

# Extracting Information from Network Data

Manish Karir, Jon Oberheide  
Merit Network, Inc.

Dionysus Blazakis  
University of Maryland

# Outline

- The Problem
- BGP/Routing Information
  - BGP-Inspect – Information Extraction from BGP Update messages
  - VAST – Internet AS topology Visualization
- Netflow/Traffic Information
  - Flamingo – Internet Traffic Exploration
- Conclusions

# The Problem

- Large amounts of data are now, or soon will be available:
  - Route Views, RIPE Archives, PREDICT, etc
- The problem is no longer access to raw data but how to extract useful information from the raw data
- Need tools that can:
  - Scale to large input datasets
  - Provide useful data summarizations
  - Are easy to use
  - Provide useful information
- BGP-Inspect, VAST, Flamingo are tools that we have implemented that attempt to address this problem

# BGP-Inspect: Why and What

- Analyzing MRT Data:
  - Large volumes of data ~RV-66G compressed
  - Extracting useful information requires writing custom parsers even for basic information
  - Lots and lots of redundancy
- Approach:
  - Preprocess Route Views data
  - Remove redundancy as much as possible
  - Use data compression to the extent possible
  - Build efficient indices to help queries
  - Pre-compute and store commonly used statistics at data load time not at query time
  - Build easy to use interface

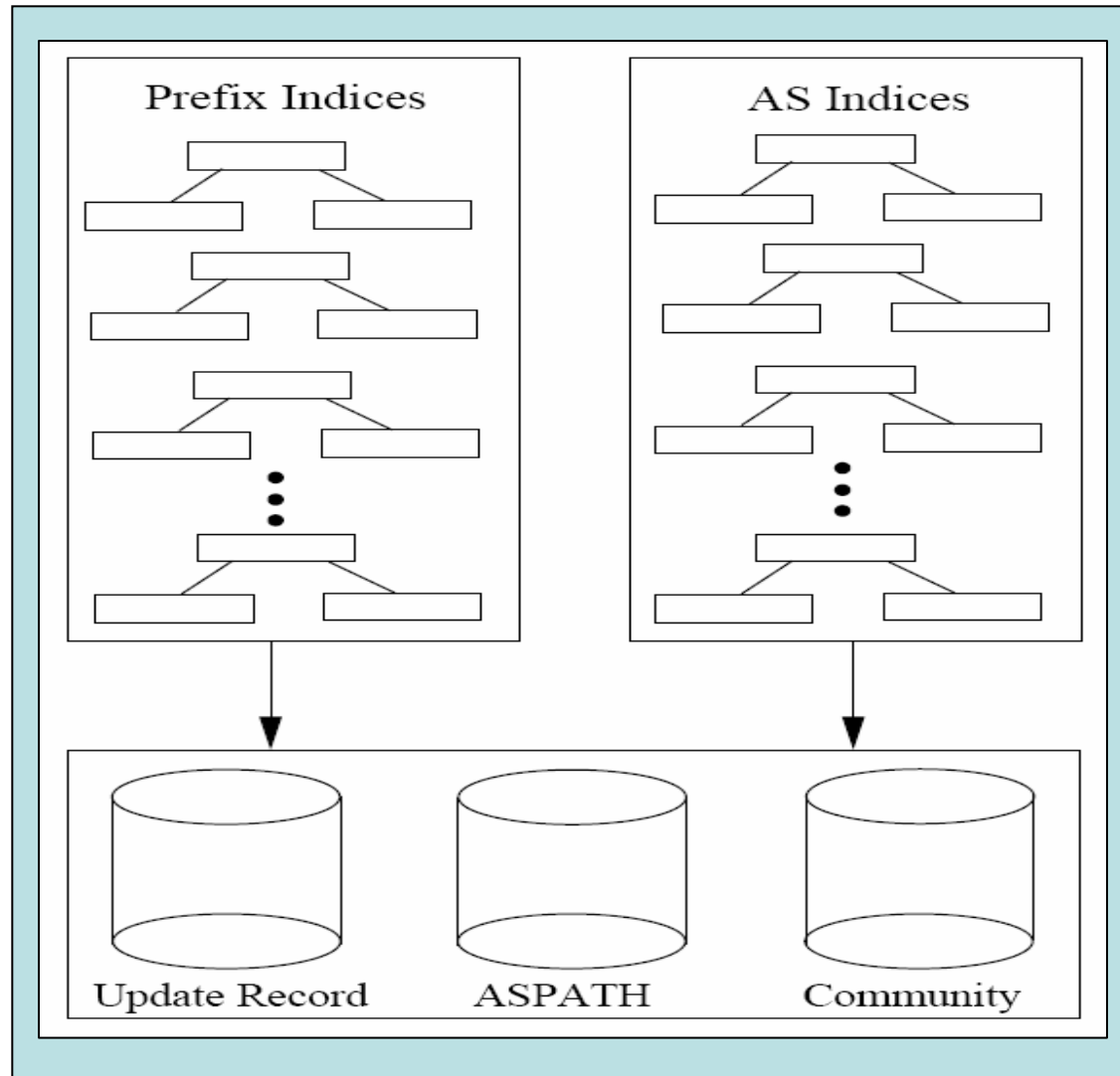
# BGPdb vs. BGP-Inspect

- BGPdb is the core of the BGP-Inspect system
- BGPdb represents the pre-processed database, which is queried by the BGP-Inspect interface

# BGPdb – Techniques and Algorithms

- Removing redundancy from BGP datasets
  - ASPATH, COMMUNITY, UPDATE Msgs are repeated over and over, only time changes
- Compressed-Chunked Files
  - Compromise between size and usability
- B+ Tree indices
  - Indexing based on time, this enables fast time-range queries
- Caching while processing input datasets
  - Messages are repetitive, so keep cache of previous processing for speedup

# BGPdb – System Architecture



# BGP-Inspect Interface

The screenshot shows a Mozilla browser window with the address bar containing `http://bgpinspect.merit.edu/index.php`. The page title is "BGP-Inspect - bgpinspect.merit.edu - Mozilla". The main content area has a blue header with the text "BGP-Inspect-Routeviews" in yellow. To the right of the header, it says "First DB Update: Aug 1, 2005, 12:01 am +0000" and "Last DB Update: June 28, 2006, 12:11 am +0000". Below the header is a navigation menu with buttons for "Home", "Reports", "Documentation", "FAQ", and "About".

The main content is divided into two sections:

- Global Summary Queries:** (Please select a peer, query type and duration)
  - Peer:** A dropdown menu with options: 12.0.1.63 - ATT, 4.68.0.243 - Level 3, 66.185.128.1 - AOL, 144.228.241.81 - Sprint, 208.51.134.253 - Global X.
  - Query Type:** A dropdown menu with options: Most Active ASes, Most Active Prefixes, Prefixes Most Announced, Prefixes Most Withdrawn, Prefixes with Most AS Changes.
  - Duration:** A dropdown menu with options: Last 1 Days, Last 2 Days, Last 3 Days, Last 7 Days, Last 10 Days.
  - A "Submit Query" button is located below the dropdowns.
- Raw Data Analysis:** (Please select a peer, query type, AS/prefix, and time range)
  - Peer:** A dropdown menu with options: 12.0.1.63 - ATT, 4.68.0.243 - Level 3, 66.185.128.1 - AOL, 144.228.241.81 - Sprint, 208.51.134.253 - Global X.
  - Query Type:** A dropdown menu with options: AS, Prefix-Exact, Prefix-More Specific.
  - Query:** A text input field with the placeholder "(ASN or a.b.c.d/len)".
  - Start Date:** A date selector showing 2006 Jun 21 00 : 00.
  - End Date:** A date selector showing 2006 Jun 28 00 : 00.
  - A "Submit Query" button is located below the date selectors.

At the bottom of the page, there is a footer with the text "Copyright © Merit Network Inc." and "Copyright © University of Maryland". The browser's status bar at the very bottom shows various icons and the address bar.



# Global Queries – Most Active ASes

BGP-Inspect - bgpinspect.merit.edu - Mozilla

File Edit View Go Bookmarks Tools Window Help

http://bgpinspect.merit.edu/index.php

## BGP-Inspect-Routeviews

First DB Update: Aug 1, 2005,12:01 am +0000  
Last DB Update: June 28, 2006, 12:11 am +0000

Home Reports Documentation FAQ About

**Global Summary Queries:** (Please select a peer, query type and duration)

Peer: 12.0.1.63 - ATT  
4.68.0.243 - Level 3  
66.185.128.1 - AOL  
144.228.241.81 - Sprint  
208.51.134.253 - GlobalX

Query Type: Most Active ASes  
Most Active Prefixes  
Prefixes Most Announced  
Prefixes Most Withdrawn  
Prefixes with Most AS Changes

Duration: Last 1 Days  
Last 2 Days  
Last 3 Days  
Last 7 Days  
Last 10 Days

Submit Query

**Raw Data Analysis:** (Please select a peer, query type, AS/prefix, and time range)

Peer: 12.0.1.63 - ATT  
4.68.0.243 - Level 3  
66.185.128.1 - AOL  
144.228.241.81 - Sprint  
208.51.134.253 - GlobalX

Query Type: AS  
Prefix-Exact  
Prefix-More Specific

Query: (ASN or a.b.c.d/len)

Start Date: 2006 Jun 21 00 : 00  
End Date: 2006 Jun 28 00 : 00

Submit Query

Copyright © Merit Network Inc.  
Copyright © University of Maryland

BGP-Inspect - bgpinspect.merit.edu - Mozilla

File Edit View Go Bookmarks Tools Window Help

http://bgpinspect.merit.edu/query.php

## BGP-Inspect-Routeviews

First DB Update: Aug 1, 2005,12:01 am +0000  
Last DB Update: June 28, 2006, 12:11 am +0000

Home	Reports	Documentation	FAQ	About
Aggregate	12.0.1.63	4.68.0.243	66.185.128.1	144.228.241.81
				208.51.134.253

Peer: 12.0.1.63

**Most Active ASes, Last 7 Days**

Rank	AS Number	AS Name	Announcements
1	5075	ATTBMG AT&T BMGS	44325
2	14169	MEADCO-1 MEAD CORPORATION	31132
3	25543	FASONET-AS ONATEL/FasoNet's Autonomous System	18259
4	4134	CHINANET-BACKBONE No.31,Jin-rong Street	17781
5	17974	TELKOMNET-AS2-AP PT TELEKOMUNIKASI INDONESIA	14822
6	7018	ATTW AT&T WorldNet Services	12779
7	5803	DNIC DoD Network Information Center	11682
8	17557	PKTELECOM-AS-AP Pakistan Telecom	11063
9	9443	INTERNETPRIMUS-AS-AP Primus Telecommunications	8652
10	3475	DEPART-52 Department of the Navy	8380
11	17451	BIZNET-AS-AP BIZNET ISP	7854
12	8452	TEDATA TEDATA	6869
13	23918	CBB-BGP-IBARAKI Connexion By Boeing Ibaraki AS	6234
14	174	COGENT Cogent/PSI	5847
15	6126	CAI-7 Computer Associates International	5590
16	19169	Telconet	5381
17	18231	EXATT-AS-AP Exatt Technologies Private Ltd.	5326
18	702	AS702 MCI EMEA - Commercial IP service provider in Europe	5239
19	11139	CWD-18 Cable & Wireless Dominica	4604
20	31200	NTK Novotelecom Ltd.	4549

Done

# Raw Data Analysis – AS Query

BGP-Inspect - bgpinspect.merit.edu - Mozilla

File Edit View Go Bookmarks Tools Window Help

http://bgpinspect.merit.edu/index.php

## BGP-Inspect-Routeviews

First DB Update: Aug 1, 2005, 12:01 am +0000  
Last DB Update: June 28, 2006, 12:11 am +0000

Home Reports Documentation FAQ About

**Global Summary Queries:** (Please select a peer, query type and duration)

Peer: 12.0.1.63 - ATT, 4.68.0.243 - Level 3, 66.185.128.1 - ADL, 144.228.241.81 - Sprint, 208.51.134.253 - Global X

Query Type: Most Active ASes, Most Active Prefixes, Prefixes Most Announced, Prefixes Most Withdrawn, Prefixes with Most AS Changes

Duration: Last 1 Days, Last 2 Days, Last 3 Days, Last 7 Days, Last 10 Days

Submit Query

---

**Raw Data Analysis:** (Please select a peer, query type, AS/prefix, and time range)

Peer: 12.0.1.63 - ATT, 4.68.0.243 - Level 3, 66.185.128.1 - ADL, 144.228.241.81 - Sprint, 208.51.134.253 - Global X

Query Type: AS, Prefix-Exact, Prefix-More-Specific

Query: (ASN or a.b.c.d/len) 145

Start Date: 2006 Jun 21 00:00

End Date: 2006 Jun 28 00:00

Submit Query

Copyright © Merit Network Inc.  
Copyright © University of Maryland

BGP-Inspect - bgpinspect.merit.edu - Mozilla

File Edit View Go Bookmarks Tools Window Help

http://bgpinspect.merit.edu/query.php

## BGP-Inspect-Routeviews

First DB Update: Aug 1, 2005, 12:01 am +0000  
Last DB Update: June 28, 2006, 12:11 am +0000

Home Reports Documentation FAQ About

Aggregate 12.0.1.63 4.68.0.243 66.185.128.1 144.228.241.81 208.51.134.253

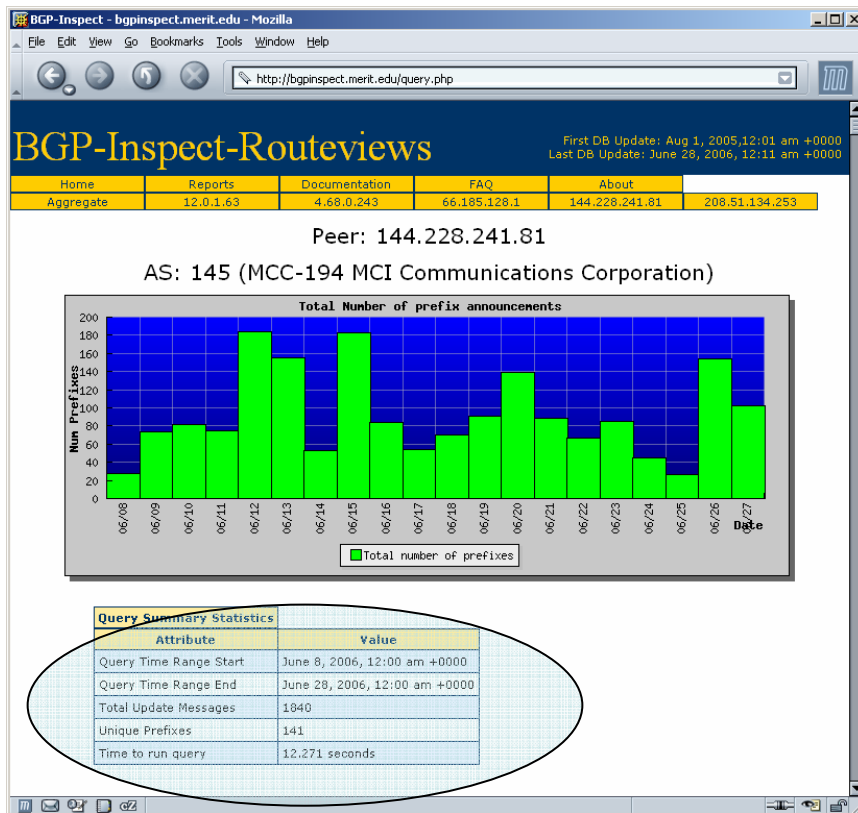
**Peer: Aggregate**

**AS: 145 (MCC-194 MCI Communications Corporation)**

**Total Number of prefix announcements**

Query Summary Statistics	Value
Query Time Range Start	June 8, 2006, 12:00 am +0000
Query Time Range End	June 28, 2006, 12:00 am +0000
Total Number of Prefixes (All Peers)	16119
Total Prefixes - 12.0.1.63	4786
Total Prefixes - 4.68.0.243	0
Total Prefixes - 66.185.128.1	3089

# Raw Data Analysis – AS Query



BGP-Inspect - bgpinspect.merit.edu - Mozilla

http://bgpinspect.merit.edu/query.php

**Query Summary Statistics**

Attribute	Value
Query Time Range Start	June 8, 2006, 12:00 am +0000
Query Time Range End	June 28, 2006, 12:00 am +0000
Total Update Messages	6428
Unique Prefixes	141
Time to run query	26.307 seconds

**Prefixes Announced**

Time	Prefix	AS Path	Communities
June 8, 2006, 12:51 am +0000	140.219.134.0/24	3549 1239 22284 688 145 145	3549:2021 3549:30840
June 8, 2006, 12:52 am +0000	140.219.134.0/24	3549 1239 22284 688 145 145	3549:2193 3549:30840
June 8, 2006, 12:52 am +0000	140.219.134.0/24	3549 1239 22284 688 145 145	3549:2141 3549:30840
June 8, 2006, 12:56 am +0000	140.219.134.0/24	3549 1239 22284 688 145 145	3549:2021 3549:30840

# Raw Data Analysis – Prefix query

BGP-Inspect - bgpinspect.merit.edu - Mozilla

File Edit View Go Bookmarks Tools Window Help

http://bgpinspect.merit.edu/index.php

## BGP-Inspect-Routeviews

First DB Update: Aug 1, 2005  
Last DB Update: June 28, 2006

Home Reports Documentation FAQ About

**Global Summary Queries:** (Please select a peer, query type and duration)

Peer: 12.0.1.63 - ATT  
4.68.0.243 - Level 3  
66.185.128.1 - AOL  
144.228.241.81 - Sprint  
208.51.134.253 - Global X

Query Type: Most Active ASes  
Most Active Prefixes  
Prefixes Most Announced  
Prefixes Most Withdrawn  
Prefixes with Most AS Changes

Submit Query

---

**Raw Data Analysis:** (Please select a peer, query type, AS/prefix, and time range)

Peer: 12.0.1.63 - ATT  
4.68.0.243 - Level 3  
66.185.128.1 - AOL  
144.228.241.81 - Sprint  
208.51.134.253 - Global X

Query Type: AS  
Prefix-Exact  
Prefix-More Specific

Query: (ASN or a.b.c.d/len)  
140.219.134.0/24

Start Date: 2006 Jun 8 00:00  
End Date: 2006 Jun 28 00:00

Submit Query

Copyright © Merit Network, Inc.

