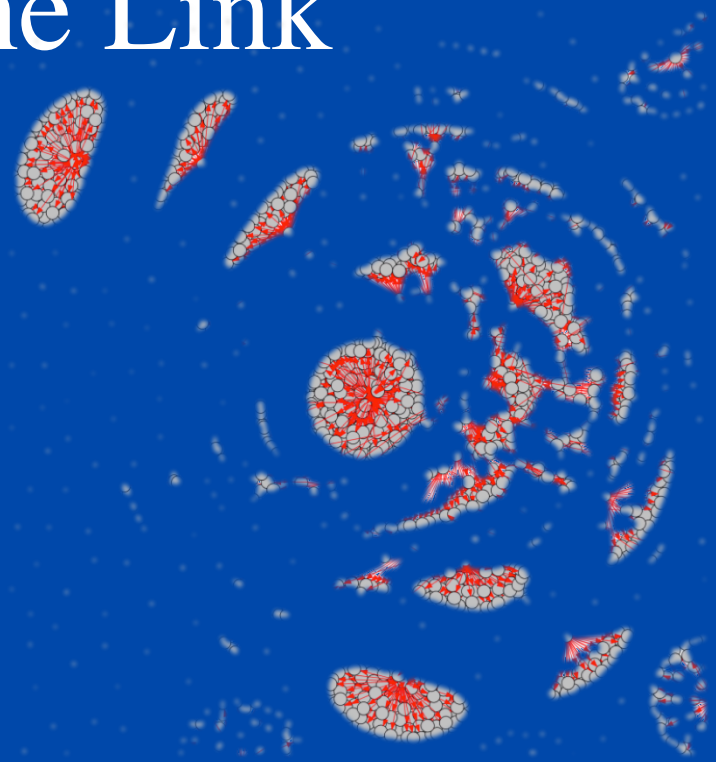
CHALMERS

Tracking Malicious Hosts on a 10Gbps Backbone Link

Magnus Almgren*
Wolfgang John



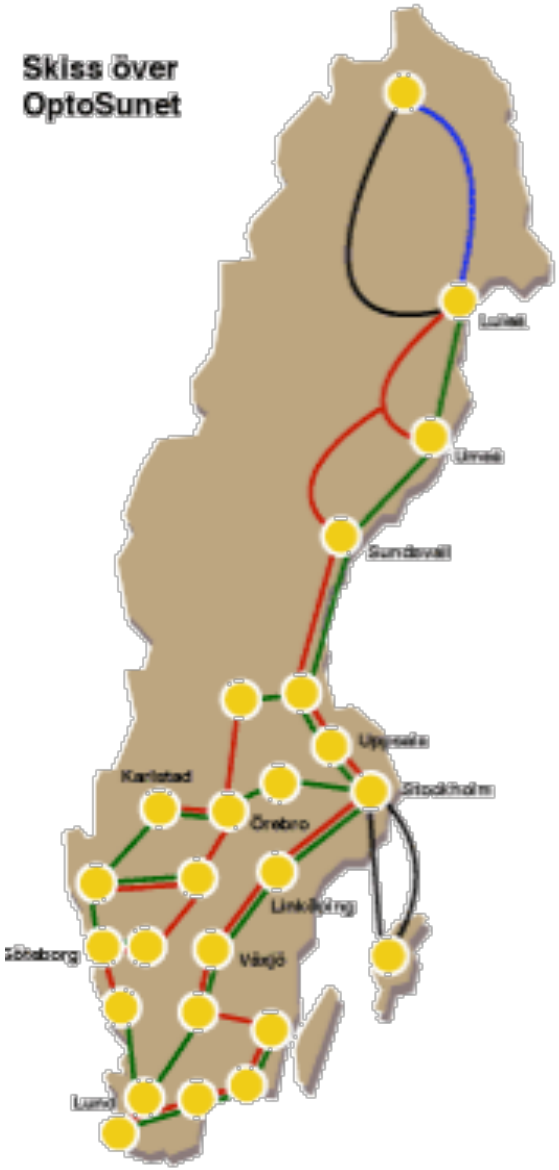
*postdoc financed by MSB

Outline

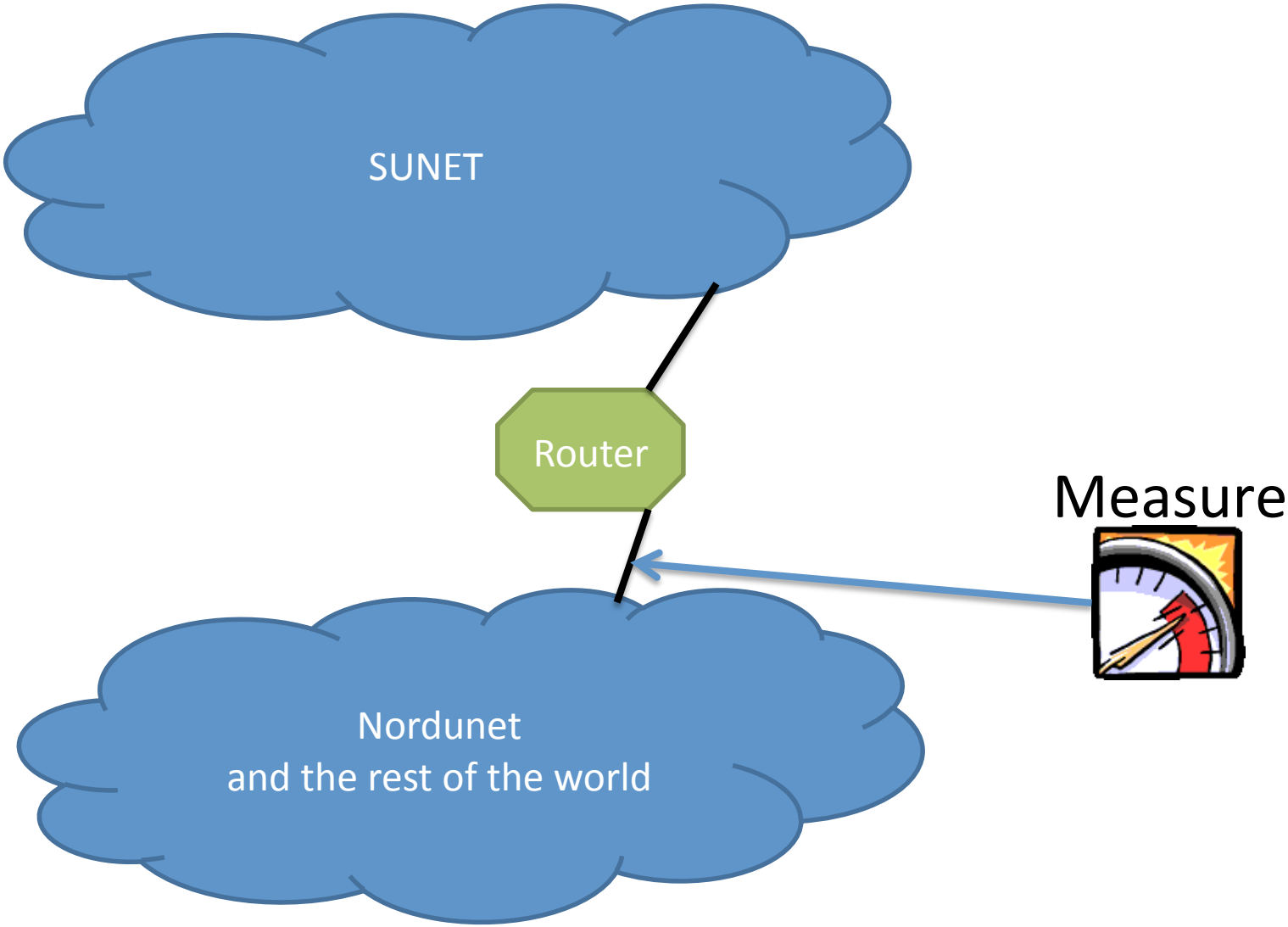
- Description of the Data Collection
- Flow Output & Overall Statistics
- Finding Malicious Hosts
- High Hitters
- The Scanner: host-A
- The Server: host-B

Analysis of malicious backbone traffic

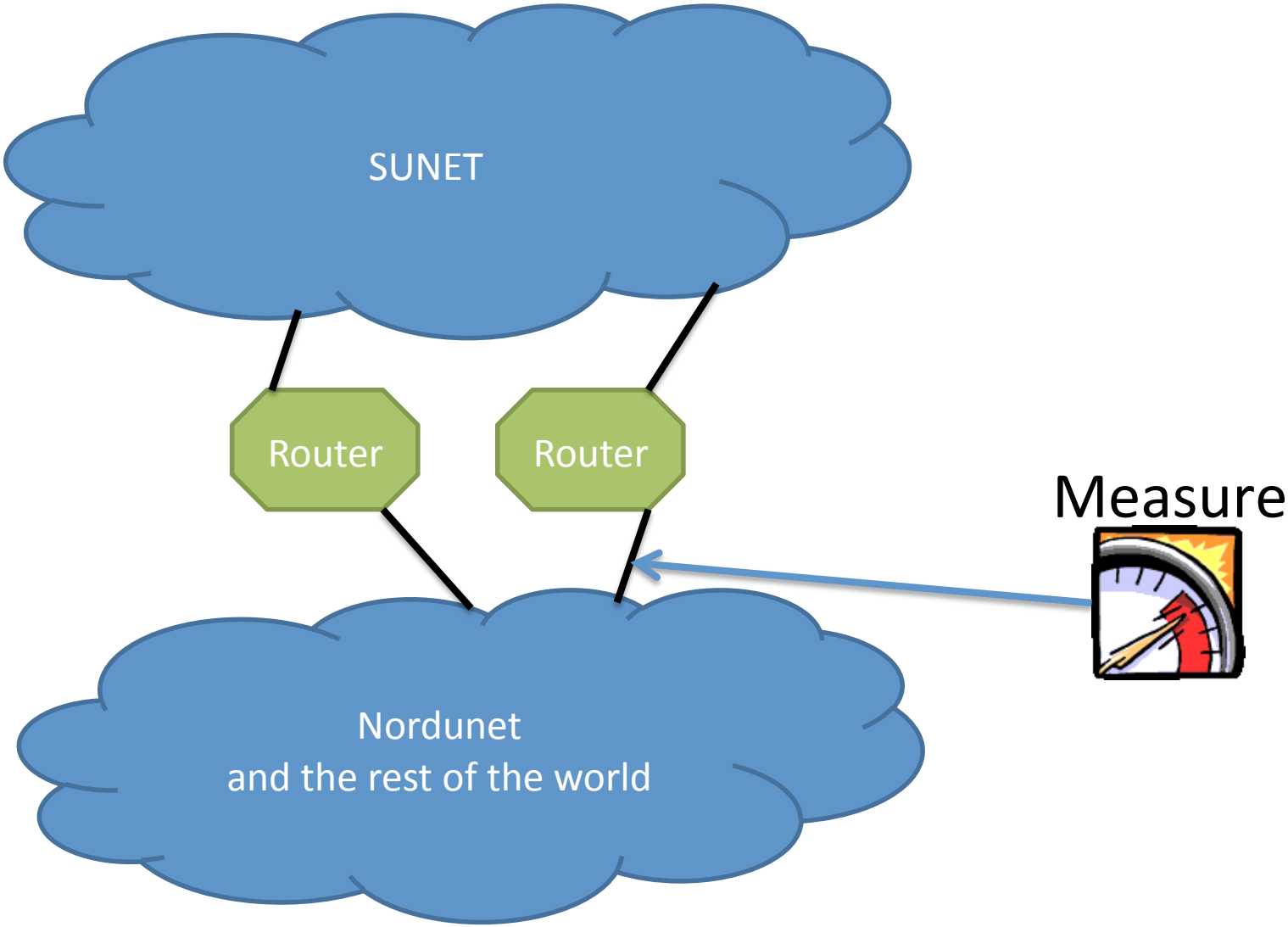
- Looking for attacks on a backbone network
  - 10 Gbps (=fast!)
  - Problems:
    - speed of network link
    - amount of data
    - routing
    - user privacy – anonymize data(!)



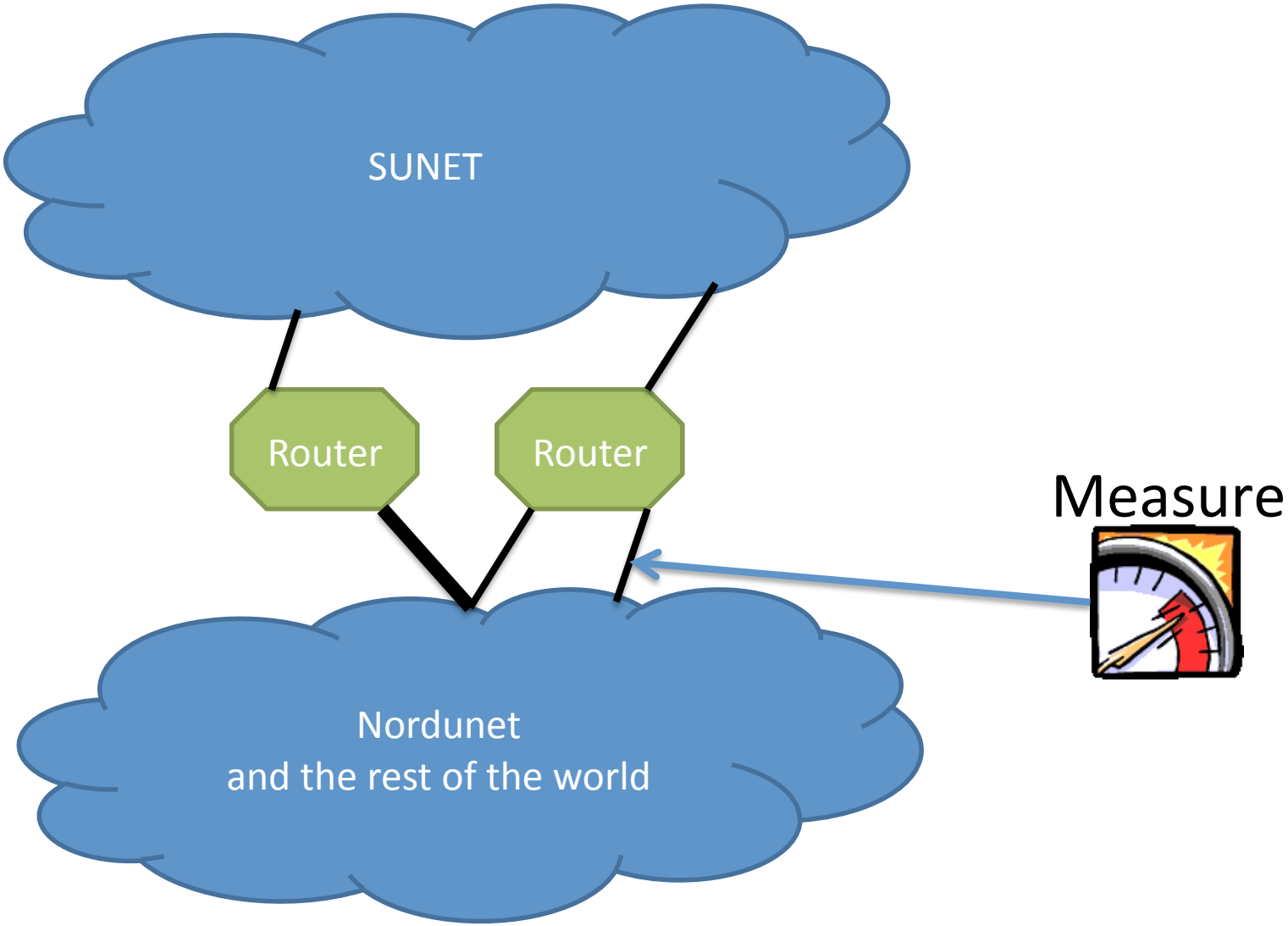
Measurement Setup (simplified)



