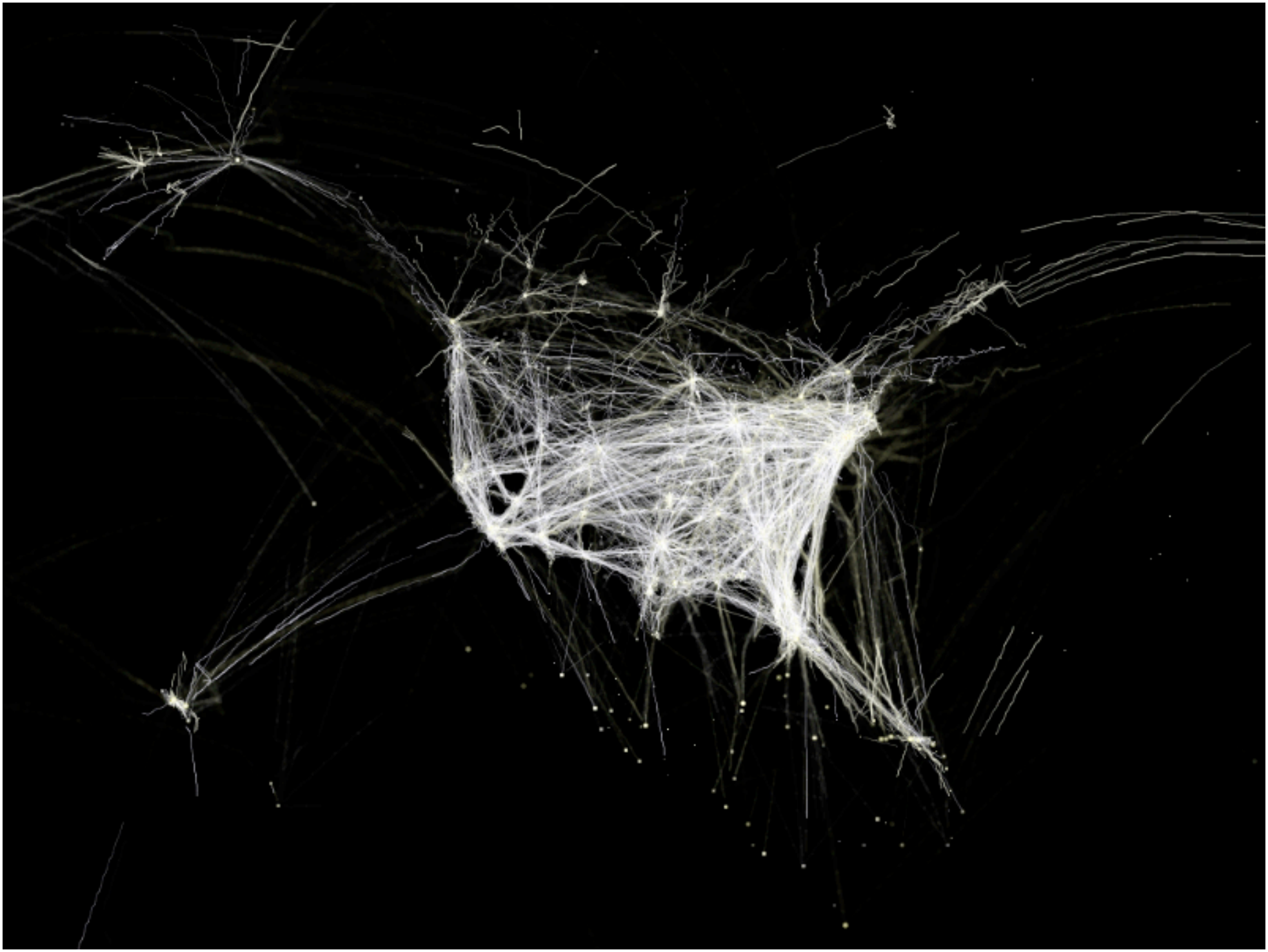
STANFORD HCI GROUP

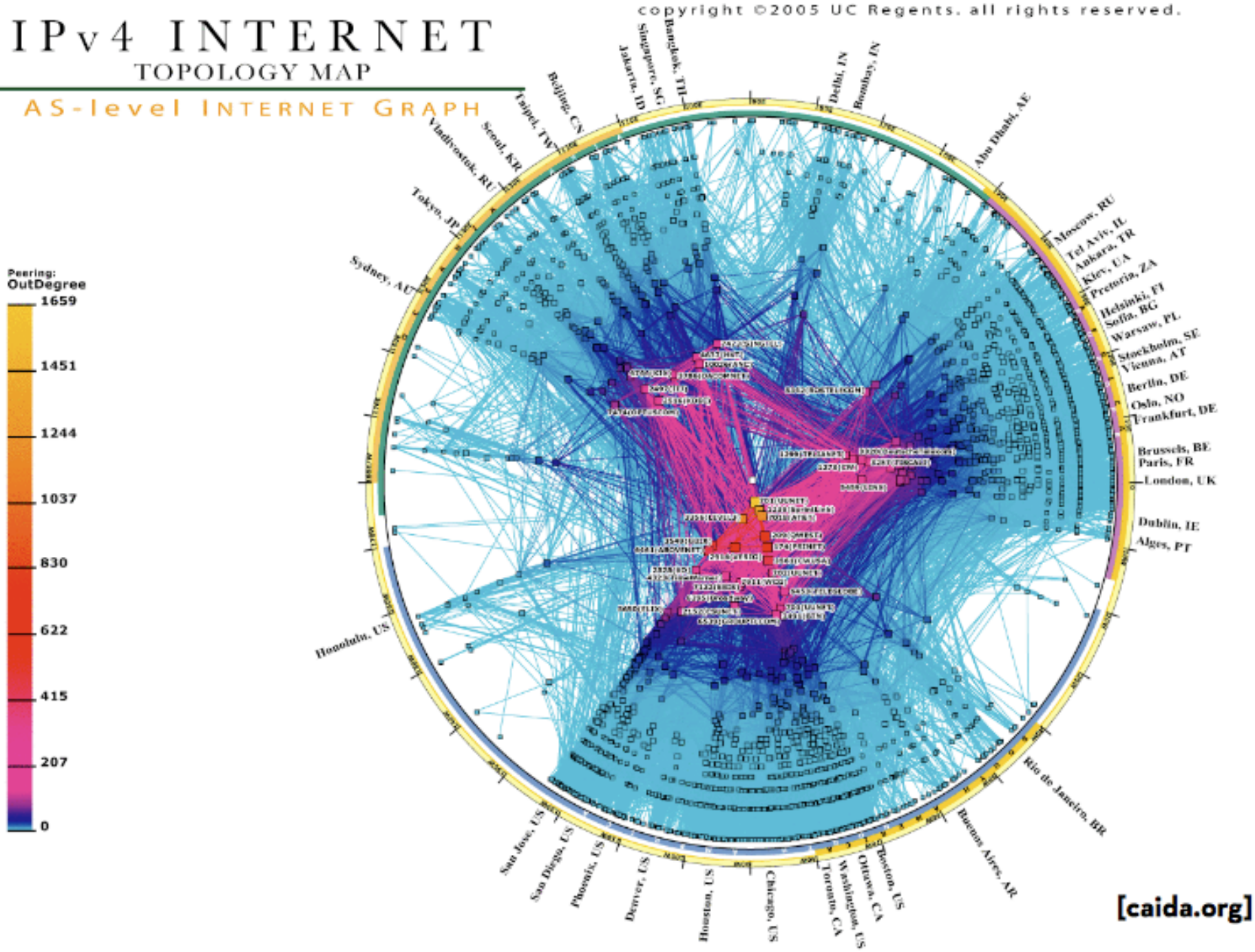
Visualization and Forensic Analysis of Network Events

Doantam Phan
Department of Computer Science
Stanford University



IPv4 INTERNET
TOPOLOGY MAP
AS-level INTERNET GRAPH

copyright ©2005 UC Regents. all rights reserved.



[caida.org]

Networks and Events

- Networks describe physical, digital, or other connections among members of a group
- Event describes a transaction between two members of the network at a specific time
  - A flight between two cities
  - A message between two computers

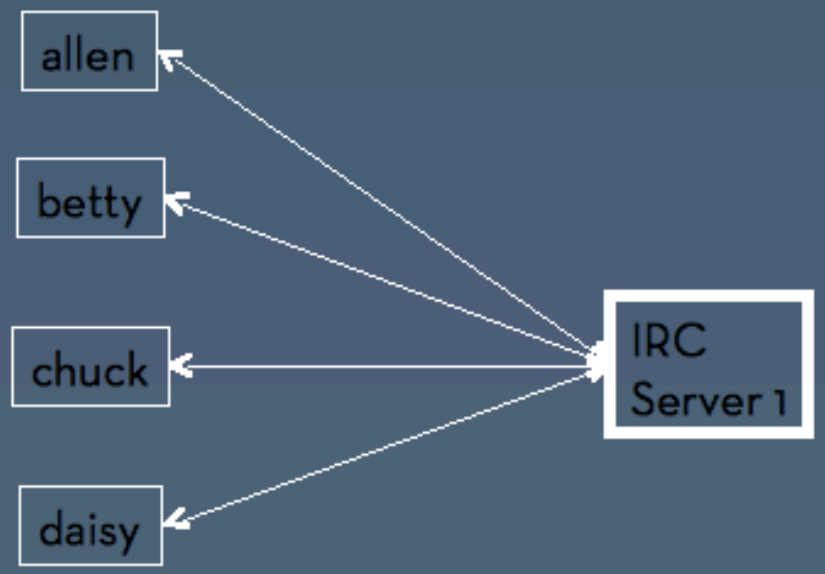
Forensic Analysis

- Unexpected events on the network
  - Flights are late
  - Computers get hacked
- Goal is to find sequence of events that explains the behavior

Chuck to Internet Relay Chat

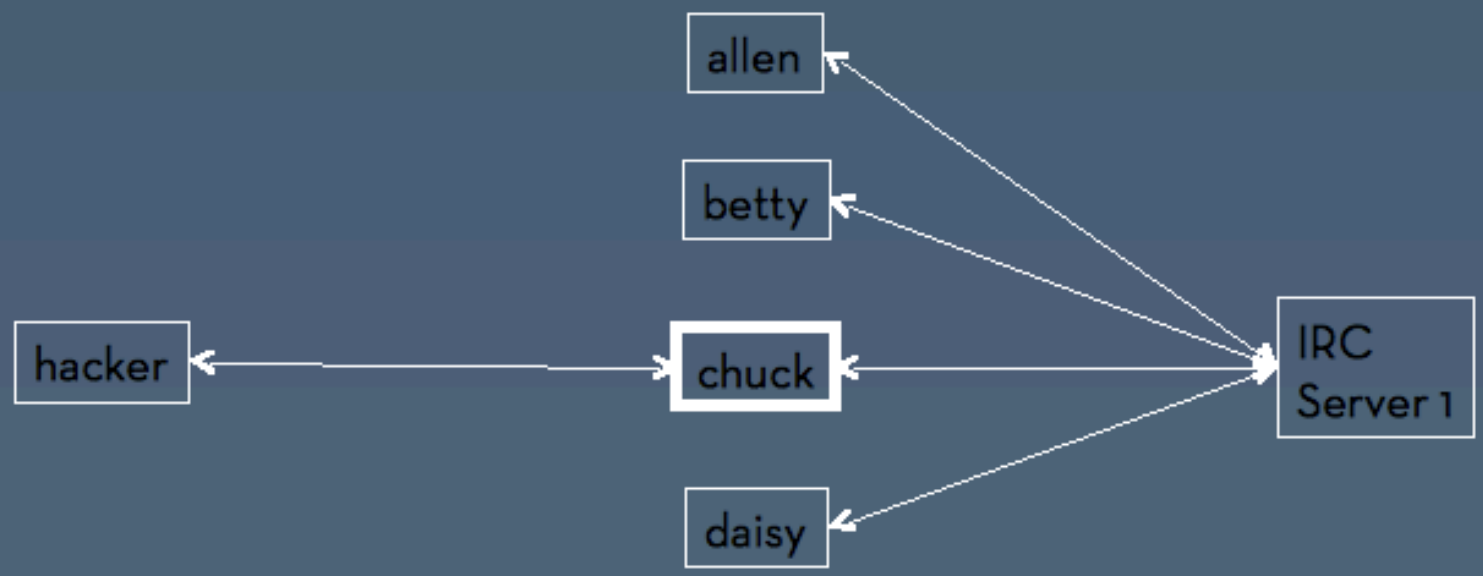
irc on port 6667

IRC to Other Computers



irc on port 6667

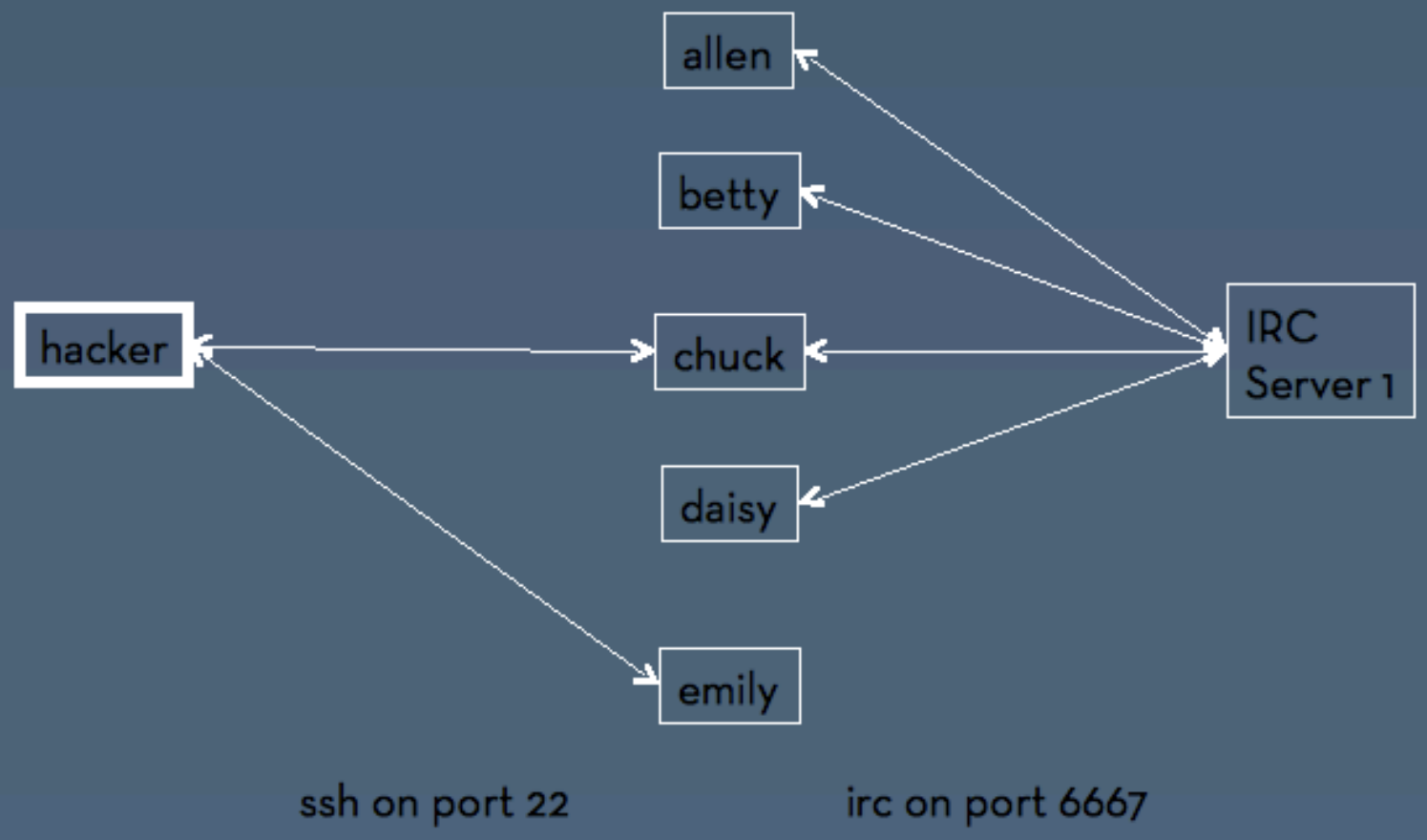
Chuck to Hacker



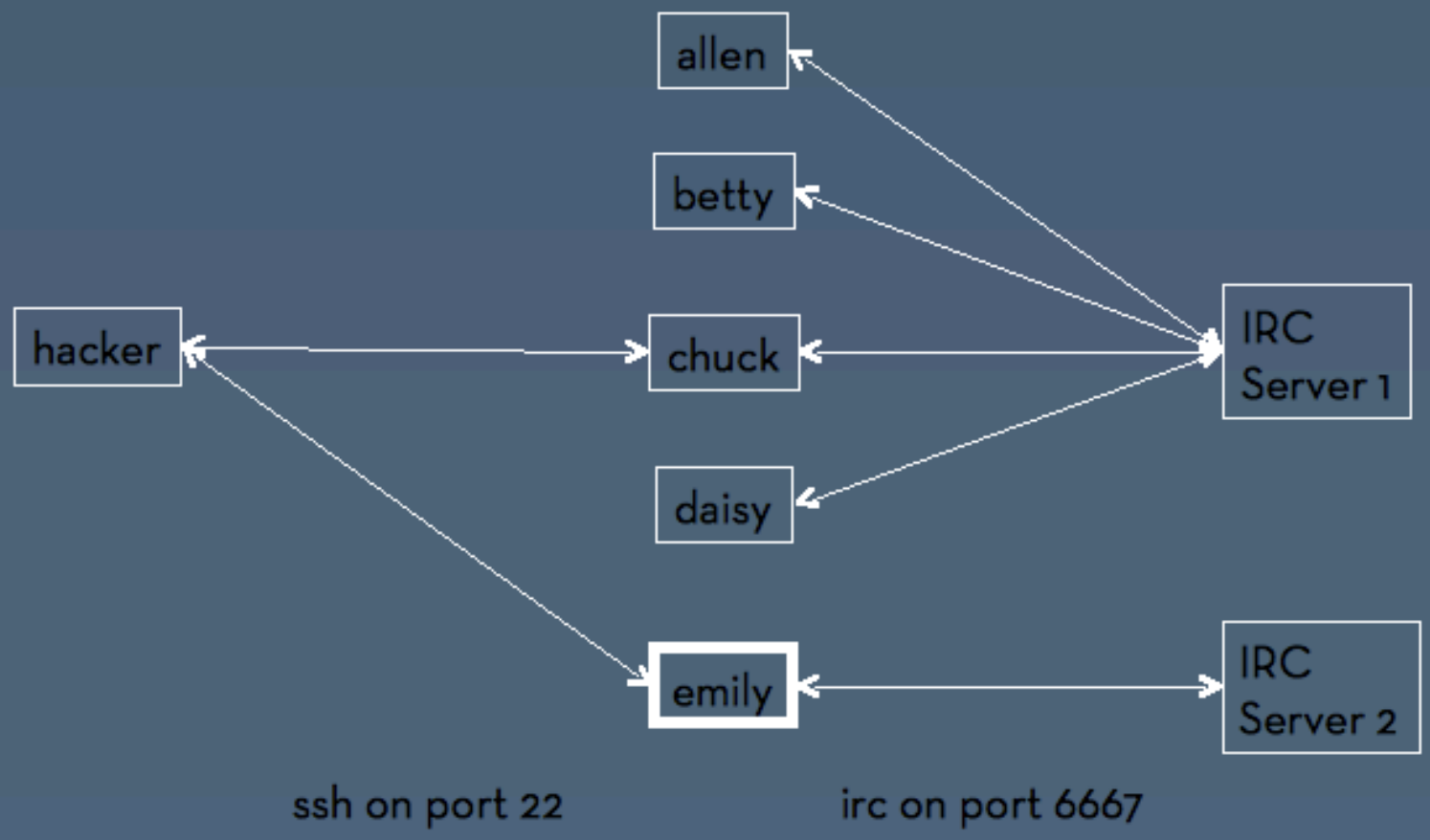
ssh on port 22

irc on port 6667

Hacker to Emily



Emily to IRC

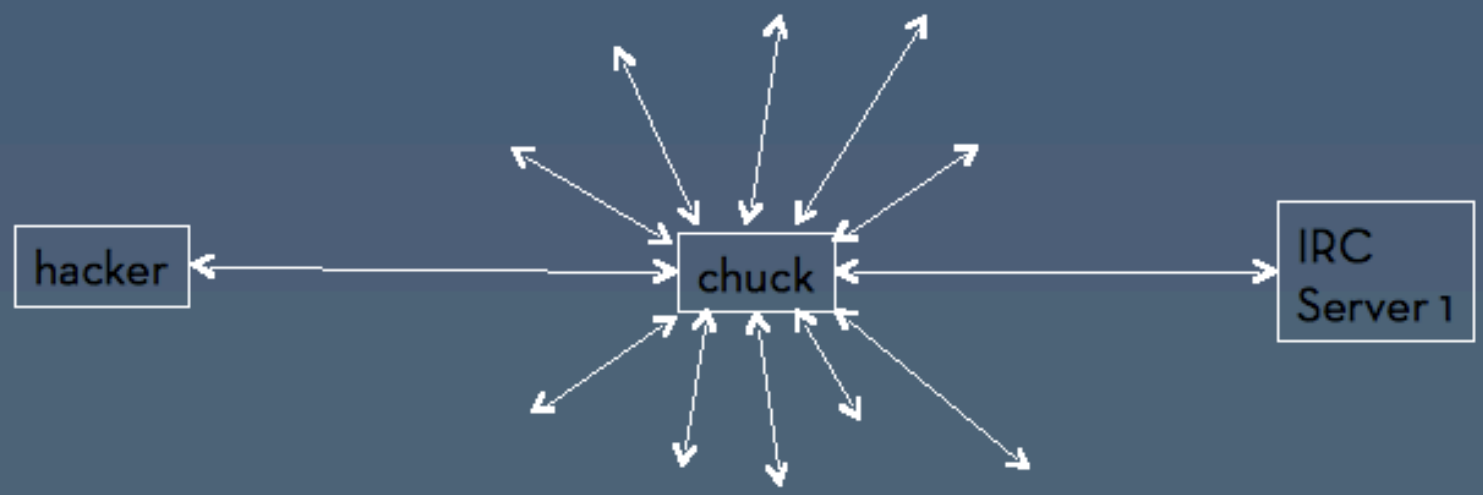

allen
betty
hacker
chuck
IRC Server 1
daisy
emily
IRC Server 2
ssh on port 22
irc on port 6667

Chuck to Hacker?

ssh on port 22

irc on port 6667

Chuck to Hacker?



ssh on port 22

irc on port 6667

THE REAL PROBLEM

Investigation is Hard

- Small number of relevant events that are hard to identify automatically
- Evidence is collected out-of-order, investigation process involves backtracking
- Hard for people to keep everything in working memory

Thesis Statement

Time-oriented visualizations of network activity that provide a visual exploration history and affordances for identifying relevant and related events can increase the effectiveness of forensic analysis.