BGP-Inspect - bgpinspect.merit.edu - Mozilla

File Edit View Go Bookmarks Tools Window Help

http://bgpinspect.merit.edu/query.php

## BGP-Inspect-Routeviews

First DB Update: Aug 1, 2005, 12:01 am +0000  
Last DB Update: June 28, 2006, 12:11 am +0000

Home Reports Documentation FAQ About

Aggregate 12.0.1.63 4.68.0.243 66.185.128.1 144.228.241.81 208.51.134.253

**Peer: Aggregate**  
**Prefix: 140.219.134.0/24**

**Prefix Activity**

Num Messages

Date

Legend: 12.0.1.63, 4.68.0.243, 66.185.128.1, 144.228.241.81, 208.51.134.253

Query Summary Statistics	
Attribute	Value
Query Time Range Start	June 8, 2006, 12:00 am +0000
Query Time Range End	June 28, 2006, 12:00 am +0000
Total Update Messages(All Peers)	320
Total Announce Messages(All Peers)	249
Total Withdraw Messages(All Peers)	71

# Raw Data Analysis – Prefix query

BGP-Inspect - bgpinspect.merit.edu - Mozilla

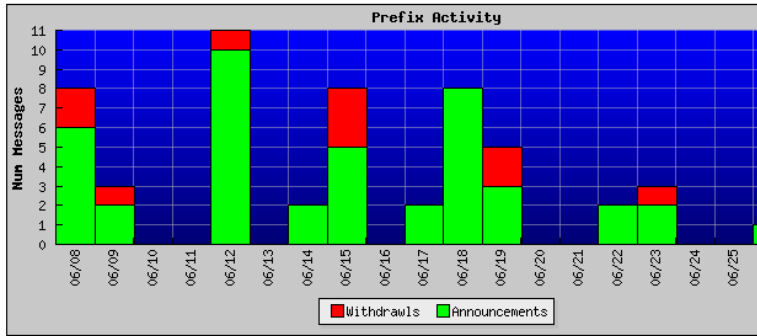
File Edit View Go Bookmarks Tools Window Help

http://bgpinspect.merit.edu/query.php

**BGP-Inspect-Routeviews** First DB Update: Aug 1, 2006  
Last DB Update: June 28, 2006

Home	Reports	Documentation	FAQ	About
Aggregate	12.0.1.63	4.68.0.243	66.185.128.1	144.228.241.81

Peer: 66.185.128.1  
Prefix: 140.219.134.0/24



Attribute	Value
Query Time Range Start	June 8, 2006, 12:00 am +0000
Query Time Range End	June 28, 2006, 12:00 am +0000
Total Update Messages	55
Total Announce Messages	44
Total Withdraw Messages	11
Maximum AS Path Length	6
Minimum AS Path Length	6
Average AS Path Length	6.000000
Origin AS Changes	0
Number of Unique ASes	1
Origin ASes List	145
Time to run query	3.102 seconds

BGP-Inspect - bgpinspect.merit.edu - Mozilla

File Edit View Go Bookmarks Tools Window Help

http://bgpinspect.merit.edu/query.php

Prefix Announcements:				
Time	Type	AS Path	Communities	
June 8, 2006, 12:51 am +0000	a	3549 1239 22284 688 145 145	3549:2021	3549:30840
June 8, 2006, 12:52 am +0000	a	3549 1239 22284 688 145 145	3549:2193	3549:30840
June 8, 2006, 12:52 am +0000	w	-	-	-
June 8, 2006, 12:52 am +0000	a	3549 1239 22284 688 145 145	3549:2141	3549:30840
June 8, 2006, 12:56 am +0000	a	3549 1239 22284 688 145 145	3549:2021	3549:30840
June 8, 2006, 12:57 am +0000	a	3549 1239 22284 688 145 145	3549:2141	3549:30840
June 8, 2006, 1:01 am +0000	a	3549 1239 22284 688 145 145	3549:2021	3549:30840
June 8, 2006, 1:02 am +0000	a	3549 1239 22284 688 145 145	3549:2141	3549:30840
June 8, 2006, 5:22 pm +0000	a	3549 1239 22284 688 145 145	3549:2523	3549:31826
June 8, 2006, 5:22 pm +0000	w	-	-	-
June 8, 2006, 5:24 pm +0000	a	3549 1239 22284 688 145 145	3549:2021	3549:30840
June 8, 2006, 5:24 pm +0000	a	3549 1239 22284 688 145 145	3549:2141	3549:30840
June 9, 2006, 2:10 pm +0000	a	3549 1239 22284 688 145 145	3549:2291	3549:30840
June 9, 2006, 2:11 pm +0000	a	3549 1239 22284 688 145 145	3549:2401	3549:30840
June 9, 2006, 2:11 pm +0000	w	-	-	-
June 9, 2006, 2:14 pm +0000	a	3549 1239 22284 688 145 145	3549:2193	3549:30840
June 9, 2006, 2:14 pm +0000	a	3549 1239 22284 688 145 145	3549:2141	3549:30840
June 12, 2006, 12:42 am +0000	w	-	-	-
June 12, 2006, 12:43 am +0000	a	3549 1239 22284 688 145 145	3549:2291	3549:30840
June 12, 2006, 12:43 am +0000	a	3549 1239 22284 688 145 145	3549:2021	3549:30840
June 12, 2006, 12:43 am +0000	a	3549 1239 22284 688 145 145	3549:2141	3549:30840
June 12, 2006, 3:00 am +0000	a	3549 1239 22284 688 145 145	3549:2021	3549:30840
June 12, 2006, 3:00 am +0000	a	3549 1239 22284 688 145 145	3549:2291	3549:30840
June 12, 2006, 3:00 am +0000	a	3549 1239 22284 688 145 145	3549:2193	3549:30840
June 12, 2006, 3:01 am +0000	a	3549 1239 22284 688 145 145	3549:2401	3549:30840
June 12, 2006, 3:01 am +0000	a	3549 1239 22284 688 145 145	3549:2021	3549:30840
June 12, 2006, 3:01 am +0000	a	3549 1239 22284 688 145 145	3549:2141	3549:30840
June 12, 2006, 3:28 am +0000	a	3549 1239 22284 688 145 145	3549:2021	3549:30840
June 12, 2006, 3:29 am +0000	a	3549 1239 22284 688 145 145	3549:2141	3549:30840
June 12, 2006, 3:39 am +0000	a	3549 1239 22284 688 145 145	3549:2021	3549:30840
June 12, 2006, 3:39 am +0000	a	3549 1239 22284 688 145 145	3549:2141	3549:30840

# BGP-Inspect: Current State

BGP-Inspect – Beta v0.5  
<http://bgpinspect.merit.edu>

Dataset: August 1, 2005 - Present

Current BGPdb size: 170GB

Currently indexing data for 5 peers (AT&T,  
Level 3, AOL, Sprint, Global X)

- Example queries (per peer, 1,7,10 days):
  - Most active AS's
  - Most active prefixes
  - Prefixes with most OriginAS changes
- Raw Data Analysis(per peer)
  - Prefix/AS, Time Range
  - Uniques prefixes by AS
  - OriginAS changes for a prefix
  - Time to run query
  - More specific prefixes announced

The top screenshot shows the BGP-Inspect web interface with the following sections:

- Global Summary Queries:** (Please select a RouteViews Peer, Query Type and Duration)
  - RouteViews Peer: 12.0.1.63 - AT1, 4.68.0.243 - Level 3, 66.185.120.1 - AOL, 144.228.241.01 - Sprint, 208.51.134.253 - GlobalX
  - Query Type: Most Active ASes, Most Active Prefixes, Prefixes Most Announced, Prefixes Most Withdrawn, Prefixes with Most AS Changes
  - Duration: Last 1 Day, Last 2 Days, Last 7 Days, Last 10 Days
- Raw Data Analysis:** (Please select a RouteViews Peer, Query Type, AS/Prefix, and Time Range)
  - RouteViews Peer: 12.0.1.63 - AT1, 4.68.0.243 - Level 3, 66.185.120.1 - AOL, 144.228.241.01 - Sprint, 208.51.134.253 - GlobalX
  - Query Type: AS, Prefix Exact, Prefix More Specific
  - Query: (ASN or s.b.c.d/len)
  - Start Date: 2005 Oct 13 00:00, End Date: 2005 Oct 20 00:00

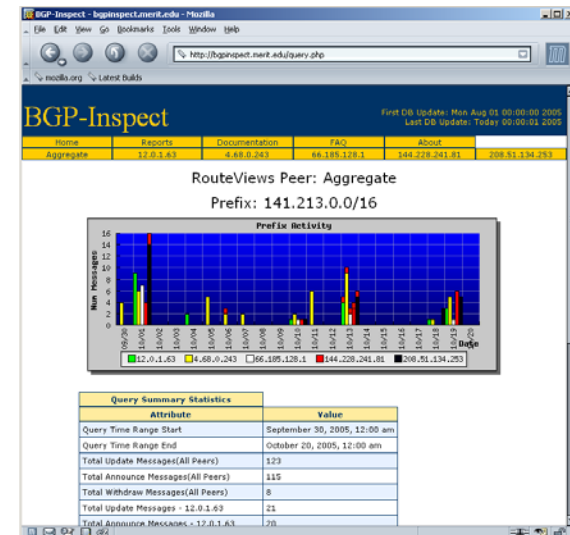
The bottom screenshot shows a detailed view for RouteViews Peer 12.0.1.63, AS 721 (DNIC DoD Network Information Center):

- RouteViews Peer: 12.0.1.63**
- AS: 721 (DNIC DoD Network Information Center)**
- Total Number of prefix announcements:** A bar chart showing the number of prefix announcements over time from 09/22 to 10/20. The y-axis ranges from 0 to 12000. The x-axis shows dates from 09/22 to 10/20. The chart shows a significant peak in announcements around 10/10.
- Query Summary Statistics:**

Attribute	Value
Query Time Range Start	September 21, 2005, 12:00 am
Query Time Range End	October 20, 2005, 12:00 am
Total Update Messages	37574
Unique Prefixes	1108
Time to run query	22.831 seconds

# BGP-Inpsect: Current State (2)

- Equipment
  - Dell 2650 - Web and DB server
  - Dell 2850, dual Xeon with NFS mounted 500GB SATA
- Traffic?
  - ~30+ unique IPs per day



RouteViews Peer: 12.0.1.63  
Overall Most Active Prefixes, Last 7 Days

Rank	Prefix	Total	Announce	Withdrawn	Origin AS Changes
1	81.213.47.0/24	10685	6202	4483	0
2	199.191.128.0/22	10316	10316	0	0
3	199.191.160.0/24	10316	10316	0	0
4	199.191.192.0/24	10316	10316	0	0
5	199.191.200.0/24	10316	10316	0	0
6	199.191.208.0/24	10316	10316	0	0
7	192.35.39.0/24	9463	8967	496	4022
8	81.212.149.0/24	4446	3373	1073	0
9	69.26.199.0/24	4103	3107	996	0
10	209.140.24.0/24	3811	3312	499	0
11	81.212.197.0/24	3779	2800	979	0
12	113.184.73.0/24	3593	1792	1761	0
13	209.144.205.0/24	2910	1455	1455	0
14	207.168.184.0/24	2725	2235	490	0
15	66.150.140.0/23	2446	2389	57	0
16	216.85.83.0/24	2135	2076	59	0
17	63.144.114.0/24	1865	1795	70	0
18	203.135.6.0/24	1815	1735	80	0

# BGP-Inspect Next Steps

- BGP-Inspect is available at <http://bgpinspect.merit.edu> and your feedback is very much appreciated.
- Future...
  - More interesting things with the multiple peer response UI (different ways of highlighting the differences between peers)
  - pyBGPdb - a python interface to the BGPdb database providing fast raw queries



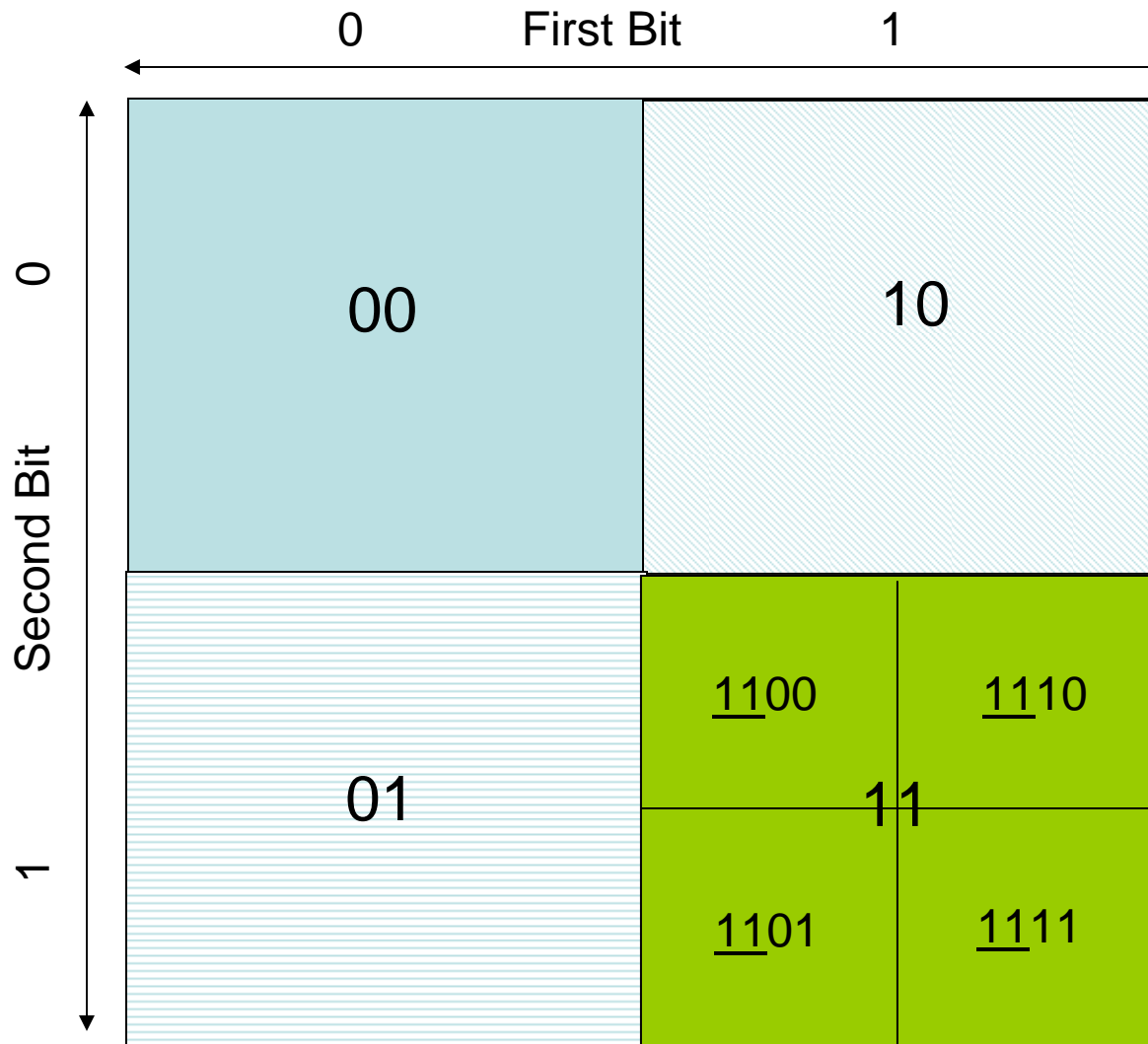
# Outline

- The Problem
- BGP/Routing Information
  - BGP-Inspect – Information Extraction from BGP Update messages
  - VAST – Internet AS topology Visualization
- Netflow/Traffic Information
  - Flamingo – Internet Traffic Exploration
- Conclusions

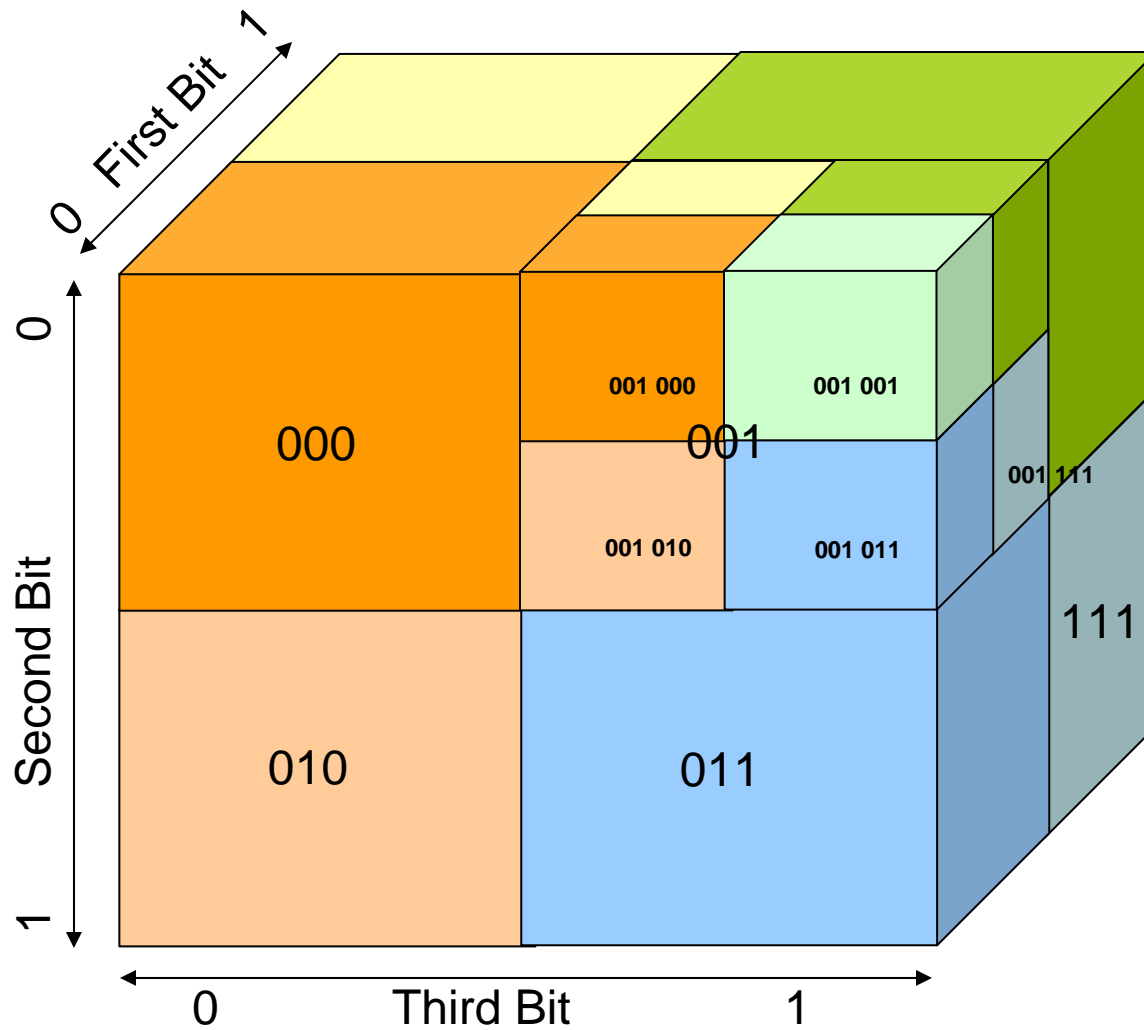
# VAST – Visualizing AS Topology

- VAST allows users to easily navigate, and explore various topological properties and features extracted from raw BGP update messages
- VAST uses both a quad-tree based algorithm as well as an octo-tree based method to build various visualization
- The ability to navigate the three dimensional space to fully explore the dataset make VAST a unique tool

# The Basic Quad-Tree



# Octo-Tree Algorithm



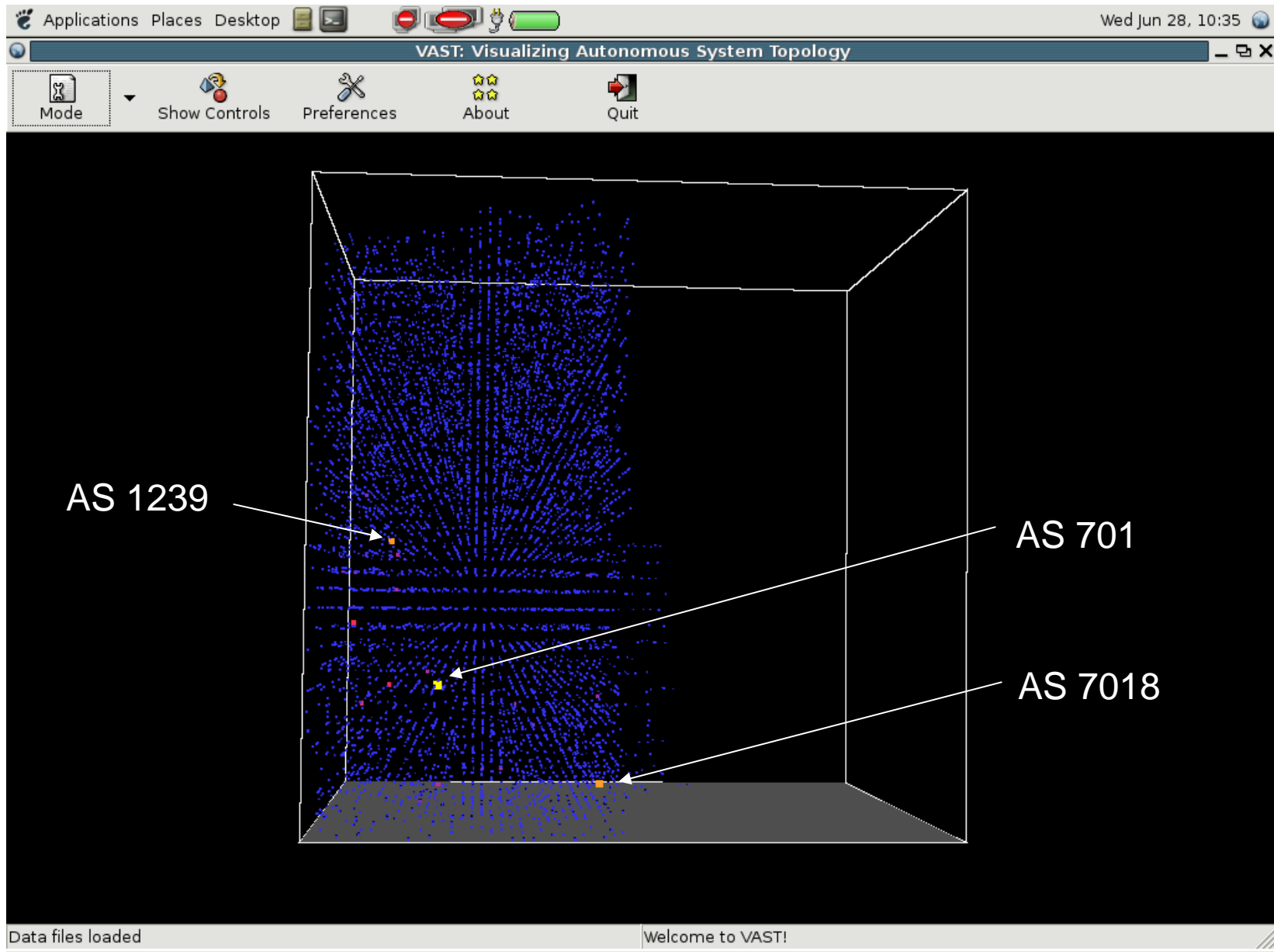
# VAST - Visualization Methods

- Out-degree per AS
- Per AS unique prefix originations
- AS topology with line scaling:
  - Peer out-degree
  - Frequency with which AS pair is seen
  - Unique prefixes with certain AS pair
  - Total address space over an AS pair

