Path Transparency Observatory
Demo of the Web User Interface

Stephan Neuhaus and Mirja Kühlewind

Brussels, Oct 21, 2016

mami

measurement and architecture for a middleboxed internet

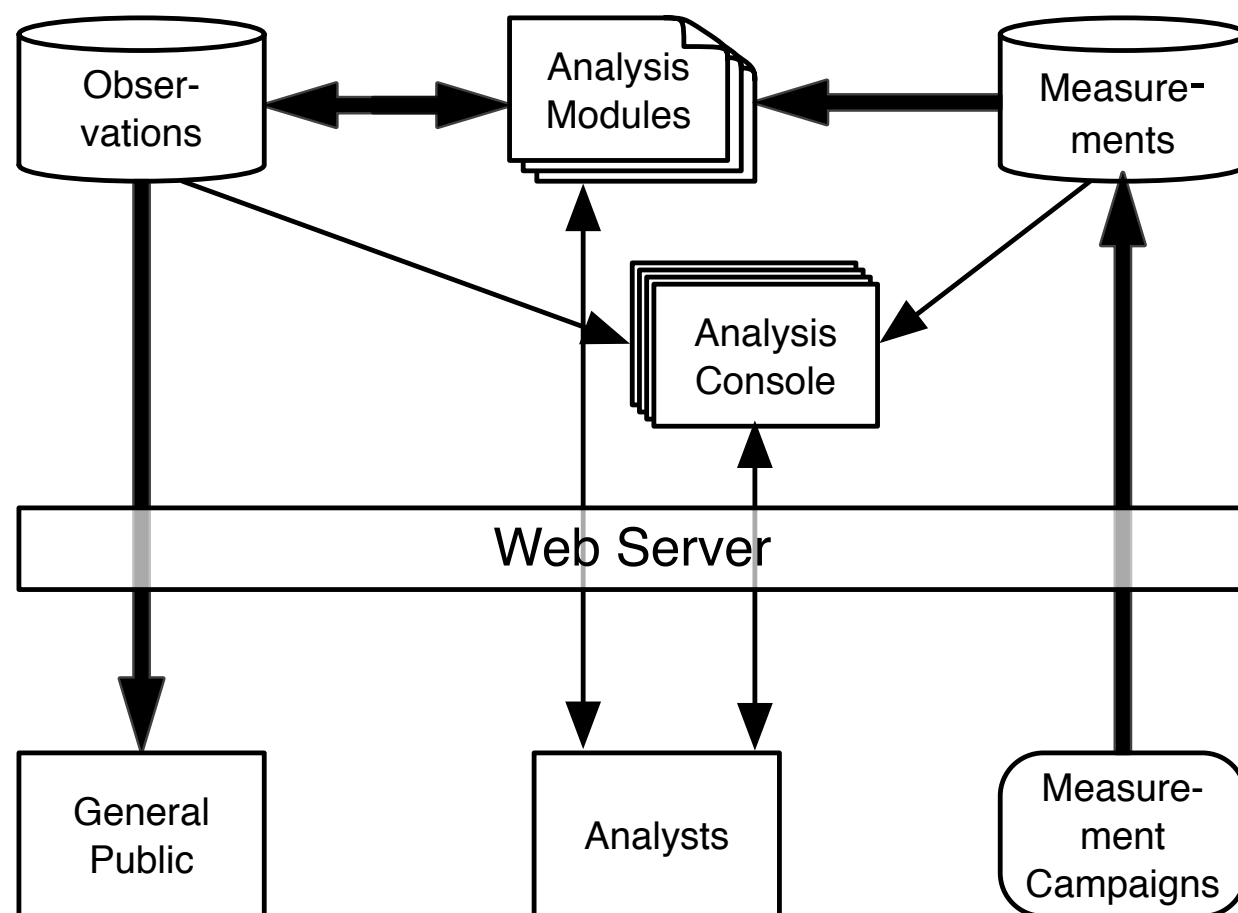measurement architecture experimentation

Path Transparency Observatory - PTO
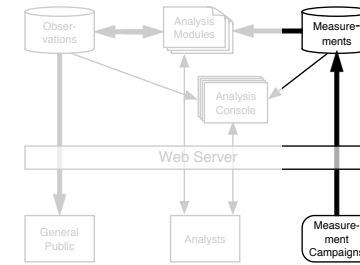
- Store

