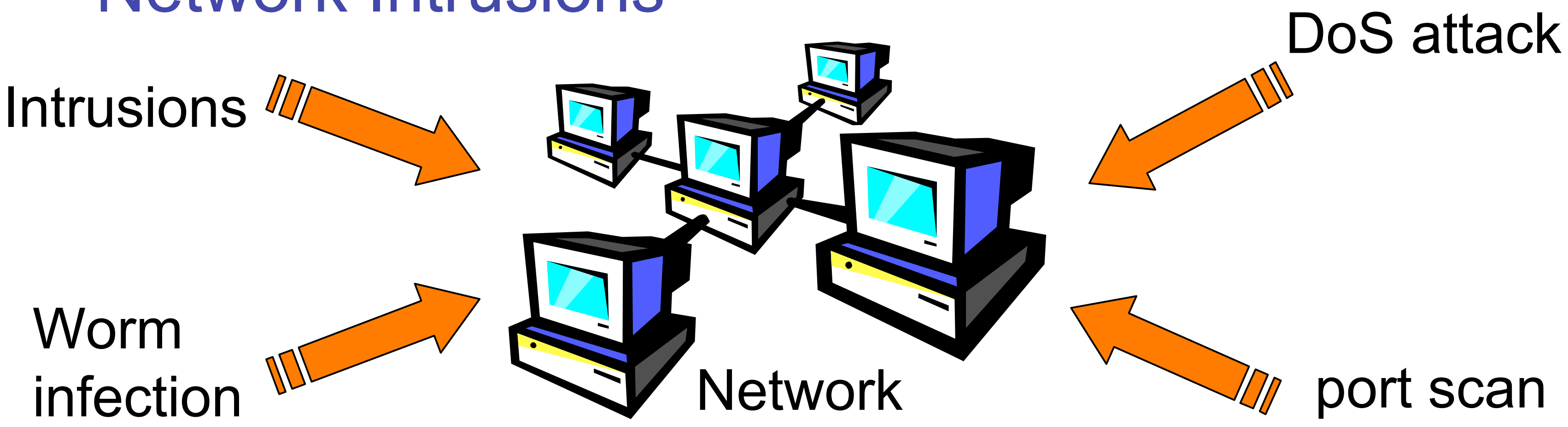
VisFlowConnect-IP: An Animated Link Analysis Tool For Visualizing Netflows