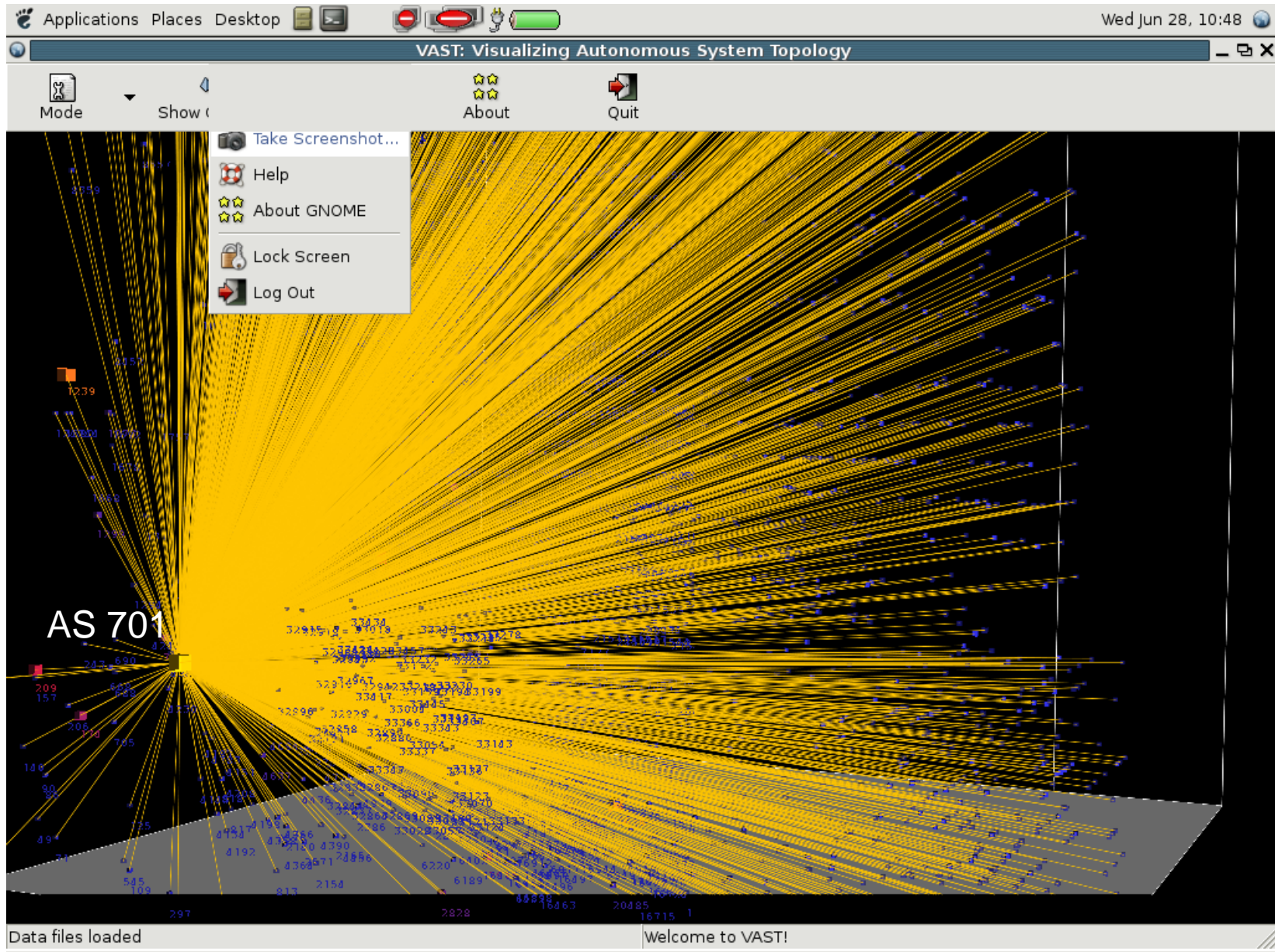
# VAST – Techniques

- Position of a node determined by quad/octo-tree
- Size determined by out-degree of node, larger out-degree -> larger size
- Color determined by out-degree of node, larger out-degree -> more yellow
- Line thickness depends on various factors(selectable), out-degree of neighbor, number of prefixes, address space size, or frequency of messages
- 3D navigation of visualization, slider bar controls, selectable listing of displayed information to control/filter what is being displayed

