Incorporating Network Flows in Intrusion Incident Handling and Analysis

John Gerth
Stanford University

gerth@stanford.edu

EE/CS Network Infrastructure

- Three buildings with one router
  - (Gates) Computer Science
  - (Packard) Electrical Engineering
  - (Allen) Center for Integrated Systems
- Composition
  - 25 VLANs controlled by disparate groups
  - 10,000 IP addresses (about half are active)
  - Eclectic mix of Windows, Linux, Solaris, OS-X, …
  - No firewall beyond minor university filters
- Analysts
  - A half-dozen people with network (and other) responsibilities

Incident Investigation Process

- Find answers to a set of classic questions…
  - Who
  - What
  - When
  - Where
  - Why
  - How
- …using an iterative process
  - Inspect events of a focus node
  - Augment, refine, filter data
  - Compare events of related nodes, looking for correlation
  - Pivot on an “interesting” node to refocus

Network Data Sources
(each step is orders of magnitude more volume)

- **Traffic counters** (SNMP, MRTG, ….)
  - Configurable in network devices
- **Event/Alert logs** (Syslog, HTTPD, SNORT, ...)
  - Collected by firewalls, IDS, individual machines and services
- **Flows** (Netflow, YAF, Argus, ….)
  - Typically collected at border routers or taps
- **Packet Headers / Traces** (tcpdump, wireshark, …)
  - Collected at switches, routers, or taps

Network Flows

- Advantages
  - Relatively uniform and increasingly available
  - Hard to subvert
  - Mitigate privacy concerns
  - Largely insensitive to encryption
- Disadvantages
  - Still voluminous compared to event logs
  - Aggregate measure
  - Lack content

Flow Capture and Data Management

- Sensor
  - Span ports from two Cisco backbone switches
  - See all layer 3 traffic for three buildings (not just external)
  - Argus capture of bidirectional ICMP, UDP, TCP flows
- Collector
  - Raw flows from sensor are multicast locally in realtime
  - Hourly files from sensor compressed and archived
  - 20-30M (peak 70M) Argus flows/day (~1G compressed)
  - Retain several months of data online for analysts to access

Support flat files and database tables

- Flat text files
  - Familiar and familiar tools
  - Extracts useful for exchange and reporting
  - Straightforward sequential processing
  - Import to other tools for aggregation and analysis
- Relational databases
  - No longer exotic
  - Suitable for large data volumes
  - Greater expressibility for queries
  - Built-in support for aggregation and analysis

Database Infrastructure

- MySQL server running on collector
  - Live flows from sensor inserted in real-time
  - Daily tables recreated from archived raw flows
  - Monthly “merge” tables
  - Anonymize extracts for research with CryptoPAN
- Flow schema tuning
  - Transform src/dst to local/remote
  - Add ASN (routeviews.org) and local VLAN metadata
  - Convenience columns for locality, local role, dst port
  - Index most dimensions (adds about 50%)
  - Tables + indices ~2G/day

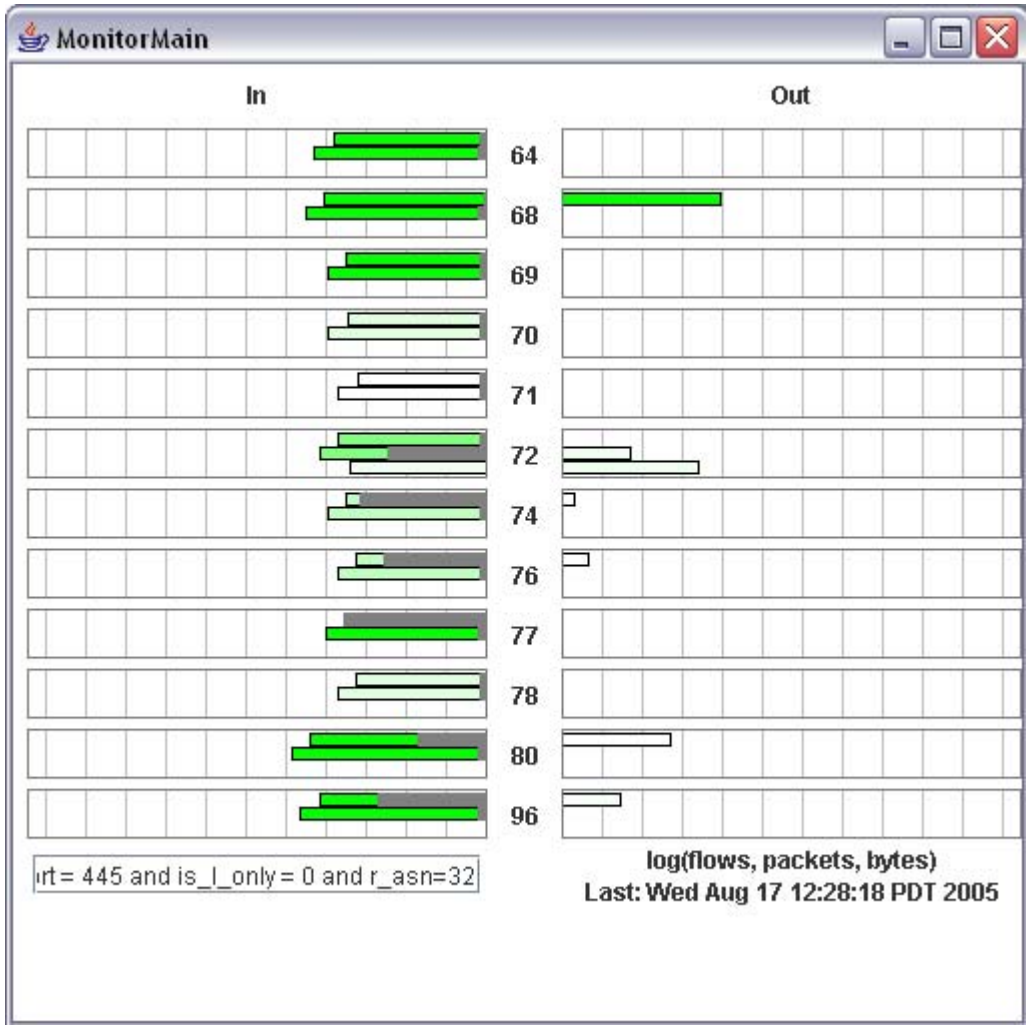
Flows in Incident Handling

- Worms and Trolls
  - Volume and promiscuity
- Immaculate Intrusions
  - Scrubbers, Keyloggers, and Remote Tunnels
- Botnets
  - Beaconing to Command+Control Hosts

Traffic Volume

- Windows Esbot worm circa 2005
  - Spread via PNP buffer overflow
  - Installed backdoor trojan
  - Victim turns into attacker
- Report
  - Overall traffic suddenly increased an order of magnitude
- Analysis
  - Flow distribution showed port 445 at 500-1000 flows/sec
  - Keyed on 445 traffic to identify attackers
  - Used “flow monitor” to reveal local compromises