Xiaoxin Yin*, William Yurcik, Adam Slagell

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

Motivations

- Network Intrusions



Intrusions

DoS attack

Worm infection

Network

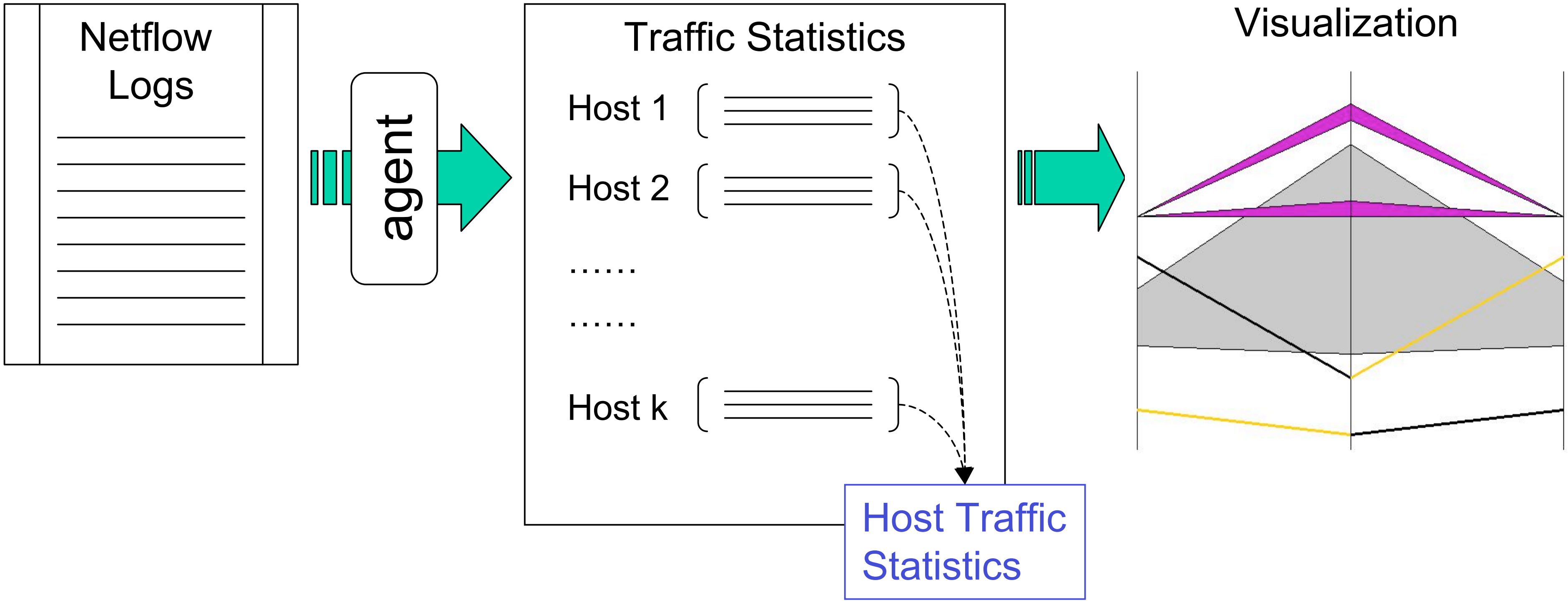
port scan

- Intrusion Detection Systems
  - Misuse detection: find signatures of intrusions
  - Anomaly detection: model normal behaviors
- Visualize network traffic
  - So that intrusions are apparent to human

Overview

- Visualizing network traffic as a graph
  - Hosts → nodes in graph
  - Traffic → flow in graph
  - other conceptual models are possible (i.e. NVisionIP)
- Visualizing by ***animation***
  - Show network dynamics by animation
  - Visualize traffic within a user adjustable time window
- High scalability
  - Run continuously for long periods
  - Uses constant storage to process large data sets or high speed streaming data.

VisFlowConnect System Design



Netflow Logs

agent

Traffic Statistics

Host 1

Host 2

......

......

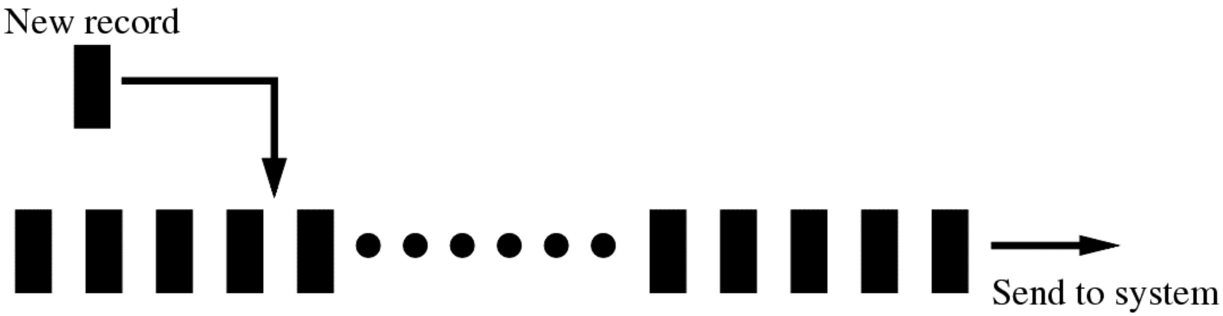
Host k

Host Traffic Statistics

Visualization

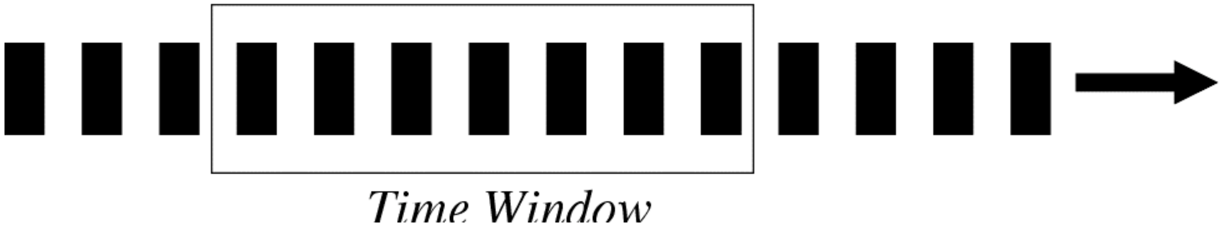
Reading Netflow Logs

- An agent reads records from file or in real time
  - Send a record to VisFlowConnect when requested
- Reorder Netflow records with record buffer
  - Records are not strictly sorted by time stamps
  - Use a record buffer



Time Window

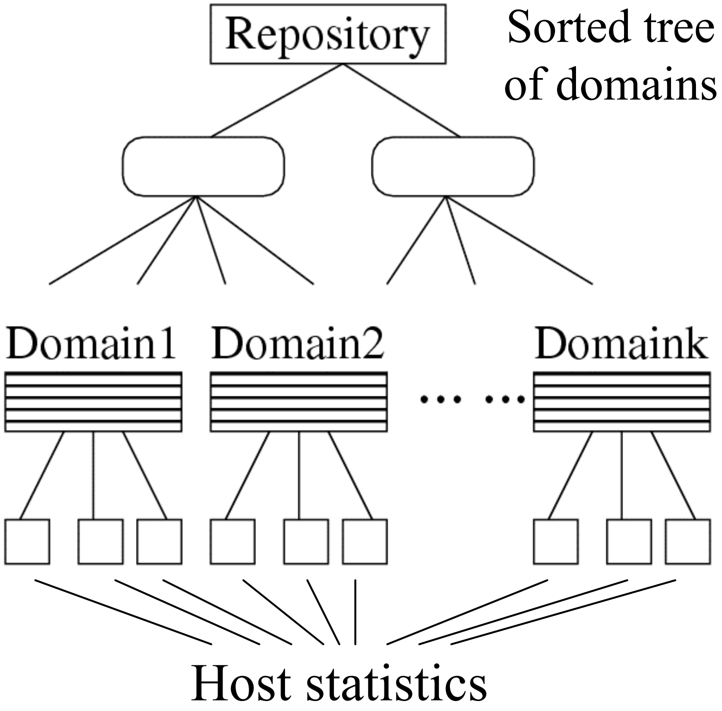
- User is usually interested in most recent traffic (e.g., in last minute or last hour)
- VisFlowConnect only visualizes traffic in a user adjustable time window



