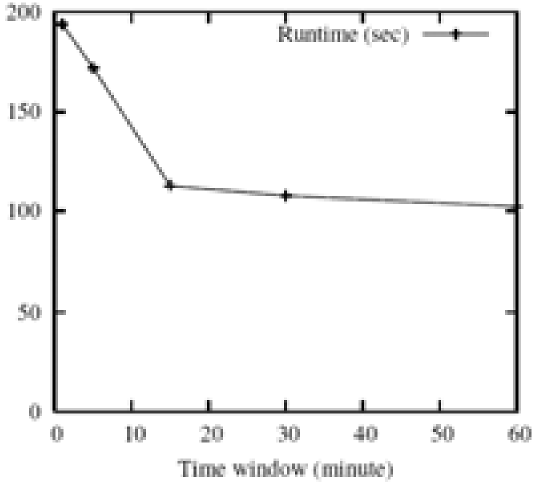
wrt records

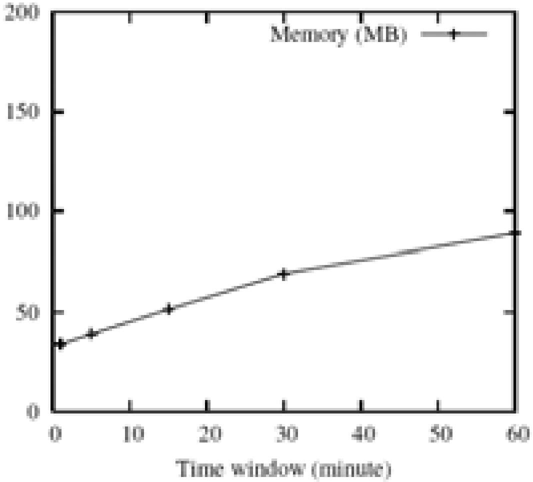




Runtime & Memory

wrt time window size

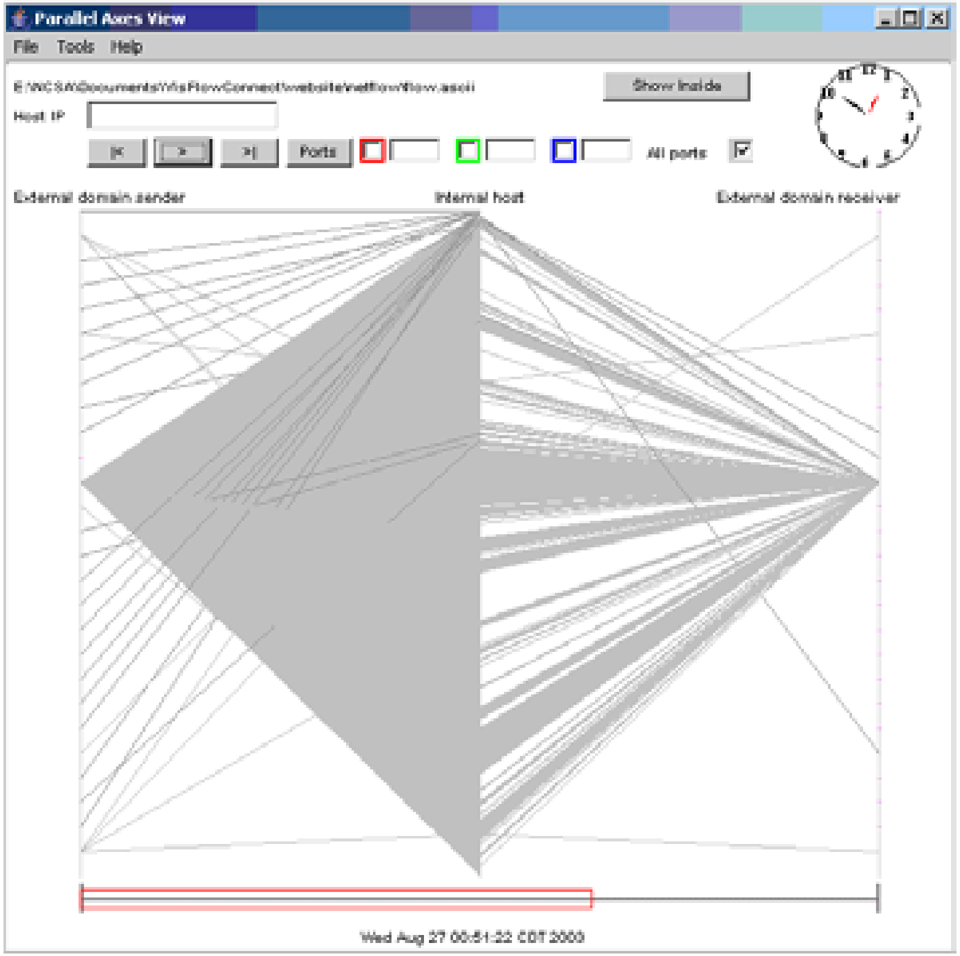




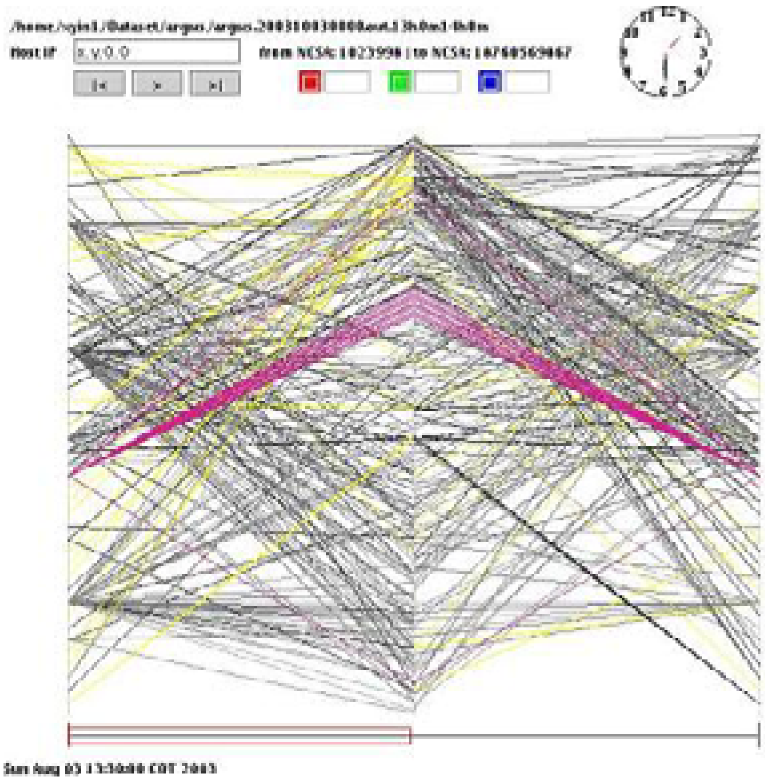
- Motivation
- Network Visualization for Security
- Our Approach: VisFlowConnect-IP
- **Use Examples**
- Future Work: Link-Based Clustering
- Summary

Example 1: MS Blaster

