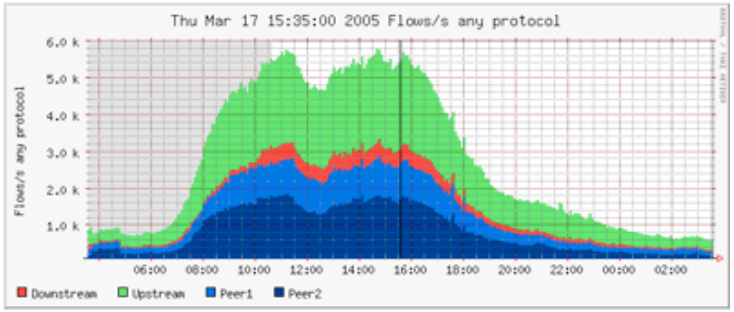
LISA'10 – Nov 7-12, 2010 – San Jose, CA

Nfsight
Netflow-based Network Awareness Tool

Robin Berthier	*(University of Illinois at Urbana-Champaign)*
Michel Cukier	*(University of Maryland)*
Matti Hiltunen	*(AT&T Research)*
Dave Kormann	*(AT&T Research)*
Dan Sheleheda	*(AT&T Security)*
Gregg Vesonder	*(AT&T Research)*

Motivation

- Current flow-based solutions to visualize network traffic:




```
** nfdump -M /netflow0/nfsen-demo/profile-data/live/peer1:peer2:gateway:site:upstream  -T  -r 2007/06/26/14/nfcapd.200706261
nfdump filter:
proto TCP
Aggregated flows 4307432
Top 20 flows ordered by flows:
Date flow start          Duration Proto      Src IP Addr:Port          Dst IP Addr:Port   Packets    Bytes Flows
2007-06-26 14:04:52.233   304.620 TCP       131.152.7.48:25000 ->  125.252.105.225:80        1276    58696   115
2007-06-26 14:04:47.723   299.707 TCP        84.16.67.133:80    ->  147.86.124.128:3136      6743    9.3 M    62
2007-06-26 14:04:47.661   307.782 TCP       194.97.52.210:8000  ->  131.152.112.160:1476     10491    9.5 M    62
2007-06-26 14:04:47.978   299.454 TCP       212.58.227.86:554   ->  131.152.84.130:44368      7385    3.5 M    61
2007-06-26 14:04:48.108   307.212 TCP       131.152.34.73:4374  ->   85.5.58.34:21           9968    1.0 M    61
2007-06-26 14:04:48.108   305.992 TCP       69.247.93.228:18376 ->  147.87.131.32:49474      5305    2.9 M    61
2007-06-26 14:04:58.195   289.820 TCP       85.158.42.174:5000  ->  129.194.97.180:4516        60     5160    60
2007-06-26 14:04:58.671   289.475 TCP       129.194.97.180:4516 ->   85.158.42.174:5000        60     2760    60
2007-06-26 14:04:48.108   305.866 TCP       131.152.164.93:49751 ->  221.9.241.96:38916       3002    3.6 M    60
2007-06-26 14:04:48.170   305.546 TCP        81.230.33.141:36220 ->  147.87.131.32:36827      9476   12.6 M    58
2007-06-26 14:04:47.981   307.337 TCP     195.176.238.195:19996 ->   69.181.19.32:57396       1887    1.7 M    57
2007-06-26 14:04:47.725   299.899 TCP       24.202.245.190:53736 ->  193.222.247.66:50515      5003    2.4 M    56
2007-06-26 13:50:30.576  1157.759 TCP       195.176.162.19:56413 ->   62.2.243.157:443         1029    71512    56
2007-06-26 14:04:48.489   298.942 TCP       131.152.55.83:4894  ->   84.125.80.128:59143       688    32004    56
2007-06-26 14:04:48.109   307.270 TCP       213.39.148.243:20784 ->  131.152.97.66:1755       7607    3.1 M    56
2007-06-26 14:04:47.978   307.468 TCP       193.222.242.13:53849 ->  205.188.215.226:8012     4057   186790    56
2007-06-26 14:05:00.357   291.634 TCP       193.222.244.196:2206  ->  82.64.151.160:6324        937    46208    55
2007-06-26 14:04:48.045   303.499 TCP        66.222.172.199:44999 ->  131.152.159.32:4164      2356    2.4 M    55
2007-06-26 14:04:47.913   304.015 TCP       193.222.243.153:2659  ->  151.203.244.128:6346     1574    76744    54
2007-06-26 14:04:47.850   299.835 TCP        84.16.67.133:80    ->  129.129.158.98:50420      6767    9.3 M    54

Summary: total flows: 6836668, total bytes: 226.6 G, total packets: 269.8 M, avg bps: 932.4 M, avg pps: 142096, avg bpp: 860
Time window: 2007-06-26 13:36:47 - 2007-06-26 14:09:58
Total flows processed: 11582548, skipped: 0, Bytes read: 602310700
Sys: 11.524s flows/second: 1005017.7   Wall: 11.521s flows/second: 1005332.2
```

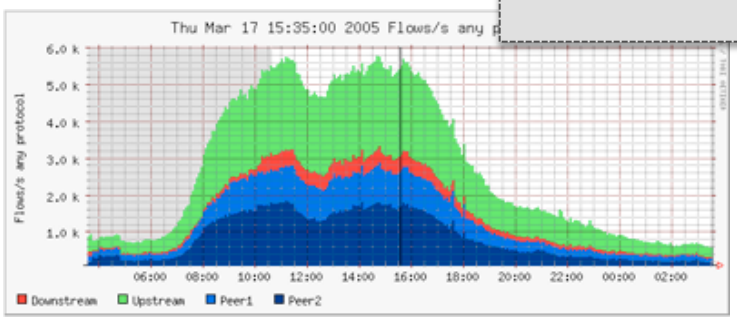
10,000 miles view

Runway

Motivation (cont.)

- Alternative solutions:

Packet-based,
Heavyweight (Java),
Offline, …



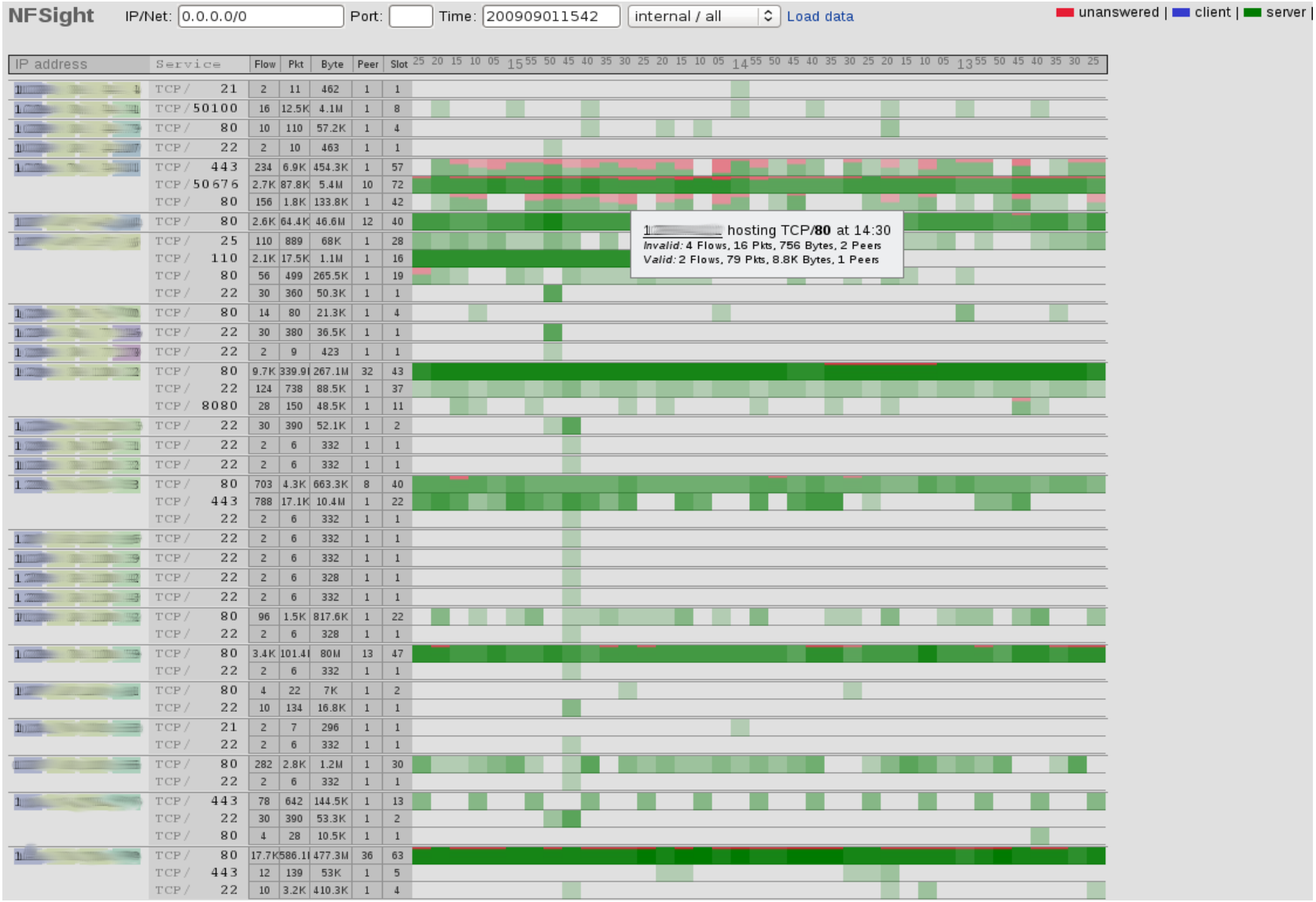
?

```
file-data/live/peer1:peer2:gateway:site:upstream  -T  -r 2007/06/26/14/nfcapd.200706261

Date flow start          Duration Proto       Src IP Addr:Port          Dst IP Addr:Port   Packets    Bytes Flows
2007-06-26 14:04:52.233   304.620 TCP      131.152.7.48:25000 ->  125.252.105.225:80        1276    58696   115
2007-06-26 14:04:47.723   299.707 TCP       84.16.67.133:80    ->  147.86.124.128:3136      6743    9.3 M    62
2007-06-26 14:04:47.661   307.782 TCP      194.97.52.210:8000  ->  131.152.112.160:1476     10491    9.5 M    62
2007-06-26 14:04:47.978   299.454 TCP      212.58.227.86:554   ->  131.152.84.130:44368      7385    3.5 M    61
2007-06-26 14:04:48.108   307.212 TCP      131.152.34.73:4374  ->    85.5.58.34:21          9968    1.0 M    61
2007-06-26 14:04:48.108   305.992 TCP      69.247.93.228:18376 ->   147.87.131.32:49474     5305    2.9 M    61
2007-06-26 14:04:58.195   289.820 TCP      85.158.42.174:5000  ->   129.194.97.180:4516       60     5160    60
2007-06-26 14:04:58.671   289.475 TCP      129.194.97.180:4516 ->    85.158.42.174:5000       60     2760    60
2007-06-26 14:04:48.108   305.866 TCP      131.152.164.93:49751 ->   221.9.241.96:38916     3002    3.6 M    60
2007-06-26 14:04:48.170   305.546 TCP       81.230.33.141:36220 ->   147.87.131.32:36827    9476   12.6 M    58
2007-06-26 14:04:47.981   307.337 TCP    195.176.238.195:19996 ->    69.181.19.32:57396     1887    1.7 M    57
2007-06-26 14:04:47.725   299.899 TCP      24.202.245.190:53736 ->   193.222.247.66:50515    5003    2.4 M    56
2007-06-26 13:50:30.576  1157.759 TCP      195.176.162.19:56413 ->    62.2.243.157:443       1029   71512    56
2007-06-26 14:04:48.489   298.942 TCP      131.152.55.83:4894  ->    84.125.80.128:59143     688   32004    56
2007-06-26 14:04:48.109   307.270 TCP      213.39.148.243:20784 ->   131.152.97.66:1755      7607    3.1 M    56
2007-06-26 14:04:47.978   307.468 TCP      193.222.242.13:53849 ->  205.188.215.226:8012     4057  186790    56
2007-06-26 14:05:00.357   291.634 TCP      193.222.244.196:2206  ->   82.64.151.160:6324       937   46208    55
2007-06-26 14:04:48.045   303.499 TCP       66.222.172.199:44999 ->   131.152.159.32:4164     2356    2.4 M    55
2007-06-26 14:04:47.913   304.015 TCP      193.222.243.153:2659  ->  151.203.244.128:6346     1574   76744    54
2007-06-26 14:04:47.850   299.835 TCP       84.16.67.133:80    ->   129.129.158.98:50420     6767    9.3 M    54

Summary: total flows: 6836668, total bytes: 226.6 G, total packets: 269.8 M, avg bps: 932.4 M, avg pps: 142096, avg bpp: 860
Time window: 2007-06-26 13:36:47 - 2007-06-26 14:09:58
Total flows processed: 11582548, skipped: 0, Bytes read: 602310700
Sys: 11.524s flows/second: 1005017.7   Wall: 11.521s flows/second: 1005332.2
```

10,000 miles view

Runway

Approach


NFSight
IP/Net: 0.0.0.0/0
Port:
Time: 200909011542
internal / all
Load data
unanswered | client | server |
IP address
Service
Flow
Pkt
Byte
Peer
Slot
hosting TCP/80 at 14:30
Invalid: 4 Flows, 16 Pkts, 756 Bytes, 2 Peers
Valid: 2 Flows, 79 Pkts, 8.8K Bytes, 1 Peers

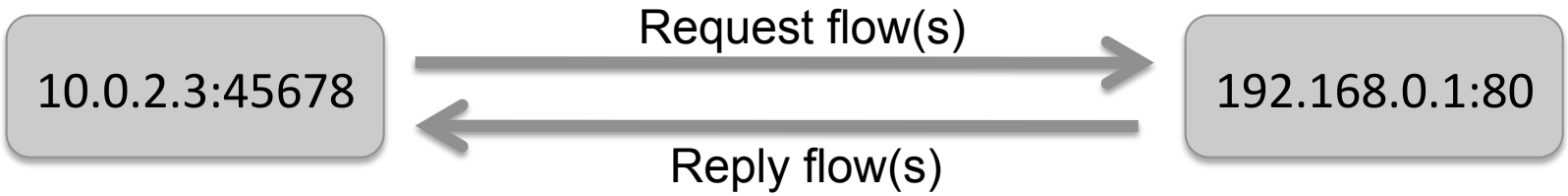
Approach (cont.)

- Flow-based (scalability, privacy, already instrumented)
- **Aggregate** network activity per host and port
- Identify **client/server** behavior
- Visualize at large scale, drill-down on demand
- Detect rogue services and intrusion attempts
- Lightweight and easy to deploy and to use

Challenge

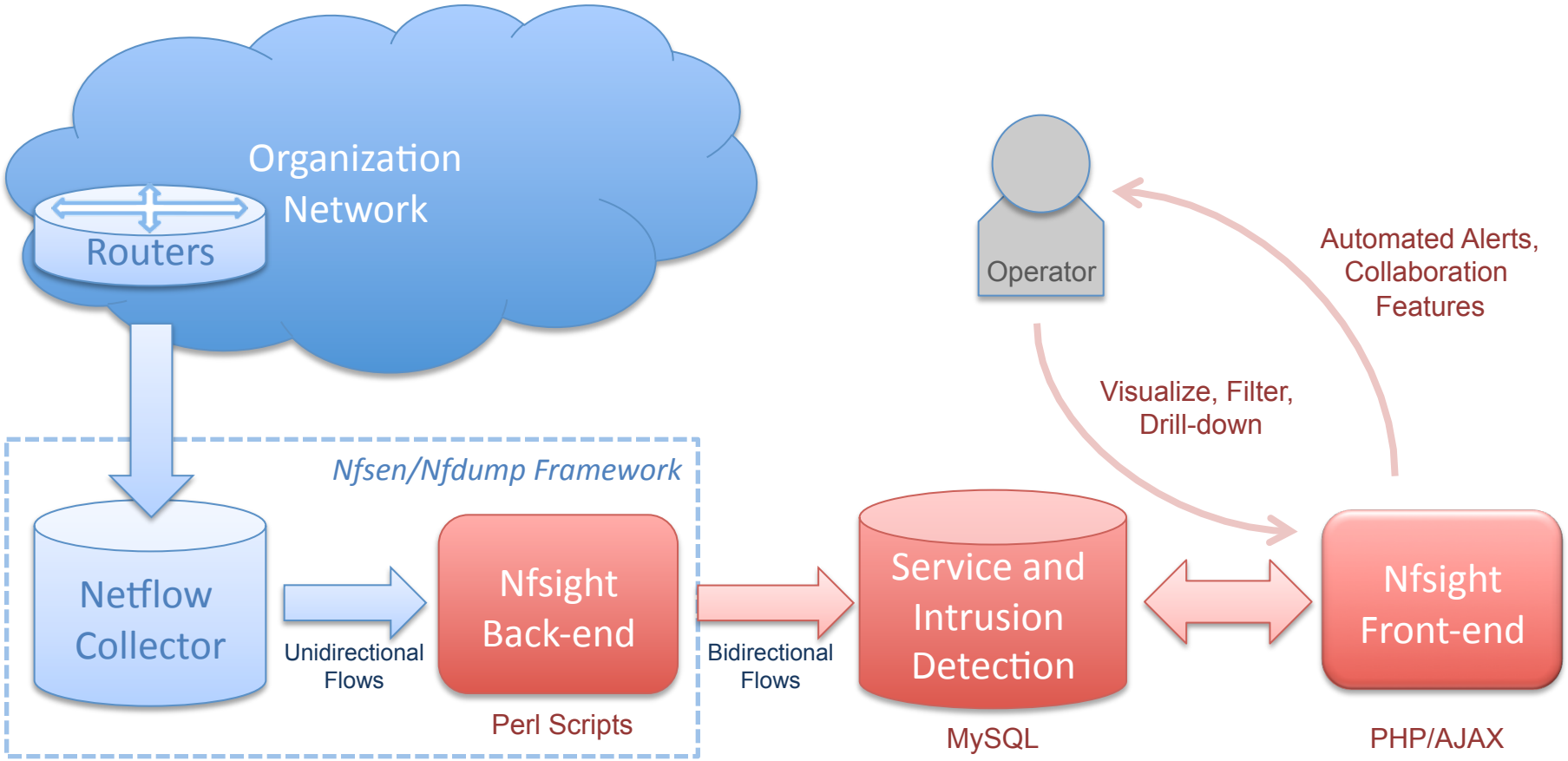
- Netflow is unidirectional

  (waiting for IPFIX…)