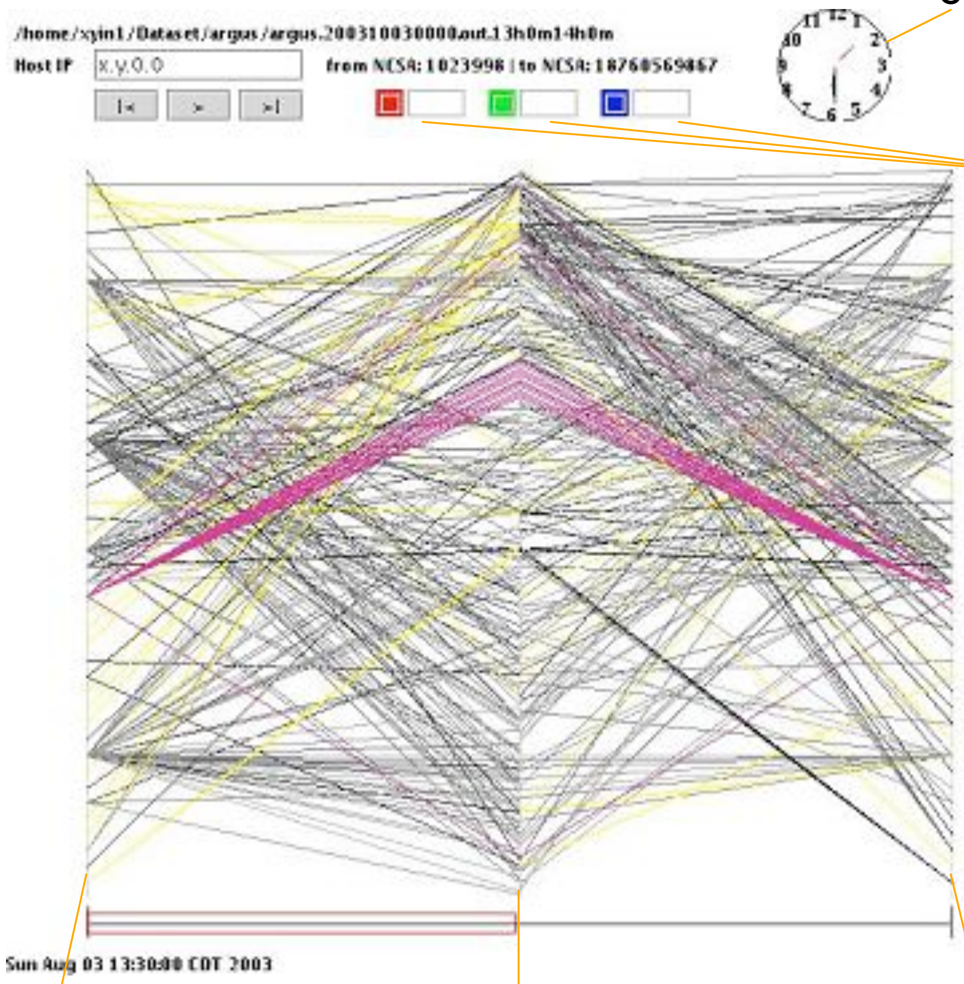
  - Update traffic statistics when
    - A record comes into time window
    - A record goes out of time window