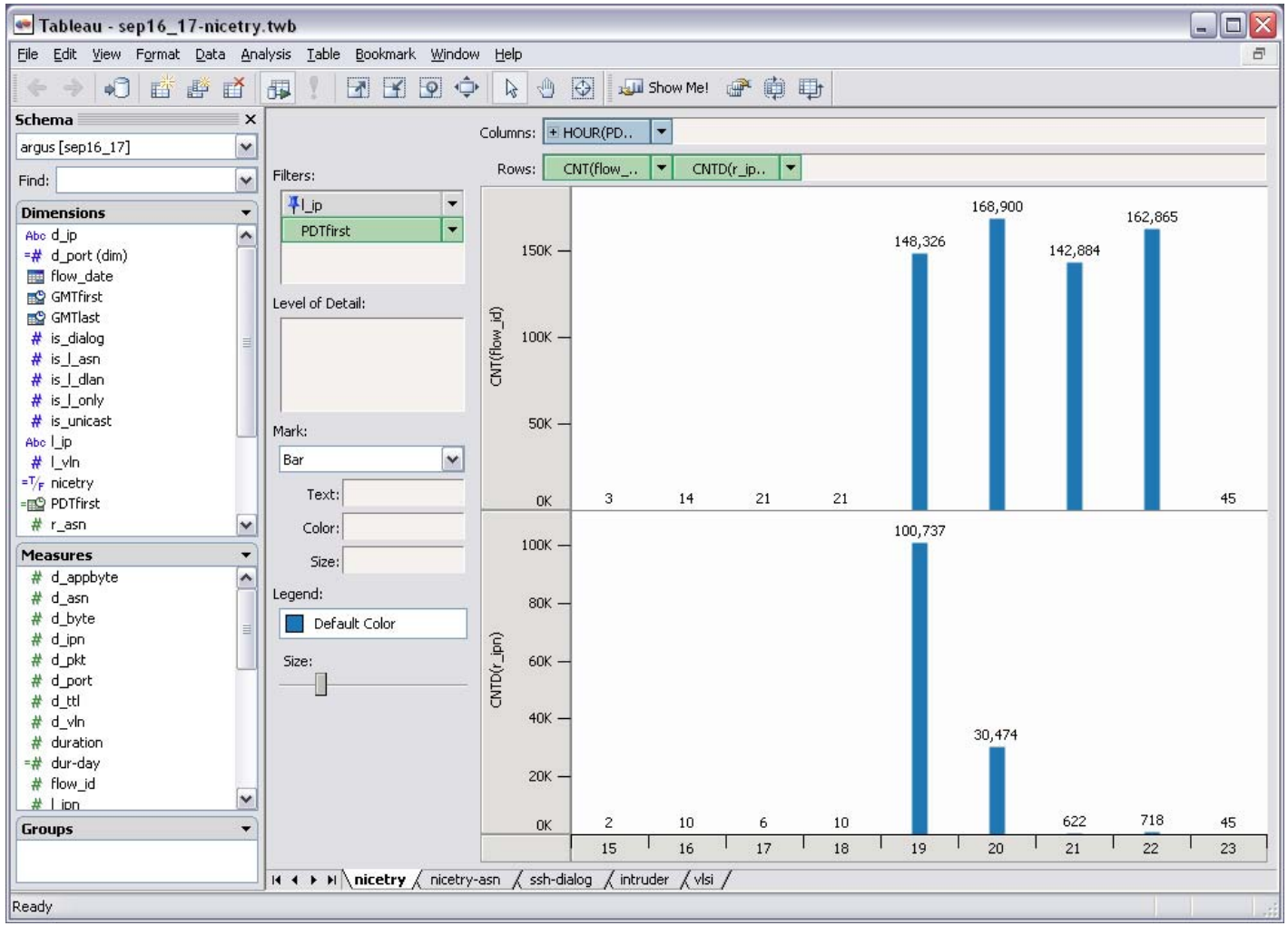
Esbot on the Flow Monitor



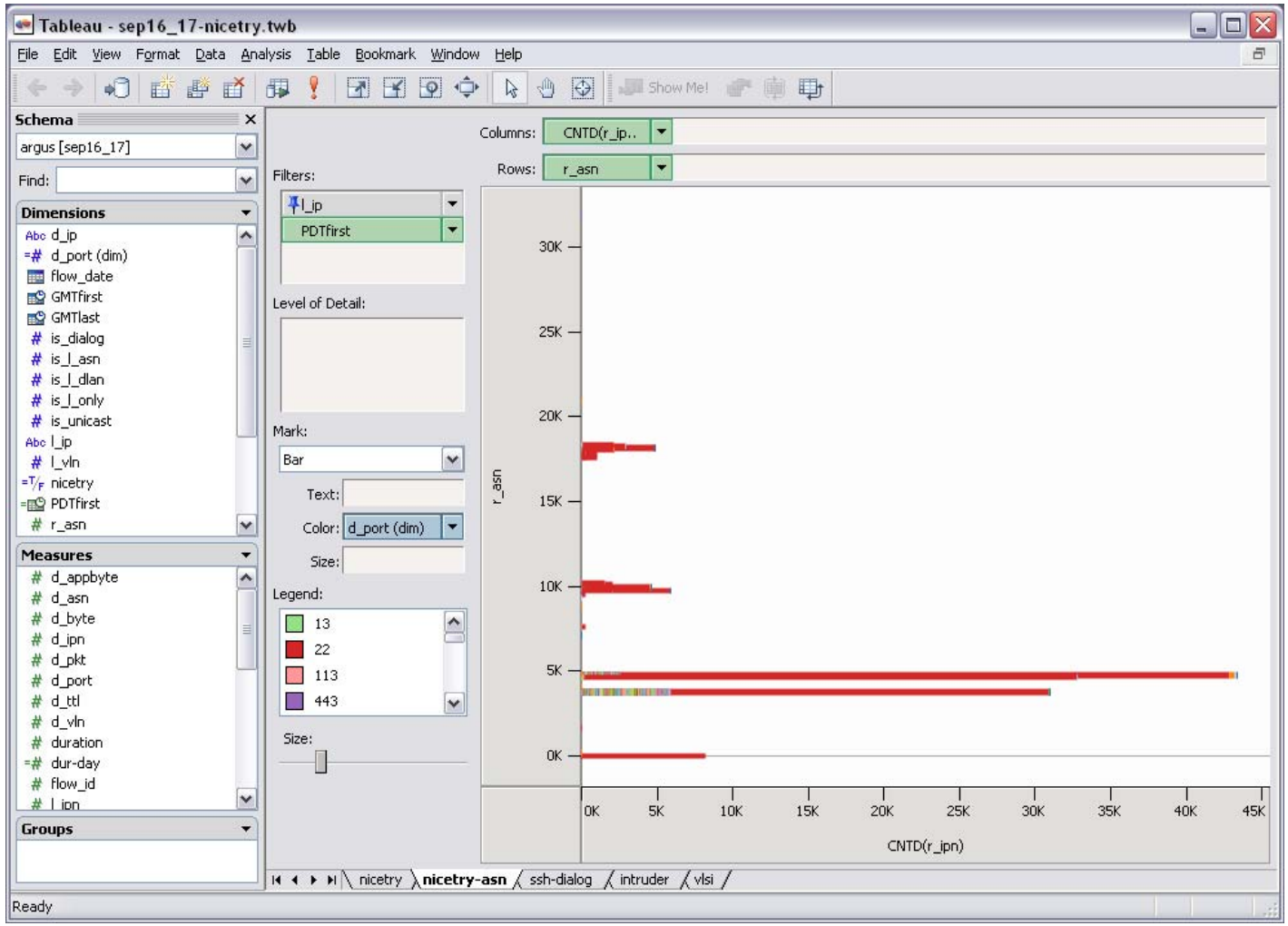
Promiscuity

- SSH Troll
  - Intruder gains access to local machine
  - Installs SSH troll
  - Launches attack on remote networks
- Report
  - Odd outbound traffic spike from local IP
- Analysis
  - Flow distribution showed many IPs, few ASNs, single port
  - Backtrack in time to find initial SSH compromise
  - Pivot reveals other victims