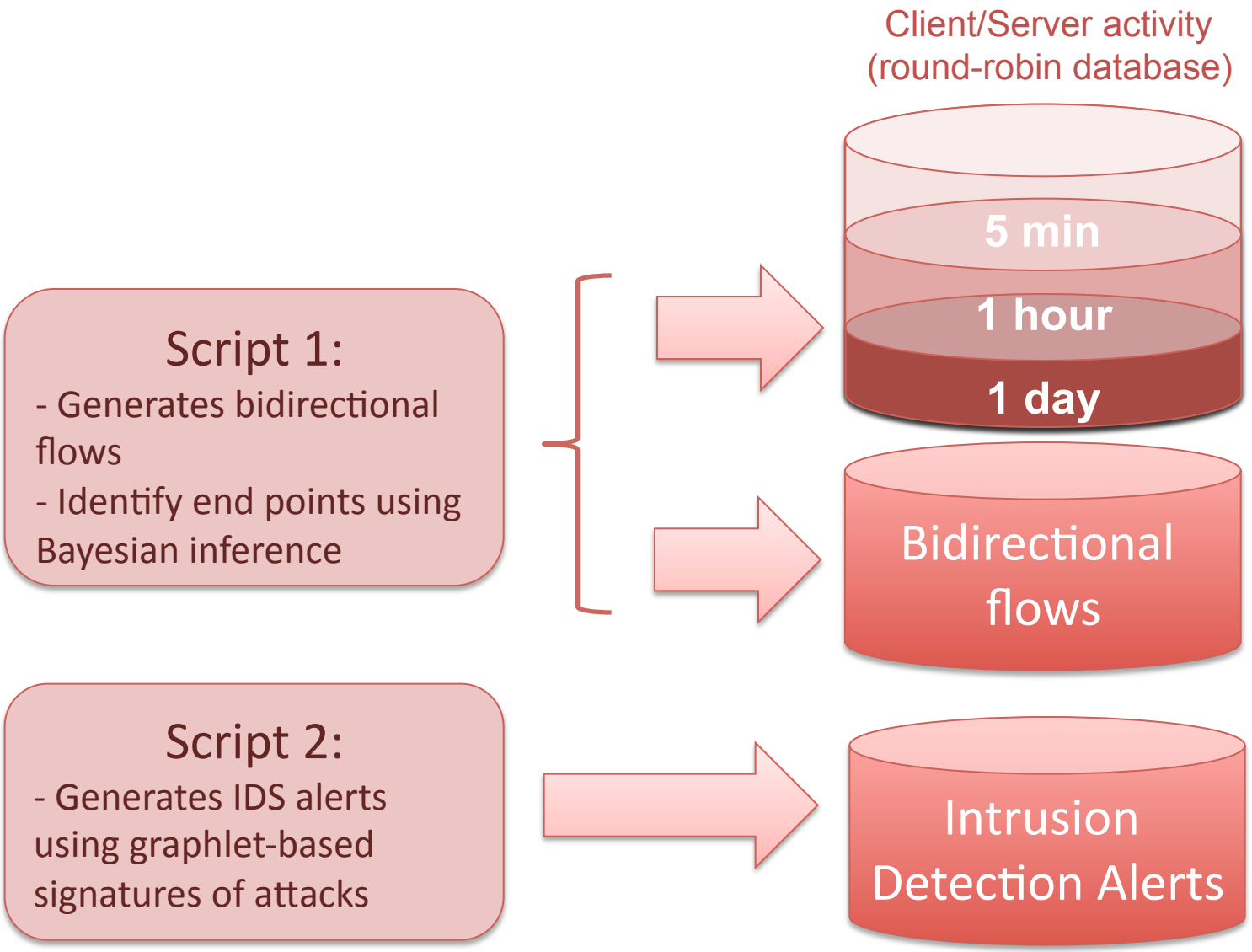
→ Requires **heuristic** to identify client/server