# VAST – ASN Distribution

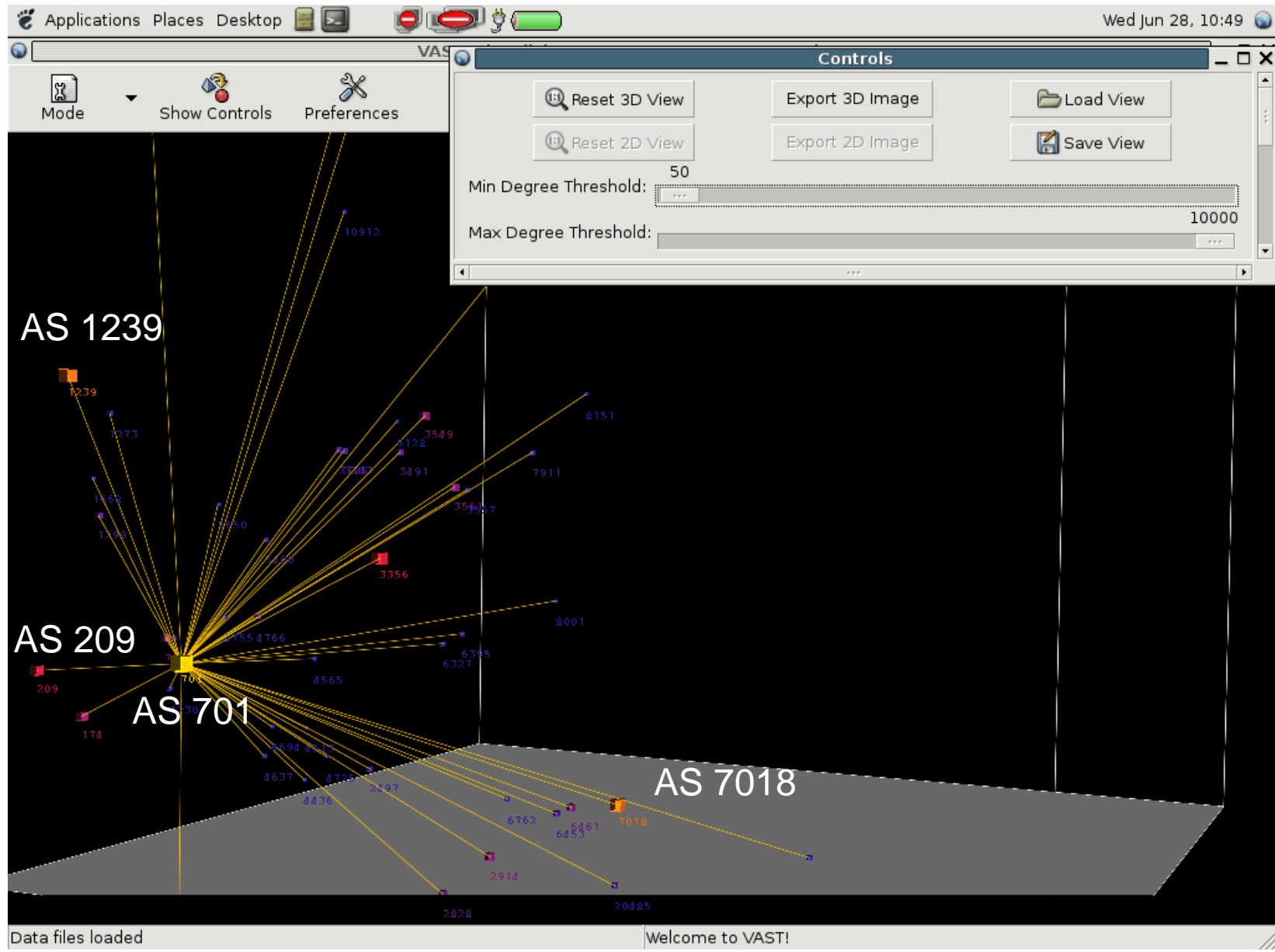


# VAST - AS CORE

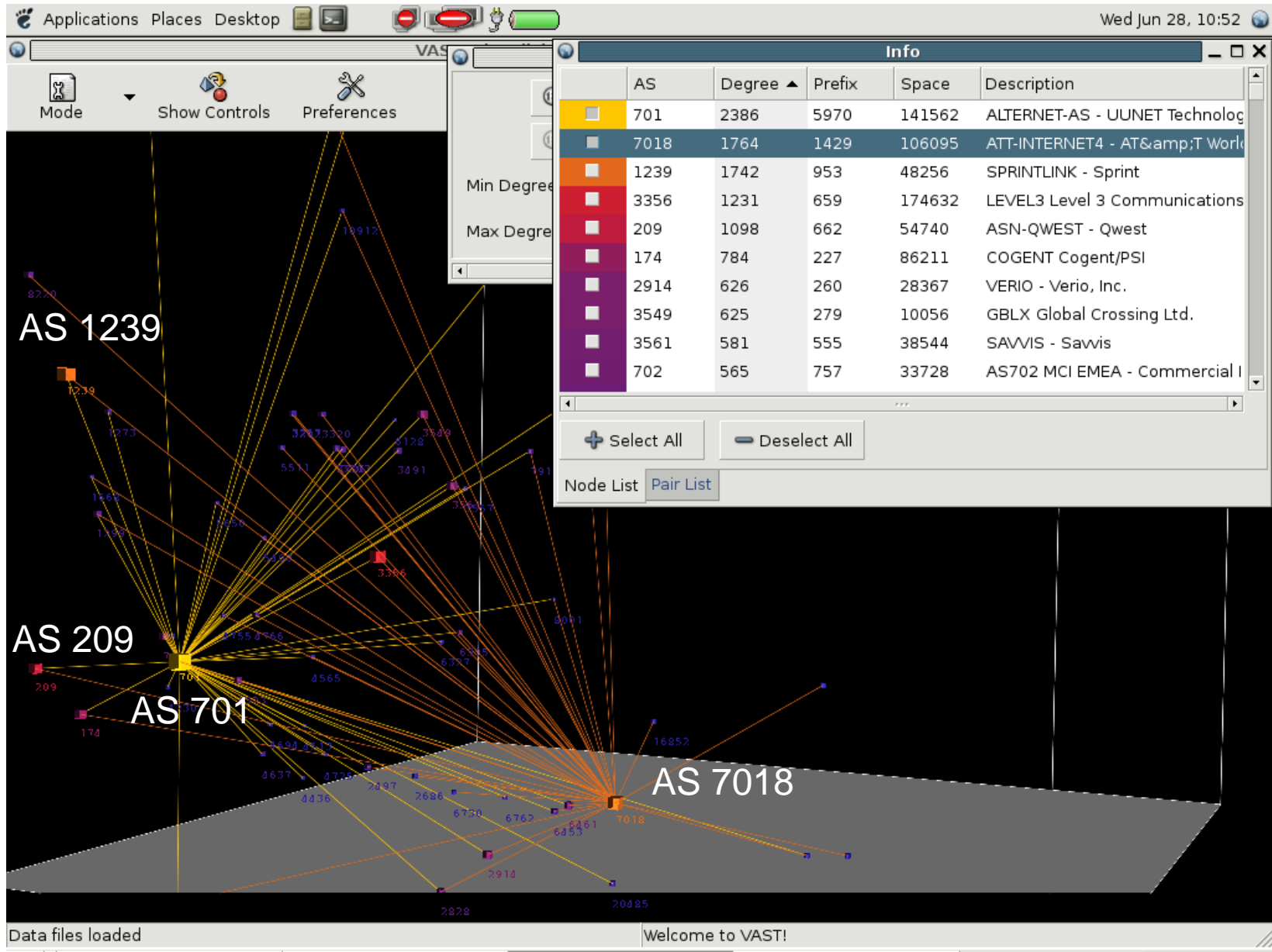




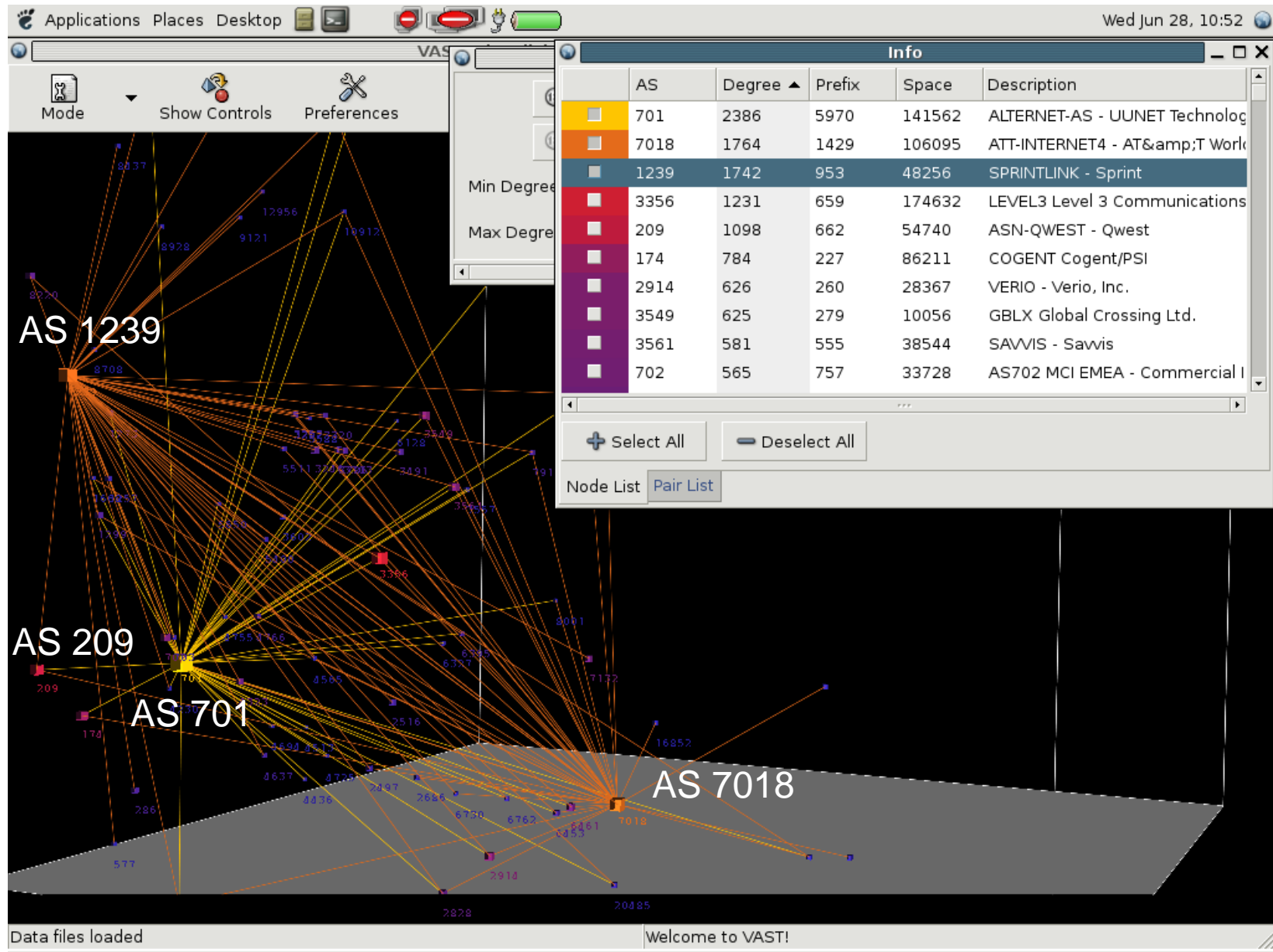
# VAST – AS CORE



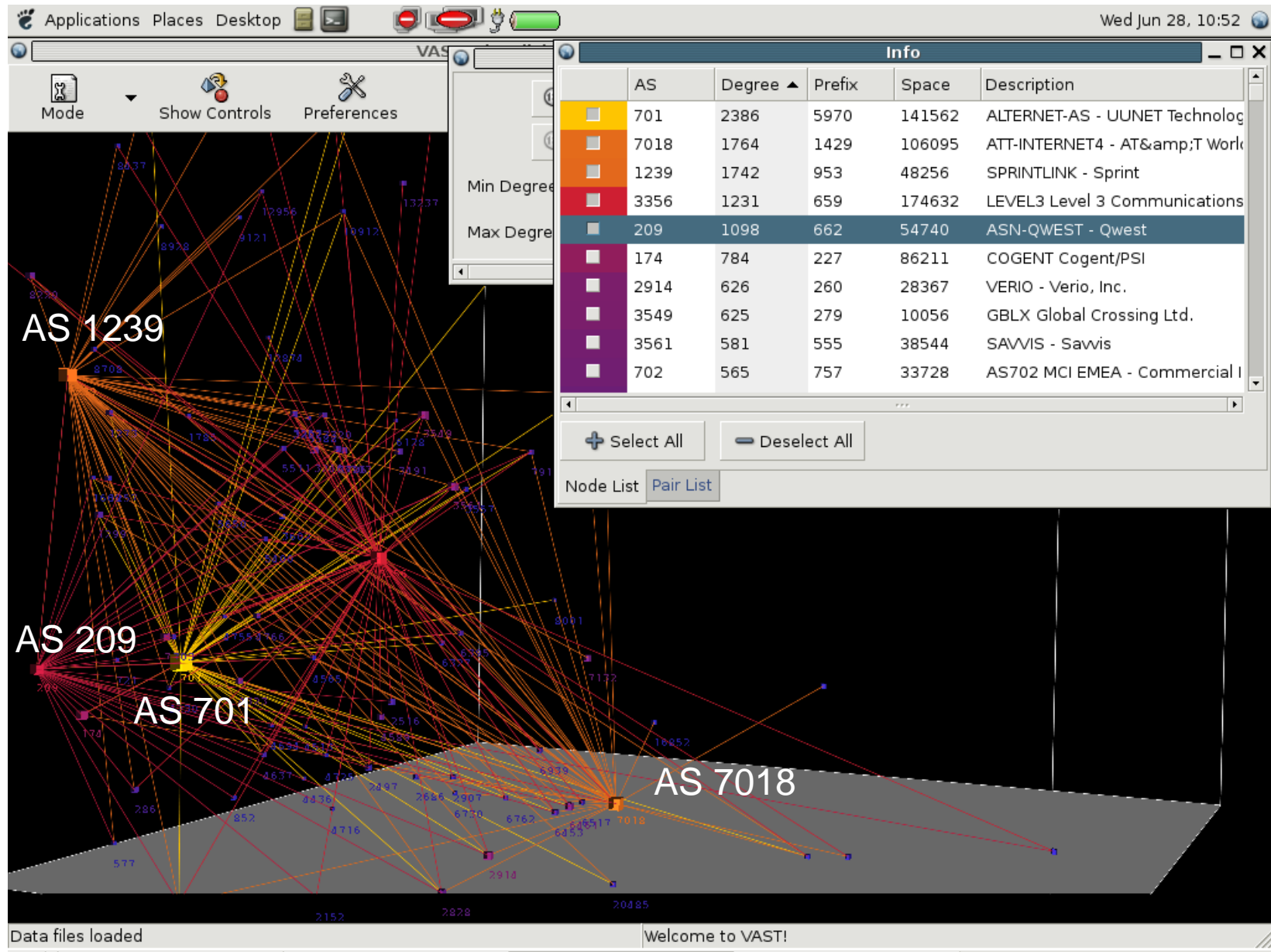
# VAST – AS CORE



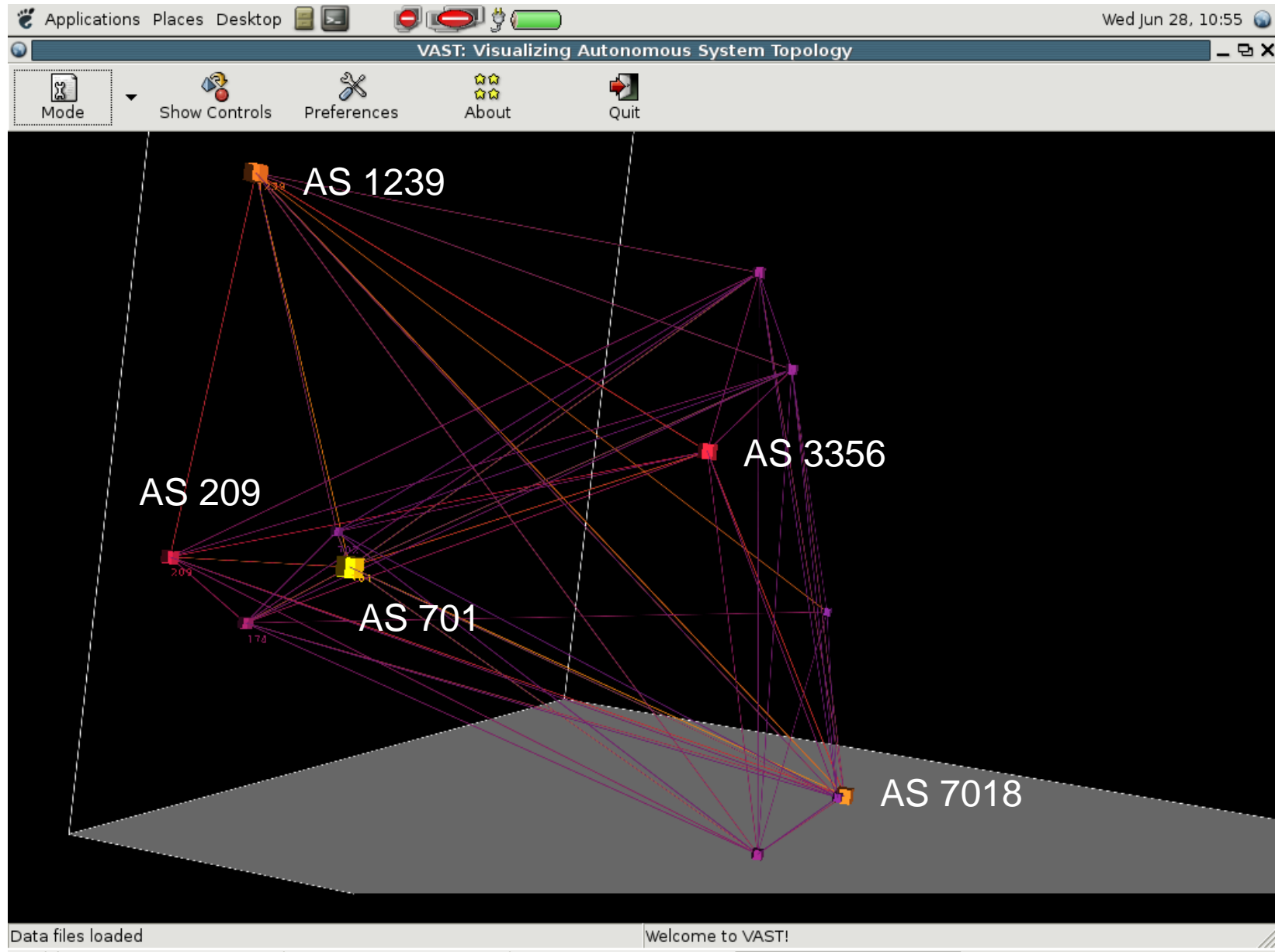
# VAST – AS CORE



# VAST – AS CORE

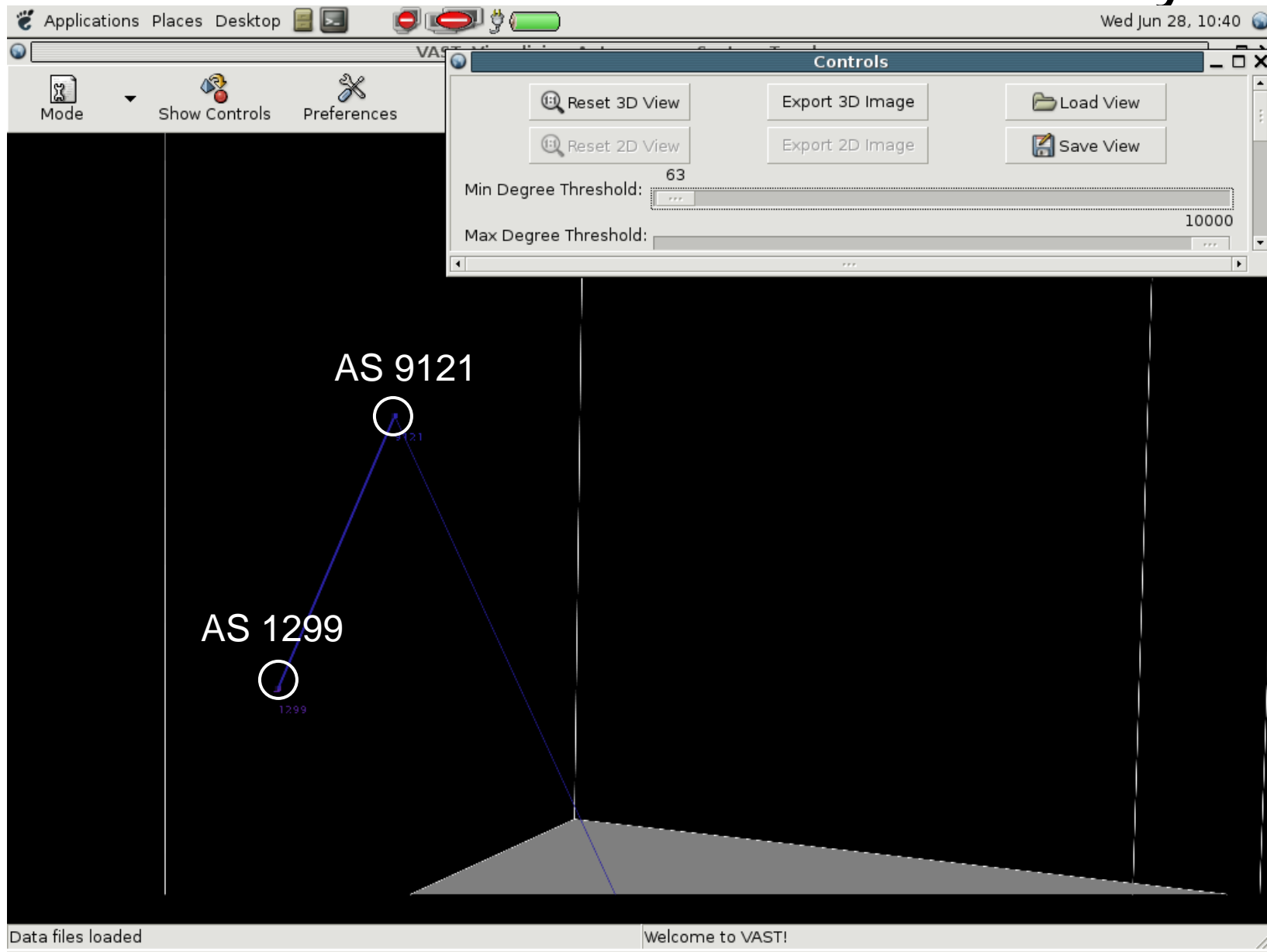


# VAST – AS CORE(500 Club)



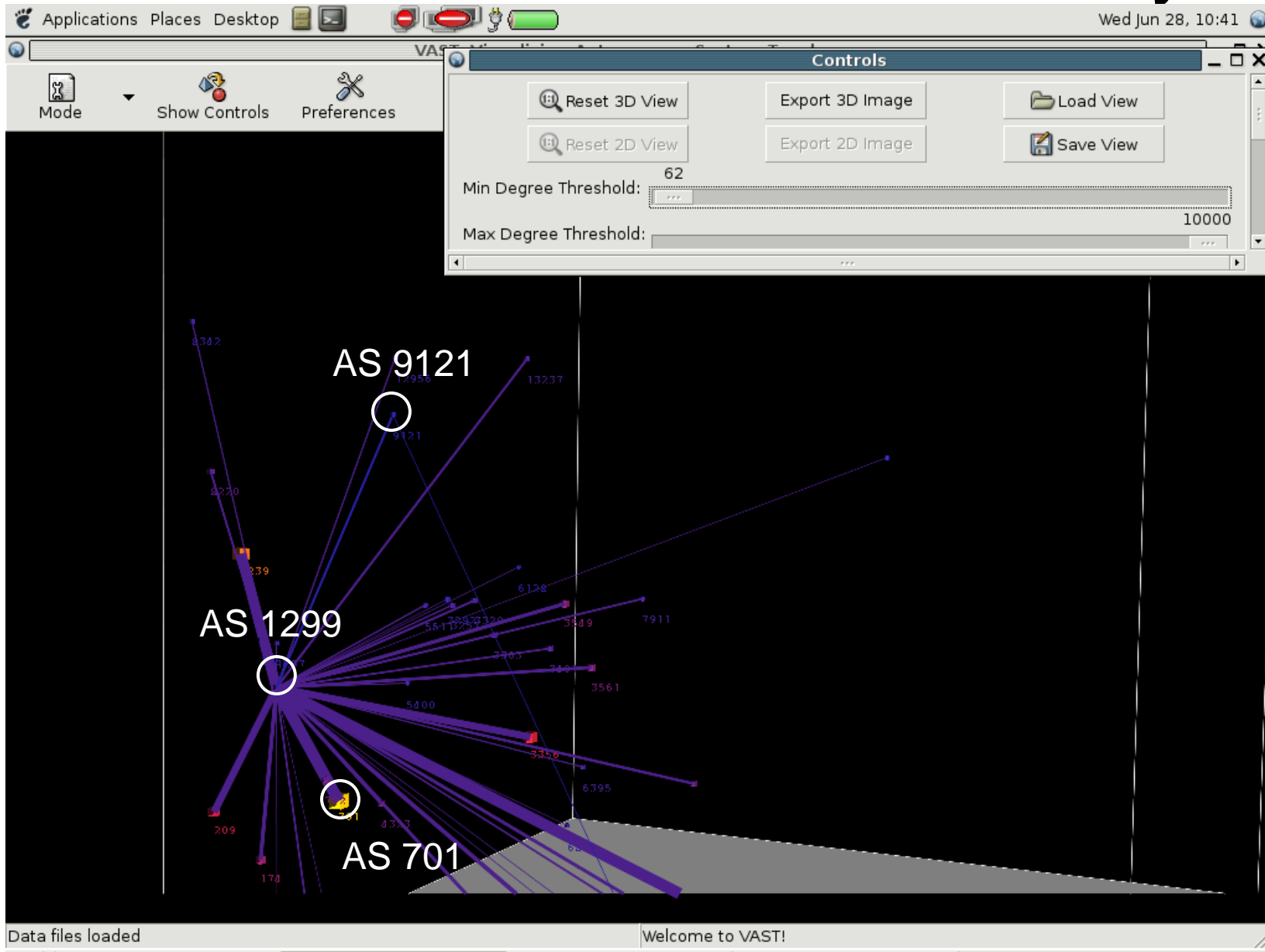


# VAST – AS9121 connectivity



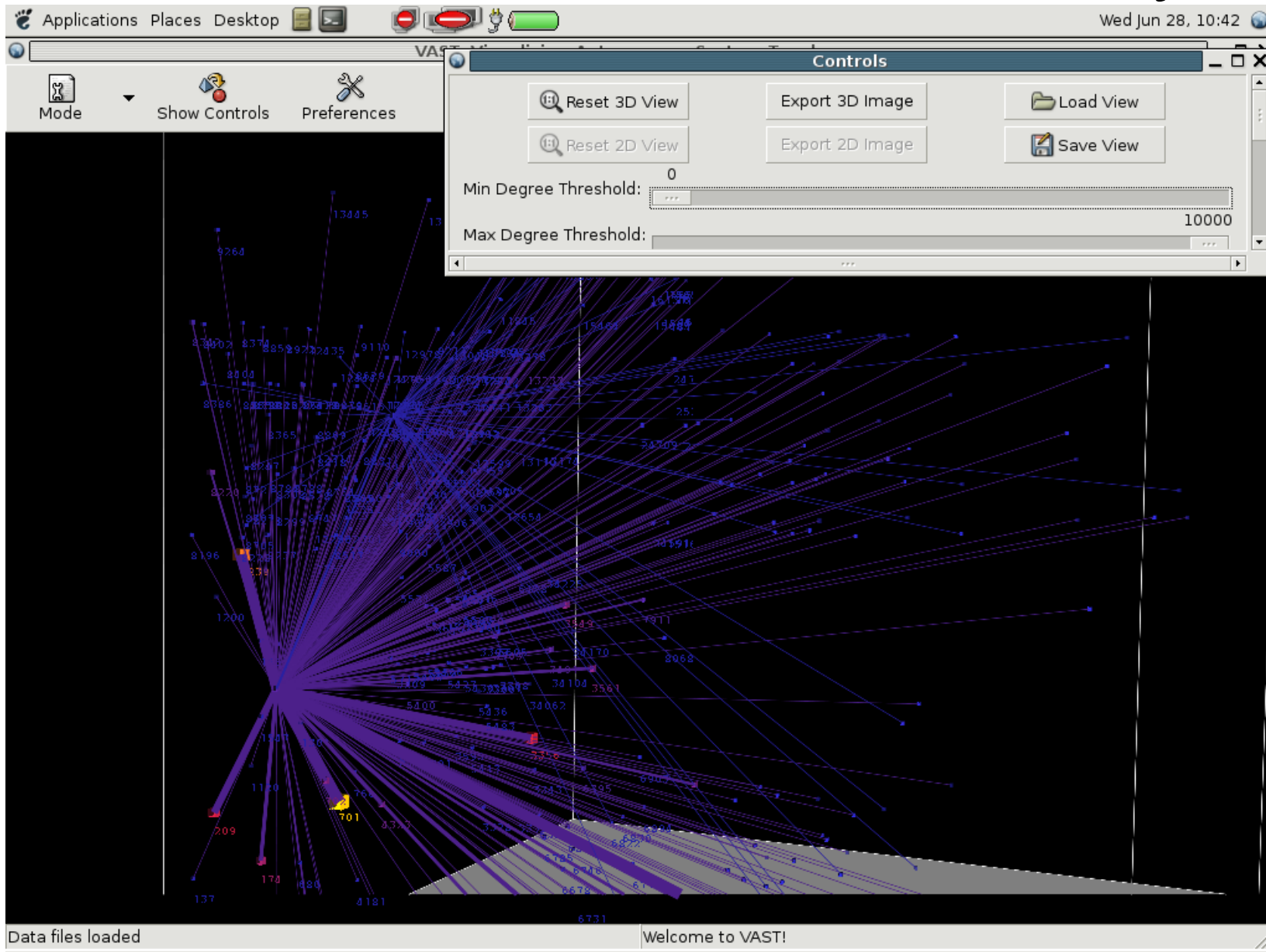


# VAST – AS9121 connectivity

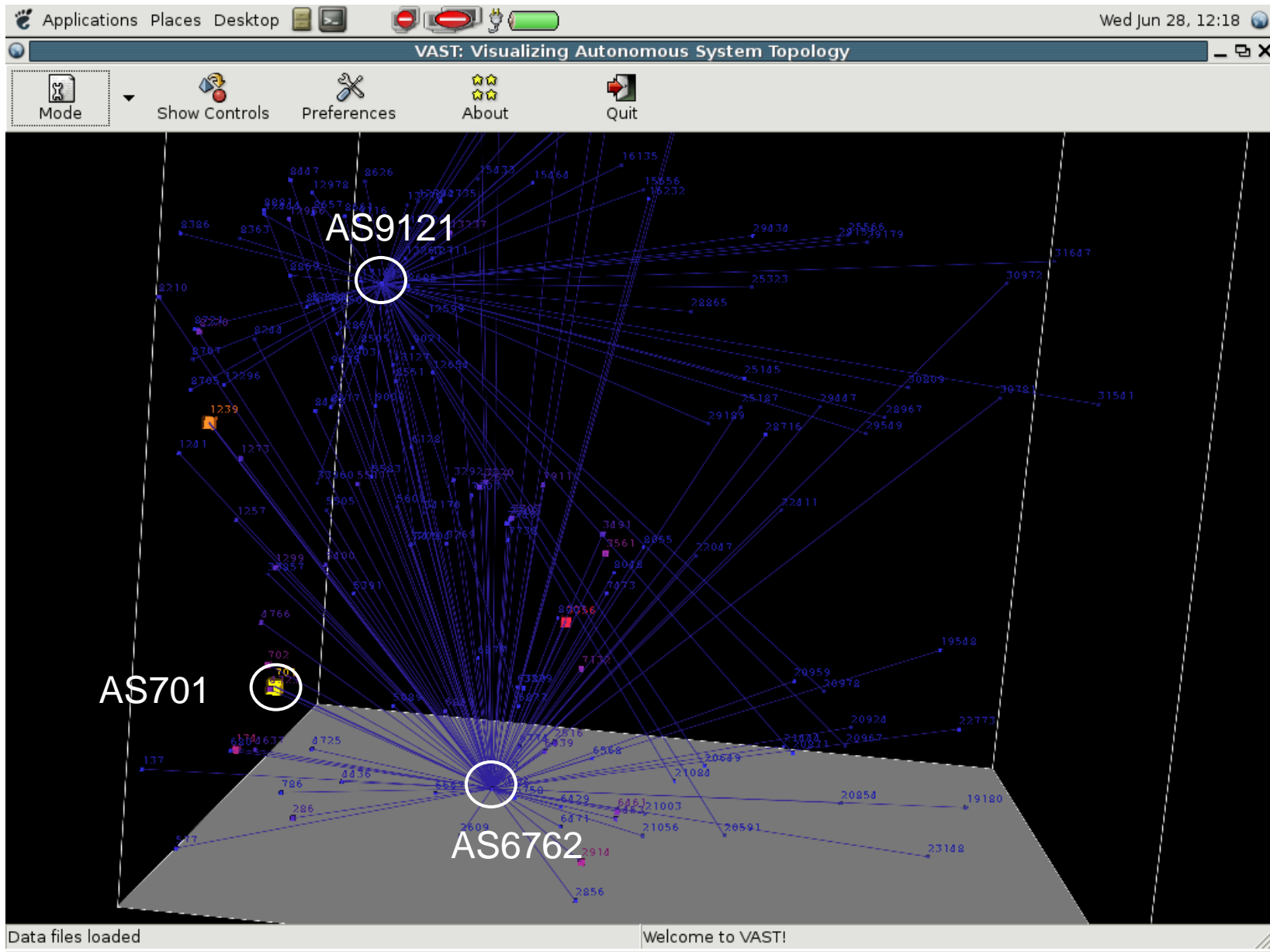




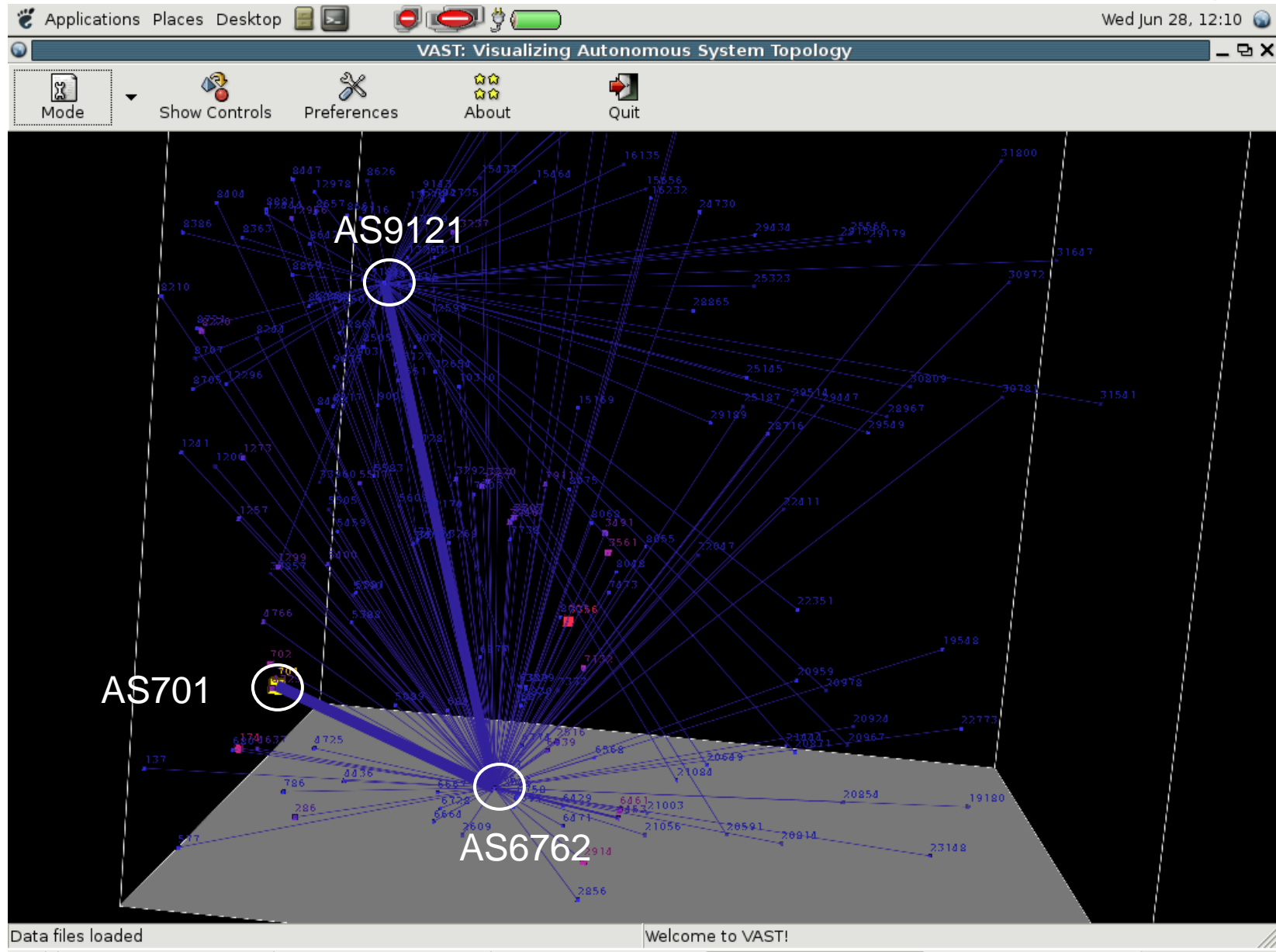
# VAST – AS9121 connectivity



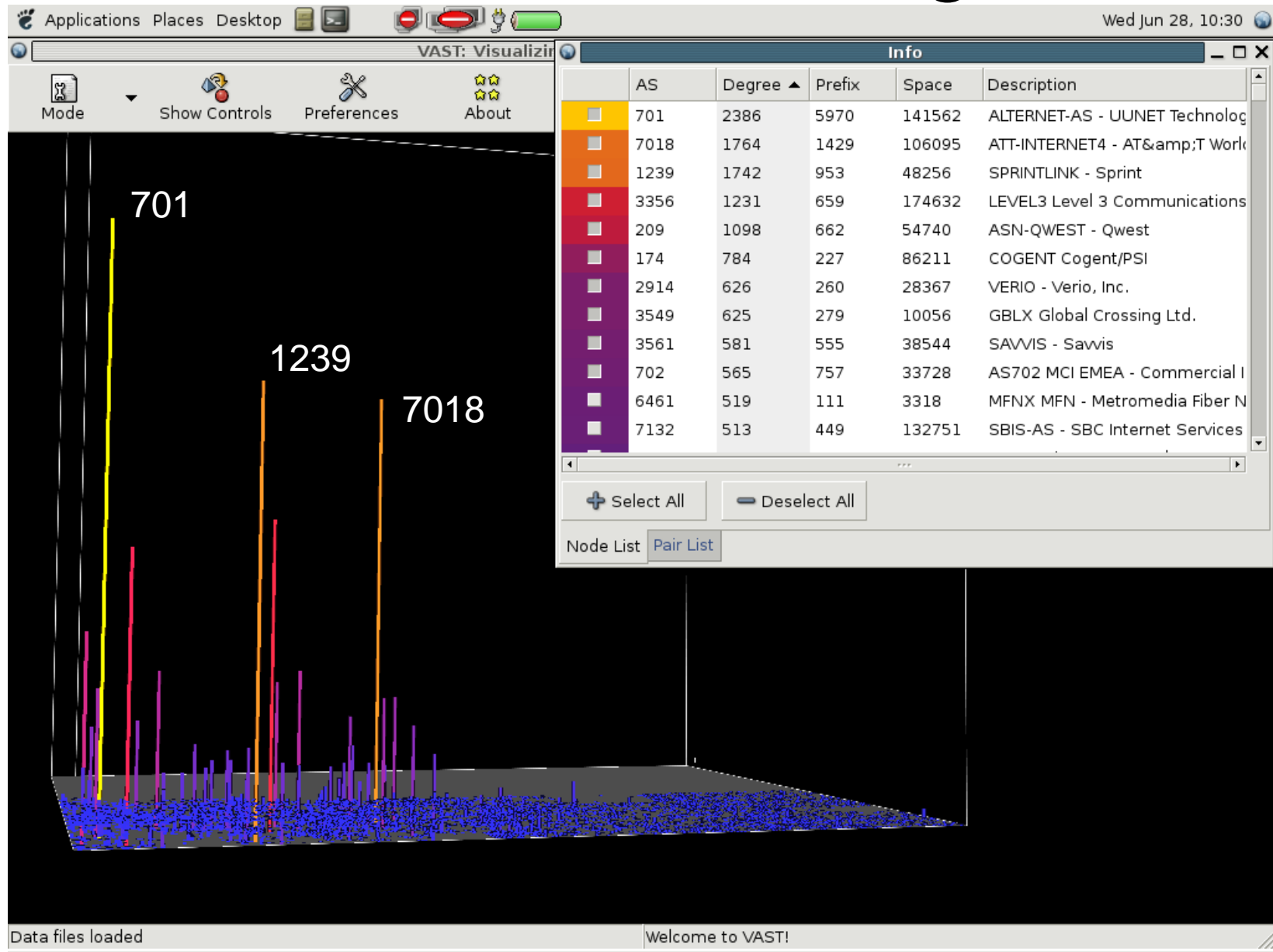
# VAST – AS9121 Route Leakage



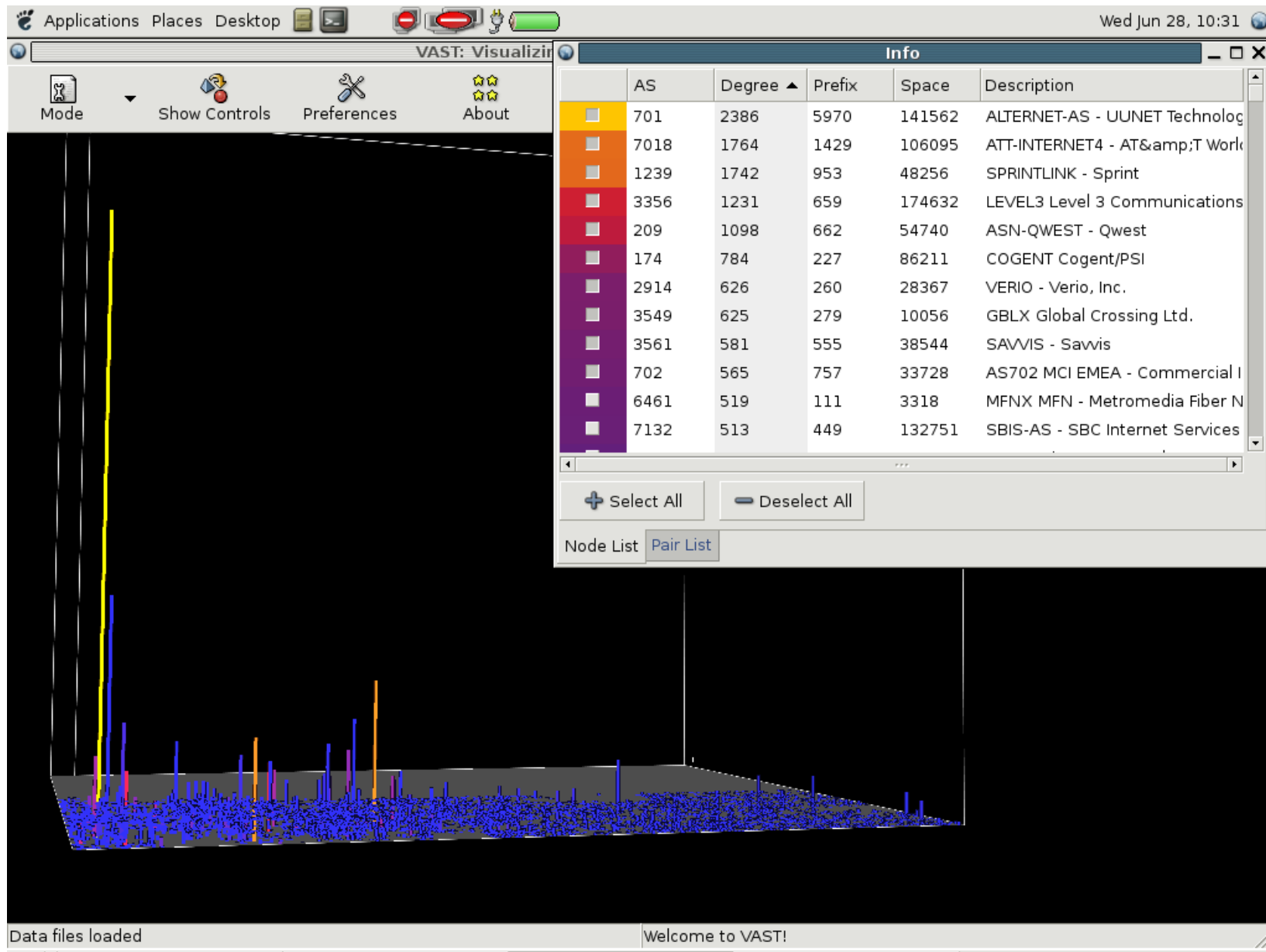