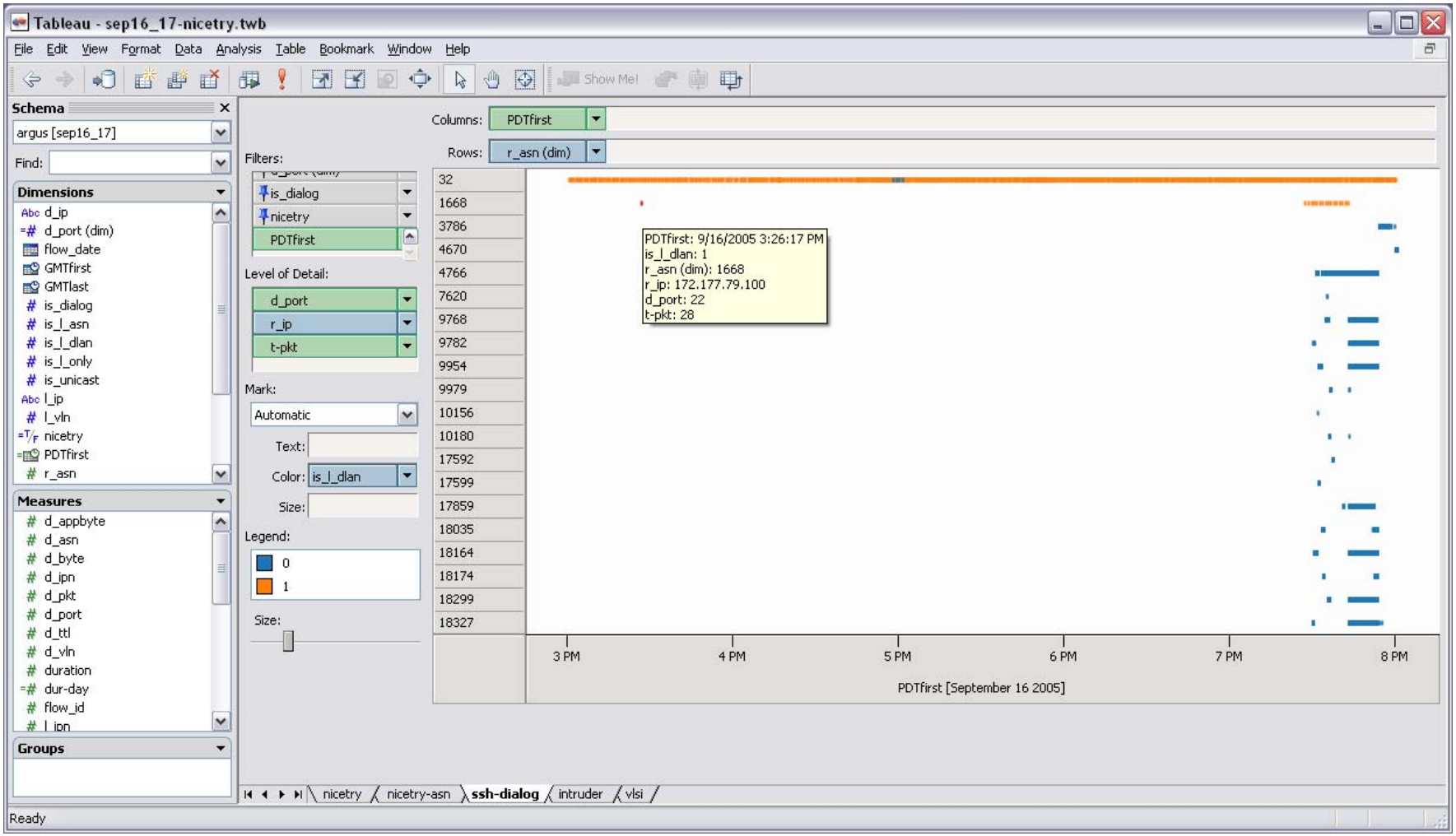
SSH Troll: Volume + Promiscuity



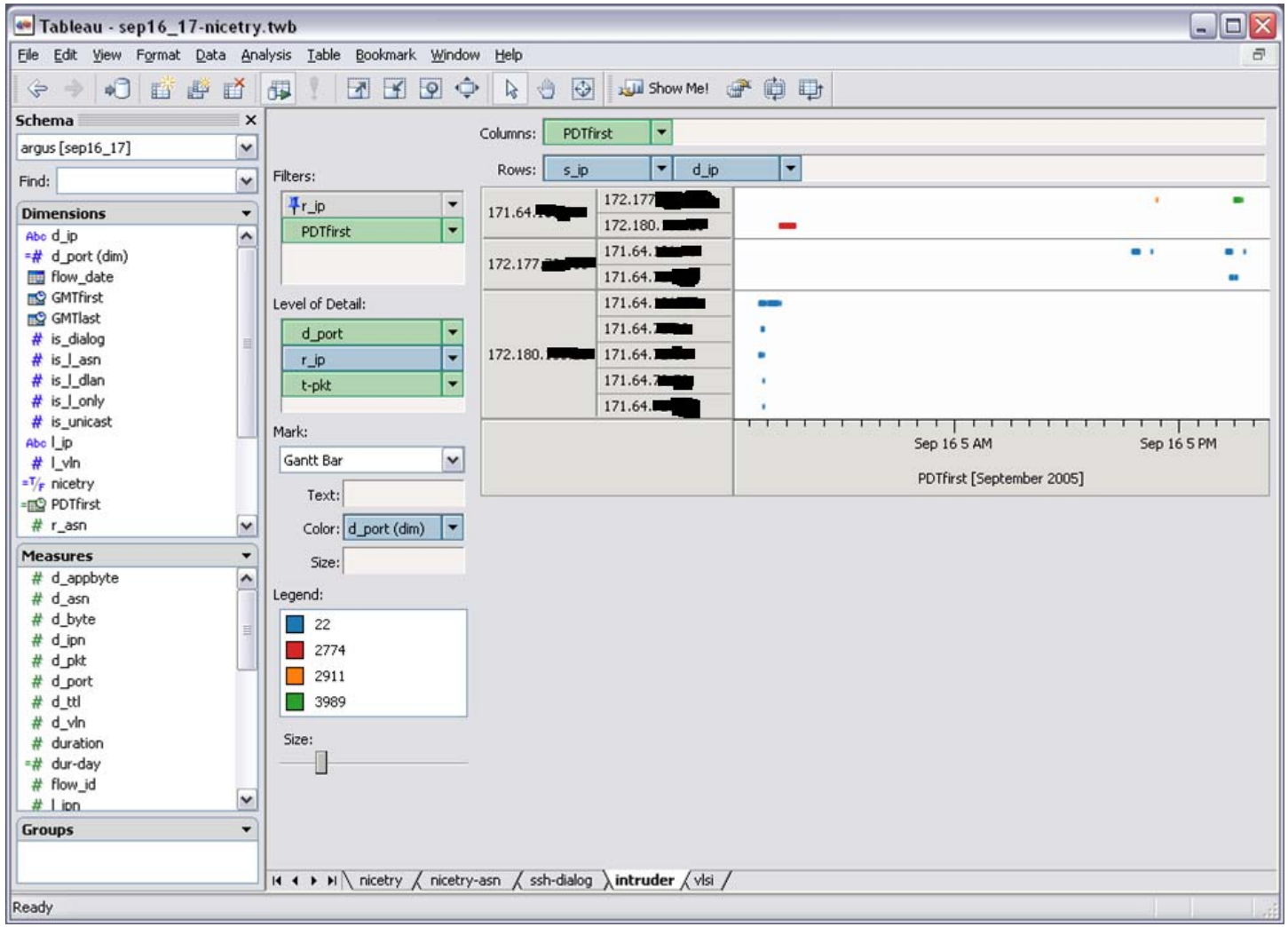
SSH Troll: Identifying targets



SSH Troll: Locate Compromise



SSH Troll: Pivot to identify other victims



Immaculate Intrusions - Keyloggers

- Unprotected X-Window server
  - Intruder maps 0x0 pixel client and signs up for keypress events
  - Steals credentials for other machines from local user
  - Uses credentials to login to experimental machine
- Report
  - Experimental machine crashes when intruder’s tools fail
- Analysis
  - Local user logged in when user not present
  - Discover open X-server on user’s desktop machine
  - Backtrack in time to find keylogger flows
  - Pivot reveals other victims

Immaculate Intrusions - Scrubbers

- Unpatched Linux machine
  - Unpatched server vulnerable to remote root compromise
  - Intruder installs backdoor, trojan binaries, and scrubs logs
  - Uses trojan ssh to steal credentials of local users
  - Uses ssh known_hosts data to attack other local machines
- Report
  - Local machine two hops away found sending spam
- Analysis
  - Backtrack of login sessions leads to compromised machine
  - Trojan binaries found, but no plausible root logins
  - Flow logs show original compromise and backdoor logins
  - Pivot reveals other victims

Immaculate Intrusions - Tunnels

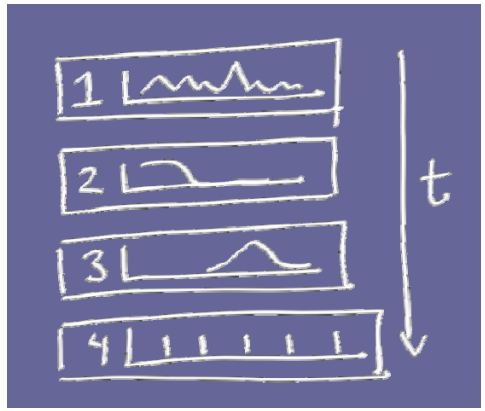
- Tunnels
  - Intruder compromises desktop machine running VNC client
  - Desktop machine has forwarded ports over ssh-tunnel
  - Intruder's traffic is tunnelled and reparented inside cluster
- Report
  - Apparent Nessus scan of ***isolated*** cluster machine
- Analysis
  - System logs of head node show no logins
  - Flow logs show massive ssh traffic from compromised machine

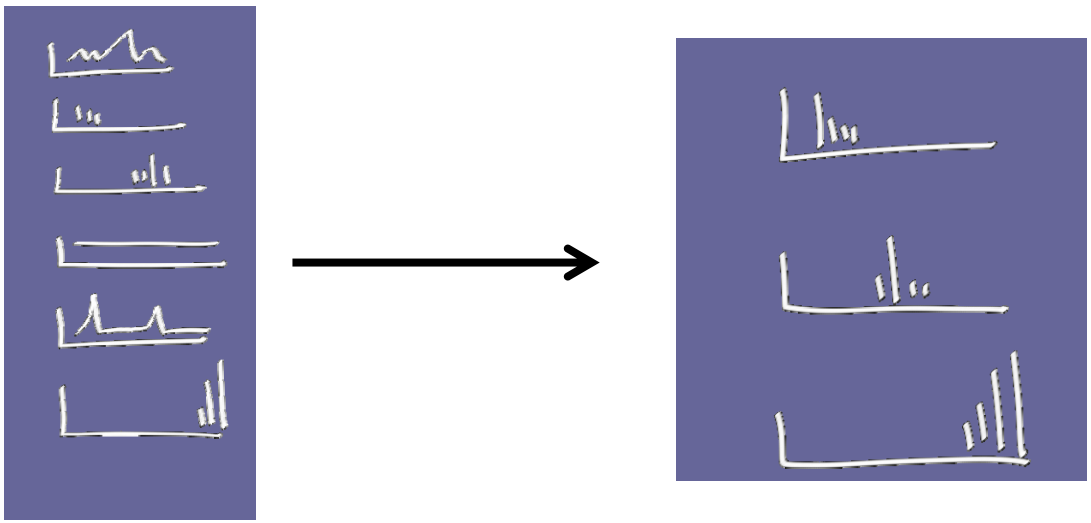
Isis:Visual Analysis of Flow Data

(see paper by Phan et al in VizSec 2007)

Progressive Multiples

- Make exploration history visible
- Reorder rows to reveal structure and event sequencing