  - raw data and

  - computed conditions

- that relate to path transparency, enabling

  - exploration and

  - analysis

- with a special focus on

  - easing replication and

  - providing insights to the general public

PTO - Architecture

- **Raw measurement storage**
  - manual or automated upload per measurement campaign

- **Analyzers**
  - generate *observations* from measurements
  - Jupyterhub for interactive analysis and analysis development

- **Web UI for general public**
  - access to selected observations from vetted analysis modules
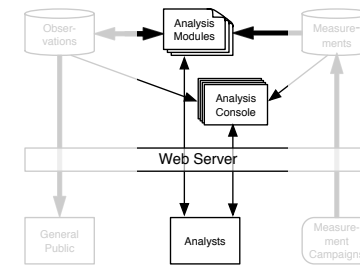
PTO - Measurements

- Raw data stored from measurement campaigns

- Data format most appropriate to the tool/campaign

  - IPFIX, tracebox, json, mPlane results …

- Stored in HDFS for easy clustering

  - metadata stored in MongoDB

PTO - Observations
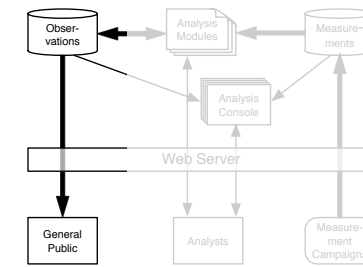
- **Path Transparency Observation {P, t, c, v}**
  - *path designator* **P:** sequence of identifiers for path along which the observation was taken (initiator, target, and other point on along path between them, as e.g. network-layer addresses, prefixes, or AS numbers)
  - *temporal scope* **t:** when a observation was taken and is considered valid
  - *condition* **c:** condition observed along the path (e.g. whether ECN causes connectivity failure, whether a TFO cookie request succeeds); inherently boolean, but may contain associated values **v**.

- Conditions have symbolic names, arranged in hierarchy, e.g.,
  - `ecn.connectivity.works`
  - `ecn.negotiation_attempt.succeeded`
- Stored in MongoDB for ease of clustering and querying

PTO - Analyzers

- Analyzers produce Observations

  - based on measurements in the raw data or other Observations

- Observations contain the name of the analyzer…

  - …and the commit identifier of the exact analyzer version

- This enables reproducibility and also easier filtering:

  - "Give me all observations from analyser x, excluding (buggy) commit 1a2b3c4d"

- Analysis console is Jupyterhub, enables easy exploratory analysis of measurements and observations

Automatic Analysis and Reanalysis

- Research is inherently exploratory, but
  requires immutable raw data for repeatability.

- Observations contain all information required to regenerate them.

- Sensors trigger on new data uploads, rerun analysis on data replacement.