# VAST – AS9121 Route Leakage



# VAST – AS Out-degree



# VAST – Unique Prefix



# Outline

- The Problem
- BGP/Routing Information
  - BGP-Inspect – Information Extraction from BGP Update messages
  - VAST – Internet AS topology Visualization
- Netflow/Traffic Information
  - Flamingo – Internet Traffic Exploration
- Conclusions

# Flamingo – Visualizing Internet Traffic Data

- 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!

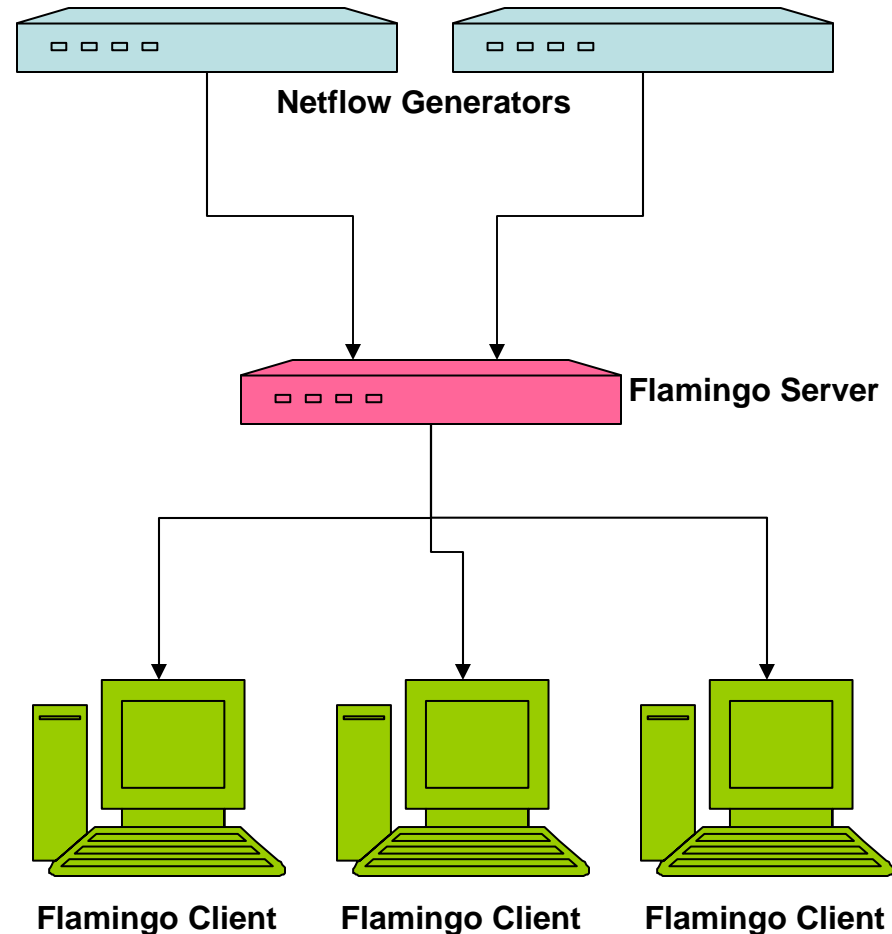
# Flamingo - Introduction

- 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

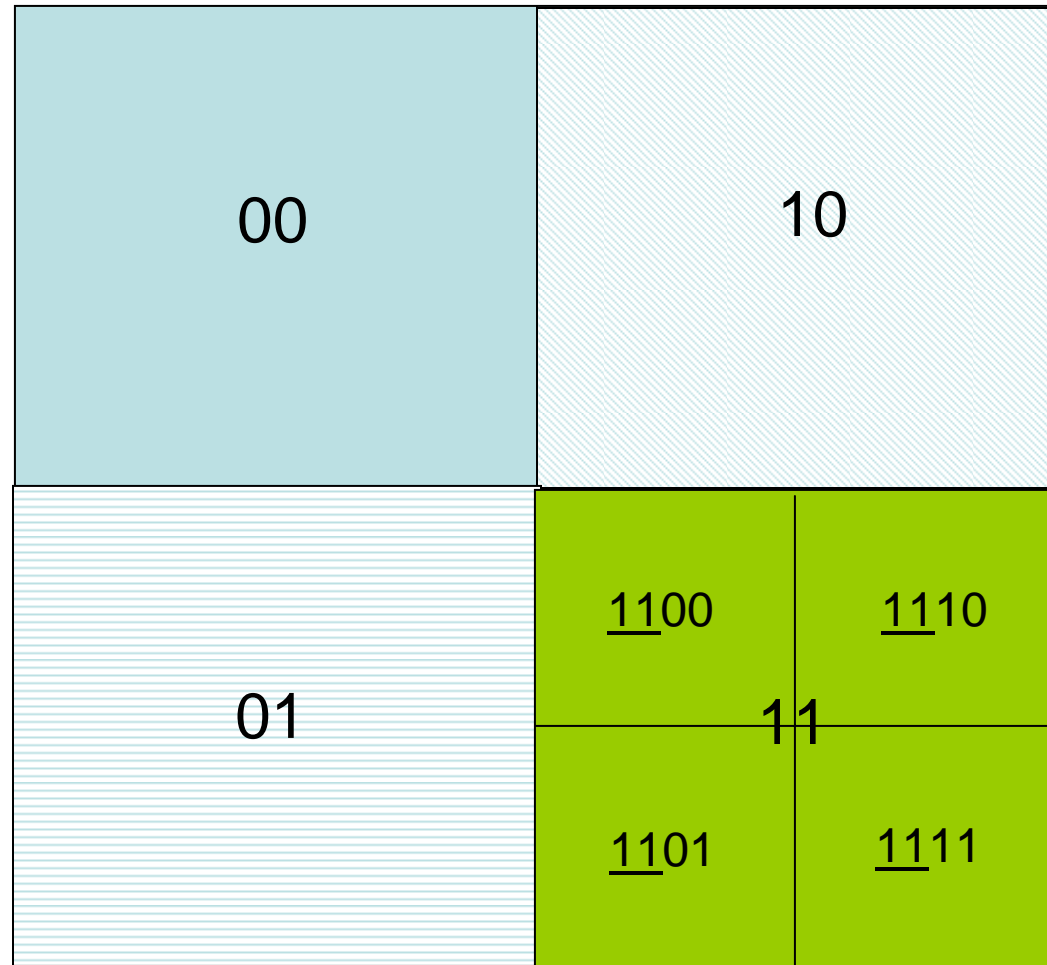
- Client/Server Architecture
- A single server can support multiple clients
- A single server can act as collector for multiple netflow feeds
- Supports both aggregation as well as non-aggregation mode