- MS Blaster virus causes machines to send out 92 byte pakcets to many machines



Example 2: ?



/home/vyin1/Dataset/argus/argus.2003100300Host.13h0m1-0h0m

Host IP x.y.0.0

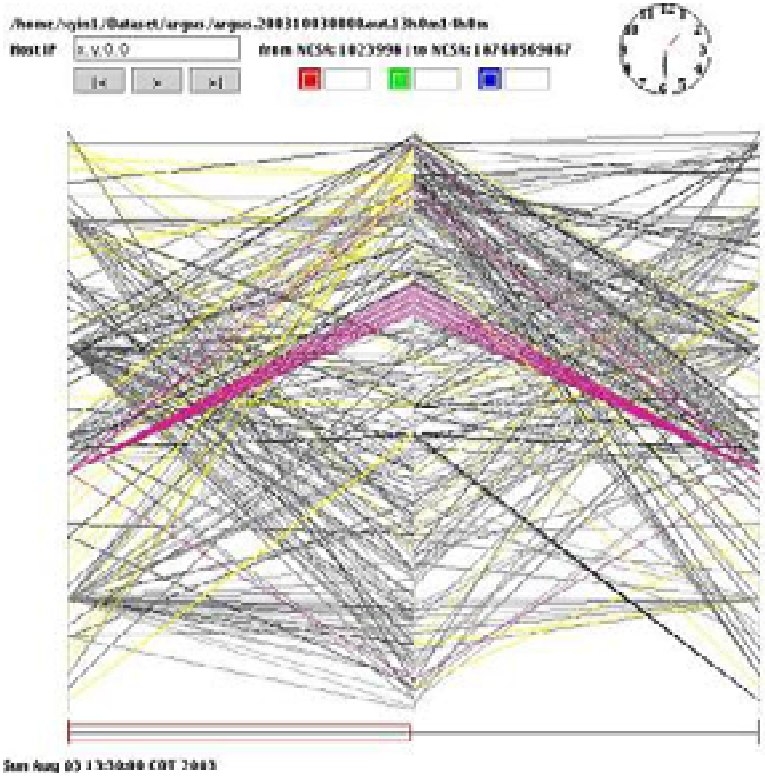
from NCSA: 1023998 | to NCSA: 16760569867

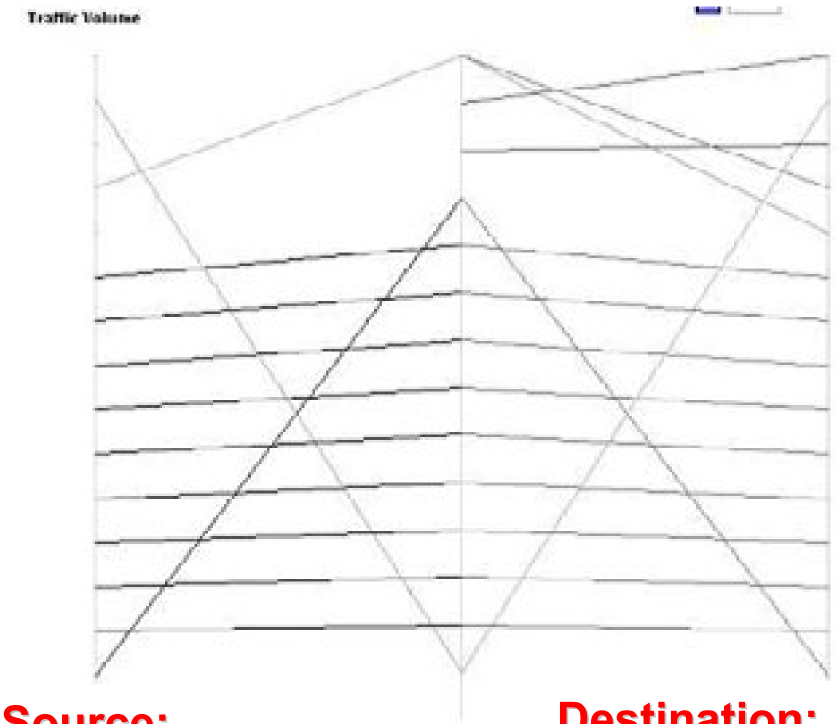
Sun Aug 03 13:30:00 CDT 2003

multiple connections to NCSA cluster from same domain

(scan?, DoS?)

Example 2: ?