Contributions

Interactive visualization techniques to analyze, organize, and present network events

1) Flow Map Layout for presentation

2) Progressive Multiples for analysis, and Isis, a system for network security

Overview

- Network Security
- Flow Maps
- Progressive Multiples
- Using Isis for Investigation
- Evaluation

NETWORK SECURITY

THE AT&T GLOBAL NETWORK
OPERATIONS CENTER

The Stanford Network

15-30 million flows per day to EE and CS buildings in MySQL database (1.5GB to 3GB)

Headers saved, not packet contents, since they can be encrypted

ID	Time	Local/Remote IP	Local/Remote Port	Measures

Network Security Tasks

Monitoring

Investigation

Reporting

Intrusion Detection Systems

Monitoring

- Signature Based
  - Bro [Paxson 1999]
  - Snort [Roesch 1999]
- Statistical Classification
  - BLINC *[Karagiannis et al. 2005]*

Network Security Tasks

Monitoring

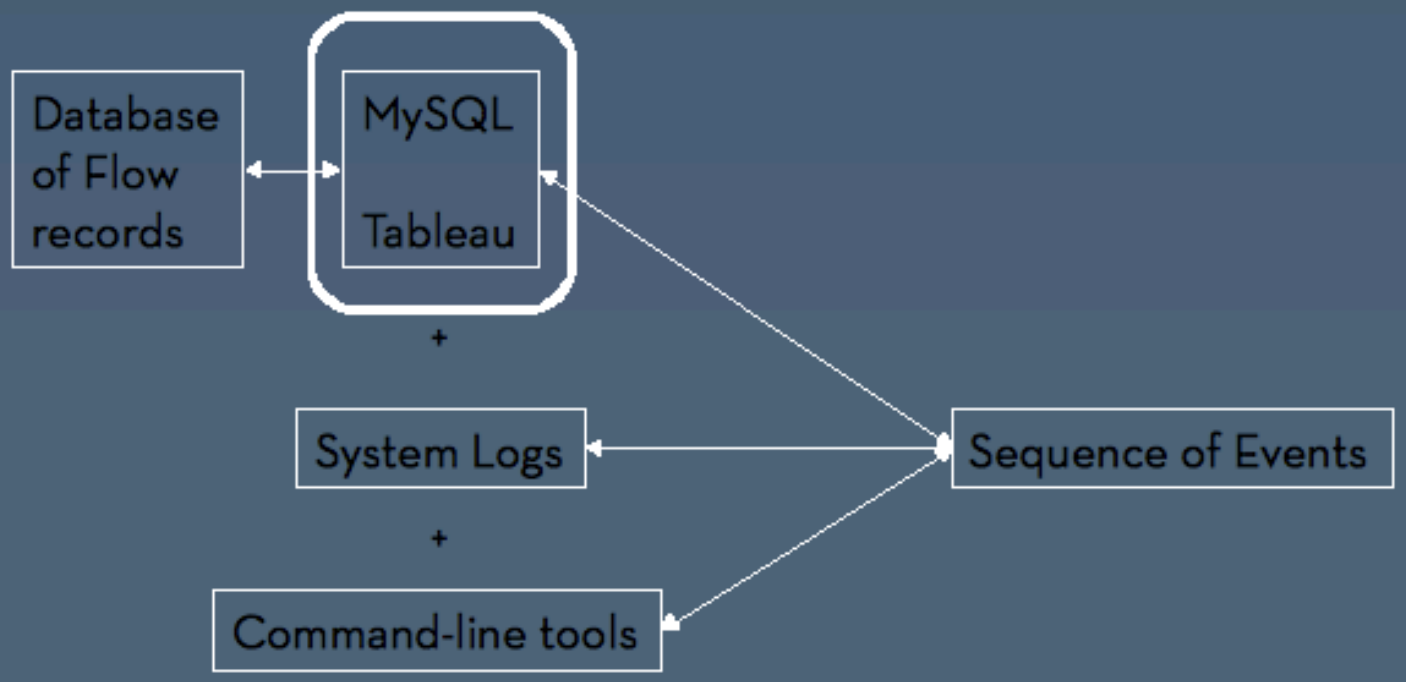
Investigation

Reporting

Case Study on Network Security

- Close collaboration for over 2 years
- 43 intrusions from 9/2005 to 11/2007; early observations to produce task analysis
- Deployed and refined expert system, Isis, for network security investigation

Previous Investigation Tools



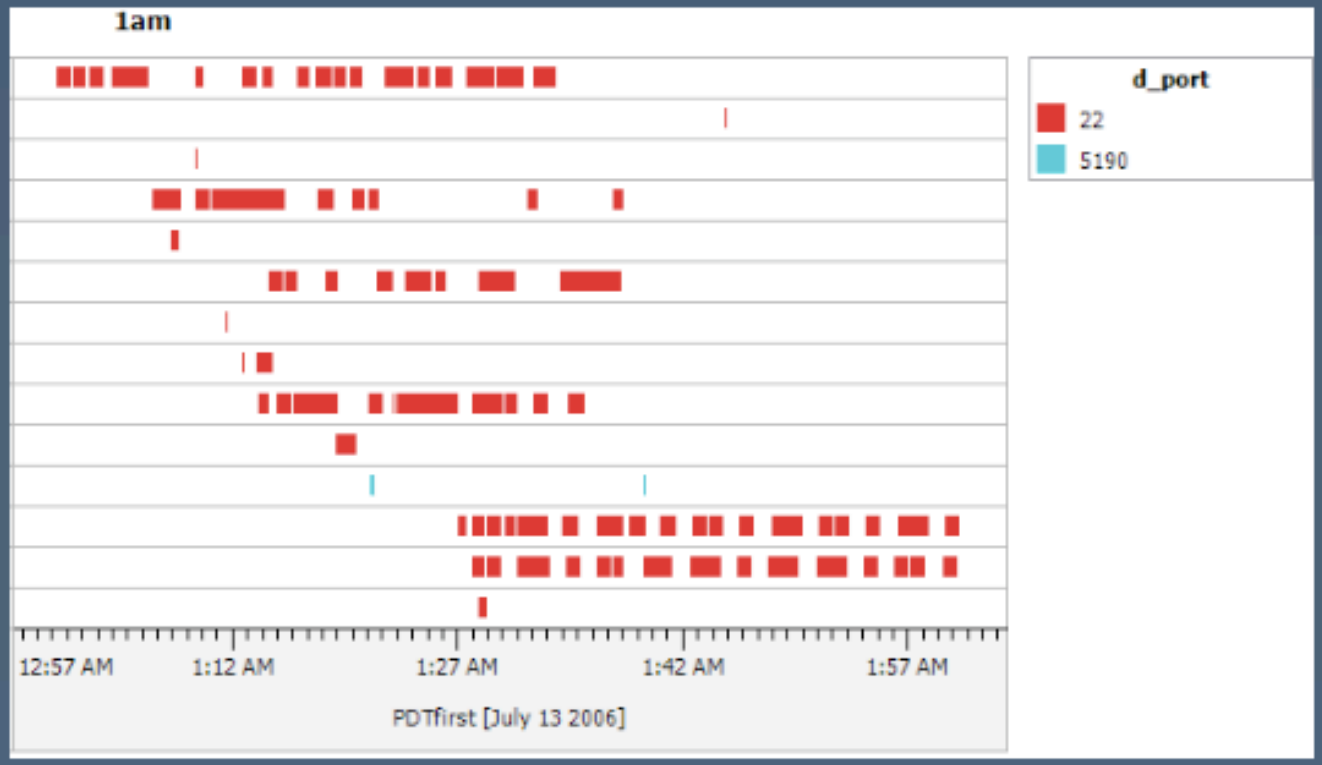
```
mysql> select min(gmtfirst), inet_ntoa(l_ipn), count(seq_id) , avg(l_abyte+r_abyte),
 unix_timestamp(gmtfirst) div 15 as b  from ee_2006_05_01   group
by b, l_ipn  with rollup having  b <  76429442 ;
```

min(gmtfirst)	inet_ntoa(l_ipn)	count(seq_id)	avg(l_abyte+r_abyte)	b
2006-04-30 23:59:59	75.64.90.151	2	103230.0000	76429439
2006-04-30 23:59:58	75.64.90.160	3	8813.3333	76429439
2006-04-30 23:59:57	75.64.92.11	1	80.0000	76429439
2006-04-30 23:59:56	75.64.92.13	1	98.0000	76429439
2006-04-30 23:59:59	75.64.93.172	1	180.0000	76429439
2006-04-30 23:59:57	75.64.94.43	2	338.0000	76429439
2006-04-30 23:59:59	75.64.94.64	1	0.0000	76429439
2006-04-30 23:59:59	77.232.66.45	1	2273.0000	76429439
2006-04-30 23:59:59	77.232.89.231	1	96.0000	76429439
2006-04-30 23:59:59	77.232.91.176	1	311.0000	76429439
2006-04-30 23:59:56	NULL	14	16901.0000	76429439
2006-05-01 00:00:00	75.64.73.64	8	367.5000	76429440
2006-05-01 00:00:14	75.64.88.69	1	0.0000	76429440
2006-05-01 00:00:10	75.64.89.188	1	25633.0000	76429440
2006-05-01 00:00:13	75.64.90.20	1	197.0000	76429440
2006-05-01 00:00:02	75.64.90.51	1	402.0000	76429440
2006-05-01 00:00:01	75.64.90.57	3	84349.3333	76429440
2006-05-01 00:00:05	75.64.90.74	1	1827960.0000	76429440
2006-05-01 00:00:12	75.64.90.133	2	89440.0000	76429440
2006-05-01 00:00:09	75.64.90.138	3	2831.6667	76429440
2006-05-01 00:00:01	75.64.90.141	1	0.0000	76429440
2006-05-01 00:00:10	75.64.90.142	1	1284.0000	76429440
2006-05-01 00:00:00	75.64.90.151	17	50202.0588	76429440
2006-05-01 00:00:01	75.64.90.160	3	233.3333	76429440
2006-05-01 00:00:03	75.64.90.165	9	233.3333	76429440
2006-05-01 00:00:05	75.64.90.173	2	144.0000	76429440
2006-05-01 00:00:02	75.64.90.194	5	166.0000	76429440
2006-05-01 00:00:03	75.64.90.222	2	192.0000	76429440
2006-05-01 00:00:07	75.64.90.244	3	5049.3333	76429440
2006-05-01 00:00:08	75.64.90.248	1	38.0000	76429440
2006-05-01 00:00:03	75.64.91.26	4	187734.0000	76429440
2006-05-01 00:00:08	75.64.91.31	1	768.0000	76429440
2006-05-01 00:00:08	75.64.91.81	3	2227.3333	76429440
2006-05-01 00:00:01	75.64.91.88	2	2334.0000	76429440
2006-05-01 00:00:08	75.64.91.123	1	121360.0000	76429440
2006-05-01 00:00:03	75.64.91.137	3	310881.3333	76429440
2006-05-01 00:00:04	75.64.91.209	1	120.0000	76429440
2006-05-01 00:00:01	75.64.91.211	1	672.0000	76429440
2006-05-01 00:00:01	75.64.91.214	1	96.0000	76429440
2006-05-01 00:00:18	77.232.94.201	1	159.0000	76429441
2006-05-01 00:00:15	NULL	265	30368.2189	76429441

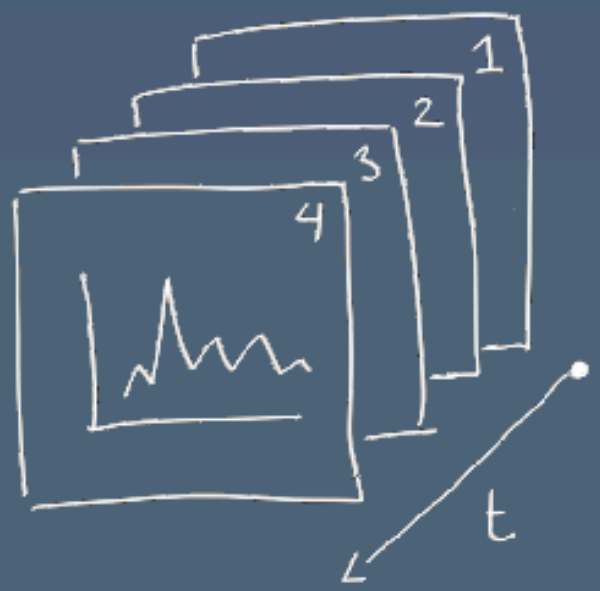
Tableau

[Stolte et al. 2000]

IP Addresses

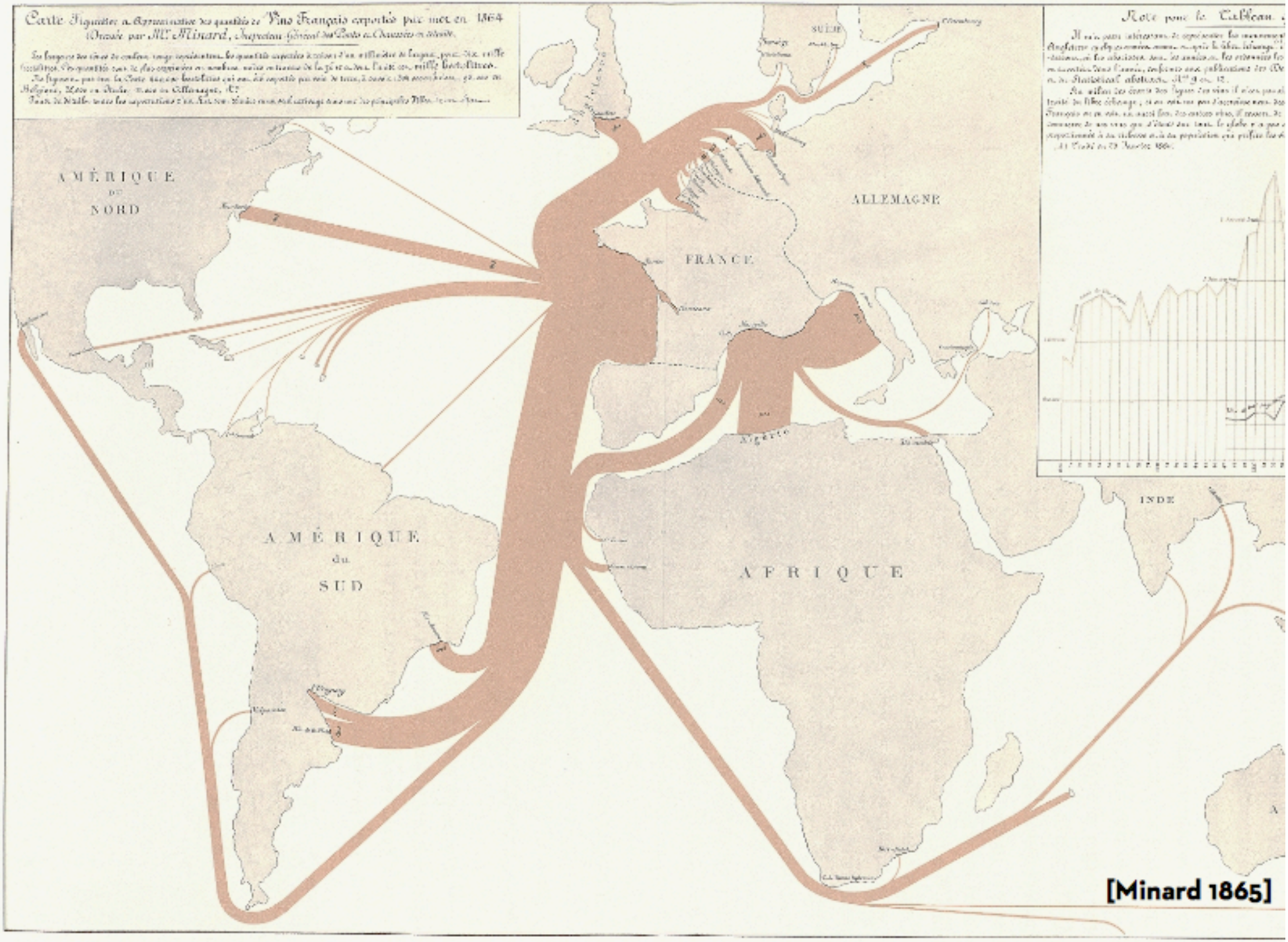


Remembering Previous States



Previous visualization states

FLOW MAPS

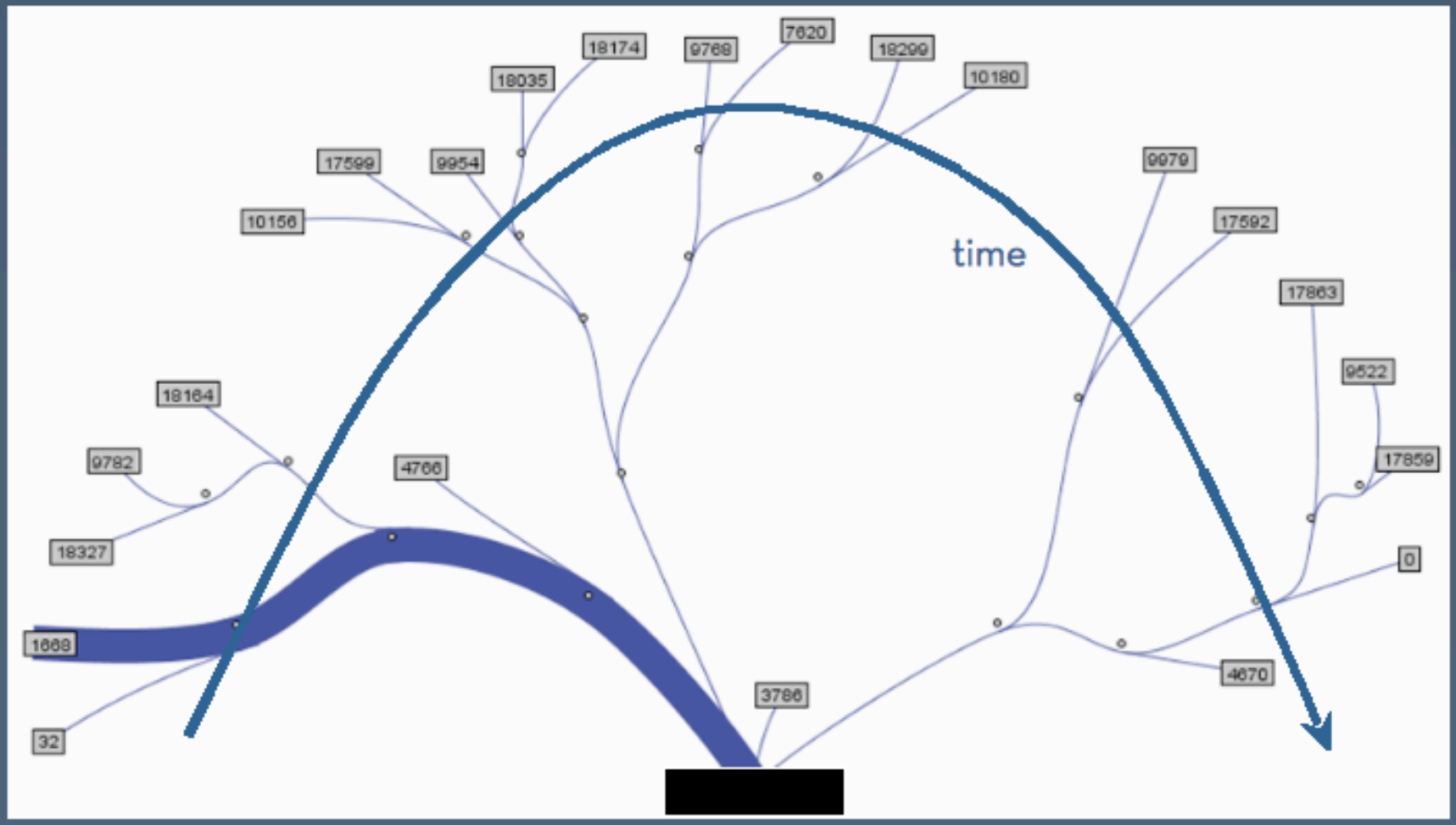

AMÉRIQUE DU NORD
AMÉRIQUE du SUD
FRANCE
ALLEMAGNE
SUÈDE
AFRIQUE
INDE

[Minard 1865]

