Flamingo: Visualizing Internet Traffic Data

Manish Karir, Jon Oberheide, Mike Goff

Networking R&D, Merit Network Inc.

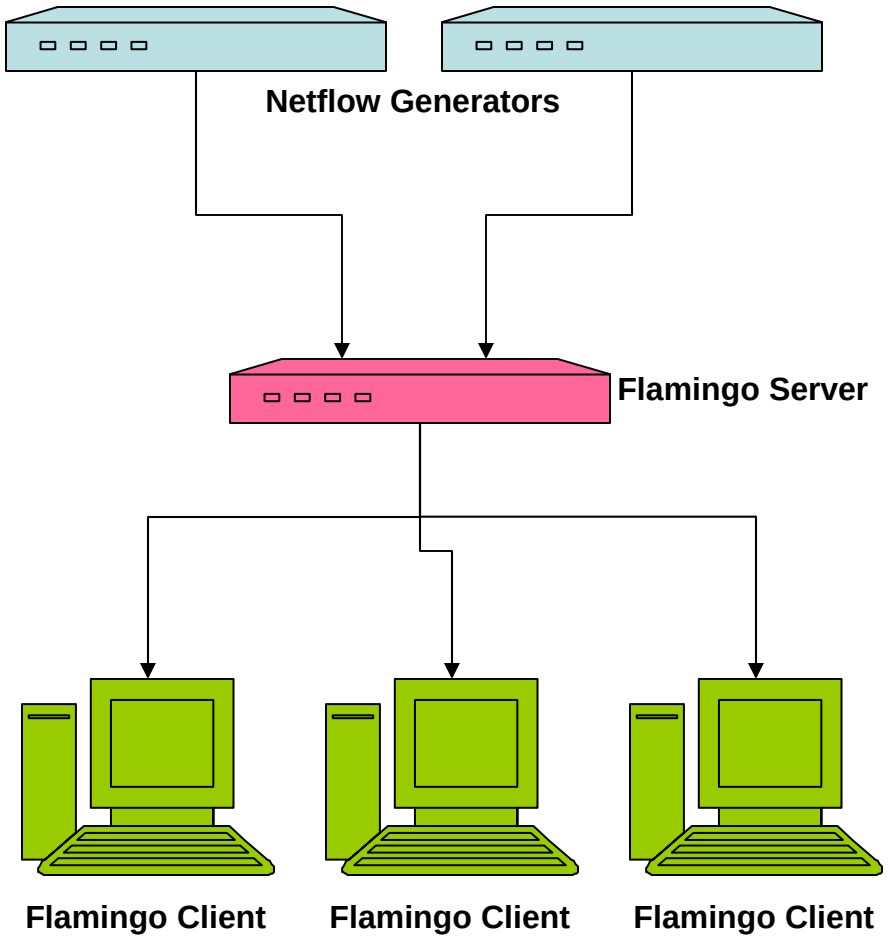
- Introduction: What is Flamingo
- Visualizations
- The Flamingo Tool
  - Combining visualization with controls
- Case Studies
  - Traffic Anomaly
  - Network Scans
  - Worm traffic
  - P2P traffic
  - The Slashdot effect!

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

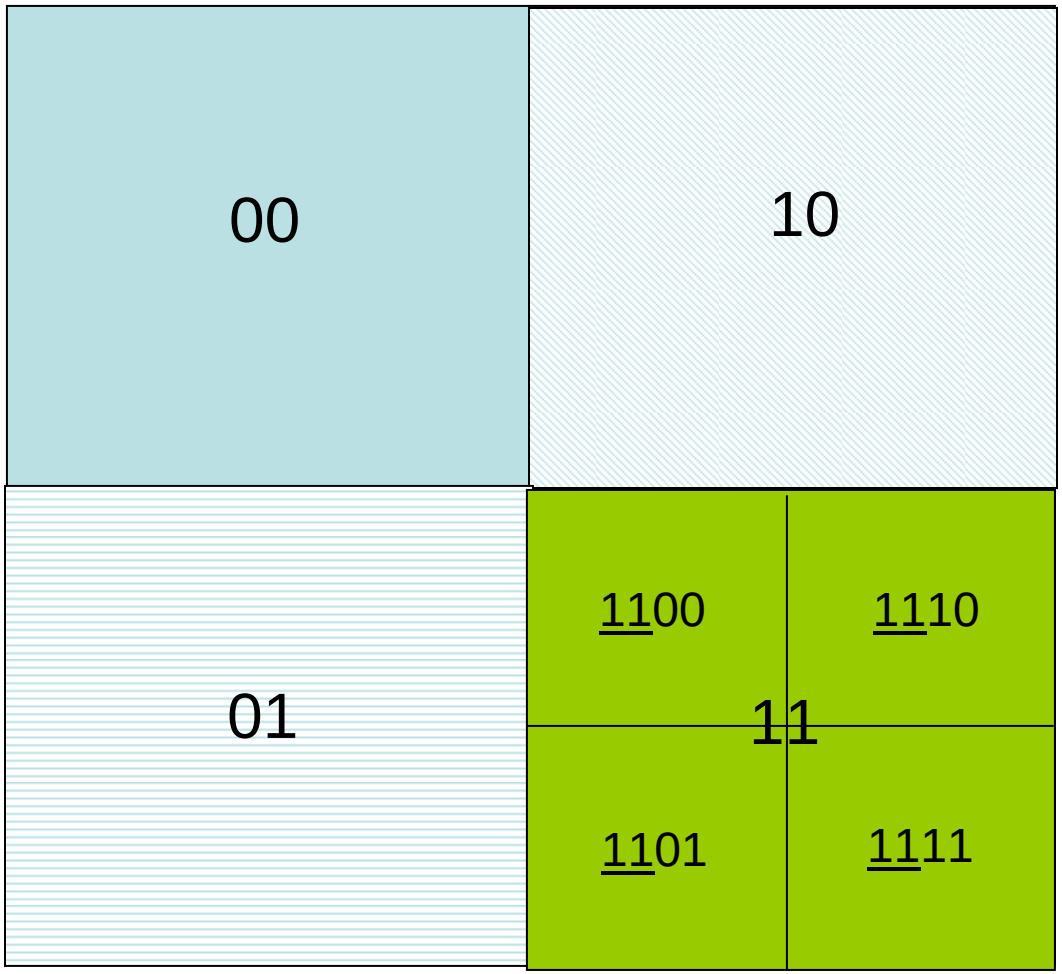
- Client/Server Architecture
- A single server can support multiple clients
- A single server can act as collector for multiple netflow feeds
- Supports both aggregation as well as non-aggregation mode



