VISUALIZATION ANALYSES & INFORMATION SECURITY

-GLSP meeting, 4/11/13

What is Visual Analysis (VA)?

- "The science of reasoning through visualization"

---- A picture worth thousand lines of logs

General VA Approach

- Overview first
- Change graph attributes
- Zoom and filter
- Details on demand

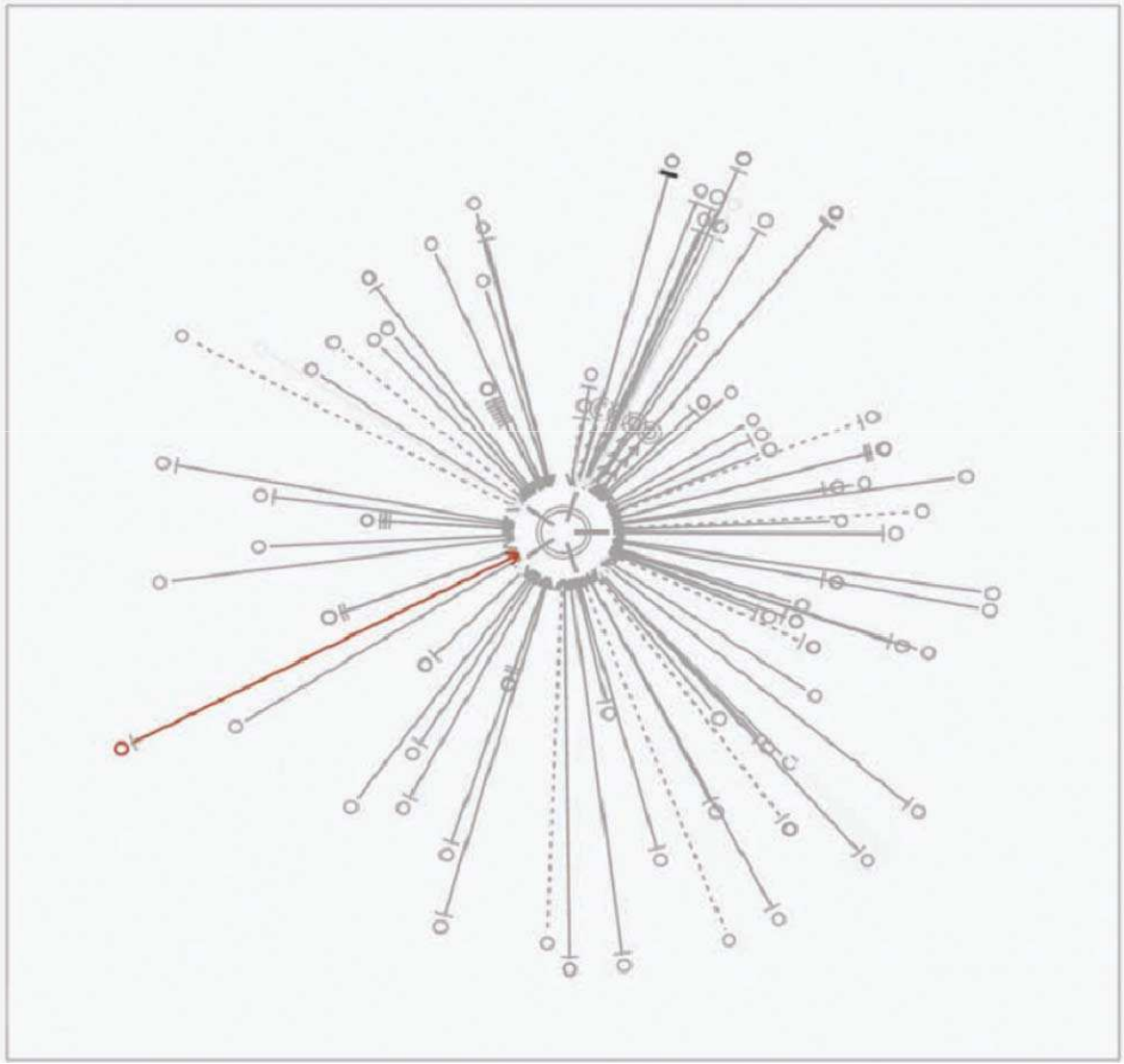
Why Security Visualization (VizSec)

- The security problem we are facing:
  - Multiple distributed data sources
  - High-Dimensional
    - Spatio-temporal
    - Activities generated by human
    - Log generated by machine
  - Big data
  - Unpredictable
  - Often need immediate response
  - Root cause analyses