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



VisFlowConnect External View



/home/xyin1/Dataset/argus/argus.200310030000.out.13h0m14h0m

Host IP x.y.0.0

from NCSA: 1023998 | to NCSA: 18760569867

Sun Aug 03 13:30:00 CDT 2003

clock

highlighted ports

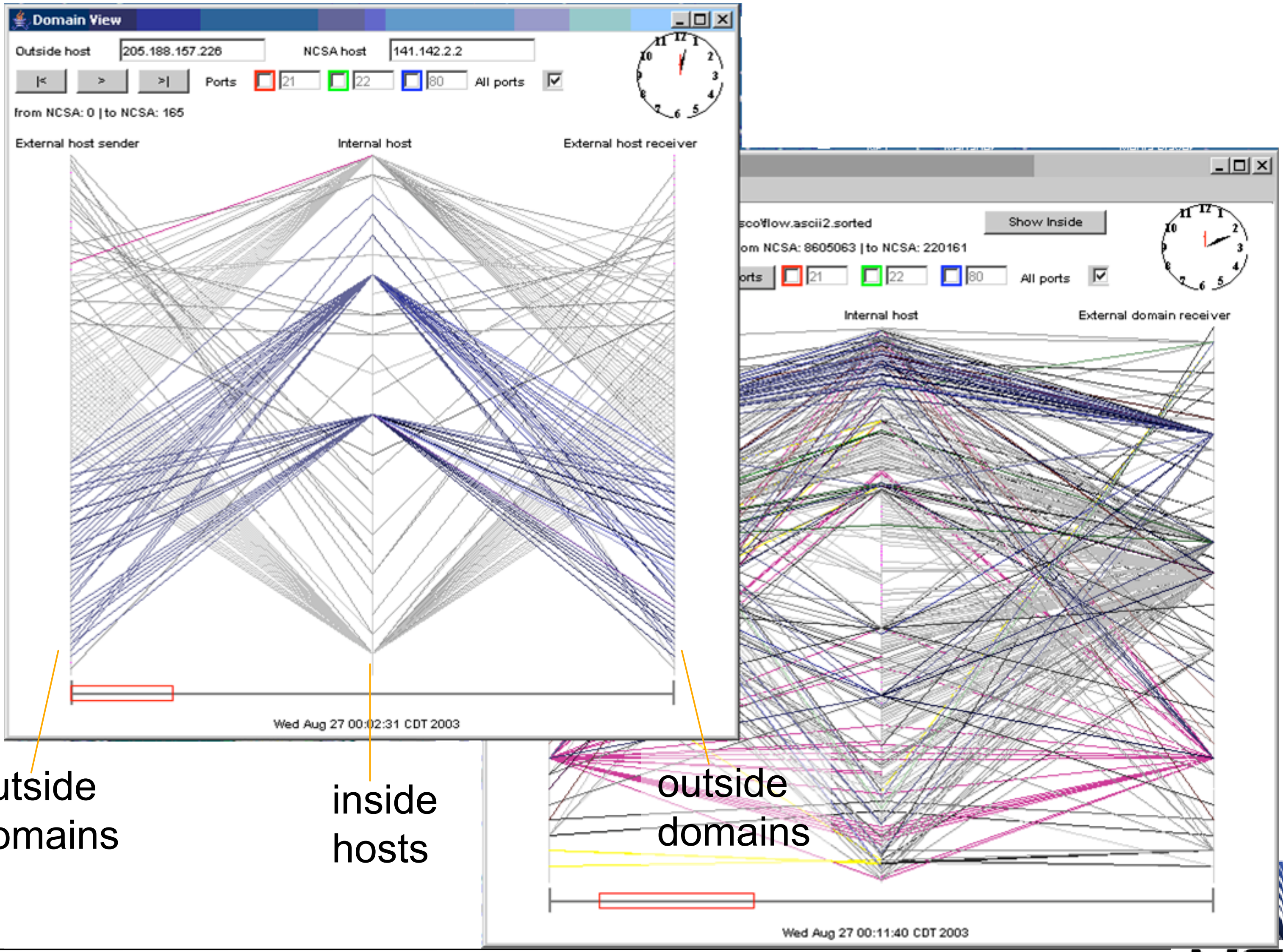
time axis

outside domains

inside hosts

outside domains

VisFlowConnect Domain View



Domain View

Outside host 205.188.157.226 NCSA host 141.142.2.2

|< > >| Ports □ 21 □ 22 □ 80 All ports ☑

from NCSA: 0 | to NCSA: 165

Wed Aug 27 00:02:31 CDT 2003

scoflow.ascii2.sorted Show Inside

om NCSA: 8605063 | to NCSA: 220161

orts □ 21 □ 22 □ 80 All ports ☑

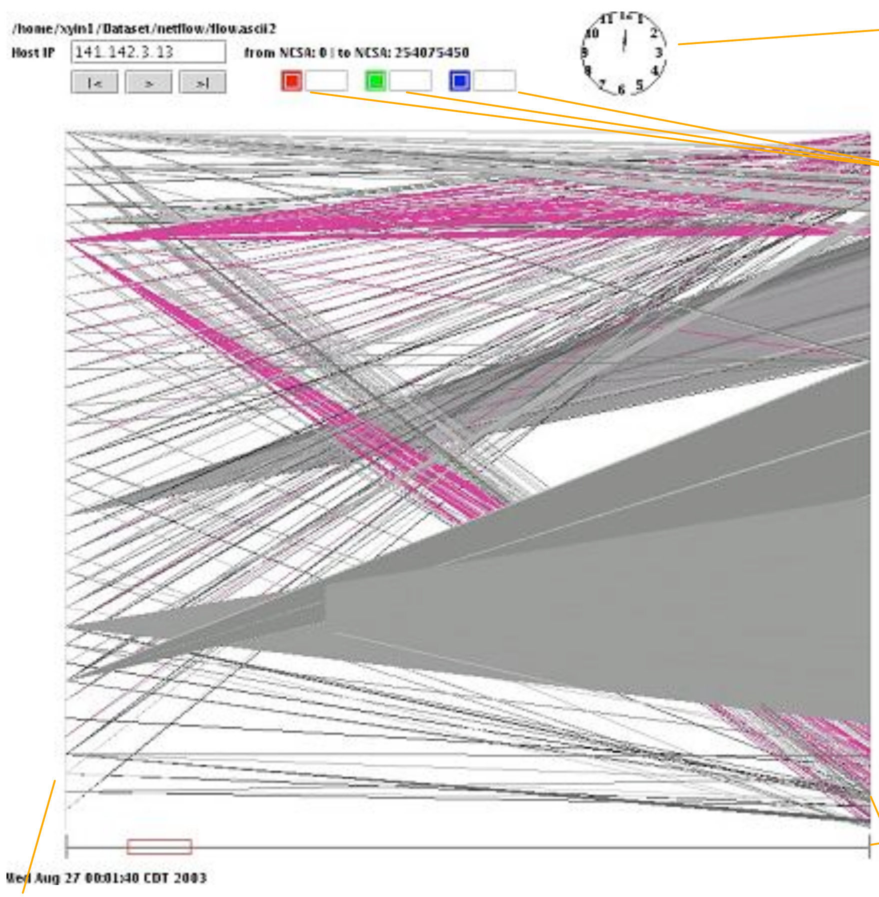
Wed Aug 27 00:11:40 CDT 2003

outside domains

inside hosts

outside domains

VisFlowConnect Internal View



/home/xyin1/Dataset/netflow/flow.ascii2