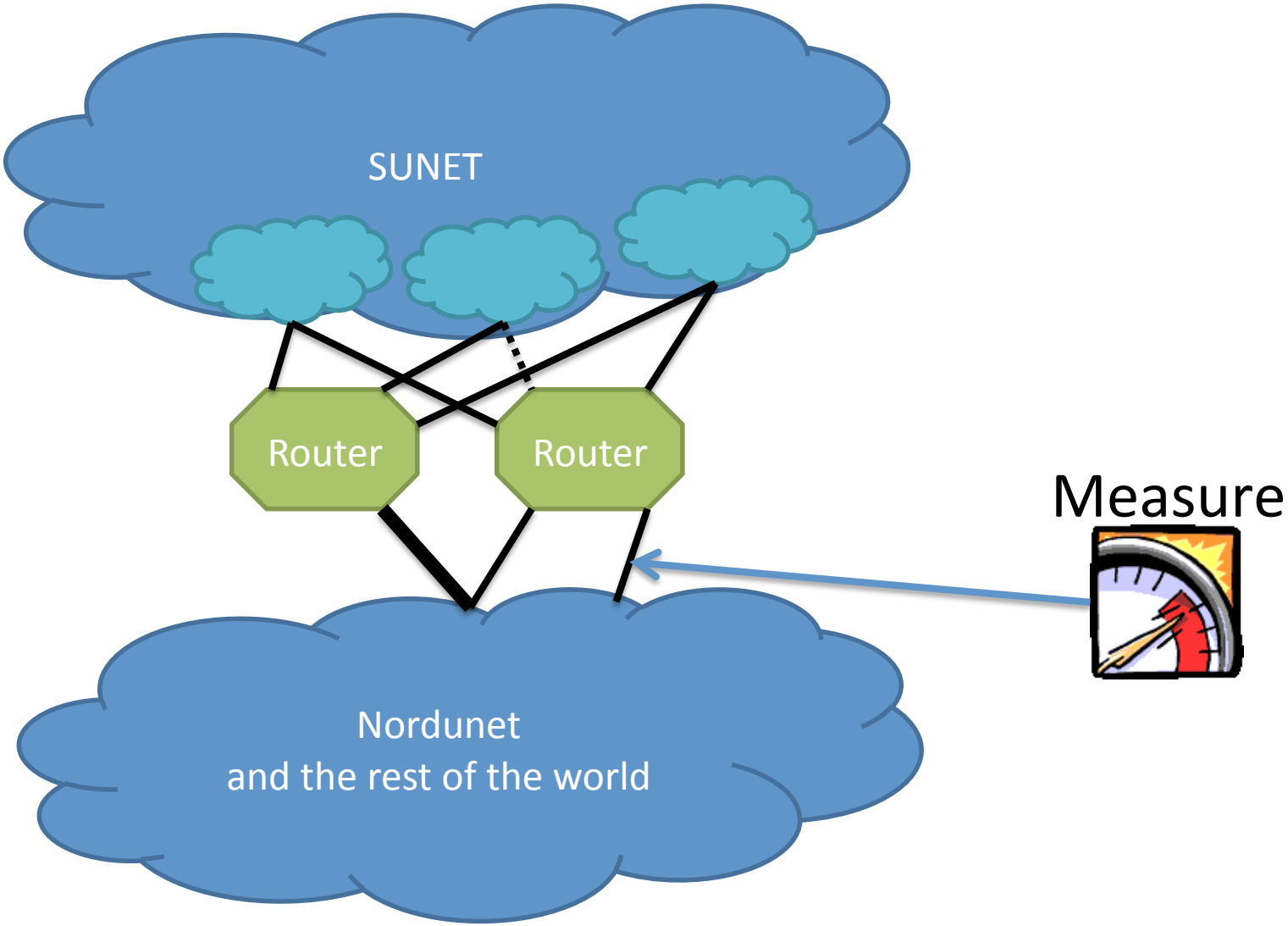
Measurement Setup (simplified)



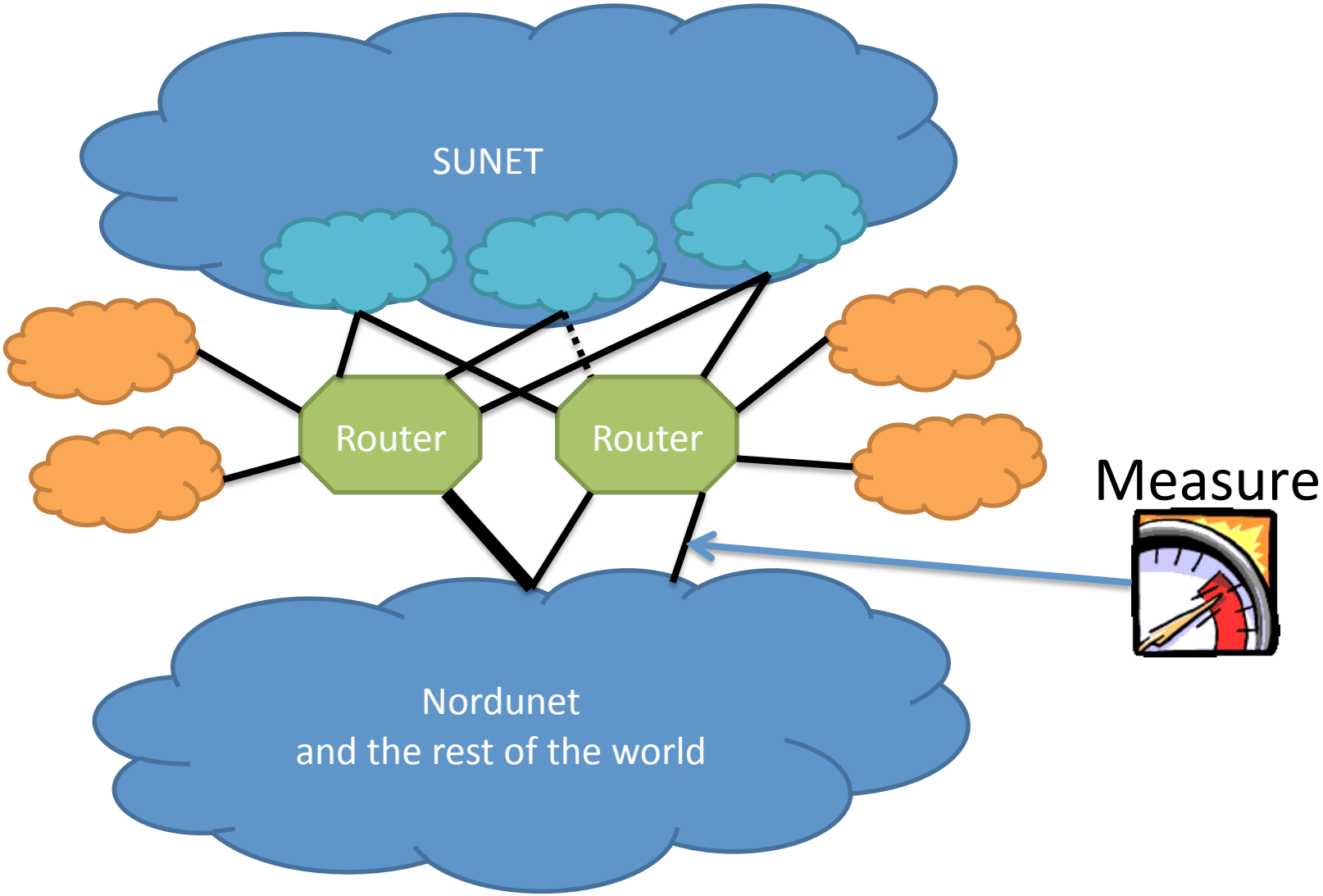
Measurement Setup (simplified)



Measurement Setup (simplified)



Measurement Setup (simplified)



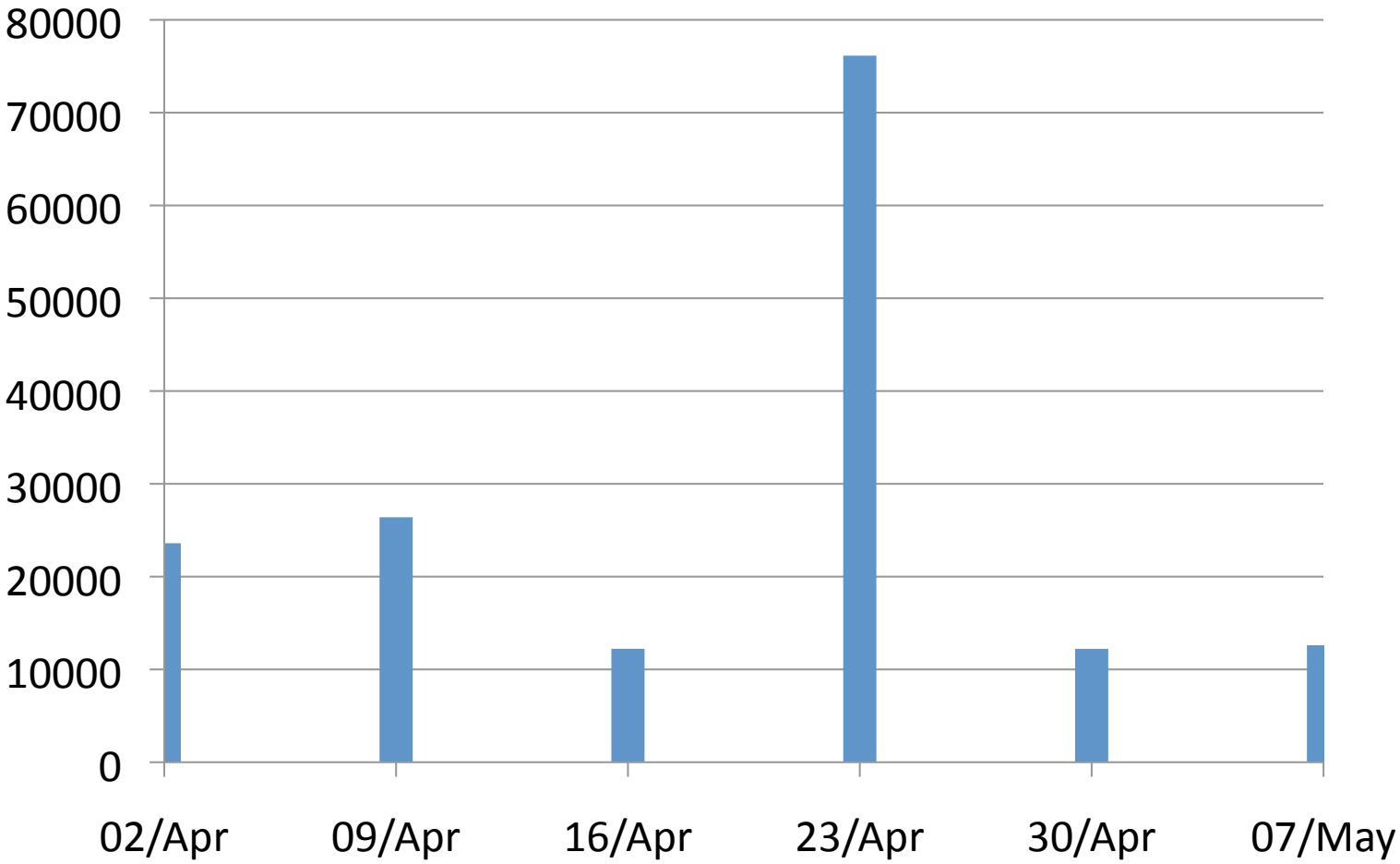
Flow Output from CoralReef

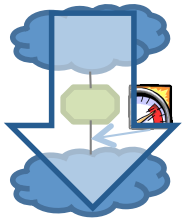
- Unidirectional flows: 5-tuple

srcIP	destIP	proto	ok	sport	dport	pkts	bytes	flows	firstTS	lastTS
src^1	dst^1	6	1	445	3995	3	120	1	t_0^1	t_n^1
src^2	dst^2	1	1	3	1	1	56	1	t_0^2	t_n^2

- Data collected over a 6-week period (spring)
- IP addresses anonymized

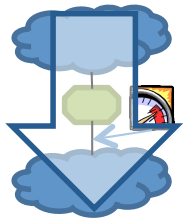
Unique Source IPs (outgoing link)



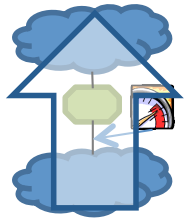


Overall Statistics

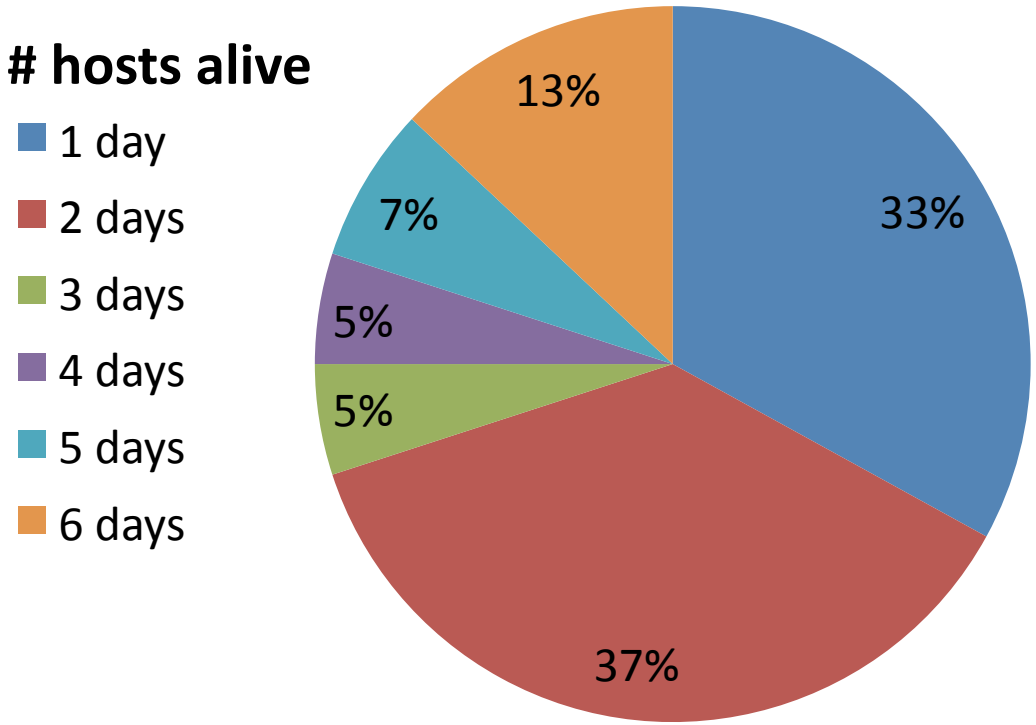
- 23,600 ***inside hosts*** initiating communication with 18,780,894 on the outside.



- 24,587,096 ***outside hosts*** trying to reach (scan) 970,149 inside hosts.