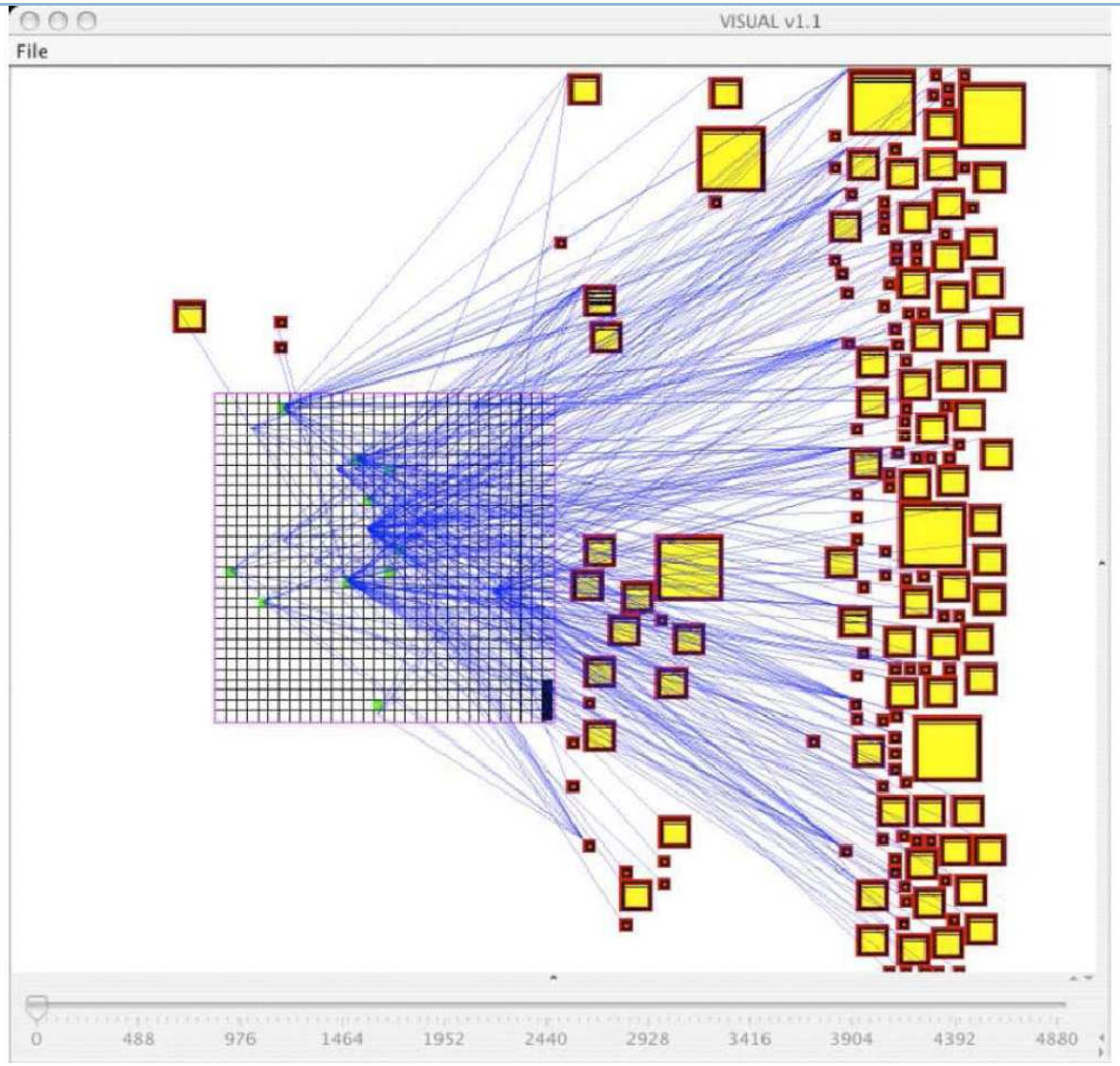
Current Work in VizSec

- Research Papers focus on applying the best visualization techniques to help analyzing data, often from single source