Architecture

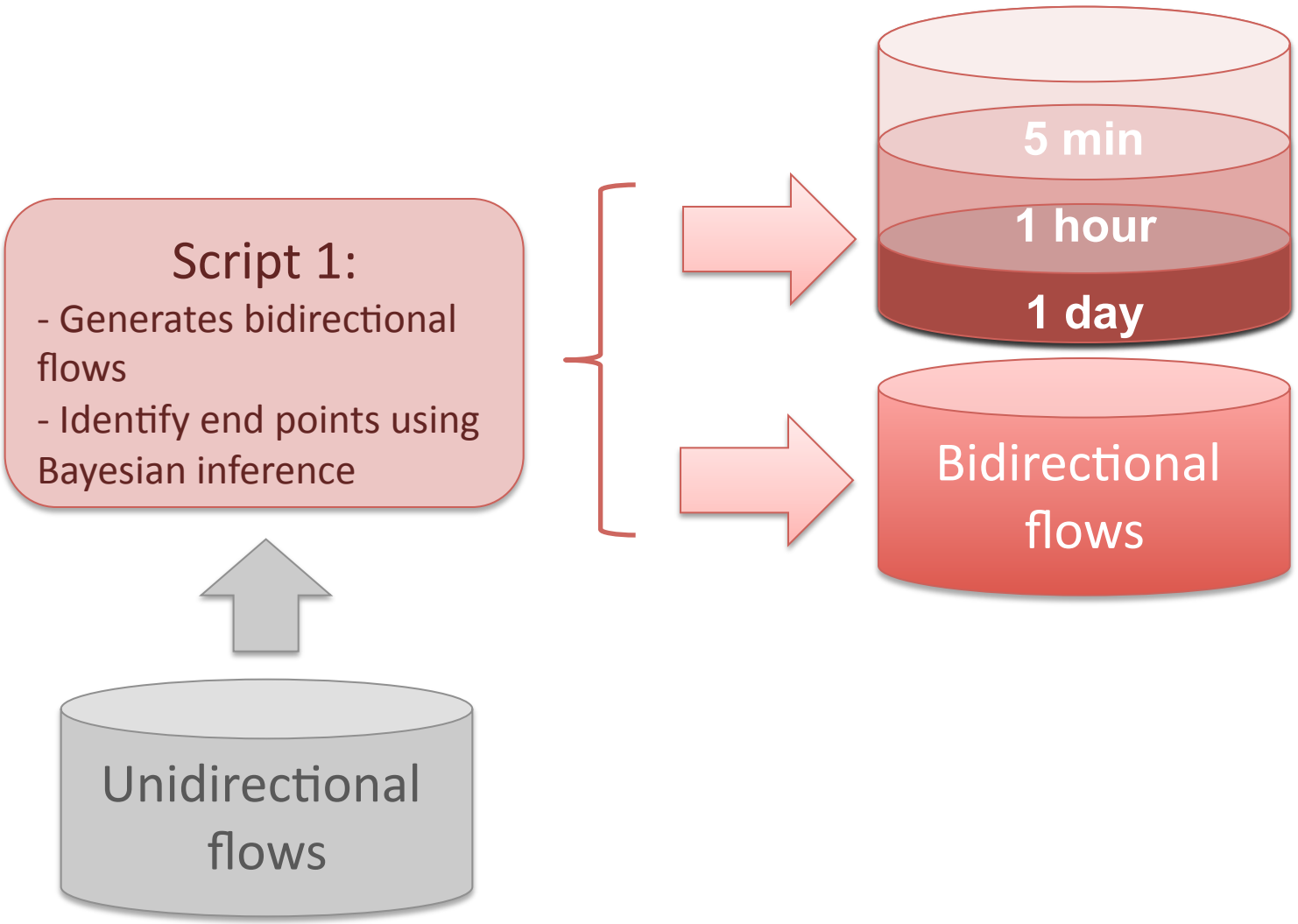


Back-end

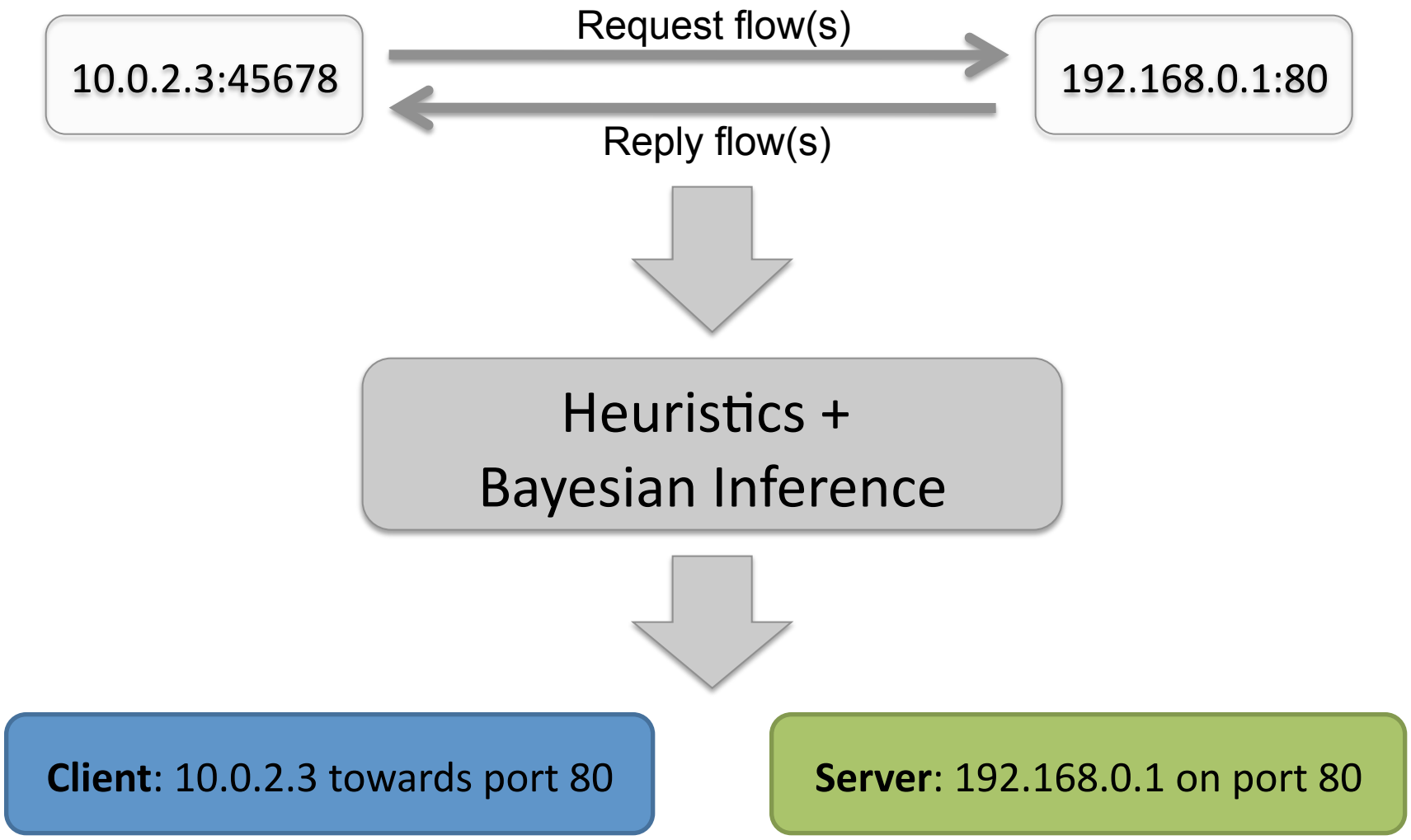
Back-end Scripts



Identifying Client/Server



Generating Bidirectional Flows



Heuristics

Heuristic ID	*Features and Formula Used*	*Output Values*
Timing:		
Heuristic 0	Timestamp of request < Timestamp of reply	[0, …]
Port numbers:		
Heuristic 1	Src port > Dst port	{0, 0.5, 1}
Heuristic 2	Src port > 1024 > Dst port	{0, 0.5, 1}
Heuristic 3	Port in /etc/services	{0, 0.5, 1}
Fan in/out relationships:		
Heuristic 4	# ports related	[0, …]
Heuristic 5	# IP related	[0, …]
Heuristic 6	# Tuples related	[0, …]

Bayesian Inference

Initial Prior:

SRC is server	*DST is server*
0.5	0.5

Measurement 1: DST is server

Heuristic 0	1
Heuristic 1	0.5
Heuristic 2	0.5
Heuristic 3	0.5
Heuristic 4	0.5
Heuristic 5	0.5

Posterior 1

SRC is server	*DST is server*
30.00%	**70.00%**

Measurement 2: DST is server

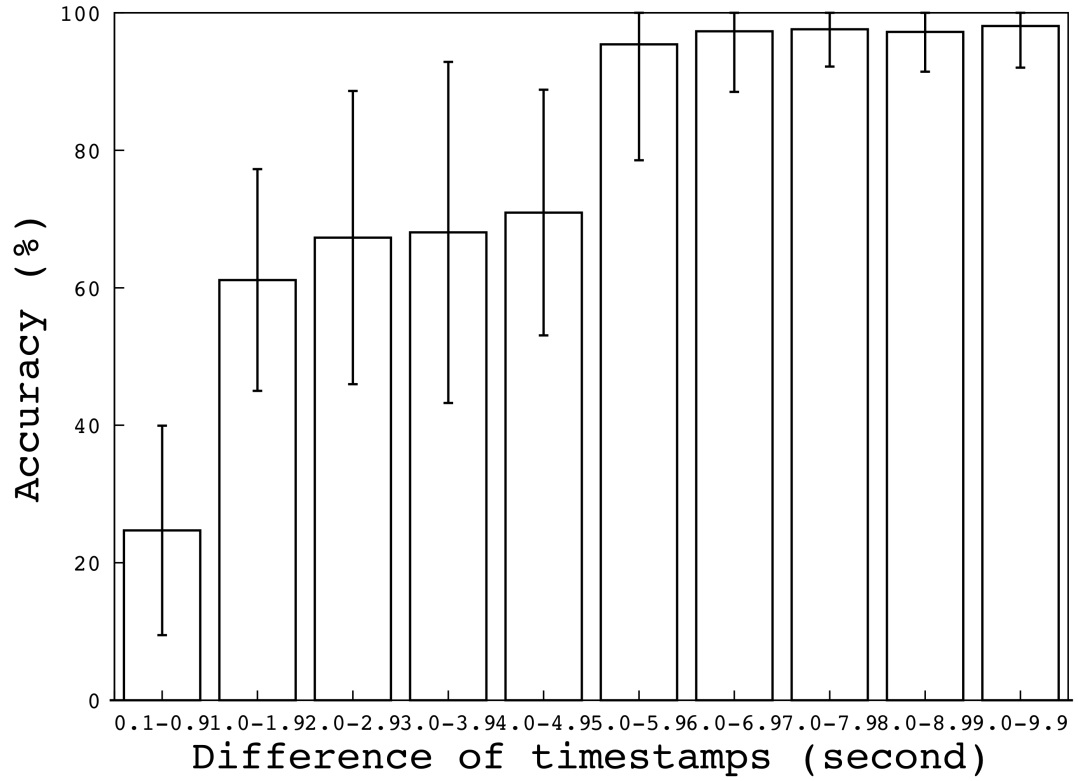
Heuristic 0	0
Heuristic 1	0.5
Heuristic 2	0.5
Heuristic 3	0.8
Heuristic 4	0.7
Heuristic 5	0.5

Posterior 2

SRC is server	*DST is server*
15.33%	**84.67%**

Heuristic Evaluation

Distribution accuracy of heuristic 0:
(Timestamp of request < Timestamp of reply)