- Observations marked deprecated due to invalidated raw data remain available.
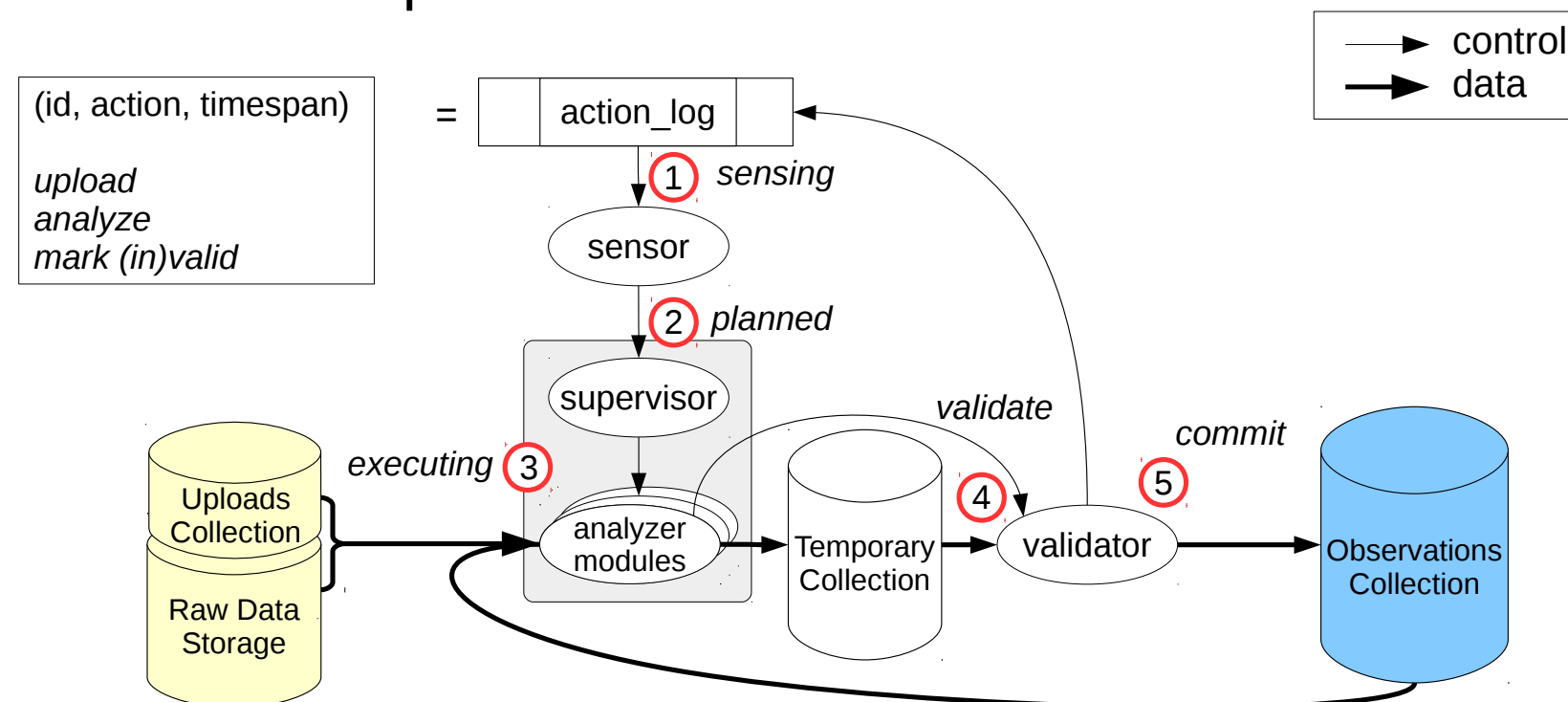
PTO - State of Development

- PTO backend infrastructure up since *~March*

- PTO measurement upload software since *~April*

- PTO condition-querying REST service since *~May*

- PTO query API working since *~June*

  - …but being constantly optimised

- PTO Web UI in active development since *~July*

  - …now in Beta testing inside MAMI

  - …about to go live in December

Example measurement study (1): ECN connectivity testing

- **Explicit Congestion Notification (ECN):** TCP extension to signal congestion before dropping packets