Host IP 141.142.3.13

from NCSA: 0 | to NCSA: 254075450

Wed Aug 27 00:01:40 CDT 2003

clock

highlighted ports

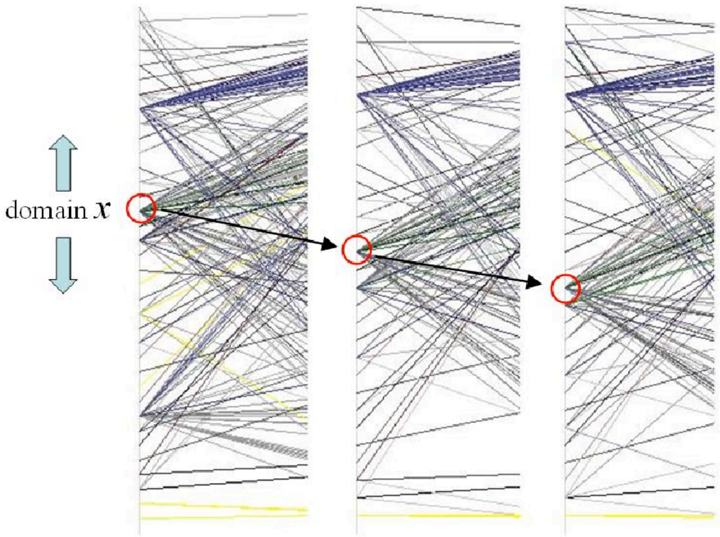
time axis

inside senders

inside receivers

Creating Animation

- Visualizing traffic statistics with time
  - Update visualization after each time unit
- How to arrange domains/hosts?
  - Only hundreds of domains/hosts can be put on one axis
  - Domains/hosts may be added or removed with time
  - The position of each domain/host must be fairly stable
- Solution: sort domains/hosts by IP
  - Each domain/IP may move up or down



Filtering Capability

- Filter out regular traffic
  - E.g., DNS traffic, common HTTP traffic
- Work like a spam-mail filter
  - 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

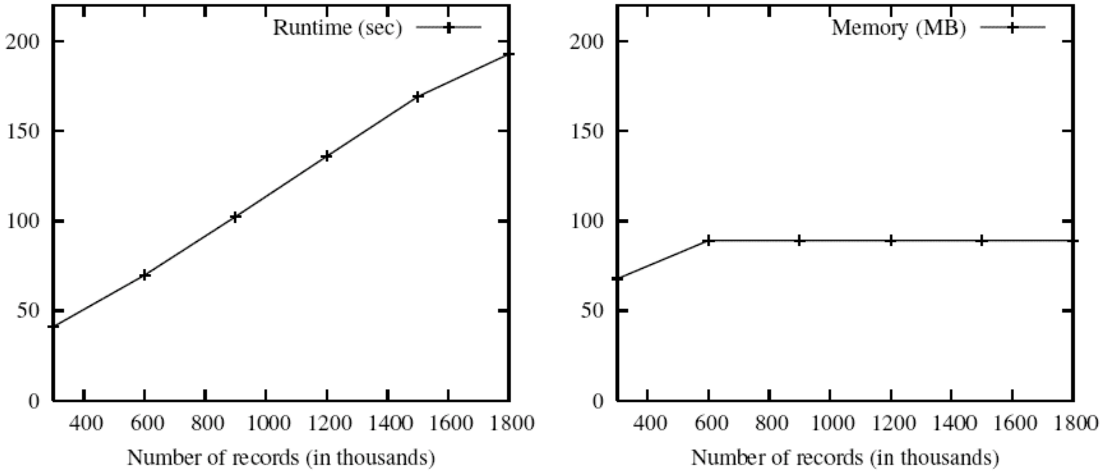
    −: (SrcPort=80)