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

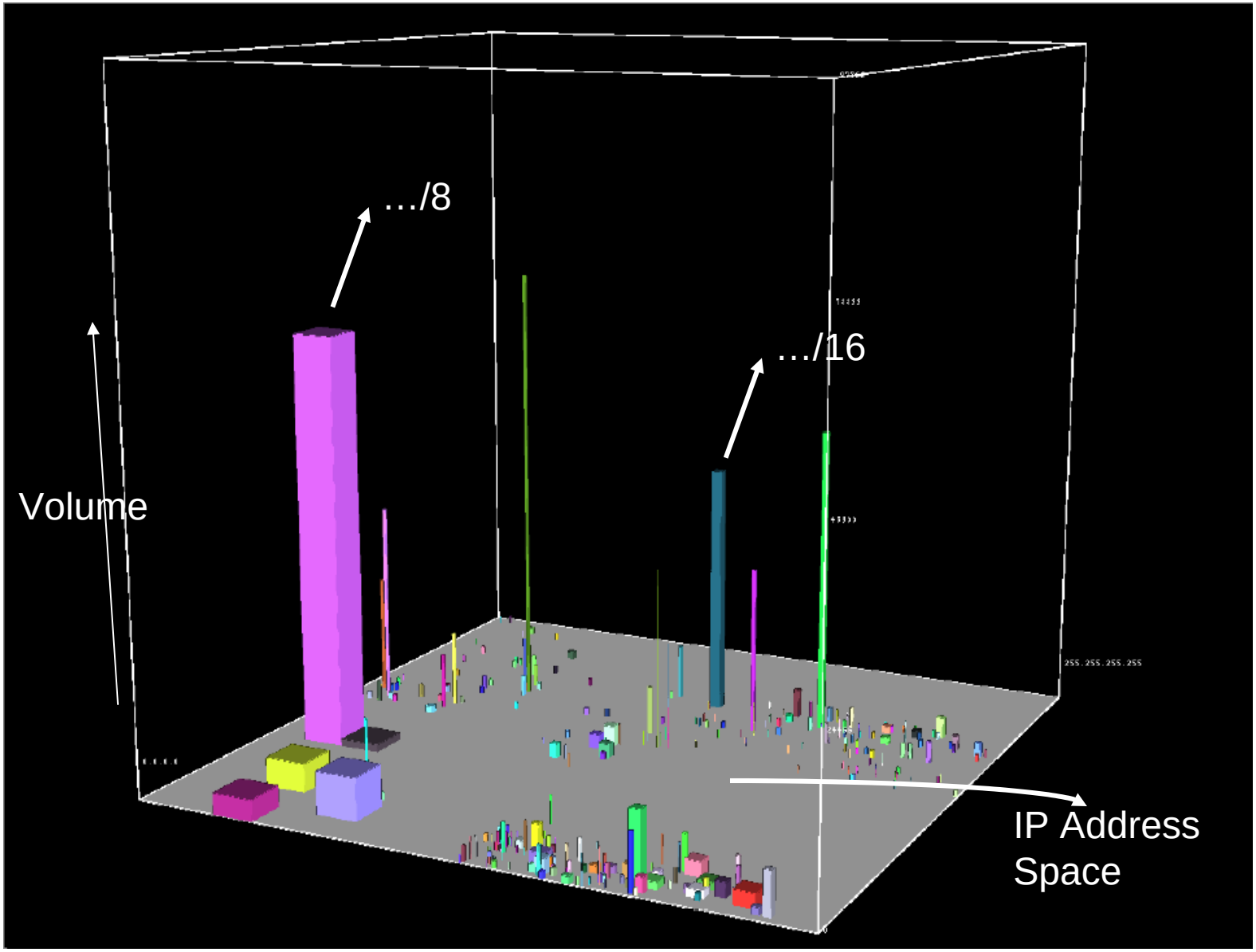
The Basic Quad-Tree



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

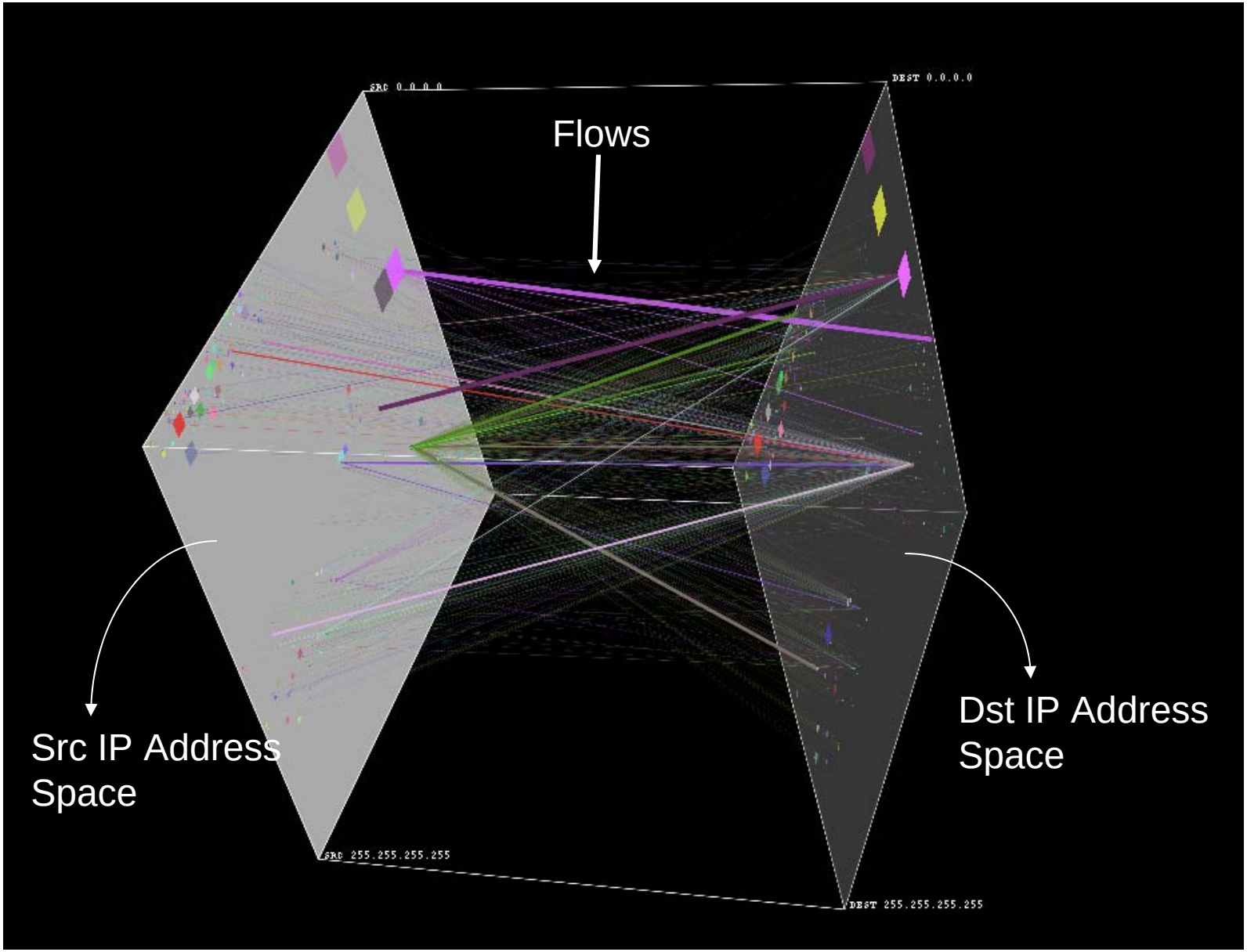
Traffic Volume by Src/Dst IP



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

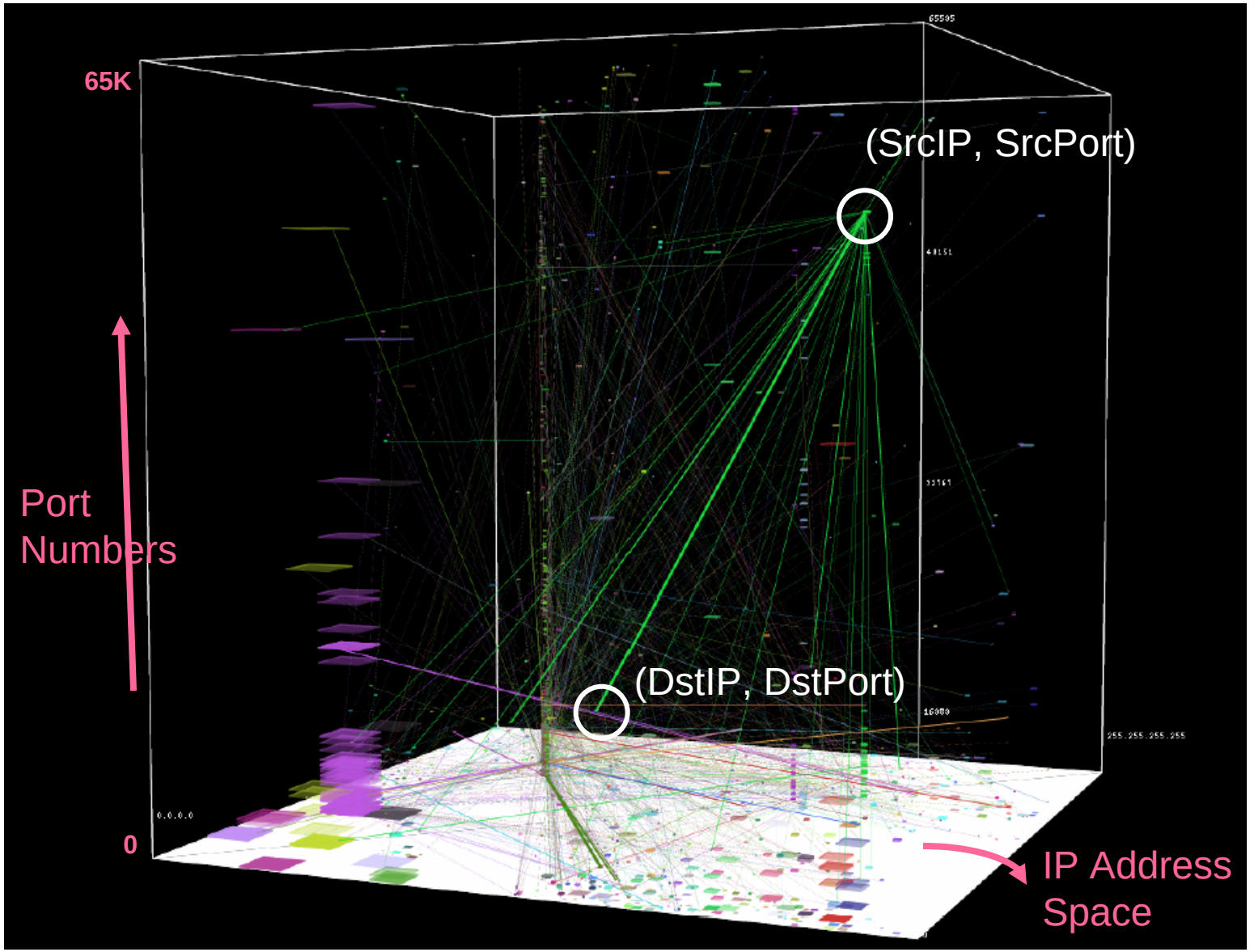
Traffic Flows by Aggregated Src/Dst IP



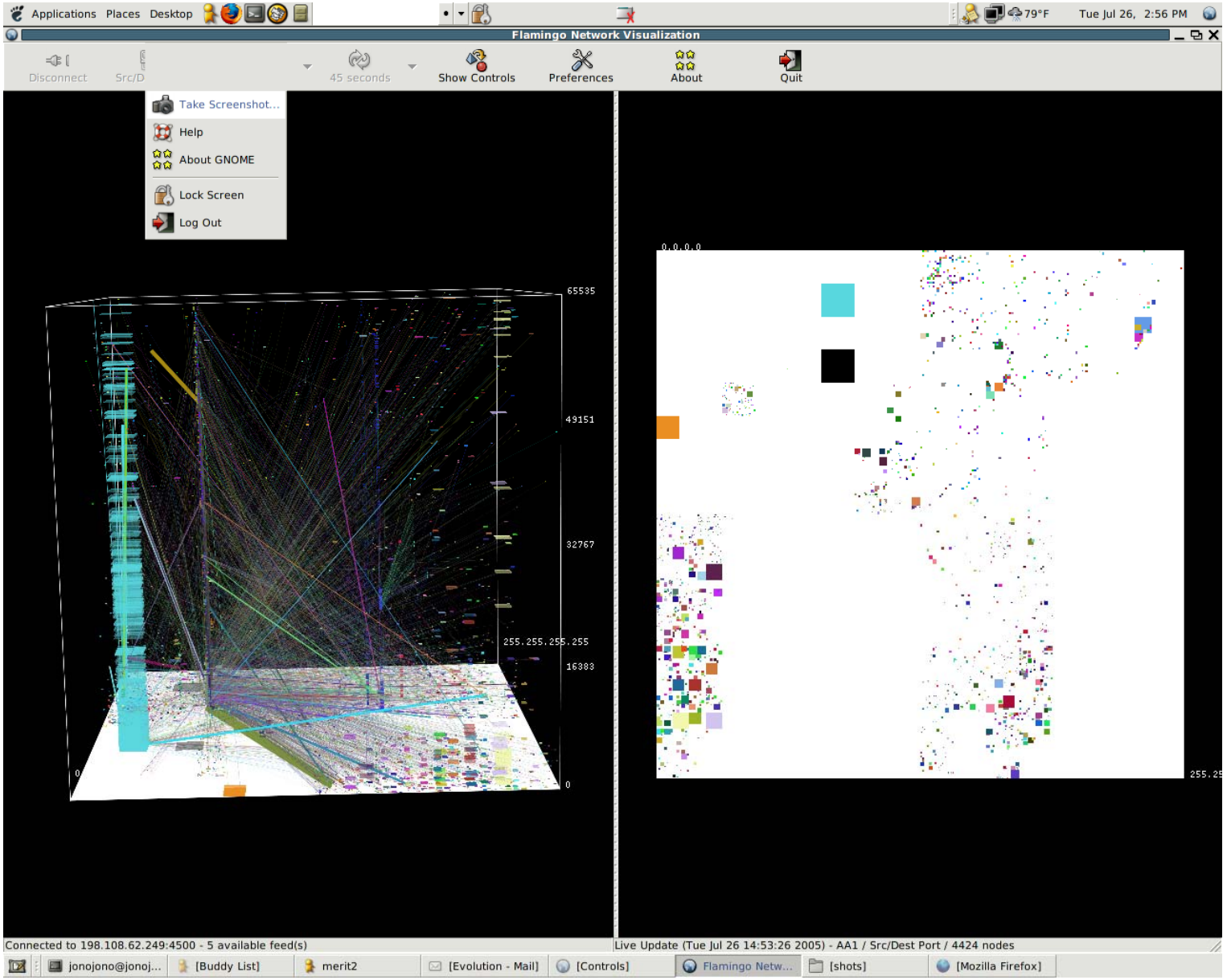
Traffic Flows by Src/Dst IP and Port

- 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
  - (srcIP, srcPort) >>>>>>>>>>> ((x1,y1), z1)
  - (dstIP, dstPort) >>>>>>>>>>> ((x2,y2), z2)
  - Line between these 2 points in 3D space represents a flow from (srcIP, srcPort) to (dstIP, 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