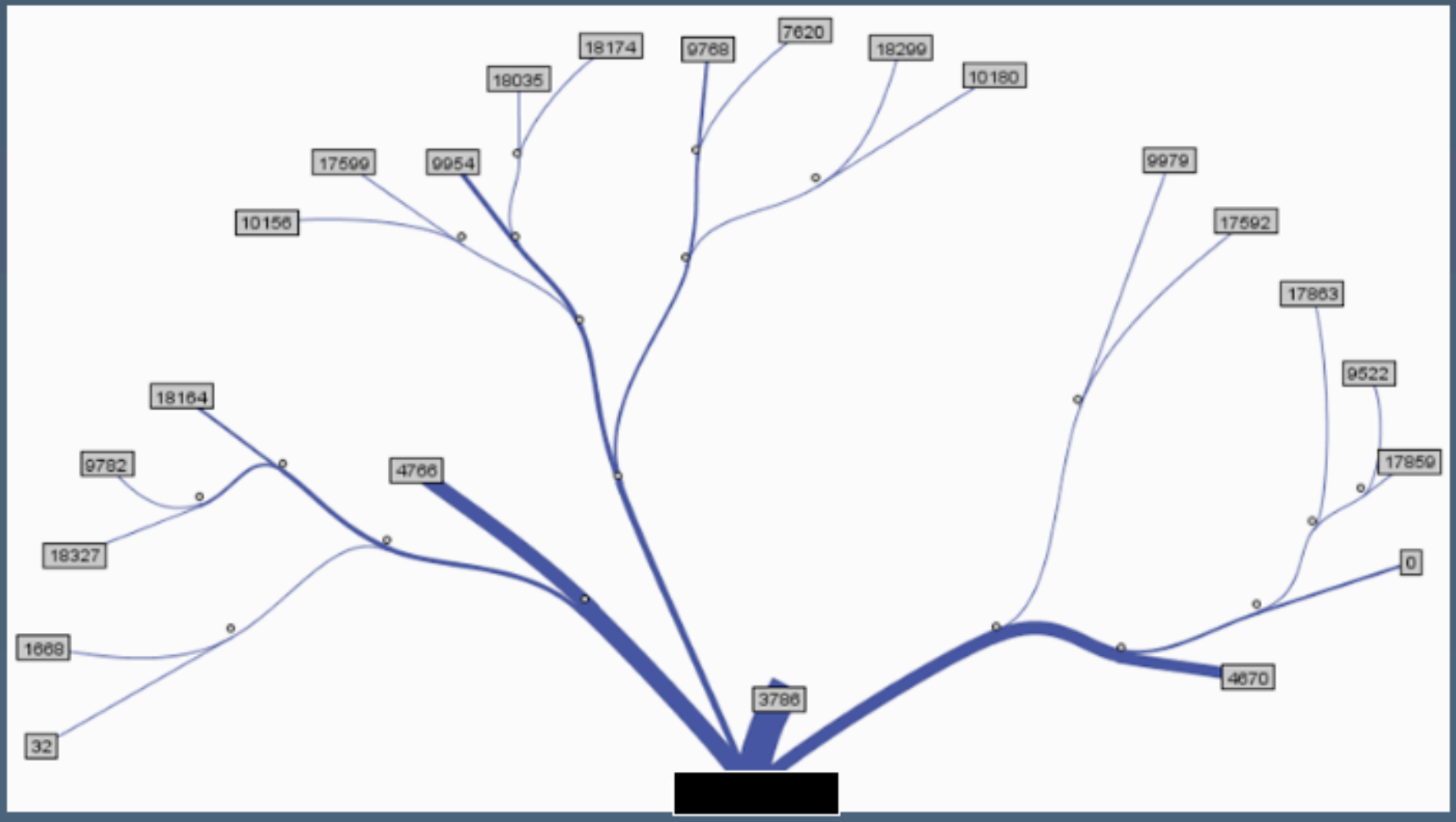
Flow Maps

- Node-link diagrams where the edges have varying width
- Instead of drawing by hand, generate automatically using hierarchical clustering [Phan et al. 2005]

Incoming Attack



Outgoing Scan from Victim



Flow Map Lessons Learned

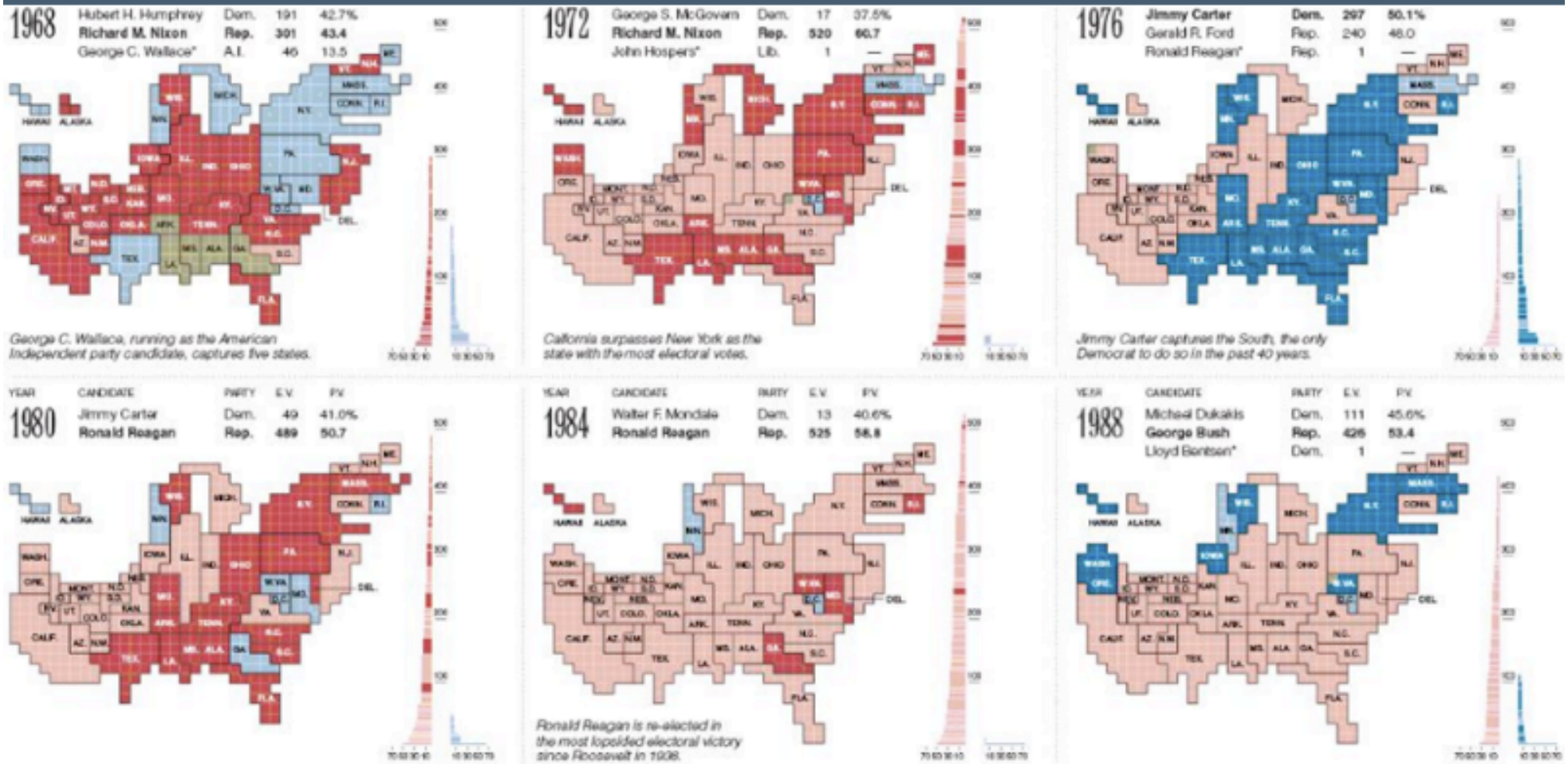
+ Good for presentation after-the-fact

- Bad for investigation
  - Time not emphasized
  - No exploration context

PROGRESSIVE MULTIPLES

Small Multiples

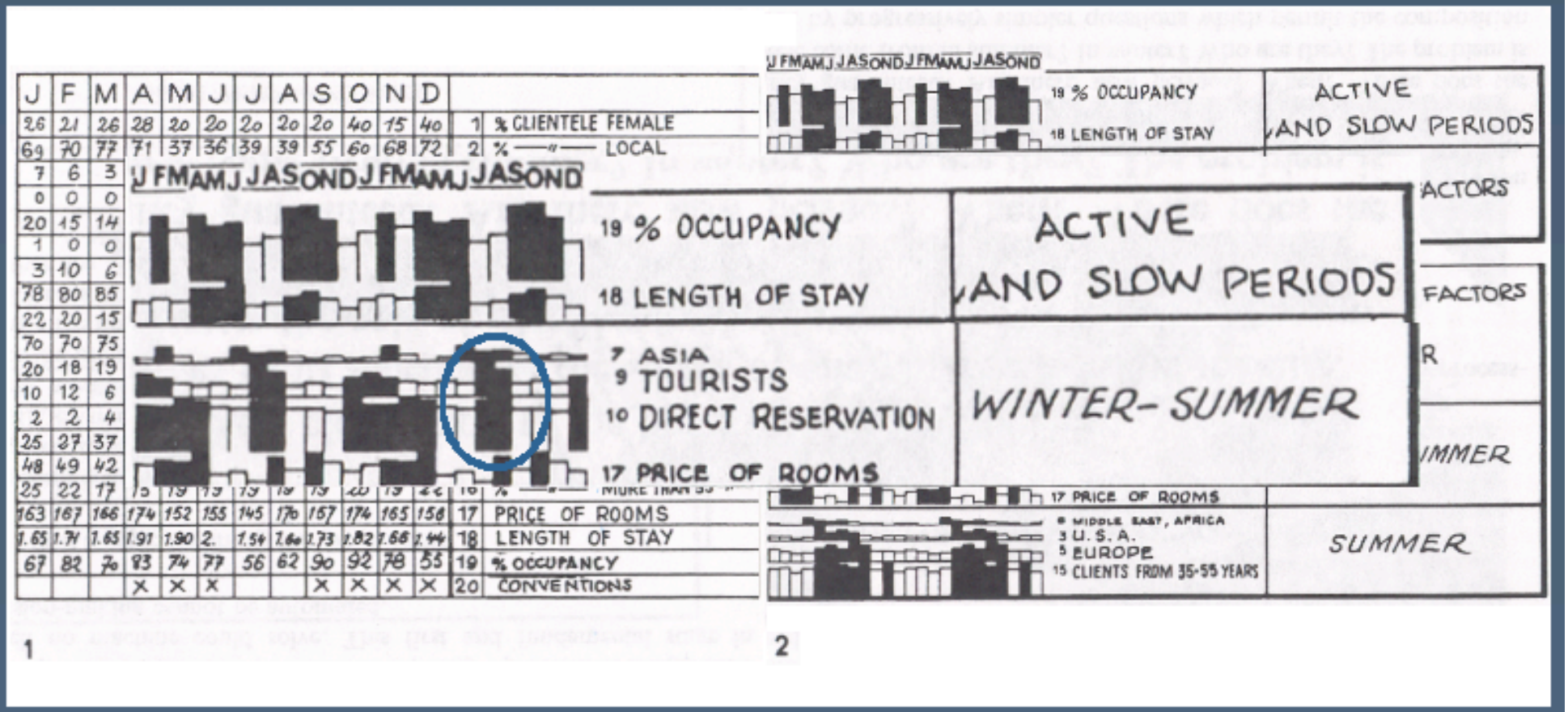
[Tufte 1983]



[New York Times, November 3, 2004]

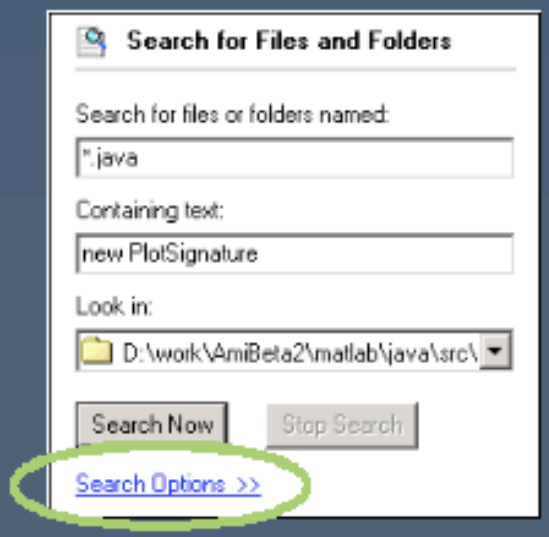
Reorderable Matrices

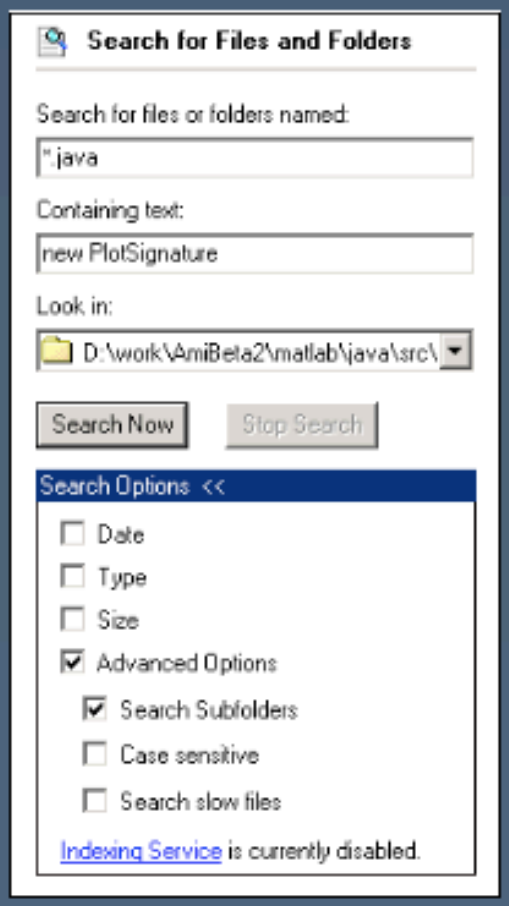
[Bertin 1977]



Progressive Disclosure

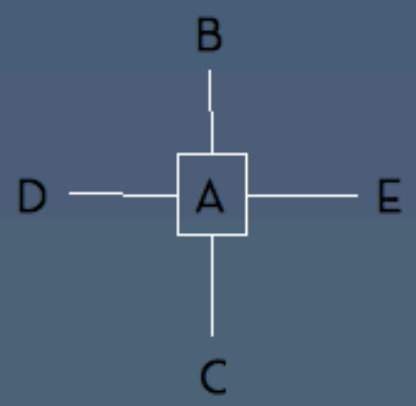
[Xerox Star 1981]



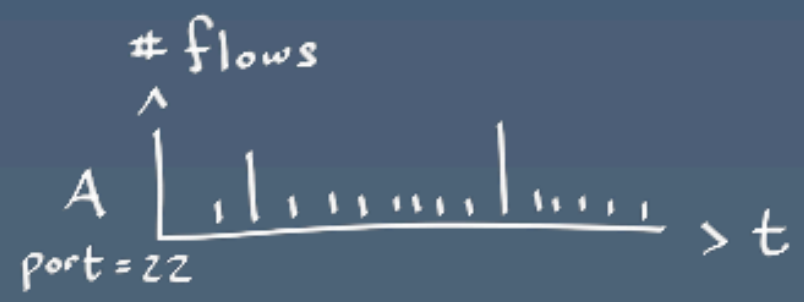


PROGRESSIVE MULTIPLES

Use row-based timelines



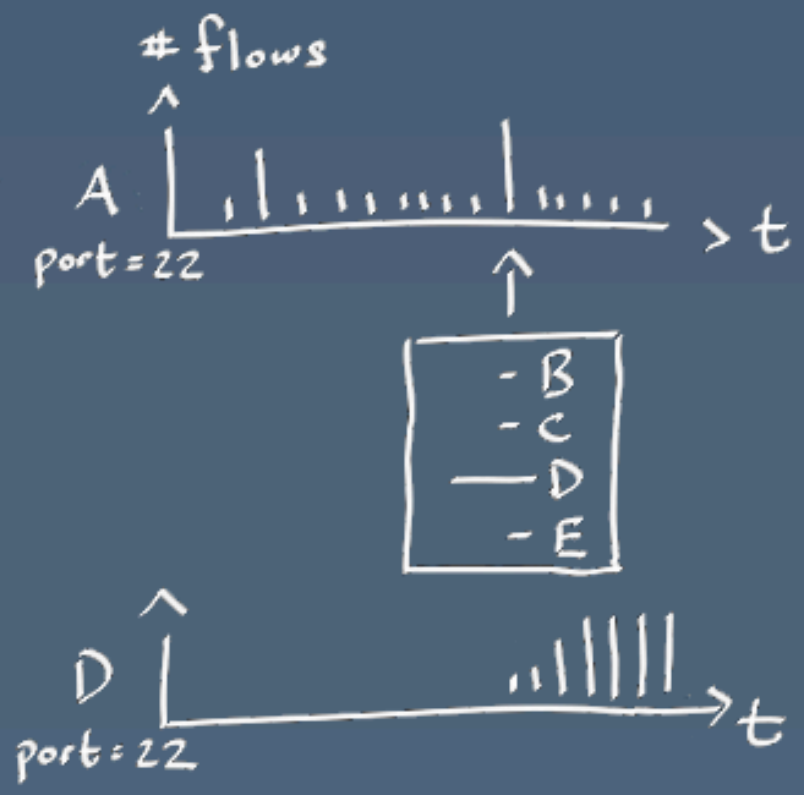
Traffic involving node A
as node-link diagram



Traffic involving A
as a timeline

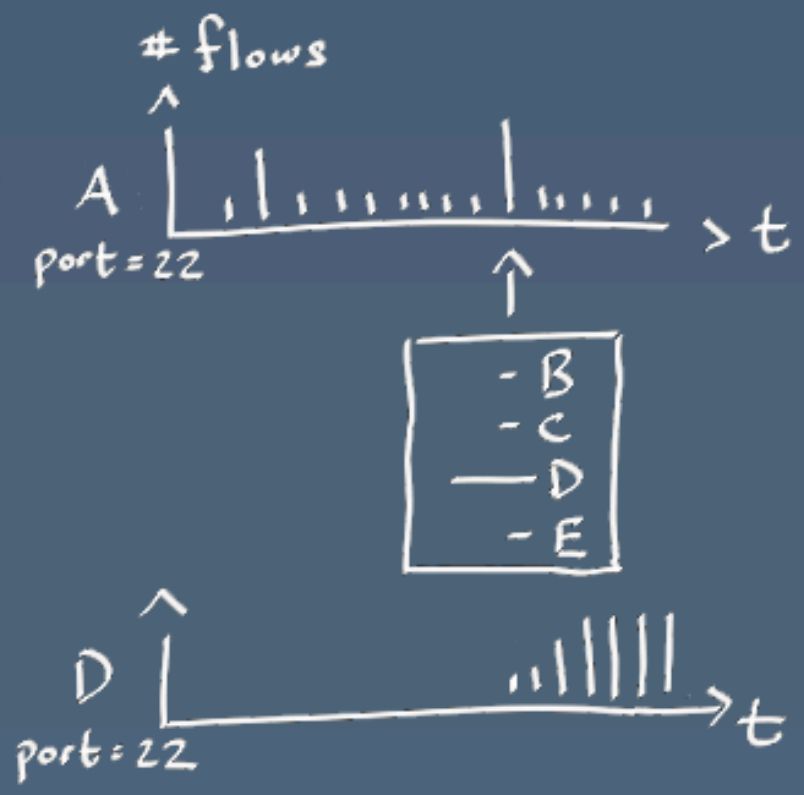
PROGRESSIVE MULTIPLES

Small multiples for comparison



PROGRESSIVE MULTIPLES

Find related events by pivoting

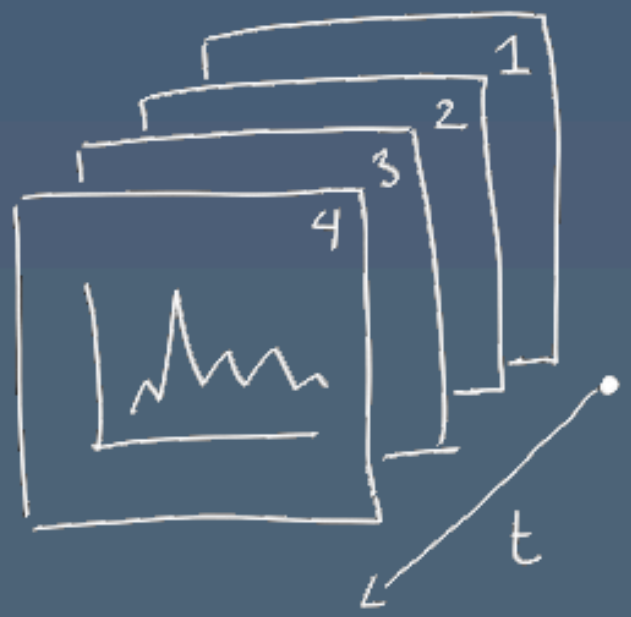


PROGRESSIVE MULTIPLES

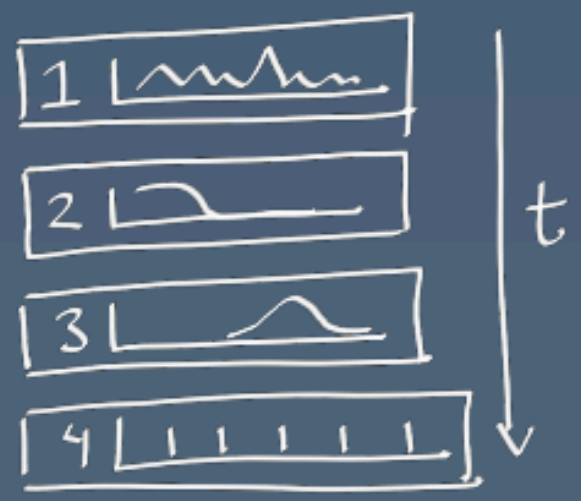
Reorder to reveal structure

PROGRESSIVE MULTIPLES

Support iterative investigation



Previous visualization states



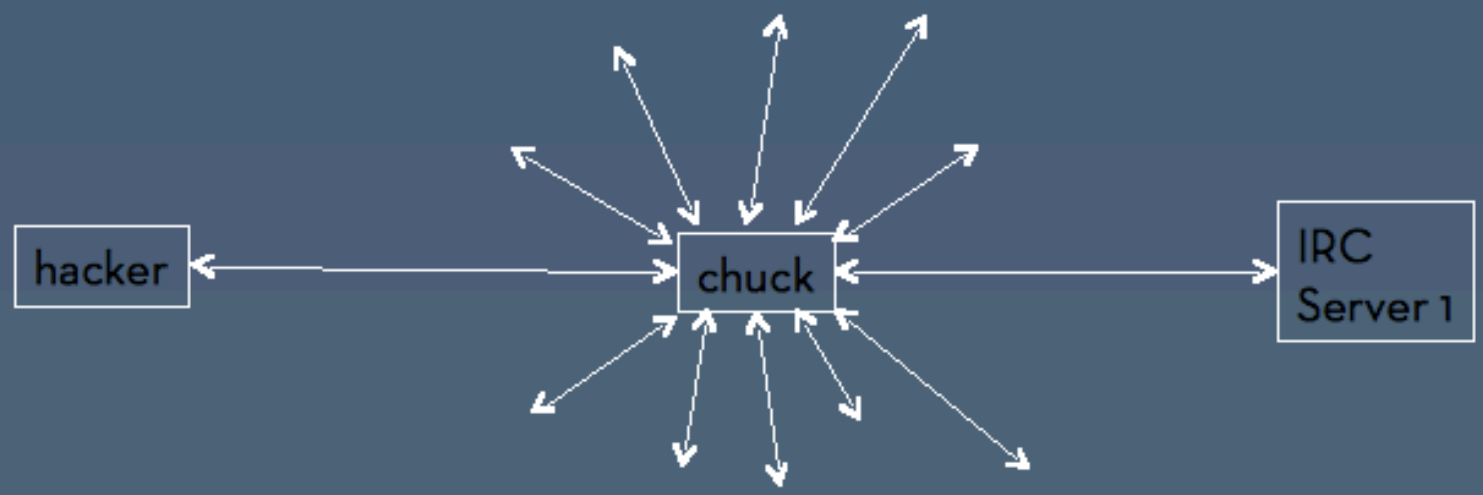
Making exploration history visible

External Cognition

- Human working memory is 7 ± 2 chunks
  [Miller 1956]
- Cognitive artifacts augment abilities
- Make investigation steps visible as rows
- Create new rows by pivoting

USING ISIS TO INVESTIGATE

The Scenario



ssh on port 22

irc on port 6667

Timelines for Overview

Analyst provides

- Focus IP Address – chuck (75.64.71.22)
- Time period – 1 day
- Height of bars - # of connections

When did the IRC traffic start?