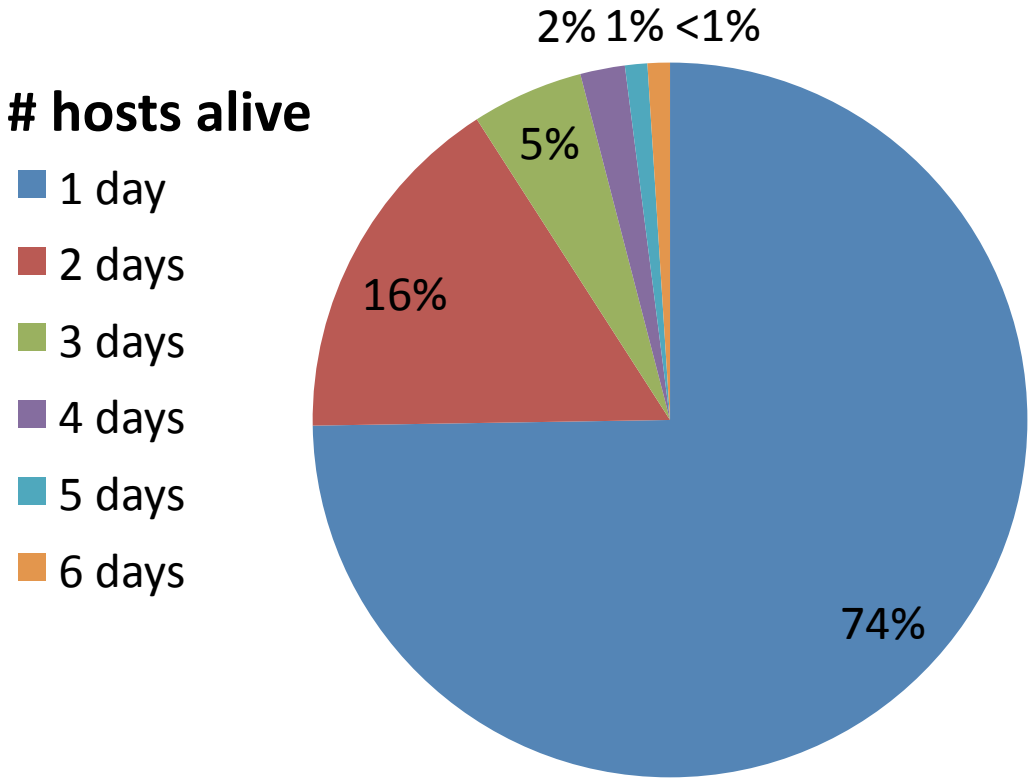


#days a particular host is active



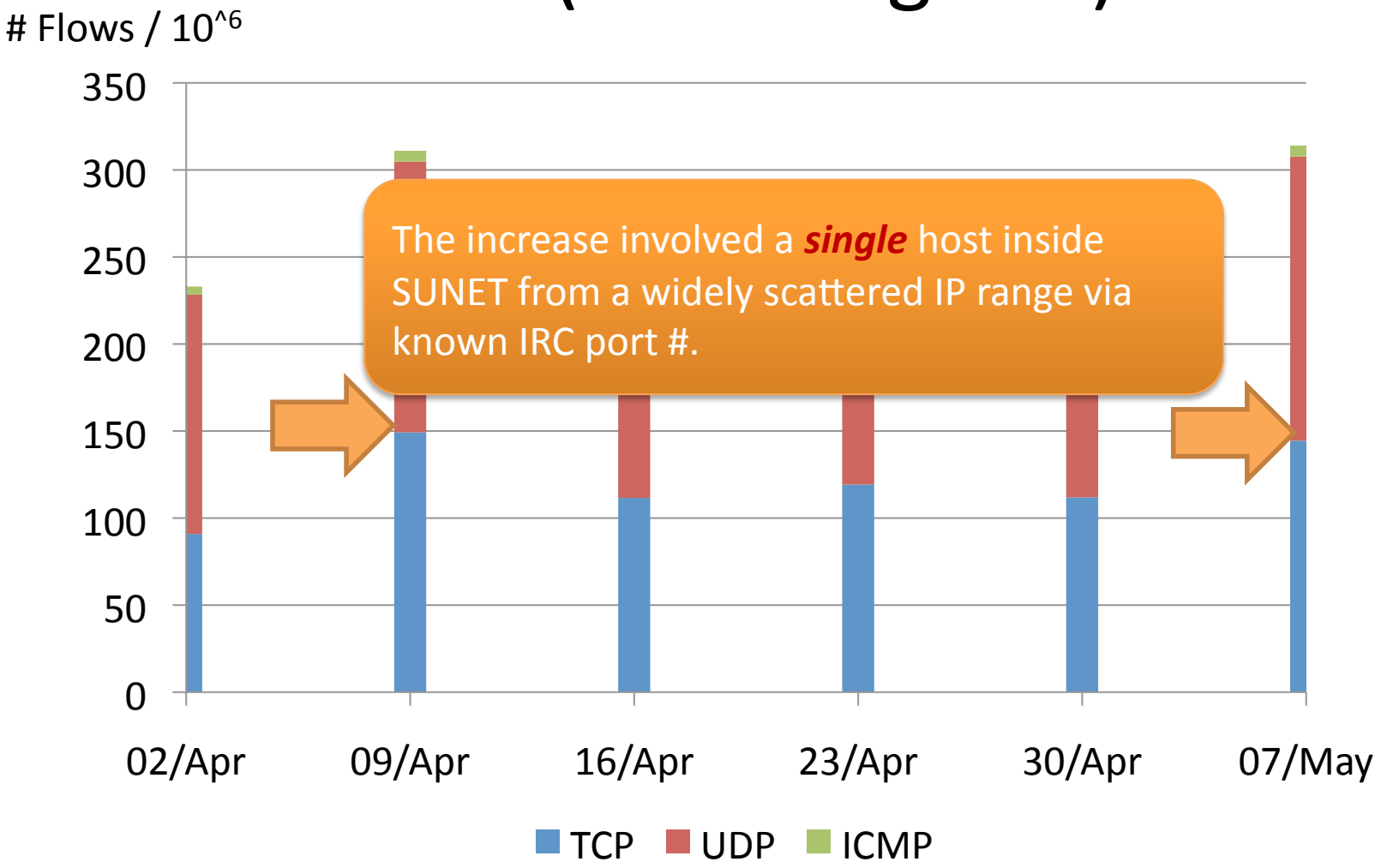
src = inside SUNET

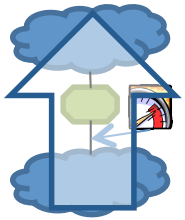
#days a particular host is active



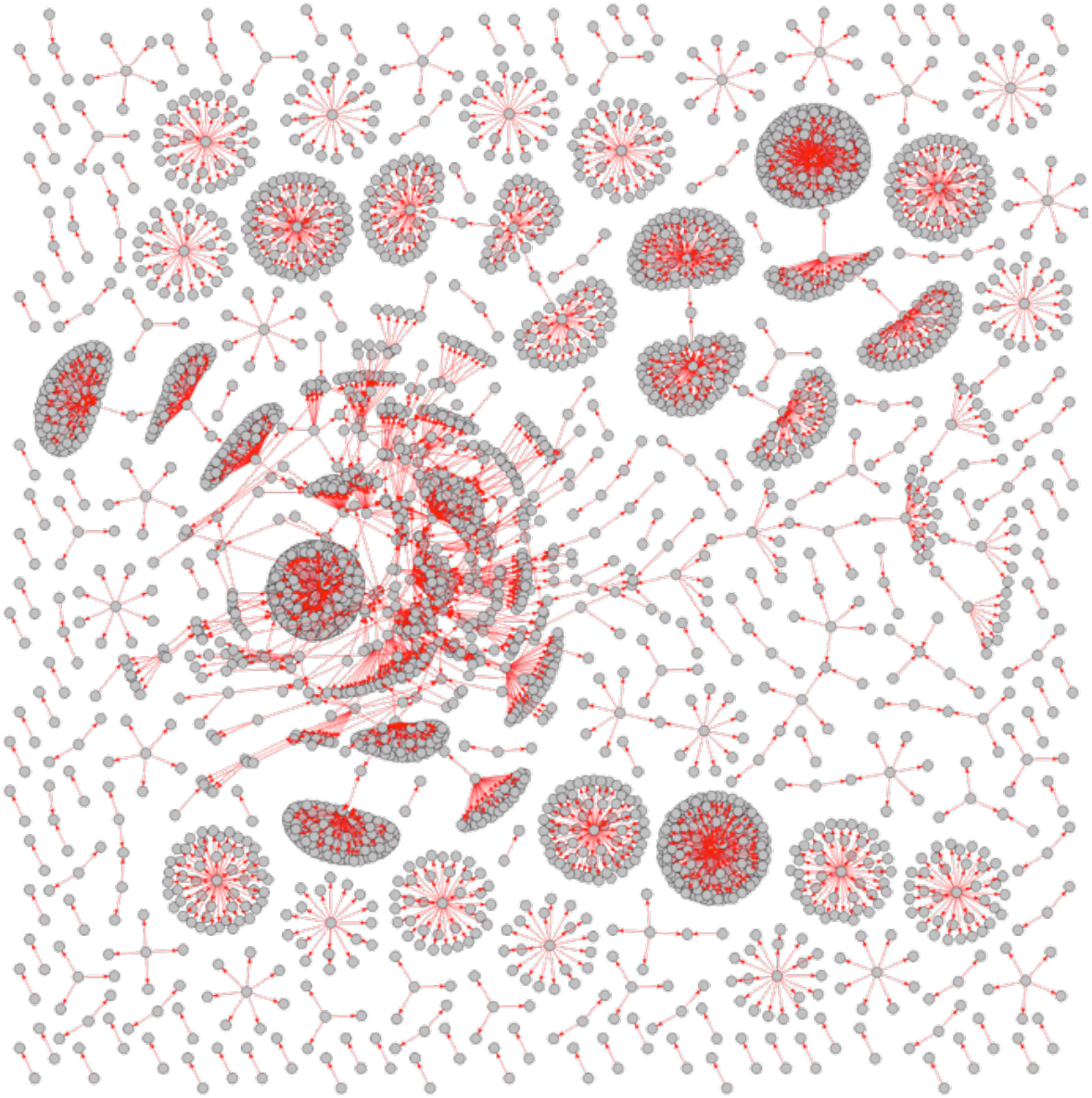
dst = outside SUNET

Outstanding Patterns: # Flows (incoming link)