  - Negotiation between both endpoints and network support needed

- **ECNspider:** A/B testing using PATHspider for ECN connectivity testing:
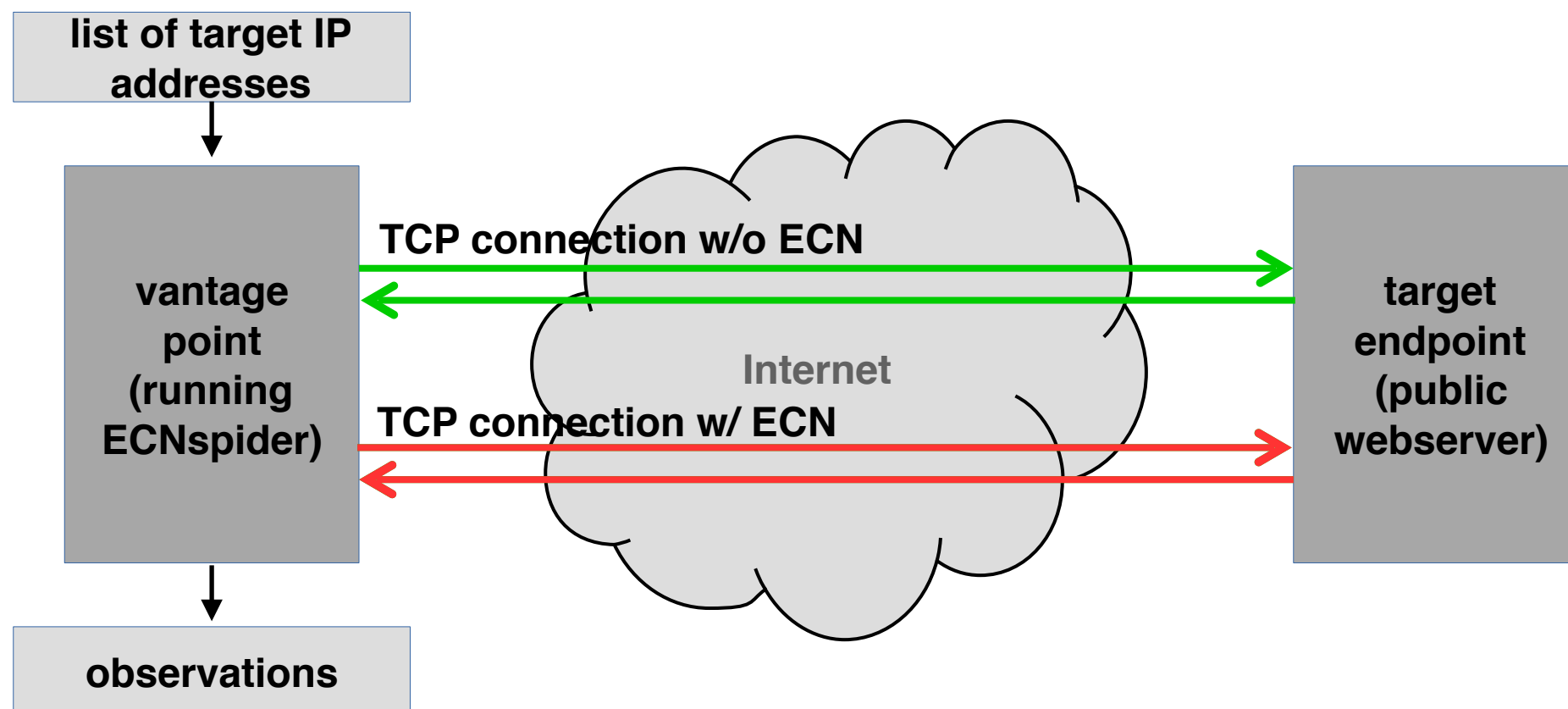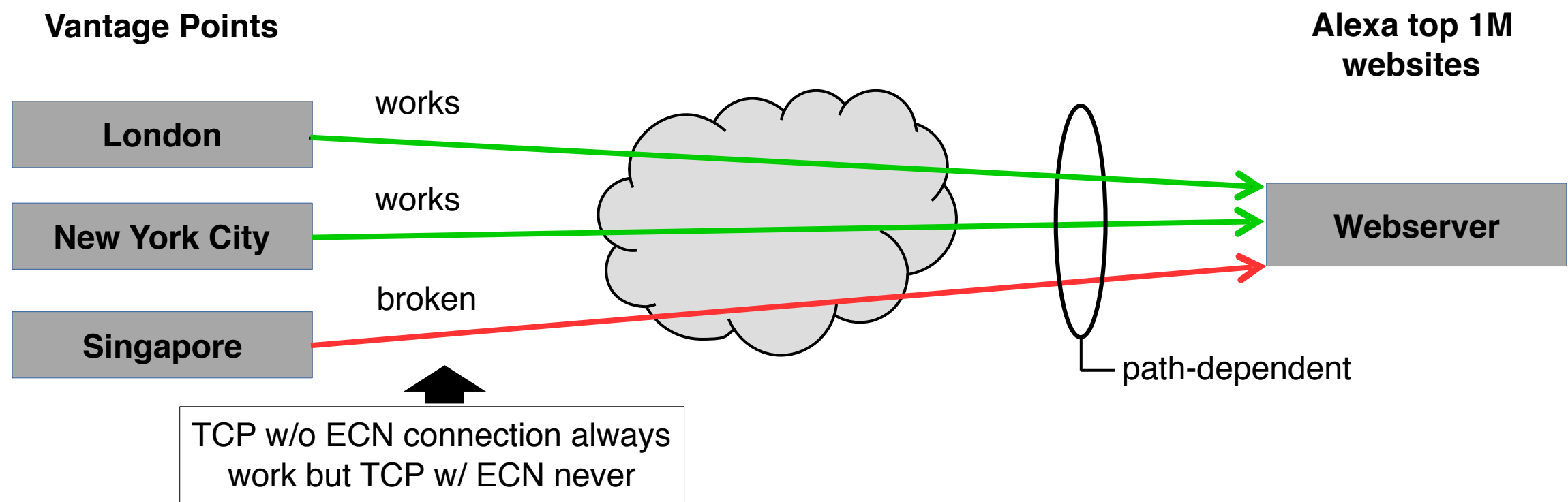
Example measurement study (1): ECN observations