- Brushing to Drill Down

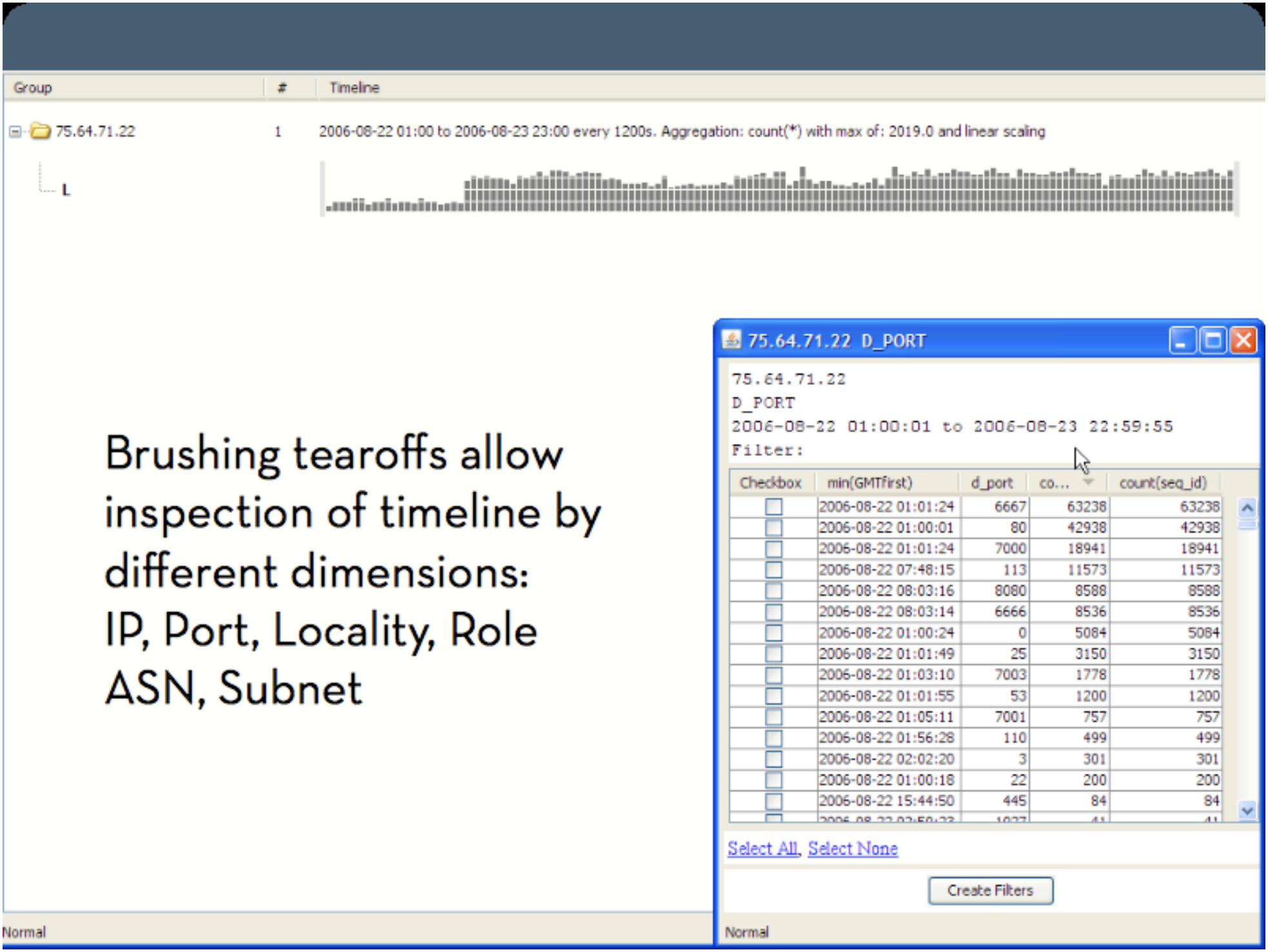
Title isis Comment Enter Window Comment Here

IP Query | Attribute Query | New Folder | New Window

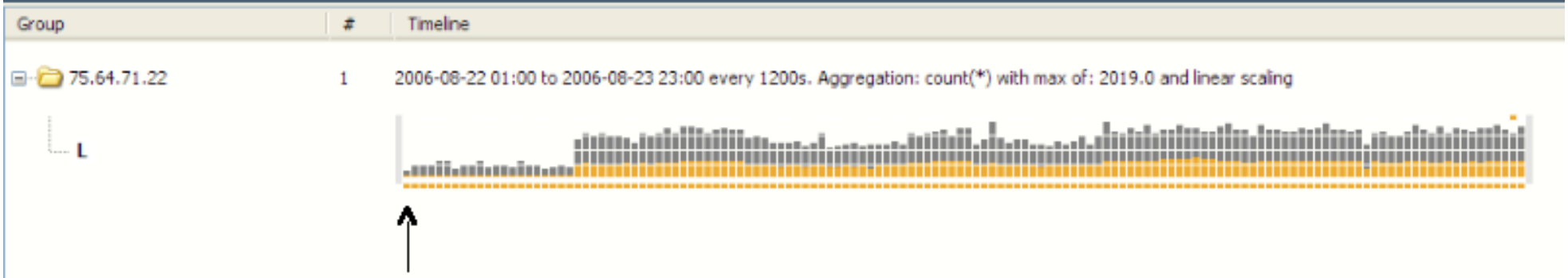
Group	#	Timeline
75.64.71.22	1	2006-08-22 01:00 to 2006-08-23 23:00 every 1200s. Aggregation: count(*) with max of: 2019.0 and linear scaling

2006-08-22 01:00:01 to 2006-08-23 22:59:55
0.0

- Switch to Log
- Make Graph
- Make Event Plot
- Filter Entire Row
- Selection
- See All Data
- See IP Tearoff (5797)
- See ASN Tearoff
- See D_PORT Tearoff
- See ROLE Tearoff
- See LOCALITY Tearoff
- See VLN Tearoff

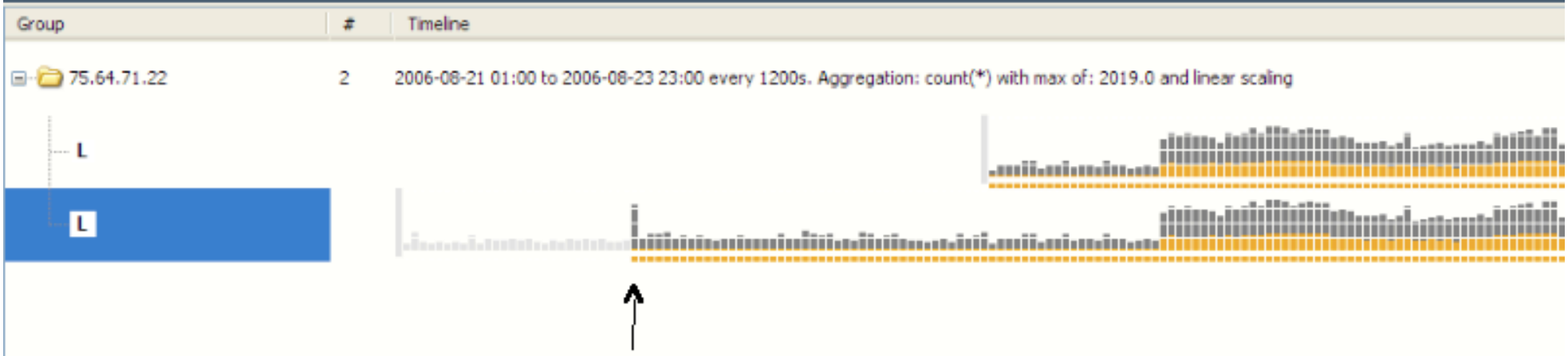


Brushing tearoffs allow inspection of timeline by different dimensions:
IP, Port, Locality, Role
ASN, Subnet



Brushing port 6667,
an IRC Port

Froze D_PORT 6667



Changing query timespan and brushing of port 6667 allows us to see when IRC traffic began

Froze D_PORT 6667

How was it compromised?

- Use event plots to understand sequences

The Event Table

75.64.71.22 and not(l_role > 0 and d_port=80)

2006-08-21 10:00:01 to 2006-08-21 10:50:28
0.0
75.64.71.22
and not(l_role > 0 and d_port=80)

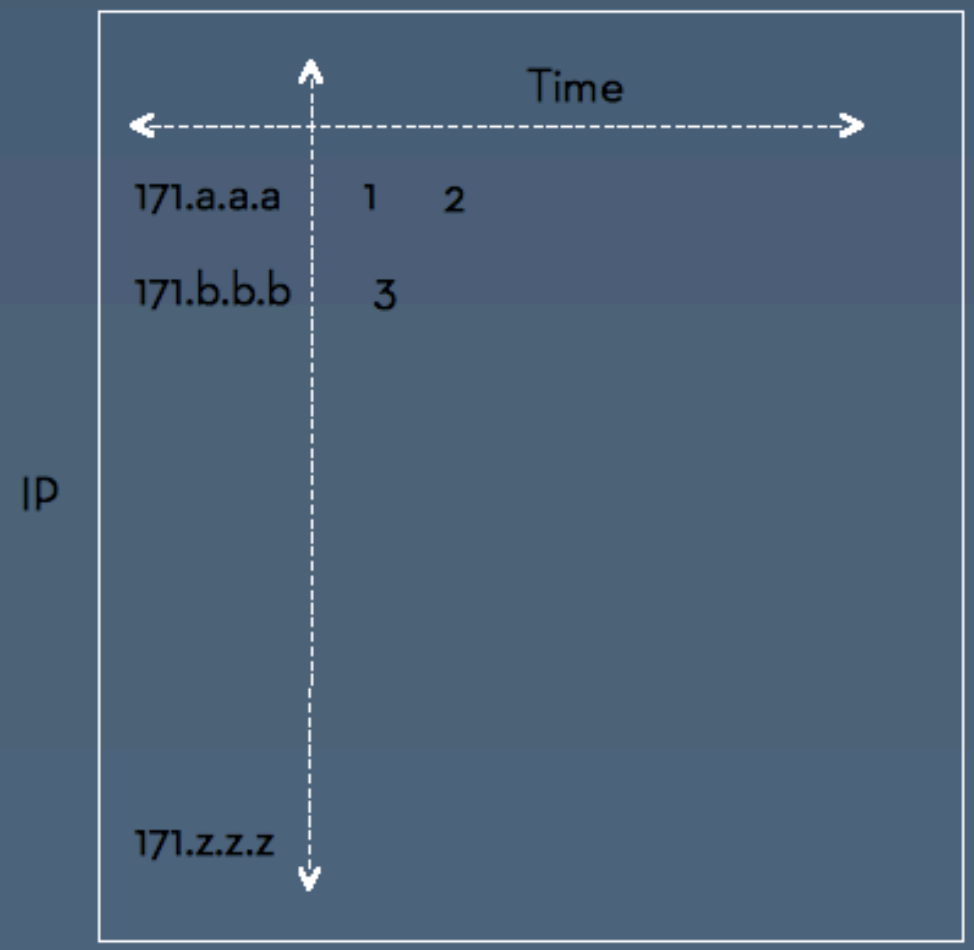
l_weekday	l_hour	GMTfirst	duration	locality	l_role	proto	l_asn	l_vlr	inet_ntoa(l_ipn)	l_port	r_asn	r_vlr	inet_ntoa(r_ipn)	r_port	d_port	l_pkt	l_byte	l_abyte	r_pkt	r_byte	r_abyte
2	3	2006-08-21 10:00:01	0.002	1	-3	1	32	71	75.64.71.22	8	26101	24288	66.94.230.32	8	0	2	196	128	2	196	12
2	3	2006-08-21 10:00:03	0	3	1	17	32	71	75.64.71.22	37373	32	64	75.64.67.192	37373	37373	0	0	0	1	177	13
2	3	2006-08-21 10:00:03	0.001	2	3	17	32	71	75.64.71.22	7001	32	16401	75.64.15.96	7001	7001	2	140	56	4	450	28
2	3	2006-08-21 10:00:04	0.009	1	-3	1	32	71	75.64.71.22	8	26101	24288	66.94.230.32	8	0	2	196	128	2	196	12
2	3	2006-08-21 10:00:18	0.506	2	2	6	32	71	75.64.71.22	22	32	17116	75.66.189.156	22	22	2	236	128	2	108	
2	3	2006-08-21 10:01:02	0.002	1	-3	1	32	71	75.64.71.22	8	26101	24288	66.94.230.32	8	0	2	196	128	2	196	12
2	3	2006-08-21 10:01:17	0.002	1	-3	1	32	71	75.64.71.22	8	26101	24288	66.94.230.32	8	0	2	196	128	2	196	12
2	3	2006-08-21 10:01:18	0.441	2	2	6	32	71	75.64.71.22	22	32	17116	75.66.189.156	22	22	2	236	128	2	108	
2	3	2006-08-21 10:01:45	0.001	1	3	1	32	71	75.64.71.22	0	15243	7936	147.31.67.106	0	0	2	184	116	2	184	11
2	3	2006-08-21 10:01:49	0.109	2	-3	6	32	71	75.64.71.22	45075	32	17174	75.67.9.109	45075	25	8	551	111	14	1202	43
2	3	2006-08-21 10:01:52	0.001	2	3	17	32	71	75.64.71.22	7001	32	16401	75.64.15.111	7001	7001	2	140	56	4	450	28
2	3	2006-08-21 10:01:54	27.002	2	-3	17	32	71	75.64.71.22	37396	32	7	75.64.24.227	37396	53	2	148	64	4	296	12
2	3	2006-08-21 10:01:59	19.998	2	-3	17	32	71	75.64.71.22	37396	32	7	75.64.24.201	37396	53	2	148	64	4	296	12
2	3	2006-08-21 10:02:03	0.002	1	-3	1	32	71	75.64.71.22	8	26101	24288	66.94.230.32	8	0	2	196	128	2	196	12
2	3	2006-08-21 10:02:06	4.99	2	-1	17	32	71	75.64.71.22	7001	32	14	75.64.22.185	7001	7000	8	658	322	0	0	
2	3	2006-08-21 10:02:13	7	1	-1	17	32	71	75.64.71.22	7001	3	17920	18.70.0.6	7001	7003	10	872	452	0	0	
2	3	2006-08-21 10:02:18	0.379	2	2	6	32	71	75.64.71.22	22	32	17116	75.66.189.156	22	22	2	236	128	2	108	
2	3	2006-08-21 10:02:19	0	2	-1	1	32	71	75.64.71.22	3	32	7	75.64.24.201	3	3	2	204	136	0	0	
2	3	2006-08-21 10:02:19	2.004	2	-3	17	32	71	75.64.71.22	37401	32	7	75.64.24.227	37401	53	1	74	32	2	148	6
2	3	2006-08-21 10:02:20	5.5	1	-1	17	32	71	75.64.71.22	7001	3	37120	18.145.0.25	7001	7003	8	724	388	0	0	
2	3	2006-08-21 10:02:21	0	2	-1	17	32	71	75.64.71.22	37402	32	7	75.64.24.201	37402	53	1	74	32	0	0	
2	3	2006-08-21 10:02:21	0	2	-1	1	32	71	75.64.71.22	3	32	7	75.64.24.227	3	3	2	204	136	0	0	
2	3	2006-08-21 10:02:27	0	2	-1	17	32	71	75.64.71.22	37403	32	7	75.64.24.201	37403	53	1	74	32	0	0	
2	3	2006-08-21 10:02:30	0.002	1	-3	1	32	71	75.64.71.22	8	26101	24288	66.94.230.32	8	0	2	196	128	2	196	12
2	3	2006-08-21 10:02:51	0	2	-1	1	32	71	75.64.71.22	3	32	7	75.64.24.201	3	3	1	102	68	0	0	
2	3	2006-08-21 10:02:51	0	2	-1	1	32	71	75.64.71.22	3	32	7	75.64.24.201	3	3	1	102	68	0	0	
2	3	2006-08-21 10:02:51	0	2	1	17	32	71	75.64.71.22	37402	32	7	75.64.24.201	37402	37402	0	0	0	2	148	6
2	3	2006-08-21 10:02:51	0	2	1	17	32	71	75.64.71.22	37403	32	7	75.64.24.201	37403	37403	0	0	0	2	148	6
2	3	2006-08-21 10:03:04	0.002	1	-3	1	32	71	75.64.71.22	8	26101	24288	66.94.230.32	8	0	2	196	128	2	196	12
2	3	2006-08-21 10:03:16	0.546	3	3	6	32	71	75.64.71.22	25	32	64	77.232.79.23	25	25	9	937	331	9	691	10
2	3	2006-08-21 10:03:18	0.548	2	2	6	32	71	75.64.71.22	22	32	17116	75.66.189.156	22	22	2	236	128	2	108	
2	3	2006-08-21 10:03:25	0.224	3	3	6	32	71	75.64.71.22	25	32	64	77.232.79.23	25	25	13	1453	583	13	3643	280
2	3	2006-08-21 10:03:32	4.285	2	-3	6	32	71	75.64.71.22	45075	32	17174	75.67.9.109	45075	25	33	2391	609	54	4072	115
2	3	2006-08-21 10:03:43	0.002	1	-3	1	32	71	75.64.71.22	8	26101	24288	66.94.230.32	8	0	2	196	128	2	196	12
2	3	2006-08-21 10:04:05	0.002	1	-3	1	32	71	75.64.71.22	8	26101	24288	66.94.230.32	8	0	2	196	128	2	196	12
2	3	2006-08-21 10:04:18	0.267	2	2	6	32	71	75.64.71.22	22	32	17116	75.66.189.156	22	22	2	236	128	2	108	
2	3	2006-08-21 10:04:56	0.002	1	-3	1	32	71	75.64.71.22	8	26101	24288	66.94.230.32	8	0	2	196	128	2	196	12
2	3	2006-08-21 10:05:06	0.002	1	-3	1	32	71	75.64.71.22	8	26101	24288	66.94.230.32	8	0	2	196	128	2	196	12
2	3	2006-08-21 10:05:18	0.332	2	2	6	32	71	75.64.71.22	22	32	17116	75.66.189.156	22	22	2	236	128	2	108	
2	3	2006-08-21 10:05:27	5	2	-1	17	32	71	75.64.71.22	7001	32	14	75.64.22.185	7001	7000	8	658	322	0	0	