/home/yin1/Dataset/argus/argus.20031003000host.13h0m1-0h0m
Host IP
x.y.0.0
from NCSA: 1023998 | to NCSA: 1074056986?
|<
>
>|
Sun Aug 03 13:30:00 CDT 2003



multiple connections to NCSA cluster from same domain

(scan?, DoS?)

Source:

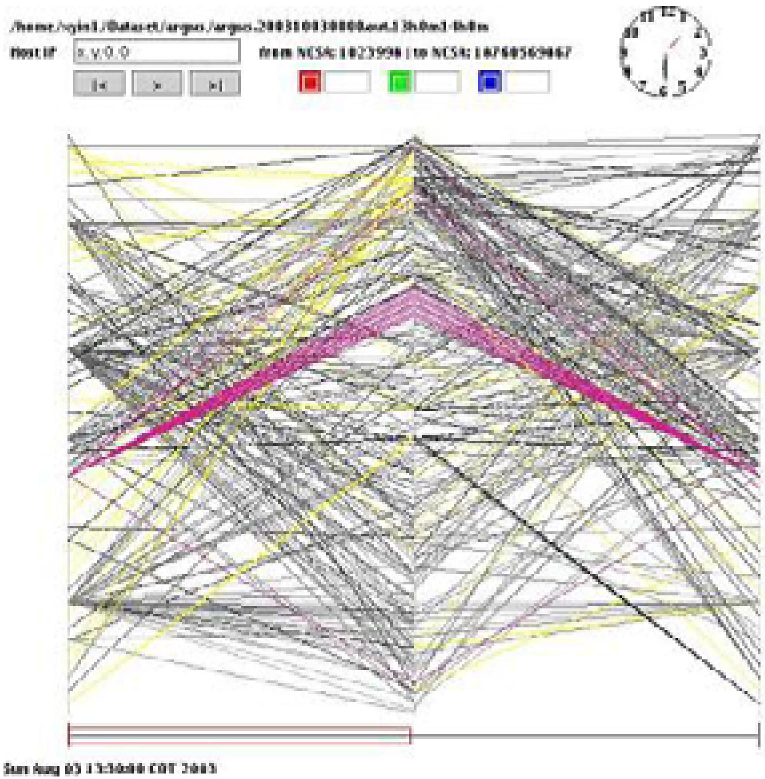
consecutive IP addresses

Destination:

consecutive IP addresses

Example 2: Grid Networking

cluster-to-cluster communications

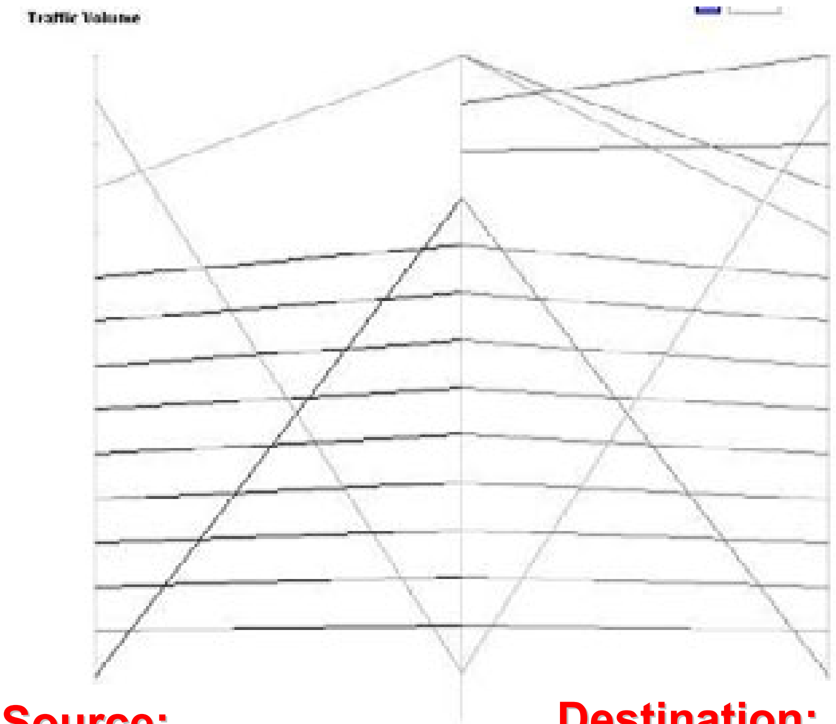


/home/cylin1/Dataset/argus/argus.20031003000[illegible]

Host IP x.y.0.0 from NCSA:1023998 | to NCSA:1676056986[illegible]

multiple connections to NCSA cluster from same domain

(scan?, DoS?)



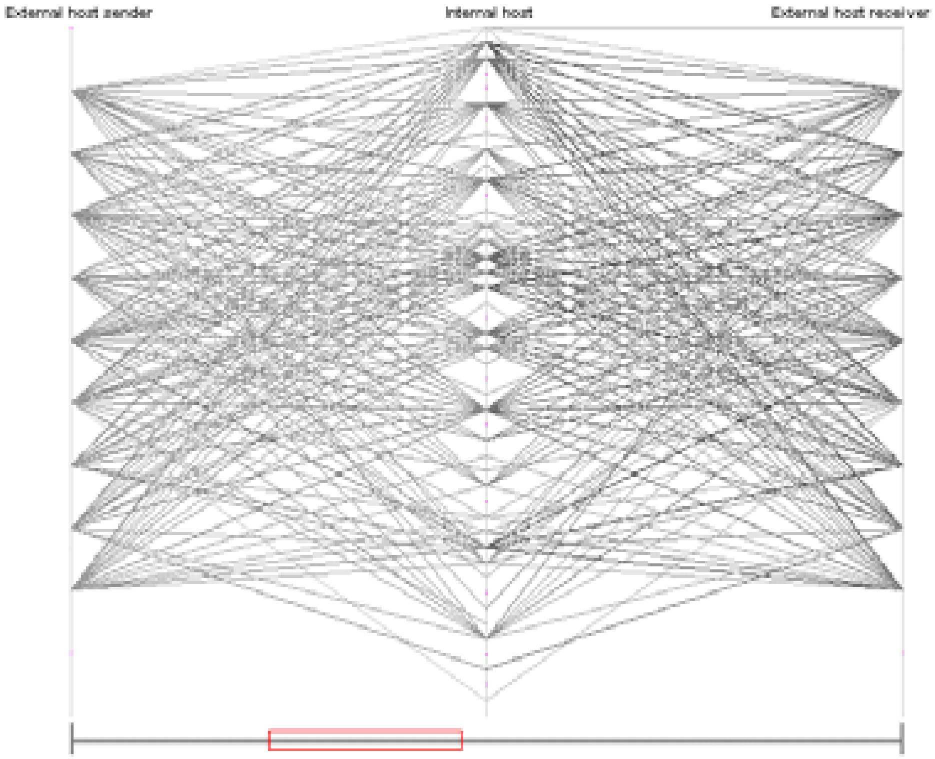
Source:

consecutive IP addresses

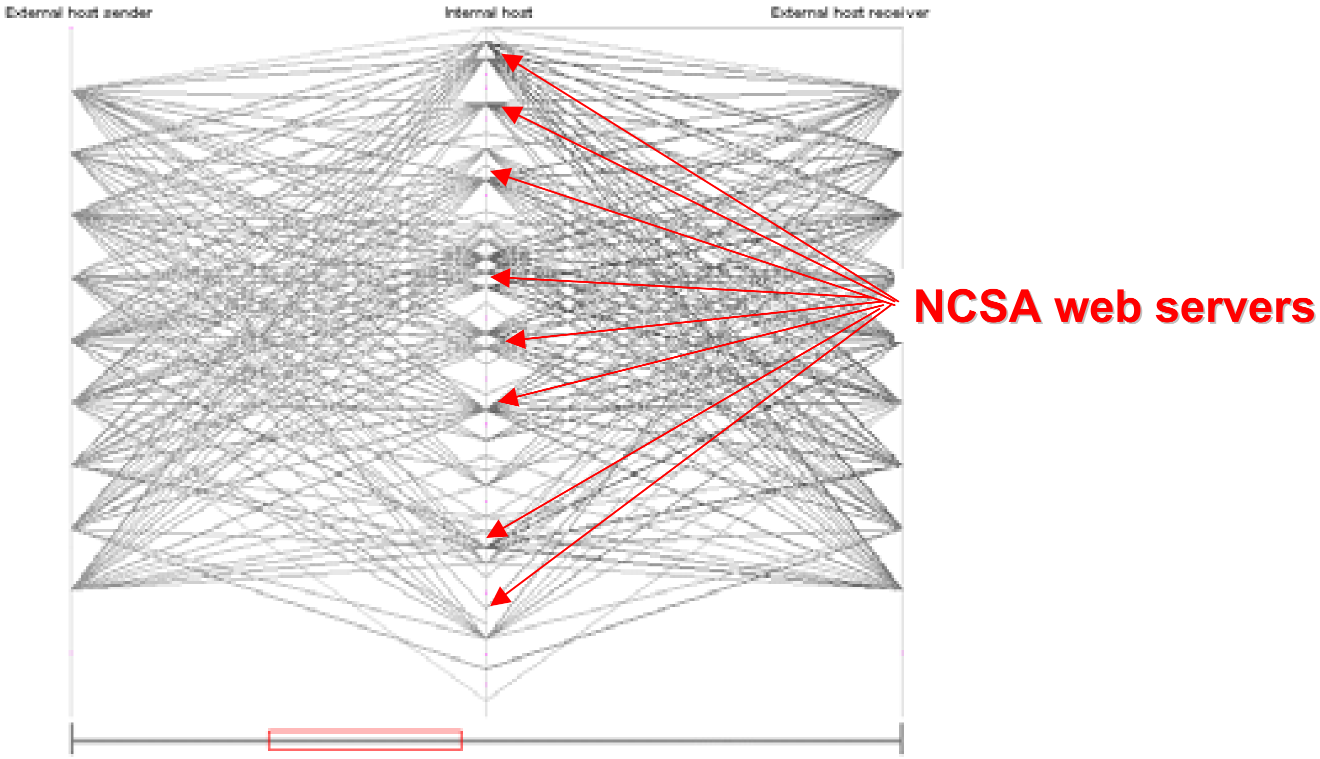
Destination:

consecutive IP addresses

Example 3: ?