    −: (DstPort=80)

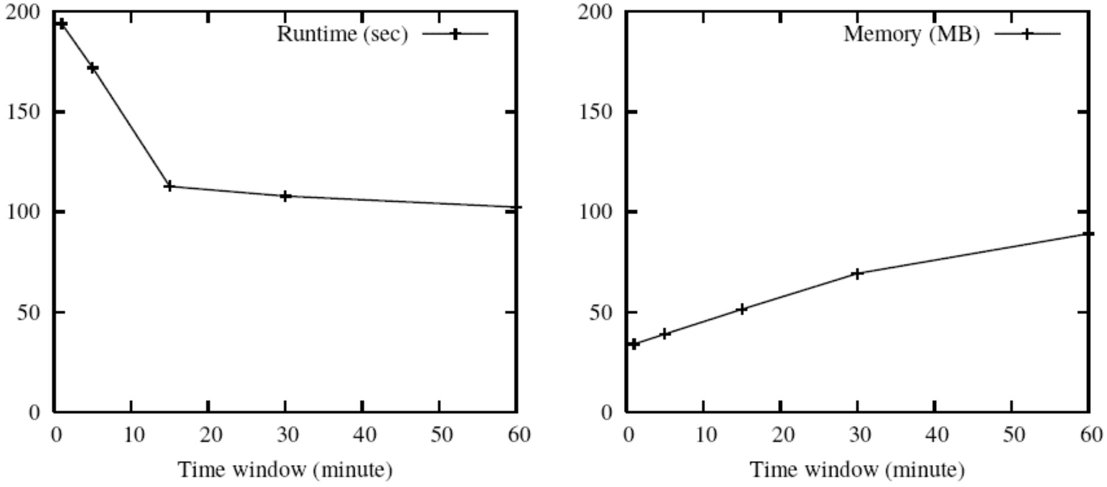
    //discard records of http traffic
  - Each record is passed through each filter
  - Last match is used to decide whether keep this record or not