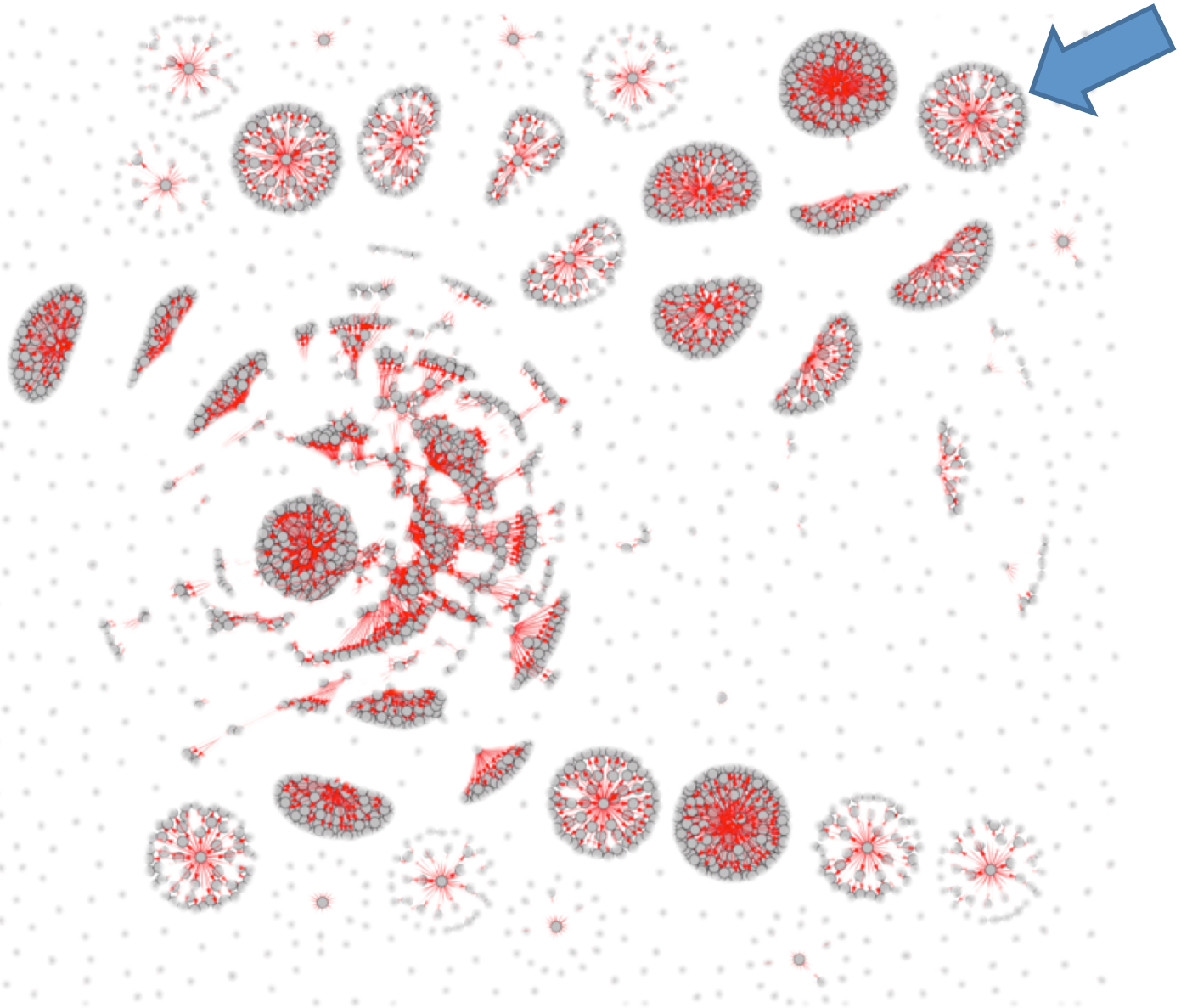




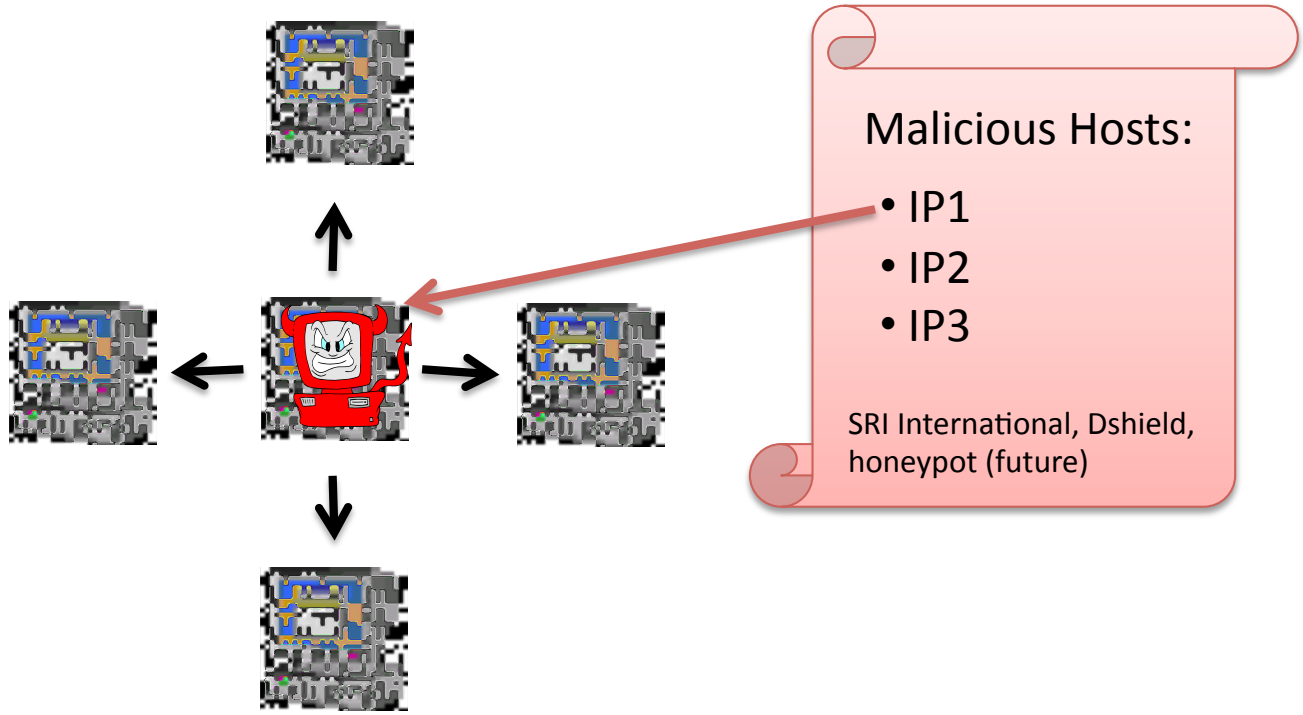
Finding Malicious Hosts



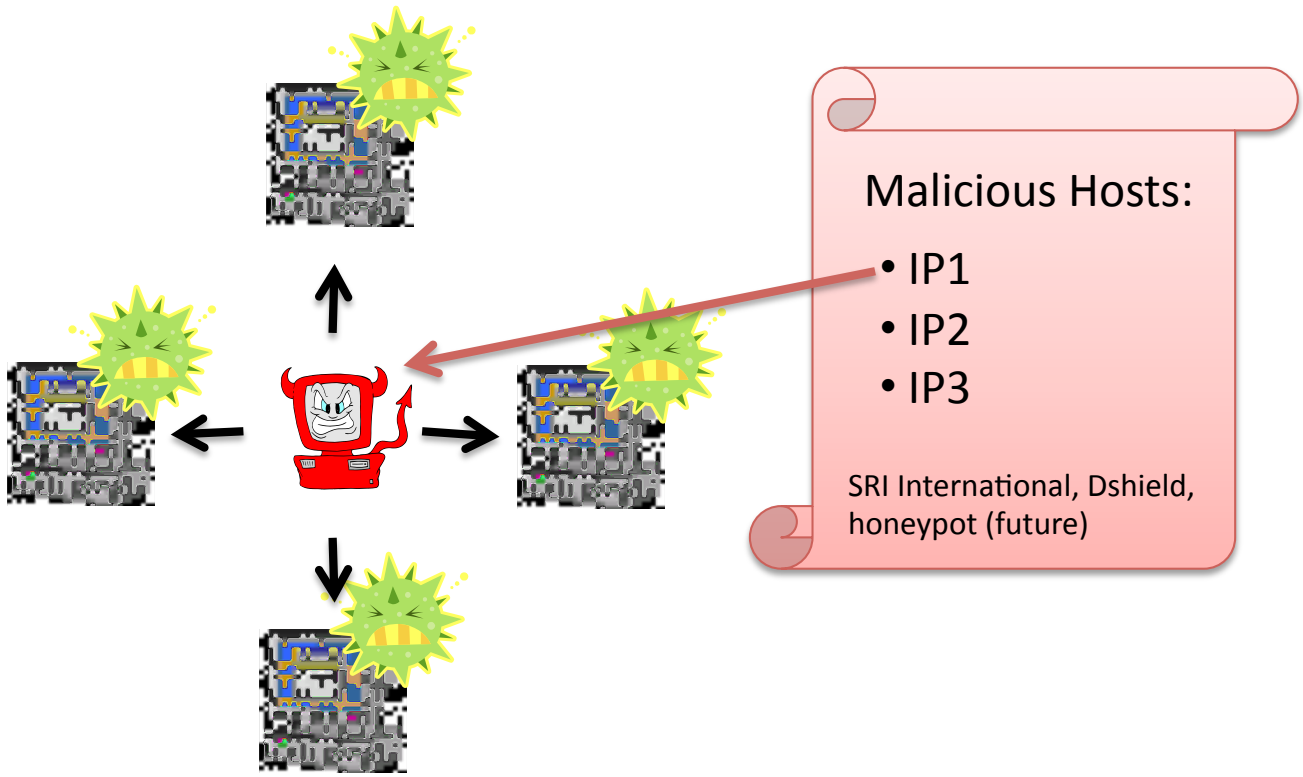
Finding Malicious Hosts



Finding Malicious Hosts



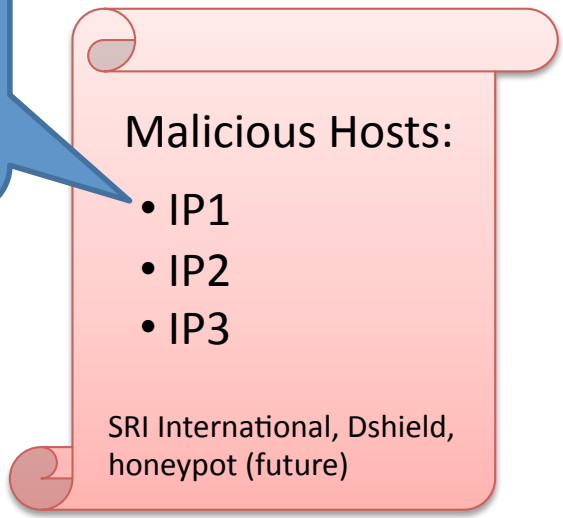
Finding Malicious Hosts



The Malicious Host List

- Currently
  - Dshield's
  - Results from SRI International Malware Threat Center
- Downloading an updated version regularly.
- Future work:
  - Connect local honeypots to dynamically analyze "local" attacks.
  - Be able to discriminate directly based on the flows.

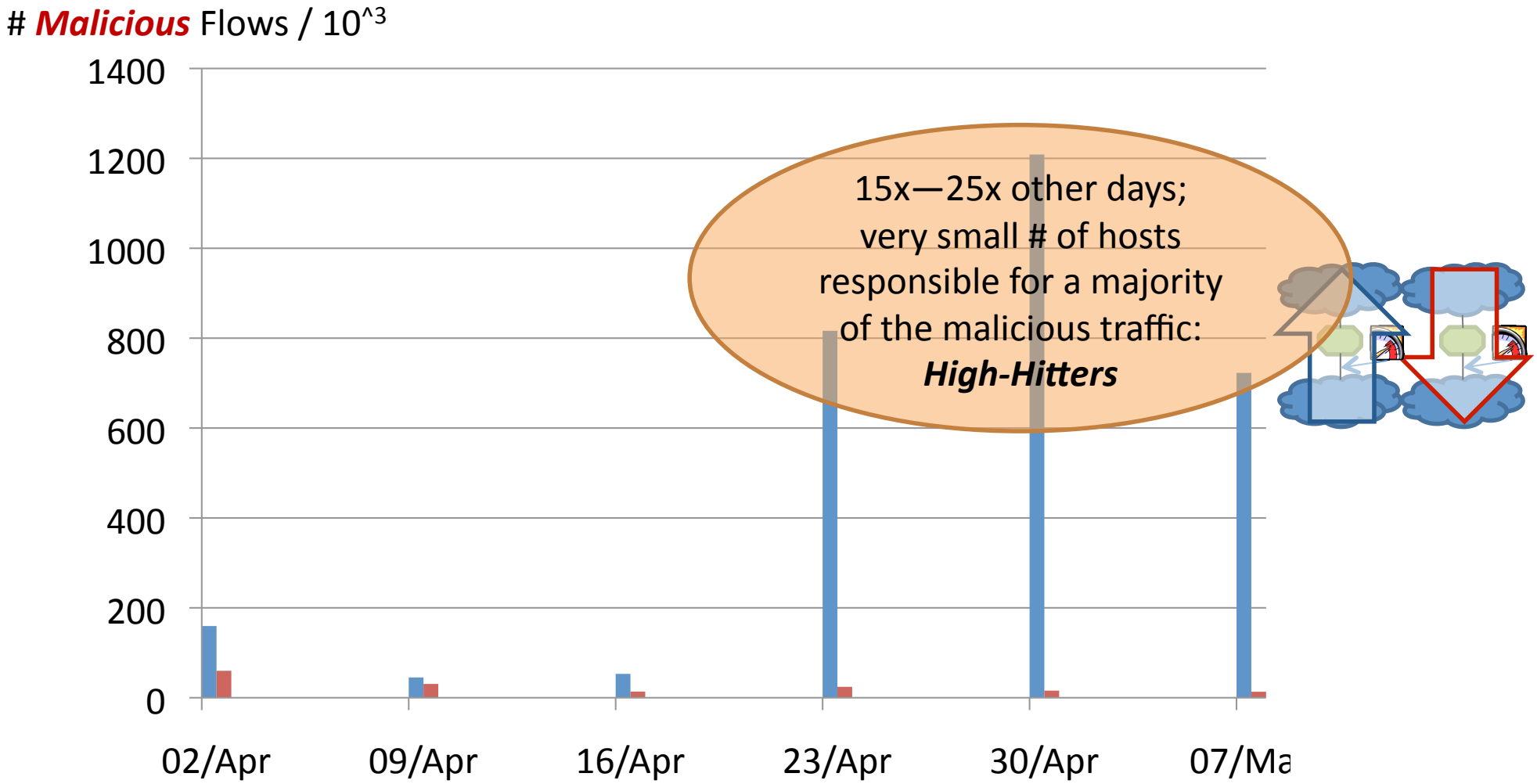
We saw activity from about 5% of the hosts in this set.



Outline

- Description of the Data Collection
- Flow Output & Overall Statistics
- Finding Malicious Hosts
- High Hitters
- The Scanner: host-A
- The Server: host-B

Malicious Hosts: High-Hitters



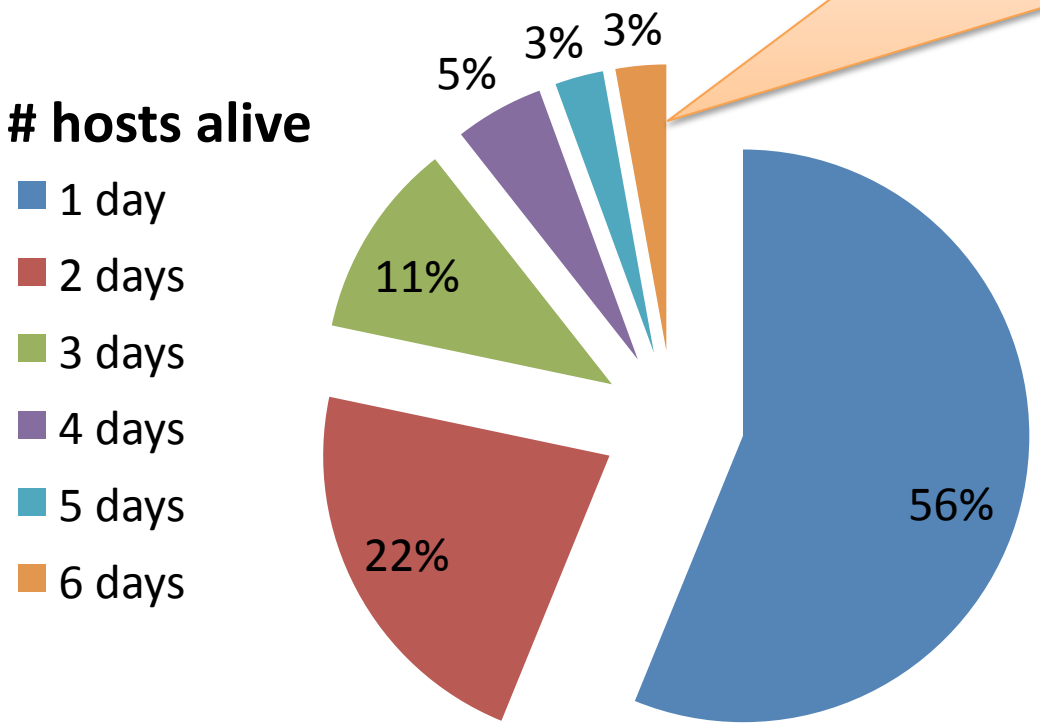
Analysis of High-Hitters

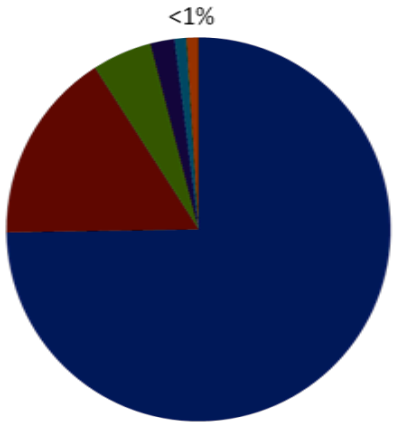
- April 1:
  - 1H-H: 83% of the malicious traffic
  - Port # + packet size ➔ spreading Sapphire worm
- April 22:
  - 5 H-H: 97% of the malicious traffic
    - 3 H-H: Tried to log into 300,000 hosts over SSH or VNC
    - 2 H-H: contacted about 60,000 hosts (varying port #)
- May 6:
  - 1 H-H: 96% of the malicious traffic
  - Scanning port 1433 (MySQL) always with the same sport 6000

The Ubiquitous Malicious Hosts

Malicious hosts present all days.

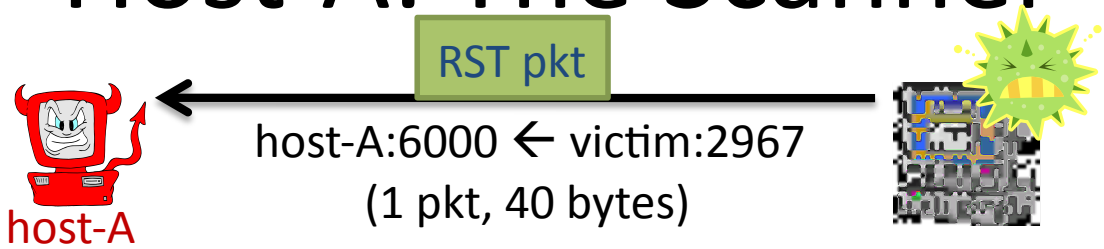
Only 3% of all M-hosts, but have 26% of the traffic.