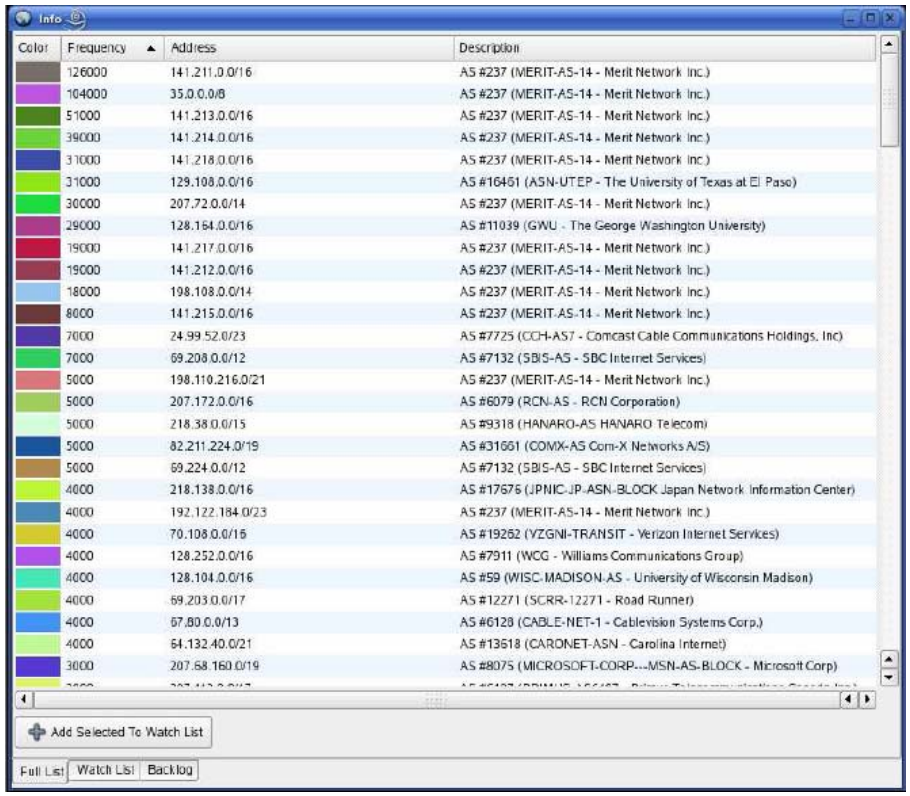


Flamingo Visualization Tool

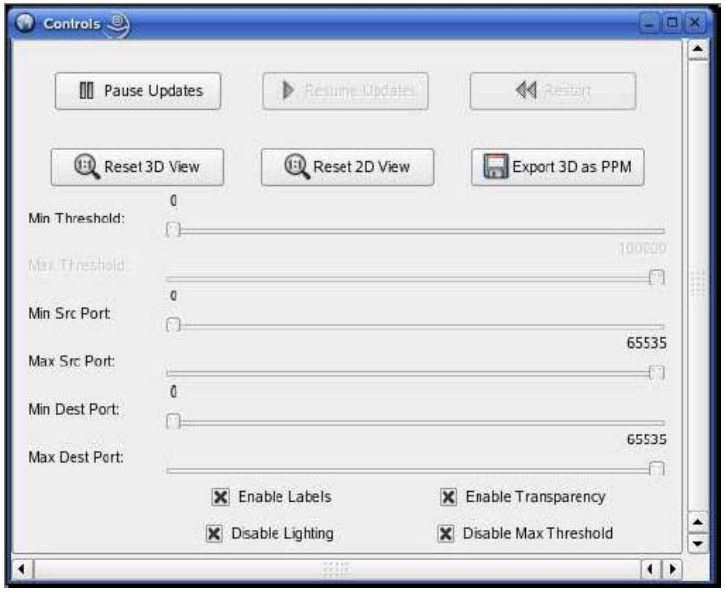


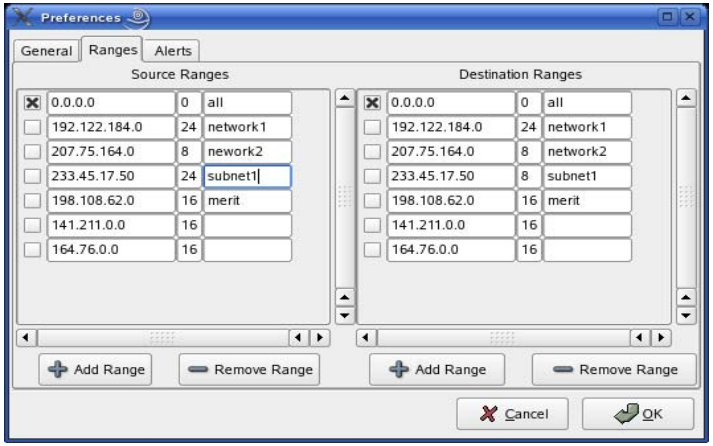
Flamingo Controls

Text Representation of Visualized Information



Slider Bar Controls



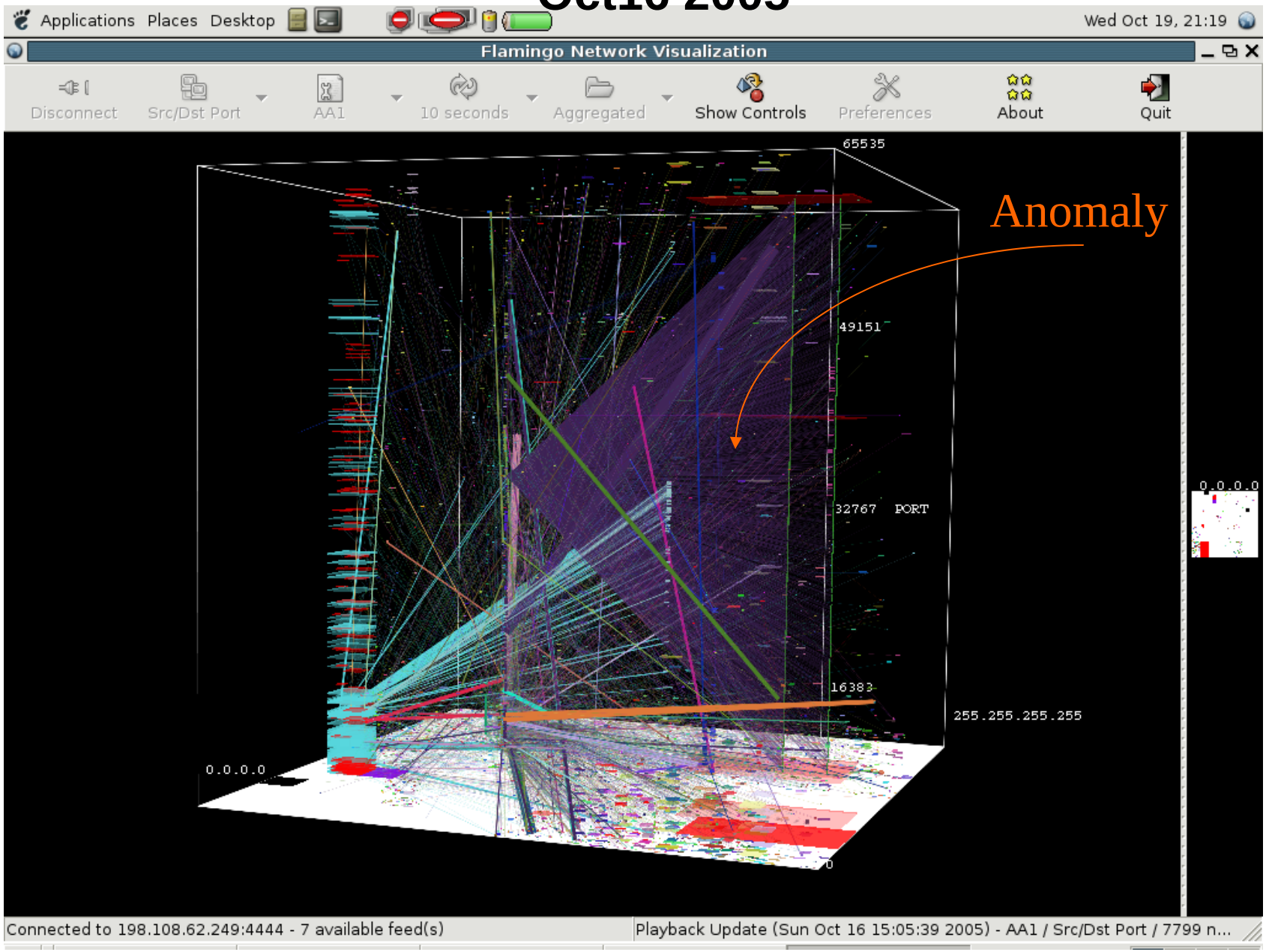