Scalability Experiments

Runtime and memory w.r.t. number of records

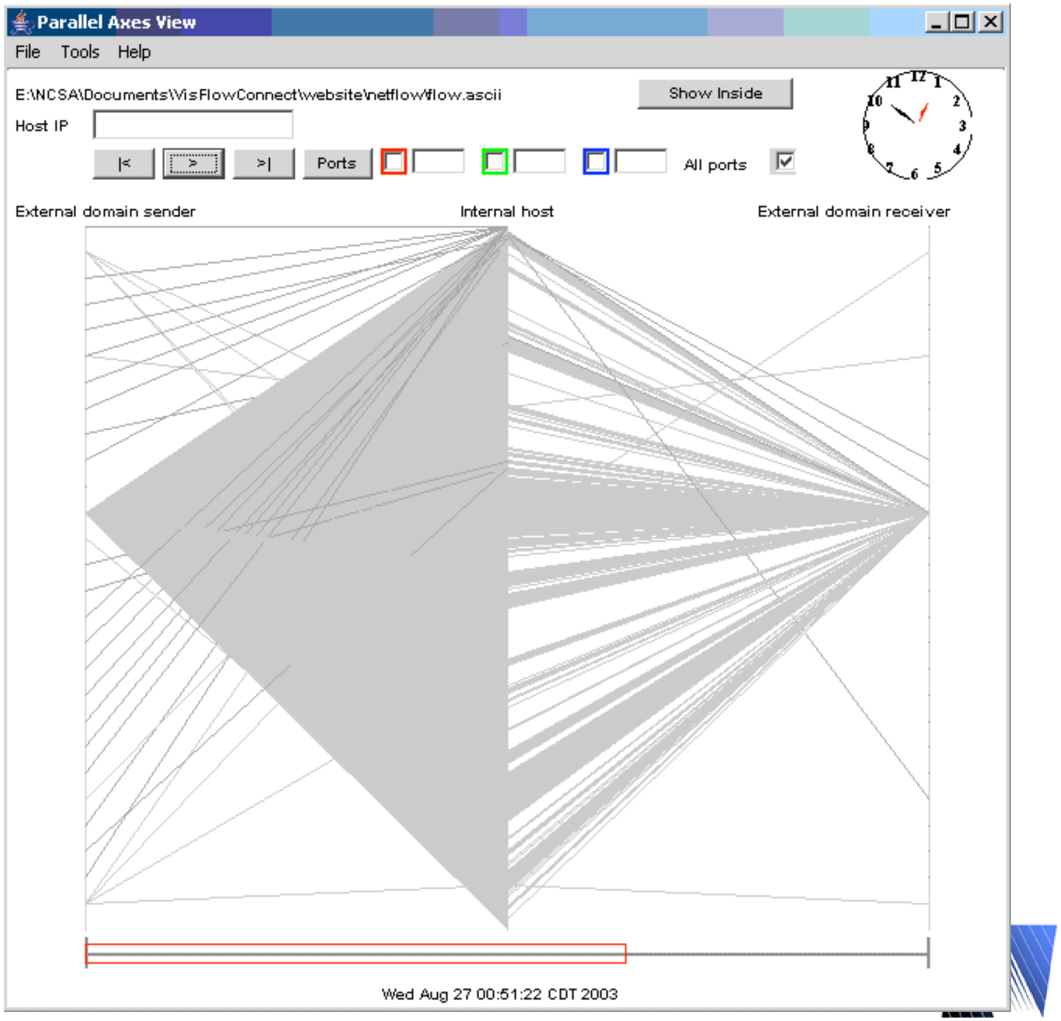


Runtime and memory w.r.t. size of time window



Example 1: MS Blaster Virus

- MS Blaster virus causes machines to send out packets of size 92 to many machines

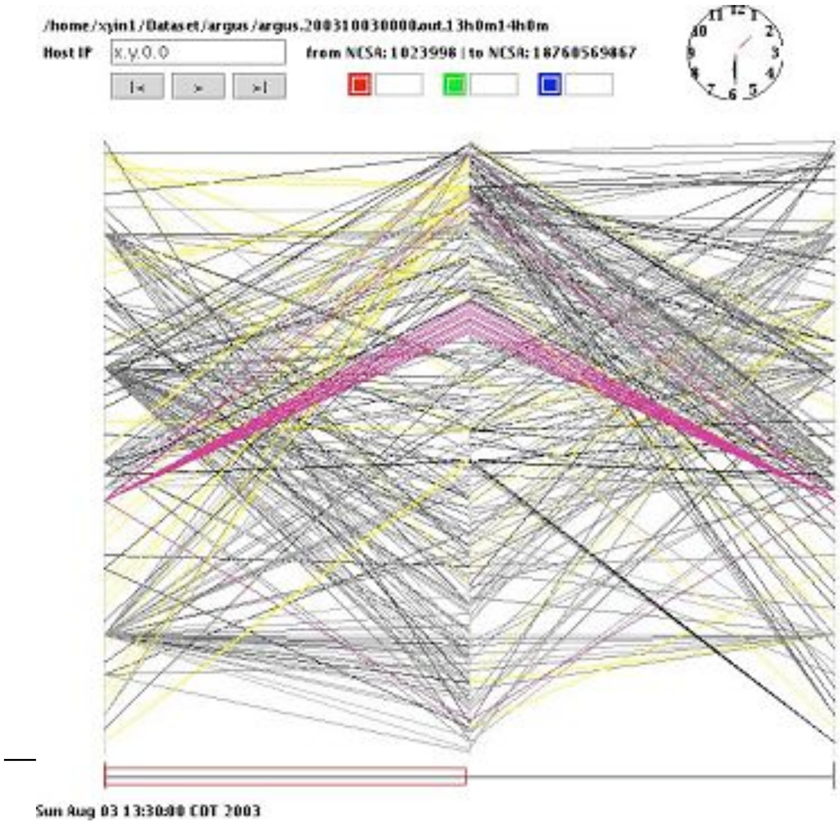


E:\NCSA\Documents\VisFlowConnect\website\netflow\flow.ascii

Wed Aug 27 00:51:22 CDT 2003

Example 2: 1-to-1 Communication of Clusters

- We found there are two sets of hosts of continuous IPs have 1-to-1 communications with each other. Finally we found they are two clusters.

