NATed hosts identification

Phd Student at LSV, ENS-Cachan, CNRS, INRIA

Version 1.0

10 oct 2007

®Elie Bursztein (eb@lsv.ens-cachan.fr)

Plan

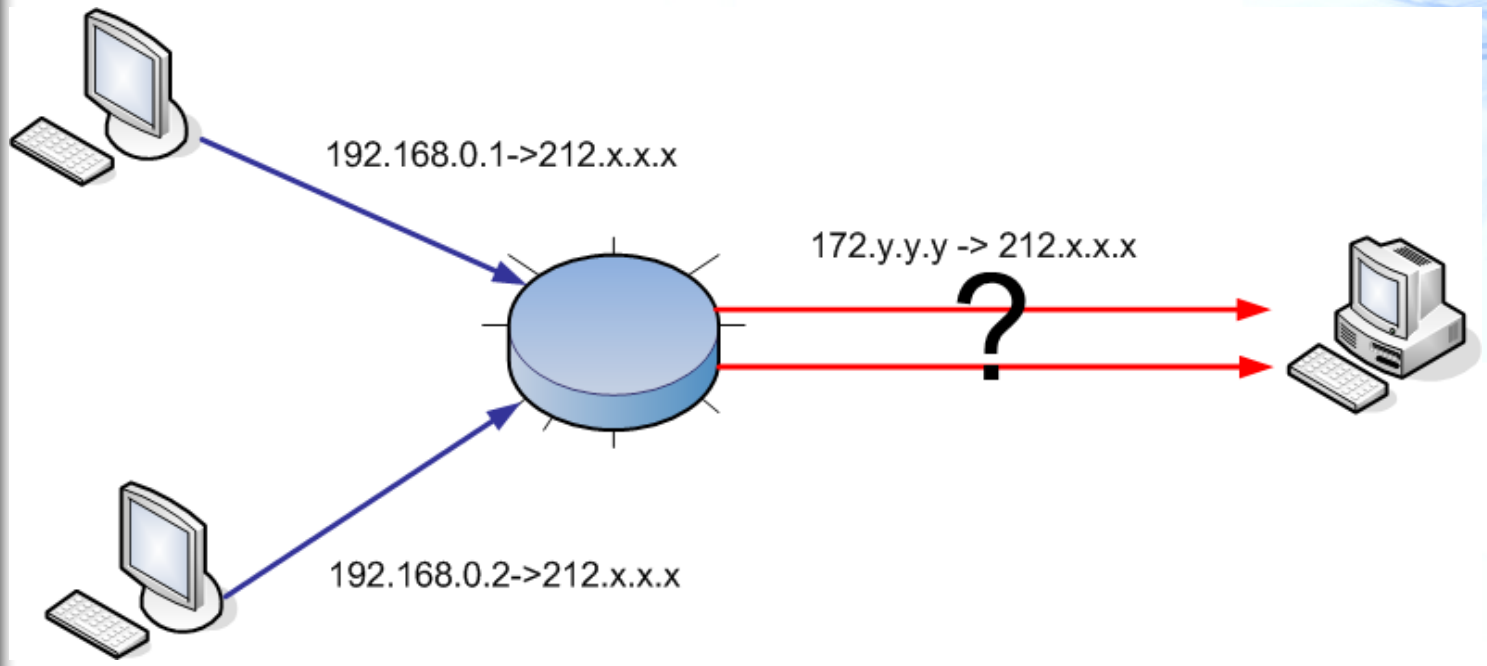
Introduction

- ✓ TCP/IP network is meant to have end to end connection
- ✓ This is assumed in network models and Specification
- ✓ However in real world is not the case any more

®Elie Bursztein (eb@lsv.ens-cachan.fr)

What is NAT

- ✓ Used to Share the same public IP
- ✓ Rewrite all IP address field of internal hosts to the same IP
- ✓ Tracking and Flow analysis are deceived



®Elie Bursztein (eb@lsv.ens-cachan.fr)