Address Range Configuration

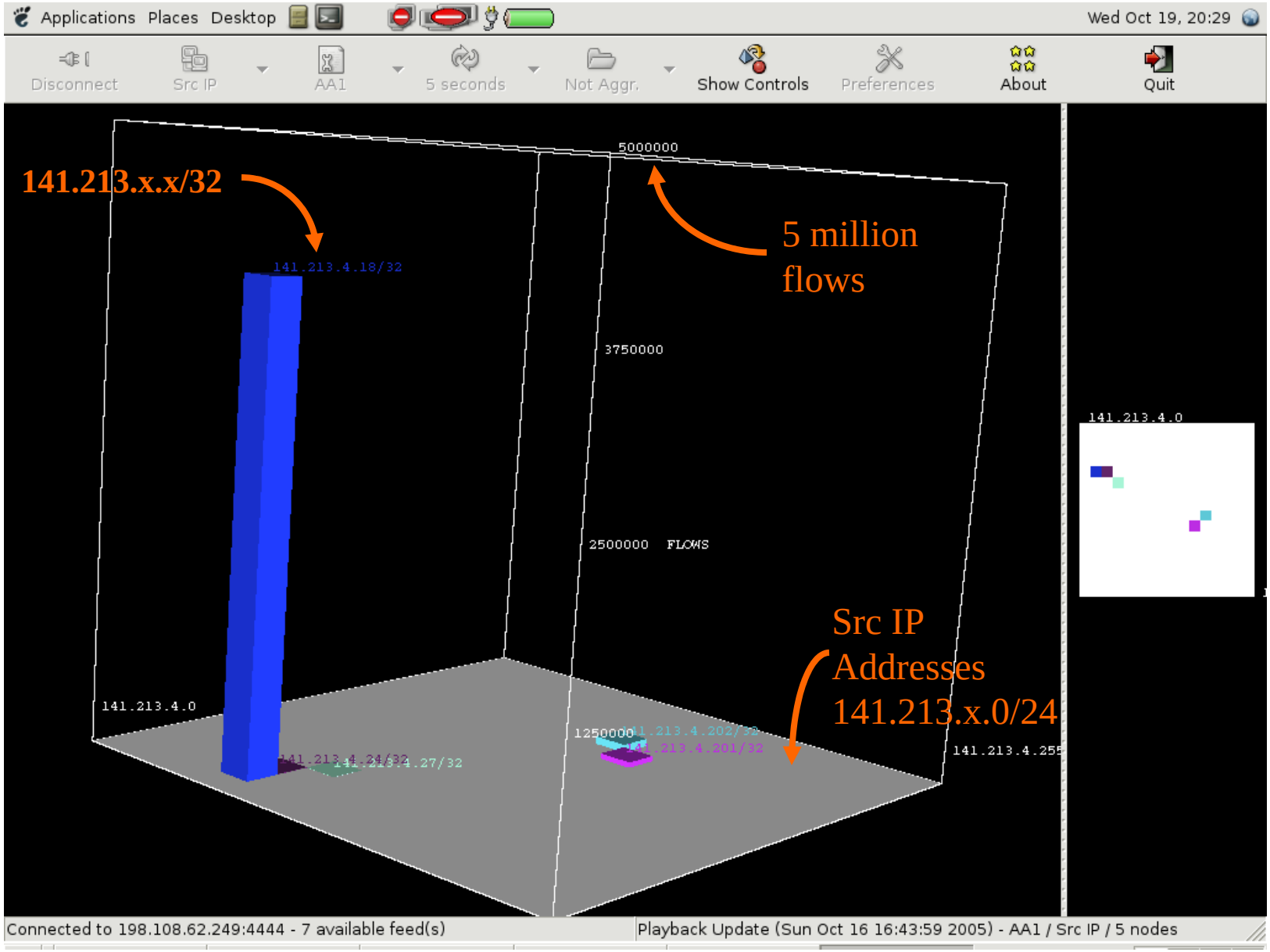
Case Study: Traffic Anomaly Sunday- Oct 16, 2005

- 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

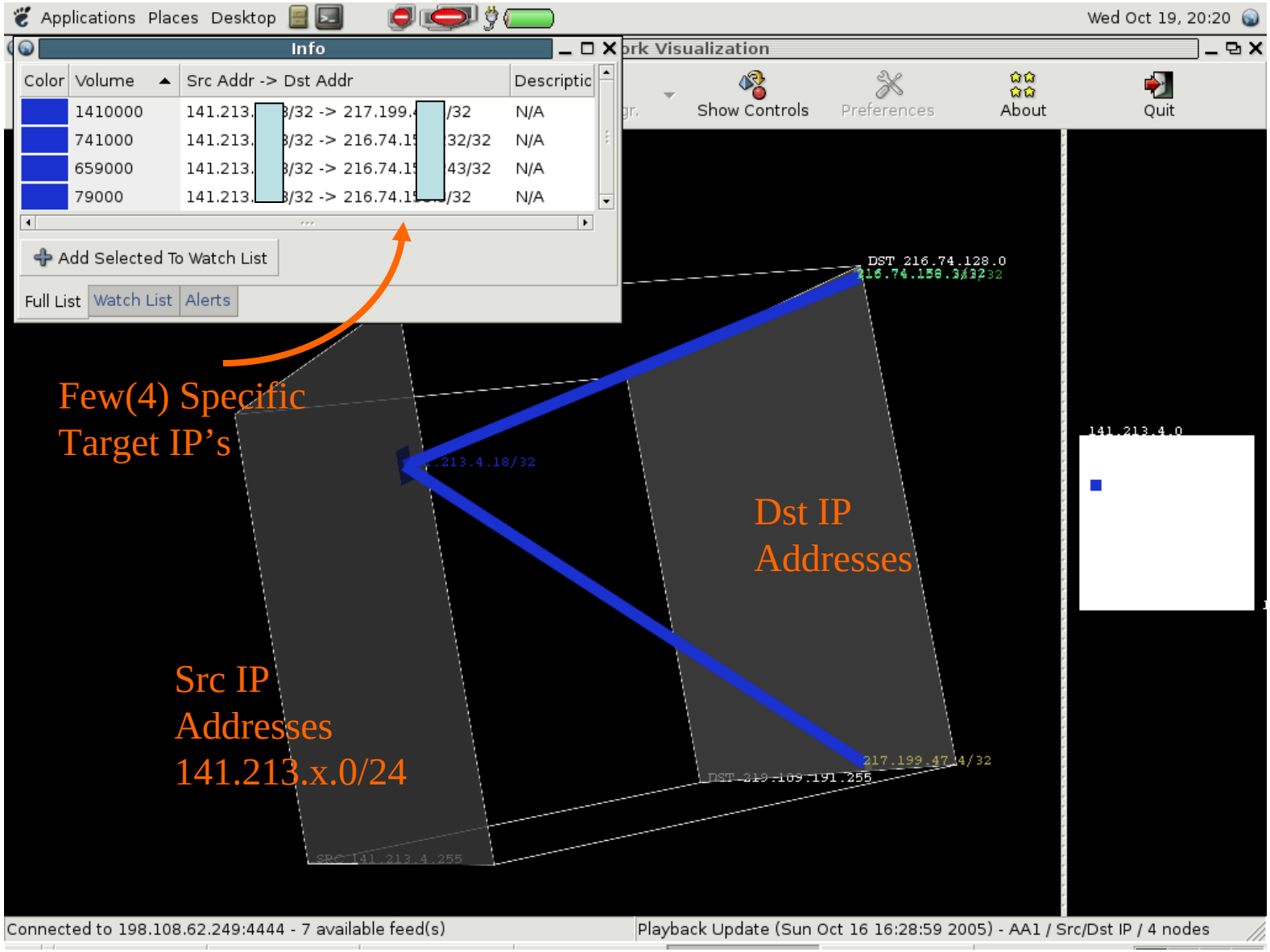
Overall Traffic Pattern at Primary Router Sunday, Oct16 2005