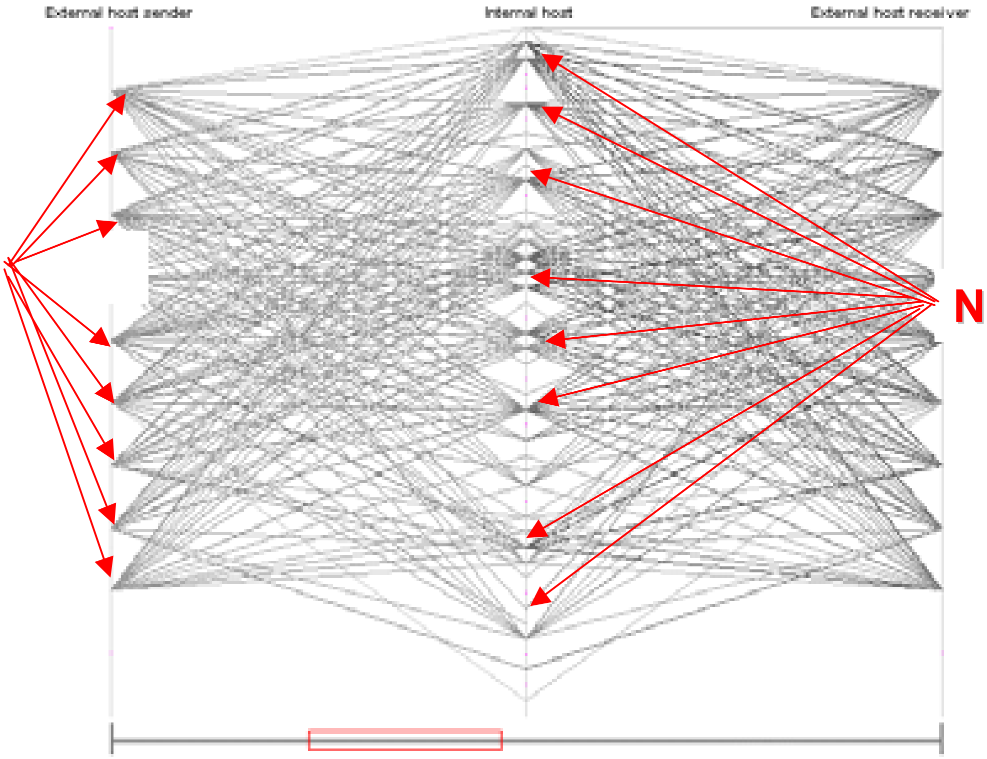

External host sender
Internal host
External host receiver

Example 3: ?


External host sender
Internal host
External host receiver
NCSA web servers

Example 3: Web Crawlers

muitiple crawlers indexing NCSA web server content



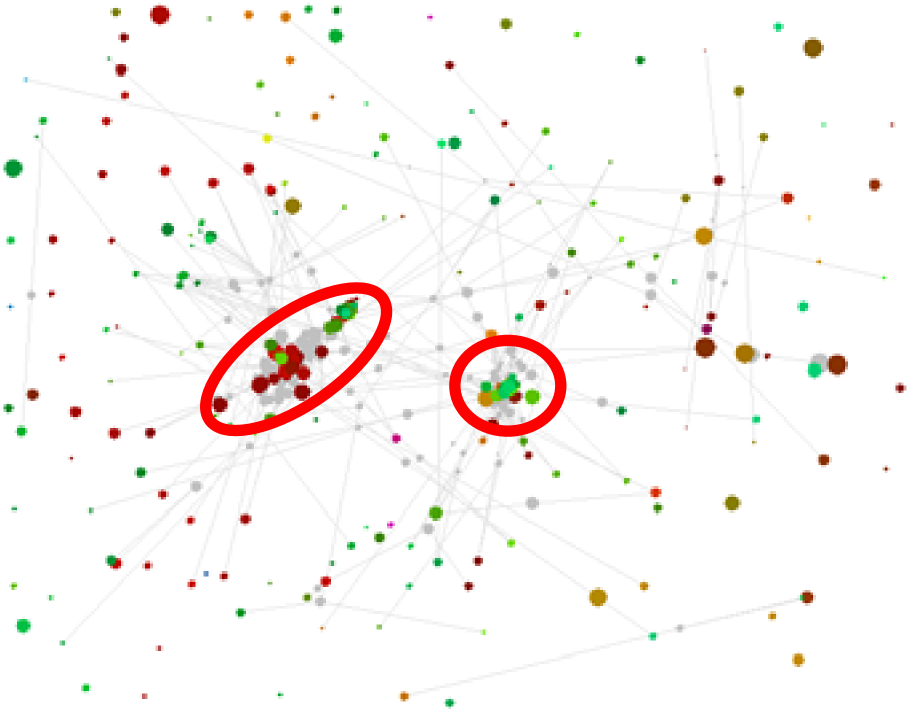
Web crawlers

NCSA web servers

- Motivation
- Network Visualization for Security
- Our Approach: VisFlowConnect-IP
- Use Examples
- **Future Work: Link-Based Clustering**
- Summary

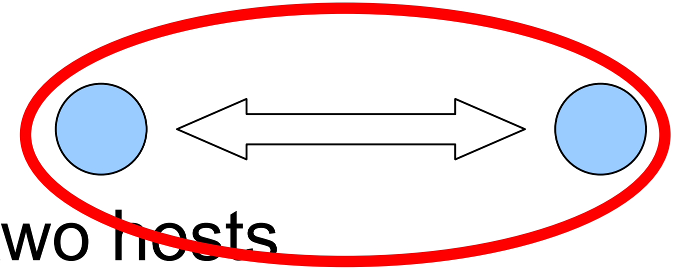
Visual Clustering of Hosts

- Visual clustering of hosts by link analysis
  - – represent each host by a point
  - – arrange hosts so related hosts are clustered

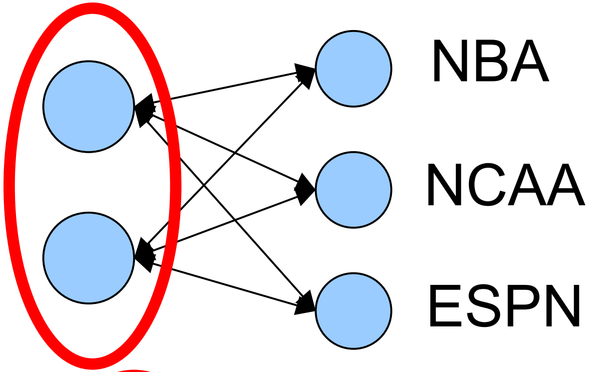


Relationships between Hosts

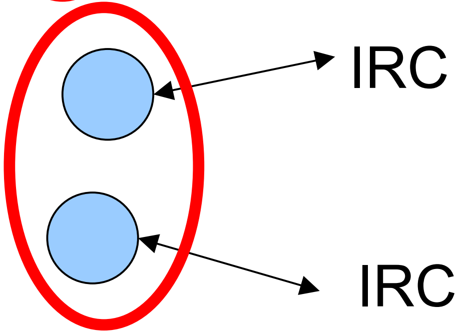
- Direct communications
  - – traffic intensity between two hosts