# Flamingo - 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

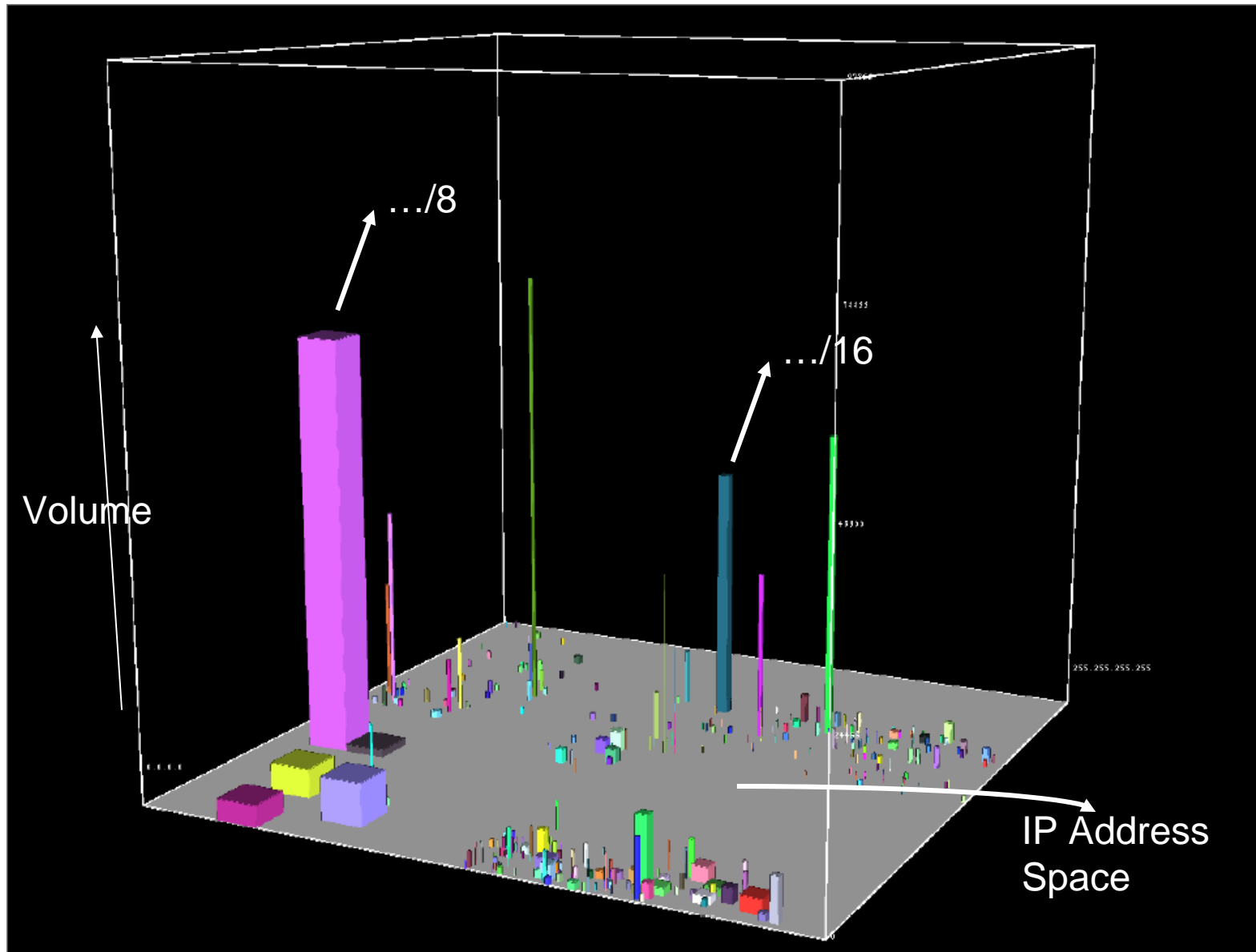
# The Basic Quad-Tree



# 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

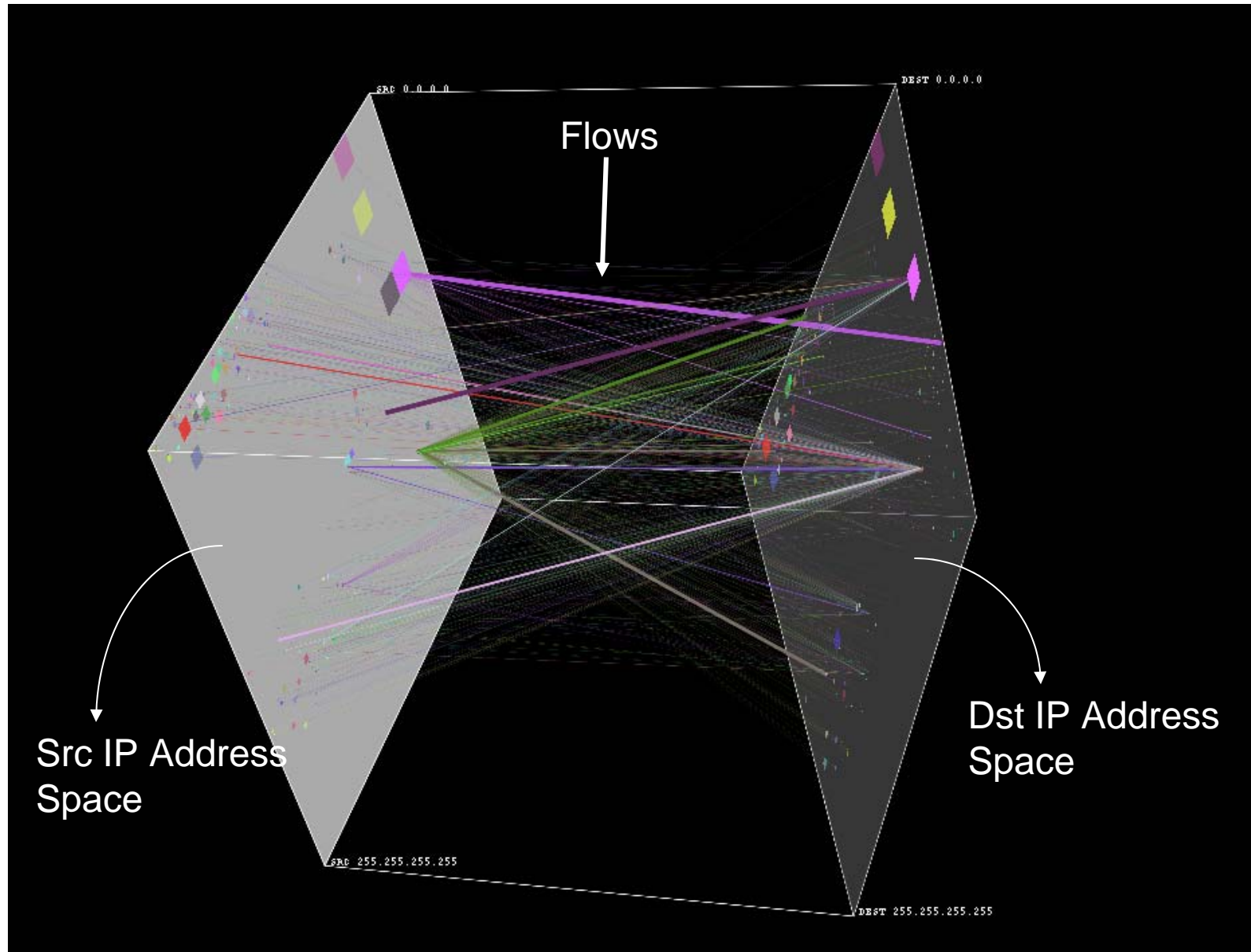
# Traffic Volume by Src/Dst IP



## 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

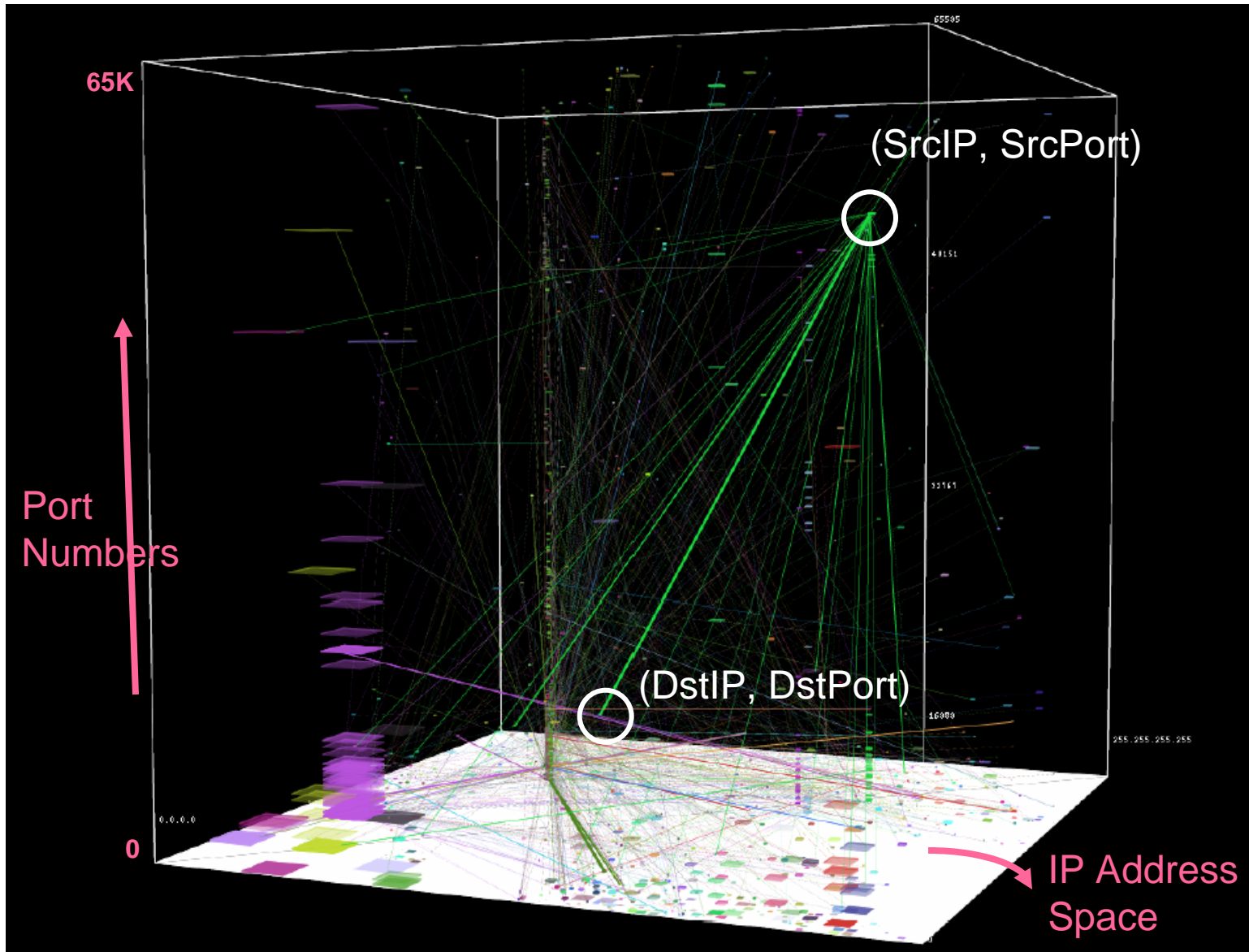


# Traffic Flows by Src/Dst IP and Port

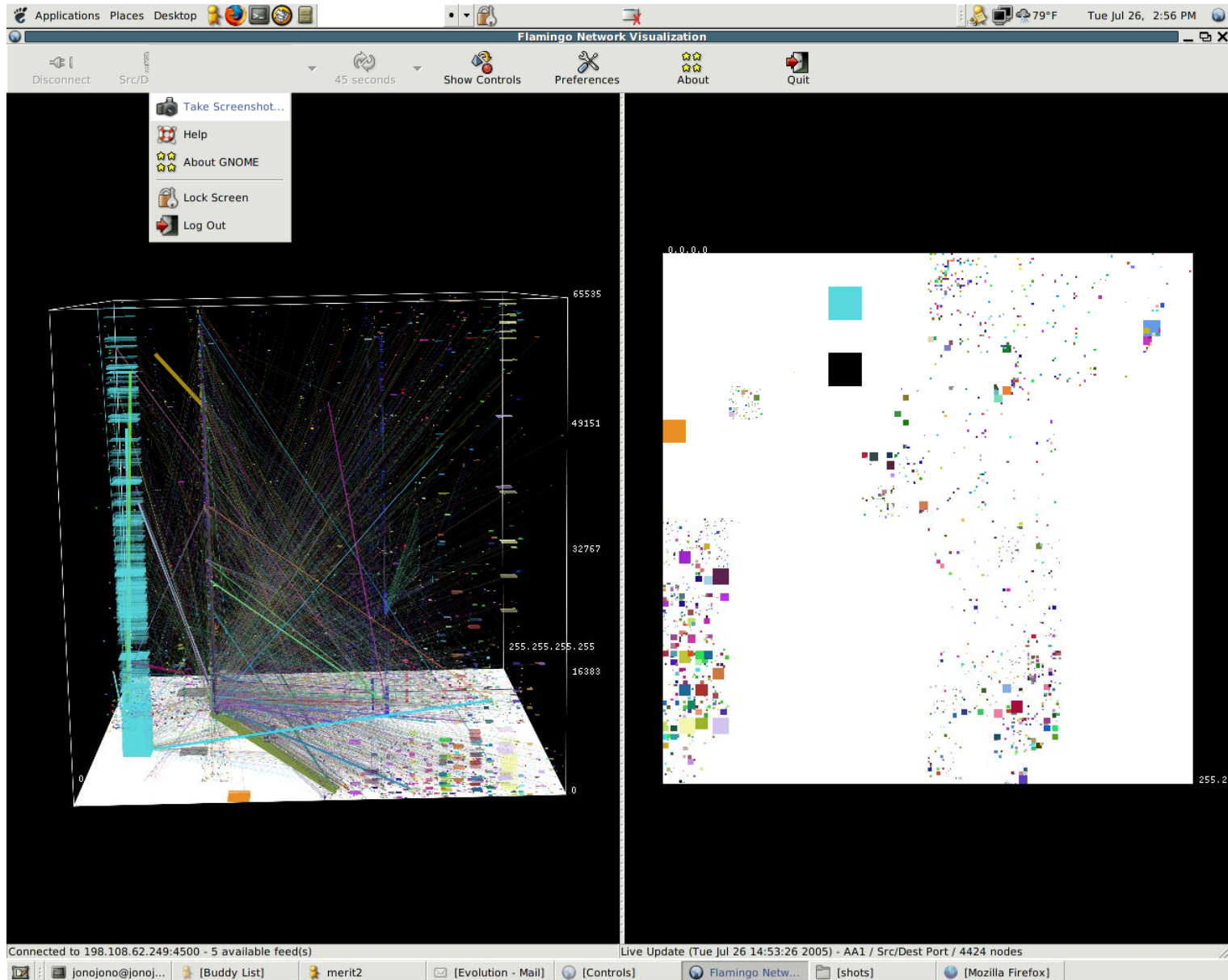
- Flows contain source/destination port number information as well
- Solution:
  - Use base of cube to represent prefixes, both source and destination are on the same base
  - The z-axis is used to represent port numbers, source and destination
  - (srcIP, srcPort) >>>>>>>>>> ((x1,y1), z1)
  - (dstIP, 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

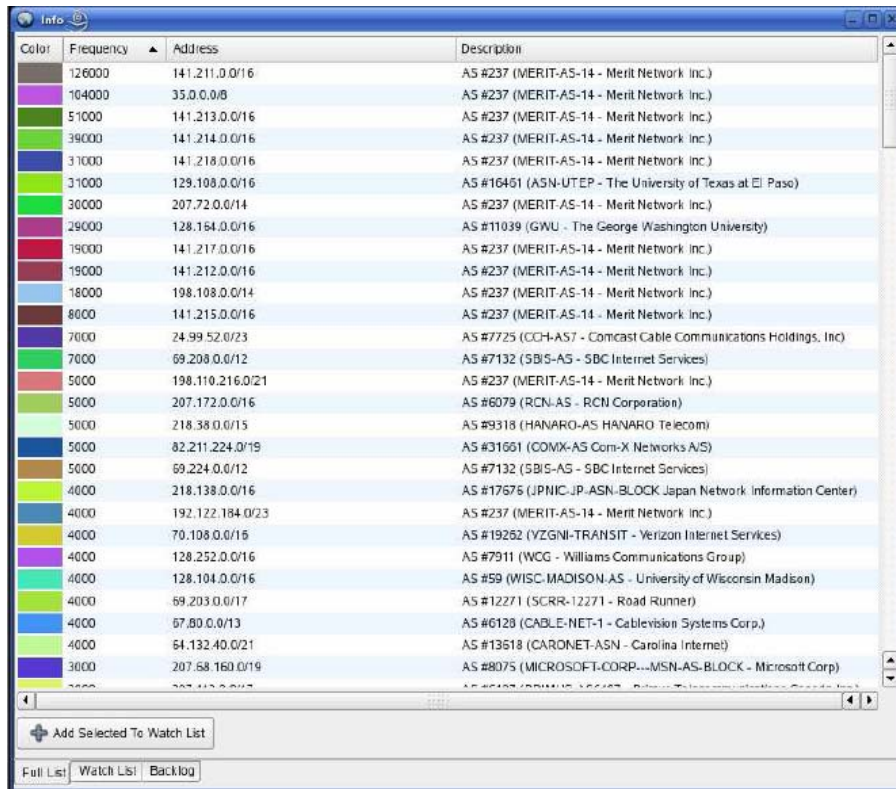


# Flamingo Visualization Tool



# Flamingo Controls

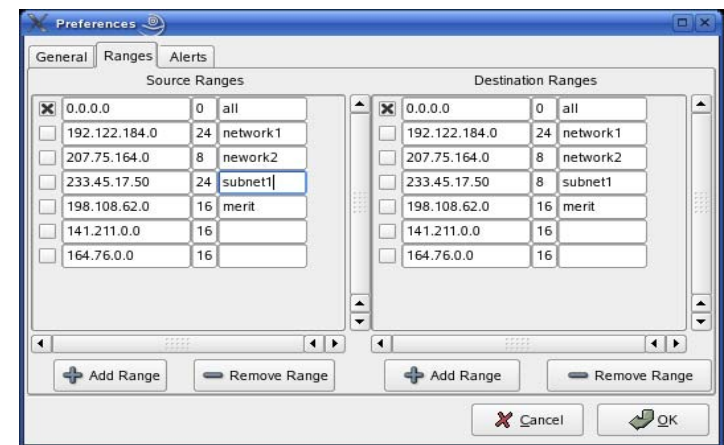
## Text Representation of Visualized Information



The 'Info' window displays a table with the following columns: Color, Frequency, Address, and Description. The table lists various network entries with their corresponding colors, frequencies, addresses, and descriptions.

Color	Frequency	Address	Description
126000	141.211.0.0/16	AS #237 (MERIT-AS-14 - Merit Network Inc.)	
104000	35.0.0.0/8	AS #237 (MERIT-AS-14 - Merit Network Inc.)	
51000	141.213.0.0/16	AS #237 (MERIT-AS-14 - Merit Network Inc.)	
39000	141.214.0.0/16	AS #237 (MERIT-AS-14 - Merit Network Inc.)	
31000	141.218.0.0/16	AS #237 (MERIT-AS-14 - Merit Network Inc.)	
31000	129.108.0.0/16	AS #16451 (ASN-UTEP - The University of Texas at El Paso)	
30000	207.72.0.0/14	AS #237 (MERIT-AS-14 - Merit Network Inc.)	
29000	128.184.0.0/16	AS #11039 (GWU - The George Washington University)	
19000	141.217.0.0/16	AS #237 (MERIT-AS-14 - Merit Network Inc.)	
19000	141.212.0.0/16	AS #237 (MERIT-AS-14 - Merit Network Inc.)	
18000	198.108.0.0/14	AS #237 (MERIT-AS-14 - Merit Network Inc.)	
8000	141.215.0.0/16	AS #237 (MERIT-AS-14 - Merit Network Inc.)	
7000	24.99.52.0/23	AS #7725 (CCH-AS7 - Comcast Cable Communications Holdings, Inc)	
7000	69.208.0.0/12	AS #7132 (SBS-AS - SBC Internet Services)	
5000	198.110.216.0/21	AS #237 (MERIT-AS-14 - Merit Network Inc.)	
5000	207.172.0.0/16	AS #6079 (RCN-AS - RCN Corporation)	
5000	218.38.0.0/15	AS #9318 (HANARO-AS HANARO Telecom)	
5000	82.211.224.0/19	AS #31651 (COMX-AS Com-X Networks A/S)	
5000	69.224.0.0/12	AS #7132 (SBS-AS - SBC Internet Services)	
4000	218.138.0.0/16	AS #17676 (JPNIC-JP-ASN-BLOCK Japan Network Information Center)	
4000	192.122.184.0/23	AS #237 (MERIT-AS-14 - Merit Network Inc.)	
4000	70.108.0.0/15	AS #19262 (YZGNI-TRANSIT - Verizon Internet Services)	
4000	128.252.0.0/16	AS #7911 (WCG - Williams Communications Group)	
4000	128.104.0.0/16	AS #59 (WISC-MADISON-AS - University of Wisconsin Madison)	
4000	69.203.0.0/17	AS #12271 (SCRR-12271 - Road Runner)	
4000	67.80.0.0/13	AS #6128 (CABLE-NET-1 - Cablevision Systems Corp.)	
4000	64.132.40.0/21	AS #13618 (CARONET-ASN - Carolina Internet)	
3000	207.68.160.0/19	AS #8075 (MICROSOFT-CORP...-MSN-AS-BLOCK - Microsoft Corp)	