From Event Table to Event Plot

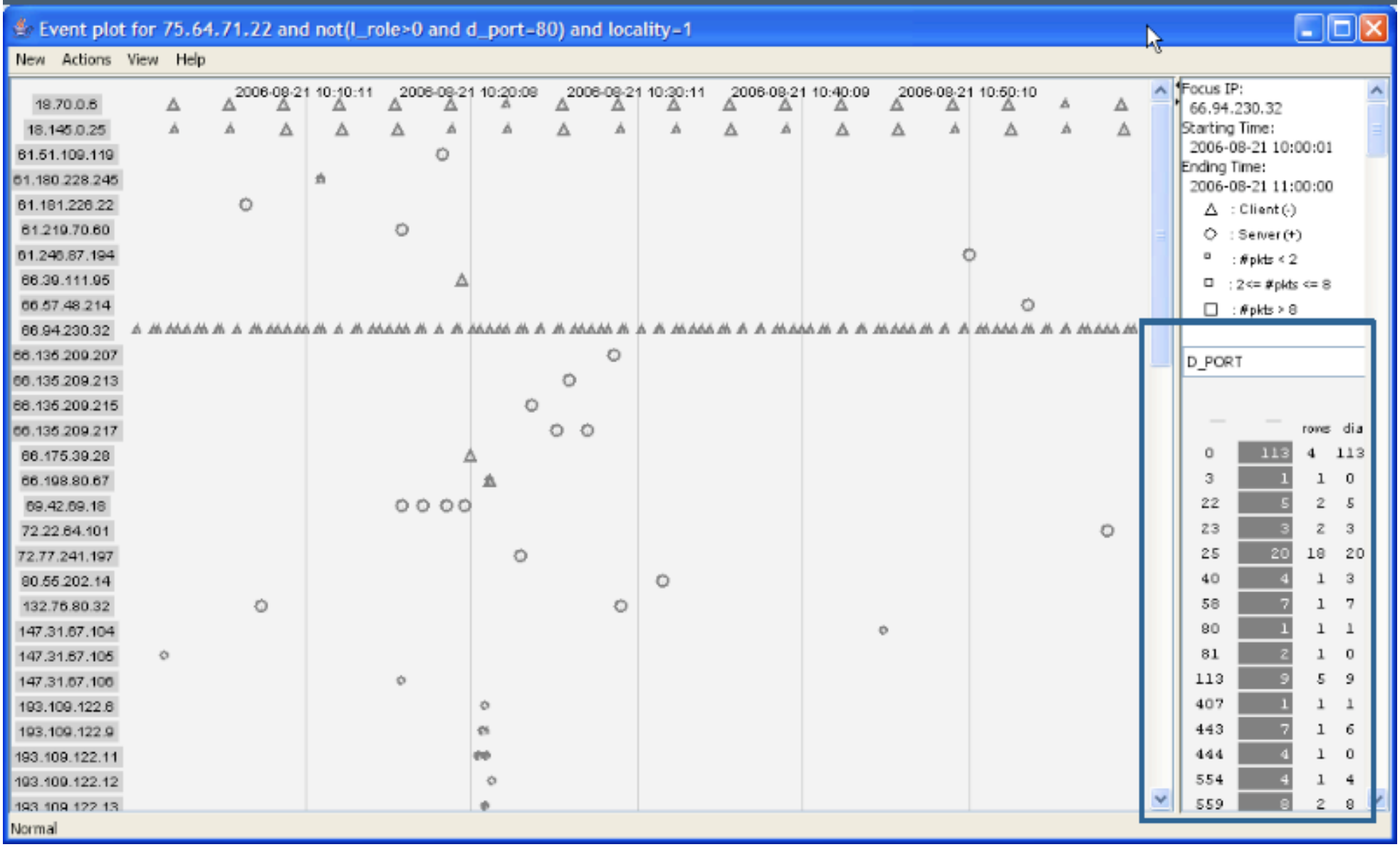
Event Table

1	Time	171.a.a.a	...
2	Time	171.a.a.a	...
3	Time	171.b.b.b	...

Event Plot



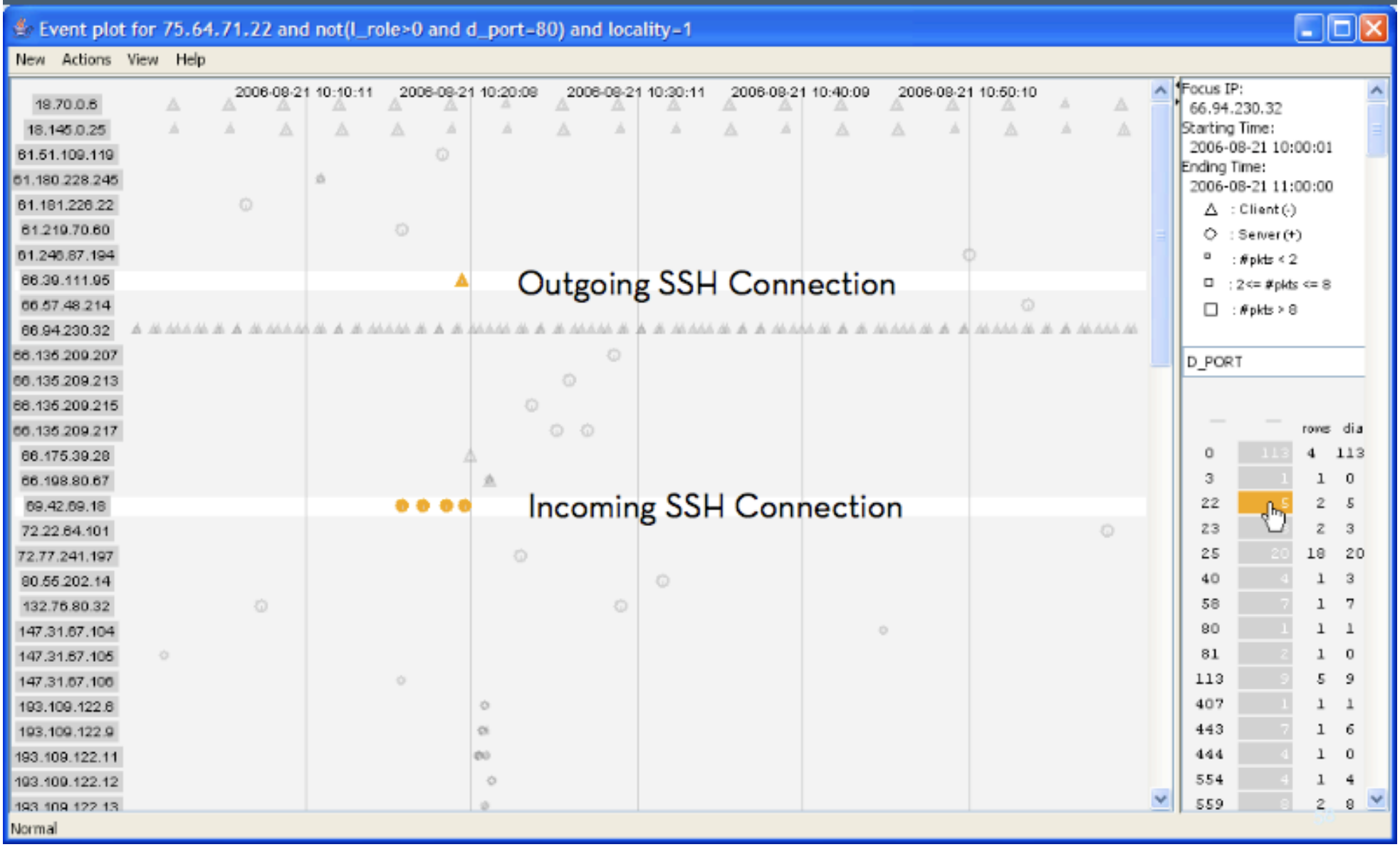
Event Plot


Event plot for 75.64.71.22 and not(l_role>0 and d_port=80) and locality=1
New Actions View Help
2006-08-21 10:10:11
2006-08-21 10:20:09
2006-08-21 10:30:11
2006-08-21 10:40:09
2006-08-21 10:50:10
18.70.0.6
18.145.0.25
61.51.109.119
61.180.228.245
61.181.226.22
61.219.70.60
61.246.67.194
66.39.111.95
66.57.48.214
66.94.230.32
66.135.209.207
66.135.209.213
66.135.209.215
66.135.209.217
66.175.39.28
66.198.80.67
69.42.69.18
72.22.64.101
72.77.241.197
80.55.202.14
132.76.80.32
147.31.67.104
147.31.67.105
147.31.67.106
193.109.122.8
193.109.122.9
193.109.122.11
193.109.122.12
193.109.122.13
Focus IP:
66.94.230.32
Starting Time:
2006-08-21 10:00:01
Ending Time:
2006-08-21 11:00:00
: Client (-)
: Server (+)
: #pkts < 2
: 2<= #pkts <= 8
: #pkts > 8
D_PORT
rows dia
0 113 4 113
3 1 1 0
22 5 2 5
23 3 2 3
25 20 18 20
40 4 1 3
58 7 1 7
80 1 1 1
81 2 1 0
113 9 5 9
407 1 1 1
443 7 1 6
444 4 1 0
554 4 1 4
559 8 2 8
Normal

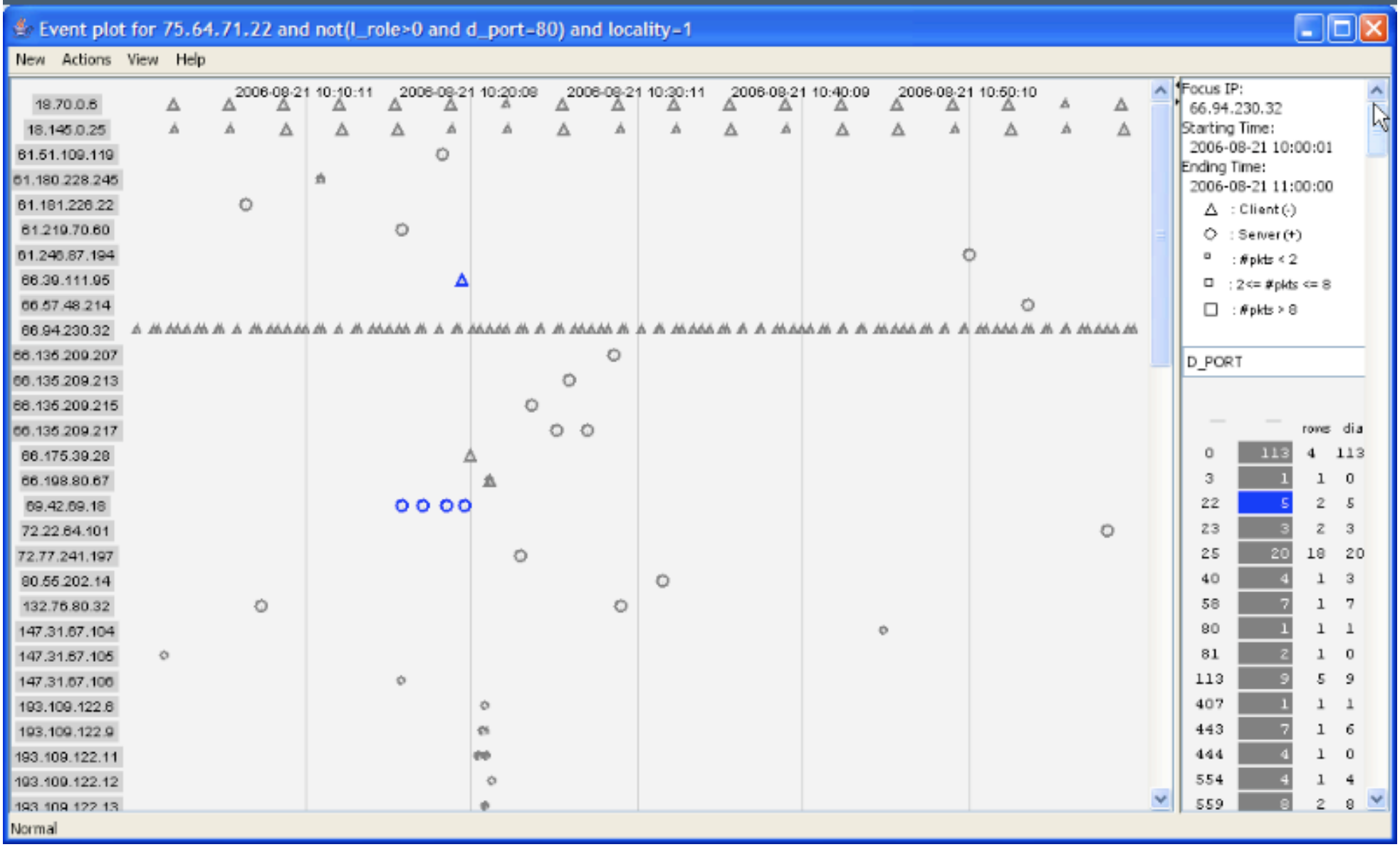
What are the intrusion stages?

- Initial SSH Connection
- Download of Client Tools
- IRC Servers Contacted

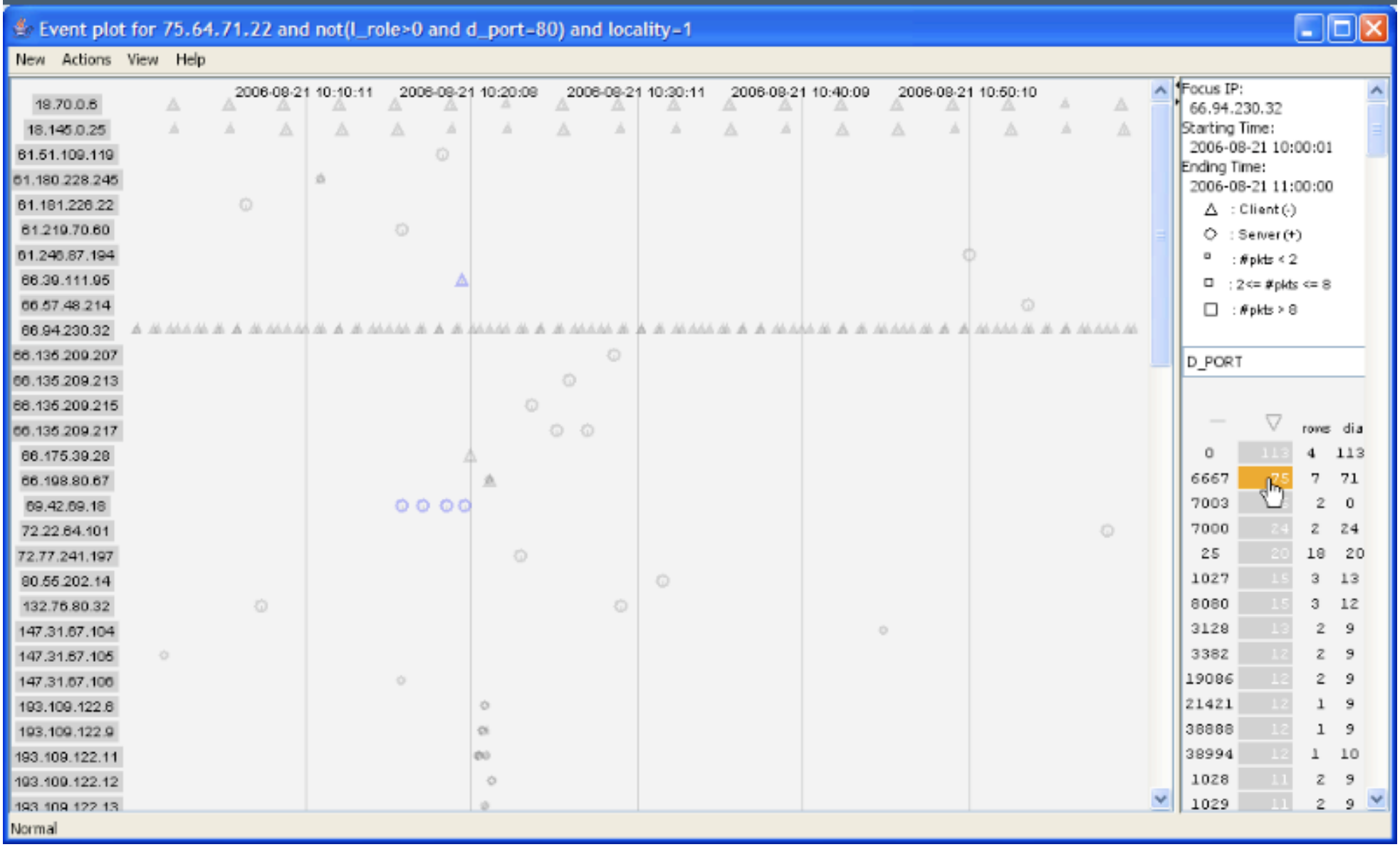
Initial SSH Connection



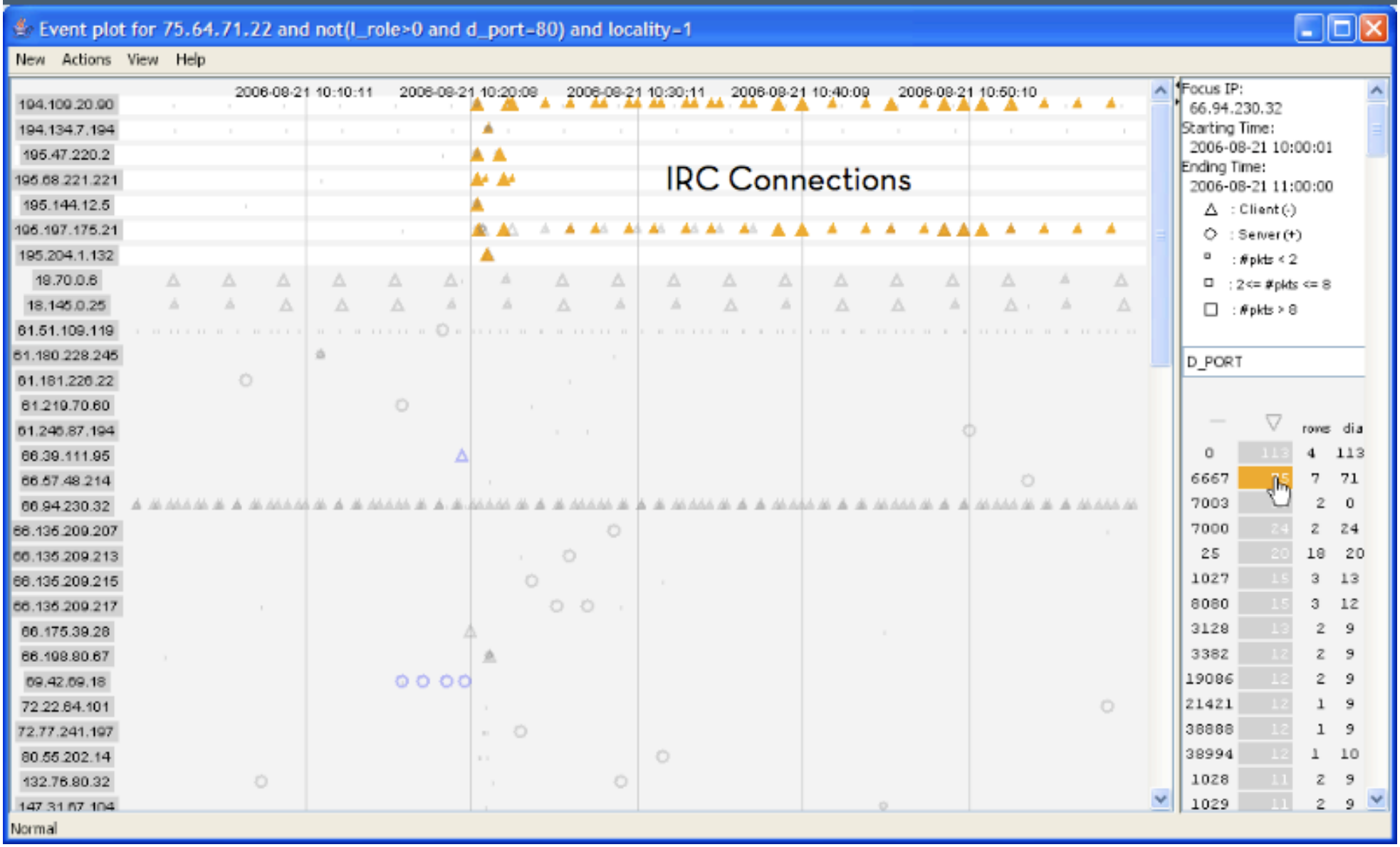
Initial SSH Connection



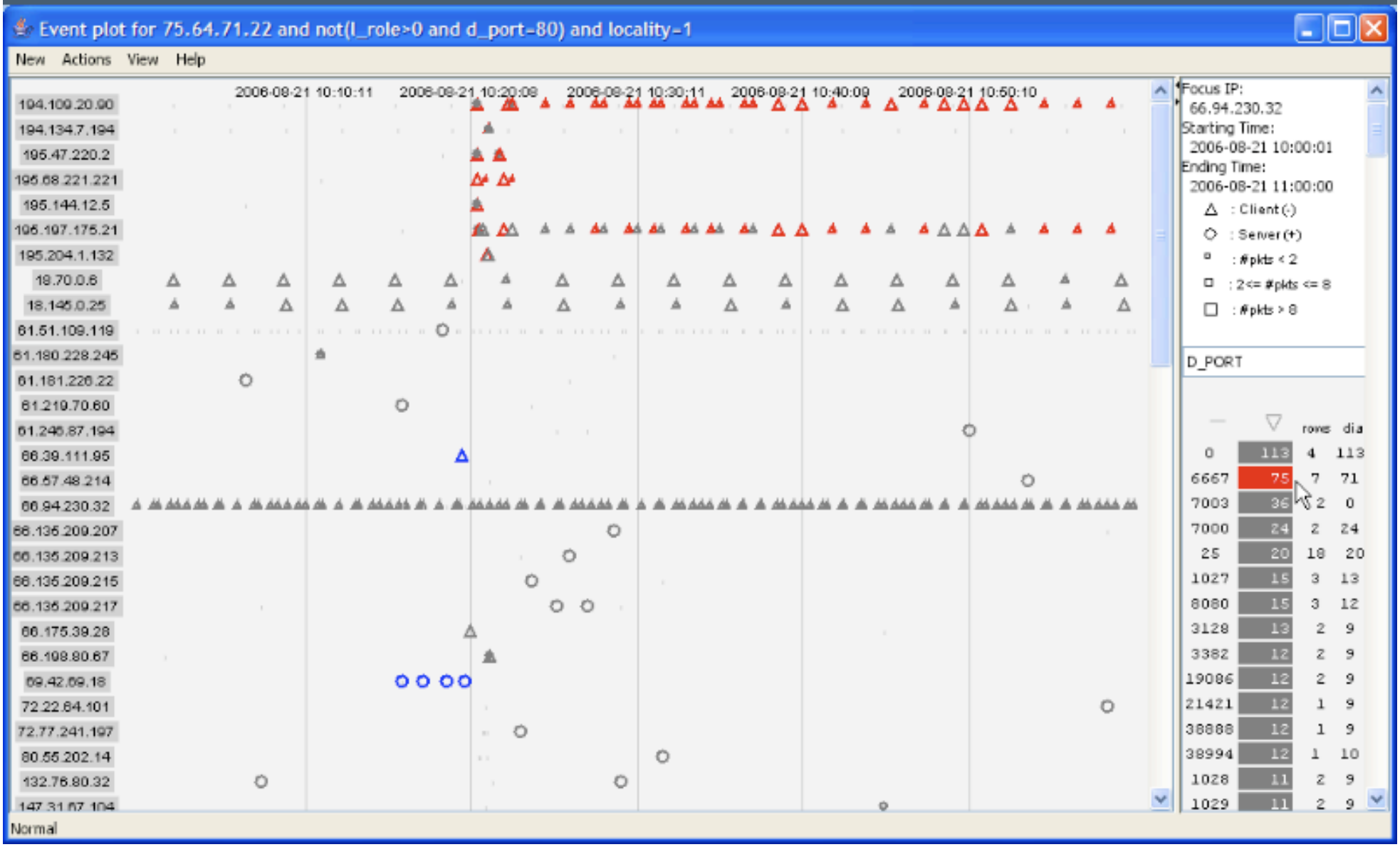
IRC Traffic


Event plot for 75.64.71.22 and not(l_role>0 and d_port=80) and locality=1
New Actions View Help
2006-08-21 10:10:11
2006-08-21 10:20:09
2006-08-21 10:30:11
2006-08-21 10:40:09
2006-08-21 10:50:10
18.70.0.6
18.145.0.25
61.51.109.119
61.180.228.245
61.181.226.22
61.219.70.60
61.246.67.194
66.39.111.95
66.57.48.214
66.94.230.32
66.135.209.207
66.135.209.213
66.135.209.215
66.135.209.217
66.175.39.28
66.198.80.67
69.42.69.18
72.22.64.101
72.77.241.197
80.55.202.14
132.76.80.32
147.31.67.104
147.31.67.105
147.31.67.106
193.109.122.8
193.109.122.9
193.109.122.11
193.109.122.12
193.109.122.13
Focus IP:
66.94.230.32
Starting Time:
2006-08-21 10:00:01
Ending Time:
2006-08-21 11:00:00
: Client (-)
: Server (+)
: #pkts < 2
: 2<= #pkts <= 8
: #pkts > 8
D_PORT
rows dia
0 113 4 113
6667 75 7 71
7003 2 0
7000 24 2 24
25 20 18 20
1027 15 3 13
8080 15 3 12
3128 13 2 9
3382 12 2 9
19086 12 2 9
21421 12 1 9
38888 12 1 9
38994 12 1 10
1028 11 2 9
1029 11 2 9
Normal