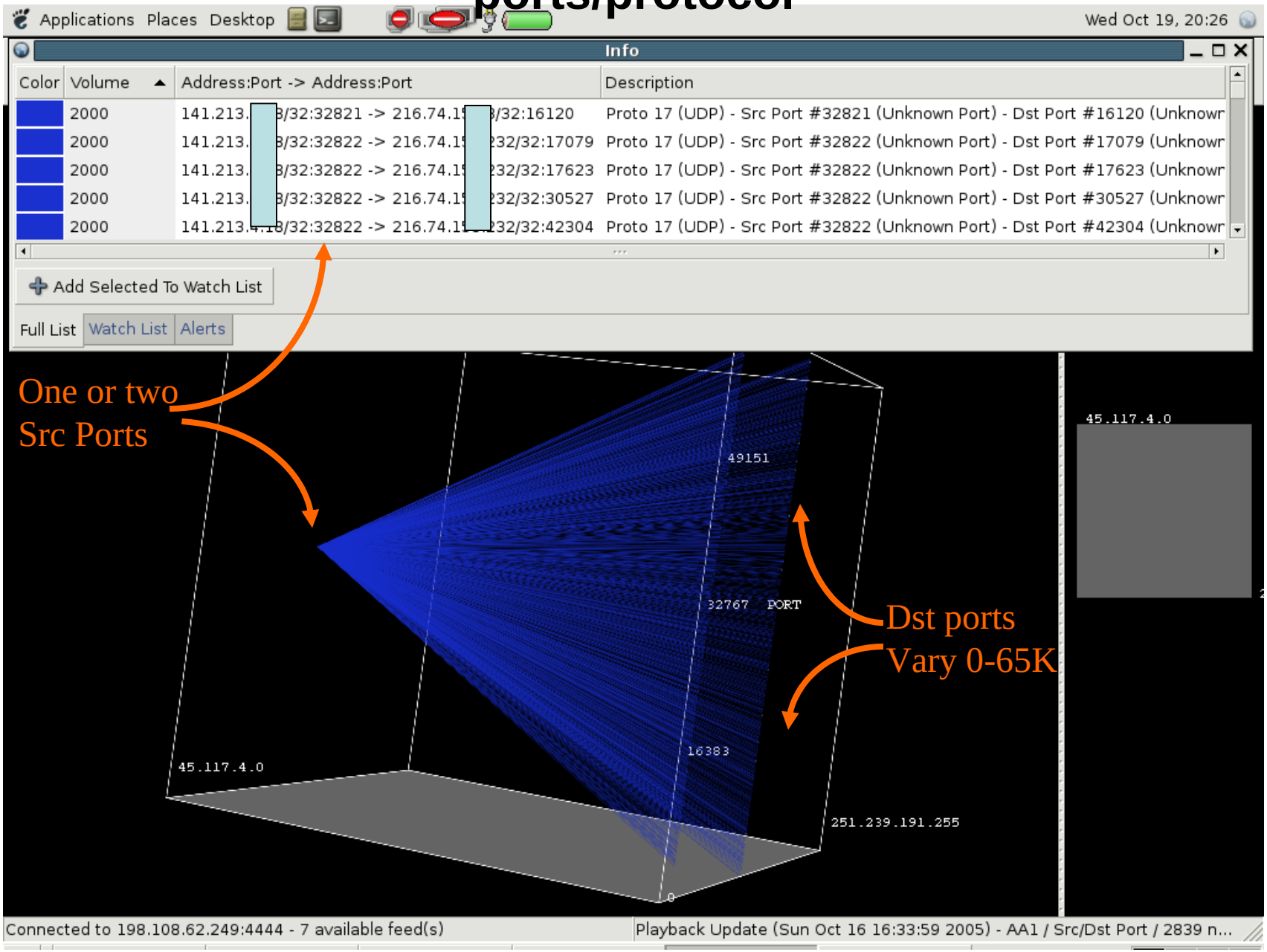
Traffic Volume sourced from /24 subnet by individual hosts