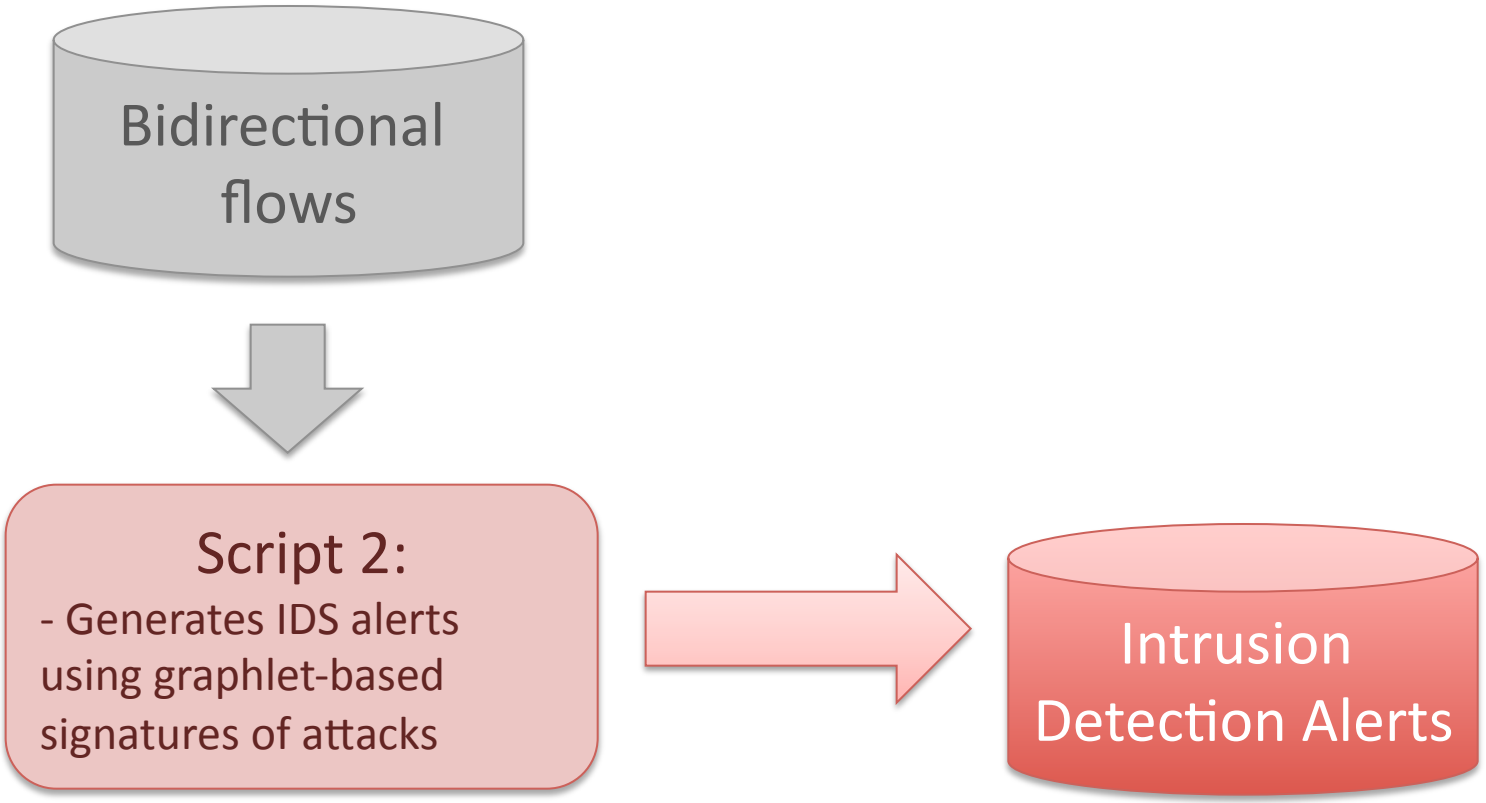


Heuristic Evaluation (cont.)

- Comparison against Argus (packet-based bidirectional flow generator)
- 30min dataset collected at the Univ. of Maryland (40,000+ nodes, 3.6 million of bidirectional flows)
- Results:

Heuristic	*Able to decide*	*Accuracy*
H.0	11.49%	94.54%
H.1	63,98%	85.54%
H.2	48.14%	98.15%
H.3	47.73%	98.17%
H.4	63.28%	93.72%
H.5	55.51%	88.76%
H.6	63.38%	92.58%

Security Monitoring



Security Monitoring (cont.)

Graphlet structure [1] (for both source and destination):

- IP address
  - Protocol
    - Peer: *the set of distinct related IP addresses*
    - Port: *the set of distinct related destination or source ports*
    - TCP flag: *the set of distinct flag combinations used*
    - Packet: *the total number of packets sent or received*
    - Byte: *the total number of bytes sent or received*
    - Flows: *the total number of bidirectional flows sent or received*
    - Failed connections: *the total number of unidirectional flows sent or received*
    - Last source end point: *the source port, IP address and TCP flag of the last flow captured*
    - Last destination end point: *the destination port, IP address and TCP flag of the last flow captured*

[1] KARAGIANNIS, T., PAPAGIANNAKI, K., AND FALOUTSOS, M. BLINC: multilevel traffic classification in the dark. In *Proc. ACM SIGCOMM Conference (2005), pp. 229–240.*

Security Monitoring (cont.)

Current set of signatures:

Id	*Name*	*Category*	*Filter*
101	Identical source and destination	Malformed flow	src_ip = dst_ip
102	Invalid ICMP flow size	Malformed flow	proto = ICMP *and* total_byte ≤ 64000
104	Invalid ICMP code	Malformed flow	proto = ICMP *and* icmp_code ∈ invalid_code
105	Invalid IP address	Malformed flow	(src_ip *or* dst_ip) ∈ invalid_ip
106	Invalid TCP flag	Malformed flow	proto = TCP *and* flag ∈ invalid_flag
201	One-to-many IP	One-to-many	failed_connection ≥ 1 *and* unique_dst_ip ≥ max_dst_ip *and* unique_flag ≤ 1
301	One-to-many Port	One-to-many	failed_connection ≥ 1 *and* unique_dst_port ≥ max_dst_port *and* unique_flag ≤ 1
401	Many-to-one IP on TCP flows	Many-to-one	proto = TCP *and* flag ∉ {19, 27, 30, 31} *and* unique_src_ip ≥ max_src_ip *and* unique_flag ≤ 1
402	Many-to-one IP on ICMP flows	Many-to-one	proto = ICMP *and* unique_src_ip ≥ max_src_ip
403	Many-to-one IP on UDP flows	Many-to-one	proto = UDP *and* unique_src_ip ≥ max_src_ip
501	Many-to-one Port on TCP flows	Many-to-one	proto = TCP *and* flag ∉ {19, 27, 30, 31} *and* unique_src_port ≥ max_src_port *and* unique_dst_port = 1 *and* unique_flag = 1
502	Many-to-one Port on ICMP flows	Many-to-one	proto = ICMP *and* unique_src_port ≥ max_src_port *and* unique_dst_port = 1
503	Many-to-one Port on UDP flows	Many-to-one	proto = UDP *and* unique_src_port ≥ max_src_port *and* unique_dst_port = 1

Intrusion Detection Evaluation

- Example of e-mail validation:

```
192.168.1.2 [One-to-many IP]

IP contacting more than 200 distinct targets in less than 5min
* Heuristic: 201
* First detected on: 2010-08-10 14:05:00
* Last detected on: 2010-08-10 16:55:00
* Number of occurrences: 52,908
* Total flows: 52,908
* Unanswered flow requests: 52,908 (100\%)
* Packets: 89,918 * Bytes: 4,316,160
* Average number of related host every 5min: 4,580
* Average number of related port every 5min: 2
* Last source port: 3317 (2,339 distinct port(s) used every 5min)
* Last related tuple: 192.168.26.198 TCP/445
* Last flag value (if TCP): 2

To visualize related Nfsight data:
https://nfsight/index.php?net=192.168.1.2&time=201008101655
--------------------------------
Please rate this alert by clicking on one of the following links:
[+] True Positive:
https://nfsight/email_validation.php?q=156505&r=1&auth=r25kfGVk
[-] False Positive:
https://nfsight/email_validation.php?q=156505&r=-1&auth=r25kfGVk
[?] Inconclusive:
https://nfsight/email_validation.php?q=156505&r=0&auth=r25kfGVk
--------------------------------
```

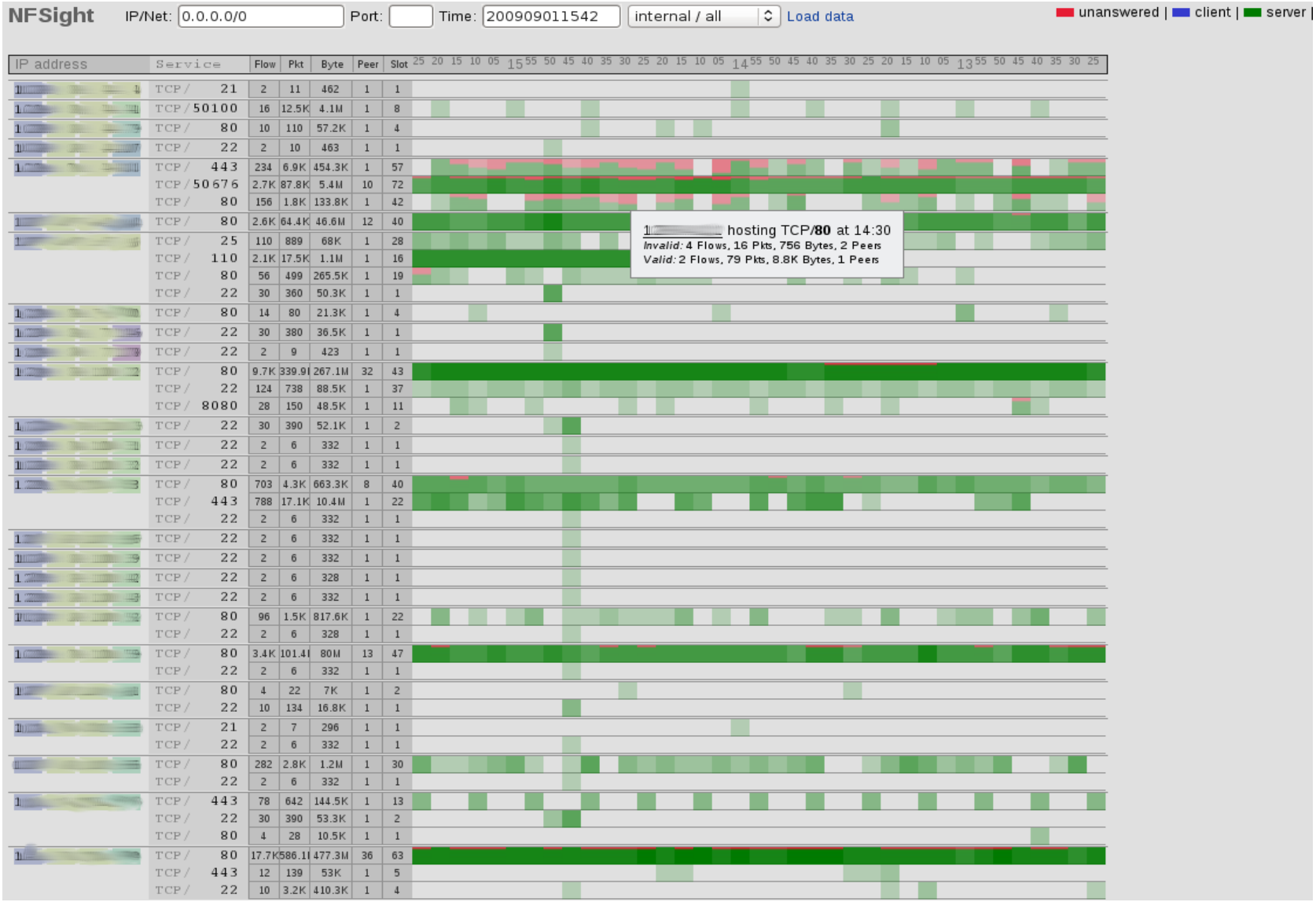
Intrusion Detection Evaluation (cont.)