- `ECN.connectivity.{works, broken, transient, offline}`

  - **works:** all TCP w/o ECN connections and all TCP connections w/ ECN worked

  - **broken:** all TCP w/o ECN connections and no TCP connections w/ ECN worked

  - **transient:** TCP connections w/ ECN did not work

  - **offline:** no TCP connections could be established at all

- `ECN.negotiation_attempt.{succeeded, failed}`

  - **succeeded:** ECN was successfully negotiated

  - **failed:** target host did not support ECN

Example measurement study (3): ECN path transparency

- Measuring path-dependency based on different vantage points

Vantage Points

Alexa top 1M websites

| London |
works

| New York City |
works

| Singapore |
broken

TCP w/o ECN connection always work but TCP w/ ECN never

Webserver

path-dependent

Example measurement study (4): ECN path transparency analyses

Additional analysis tool in PTO (analyzer-ecnspider-vp) compares observations from different vantage points and derived a path-dependent observation:

- **`ecn.path_dependent`**: from some vantage points `ecn.connectivity.works` and some `ecn.connectivity.broken`

- **`ecn.site_dependent`**: from all vantage points `ecn.connectivity.broken`
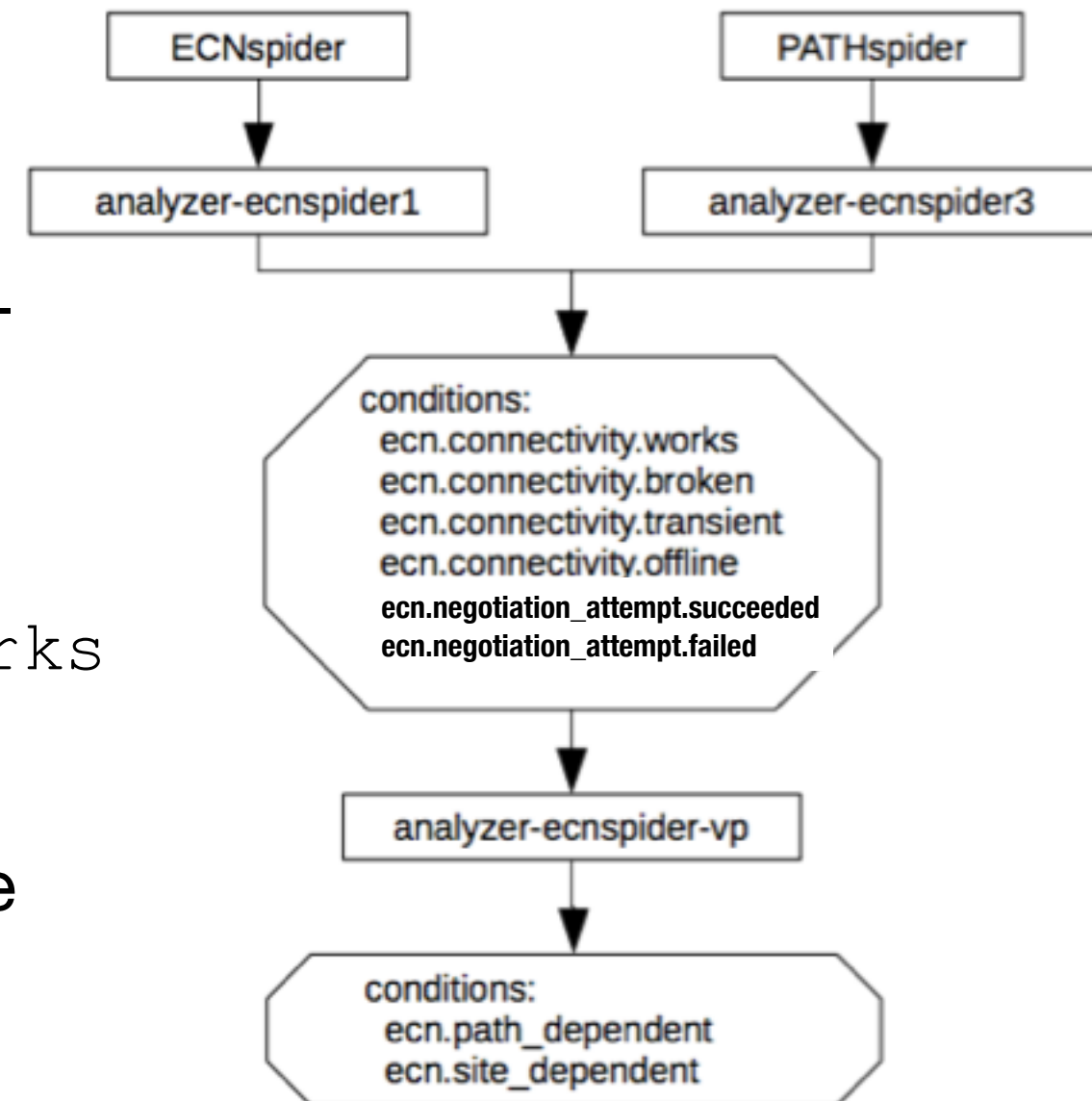
PTO Web UI: Query Parameters

PTO Web UI

- User selects criteria in PTO UI:

  - Date and time

  - Path

  - Condition(s)

- Submits query to PTO

- PTO queries mongo database according to submission

- PTO returns JSON to PTO UI

- PTO UI renders JSON

PTO Web UI: Rendering Results (1)