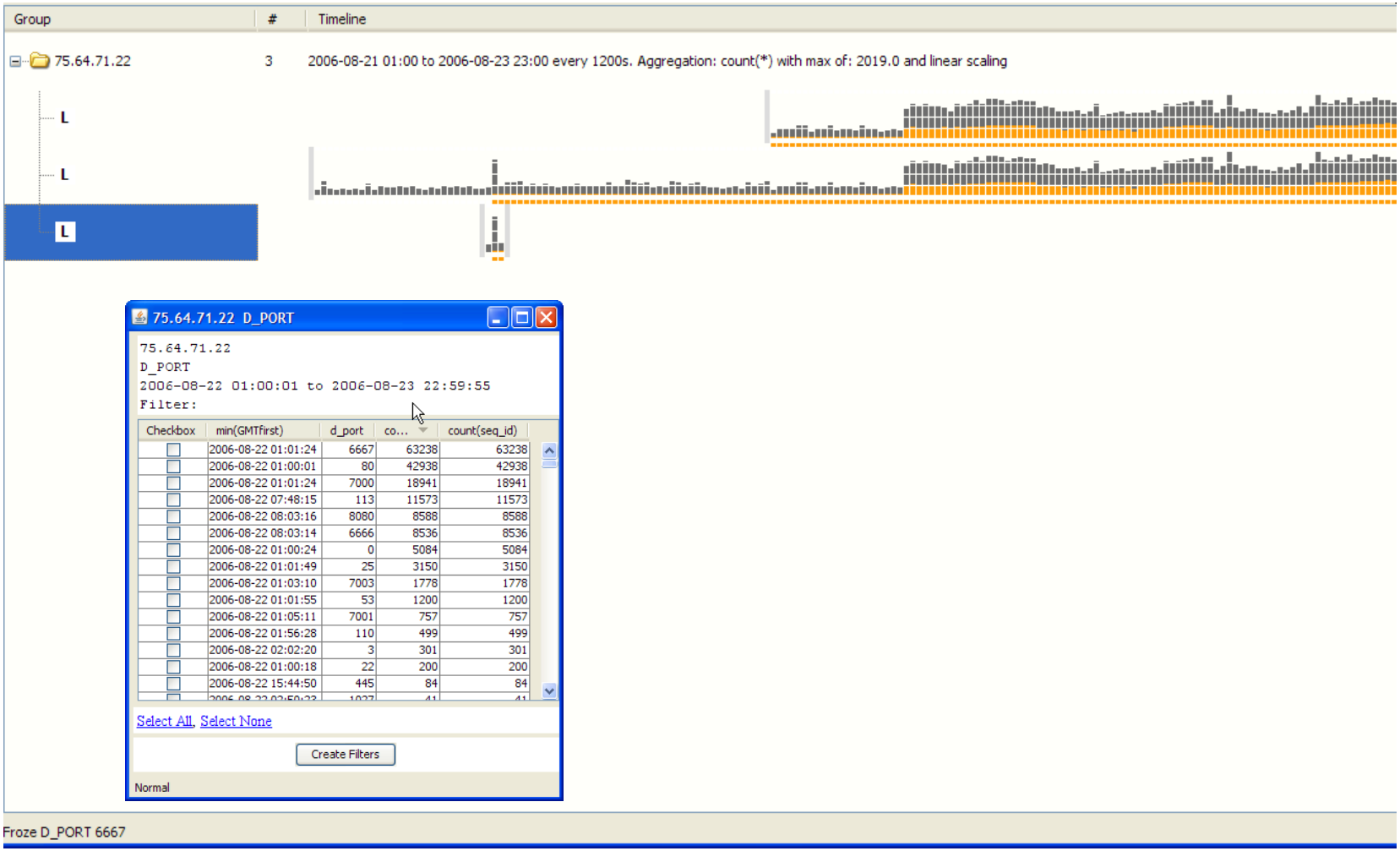




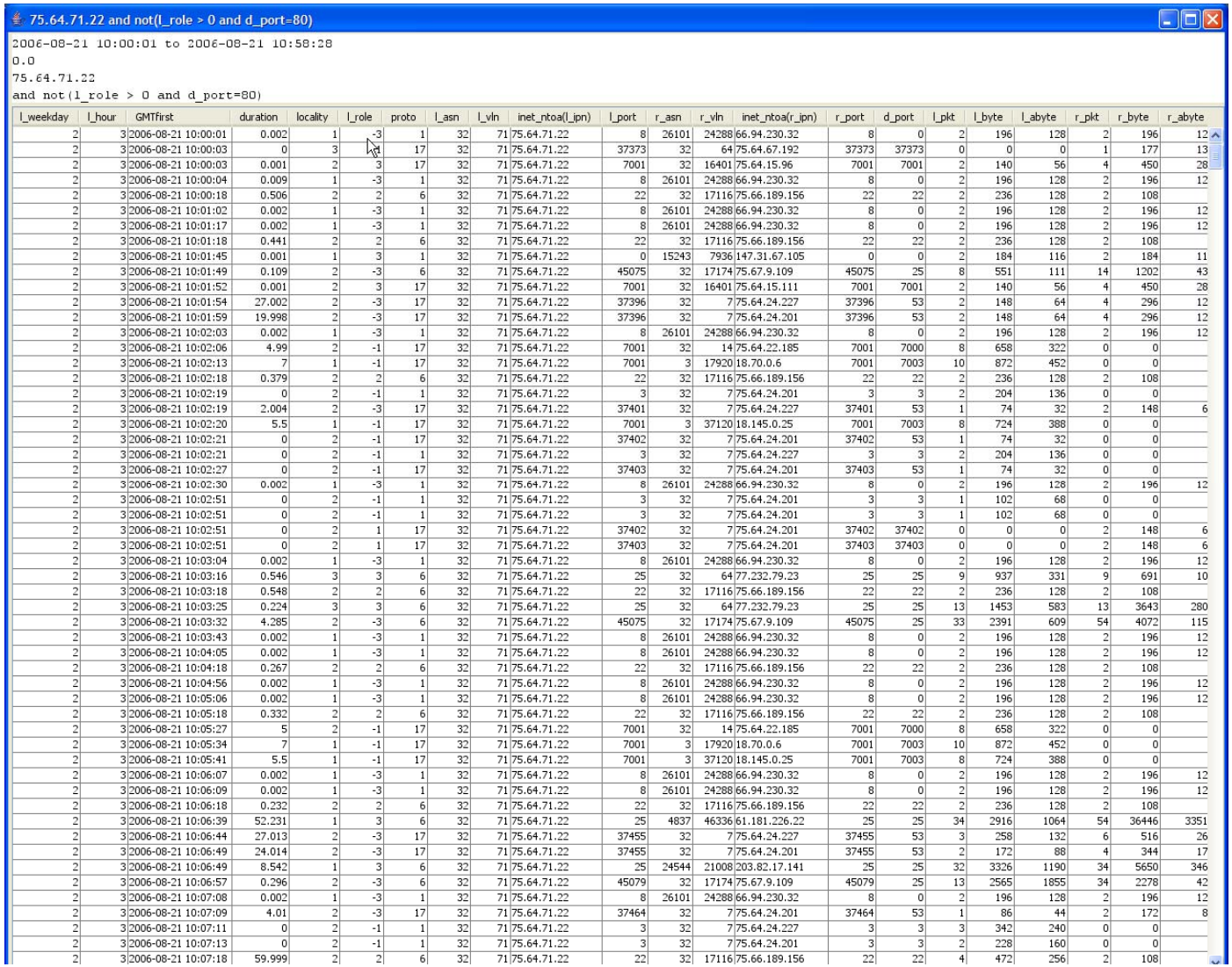
Beaconing

- Botnet zombie
  - Intruder gains access to local machine
  - Installs IRC client bot
  - zombie bot “calls home” periodically
- Report
  - Recurrent traffic to suspect IRC servers
- Analysis
  - Backtrack in time to find initial compromise
  - Observe tool download and installation
  - Pivot …

IRC bot: Timeline Investigation



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 10:58:28
0.0
75.64.71.22
and not(l_role > 0 and d_port=80)

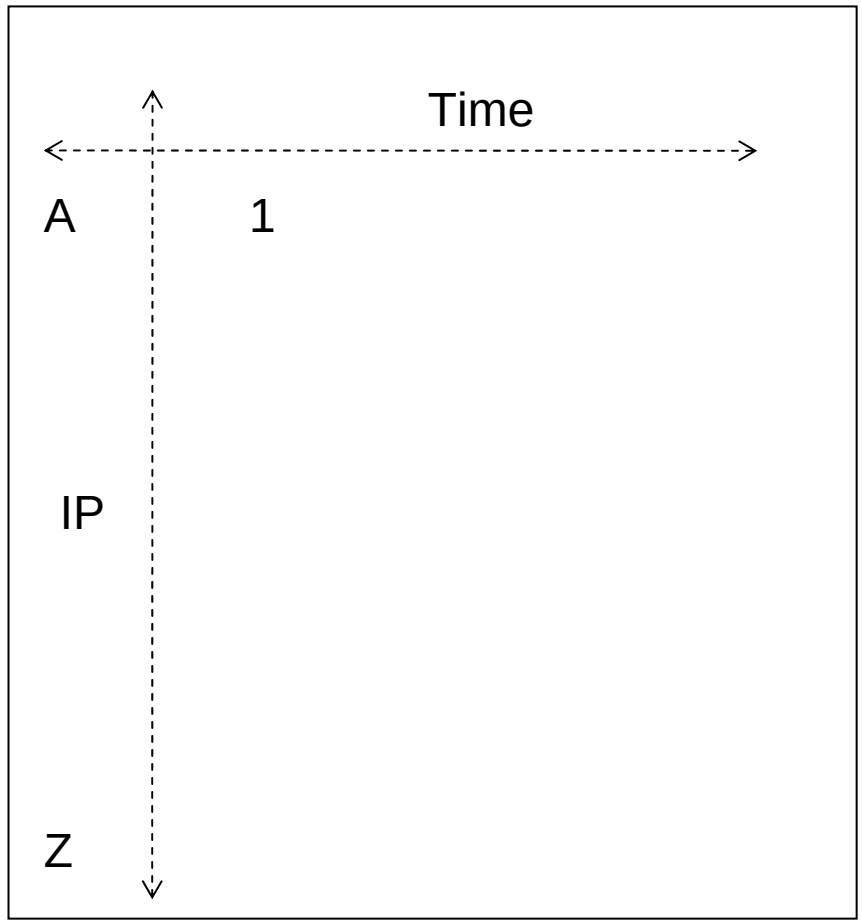
l_weekday	l_hour	GMTfirst	duration	locality	l_role	proto	l_asn	l_vln	inet_ntoa(l_ipn)	l_port	r_asn	r_vln	inet_ntoa(r_ipn)	r_port	d_port	l_pkt	l_byte	l_abyte	r_pkt	r_byte	r_abyte
2	3	2006-08-21 10:00:01	0.002	1	-3	1	32	71	75.64.71.22	8	26101	24288	66.94.230.32	8	0	2	196	128	2	196	12
2	3	2006-08-21 10:00:03	0	3	3	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	147.31.67.105	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	
2	3	2006-08-21 10:05:34	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 10:05:41	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 10:06:07	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 10:06:09	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 10:06:18	0.232	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 10:06:39	52.231	1	3	6	32	71	75.64.71.22	25	4837	46336	61.181.226.22	25	25	34	2916	1064	54	36446	3351
2	3	2006-08-21 10:06:44	27.013	2	-3	17	32	71	75.64.71.22	37455	32	7	75.64.24.227	37455	53	3	258	132	6	516	26
2	3	2006-08-21 10:06:49	24.014	2	-3	17	32	71	75.64.71.22	37455	32	7	75.64.24.201	37455	53	2	172	88	4	344	17
2	3	2006-08-21 10:06:49	8.542	1	3	6	32	71	75.64.71.22	25	24544	21008	203.82.17.141	25	25	32	3326	1190	34	5650	346
2	3	2006-08-21 10:06:57	0.296	2	-3	6	32	71	75.64.71.22	45079	32	17174	75.67.9.109	45079	25	13	2565	1855	34	2278	42
2	3	2006-08-21 10:07:08	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 10:07:09	4.01	2	-3	17	32	71	75.64.71.22	37464	32	7	75.64.24.201	37464	53	1	86	44	2	172	8
2	3	2006-08-21 10:07:11	0	2	-1	1	32	71	75.64.71.22	3	32	7	75.64.24.227	3	3	3	342	240	0	0	
2	3	2006-08-21 10:07:13	0	2	-1	1	32	71	75.64.71.22	3	32	7	75.64.24.201	3	3	2	228	160	0	0	
2	3	2006-08-21 10:07:18	59.999	2	2	6	32	71	75.64.71.22	22	32	17116	75.66.189.156	22	22	4	472	256	2	108	