- Results after 4 months of deployment:

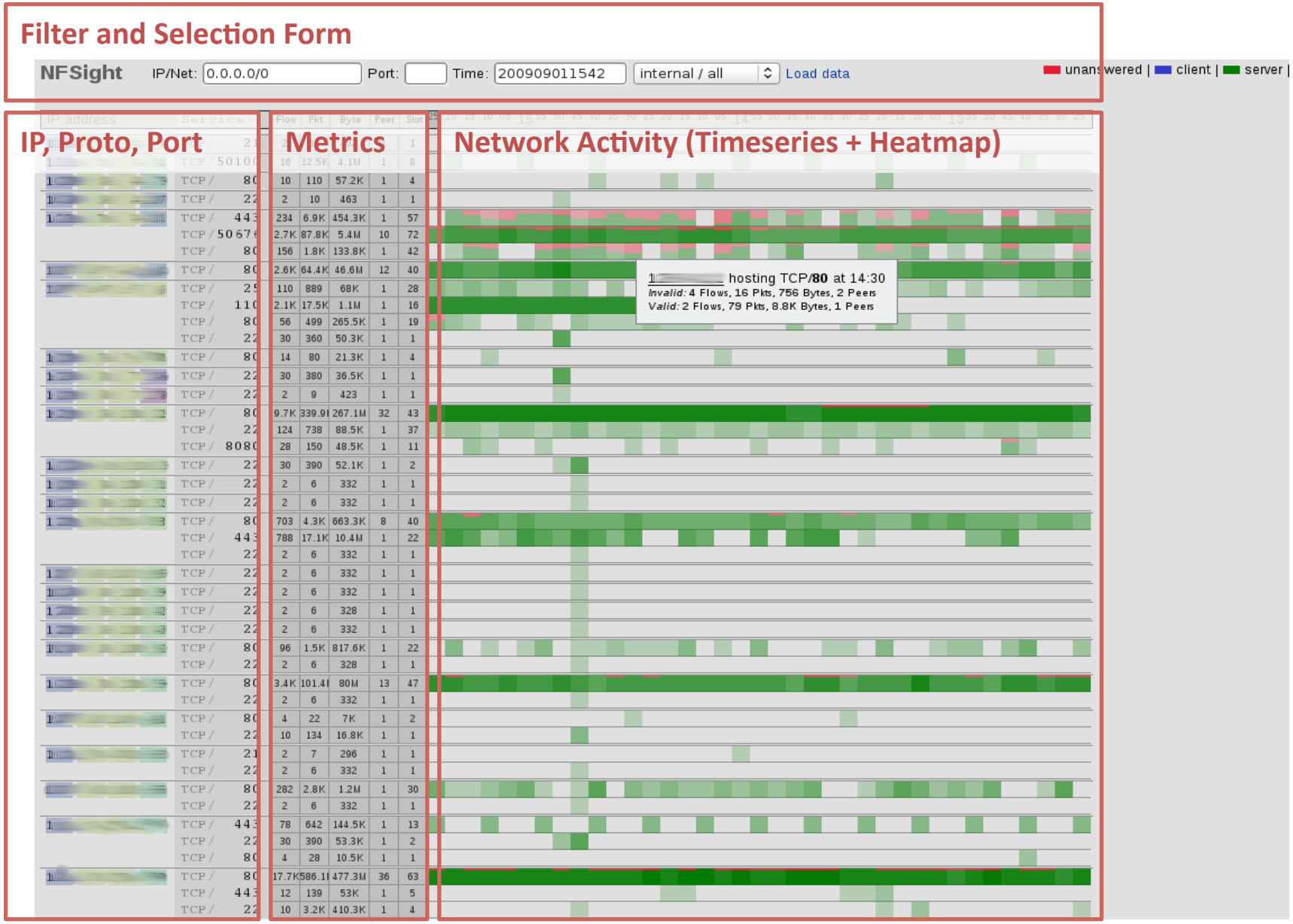
Id	*Total Validated*	*TP*	*FP*	*Unknown*	*Accuracy: TP/(FP+TP)*
105	23	11	4	8	73.3%
106	27	3	19	5	13.6%
201	68	40	21	7	65.6%
301	94	30	41	23	42.3%
501	78	21	38	19	35.6%

Front-end

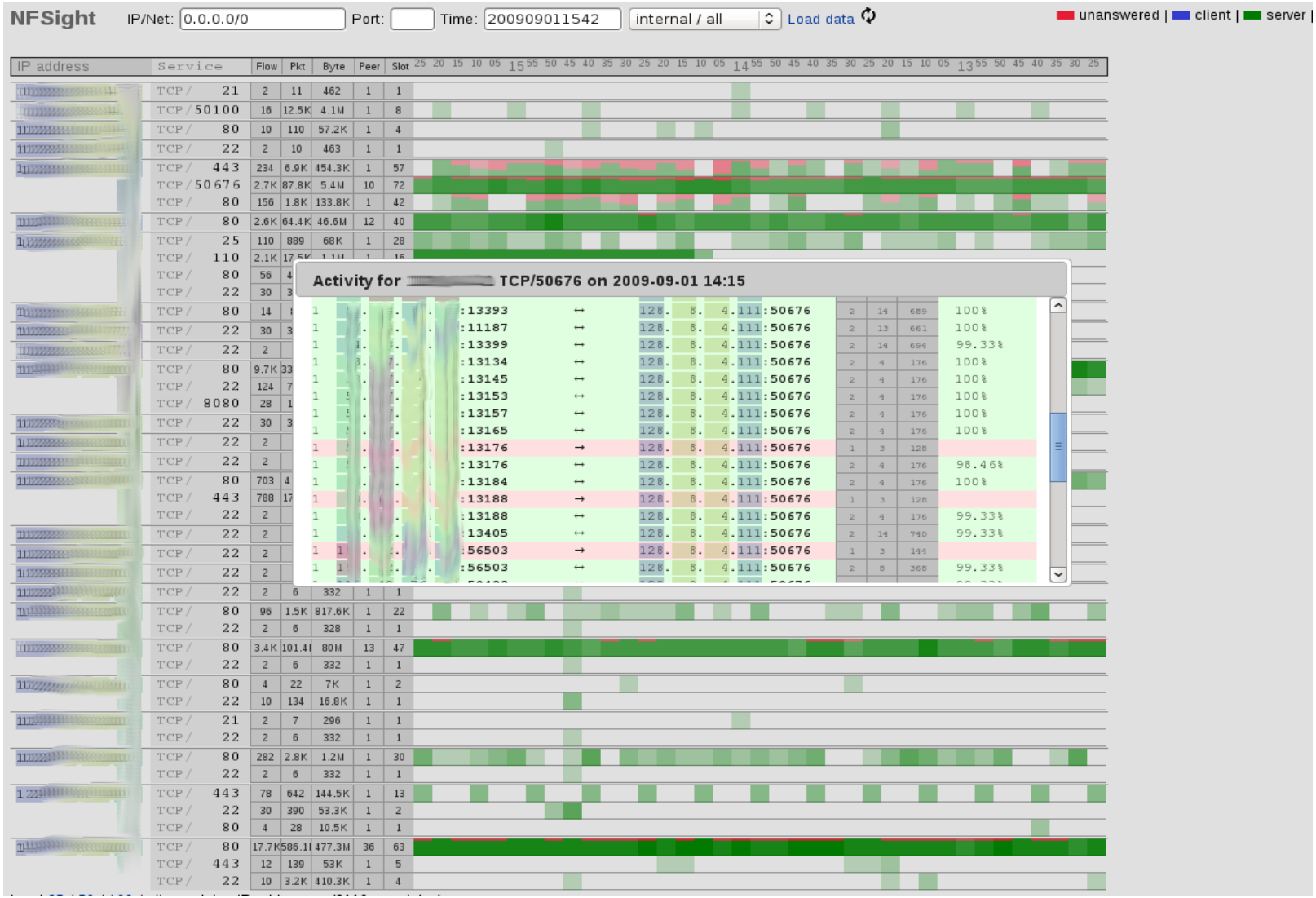
Front-end


NFSight
IP/Net: 0.0.0.0/0
Port:
Time: 200909011542
internal / all
Load data
unanswered | client | server |
IP address
Service
Flow
Pkt
Byte
Peer
Slot
hosting TCP/80 at 14:30
Invalid: 4 Flows, 16 Pkts, 756 Bytes, 2 Peers
Valid: 2 Flows, 79 Pkts, 8.8K Bytes, 1 Peers

Front-end (cont.)