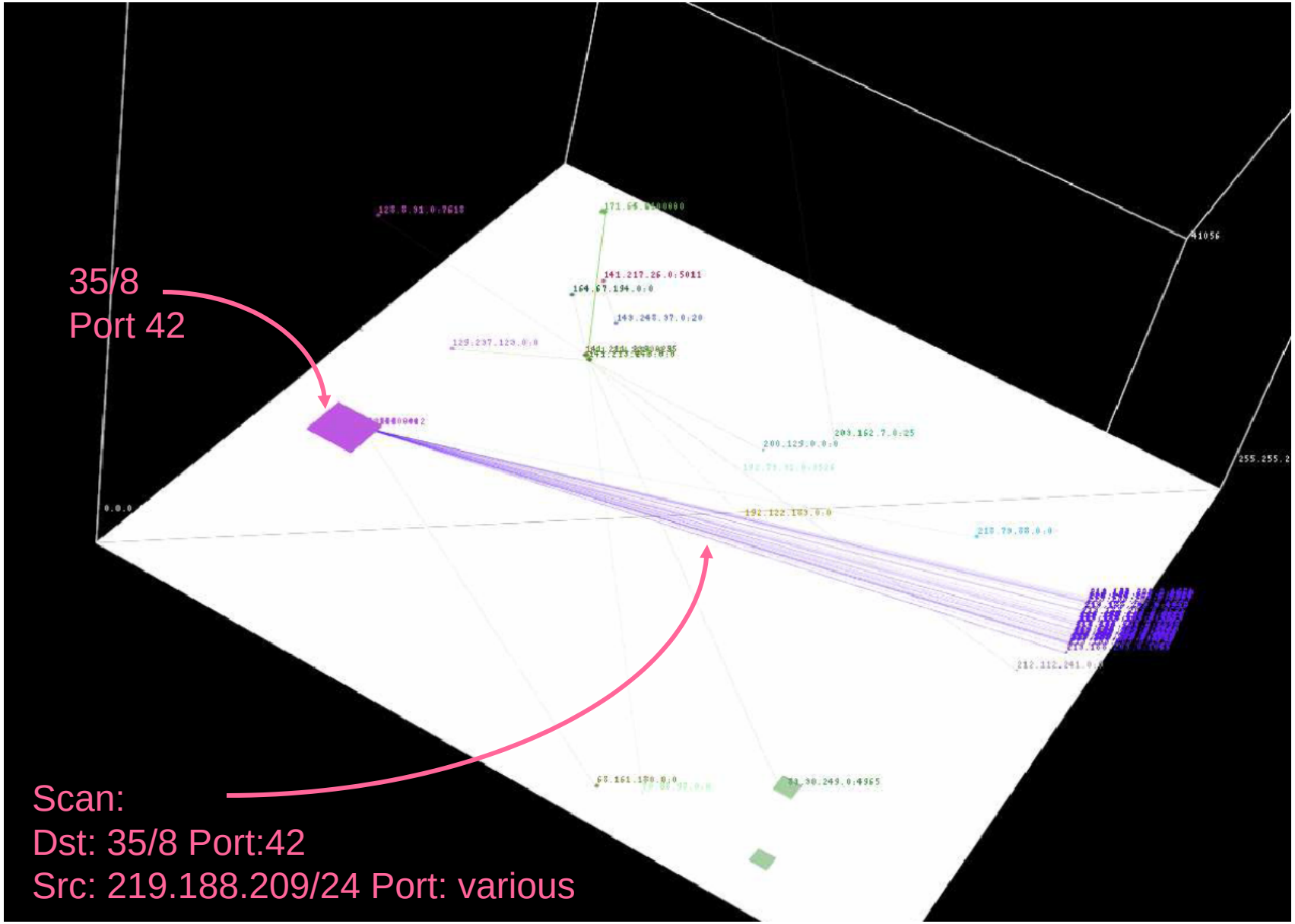
Distribution of Target IP Addresses



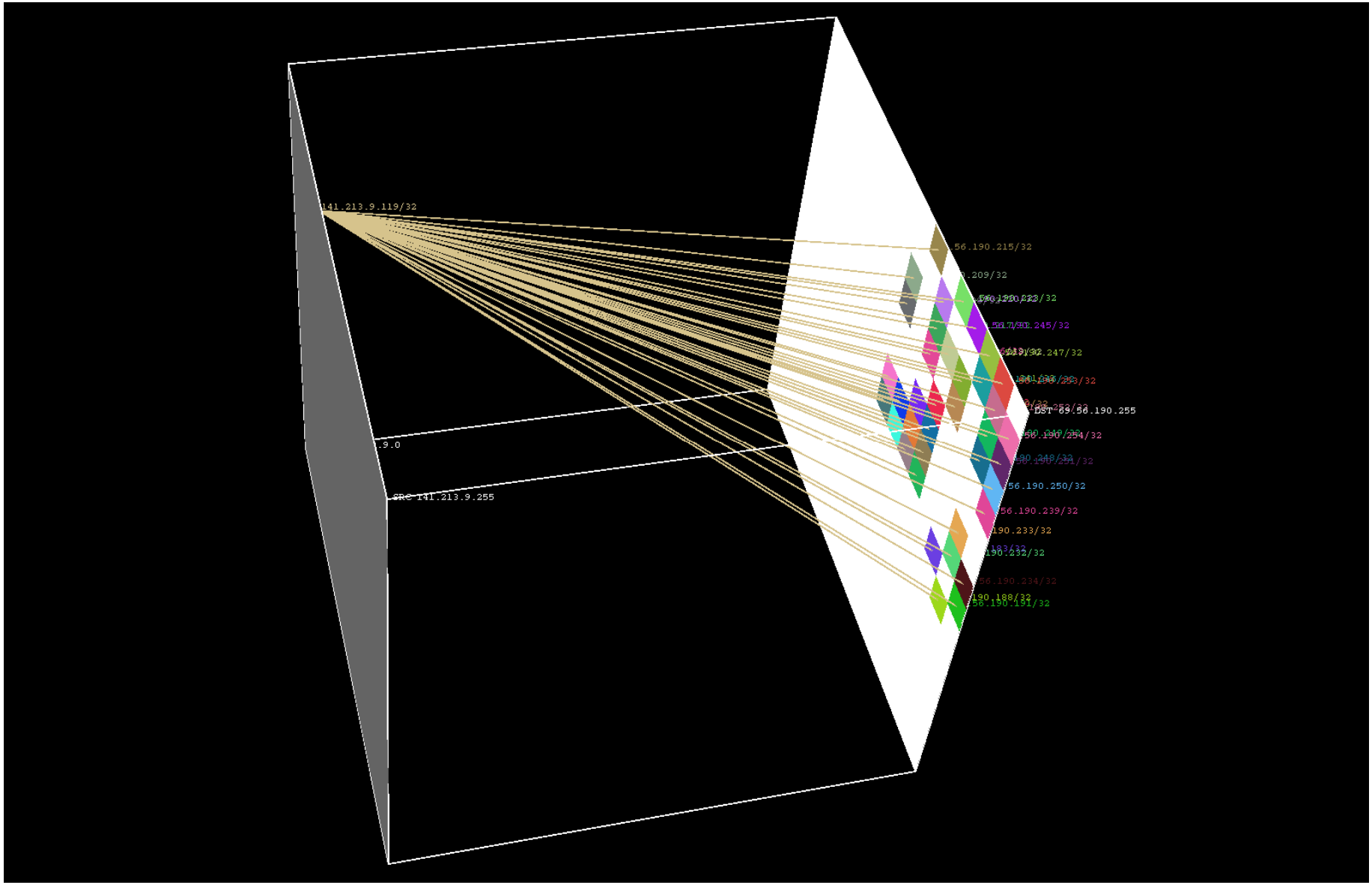
Distribution of flows in terms of src/dst ports/protocol