12%

88%

ecn.connectivity.broken 3622
ecn.connectivity.offline 474
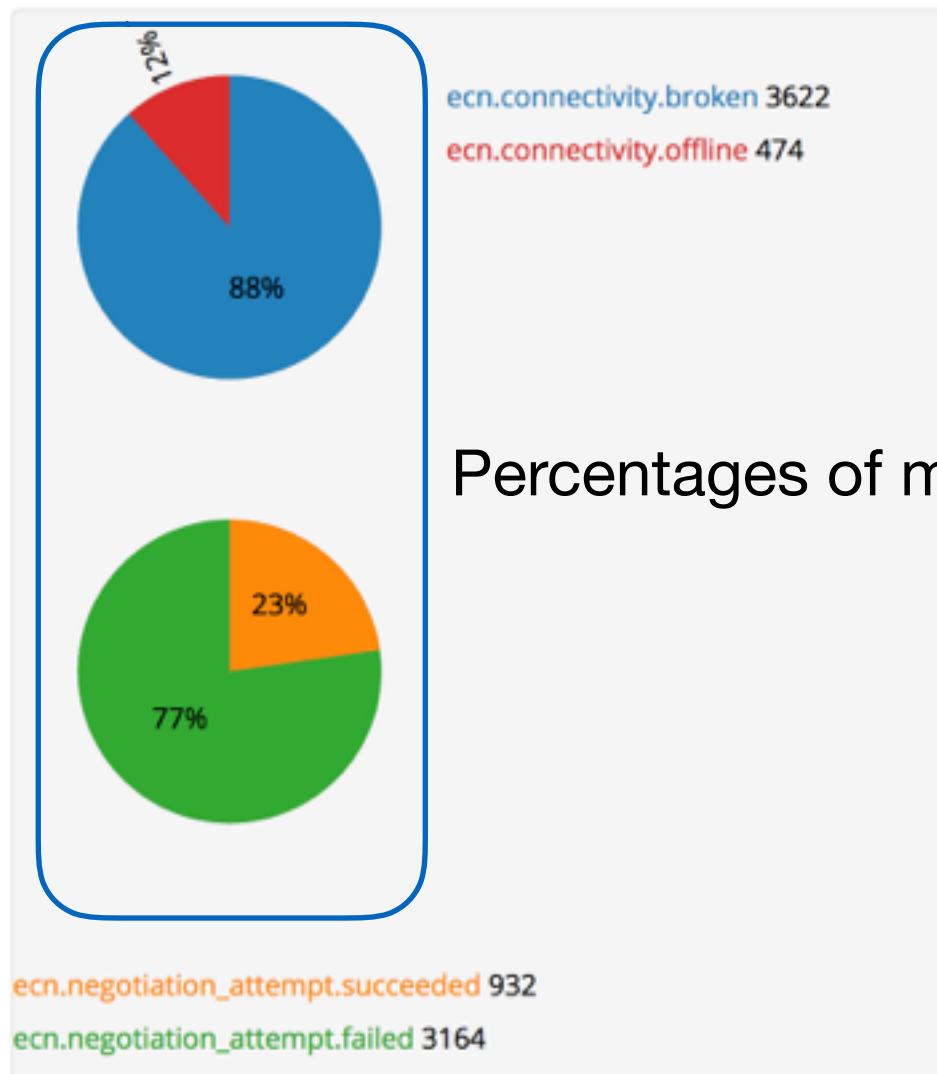
Path: - * -
Observations: 4096
From: Sun, 12 Jun 2016 16:49:59 GMT
To: Wed, 10 Aug 2016 19:09:41 GMT