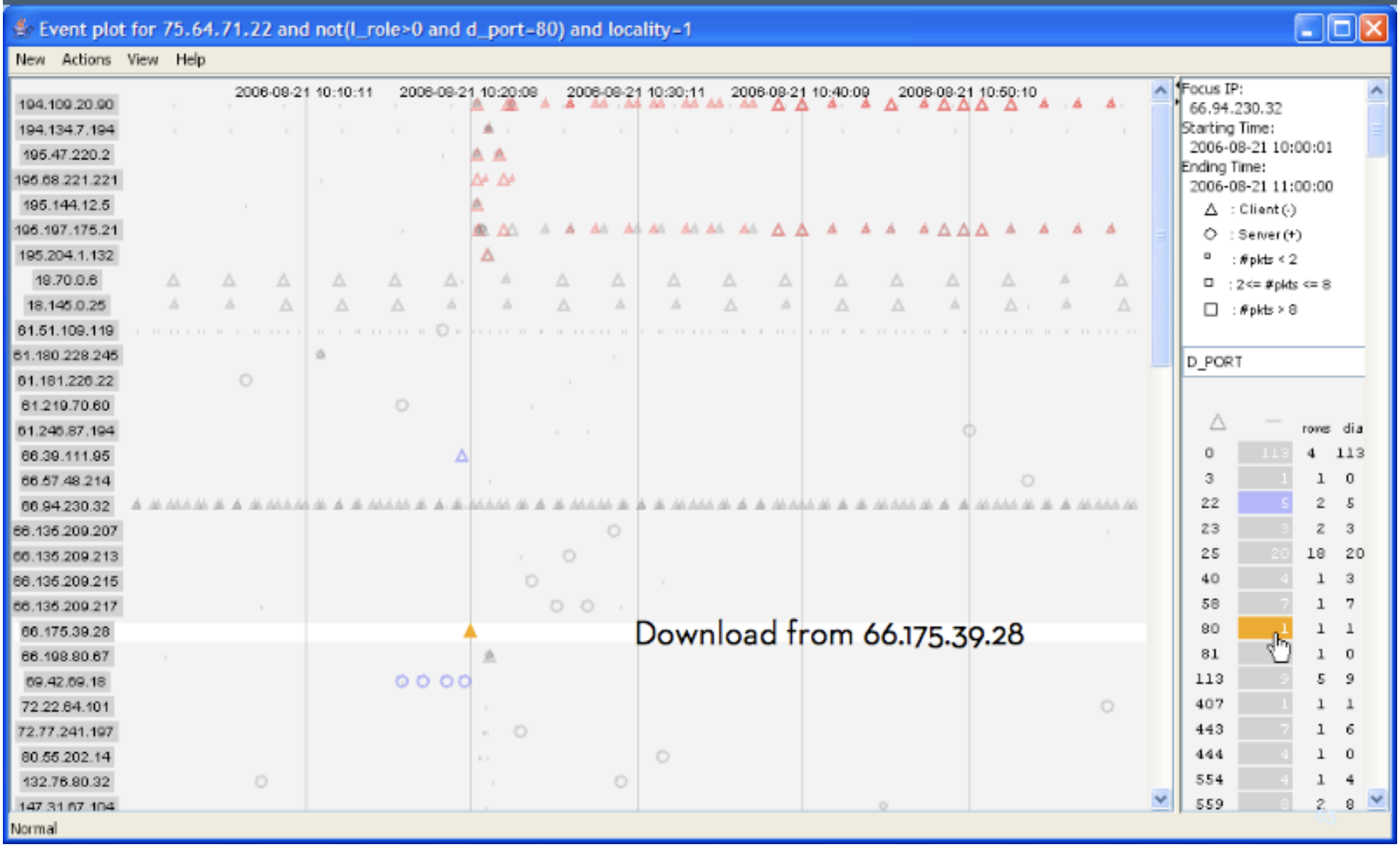
Collecting Brushed IRC Traffic



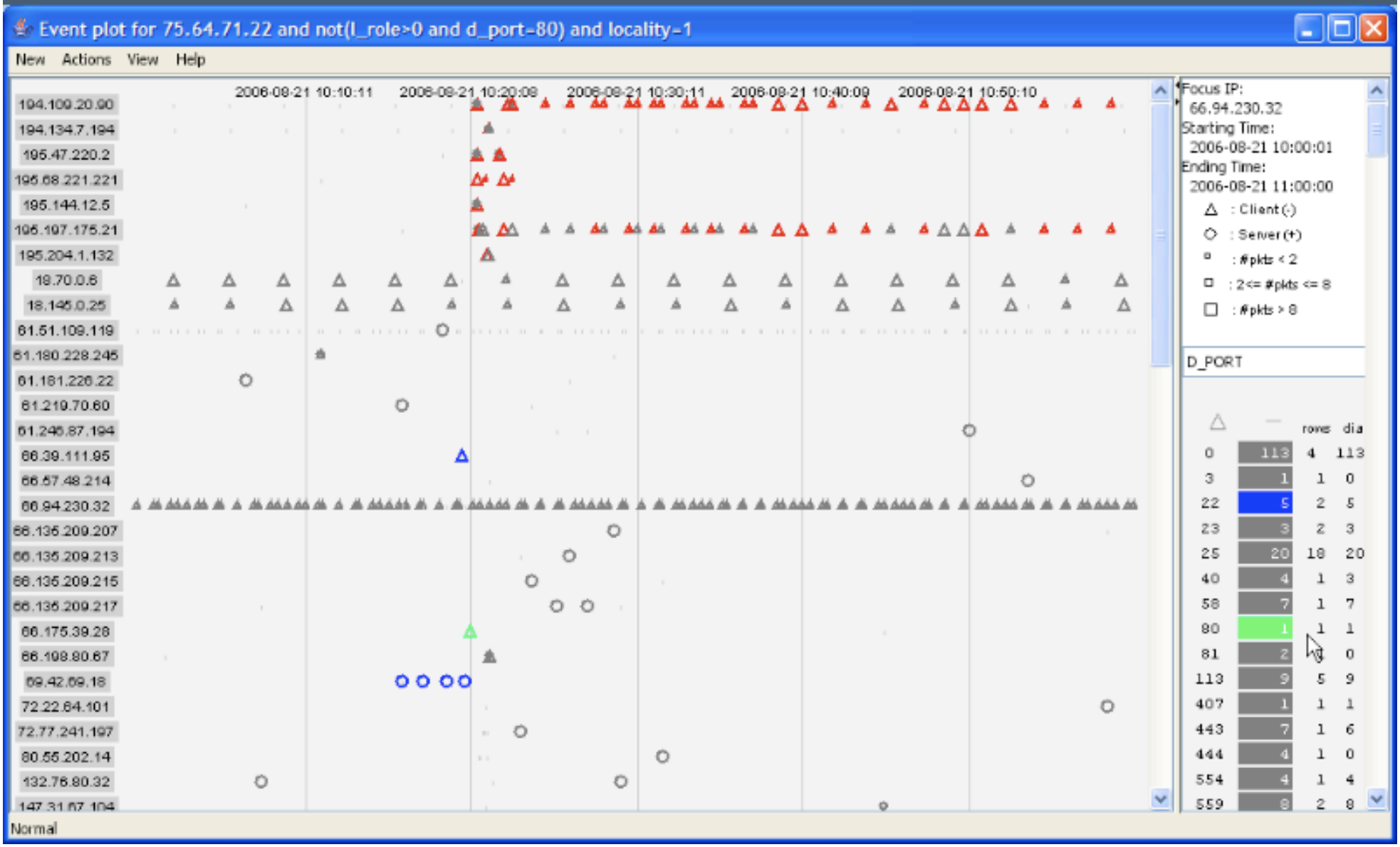
IRC Traffic


Event plot for 75.64.71.22 and not(l_role>0 and d_port=80) and locality=1
New Actions View Help
2006-08-21 10:10:11
2006-08-21 10:20:09
2006-08-21 10:30:11
2006-08-21 10:40:09
2006-08-21 10:50:10
194.109.20.90
194.134.7.194
195.47.220.2
195.68.221.221
195.144.12.5
195.197.175.21
195.204.1.132
19.70.0.6
18.145.0.25
61.51.109.119
61.180.228.245
61.181.226.22
61.219.70.60
61.245.87.194
66.39.111.95
66.57.48.214
66.94.230.32
66.135.209.207
66.135.209.213
66.135.209.215
66.135.209.217
66.175.39.28
66.198.80.67
69.42.69.18
72.22.84.101
72.77.241.197
80.55.202.14
132.76.80.32
147.31.67.104
Focus IP:
66.94.230.32
Starting Time:
2006-08-21 10:00:01
Ending Time:
2006-08-21 11:00:00
: Client (-)
: Server (+)
: #pkts < 2
: 2<= #pkts <= 8
: #pkts > 8
D_PORT
rows dia
0 113 4 113
6667 75 7 71
7003 36 2 0
7000 24 2 24
25 20 18 20
1027 15 3 13
8080 15 3 12
3128 13 2 9
3382 12 2 9
19086 12 2 9
21421 12 1 9
38888 12 1 9
38994 12 1 10
1028 11 2 9
1029 11 2 9
Normal

Download of Intrusion Tools



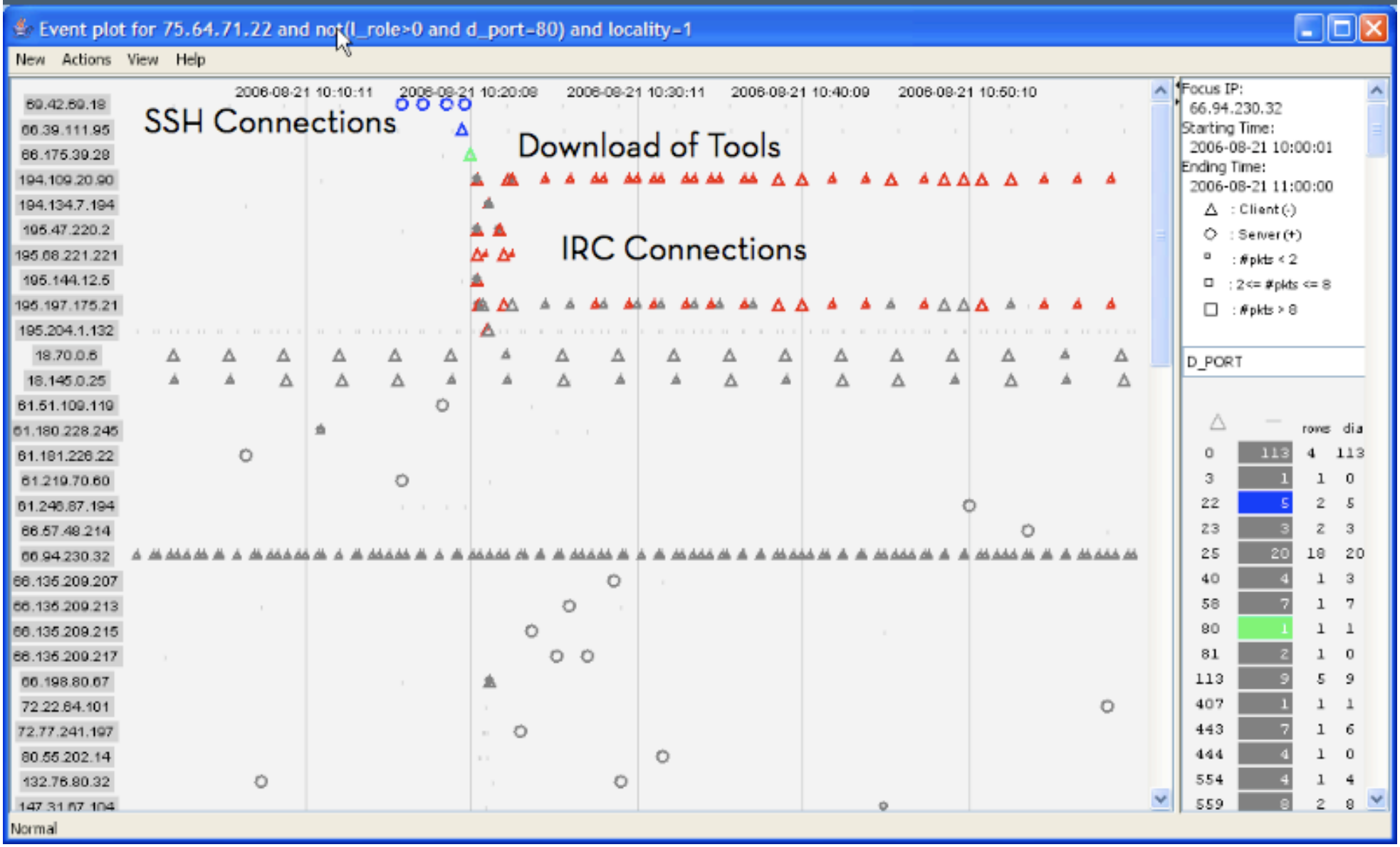
Download of Intrusion Tools



How do we make it clearer?

- Reordering to Show Sequence

Reordered Event Table



Title isis Comment Enter Window Comment Here

IP Query | Attribute Query | New Folder | New Window

Event plot for 75.64.71.22 and not(l_role>0 and d_port=80) and locality=1

New Actions View Help

2006-08-21 10:10:11 2006-08-21 10:20:08 2006-08-21 10:30:11 2006-08-21 10:40:09 2006-08-21 10:50:10

66.39.111.95
69.42.69.18
66.175.39.28
194.109.20.90
194.134.7.194
195.47.220.2
195.69.221.221
195.144.12.5
195.197.175.21
195.204.1.132
18.70.0.6
18.145.0.25
61.51.109.119
61.180.228.245
61.181.226.22
61.219.70.60
61.246.87.194
66.57.48.214
66.94.230.32
66.135.209.207
66.135.209.213
66.135.209.215
66.135.209.217
66.198.80.67
72.22.64.101
72.77.241.197
80.55.202.14
132.76.80.32

Focus IP:
75.64.71.22
Starting Time:
2006-08-21 10:00:
Ending Time:
2006-08-21 11:00:
: Client (-)
: Server (+)
: #pkts < 2
: 2 <= #pkts <=
: #pkts > 8

D_PORT

0	113	4
6667	75	7
7003	[illegible]	
7000	24	2
25	20	1
1027	15	3
8080	15	3
3128	13	
3382	12	
19086	12	
21421	12	
38888	12	
38994	12	1
1028	11	
1029	11	

Normal

Are we seeing all the data?