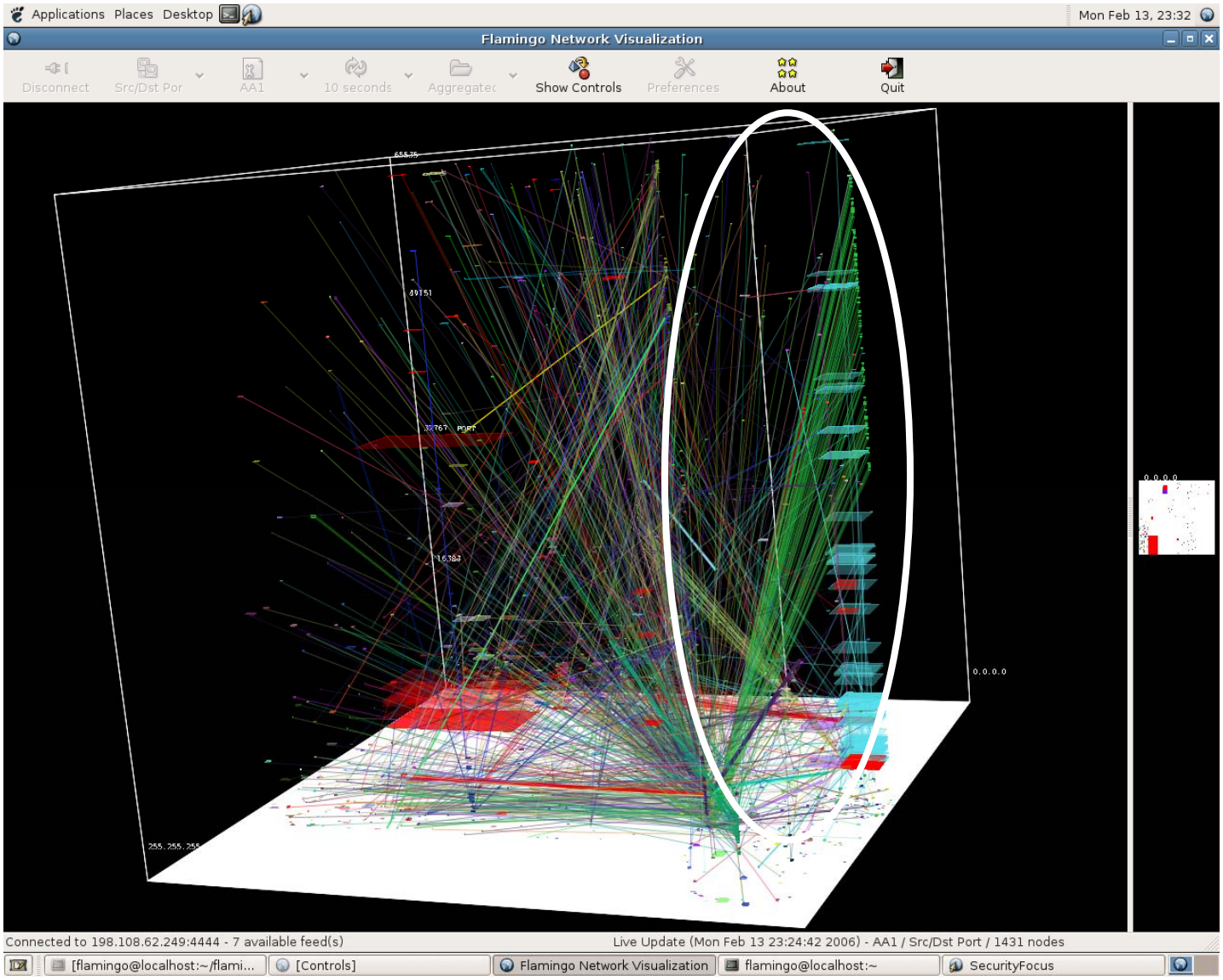
Case Study: Worm Traffic/Port 42 Scans



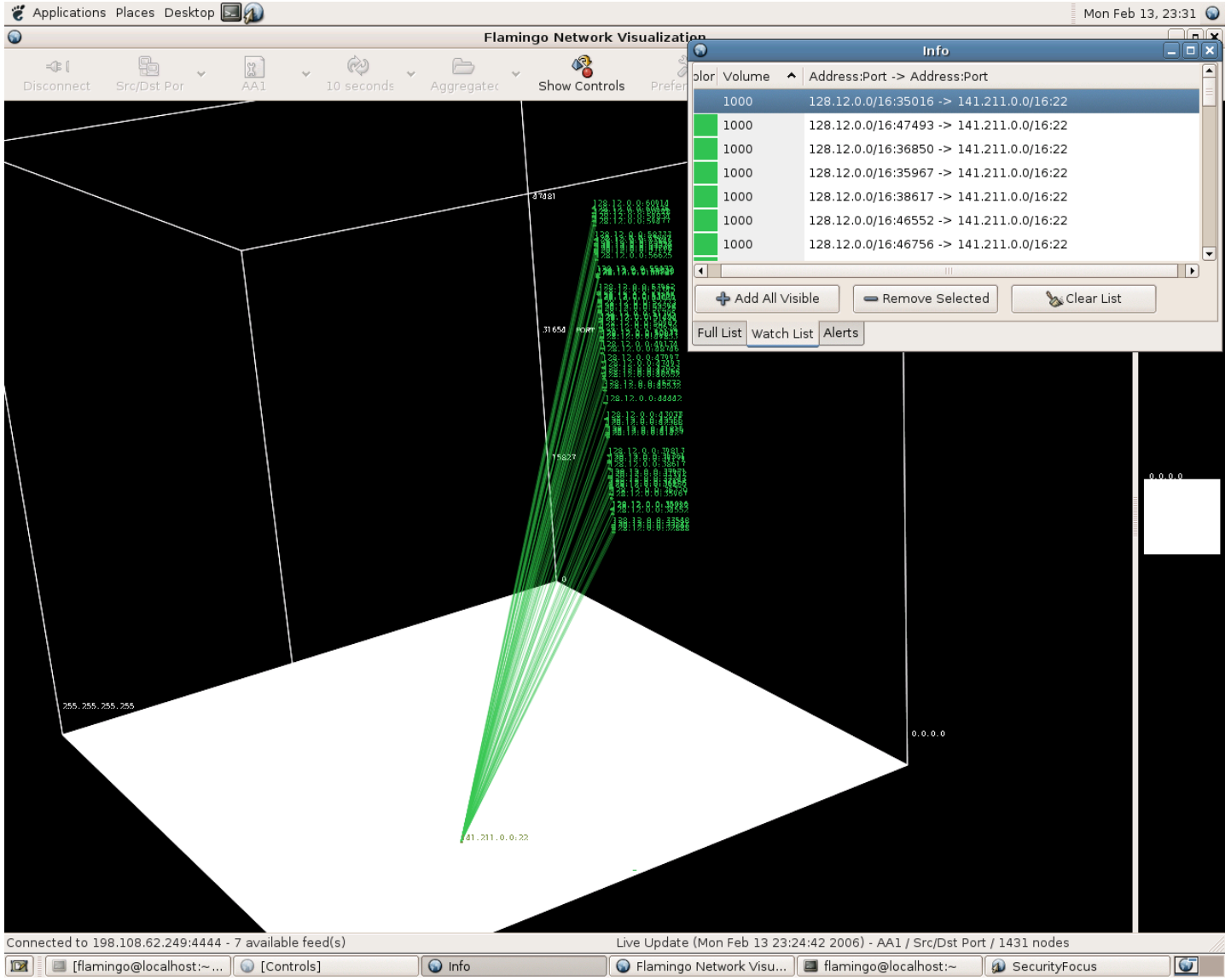
Case Study: /24 Network Scan



ssh scans

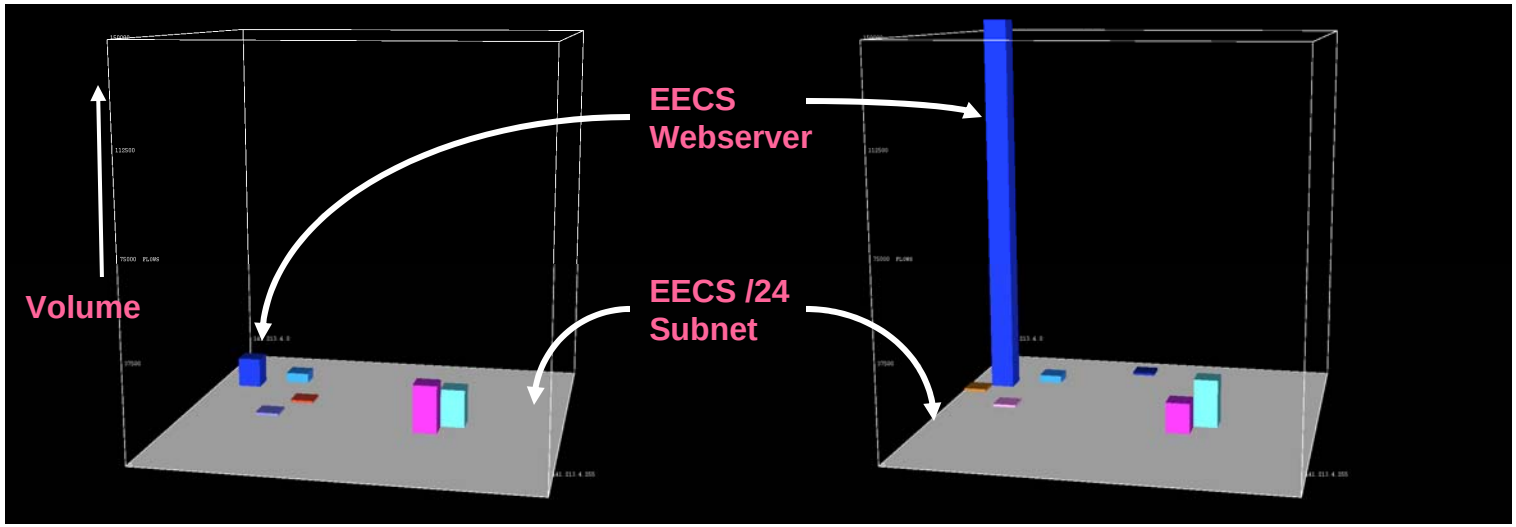

Applications Places Desktop
Mon Feb 13, 23:32
Flamingo Network Visualization
Disconnect
Src/Dst Por
AA1
10 seconds
Aggregatec
Show Controls
Preferences
About
Quit
0.0.0.0
255.255.255
Connected to 198.108.62.249:4444 - 7 available feed(s)
Live Update (Mon Feb 13 23:24:42 2006) - AA1 / Src/Dst Port / 1431 nodes
[flamingo@localhost:~/flami...
[Controls]
Flamingo Network Visualization
flamingo@localhost:~
SecurityFocus

ssh scan

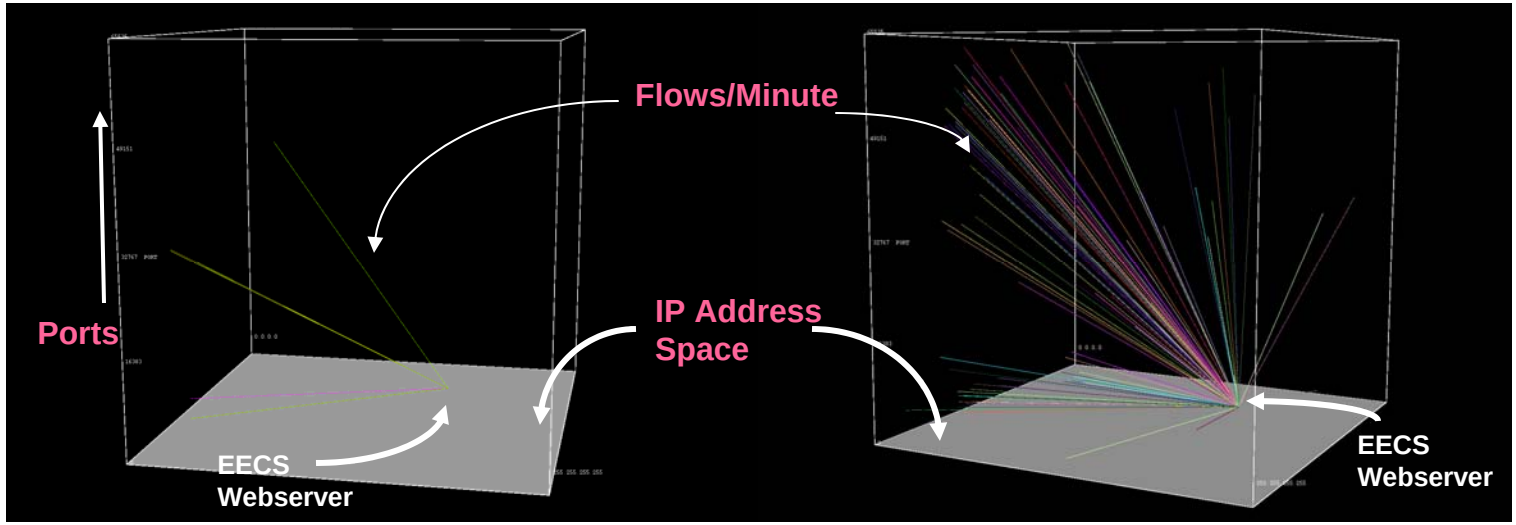

Applications Places Desktop
Mon Feb 13, 23:31
Flamingo Network Visualization
Disconnect
Src/Dst Por
AA1
10 seconds
Aggregatec
Show Controls
Info
olor Volume
Address:Port -> Address:Port
1000 128.12.0.0/16:35016 -> 141.211.0.0/16:22
1000 128.12.0.0/16:47493 -> 141.211.0.0/16:22
1000 128.12.0.0/16:36850 -> 141.211.0.0/16:22
1000 128.12.0.0/16:35967 -> 141.211.0.0/16:22
1000 128.12.0.0/16:38617 -> 141.211.0.0/16:22
1000 128.12.0.0/16:46552 -> 141.211.0.0/16:22
1000 128.12.0.0/16:46756 -> 141.211.0.0/16:22
Add All Visible
Remove Selected
Clear List
Full List
Watch List
Alerts
255.255.255.255
0.0.0.0
0.0.0.0
Connected to 198.108.62.249:4444 - 7 available feed(s)
Live Update (Mon Feb 13 23:24:42 2006) - AA1 / Src/Dst Port / 1431 nodes
[flamingo@localhost:~...
[Controls]
Info
Flamingo Network Visu...
flamingo@localhost:~
SecurityFocus