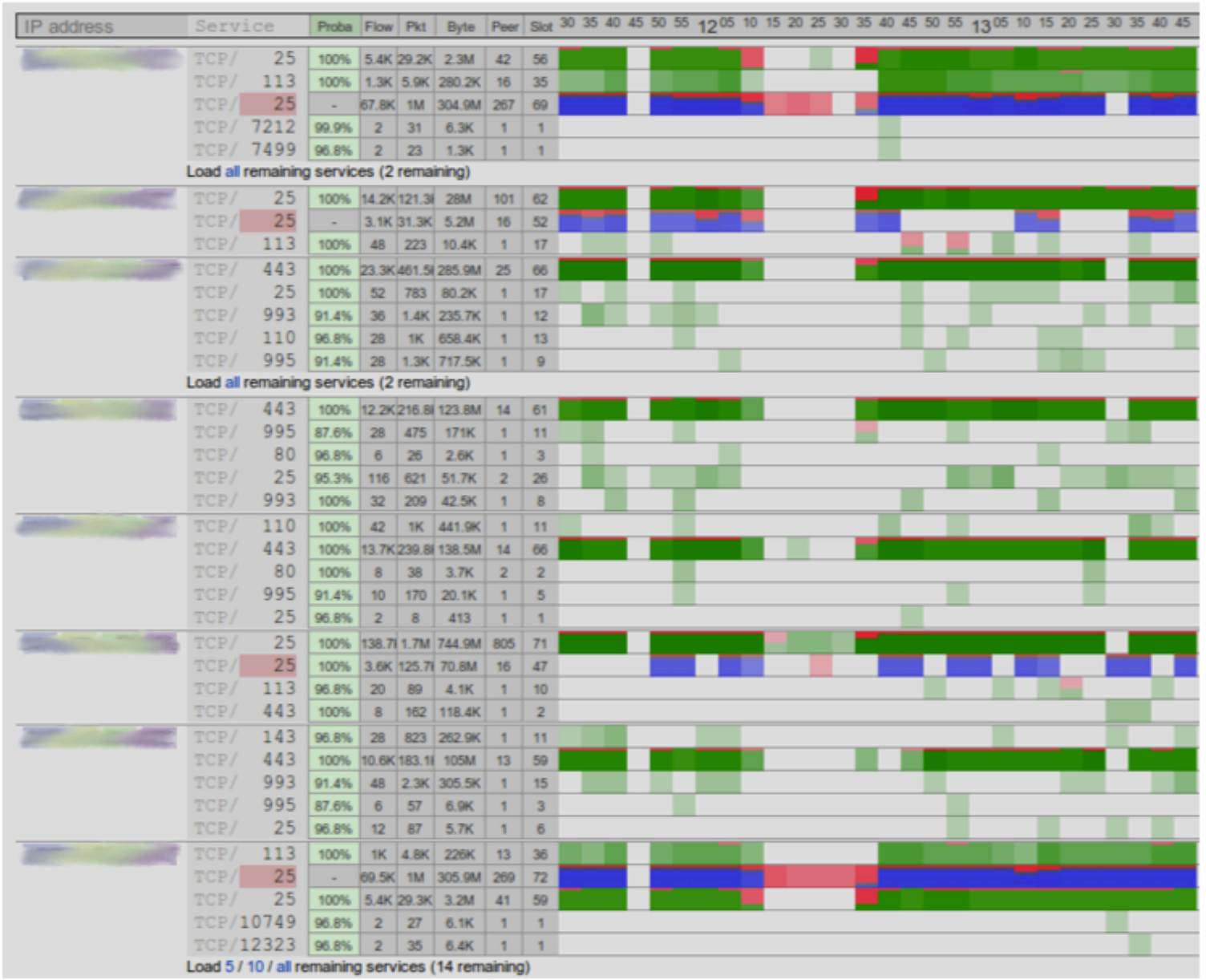


Details on Demand

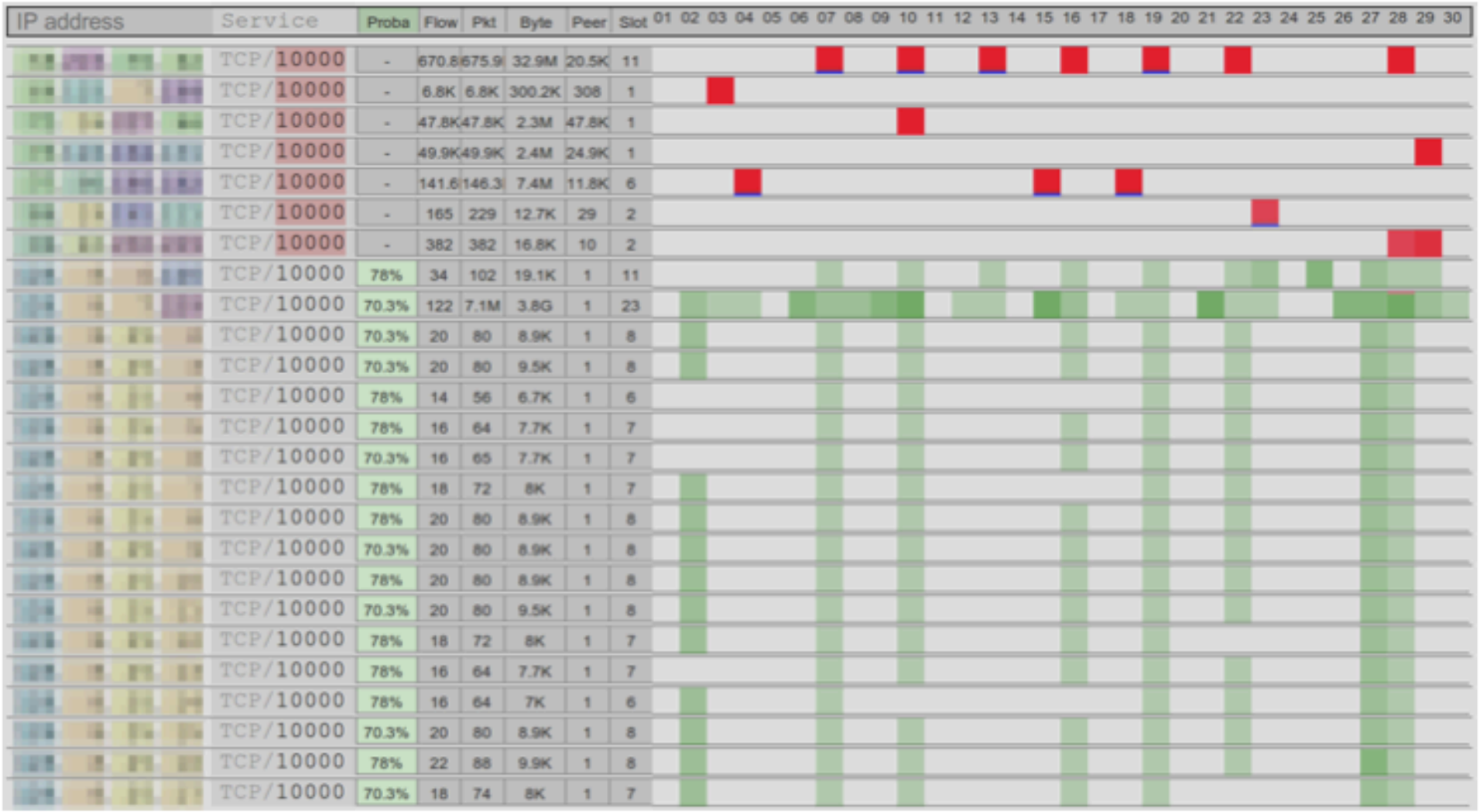


Demo / Use Cases

Use Cases Example: Effect of Power Outage



Use Cases Example: Scanning Activity



Use Cases Example: Top 20 Scanned Services

service	sources	Flow	Pkts	byte	Peer/5min
TCP/25	2617	75.6K	159.6K	8.3M	8
TCP/5900	706	213.5K	400.1K	20.3M	12
TCP/80	172	811.2K	1.1M	49.8M	652
TCP/443	113	62.6K	111.4K	5.4M	72
TCP/23	70	16.6K	44.3K	2.1M	24
TCP/22	46	801.9K	1.1M	62.9M	1.6K
TCP/3072	38	5.8K	7.6K	356.1K	39
TCP/1024	38	5.7K	7.4K	346.5K	38
TCP/1433	37	1.4M	1.5M	61.3M	15K
TCP/3389	32	307.3K	332K	16.7M	3K
TCP/8080	18	211.5K	325.4K	14.7M	372
TCP/9415	18	429.1K	498.7K	21M	2.8K
TCP/3128	17	244.2K	396K	17M	397
TCP/465	15	412	609	27.7K	6
TCP/1080	12	591.8K	603K	24.4M	3.7K
TCP/8296	9	487	573	23.1K	8
TCP/38981	9	926	1.1K	44K	11
TCP/53329	9	513	592	23.9K	9
TCP/63580	9	450	536	21.6K	8
TCP/8000	9	1M	1M	41.9M	10.8K

