A Path Transparency Observatory

Brian Trammell and Mirja Kühlewind
Networked Systems Group, ETH Zürich
EC H2020 Measurement and Architecture for a
Middleboxed Internet (MAMI)

Path Transparency

- **Path transparency**: the likelihood a packet stream that arrives at the end of the path is the one that was sent with certain properties.

- **Impairment**: something that keeps a path from being transparent for a certain kind of traffic, dependent on that traffic's properties.

  - Blocked connections, 100% packet loss

  - Increased drop rate, increased latency

  - "Bleaching": removal or rewrite of necessary headers

  - B2B proxying: replacing one e2e path with two

Observatory Requirements

- Accept data from a wide variety of sources, e.g.:

  - Raw output from tools we maintain.

  - Raw packet traces of active measurements.

  - "Here's data from a measurement study, and references to commits in GitHub for the tools and configuration used to generate it."

  - "Option foo breaks on these paths but not these paths; we're not going to tell you where they are, but this set of them belong to a major mobile carrier."

- Support path pseudonymization and aggregation for privacy.

- Support condition definition with enough precision to allow active measurements to reproduce observations on other paths.

- Integrate with existing tools, without restrictions on implementation.

Observatory Definition
(work in progress)

- An observatory collects single observations {*t, P, c*} where

  - *t*: time at which the observation was taken (and assumed valid)

  - *P*: variable-precision designator of the path on which the observation was taken

    - sequence of node/network/multi-network identifiers or pseudonyms

  - *c*: variable-definition expression of the condition observed on that path

    - Ideal: generic language with primitives for common packet structures, without regex performance/expressivity issues.

    - Initial: reference to external condition identifier linked to how it was generated (stable code and configuration reference).

- Using path and condition characterization and equivalence operations, allows more formal comparison of diverse studies.

  - Initial approach for *c* requires expert advice on condition equivalence.

Backup

Active Measurement of Path Transparency

- Basic methodology: throw a bunch of packets with certain properties at the Internet, and see what happens.

  - Ideal: two-ended A/B testing

  - More scalable: one-ended A/B testing

  - Comparison with topology to isolate on-path vs near-endpoint impairments

- Observations from platform- and application-level logs of failed attempts to use protocol features also useful.

- Integrate heterogeneous observations from many campaigns for better insight.