Case Study: Slashdot Event Oct 31, 2004

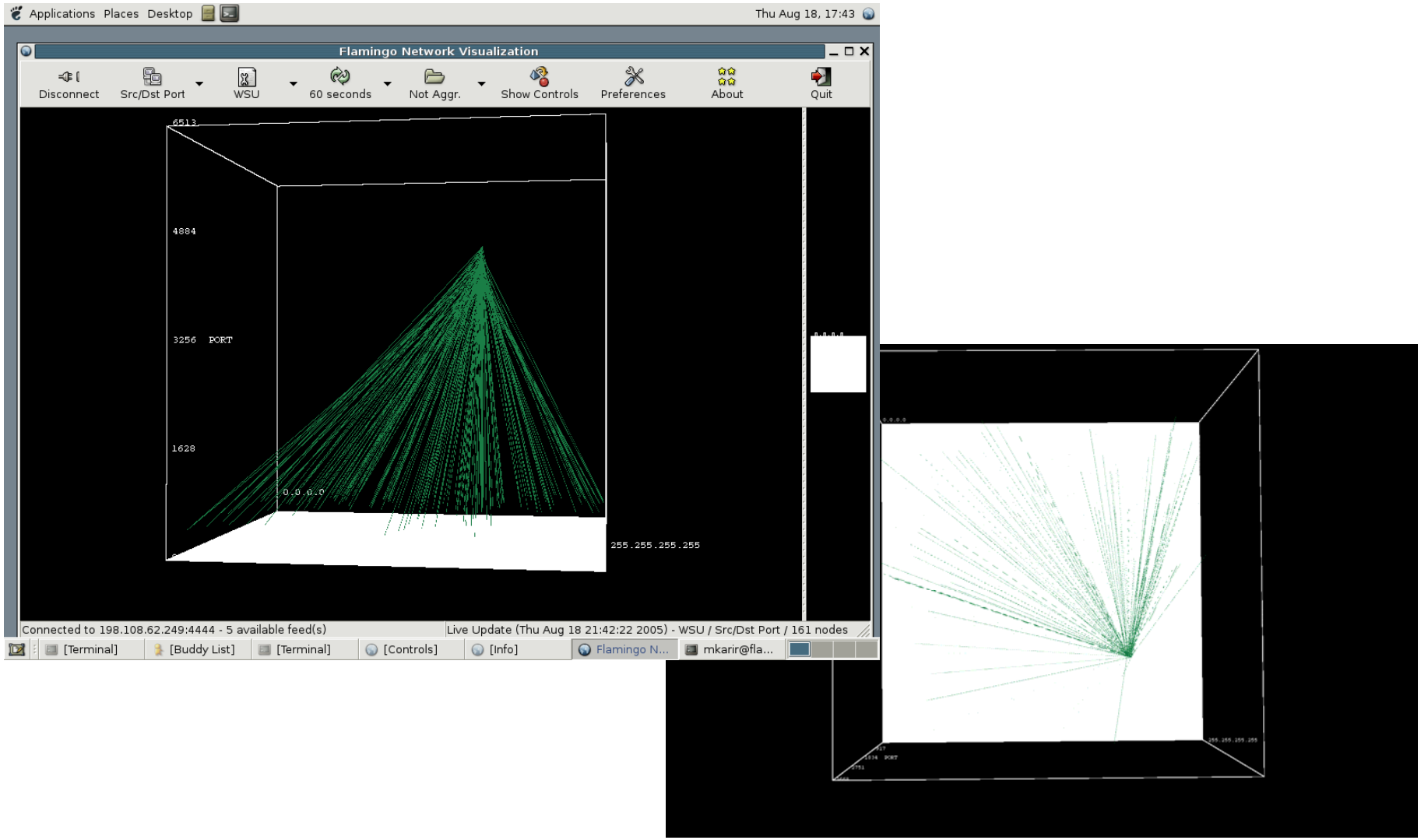
Traffic Volume



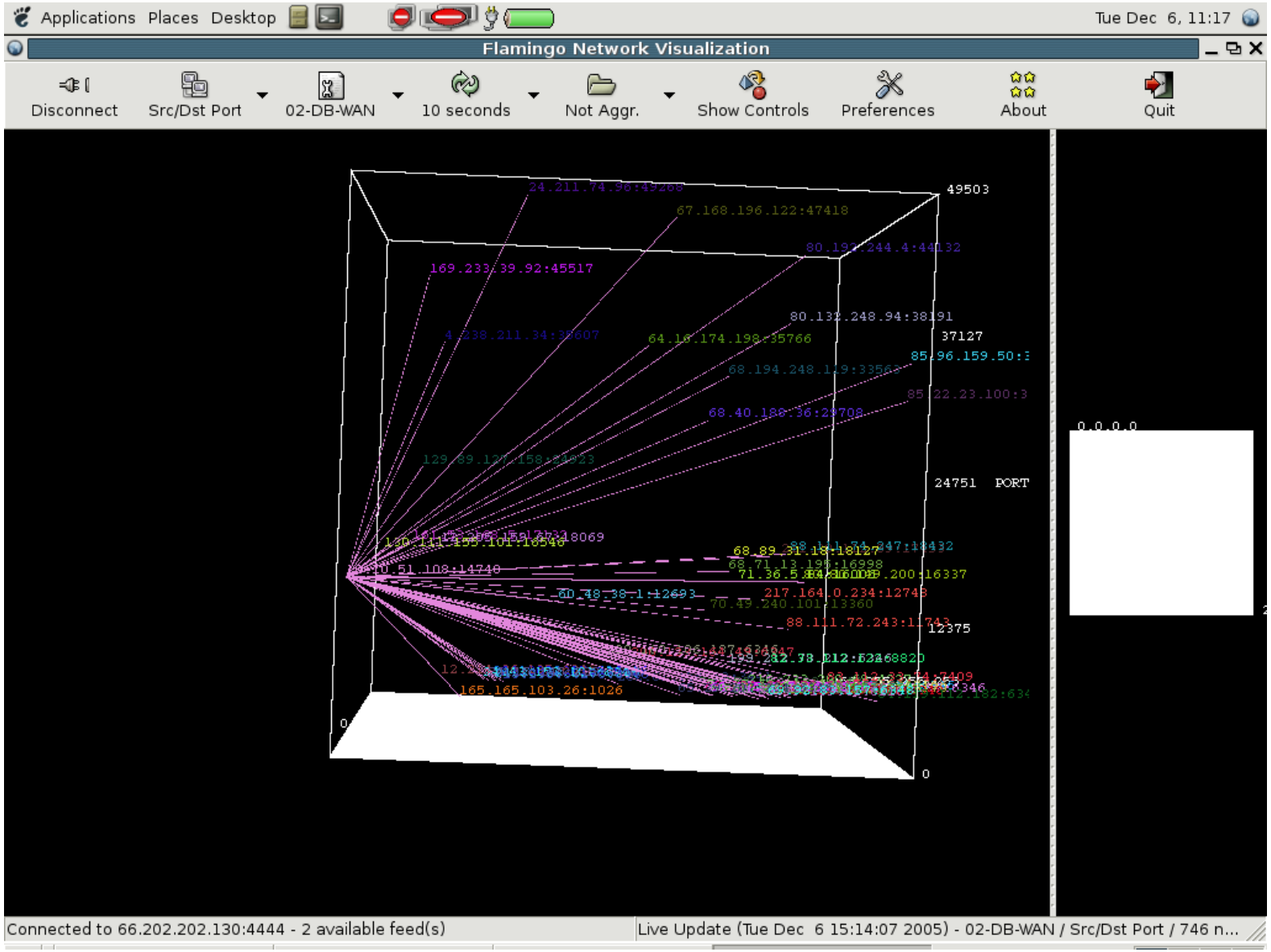
Flow Volume



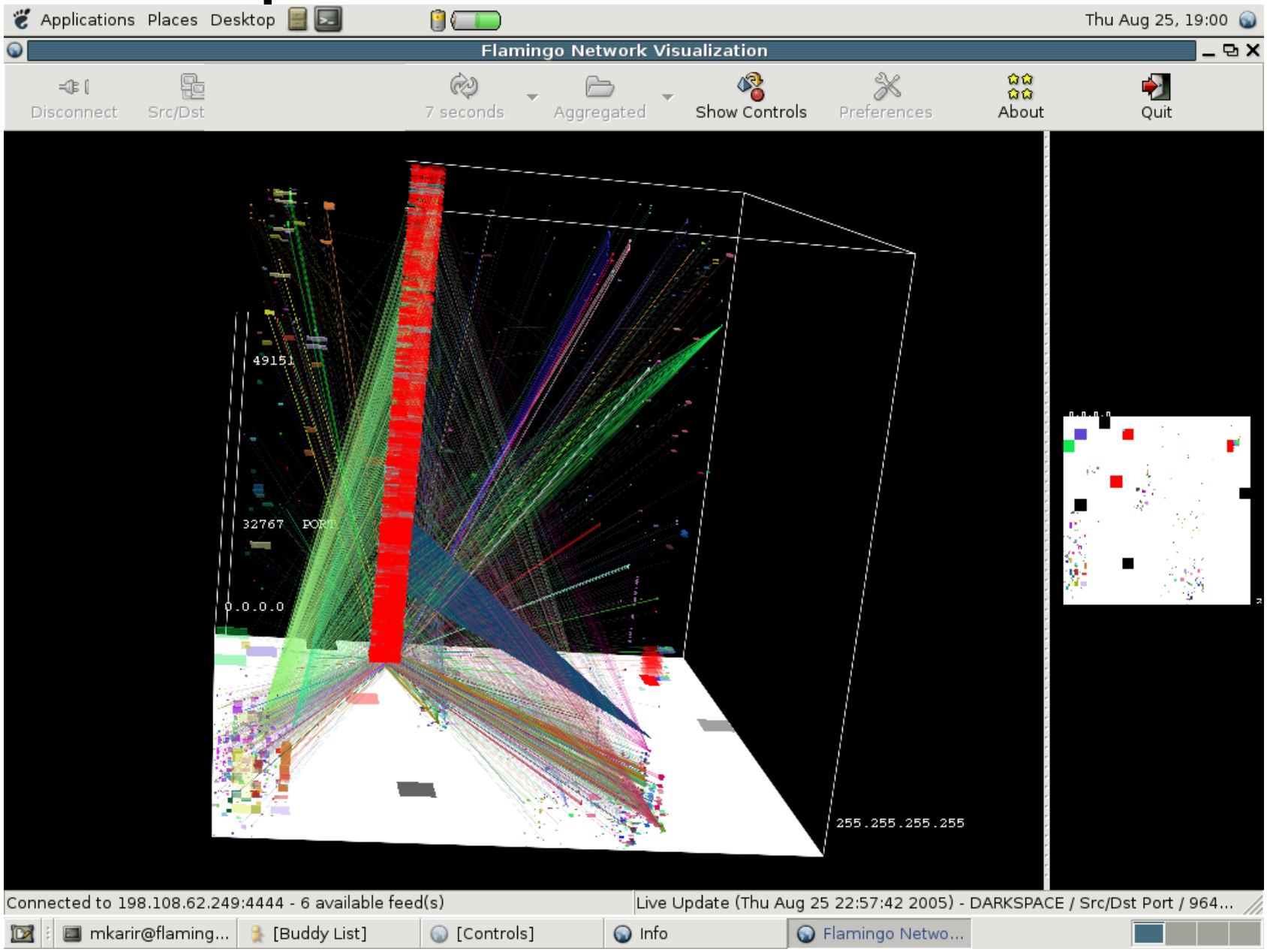
Zotob Worm Infection



P2P Traffic


Applications Places Desktop
Tue Dec 6, 11:17
Flamingo Network Visualization
Disconnect
Src/Dst Port
02-DB-WAN
10 seconds
Not Aggr.
Show Controls
Preferences
About
Quit
49503
24.211.74.96:49268
67.168.196.122:47418
169.233.39.92:45517
80.132.248.94:38191
37127
64.16.174.198:35766
68.194.248.149:33563
68.40.180.36:29708
24751 PORT
68.89.31.18:18127
68.71.13.195:16998
71.36.5.84:16337
217.164.0.234:12743
70.49.240.101:13360
88.111.72.243:11743
12375
165.165.103.26:1026
0.0.0.0
0
0
Connected to 66.202.202.130:4444 - 2 available feed(s)
Live Update (Tue Dec 6 15:14:07 2005) - 02-DB-WAN / Src/Dst Port / 746 n...

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