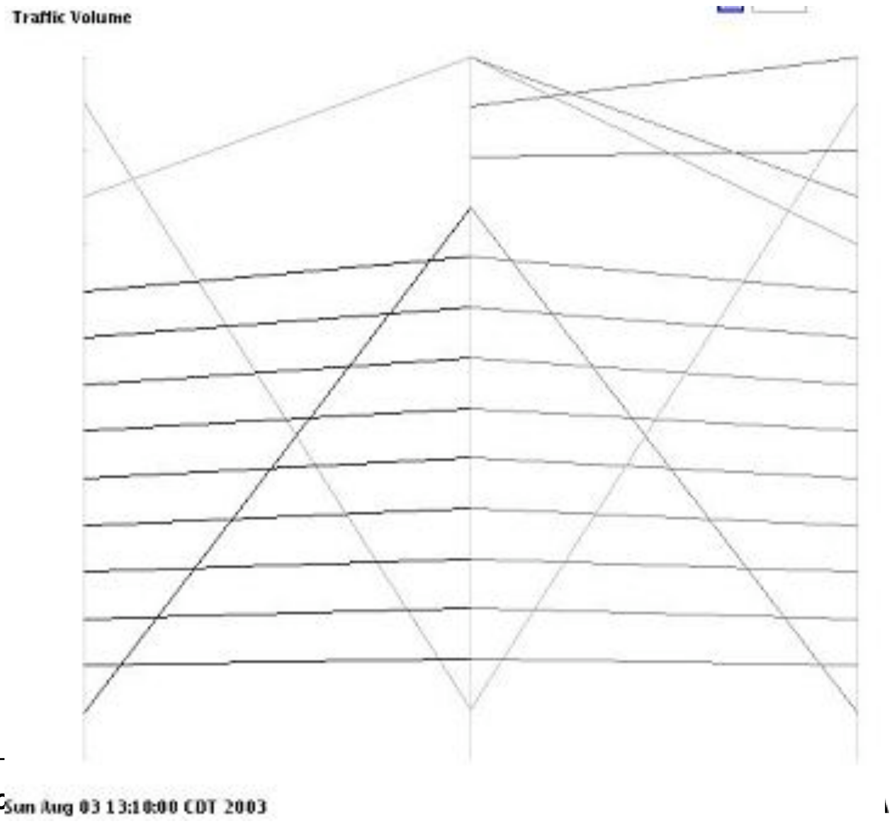


/home/xyin1/Dataset/argus/argus.200310030000out.13h0m14h0m

Host IP x.y.0.0

from NCSA: 1023998 | to NCSA: 18760569867

Sun Aug 03 13:30:00 CDT 2003

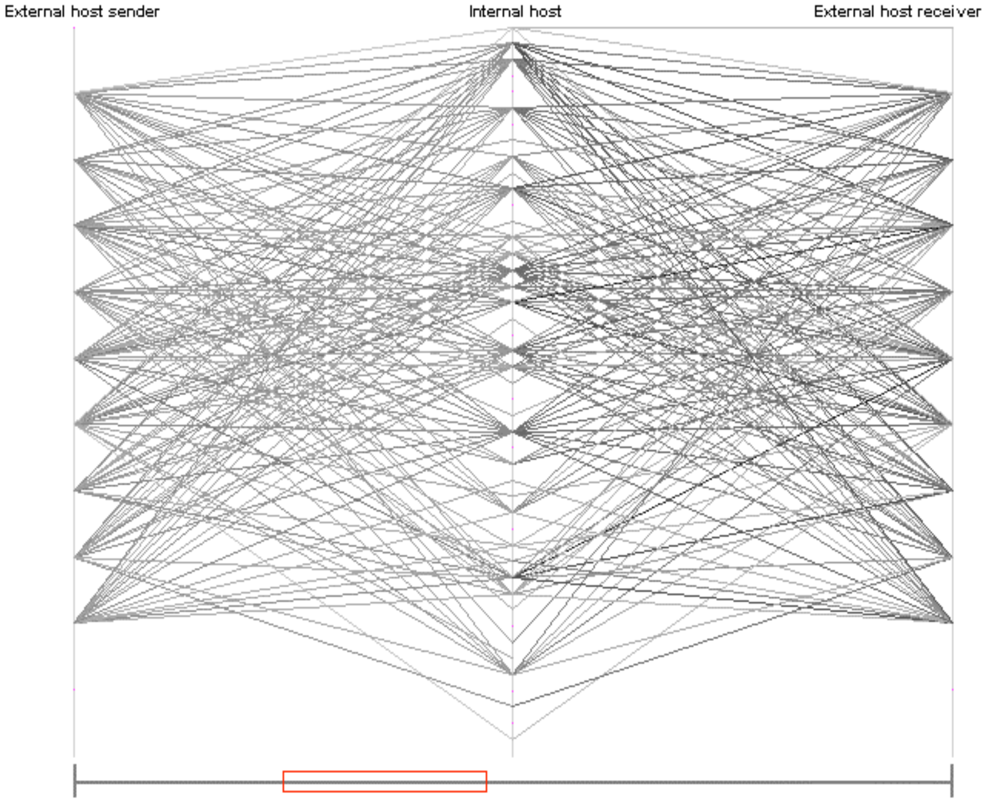


Traffic Volume

Sun Aug 03 13:10:00 CDT 2003

Example 3: Web Crawlers

- We found 9 hosts in a domain connecting to many http servers in our network
  - Their IPs are from Google.com: Web crawling



More Information

- VisFlowConnect is being ported to other specialized security domains
  - Storage security (two publications pending)
  - Cluster security
- Distribution Website
  - http://security.ncsa.uiuc.edu/distribution/VisFlowConnectDownLoad.html

    VisFlowConnect are downloadable there
- Publications of SIFT Group
  - http://www.ncassr.org/projects/sift/papers/

Thank You!

Xiaoxin Yin

<xyin1@uiuc.edu>

NCSA SIFT Group

University of Illinois