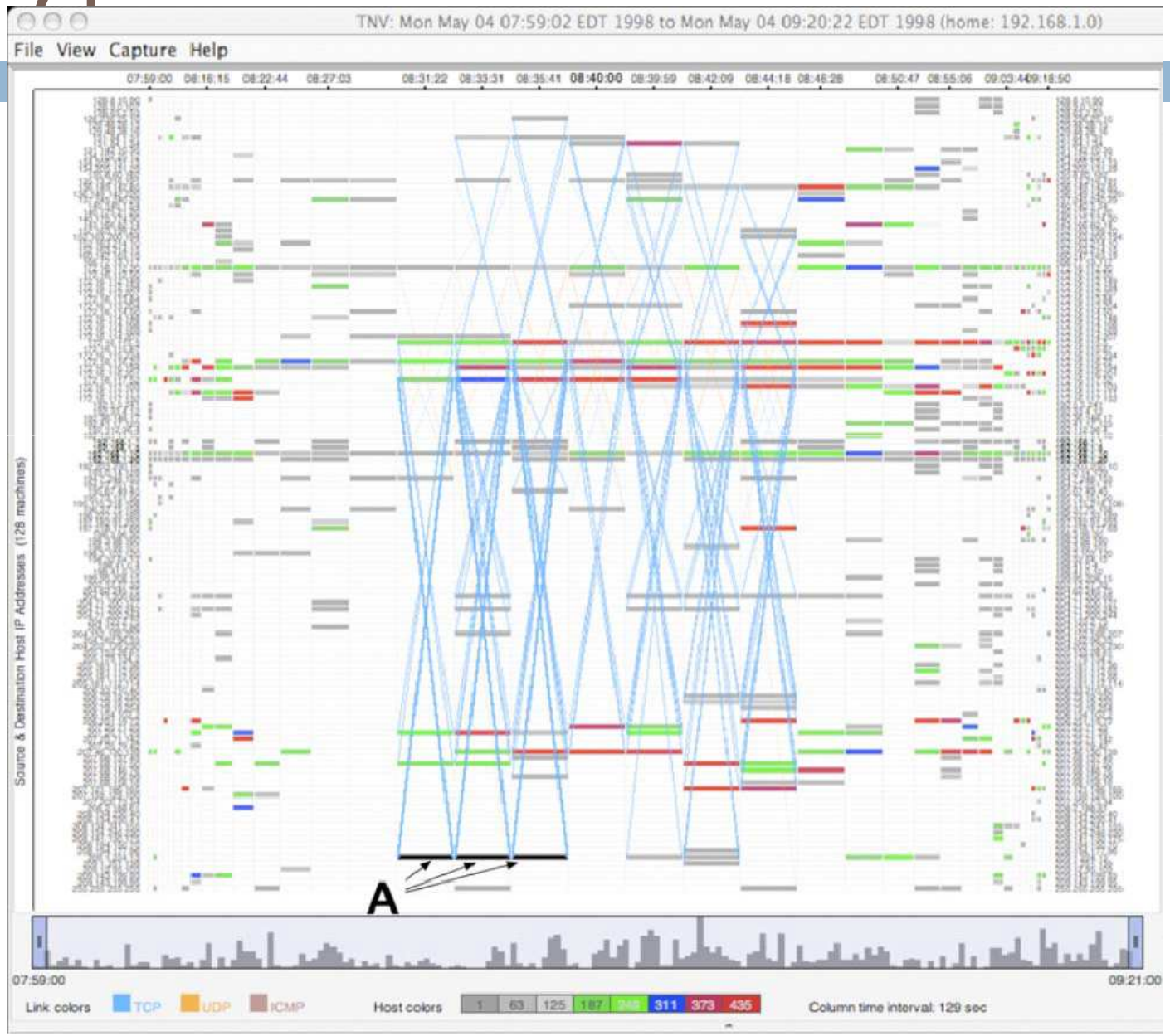
Glyph, server log



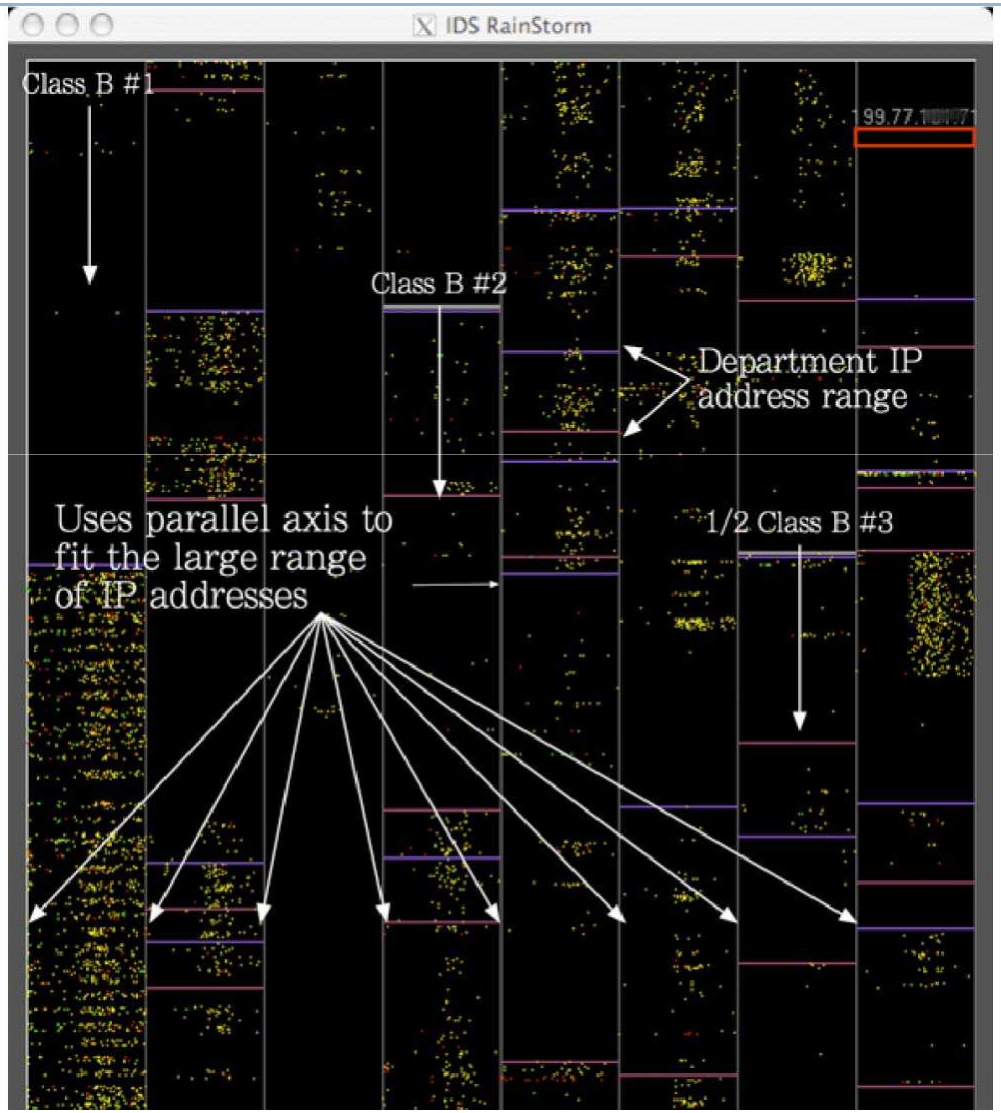
VISUAL, packet traces



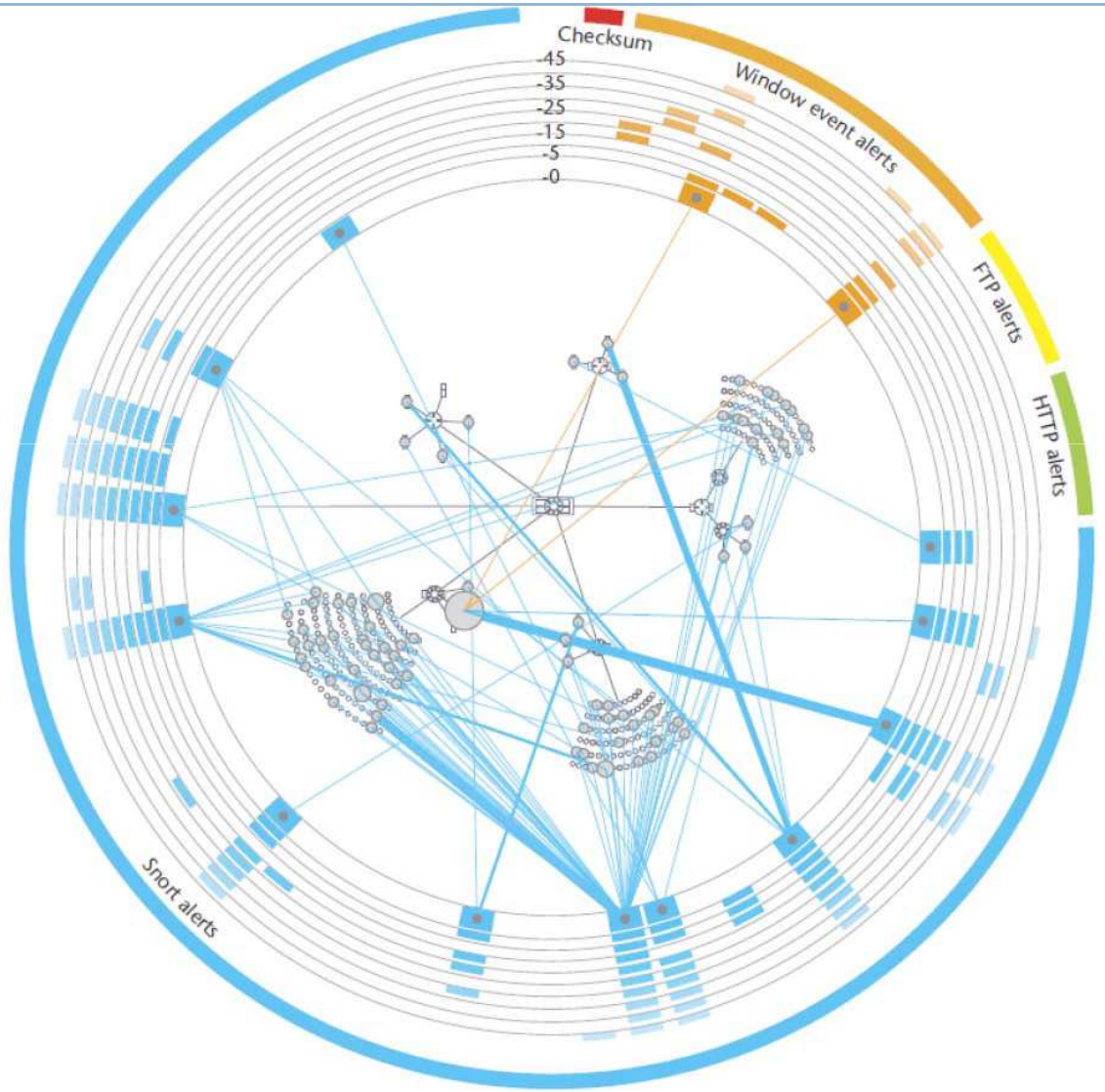
TNV, packet traces



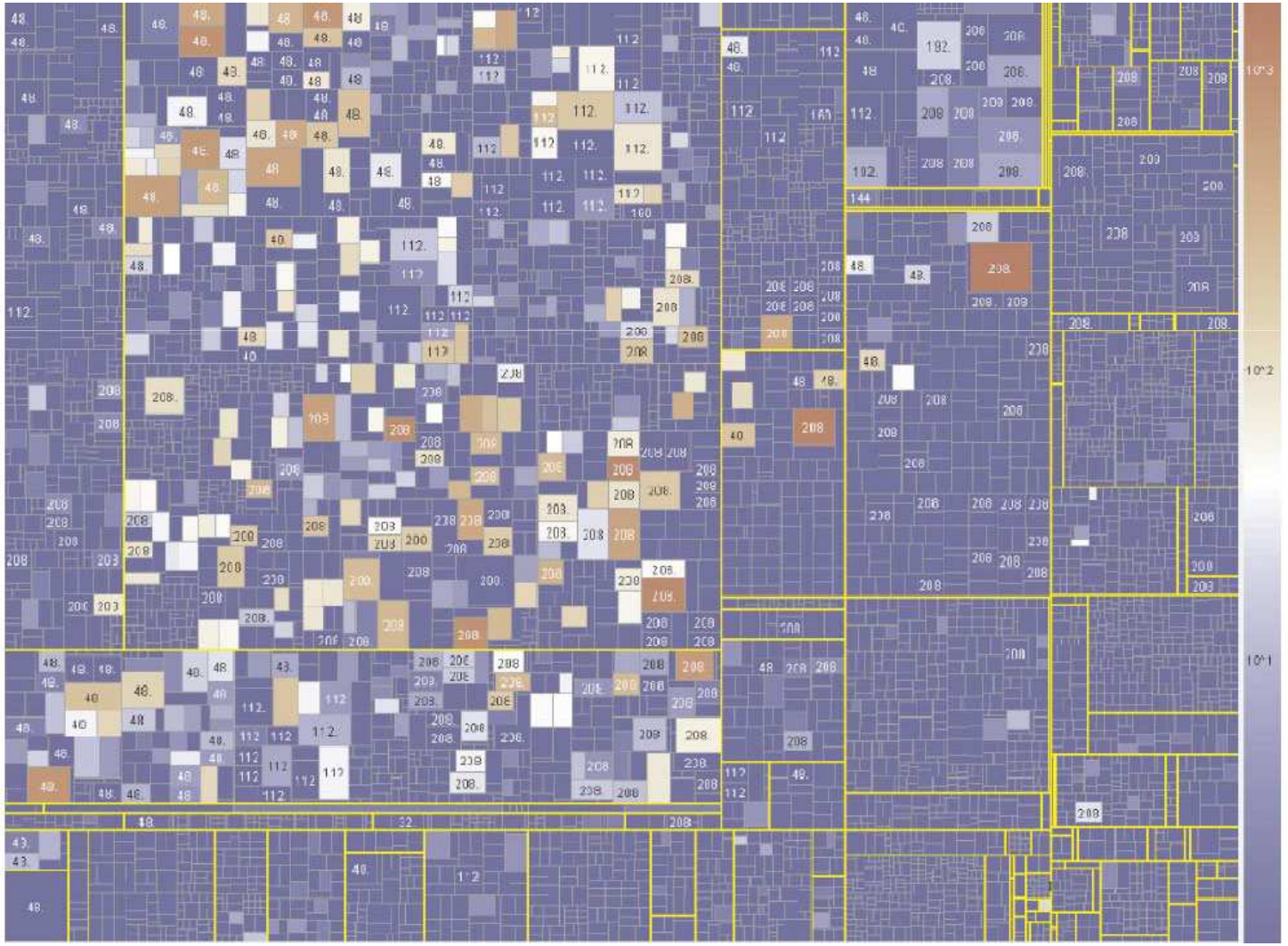
IDS RainStorm, intrusion alerts