- Continuous vs. Ordinal Time

Continuous vs. Ordinal Time

A – January 1

B – November 2

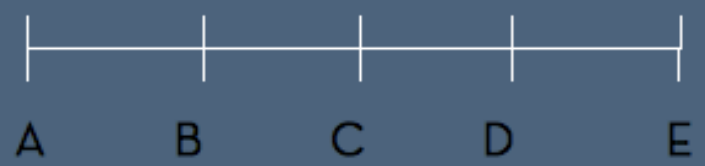
C – November 16

D – December 2

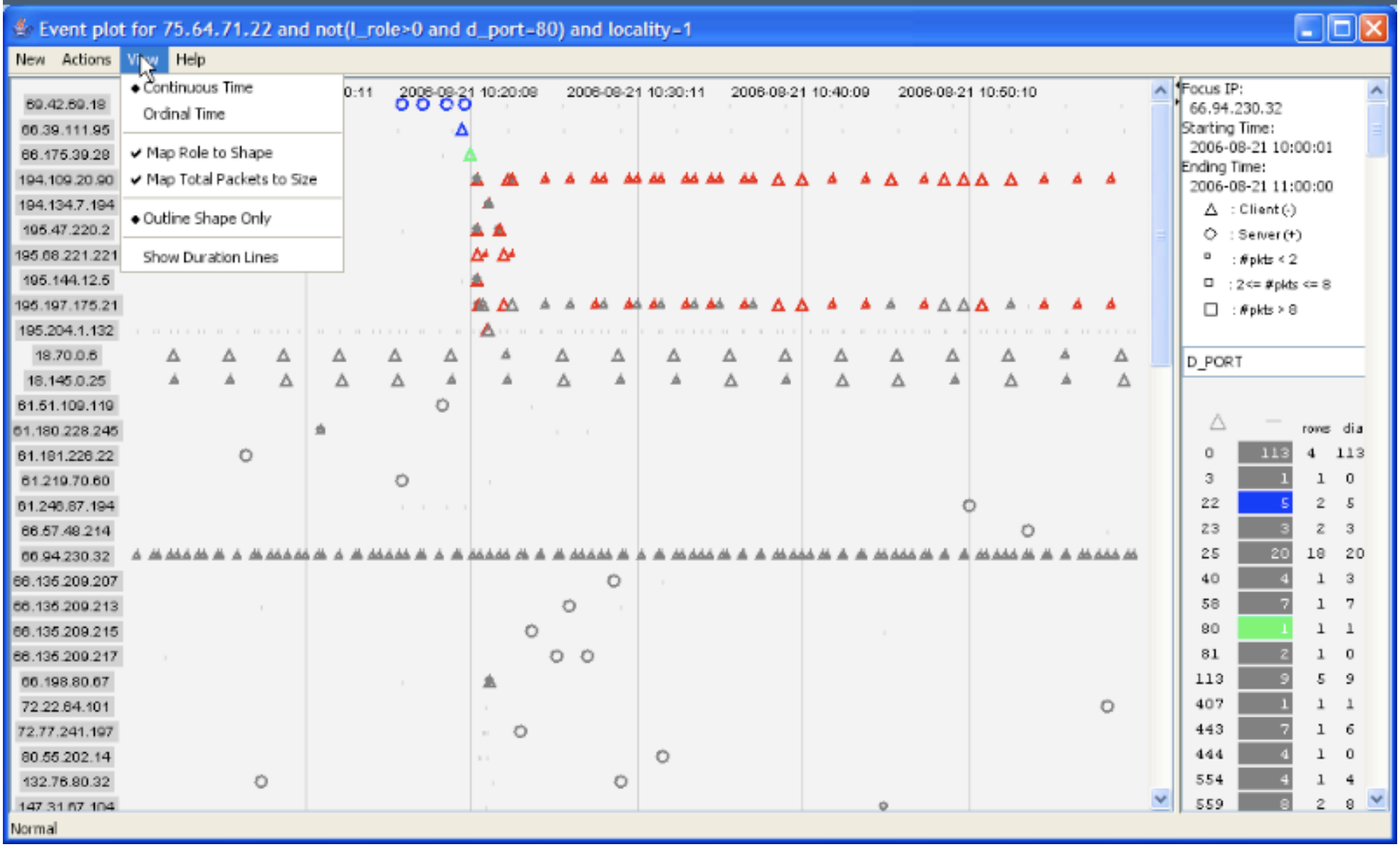
E – December 25

Continuous event plot

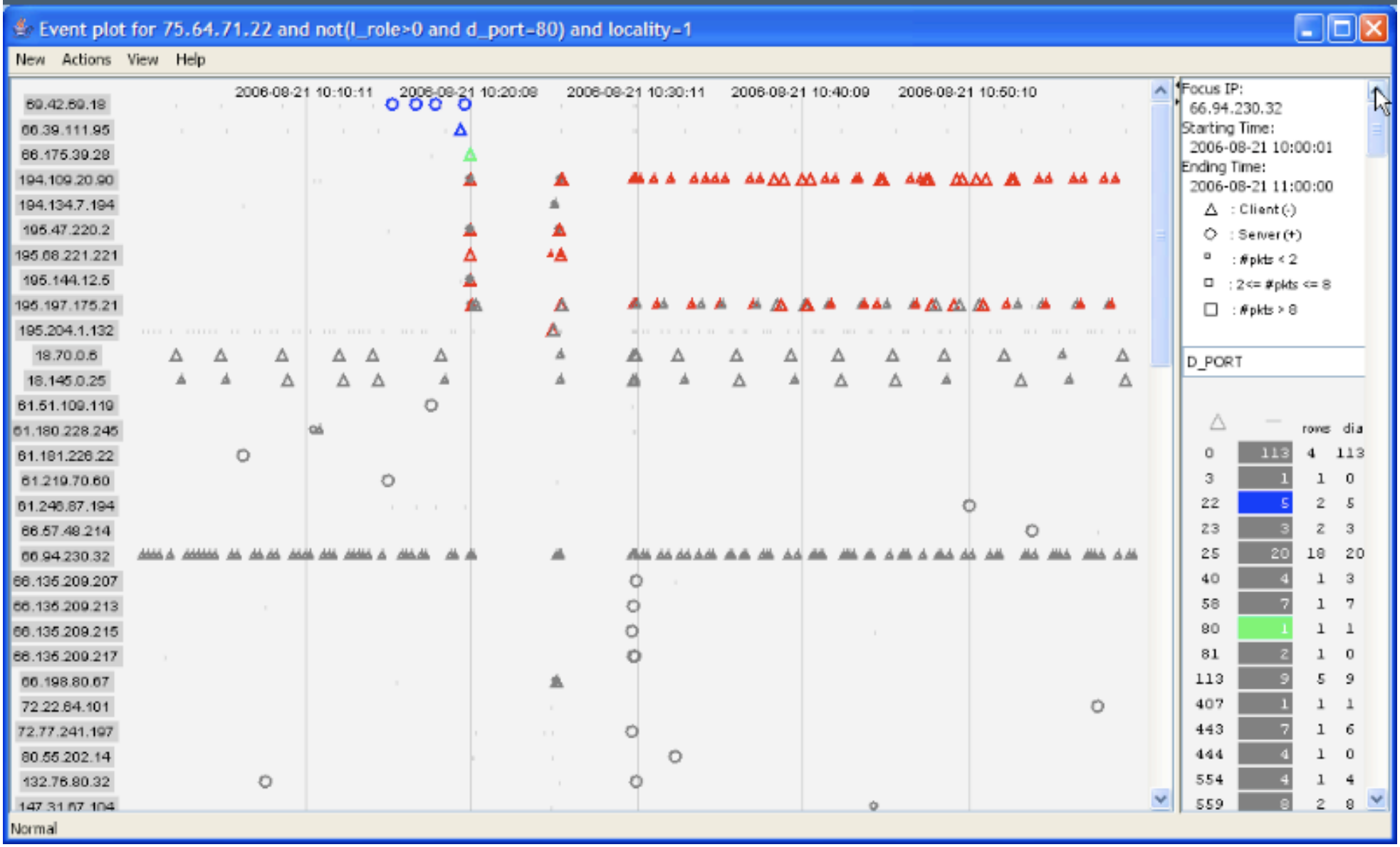
Ordinal event plot



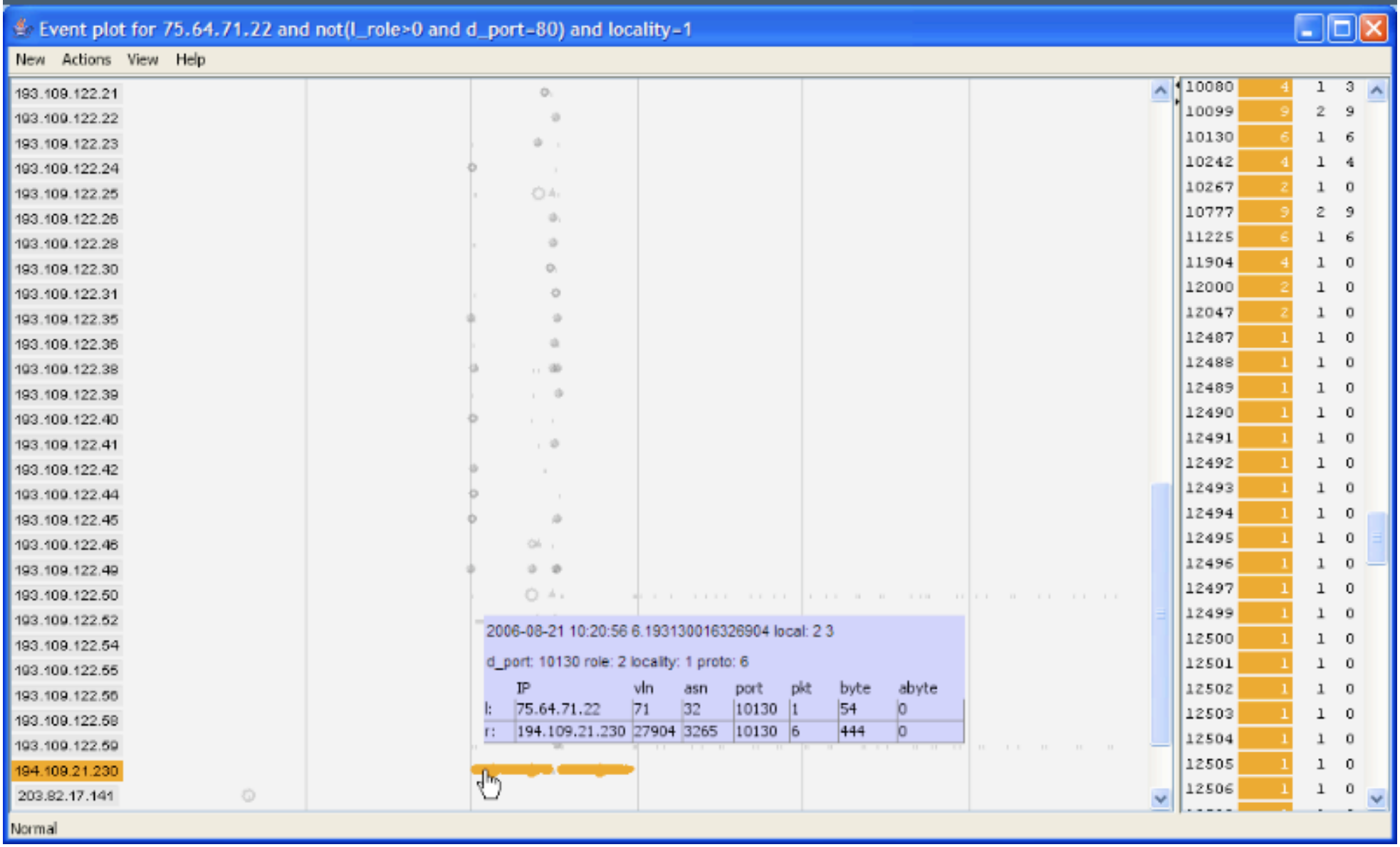
Switch to Ordinal Time



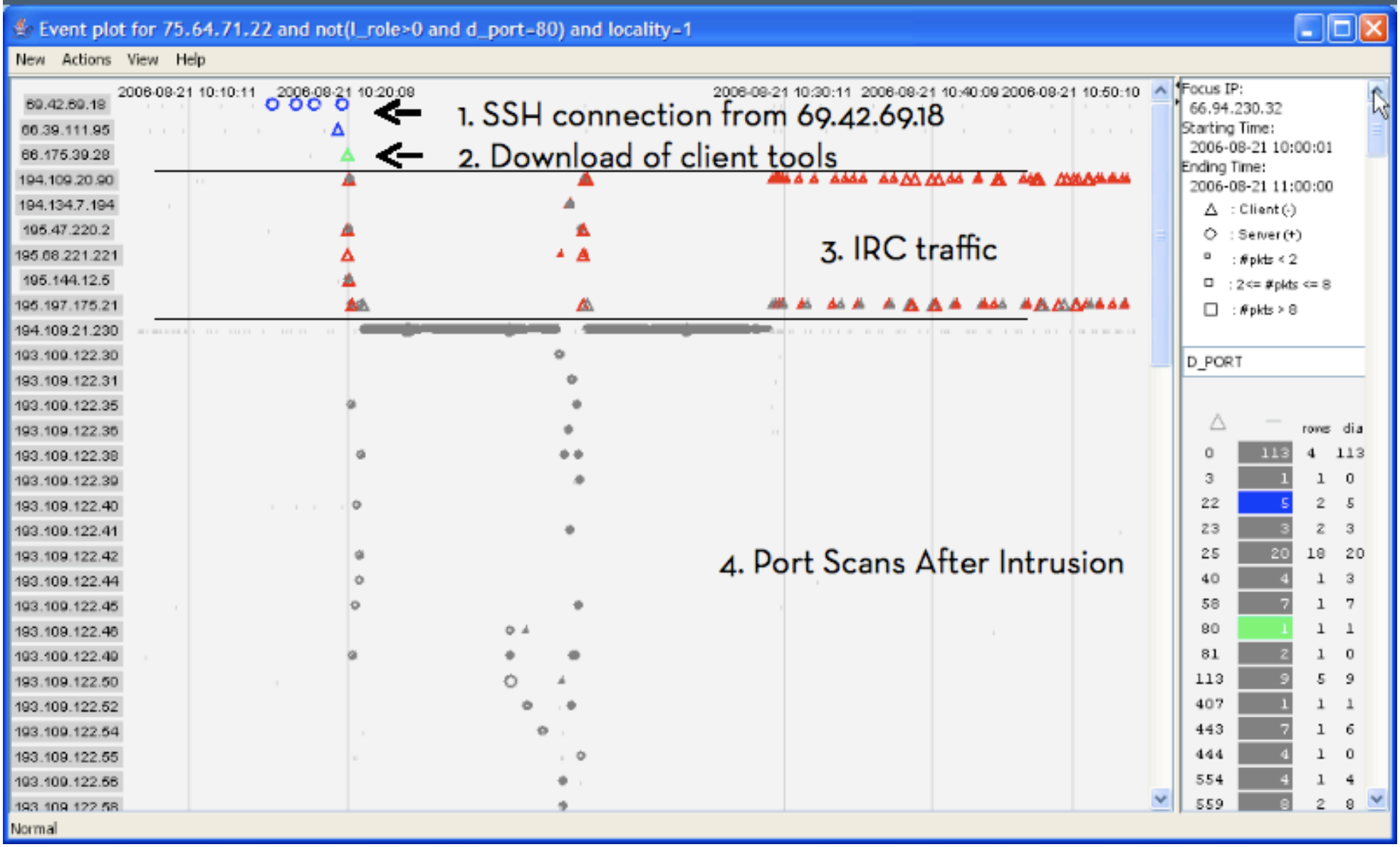
Switch to Ordinal Time



Switch to Ordinal Time

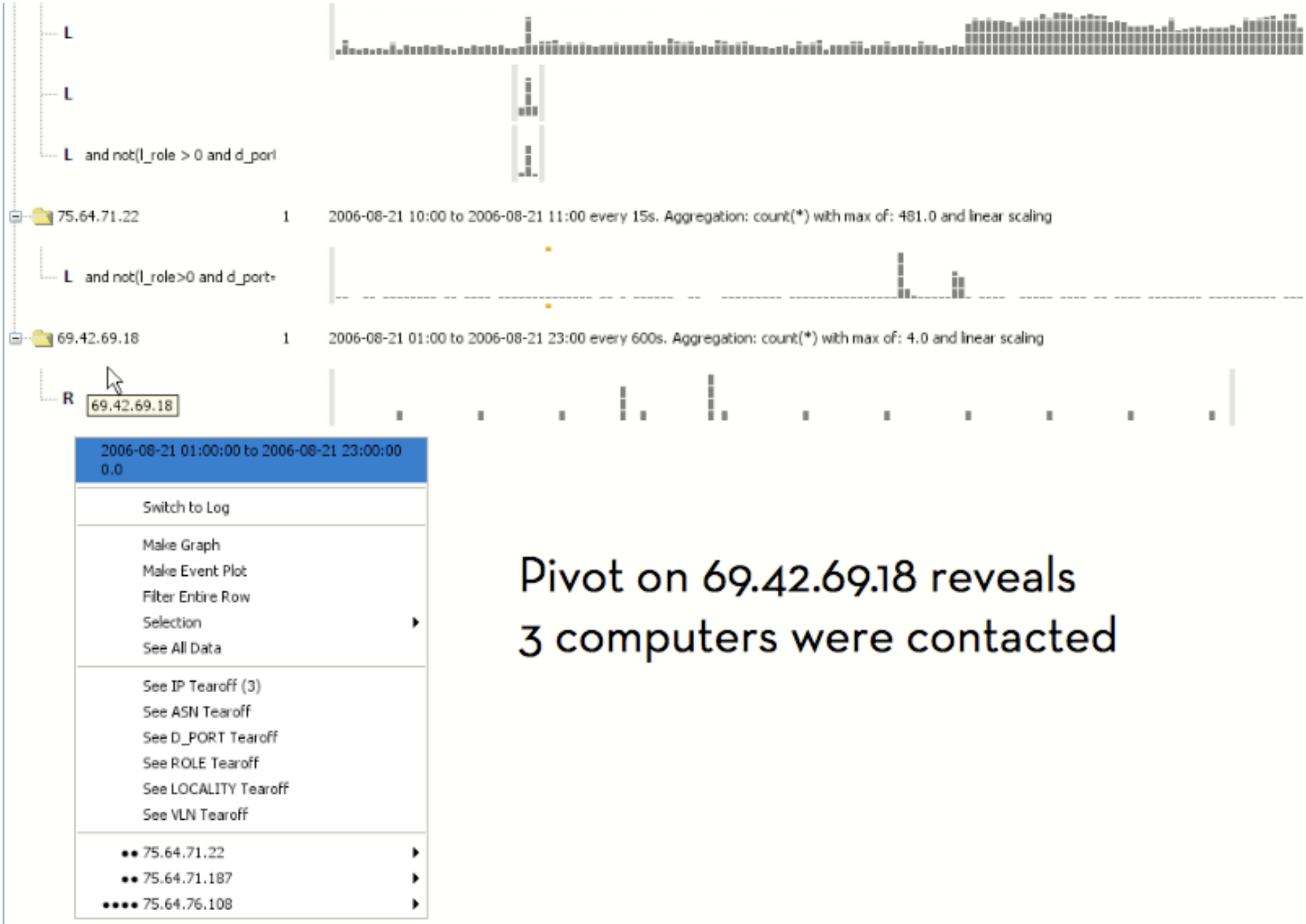


Final Sequence of Intrusion

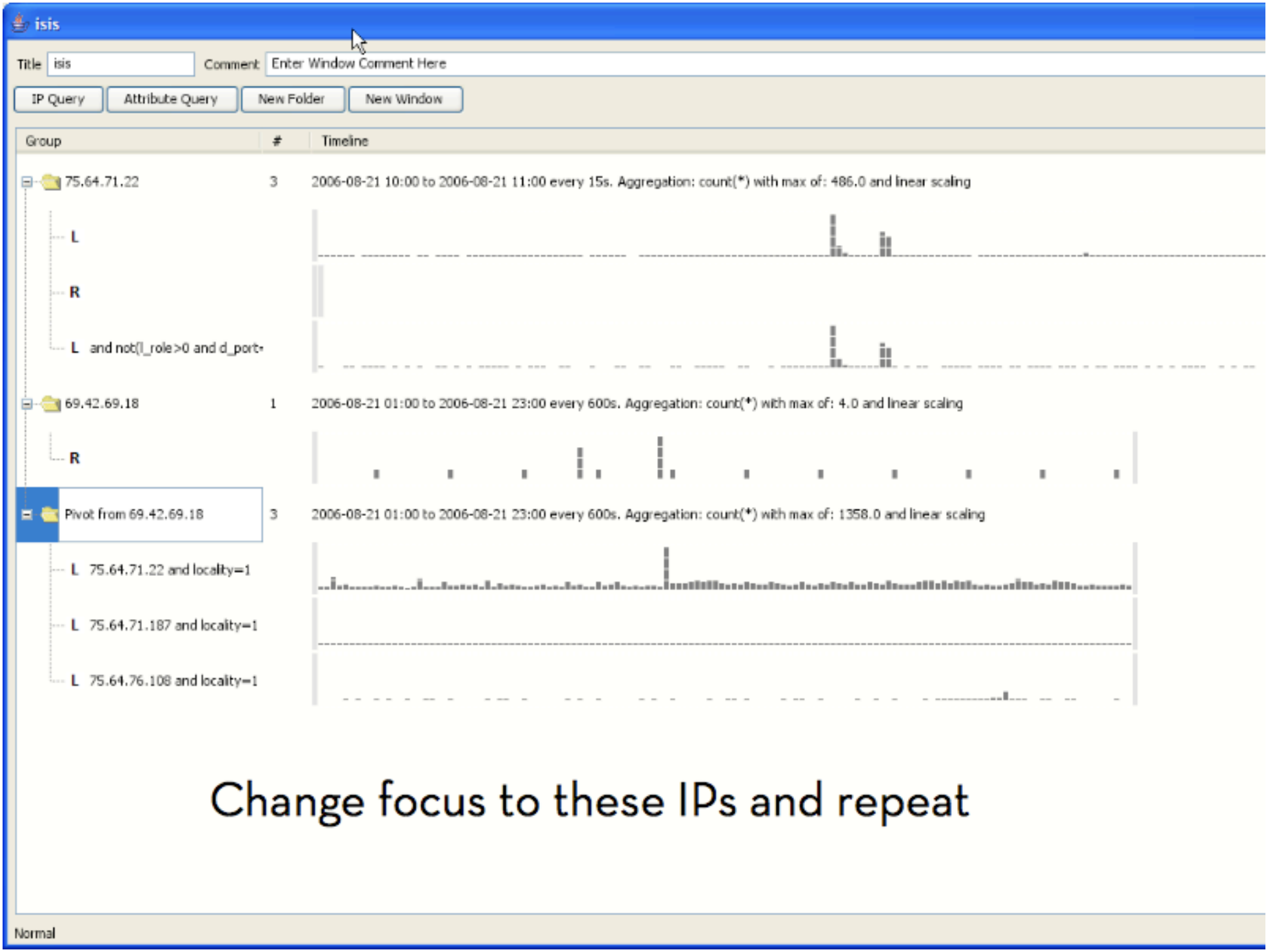


Were there other intrusions?

- Pivoting to show related events



Pivot on 69.42.69.18 reveals
3 computers were contacted



Change focus to these IPs and repeat

Feature Summary

- Timelines for overview and history
- Brushing for drilling down
- Event plots for sequencing
  - Collecting related events by brushing
  - Using ordinal time to make sequence clearer
- Pivoting to change investigative focus

EVALUATION

Evaluation Metrics

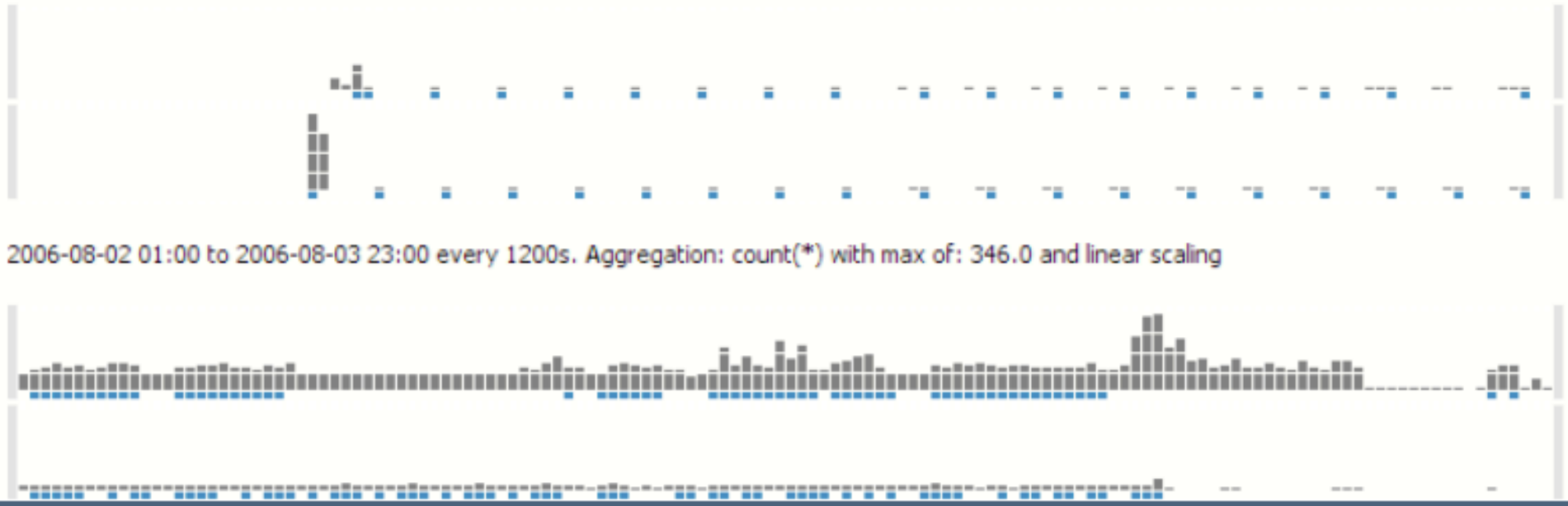
- Is investigation easier?
  - Task analysis of Isis
  - Epidemic case study
- Is Isis useful to analysts?
  - Historical intrusions
  - Observations from deployment

Task Analysis

- For each IP to investigate
  - Inspect traffic of IP
  - Compare traffic with others for correlations
  - Refine traffic query if necessary
  - Pivot to see related traffic if necessary

Timelines Accelerate Inspection

Easy to see periodicity or bursts of activity



Brushing Accelerates Inspection

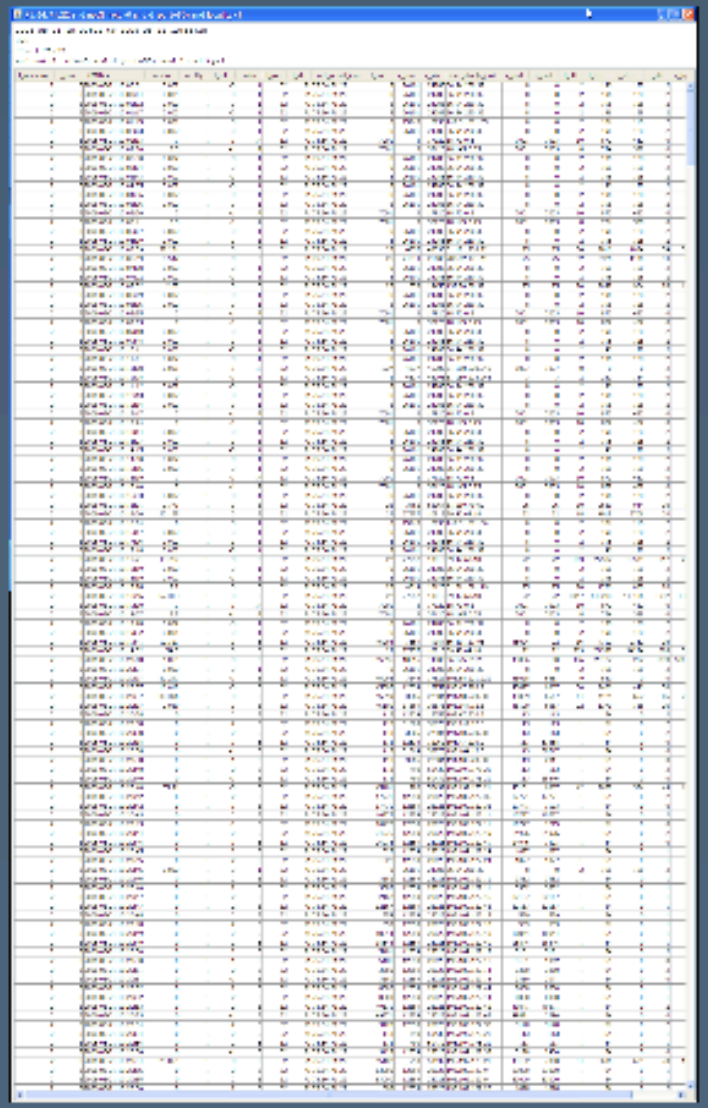
Shows temporal patterns of useful dimensions