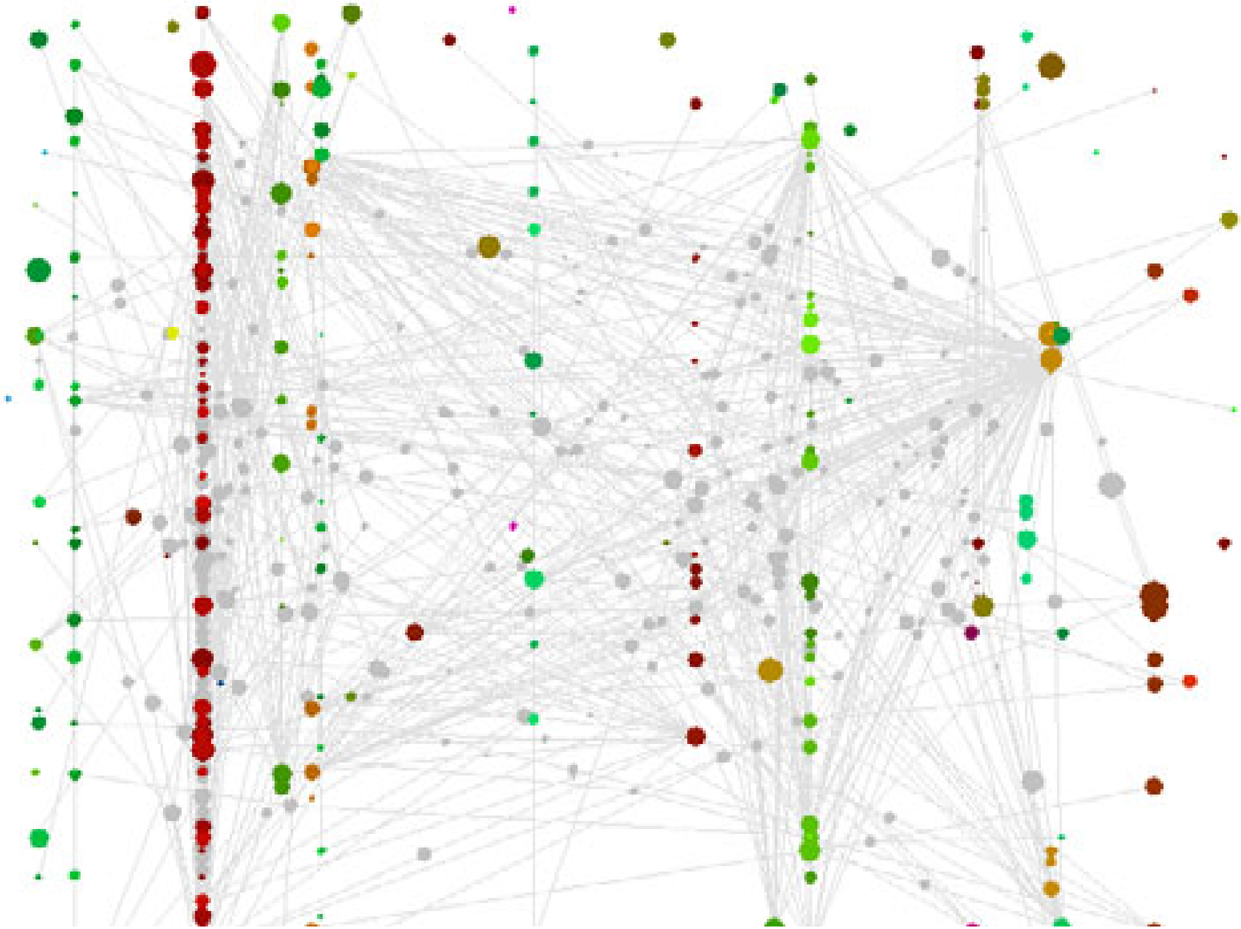
- Indirect communications
  - – eg two basketball fans