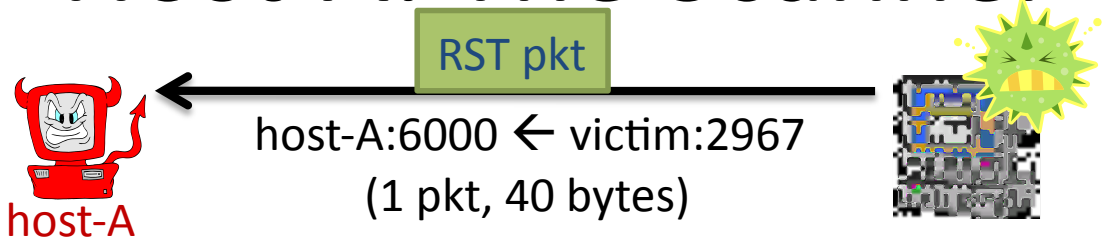


Host-A: The Scanner
Host-B: Reconnecting to the Server

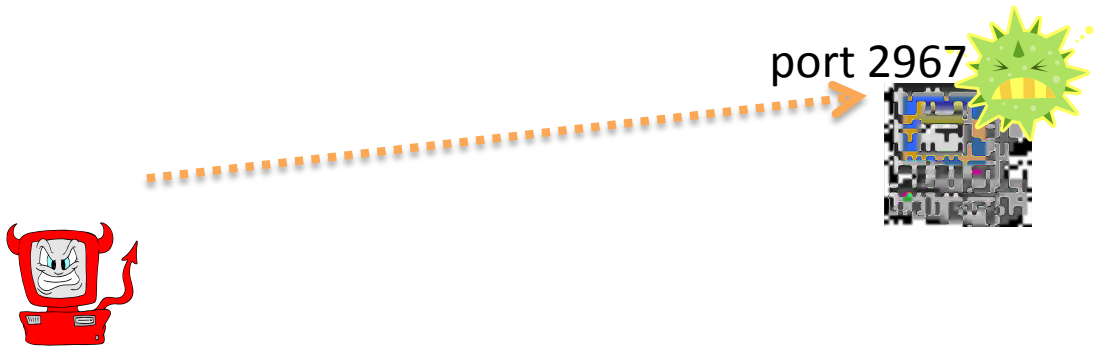
Host-A: The Scanner



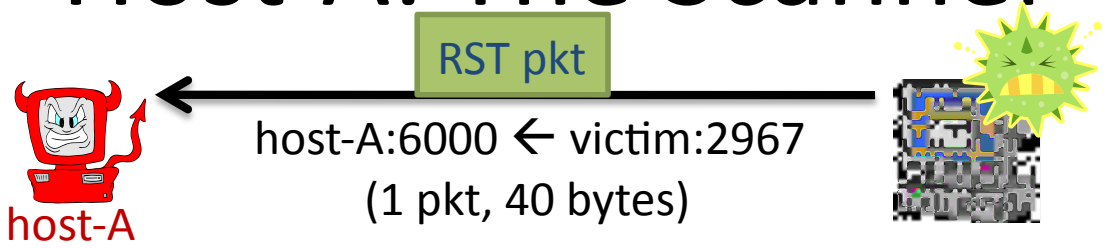
Host-A: The Scanner



April 1: 2967 probed

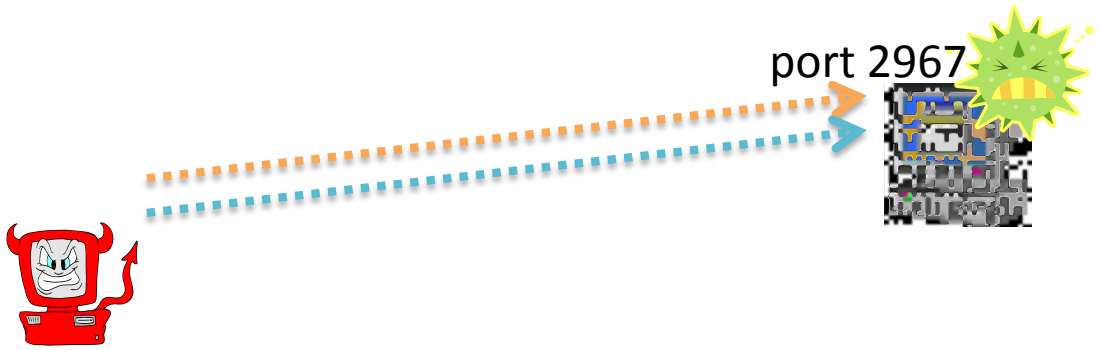


Host-A: The Scanner

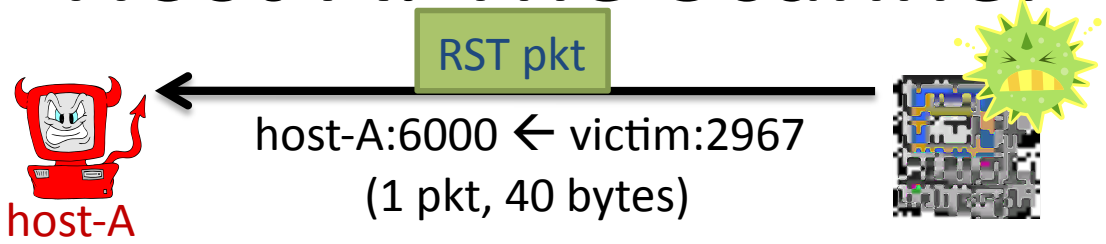


April 1: 2967 probed

April 8: 2967 probed



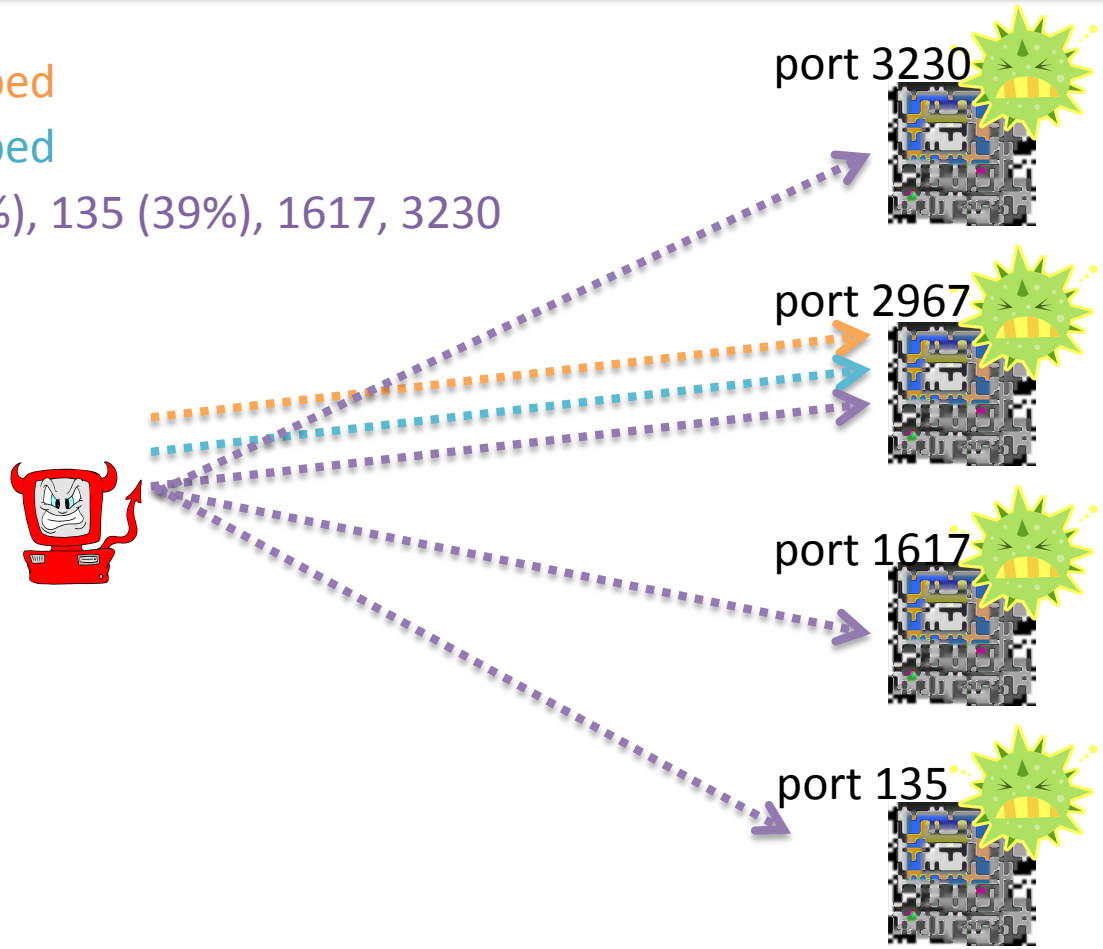
Host-A: The Scanner



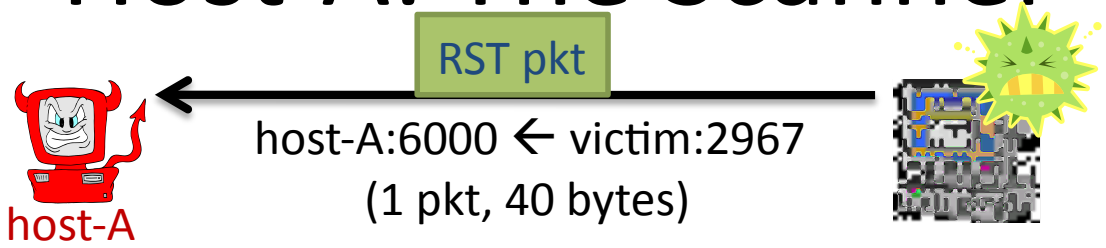
April 1: 2967 probed

April 8: 2967 probed

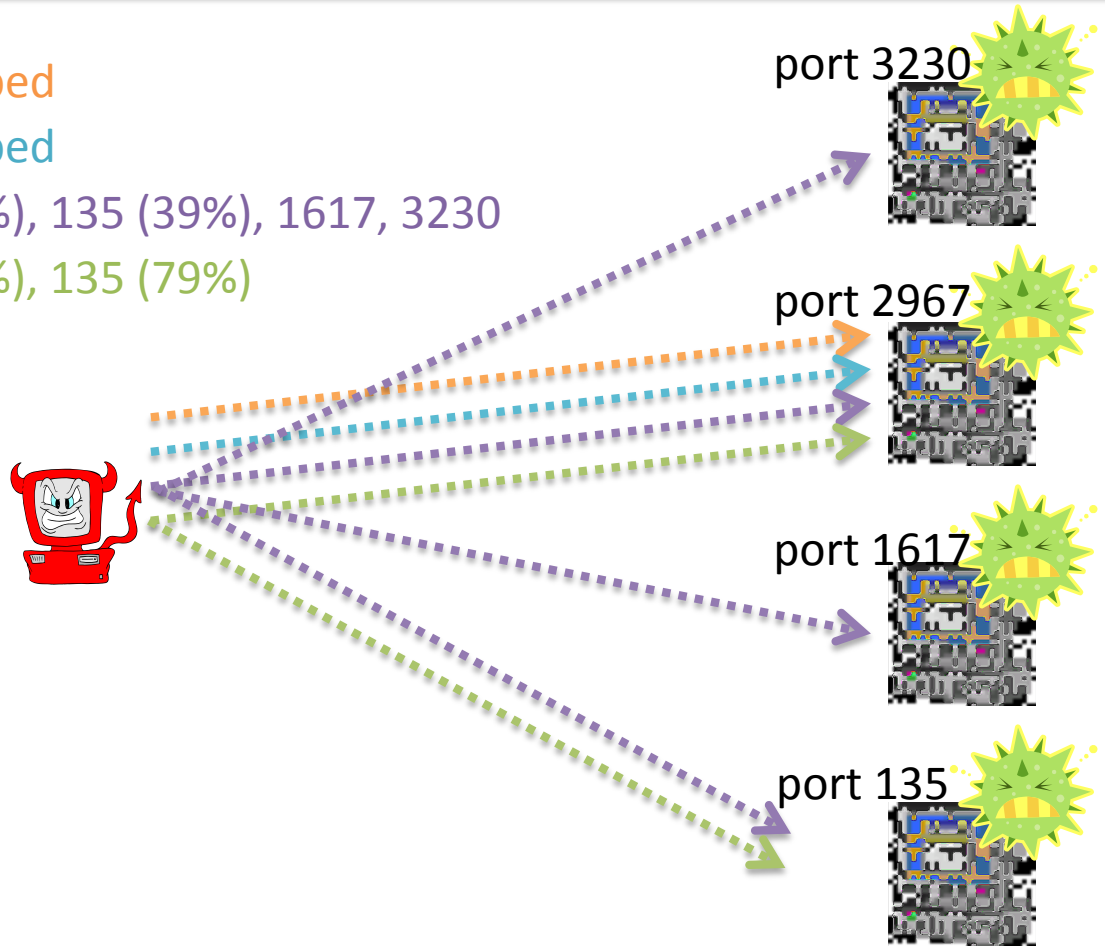
April 15: 2967 (51%), 135 (39%), 1617, 3230



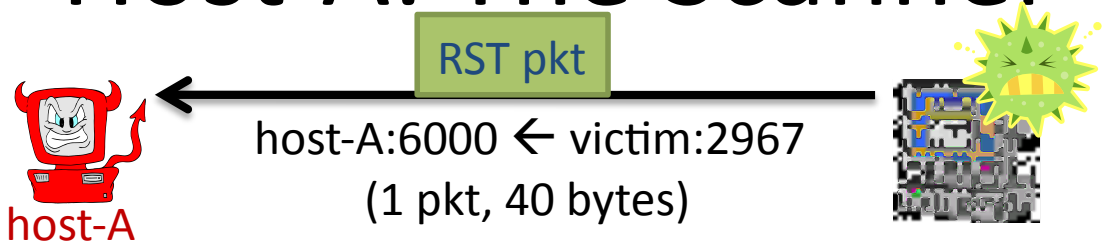
Host-A: The Scanner



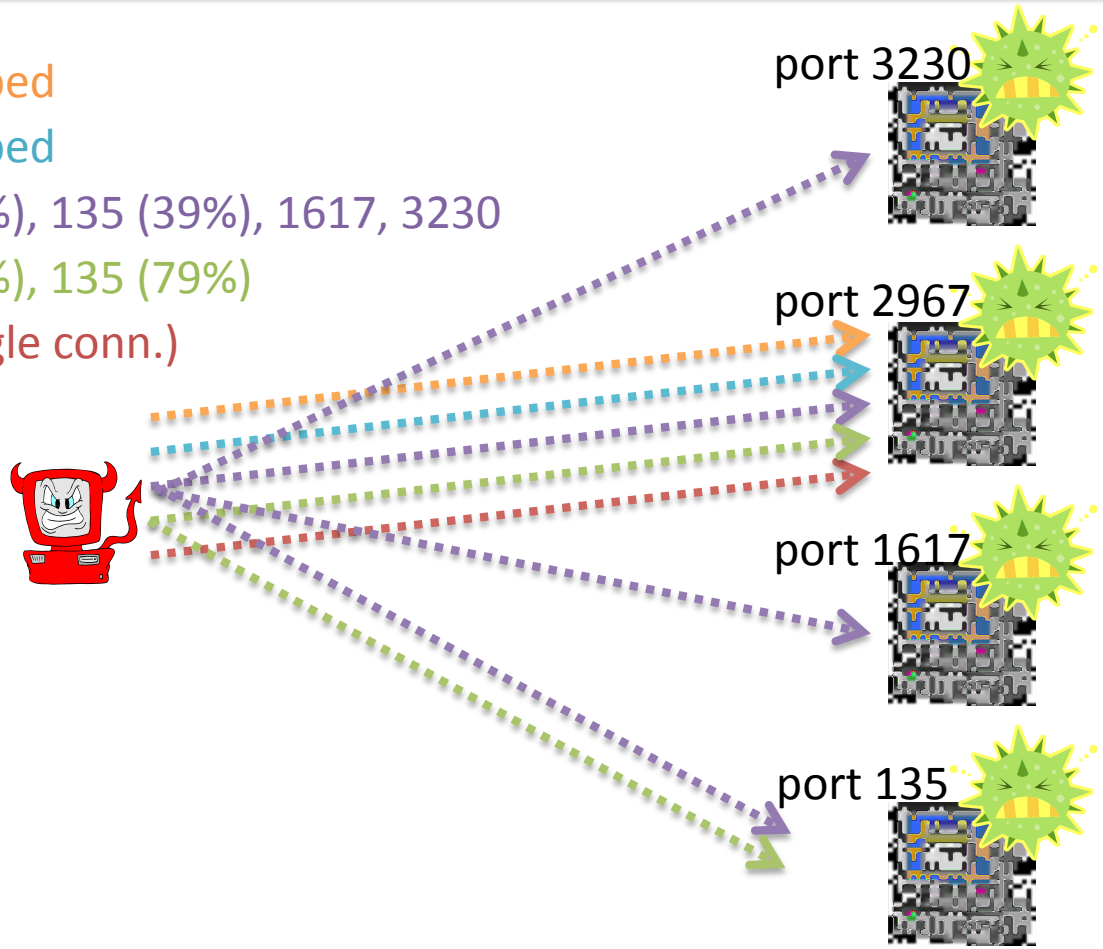
April 1: 2967 probed
April 8: 2967 probed
April 15: 2967 (51%), 135 (39%), 1617, 3230
April 22: 2967 (21%), 135 (79%)



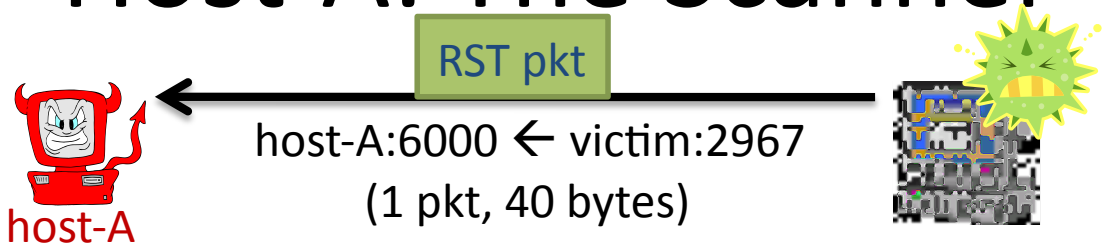
Host-A: The Scanner



April 1: 2967 probed
April 8: 2967 probed
April 15: 2967 (51%), 135 (39%), 1617, 3230
April 22: 2967 (21%), 135 (79%)
April 29: 2967 (single conn.)



