Open Source Passive DNS Replication

Robert Edmonds (edmonds@isc.org)

October 14, 2012

ISC Passive DNS and ISC DNSDB

- Sensor collects DNS response packets
- Packets parsed into DNS records
- Verification
- De-duplication
- Filtering
- Output sent to DNSDB importer
- DNSDB serves lookup results to clients

Open source components

- All software components released as open source
- Deploy your own passive DNS replication system
- See `http://rsfcode.isc.org/` for git repositories, tarballs, Debian packages
- Split between libraries (nmsg, wdns, mtbl, dnstable), language bindings (pynmsg, pywdns, pymtbl, pydnstable), and applications (nmsgtool, nmsg-dns-cache, nmsg-dns-filter, dnstli)

nmsg: network message encapsulation library

- Define a message “schema”
- Encapsulate data into payloads
- Write payloads to disk
- Send payloads (broadcast UDP, unicast TCP, UNIX socket)
- Built on top of protobuf-c, libxs
- Passive DNS sensor implemented as plugin

wdns: low-level DNS library

- Fast DNS message parsing library
- Decompose messages into sections, RRs/RRsets
- For Python users, pywdns wrapper

mtbl: immutable sorted string table library

- Stand-alone “Sorted String Table” (“SSTable”) implementation
- Also includes interfaces for sorting and merging large amounts of data
- SSTable implementation closely based on open source Google C++ code
- Other implementations in Google LevelDB, Apache Cassandra, Apache Hadoop – but internal, part of larger system

dnstable: encoding format, library, and utilities for passive DNS data

- Compact, custom serialization format tailored for passive DNS
- Wildcard searches, inverse (rdata) searches, etc.
- Built on top of `libmtbl`, `libnmsg`, `libwdns`
- This is used to power the DNSDB service
- See `dnstable-encoding(5)` manpage for details of key/value serialization format

nmsg-dns-cache: de-duplication utility

- Uses `libnmsg` to get a stream of raw DNS response messages
- Parses each message using `libwdns` to get a stream of DNS RRsets
- Builds a fixed size FIFO cache to de-duplicate the RRsets
- Passively reconstructs the DNS zone hierarchy using `NS` / `A` / `AAAA` records in order to reject out-of-bailiwick records
- Sends output stream via `libnmsg`

nmsg-dns-filter: filtering utility

- Splits out records we don’t want to keep
- (Lots of noise, don’t need to keep everything)
- Exact matches, subdomain matches, regex matches
- Reloads filter lists on the fly

dnstli: dnstable lookup interface

- Python WSGI webapp, runs behind web server
- Provides lookup service over HTTP for a set of dnstable data files
- Authenticate users with username/password or API key
- Powers `https://dnsdb.isc.org/` and `https://dnsdb-api.isc.org/`

dnstui: dnstable user interface

- Web client
- Runs in browser
- Displays results from `dnstli`

isc-dnsdb-query: dnstable lookup client

- Python and curl examples for fetching results from `dnstli` via HTTP

References

- Passive DNS Replication (Weimer; 2005)
- Passive Monitoring of DNS Anomalies (Zdrnja, Brownlee, Wessels; 2007)
- ISC Passive DNS Architecture (Edmonds; 2012)