From Event Table to Event Plot

Event Table

1	Time	A	…	Measures

Event Plot

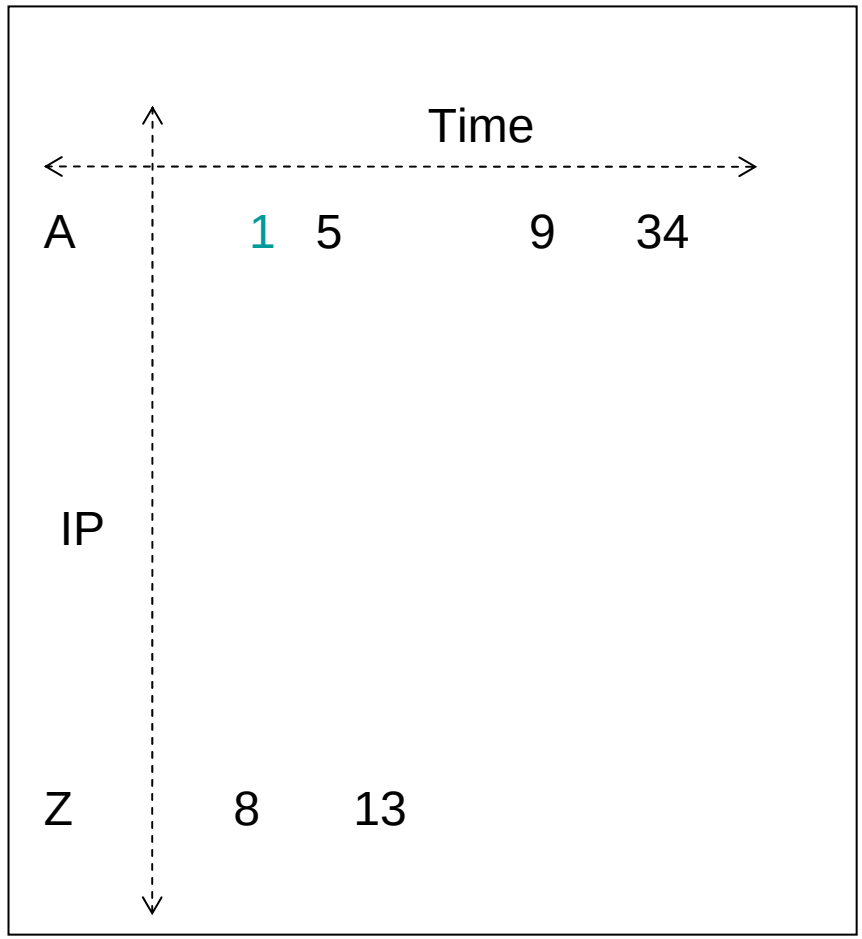


From Event Table to Event Plot

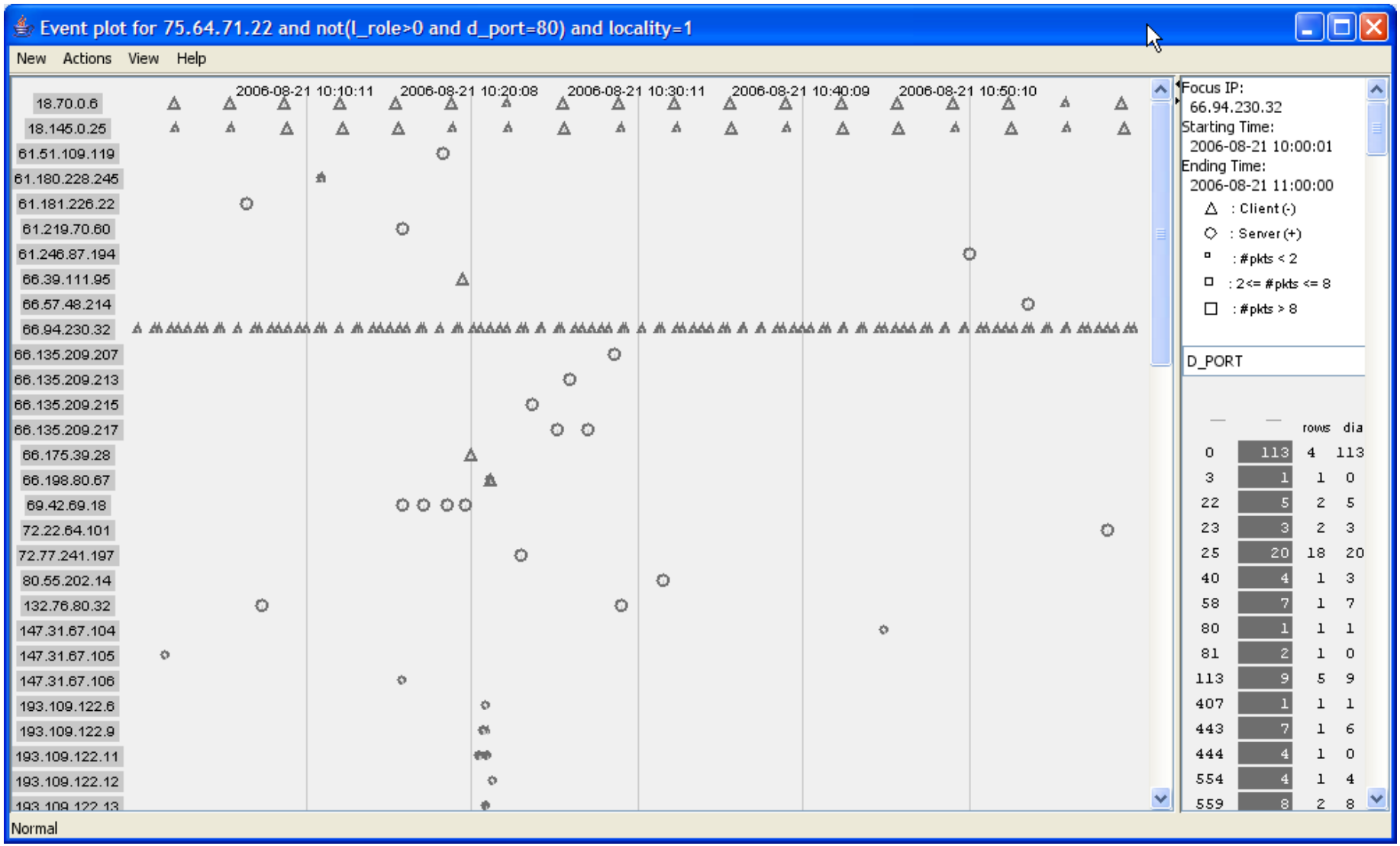
Event Table

1	Time	A	…	Measures
. .				
#	Time	IP	…	Measures
. .				
n	Time	Z	…	Measures

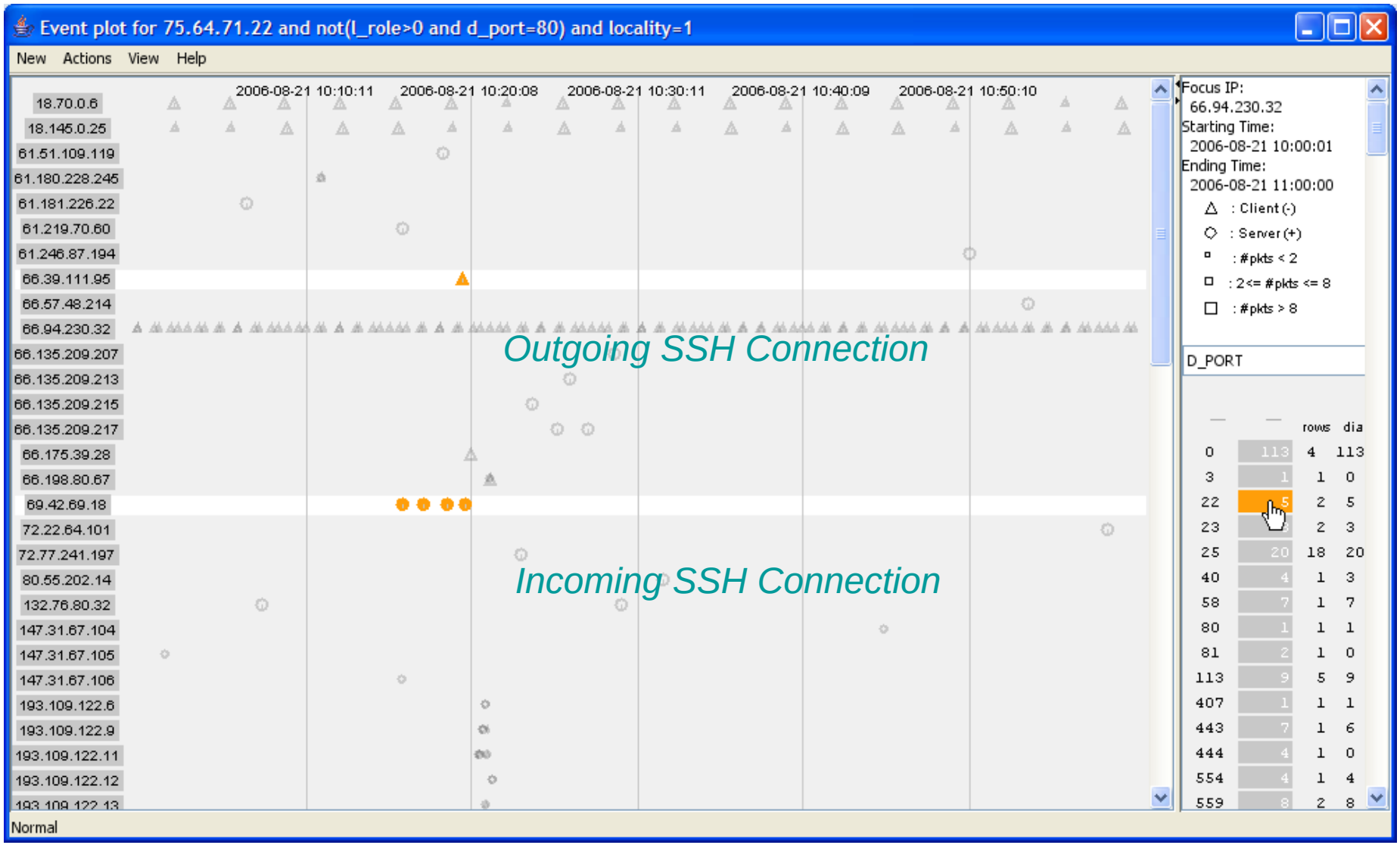
Event Plot



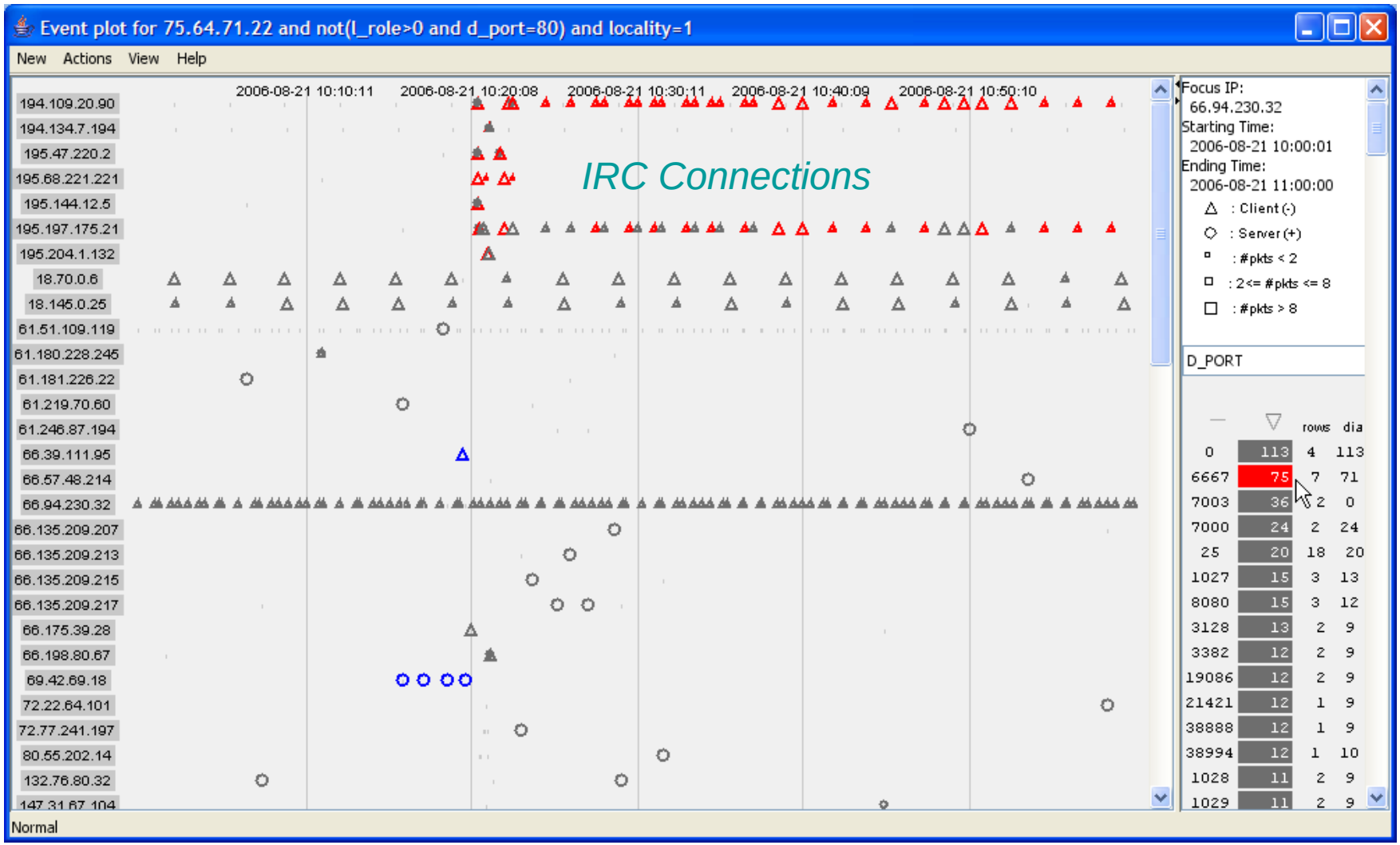
Event Plot


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 10:20:08
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
61.246.87.194
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
72.22.64.101
72.77.241.197
80.55.202.14
132.76.80.32
147.31.67.104
147.31.67.105
147.31.67.106
193.109.122.6
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