- Port Activity (Services)
  - – Eg web servers/surfers, IRC



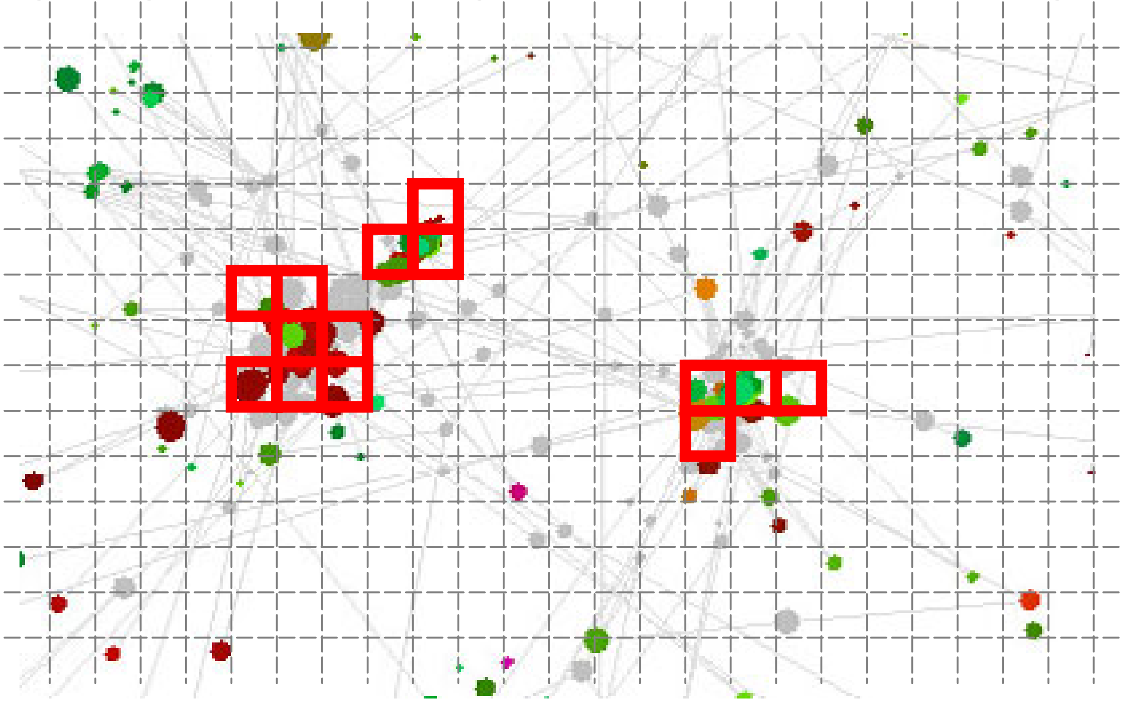
Initialization of Nodes

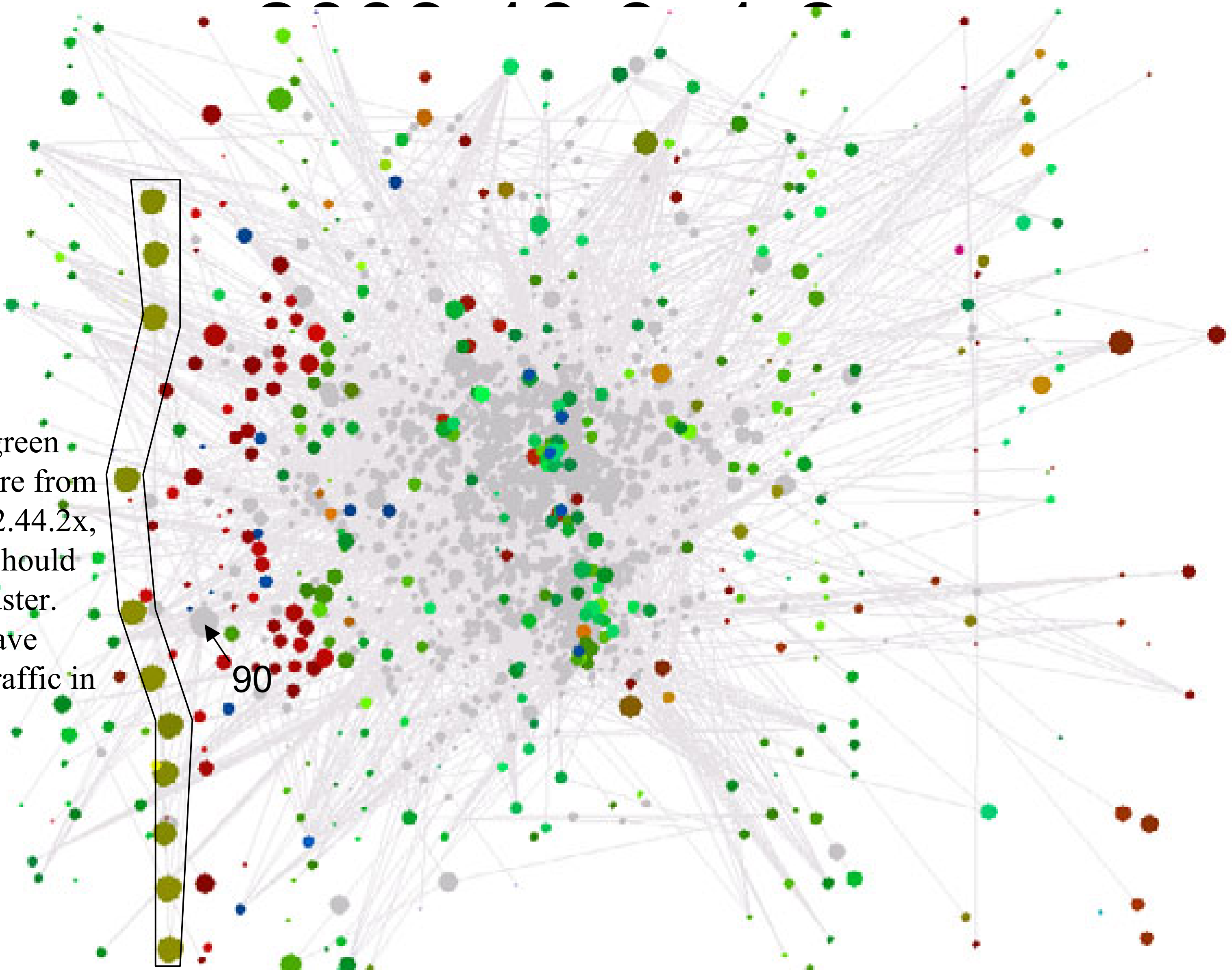


Colored points represent internal hosts, and gray points represent external ones. Size of a point is proportional to logarithm of traffic volume involving this host.

Identifying Clusters

- A cluster is a dense region in the viz space
  - – divide the space into many small grids
  - – DBSCAN to find such dense grids
  - – highlight dense grids and connect grids





These green nodes are from 141.142.44.2x, which should be a cluster. They have much traffic in port 90.

