Passive DNS at OARC

Keith Mitchell, OARC
Paul Vixie, ISC

DNS Operations Meeting
Chicago, 28th Jul 2007

History

- Florian Weimar invented this concept
- Implementation in academia
  - GNU ADA, Berkeley DB
  - Sensors in European ISPs & Universities
  - Original intent: zone content recovery
- Used today by world wide LEO community
- "Inverse directory" & botnet hunting
  - what names map to "this" address?
  - when was "this" name first used and by whom?
  - who has looked up "this" botnet C&C name?

Upcoming Alternatives

- April Lorenzen, sponsored by ISC
  - Implemented in Perl & PostgreSQL on FreeBSD
  - Uses NSF funded hardware (OARC)
  - Text-y, emphasis is on web GUI
- Florian Weimar (redux)
  - Wants to try SQL (vs. Berkeley DB)
  - Strong non-text-y schema
  - Emphasis on performance

Hazards of Decentralization

- Every new passive DNS effort has to solicit sensors (instrumented recursive NS)
- Due to ops+BW costs, no sensor can feed more than one passive DNS collector
- Thus, sensor population is bifurcated
- → Perhaps a central solution is warranted?

Hazards of Commercialization

- Huge datamining opportunity in this dataset
  - Ultimately, collectors could pay the sensors
- There might be some problems, though:
  - National privacy laws
  - ISP privacy policies
  - Competitors getting hold of it

➔ Perhaps a trusted nonprofit could help?

OARC as a Framework

- Trusted nonprofit association
  - Infrastructure for data sharing
  - Administration for resource pooling
  - Governance by membership
  - Oversight by rootops
  - ISC as secretariat
- Perfect vehicle for Passive DNS ?

Proposed Solution

- PCAP-based data capture tool (DNSCAP)
  - can hide “personally identifiable information”
- Lightweight relationship for sensor operators
  - simplified zero fee agreement
  - exchange of security keys
  - upload batches using SSH/SFTP
- Central collector operated for OARC by ISC
  - anonymizes batches, rebroadcasts on a LAN
  - each passive DNS project sits on that LAN

Business Model

- Central collector has a rebroadcast LAN
  - might also offer the raw data on DVD-ROM
  - can consider paying stipends to sensor operators
- Passive DNS projects connect to this LAN
  - nonprofit/research projects pay low fee / no fee
    - required to share collected data without fee/license
  - commercial projects pay for rack, power, & port
    - e.g. AV companies, search engines, etc
  - strong contracts, with enforcement, for ethics
    - No spamming, etc
  - extracted data can be taken offsite
    - ISC will transport but not transit “commercial” data

Business Plan

- Hardware:
  - Sun X4500 storage engine ($35K)
  - HP blade server ($50K)
- Personnel:
  - tools person ($125K)
  - ops person ($125K)
  - half a manager ($75K)
- Recovery:
  - scaled fee structure (based on company size)
  - goal is to fully support Passive DNS as of Y2
  - with some left over for cross-subsidy, new ideas