## Slider Bar Controls



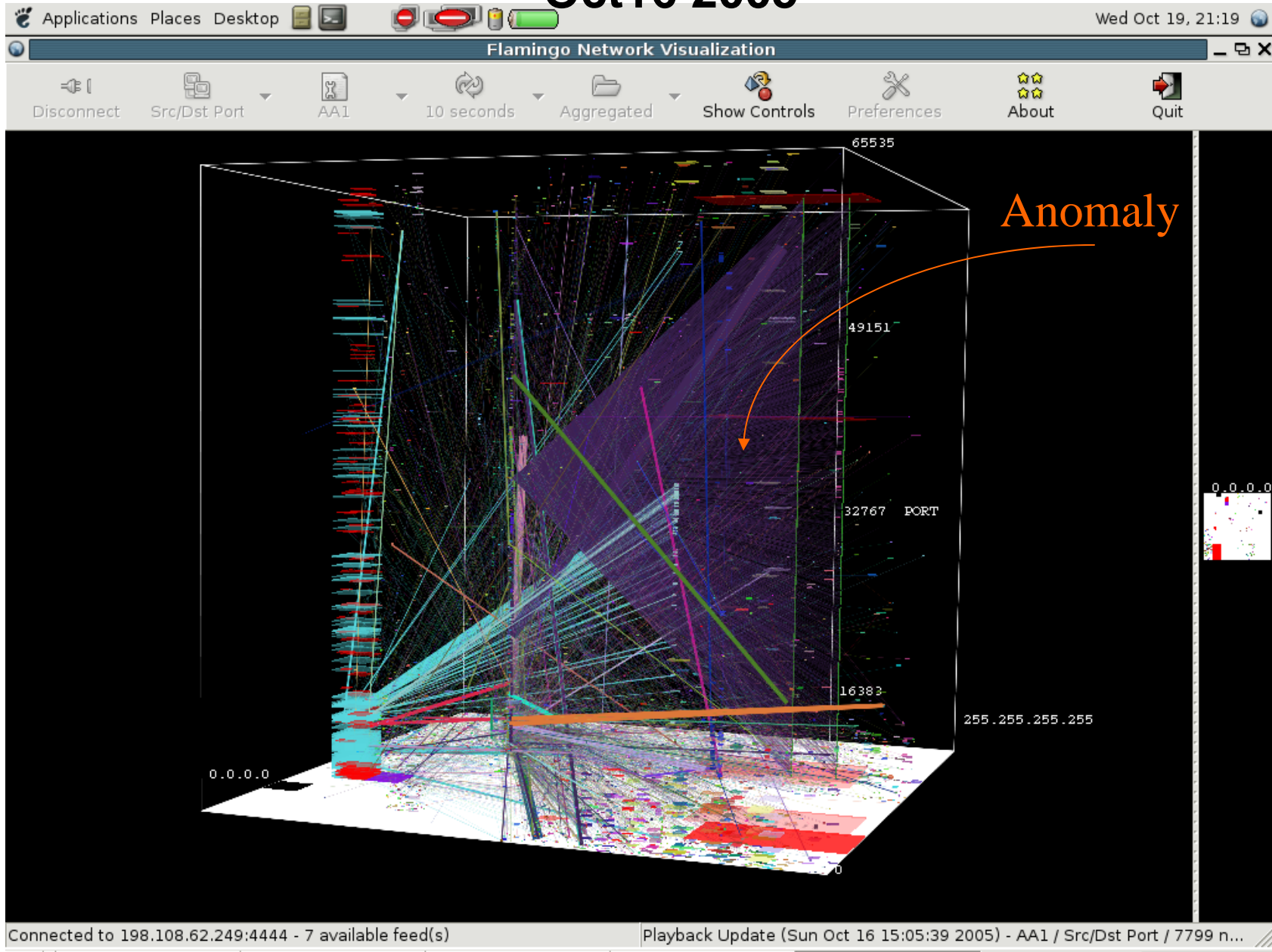
## Address Range Configuration

# Case Study: Traffic Anomaly Sunday- Oct 16, 2005

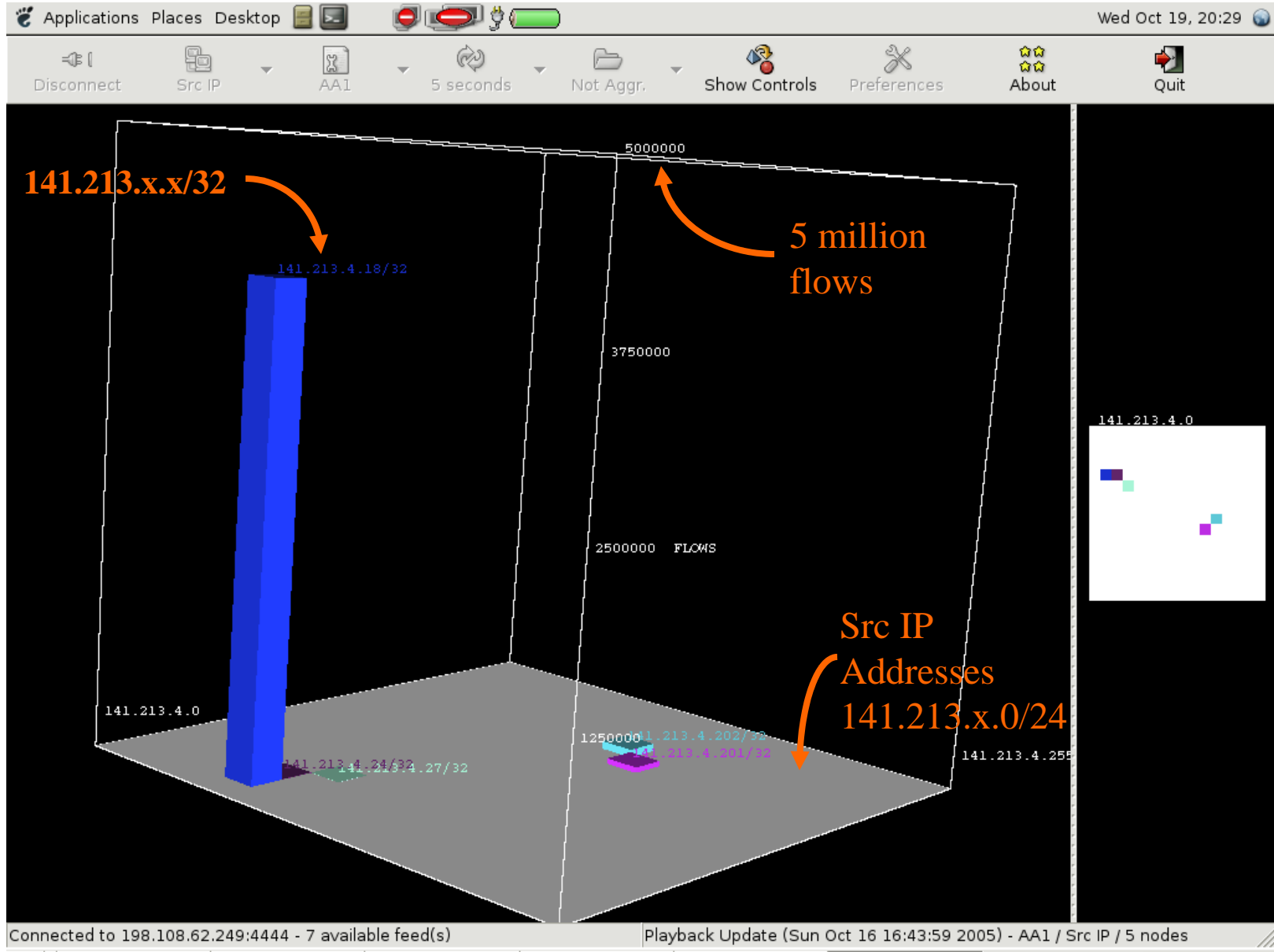
- Large burst of traffic visible outgoing from 141.213.x.x(x.x.umich.edu)
- Start time roughly – 12PM - End time roughly – 6PM
- Single srcIP/port – few(4) targetIP's/multiple ports
- UDP flows
- Traffic pattern visible in the normal clutter
- We then proceed to examine the src (141.213.x.0/24) and target prefixes (216.74.128.0/18, 217.199.32.0/19) in more detail in the following sequence of images



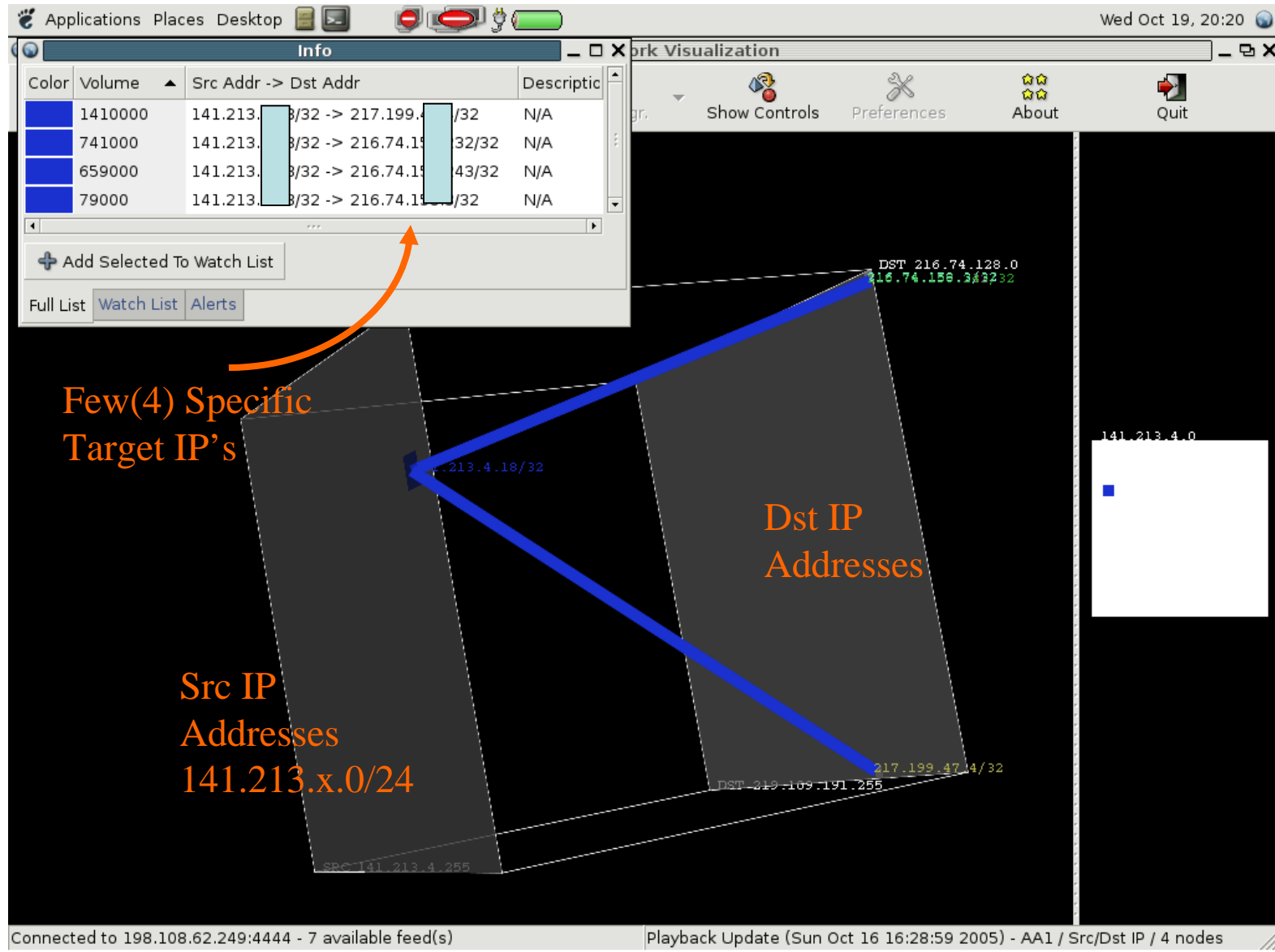
# Overall Traffic Pattern at Primary Router Sunday, Oct16 2005



# Traffic Volume sourced from /24 subnet by individual hosts



# Distribution of Target IP Addresses



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

Applications Places Desktop Wed Oct 19, 20:26

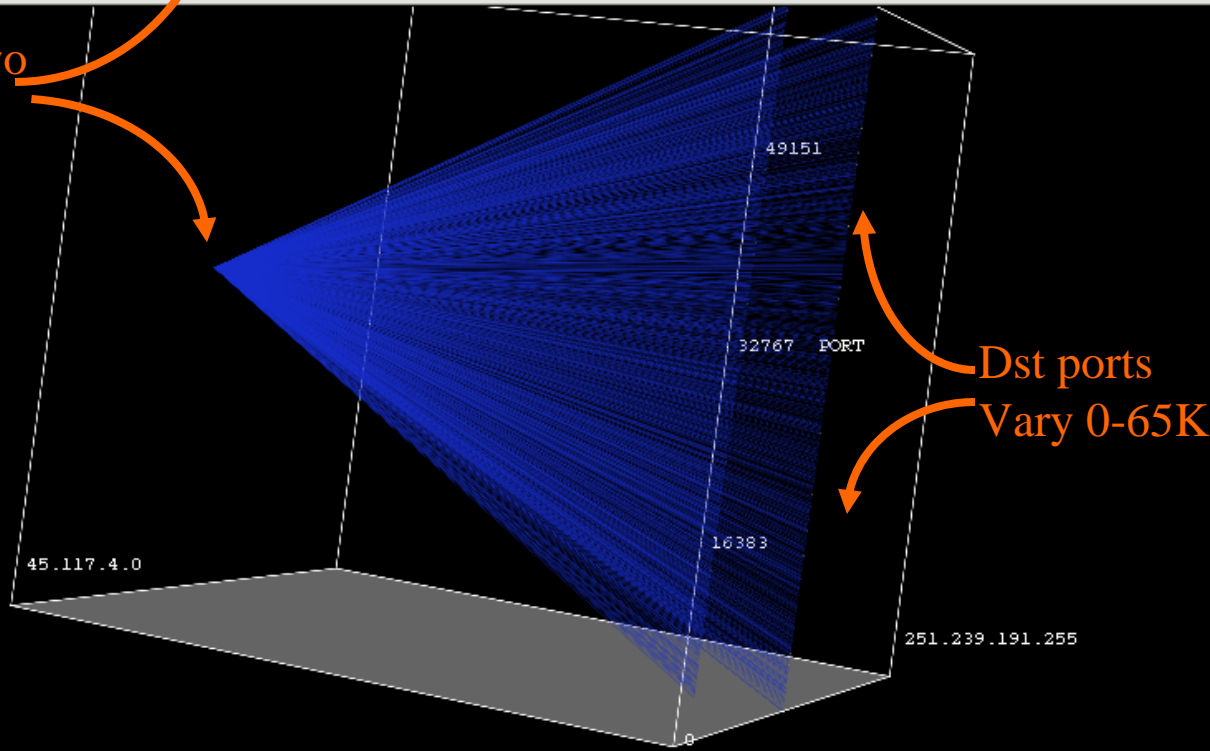
Info

Color	Volume	Address:Port -> Address:Port	Description
Blue	2000	141.213.118/32:32821 -> 216.74.115/32:16120	Proto 17 (UDP) - Src Port #32821 (Unknown Port) - Dst Port #16120 (Unknown Port)
Blue	2000	141.213.118/32:32822 -> 216.74.115/32:17079	Proto 17 (UDP) - Src Port #32822 (Unknown Port) - Dst Port #17079 (Unknown Port)
Blue	2000	141.213.118/32:32822 -> 216.74.115/32:17623	Proto 17 (UDP) - Src Port #32822 (Unknown Port) - Dst Port #17623 (Unknown Port)
Blue	2000	141.213.118/32:32822 -> 216.74.115/32:30527	Proto 17 (UDP) - Src Port #32822 (Unknown Port) - Dst Port #30527 (Unknown Port)
Blue	2000	141.213.118/32:32822 -> 216.74.115/32:42304	Proto 17 (UDP) - Src Port #32822 (Unknown Port) - Dst Port #42304 (Unknown Port)