Host-A: The Scanner



April 1: 2967 probed

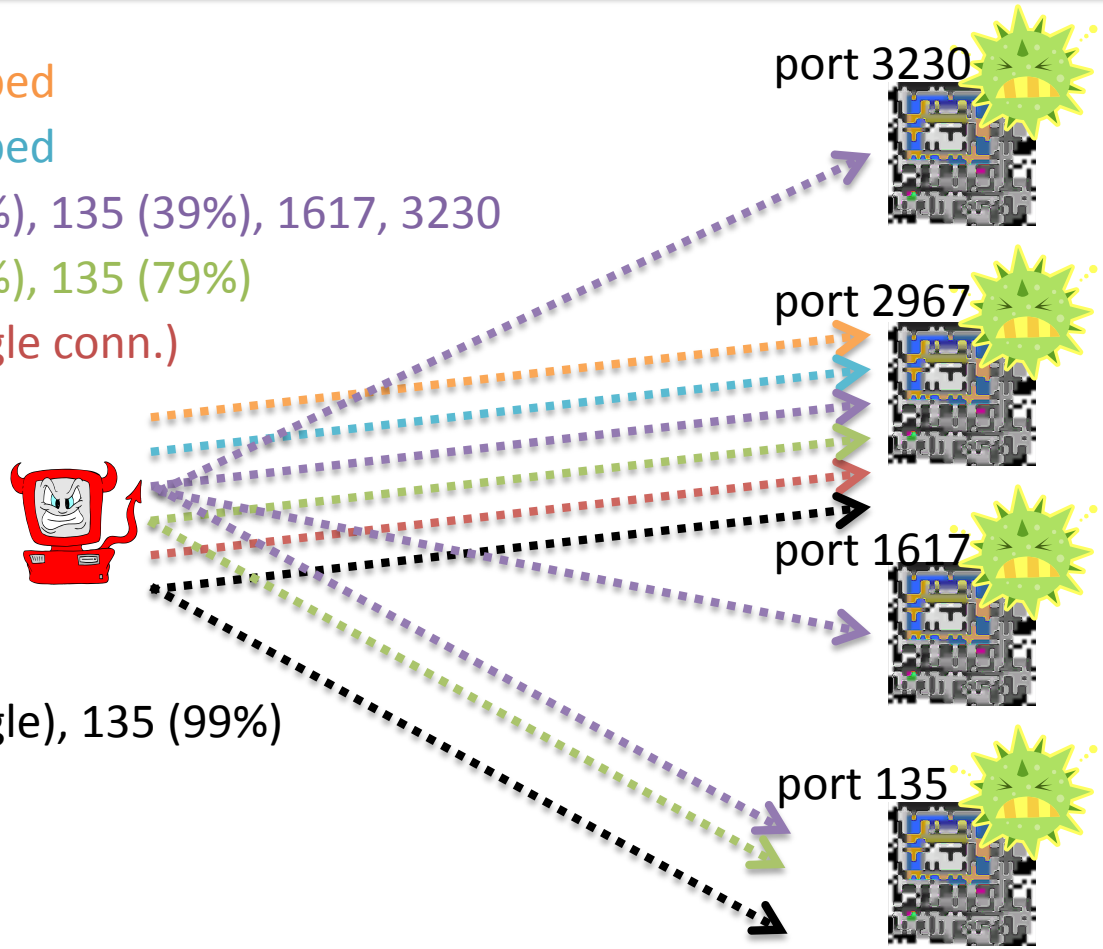
April 8: 2967 probed

April 15: 2967 (51%), 135 (39%), 1617, 3230

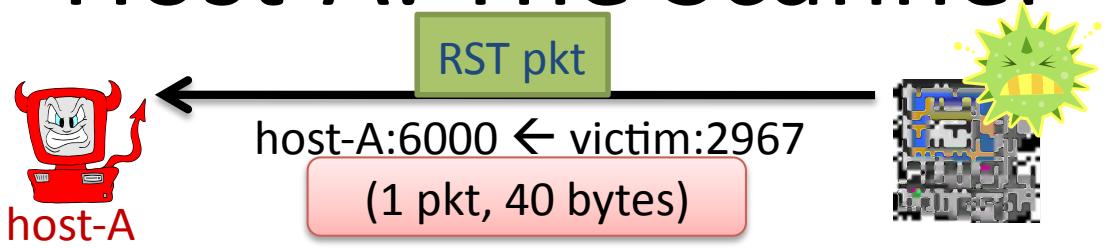
April 22: 2967 (21%), 135 (79%)

April 29: 2967 (single conn.)

May 5: 2967 (single), 135 (99%)

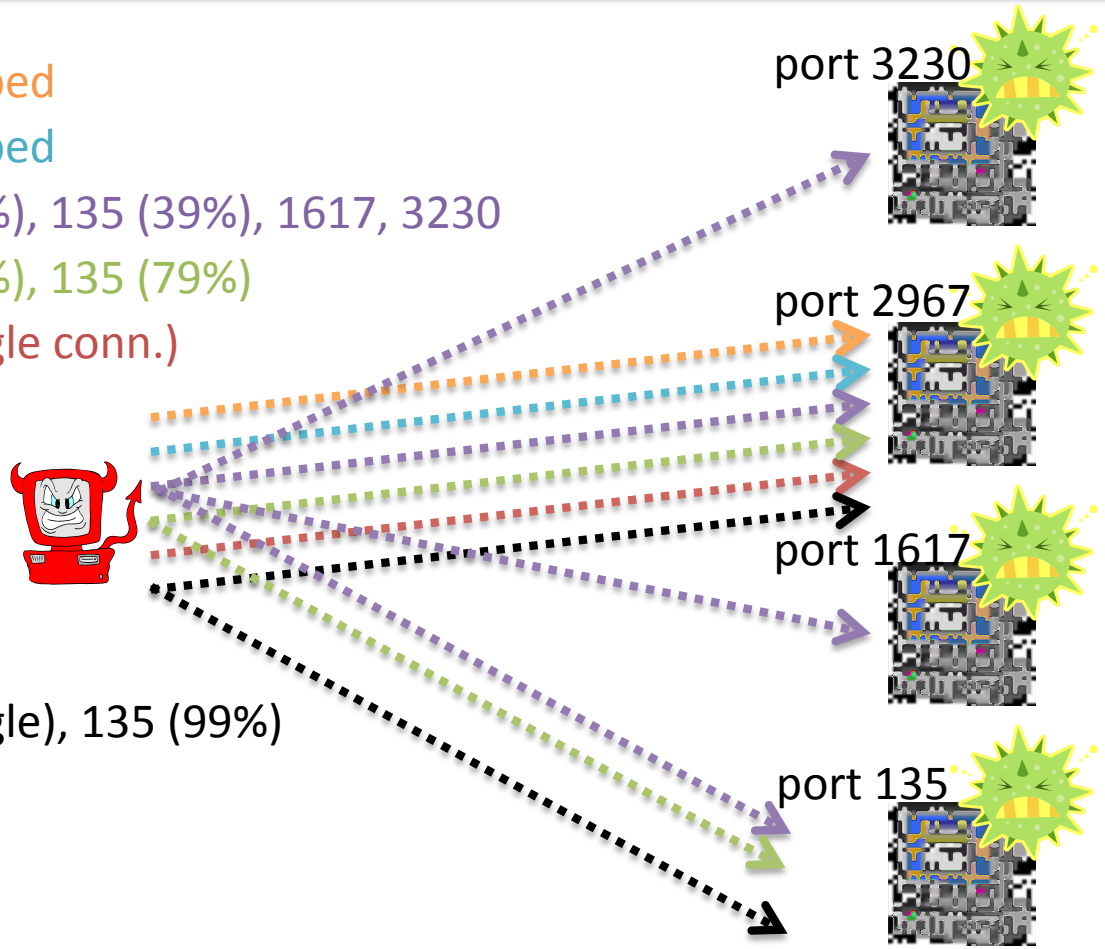


Host-A: The Scanner

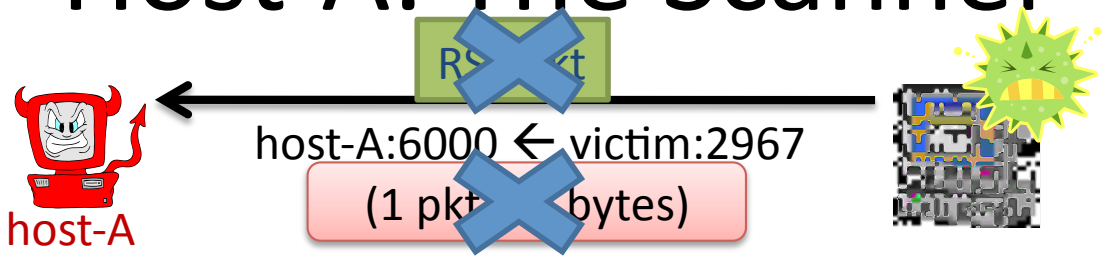


April 1: 2967 probed
April 8: 2967 probed
April 15: 2967 (51%), 135 (39%), 1617, 3230
April 22: 2967 (21%), 135 (79%)
April 29: 2967 (single conn.)

May 5: 2967 (single), 135 (99%)



Host-A: The Scanner



April 1: 2967 probed

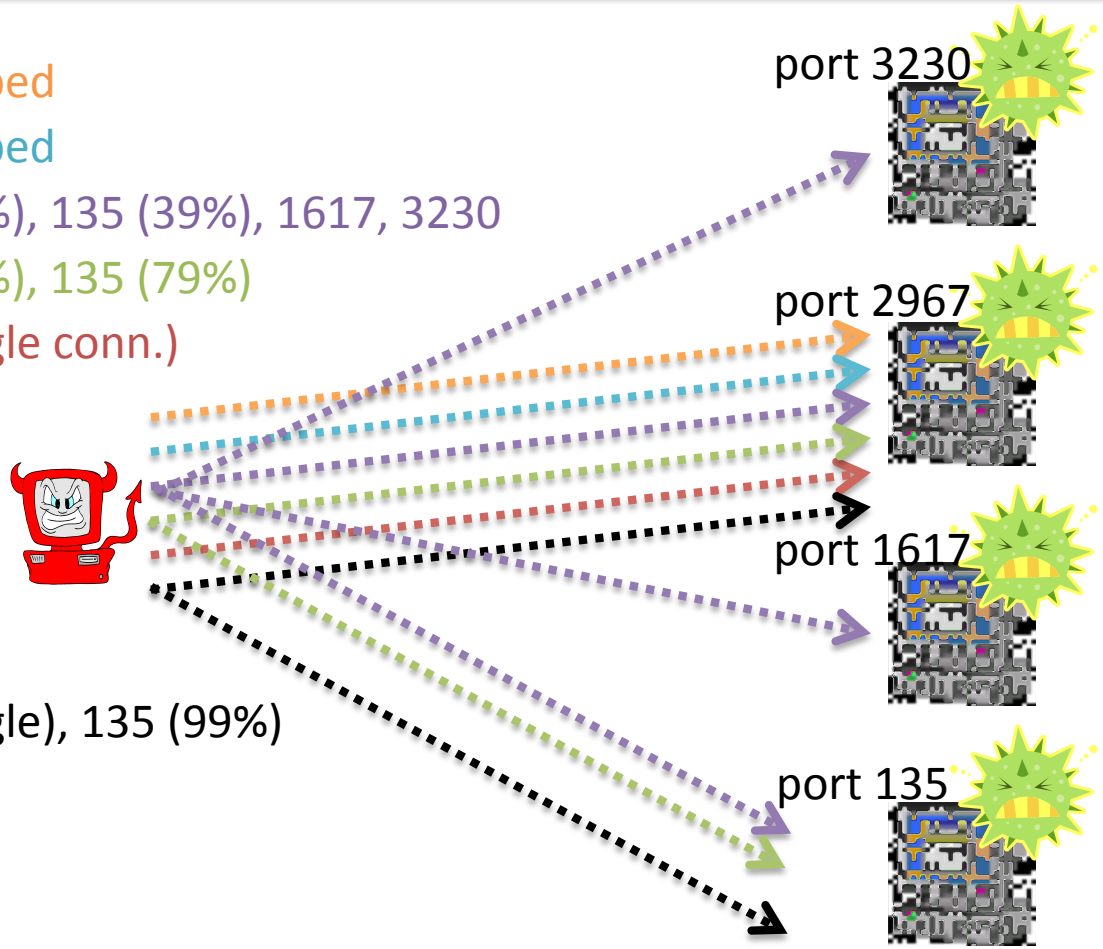
April 8: 2967 probed

April 15: 2967 (51%), 135 (39%), 1617, 3230

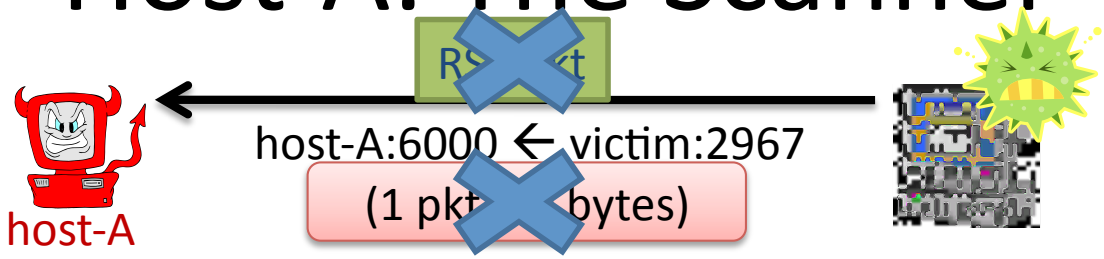
April 22: 2967 (21%), 135 (79%)

April 29: 2967 (single conn.)