- Motivation
- Network Visualization for Security
- Our Approach: VisFlowConnect-IP
- Use Examples
- Future Work: Link-Based Clustering
- **Summary**

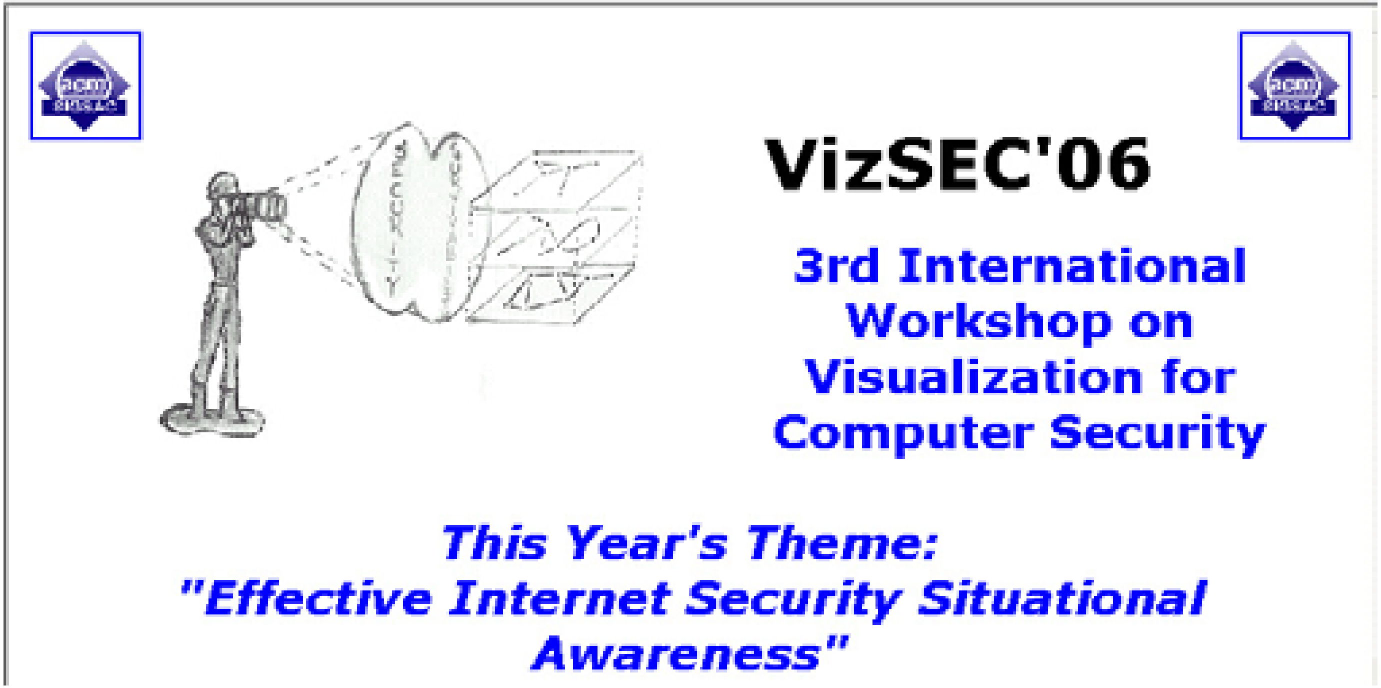
Summary

- VisFlowConnect-IP can visualize traffic in near-realtime for security monitoring purposes
- VisFlowConnect-IP is being ported to other specialized security domains
  - storage systems, linux clusters, etc.
- Distribution Website

  <http://security.ncsa.uiuc.edu/distribution/VisFlowConnectDownLoad.html>
- Publications

  <http://www.ncassr.org/projects/sift/papers/>

VizSEC Workshops



VizSEC'06

3rd International Workshop on Visualization for Computer Security

This Year's Theme:
"Effective Internet Security Situational Awareness"

<http://www.projects.ncassr.org/sift/vizsec/>

References

- William Yurcik, "Visualizing NetFlows for Security at Line Speed: The SIFT Tool Suite," 19th Usenix Large Installation System Administration Conference (LISA), San Diego, CA USA, 2005.

- Xiaoxin Yin, William Yurcik, and Adam Slagell, "VisFlowConnect-IP: An Animated Link Analysis Tool for Visualizing Netflows," FLOCON - Network Flow Analysis Workshop, Pittsburgh PA USA, 2005.

- Xiaoxin Yin, William Yurcik, and Adam Slagell, "The Design of VisFlowConnect-IP: a Link Analysis System for IP Security Situational Awareness," 3rd IEEE Intl. Workshop on Information Assurance (IWIA) University of Maryland USA, 2005.

- Xiaoxin Yin, William Yurcik, Michael Treaster, Yifan Li, and Kiran Lakkaraju " VisFlowConnect: NetFlow Visualizations of Link Relationships for Security Situational Awareness," CCS Workshop on Visualization and Data Mining for Computer Security (VizSEC/DMSEC) held in conjunction with 11th ACM Conf. on Computer and Communications Security, 2004.

- Xiaoxin Yin, William Yurcik, Yifan Li, Kiran Lakkaraju, Cristina Abad, "VisFlowConnect: Providing Security Situational Awareness by Visualizing Network Traffic Flows," 23rd IEEE Intl. Performance Computing and Communications Conference (IPCCC), 2004.

- Cristina Abad, Yifan Li, Kiran Lakkaraju, Xiaoxin Yin, and William Yurcik, "Correlation Between NetFlow System and Network Views for Intrusion Detection," Workshop on Link Analysis, Counter-terrorism, and Privacy held in conjunction with the SIAM International Conference on Data Mining (ICDM), 2004.

Q & A

VisFlowConnect-IP

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

Disclaimer:

- This material is, in part, based upon work supported by the Office of Naval Research.
- Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the Office of Naval Research.

NetFlows for Security

NetFlows can identify connection-oriented attacks like DoS, DDoS, malware distribution, worm scanning, etc…

- How many users are on the network at any given time? (upgrades)
- Top N talkers? Top N destination ports?
- How long do users surf?
- Where do they go? Where did they come from?
- Are users following the security policy?
- Is there traffic to vulnerable hosts?
- Can you identify and block bad guys?