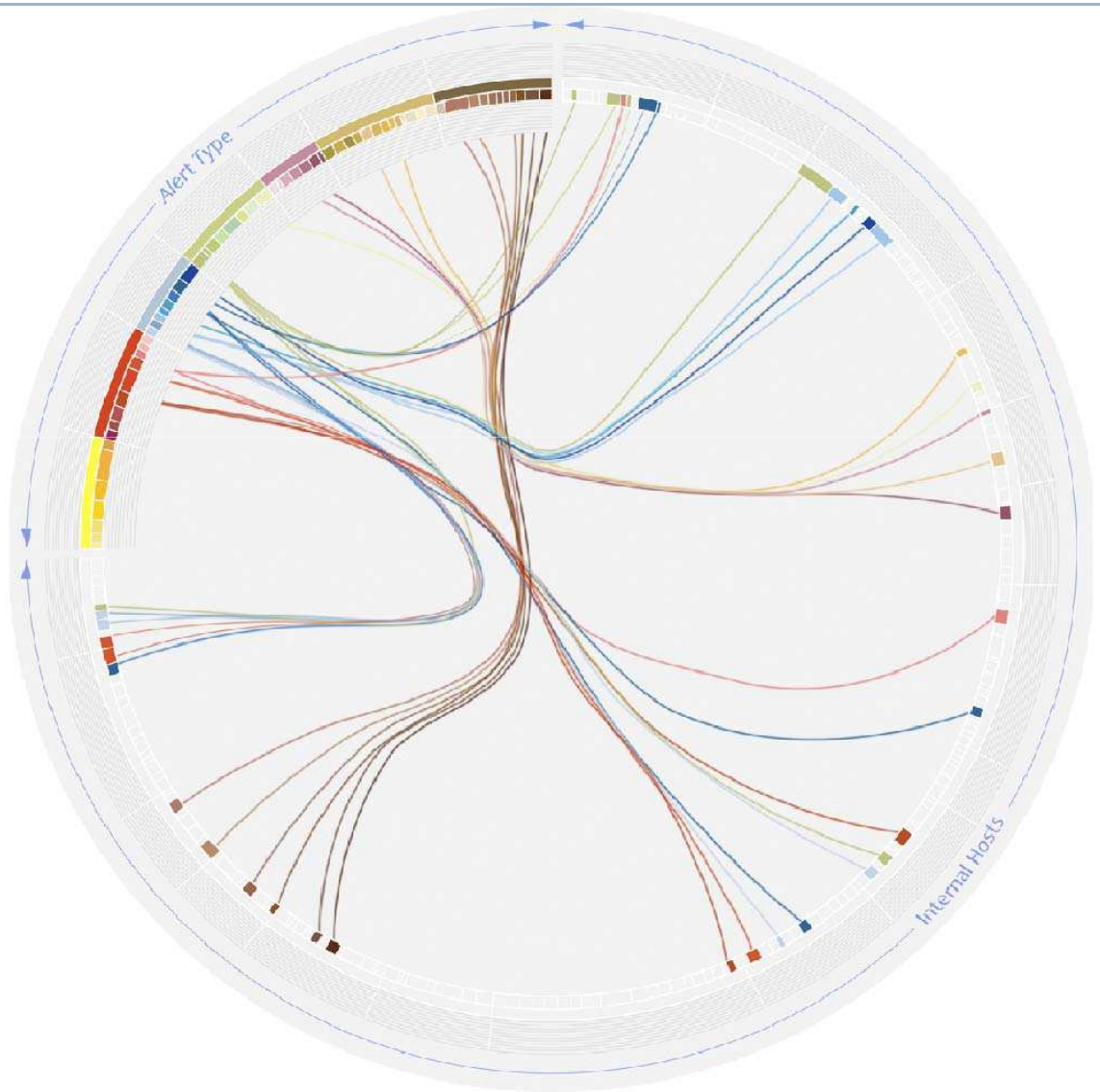
PortVis, netflows



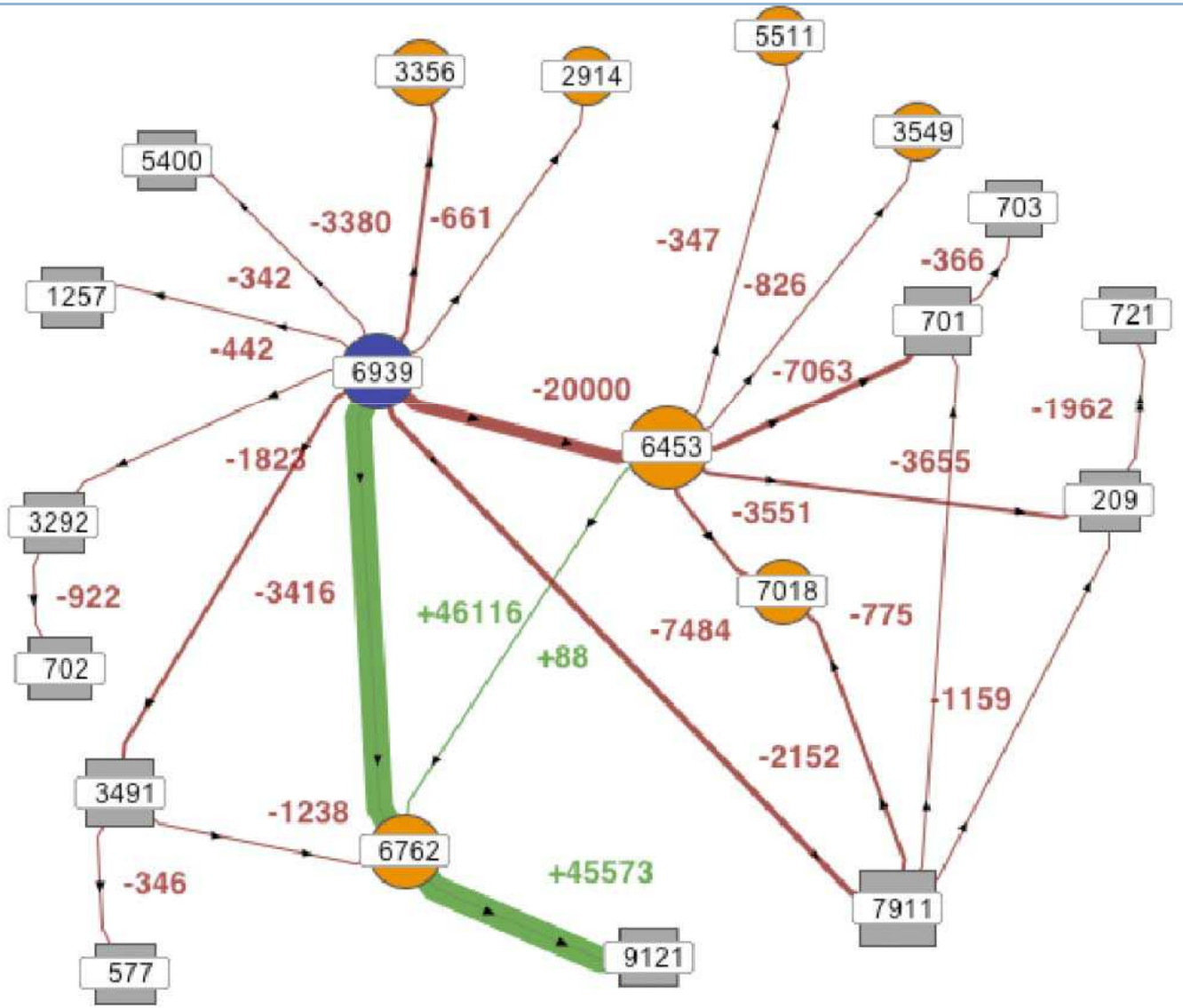
Mansmann, Packet Traces



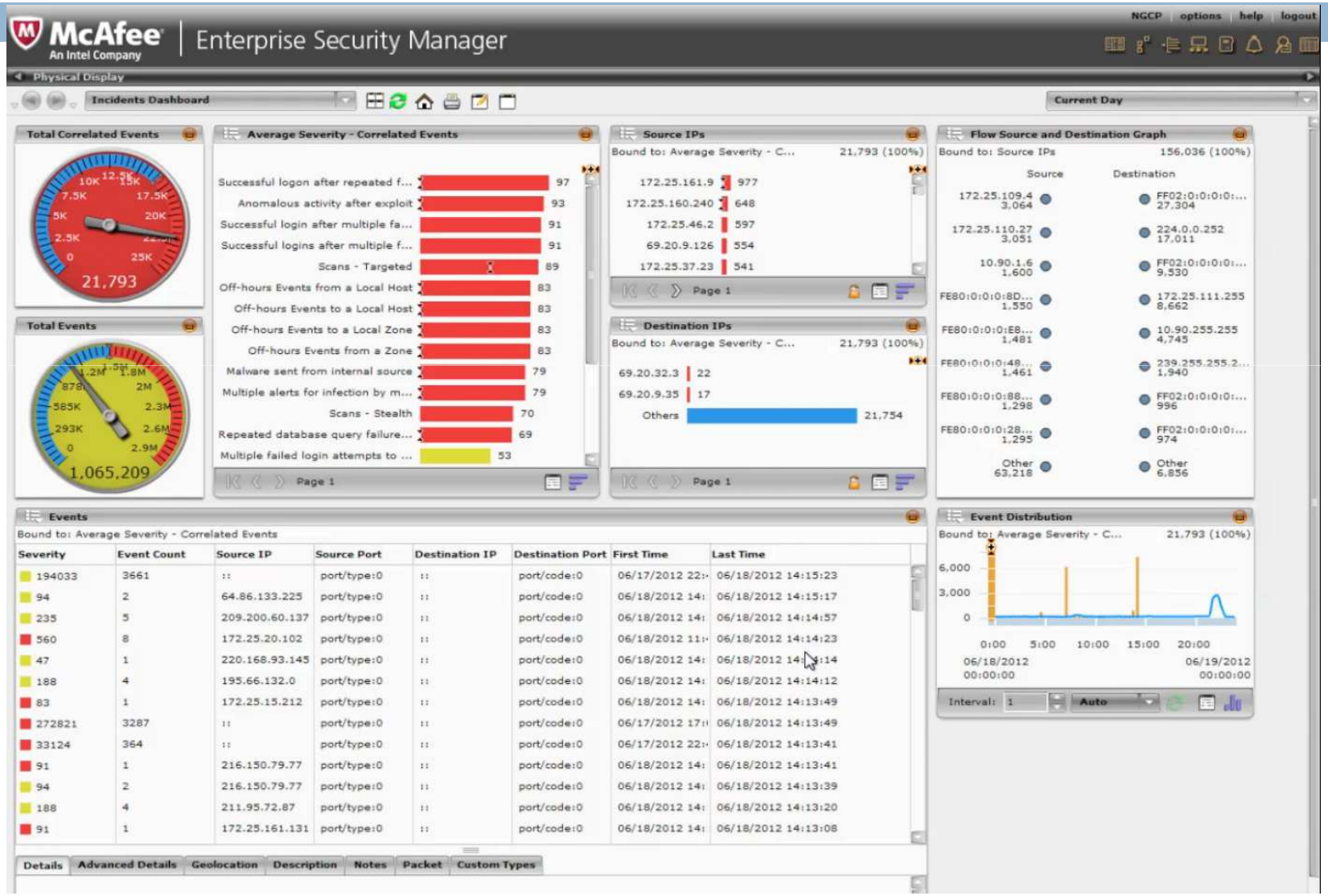
Avisa, intrusion alerts



LinkRank, BGP routing messages



Industry Approach (reporting)