- IP – the unit of investigation
- Port – can indicate traffic type
- Role – switching from client to server
- Subnet – defines administrative control
- ASN – useful for finding adjacent IPs

Making Comparisons Easier

1432 events in one hour

... as an event table



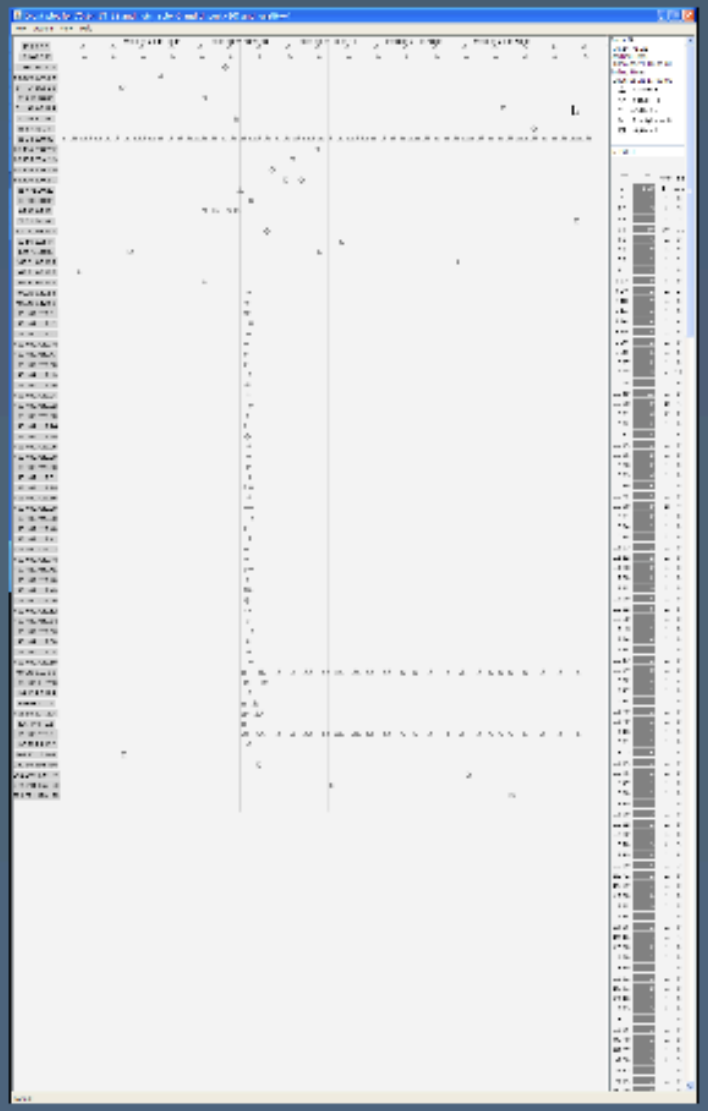
1920 pixels high

8% visible

Making Comparisons Easier

1432 events in one hour

... as an event plot



1920 pixels high

100% visible

Making Comparisons Easier

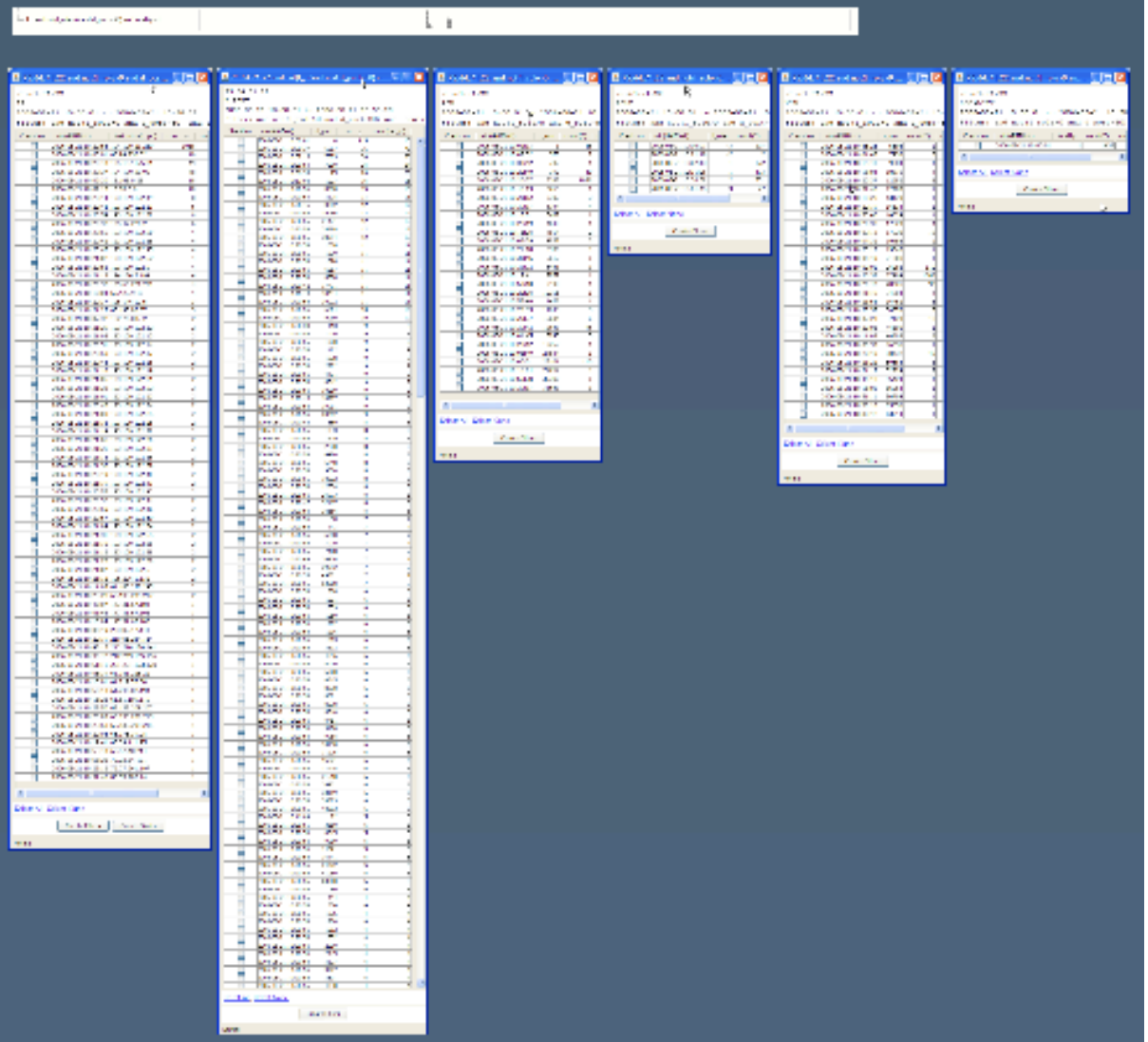
1432 events in one hour

... as a timeline

Making Comparisons Easier

1432 events in one hour

... as a timeline with tearoffs



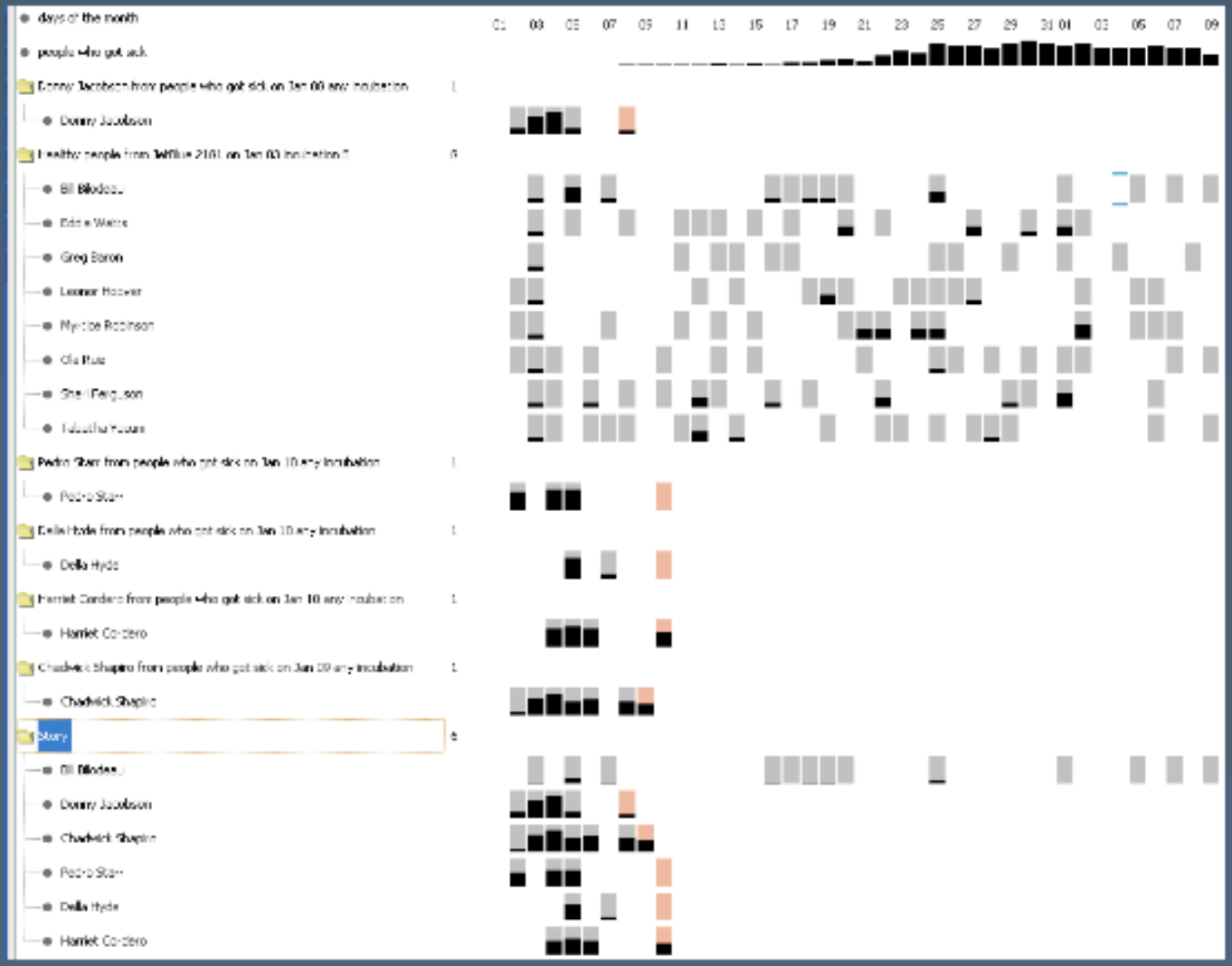
Epidemic Case Study

- Applied progressive multiples to investigation of a simulated epidemic
- Single infected person gets others sick who are on the same flight

Epidemic Study Setup

- 8 subjects, graduate students in engineering
- Analyzed two simulated epidemics using talk-aloud protocol, followed with interview
- Subjects must locate initial carrier. Limited number of tests to encourage exploration

Example Epidemic Investigation



Epidemic Study Results

- First try much slower than the second run (30-50 minutes vs. 5 minutes)
- Depth-first vs. breadth-first search behavior leads to some users getting lost
- Users found brushing useful and wanted to make its effects less transient

Evaluation Metrics

- Is investigation easier?
  - Task analysis of Isis
  - Epidemic case study
- Is Isis useful to analysts?
  - Historical intrusions
  - Observations from deployment

43 Historical Network Intrusions

Symptom	Count
IRC Usage	24
Scanning others	12
Open FTP	2
DOS or Spam	2
User noticed odd behavior	2
Not reported	1

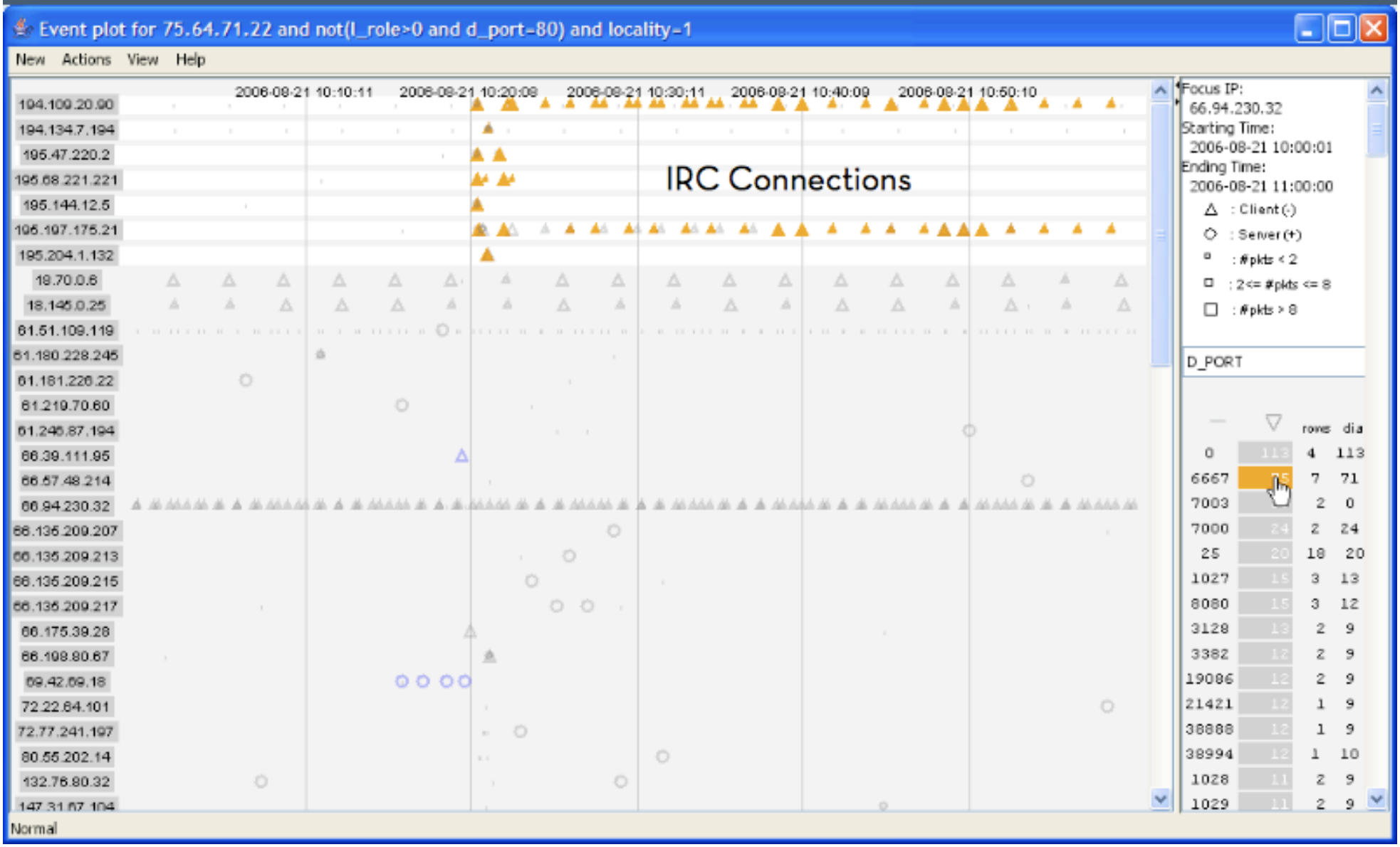
Observations from Deployment

- Used in EE and CS on live data since July 2007; investigated 5 intrusions so far
- 2/3 of time inspecting system logs, 1/3 of time inspecting flow records
- What were useful features?

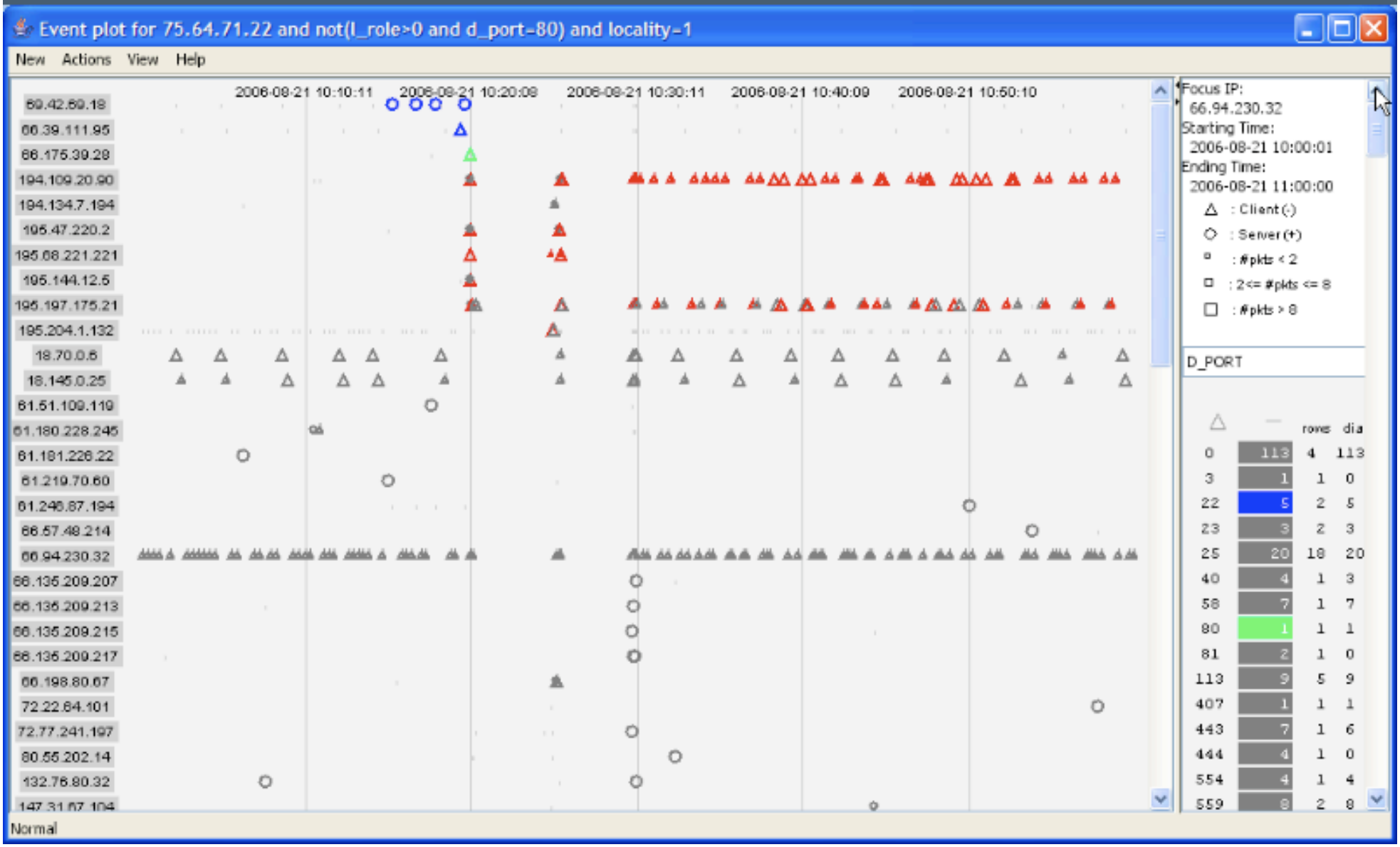
Useful Features

- History and comparison of timelines
- Pivoting and filtering of timelines
- Brushing of related events
- Collecting events that were brushed
- Ordinal time in event plots
- Filtering and aggregation in SQL

Collecting Brushed Events



Ordinal Time



Filtering and Aggregation in SQL

- Filtering
  - and not(l_role>0 and d_port=80) and locality=1
- Aggregation
  - max(l_pkt+r_pkt)
  - count(distinct(seq_id))

- Provide useful SQL examples to analyst

Design Recommendations

- Use timelines instead of node-link diagrams when sequencing events.
- Provide dynamic previews to make an exploration history more meaningful.
- Allow users to save the results of a dynamic preview in a static form.
- Support both continuous and ordinal views of time.

Isis in the Future

- Continued use in EE and CS departments
- Other security analysts
  - Medical School
  - US CERT

Making Investigation Easier

Investigations require identifying and sequencing events that are usually discovered out of order

1. Timelines make temporal relationships apparent
2. Layout provides a visible history
3. Pivoting shows more related events
4. Brushing makes existing relationships evident
5. Event plots allow sequencing and reordering

Thesis Statement

Time-oriented visualizations of network activity that provide a visual exploration history and affordances for identifying relevant and related events can increase the effectiveness of forensic analysis.

Contributions

- Demonstrated automatically generated flow maps for presentation and progressive multiples for event analysis over networks

- Designed Isis, a system for investigating network incidents, based on progressive multiples, validated with a long term deployment

Publications

Doantam Phan, John Gerth, Marcia Lee, Andreas Paepcke, and Terry Winograd. *Visual Analysis of Network Flow Data with Timelines and Event Plots*. VizSEC 2007 (Best Paper)

Doantam Phan, Andreas Paepcke, and Terry Winograd. *Progressive Multiples for Communication-Minded Visualization*. Graphics Interface 2007

Doantam Phan, Ling Xiao, Ron Yeh, Terry Winograd, and Pat Hanrahan. *Flow Map Layout*. Infovis 2005

Remarks

- Combine HCI with visualization research
- Develop and evaluate visualization techniques by working with real users who have real problems

Network Administrators

- John Gerth

- Daniel Paepcke, Charlie Orgish, Joe Little, Mike Newton, Jason Conroy, Brian Roberts, Pat Burke, Martin Frost, Miles Davis, Thoi Nguyen

Faculty and Staff

- Terry Winograd
- Pat Hanrahan
- Scott Klemmer
- Andreas Paepcke
- Barbara Tversky

- Heather Gentner
- Ada Glucksman
- NDSEG, ARCS (Sandy and Paul Otellini), Stanford RVAC

Acknowledgements

- Marcia Lee

- Beta testers: Dave Akers[2], Jeff Klingner[2], Kayvon Fatahalian, Mike Houston, Ron Yeh[2], Neil Patel, Bryan Chan, Joel Brandt, Paul Dourish, Wendy Ju, Tico Ballagas, Jeremy Sugerman, Tal Garfinkel, Ling Xiao, Christina Chan, Björn Hartmann[2], Tim Knight, Austen McDonald, Brian Lee, many others

- Friends and Family from near and far

THE END