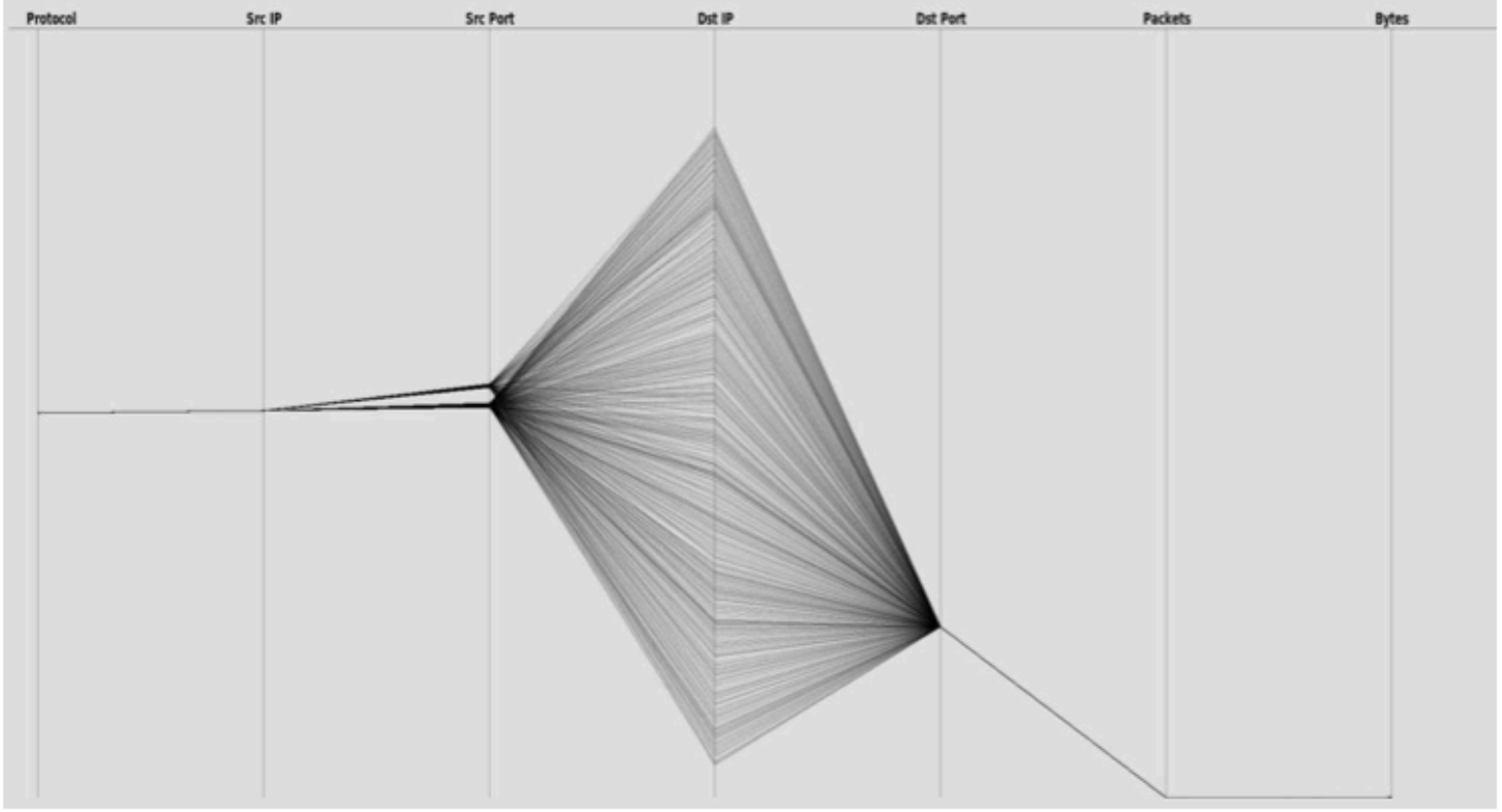
Use Cases Example: Worm Outbreak

Service activity detected between 23:35 and 02:40 of Jul 31st:

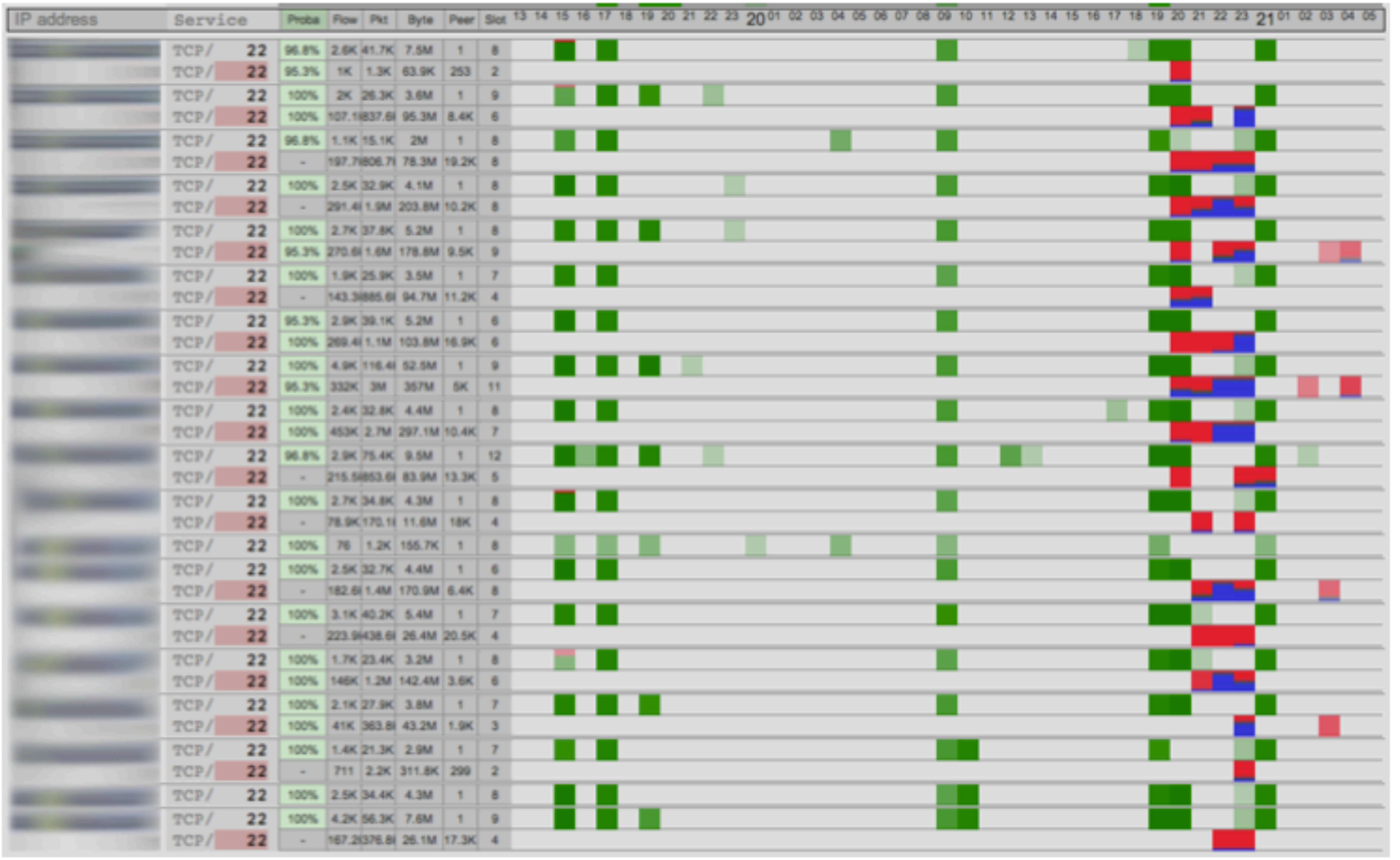
IP address	Service	Proba	Flow	Pkt	Byte	Peer	Slot
	TCP/ 445	-	768.8l	1.5M	73.8M	23.3K	33
	TCP/ 1406	99.3%	2	10	764	1	1
	TCP/ 3222	99.3%	2	10	764	1	1

35 40 45 50 55 00 05 10 15 20 25 30 35 40 45 50 55 01

Unidirectional flow collected at 02:40:



Use Cases Example: Distributed Attacks



Future Work

- Improve attack signatures
- Strengthen Bayesian inference to work in different conditions:
  - Sampling
  - Asymmetric routing
- Create heuristics to detect type of service
- Output bidirectional flows using IPFIX (RFC5103)

http://nfsight.research.att.com

Contact: rgb@illinois.edu