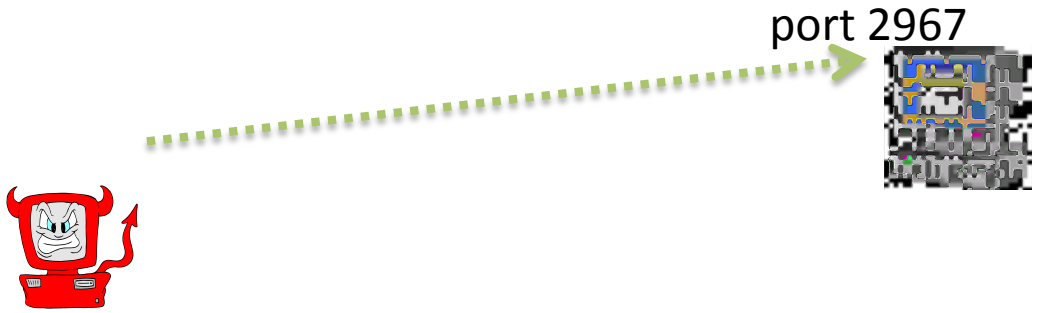
May 5: 2967 (single), 135 (99%)



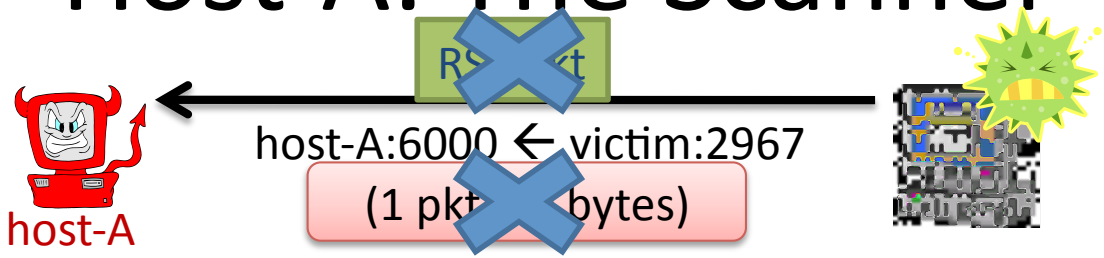
Host-A: The Scanner



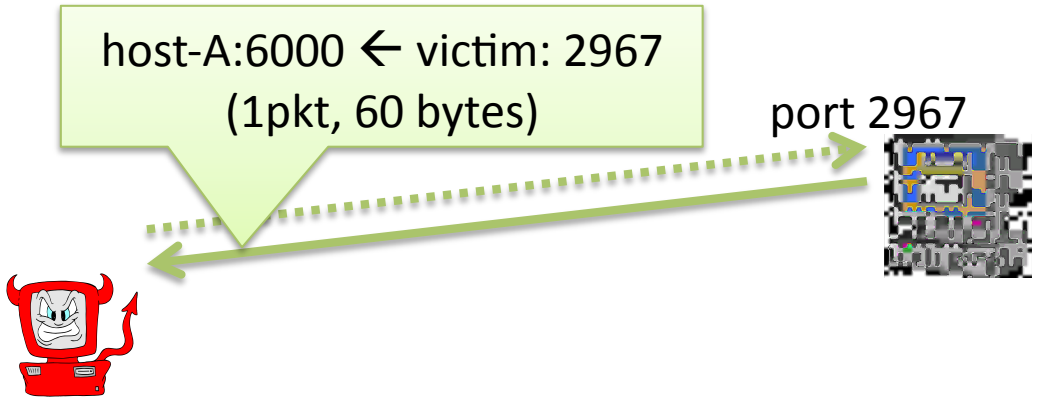
April 22:



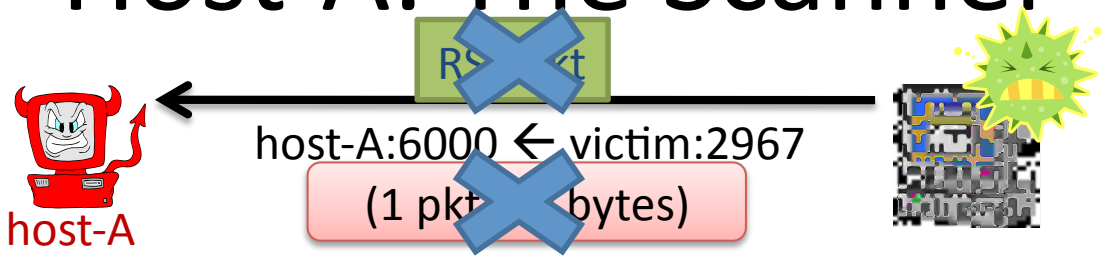
Host-A: The Scanner



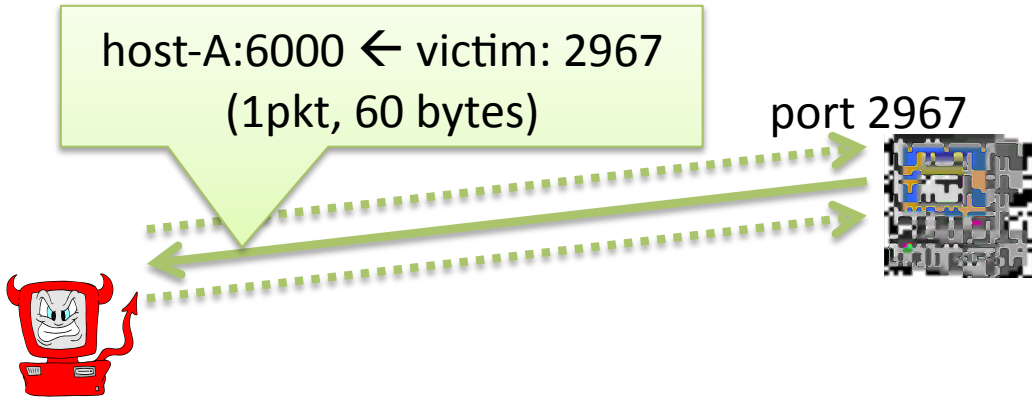
April 22:



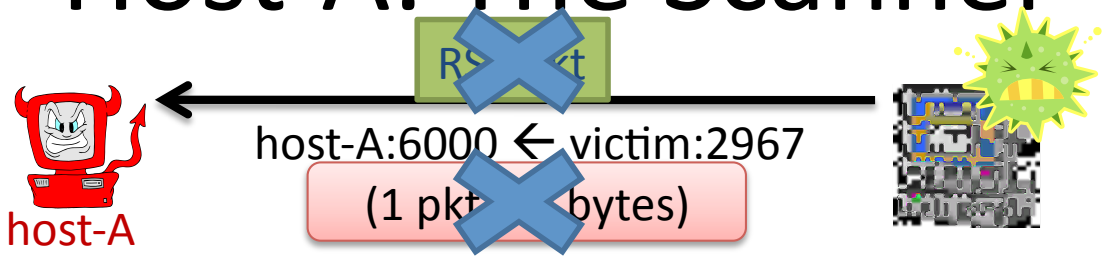
Host-A: The Scanner



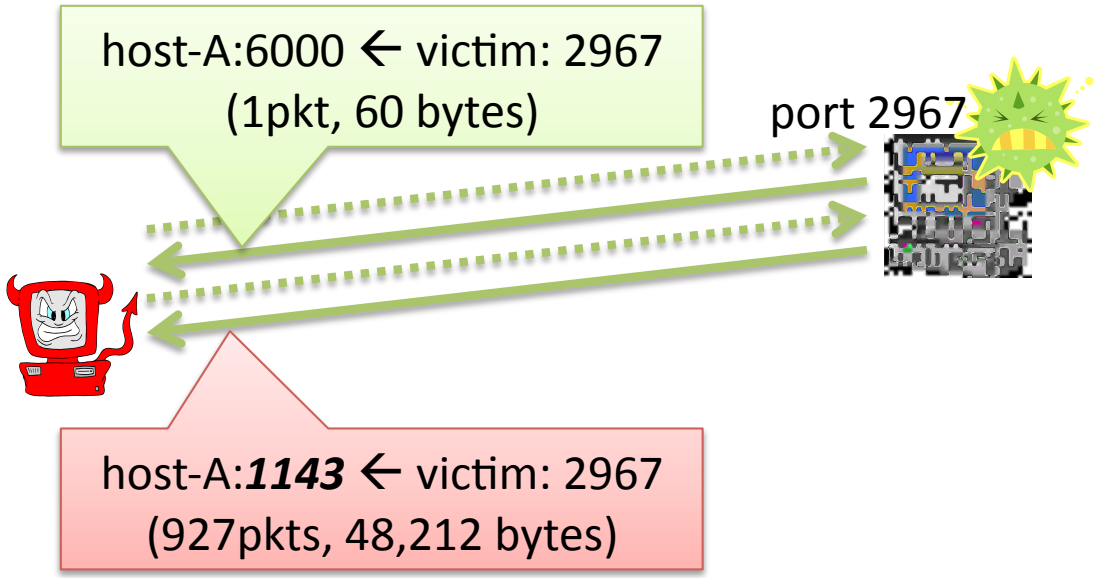
April 22:



Host-A: The Scanner

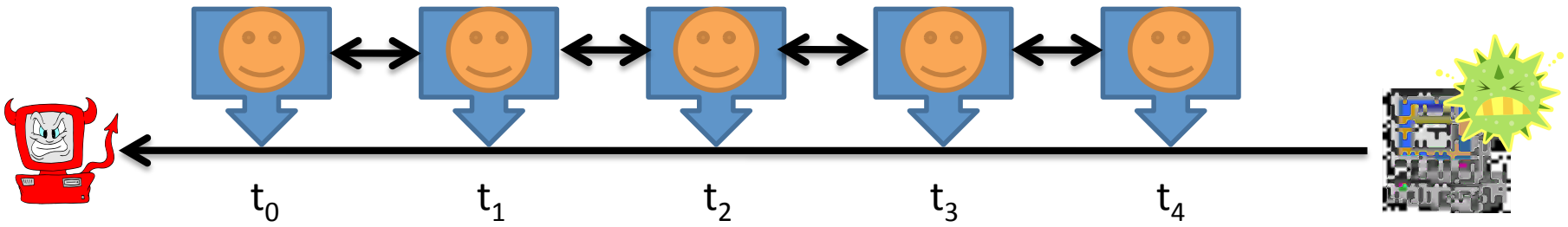


April 22:

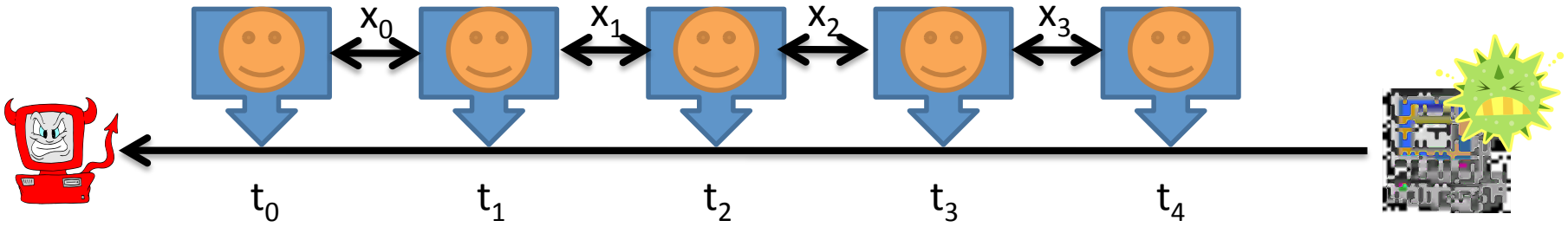


Timing Behavior of Malicious Hosts

Timing Behavior of Malicious Hosts

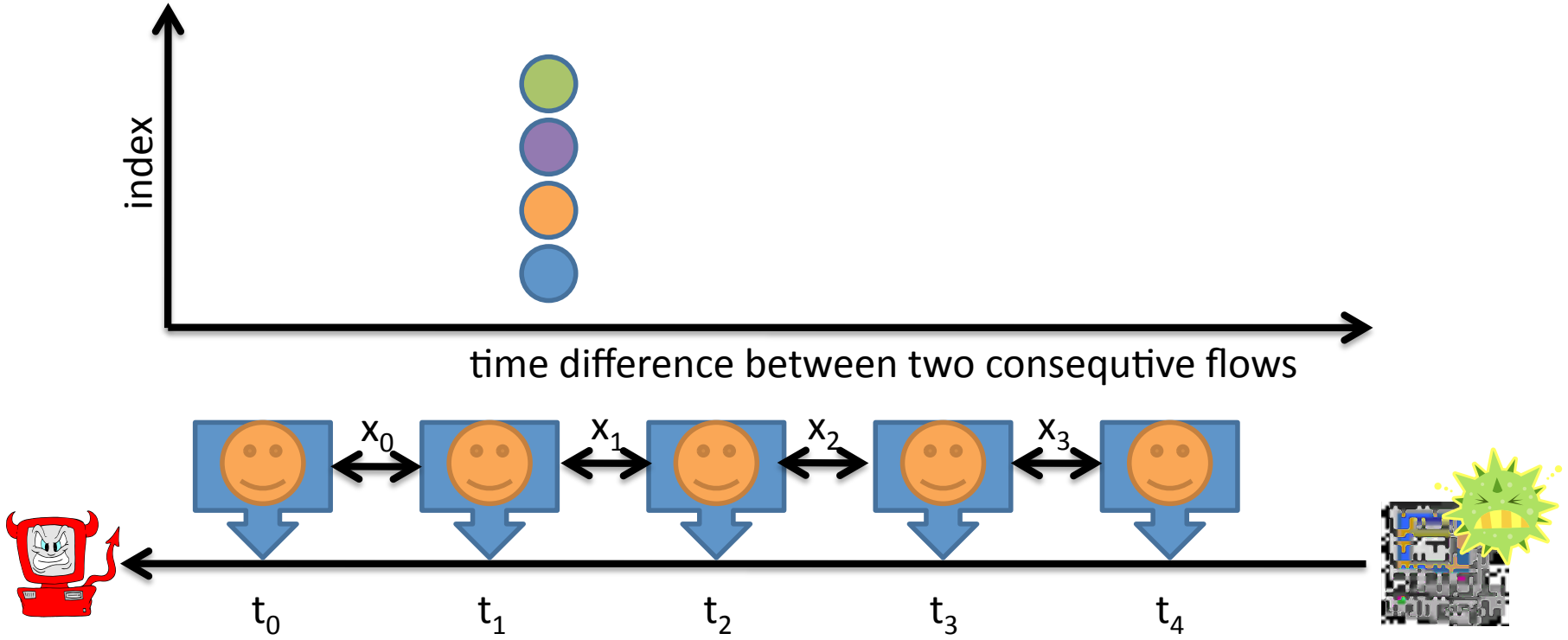


Timing Behavior of Malicious Hosts

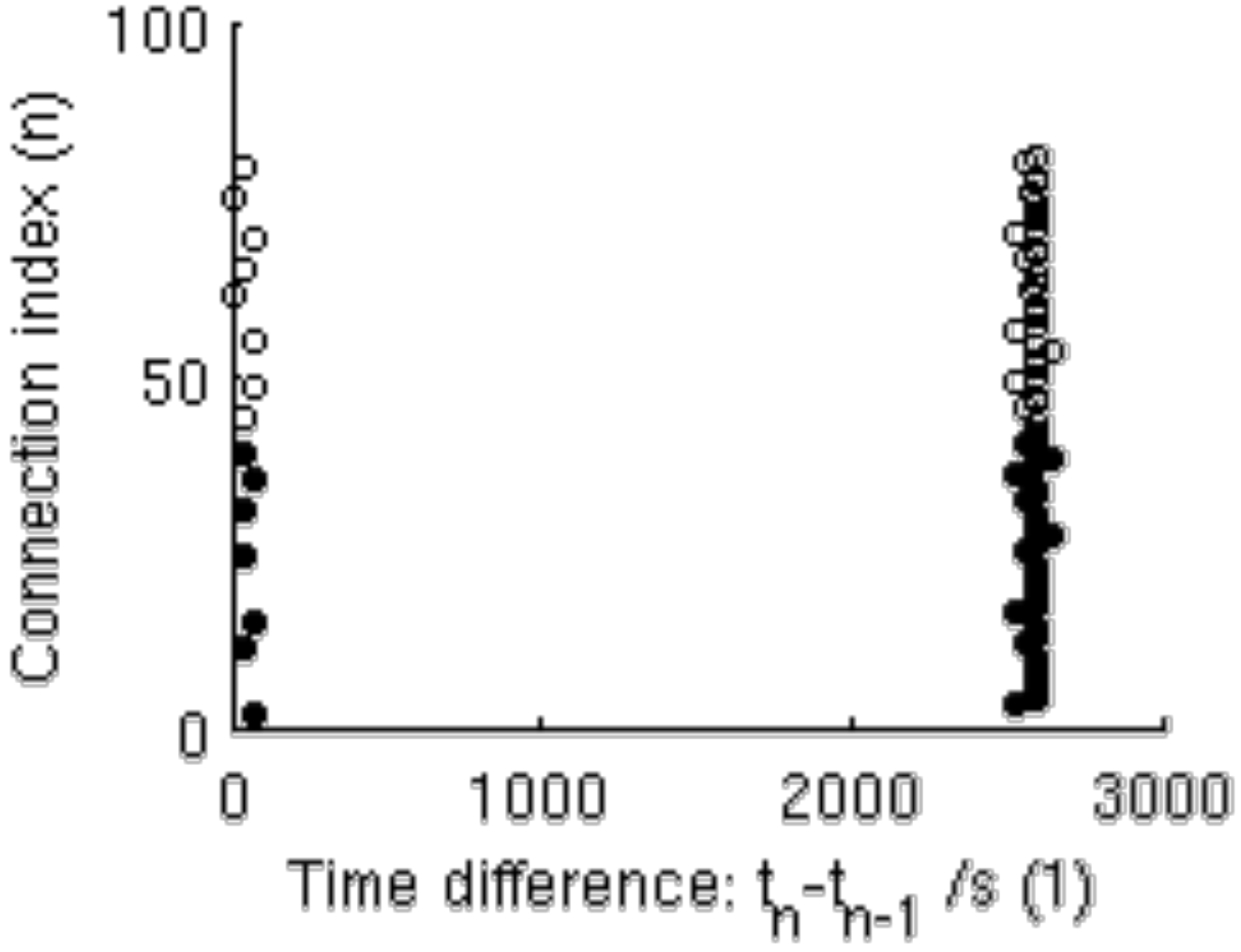


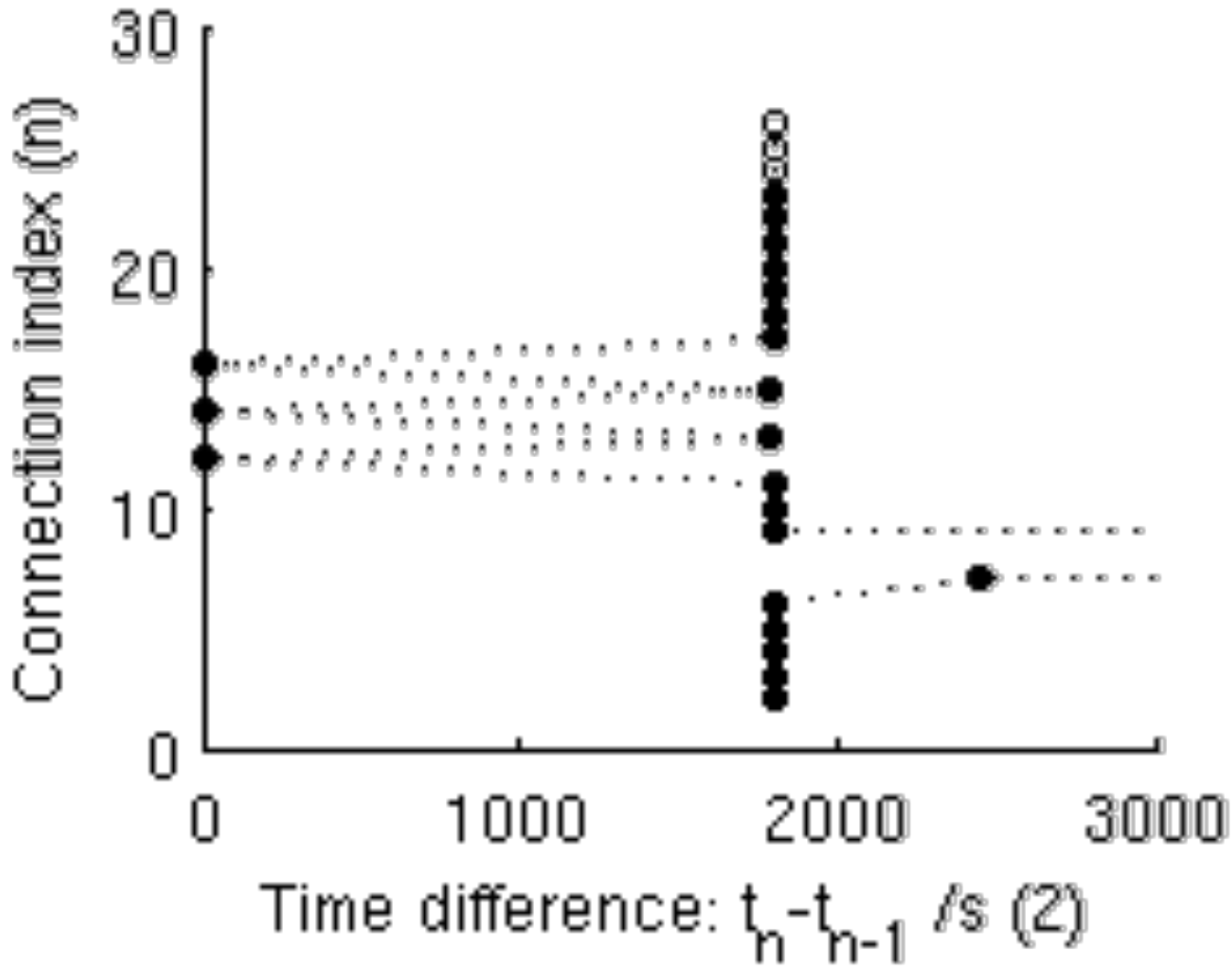
- $x_0 = t_1 - t_0$
- $x_1 = t_2 - t_1$
- $x_2 = t_3 - t_2$
- $x_3 = t_4 - t_3$

Timing Behavior of Malicious Hosts



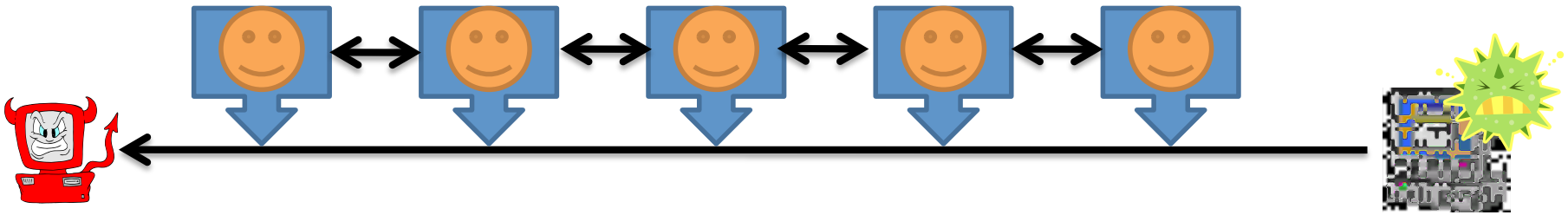
$x_0 = t_1 - t_0$
$x_1 = t_2 - t_1$
$x_2 = t_3 - t_2$
$x_3 = t_4 - t_3$


100
50
0
Connection index (n)
0
1000
2000
3000
Time difference: $t_n - t_{n-1}$ /s (1)

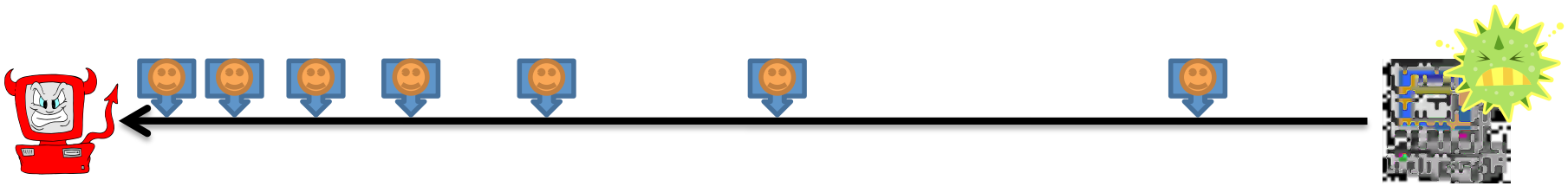

Connection index (n)
30
20
10
0
0
1000
2000
3000
Time difference: $t_n - t_{n-1}$ /s (2)

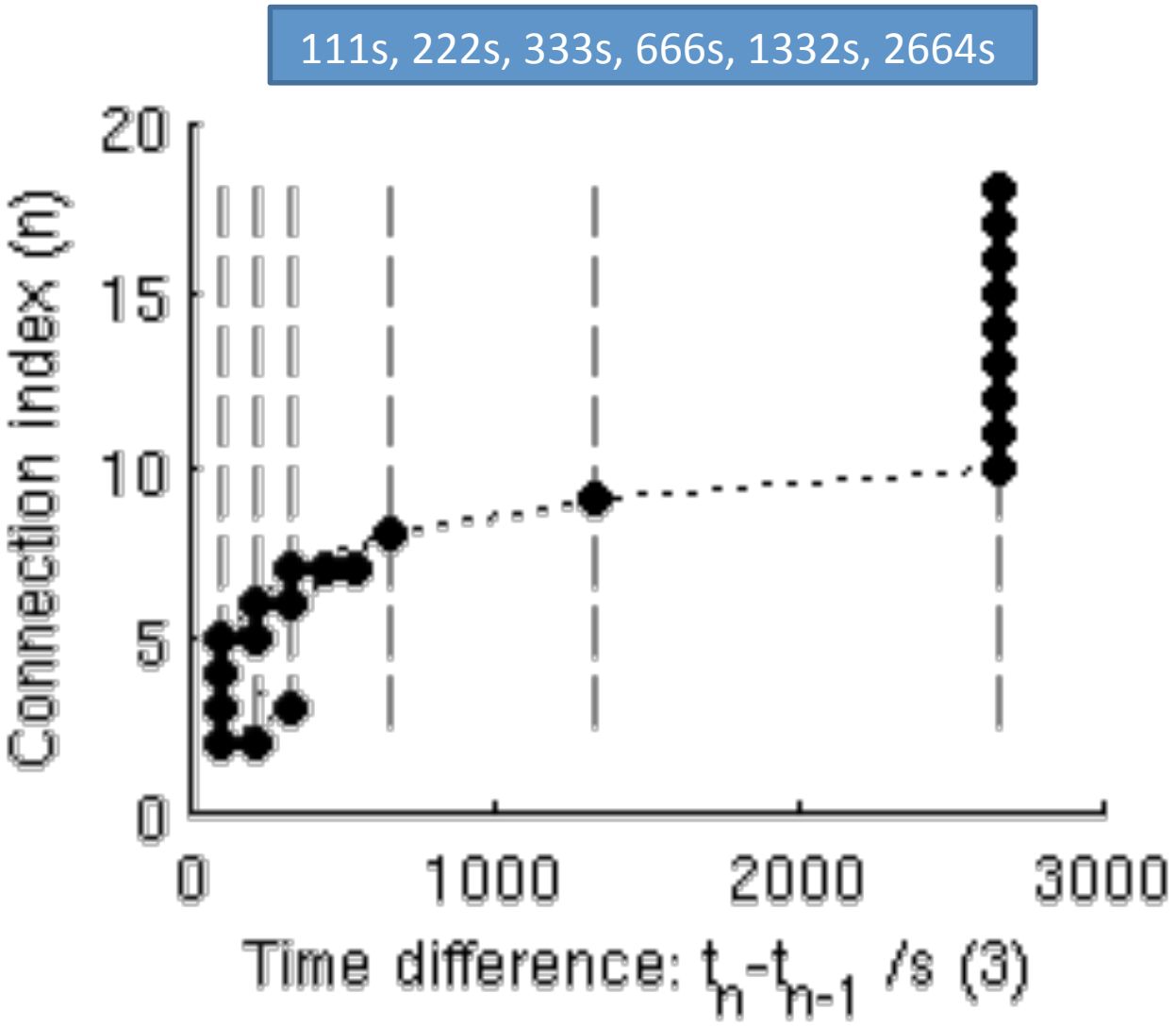
Timing Behavior of Malicious Hosts

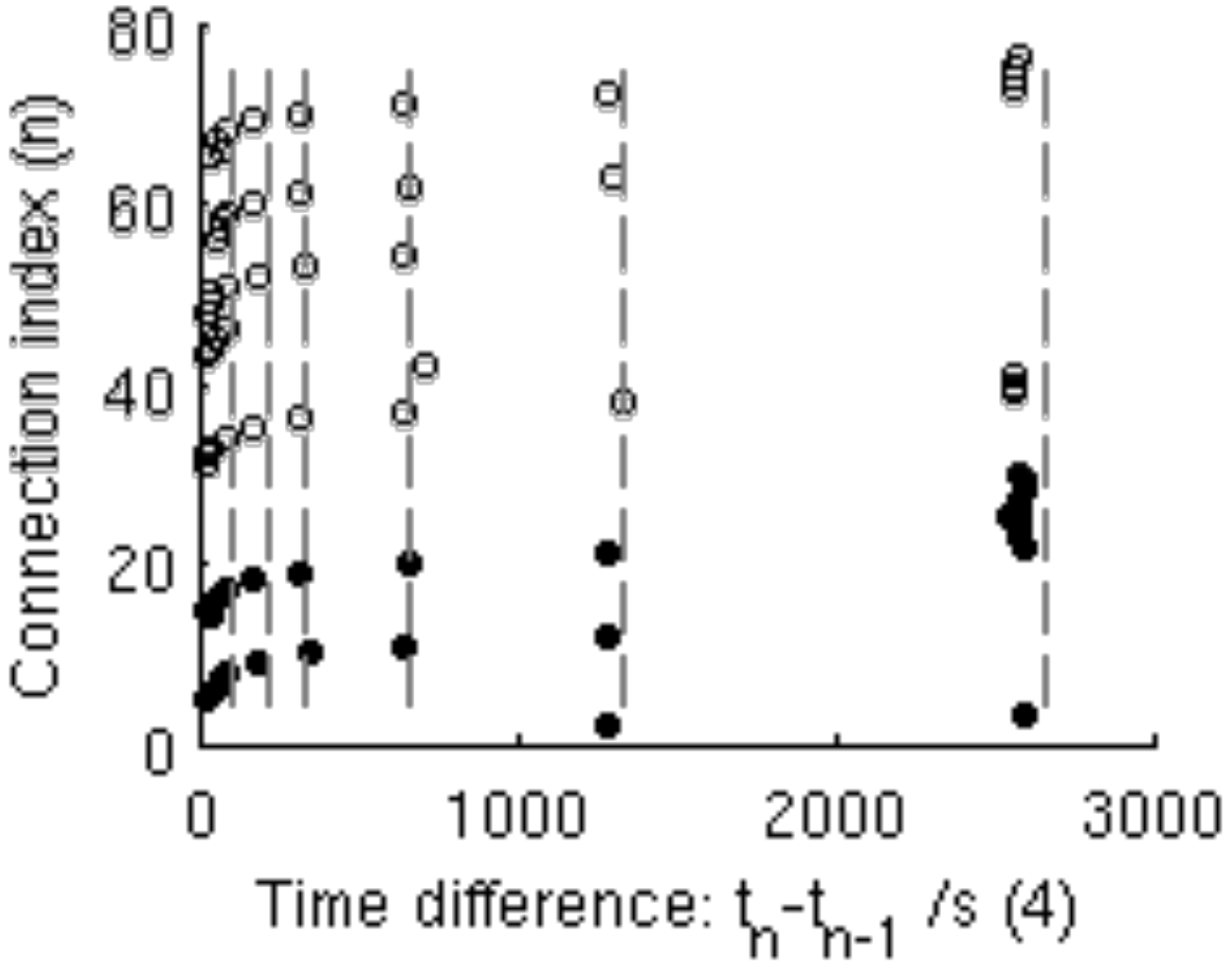
Simple refresh / keep-alive: once every 30 min or 43min



Timing Behavior of Malicious Hosts




111s, 222s, 333s, 666s, 1332s, 2664s
Connection index (n)
20
15
10
5
0
0
1000
2000
3000
Time difference: $t_n - t_{n-1}$ /s (3)


Connection index (n)
80
60
40
20
0
0
1000
2000
3000
Time difference: $t_n - t_{n-1}$ /s (4)

Summary

- We have analyzed malicious flow patterns.
- Our analysis is ***bootstrapped*** with a set of externally defined malicious hosts.
  - It allowed more obscure patterns to be analyzed.
- Some examples from our analysis:
  - blacklisting is sometimes ignored by malware owner (effective anyways)
  - malware host is actively monitored / supervised
  - comm. patterns may last more than 1 collection date.