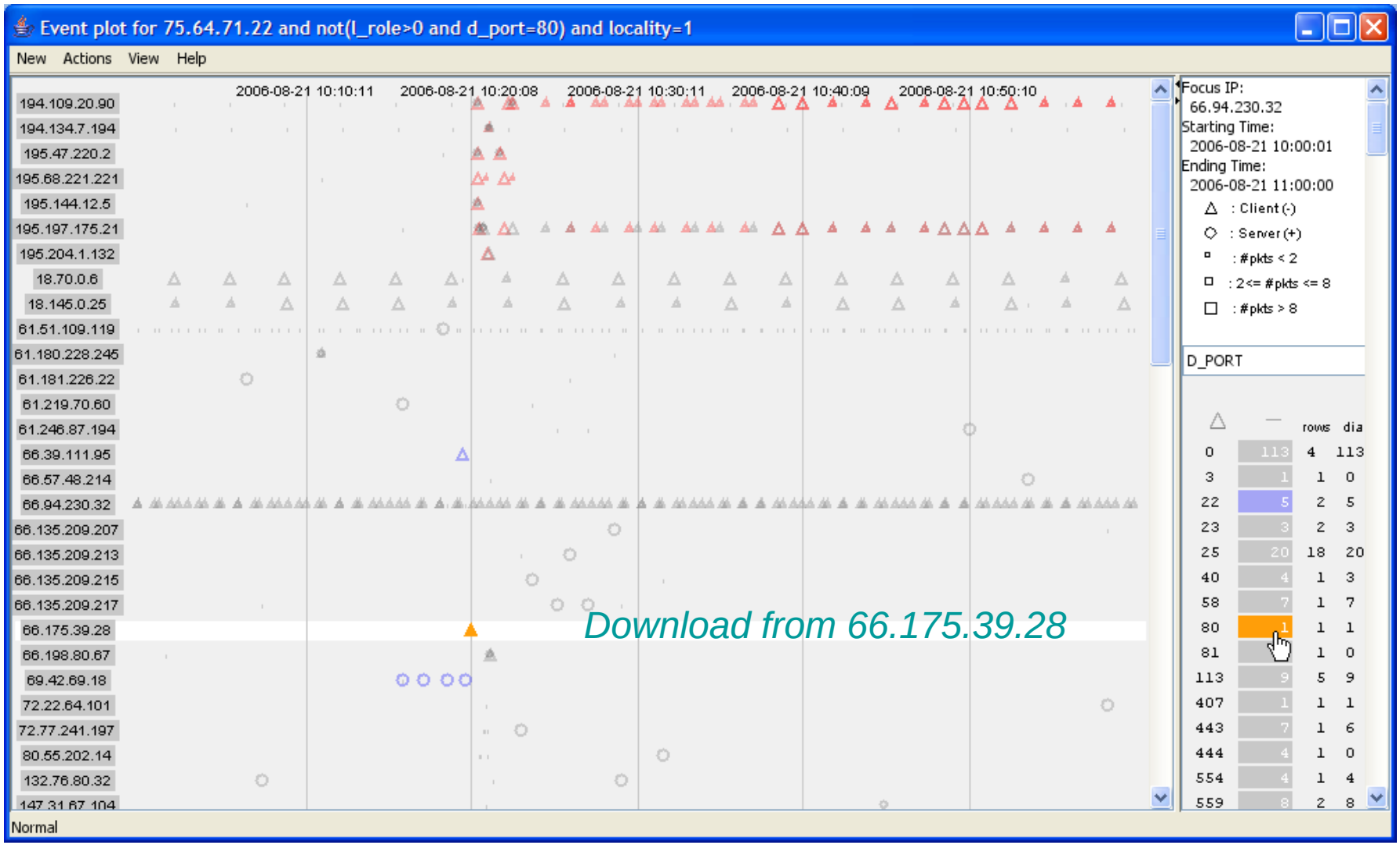
IRC Bot: Initial SSH Connection



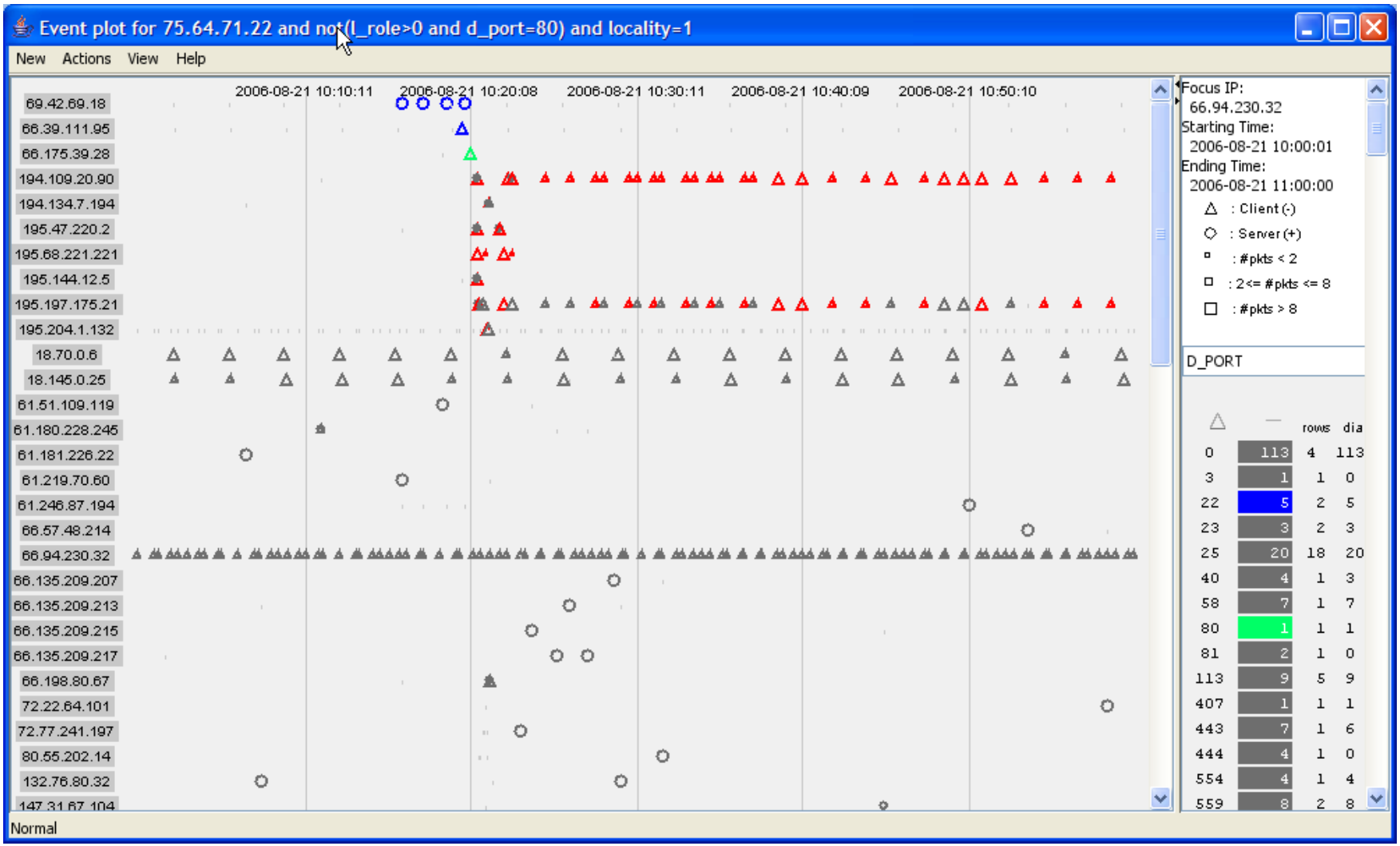
IRC Traffic on port 6667



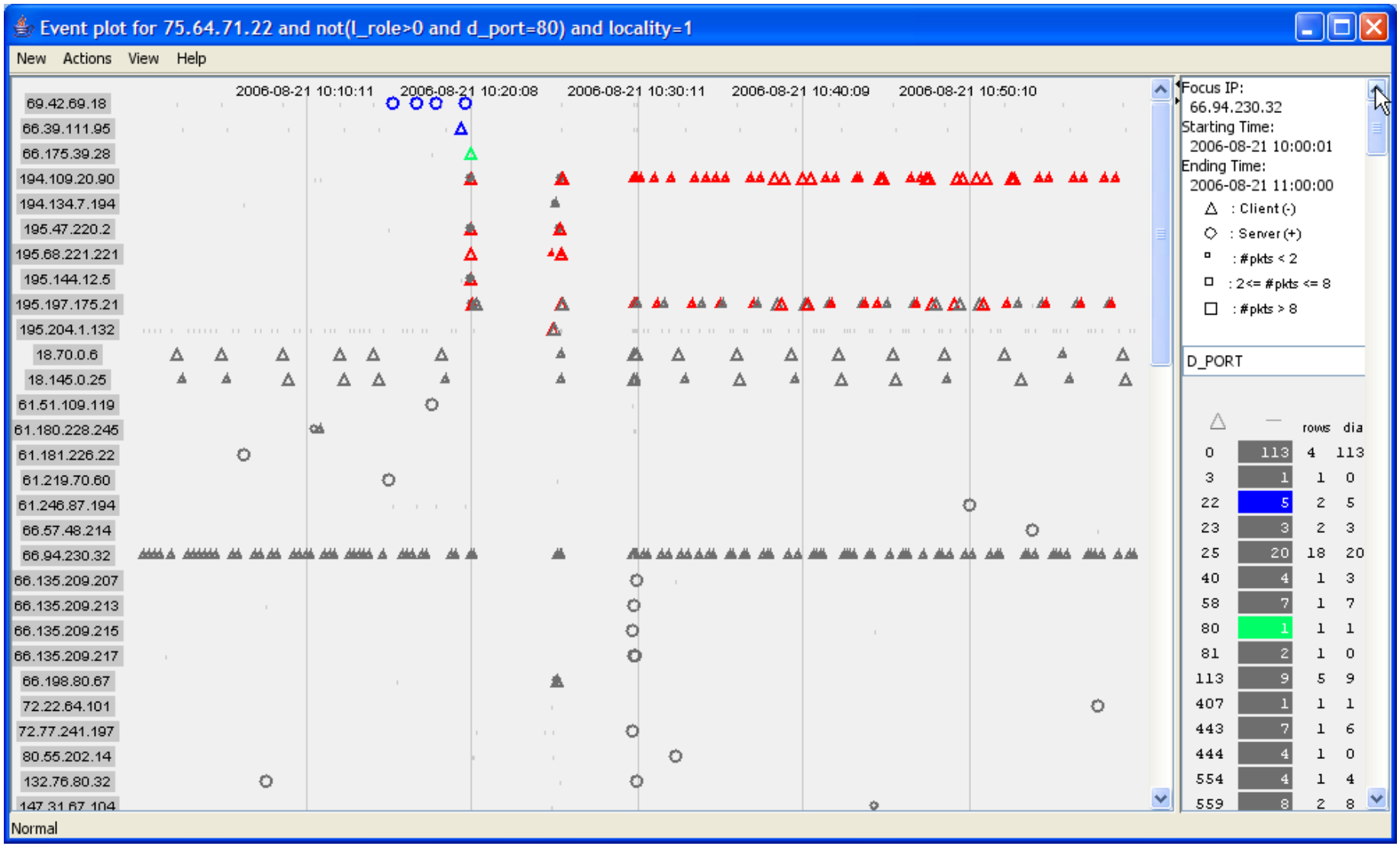
Download of Intrusion Tools



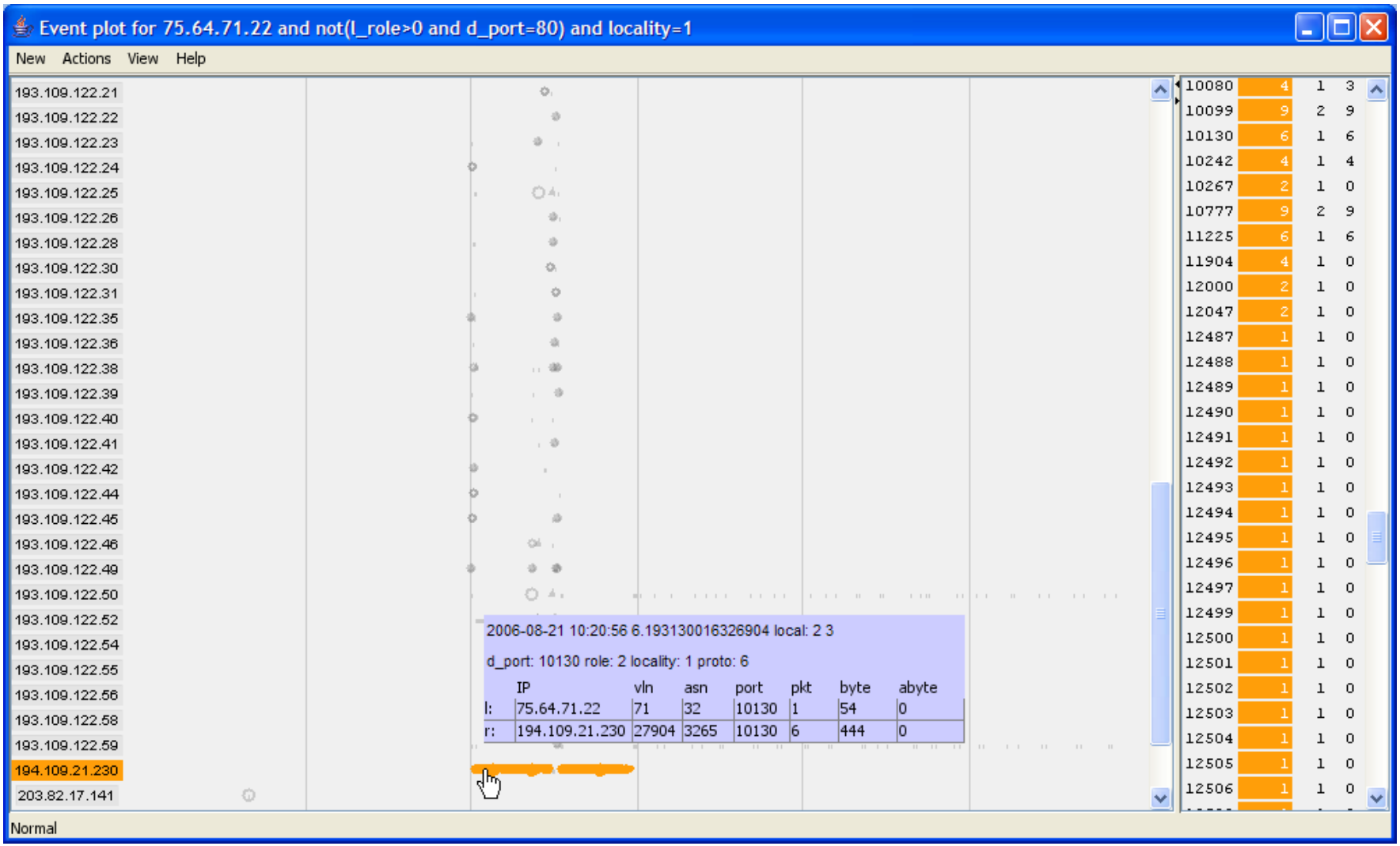
Reordered Rows



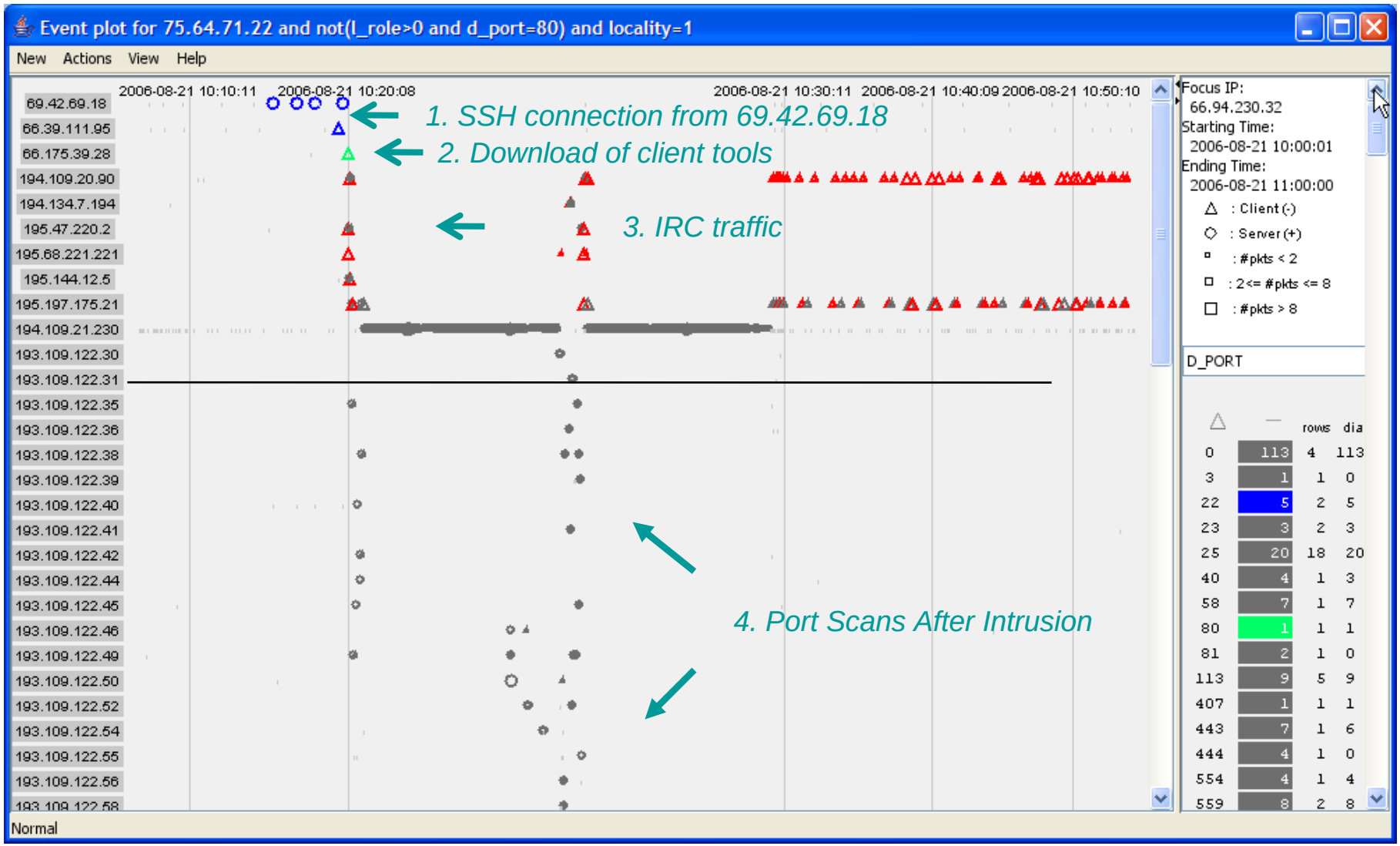
Switch to Ordinal Time



Mine the Gap


Event plot for 75.64.71.22 and not(l_role>0 and d_port=80) and locality=1
New Actions View Help
193.109.122.21
193.109.122.22
193.109.122.23
193.109.122.24
193.109.122.25
193.109.122.26
193.109.122.28
193.109.122.30
193.109.122.31
193.109.122.35
193.109.122.36
193.109.122.38
193.109.122.39
193.109.122.40
193.109.122.41
193.109.122.42
193.109.122.44
193.109.122.45
193.109.122.46
193.109.122.49
193.109.122.50
193.109.122.52
193.109.122.54
193.109.122.55
193.109.122.56
193.109.122.58
193.109.122.59
194.109.21.230
203.82.17.141
2006-08-21 10:20:56 6.193130016326904 local: 2 3
d_port: 10130 role: 2 locality: 1 proto: 6