+ Add Selected To Watch List

Full List Watch List Alerts

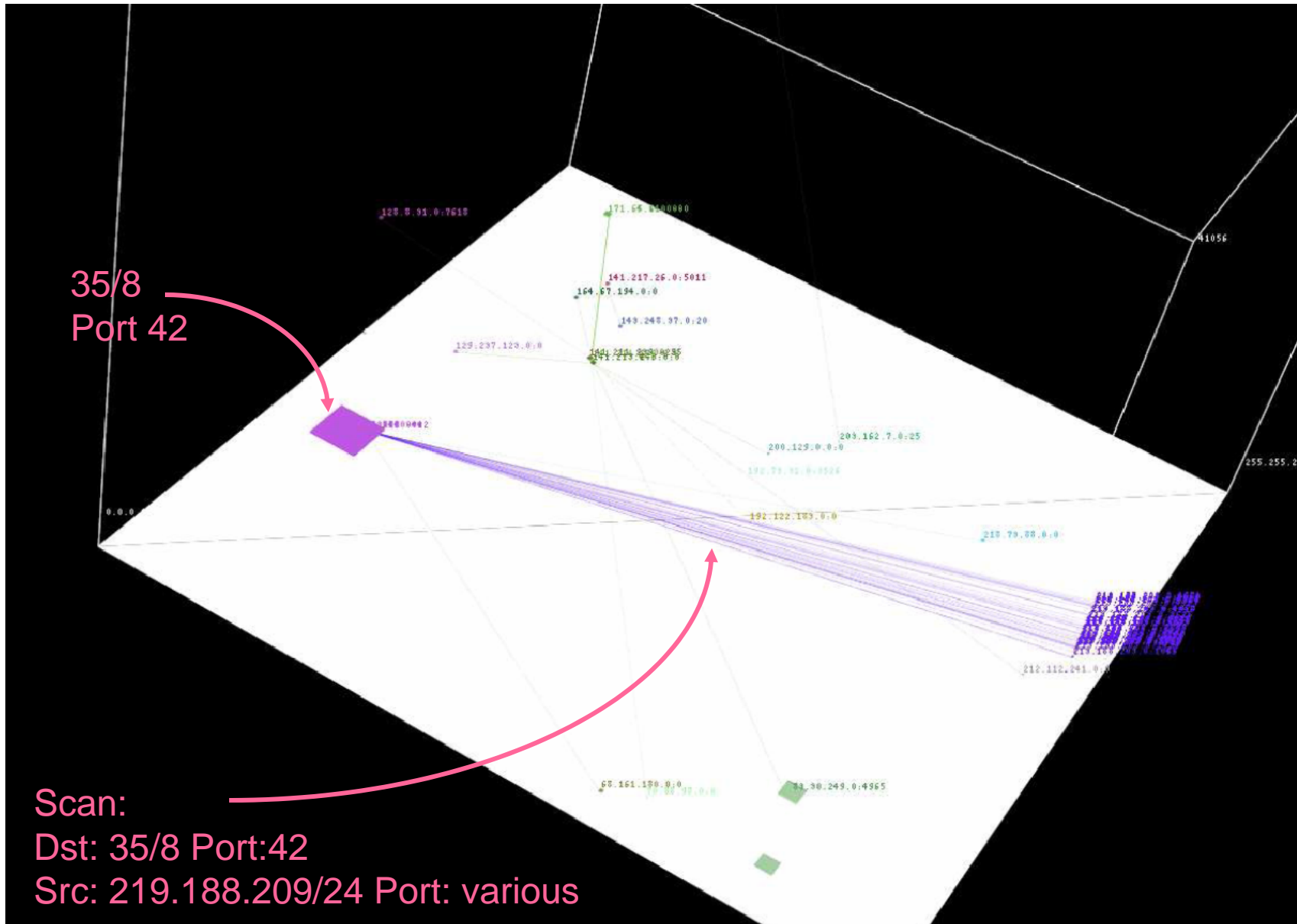
One or two Src Ports



Dst ports Vary 0-65K

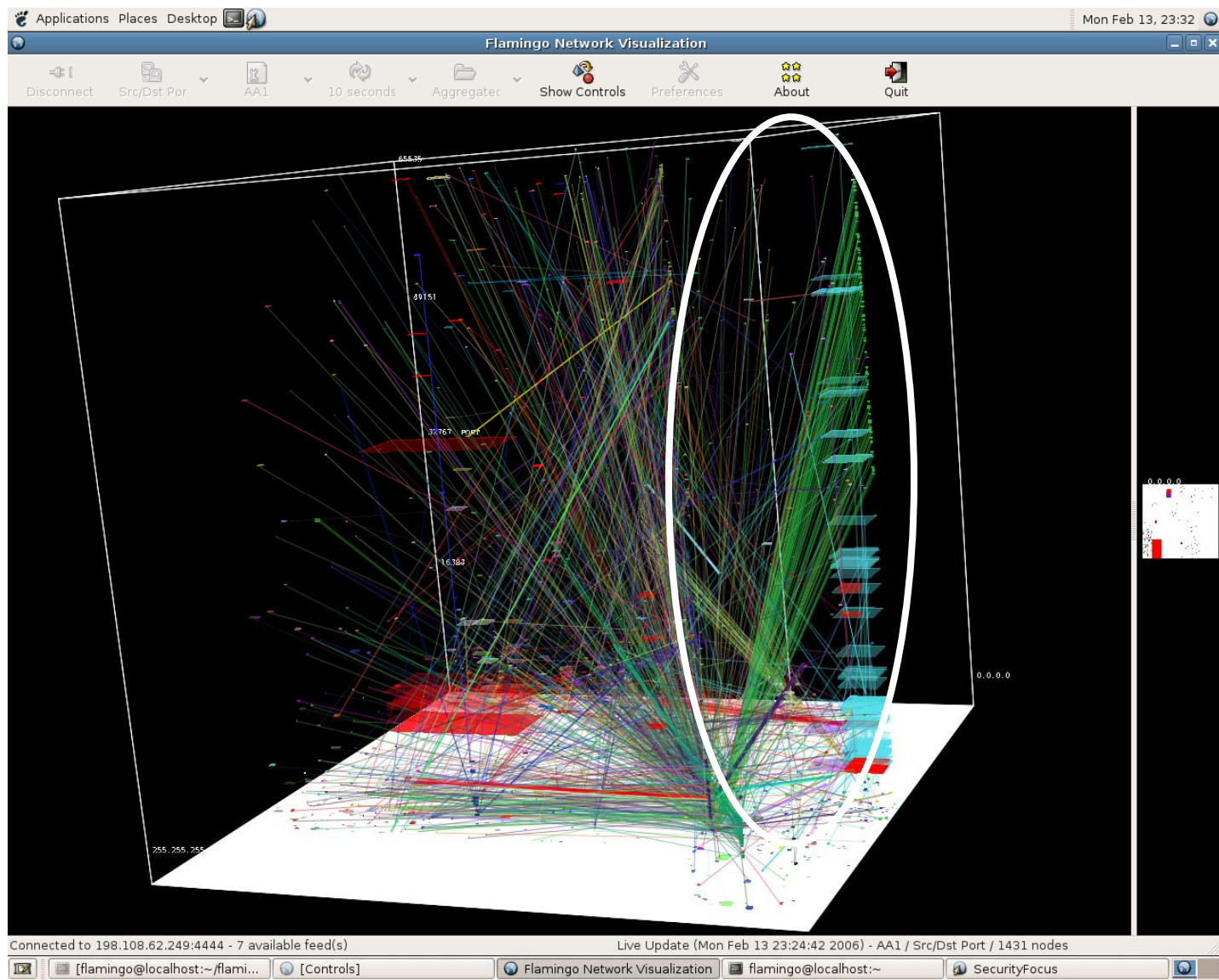


# Case Study: Worm Traffic/Port 42 Scans





# ssh scans



# ssh scan

Flamingo Network Visualization

Info

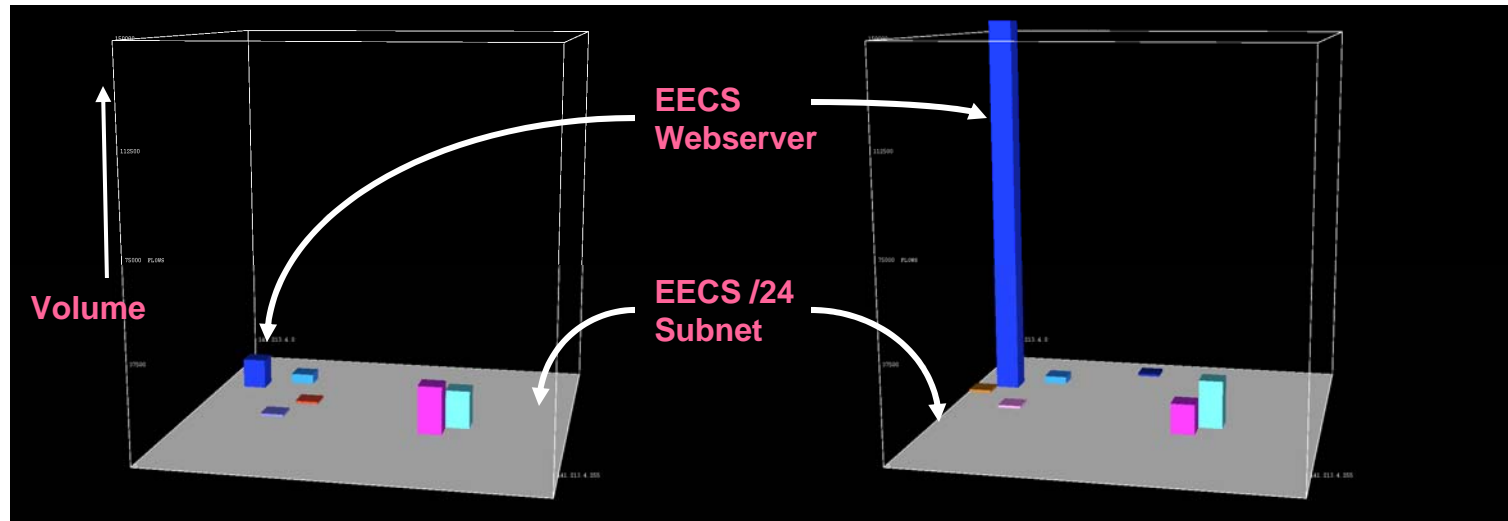
Color	Volume	Address:Port	Address:Port
1000	1000	128.12.0.0/16:35016 ->	141.211.0.0/16:22
1000	1000	128.12.0.0/16:47493 ->	141.211.0.0/16:22
1000	1000	128.12.0.0/16:36850 ->	141.211.0.0/16:22
1000	1000	128.12.0.0/16:35967 ->	141.211.0.0/16:22
1000	1000	128.12.0.0/16:38617 ->	141.211.0.0/16:22
1000	1000	128.12.0.0/16:46552 ->	141.211.0.0/16:22
1000	1000	128.12.0.0/16:46756 ->	141.211.0.0/16:22

Connected to 198.108.62.249:4444 - 7 available feed(s)

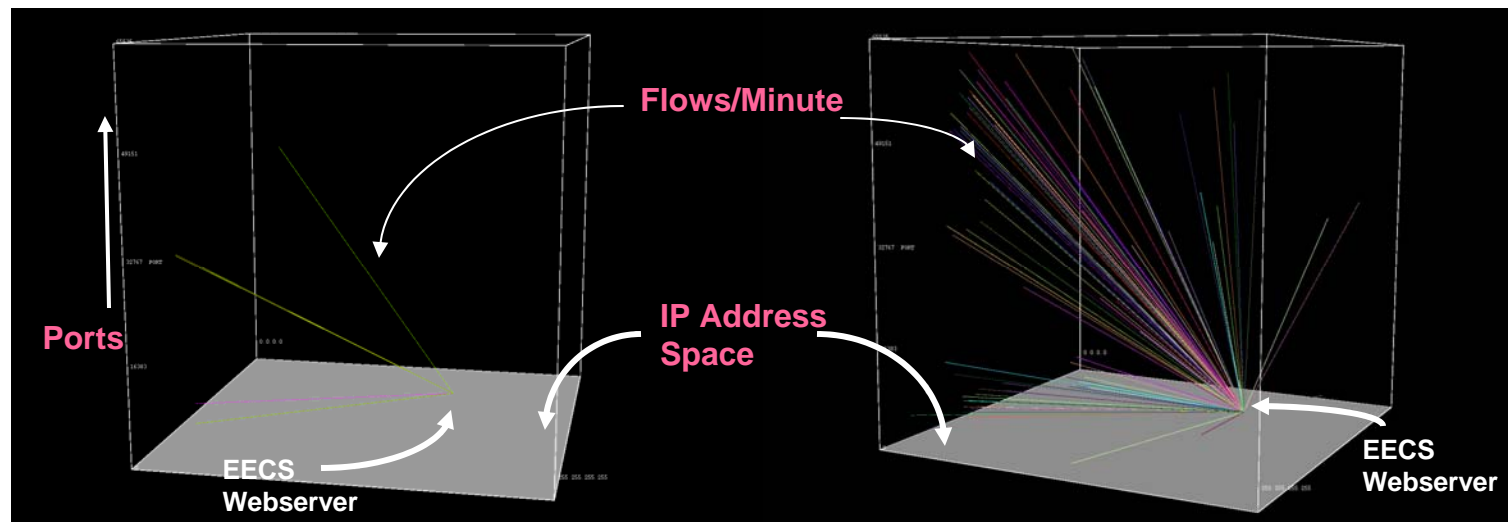
Live Update (Mon Feb 13 23:24:42 2006) - AA1 / Src/Dst Port / 1431 nodes

# Case Study: Slashdot Event Oct 31, 2004

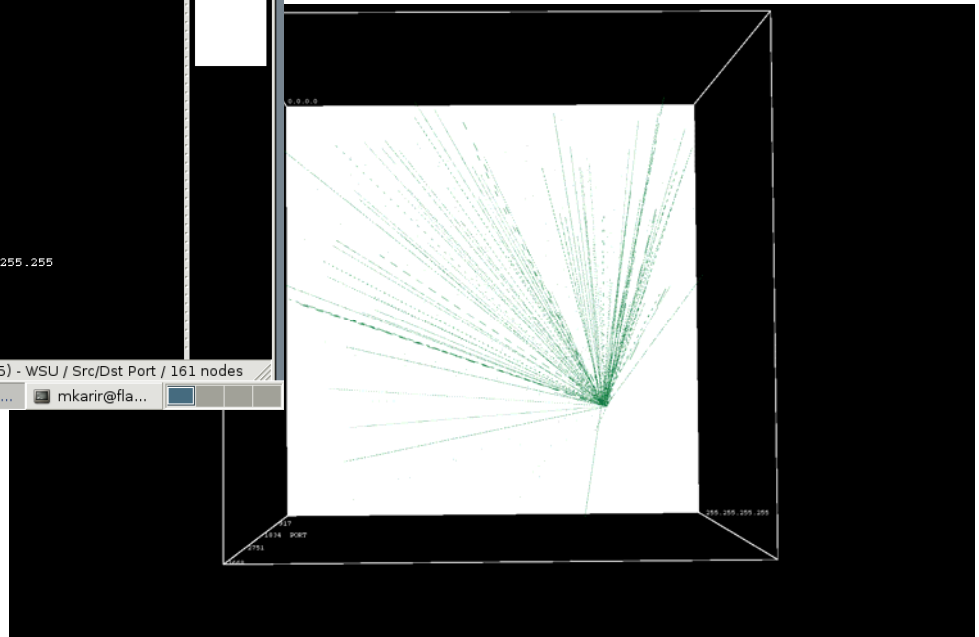
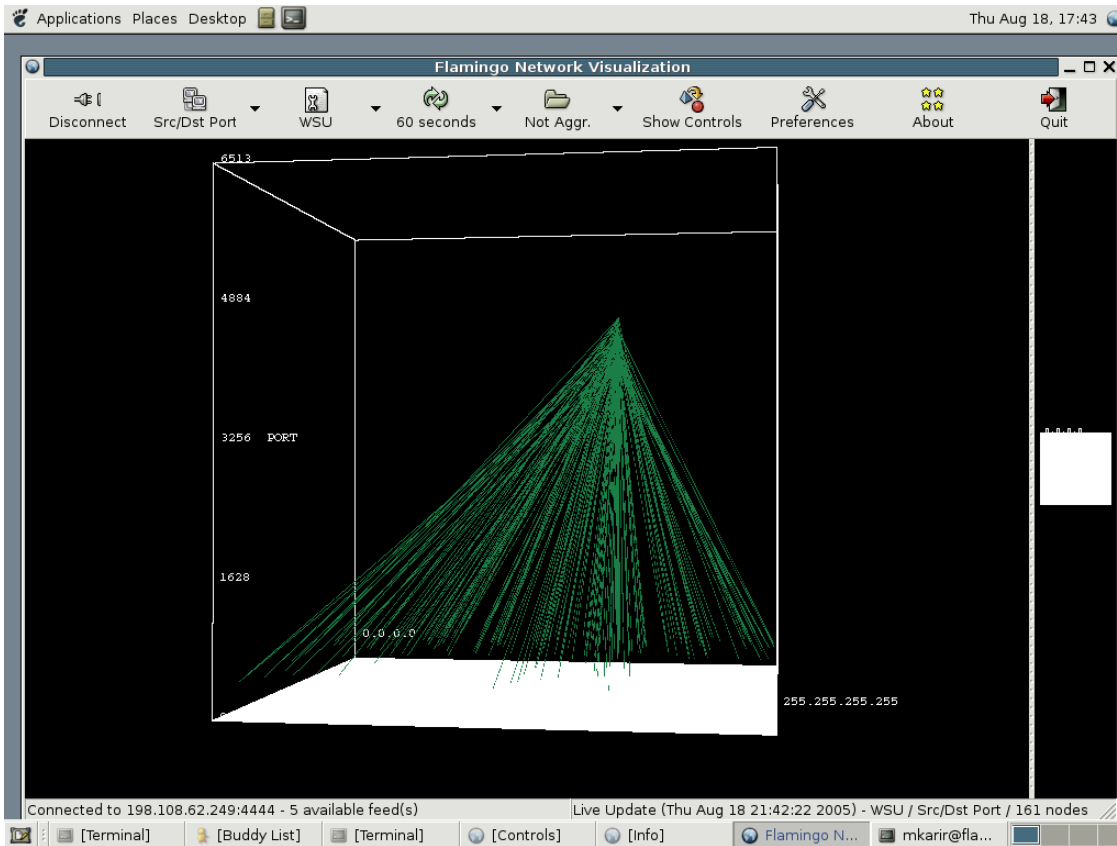
## Traffic Volume



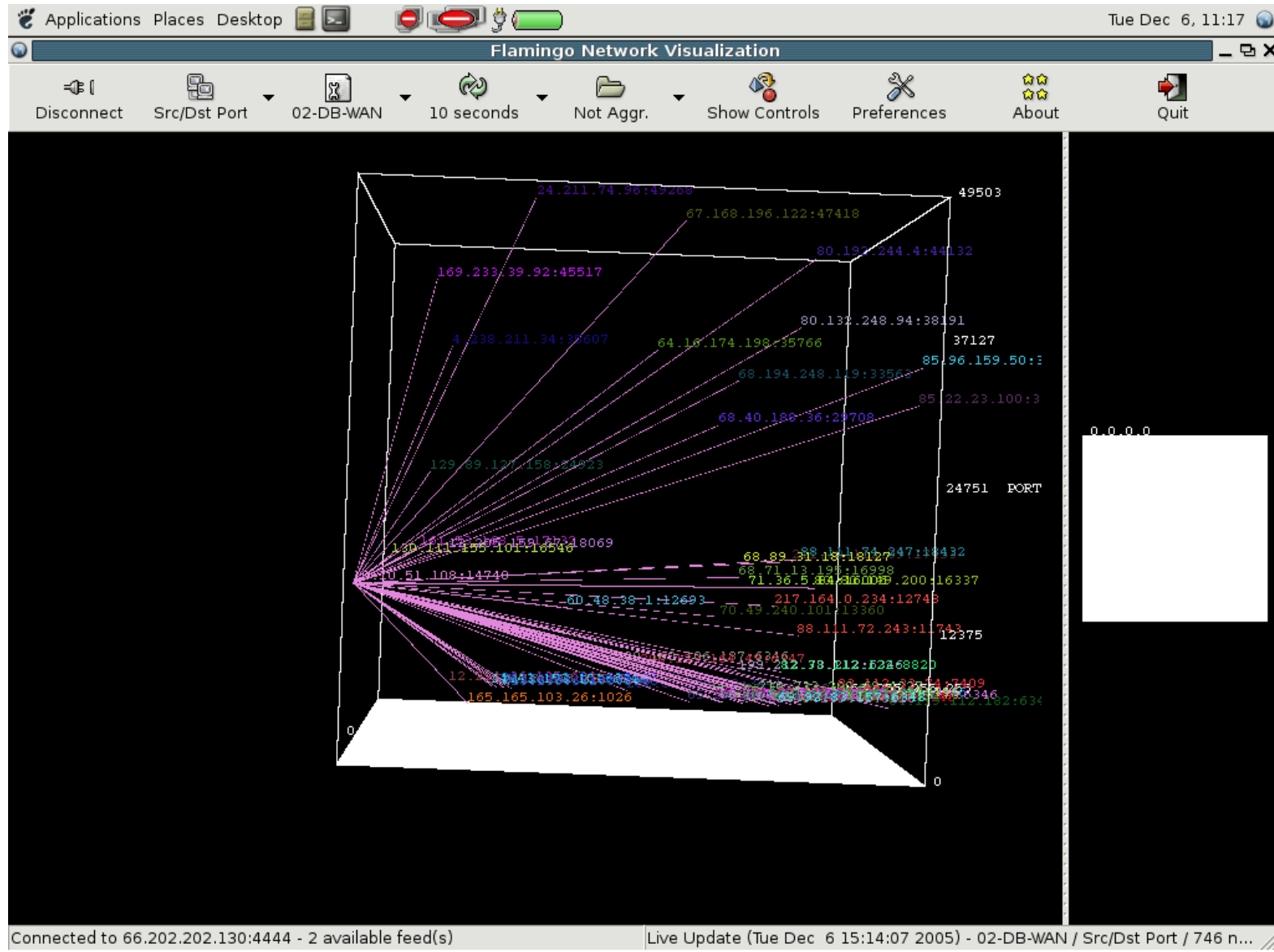
## Flow Volume



# Zotob Worm Infection

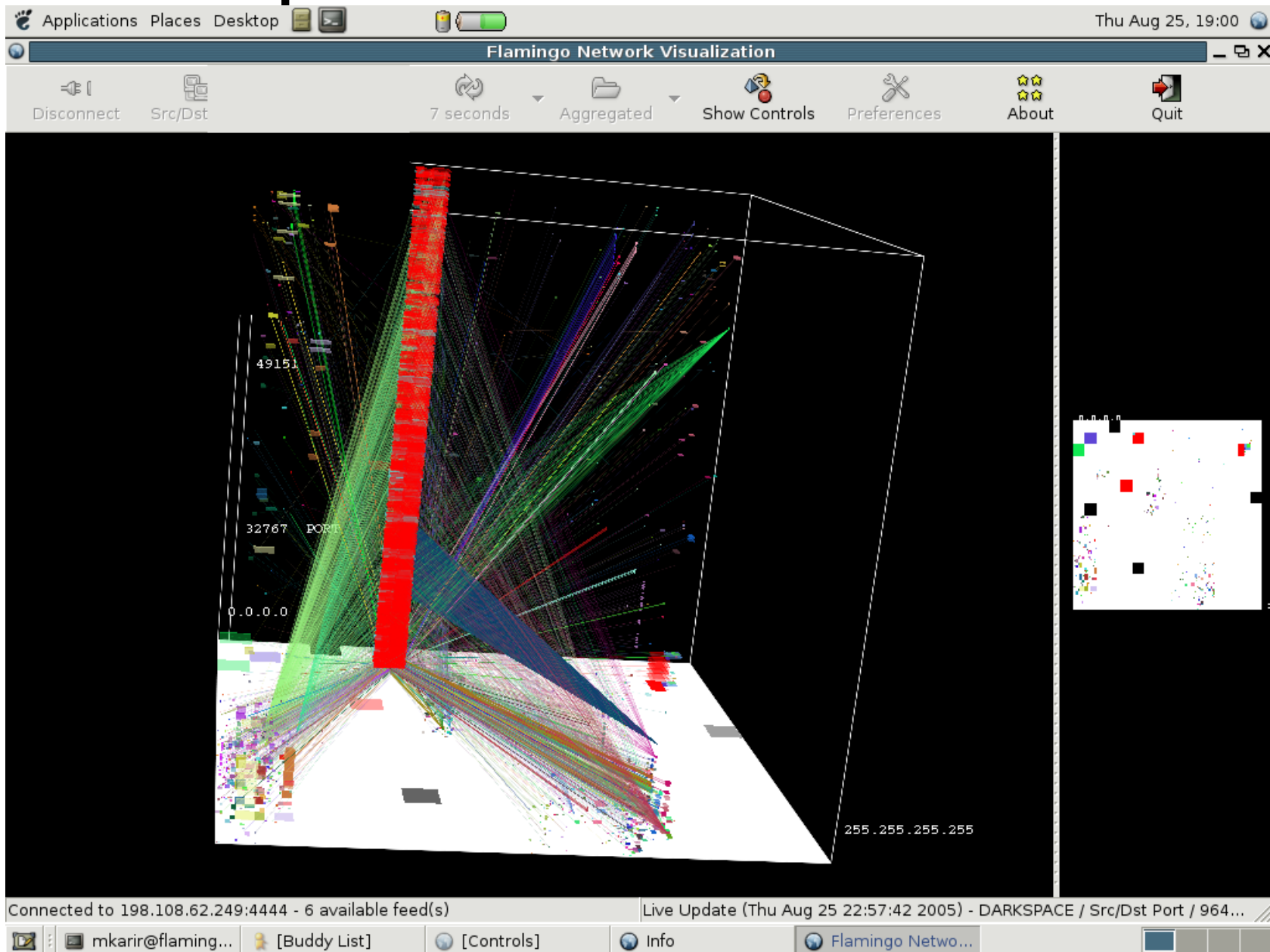


# P2P Traffic





# Dark-space Traffic Visualization





# Conclusion

- Obtaining raw data from networks is becoming easier, however using it effectively continues to be challenging
- Tools are the key in making raw data useful
- Operators often know what they are looking for, we should just make it easier for them to quickly get the information they need; there is some incident that is brought to their attention etc.
- Visualization can help to understand complex multi-dimensional data, instead of a database based query-response system which does not allow you to easily “explore” your data