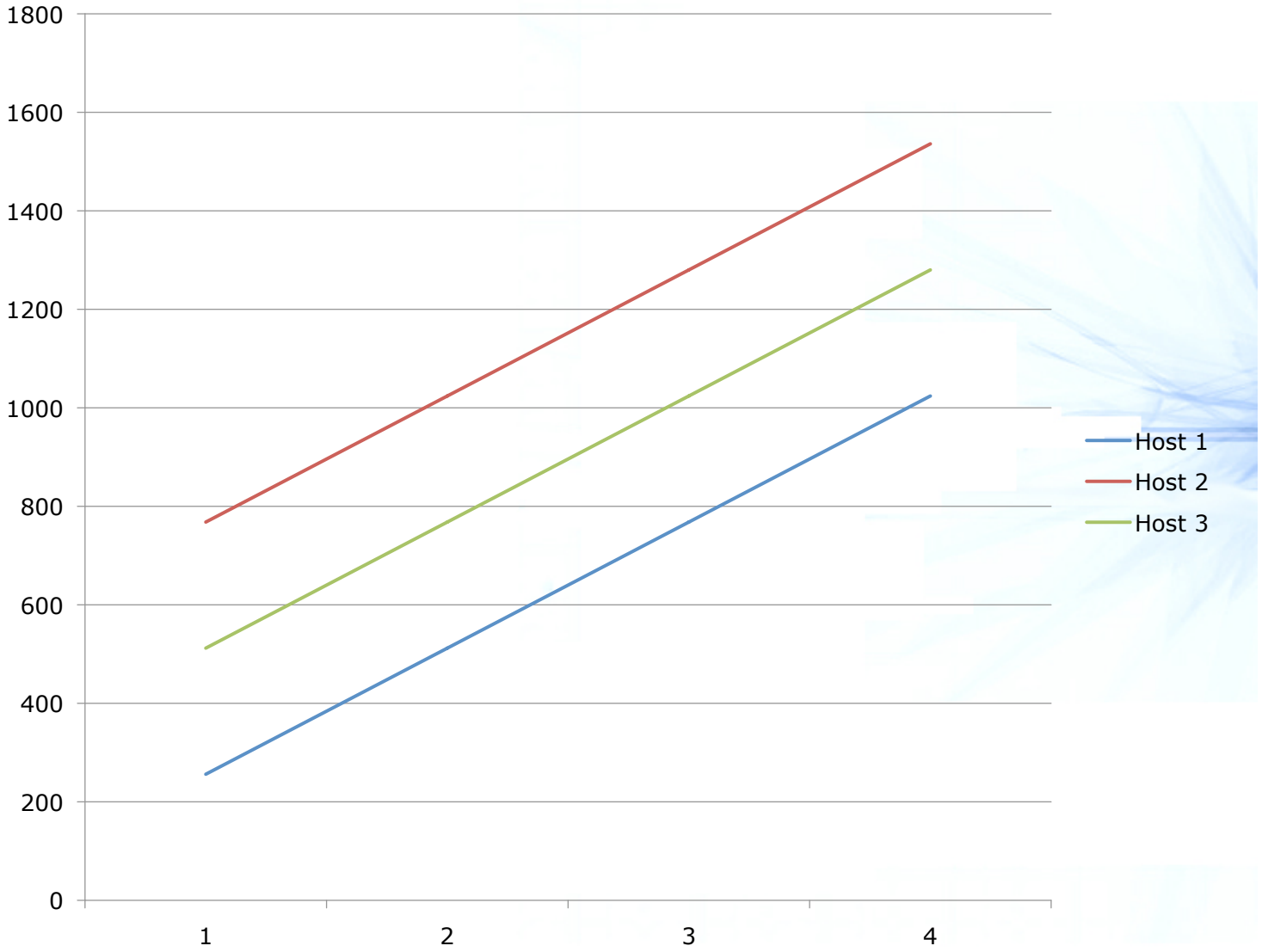
IPID field

- ✓ IPID field is used when fragmentation
- ✓ On some OS (Windows mainly) it is implemented as a simple counter.

®Elie Bursztein (eb@lsv.ens-cachan.fr)

Plan

IPID identification idea



®Elie Bursztein (eb@lsv.ens-cachan.fr)

Drawback of IPID identification

- ✓ IPID warp-around
- ✓ Unix based hosts (Linux, FreeBSD, OSX…) does not implement it at counter
  - ✓ Example : On Linux it is always 0.
- ✓ Hard to use if you see only a part of the traffic

®Elie Bursztein (eb@lsv.ens-cachan.fr)

TCP Timestamp option

- ✓ Added in RFC 1323
- ✓ TCP timestamp
  - ✓ Incremented on time basis not traffic
  - ✓ Increment value is well known
  - ✓ Increment value is OS dependent

®Elie Bursztein (eb@lsv.ens-cachan.fr)

Advantages

- ✓ Work with computed that do not use IPID counter
- ✓ Allows to know the NATed hosts OS family
- ✓ Is not sensitive to computer traffic
- ✓ Work well even if only partial traffic is seen
- ✓ Less sensitive to warp around
  - ✓ Linux warp-around occurs every 248 days, 13:13:56

®Elie Bursztein (eb@lsv.ens-cachan.fr)

Timestamp details

	Cisco	Linux	BSD
Increment/s	1000	100	2

Ts2 – Ts1 / t2 – t1 = slope

Predicted Timestamp = tx – t1 * slope

Note : Initial value is random for Windows .

®Elie Bursztein (eb@lsv.ens-cachan.fr)

PAT host count

- ✓ PAT port address translation
  - ✓ Used for load balancing

®Elie Bursztein (eb@lsv.ens-cachan.fr)

Drawbacks

- ✓ Windows hosts does not use timestamp by default
  - ✓ Use it when requested
  - ✓ Can be activated
- ✓ Network congestion/lag may misled the analysis

®Elie Bursztein (eb@lsv.ens-cachan.fr)

Conclusion

- ✓ Combine well with previous technique
- ✓ Detection is suitable for online analysis
- ✓ Can be improved by adding passive software detection to spot inconsistency

®Elie Bursztein (eb@lsv.ens-cachan.fr)