IP vln asn port pkt byte abyte
l: 75.64.71.22 71 32 10130 1 54 0
r: 194.109.21.230 27904 3265 10130 6 444 0
10080 4 1 3
10099 9 2 9
10130 6 1 6
10242 4 1 4
10267 2 1 0
10777 9 2 9
11225 6 1 6
11904 4 1 0
12000 2 1 0
12047 2 1 0
12487 1 1 0
12488 1 1 0
12489 1 1 0
12490 1 1 0
12491 1 1 0
12492 1 1 0
12493 1 1 0
12494 1 1 0
12495 1 1 0
12496 1 1 0
12497 1 1 0
12499 1 1 0
12500 1 1 0
12501 1 1 0
12502 1 1 0
12503 1 1 0
12504 1 1 0
12505 1 1 0
12506 1 1 0
Normal

Sequence of Intrusion



Future Work

- Scalable query performance
  - Want to query billion row tables at interactive speeds
  - Column-oriented database
  - Distribute across commodity cluster
- Finding network signatures
  - Bottom up capture of analyst domain knowledge
    (see our paper by Xiao in VAST 2006)
  - Top down search for frequent patterns
  - Build disparate flows into behaviors (boot, logon, mail, print, surf, …)
- Modeling Local Machine Behavior
  - Shift the burden to the attacker?