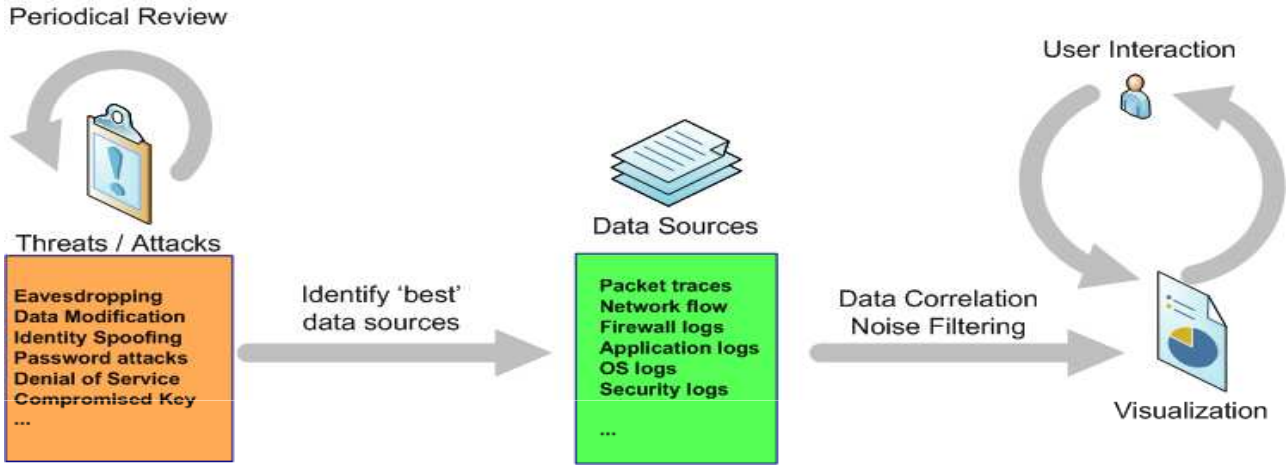


DAVIX, a collection of SecViz tools

- AfterGlow
- GraphViz
- NVisionIP
- GUESS
- TreeMap
- InteViz
- GnuPlot
- Ggobi
- Parvis
- GITail
- Rumint
- MRTG/RRD

http://secviz.org/content/the-davix-live-cd

Our Approach



- Integrate data from multiple sources
- Synthesize data based on events
- Inference analysis to remove background noises (e.g normal traffic data)
- OLAP (Online Analytical processing) approach to let the user overview the status from different perspectives at different level of granularities
- Semantic zooming interaction to connection overview to details

Related Research Questions

- Log format and correlation
  - No standard on Log
  - Log may not be available at all
  - Application log will be helpful, but too chatty
  - Time may not synchronized
- NAT/PAT, IP/Mac Spoofing complicate the process
- Data reduction (Summarization, noise filtering)
- …

Also, would Virtualization change anything?

Need your help

- Use case and threats
- Data