Percentages of mutually exclusive conditions

23%

77%

ecn.negotiation_attempt.succeeded 932
ecn.negotiation_attempt.failed 3164

Clickable measurement
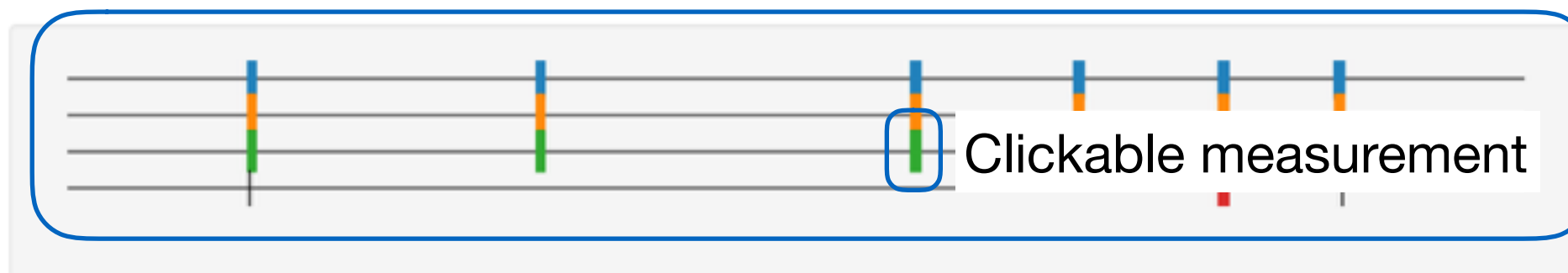
Timeline

PTO Web UI: Rendering Results (2)

Analyzer	Path IP Src	Path IP Dst	Conditions
analyzer-ecnspider1	188.166.3.245	85.26.148.161	ecn.connectivity.broken, ecn.negotiation_attempt.failed
analyzer-ecnspider1	188.166.3.245	119.10.36.27	ecn.connectivity.broken, ecn.negotiation_attempt.failed
analyzer-ecnspider1	188.166.3.245	114.80.207.205	ecn.connectivity.broken, ecn.negotiation_attempt.failed
analyzer-ecnspider1	188.166.3.245	14.63.164.41	ecn.connectivity.broken, ecn.negotiation_attempt.failed
analyzer-ecnspider1	188.166.3.245	149.202.112.33	ecn.connectivity.broken, ecn.negotiation_attempt.succeeded
analyzer-ecnspider1	188.166.3.245	103.35.151.41	ecn.connectivity.broken, ecn.negotiation_attempt.failed
analyzer-ecnspider1	188.166.3.245	204.141.11.38	ecn.connectivity.broken, ecn.negotiation_attempt.failed
analyzer-ecnspider1	188.166.3.245	69.87.205.26	ecn.connectivity.broken, ecn.negotiation_attempt.succeeded
analyzer-ecnspider1	188.166.3.245	192.187.111.202	ecn.connectivity.broken, ecn.negotiation_attempt.succeeded
analyzer-ecnspider1	188.166.3.245	63.141.224.83	ecn.connectivity.broken, ecn.negotiation_attempt.succeeded
analyzer-ecnspider1	188.166.3.245	61.91.120.146	ecn.connectivity.broken, ecn.negotiation_attempt.failed
analyzer-ecnspider1	188.166.3.245	185.22.234.105	ecn.connectivity.broken, ecn.negotiation_attempt.succeeded
analyzer-ecnspider1	188.166.3.245	119.254.65.178	ecn.connectivity.broken, ecn.negotiation_attempt.failed
analyzer-ecnspider1	188.166.3.245	219.163.33.35	ecn.connectivity.broken, ecn.negotiation_attempt.failed
analyzer-ecnspider1	188.166.3.245	14.49.24.71	ecn.connectivity.broken, ecn.negotiation_attempt.failed
analyzer-ecnspider1	188.166.3.245	208.110.70.75	ecn.connectivity.broken, ecn.negotiation_attempt.failed

Clickable row

Table View

PTO Web UI: Rendering Results (3)

Observation Details 5790f2a3fed8f54132e425ab ✖

JSON Raw

{"sources":{"upl":[0]},"_id":{"$oid":"5790f2a3fed8f54132e425ab"},"action_ids

Properties

Time From	Sun, 12 Jun 2016 16:49:59 GMT
Time To	Sun, 12 Jun 2016 16:50:04 GMT
Analyzer	analyzer-ecnspider1
Value	(empty)
Sources	
Action IDs	67:true
Conditions	ecn.connectivity.broken, ecn.negotiation_attempt.failed
Path	188.166.3.245, *, 85.26.148.161

Uploads

Measurement Campaign	Start Time	Stop Time	Sequence	Format
ecn-june16	Sun, 12 Jun 2016 00:00:00 GMT	Mon, 13 Jun 2016 00:00:00 GMT	0000	ecnspider1-zip-csv-ipfix

PTO on the web

- Path Transparency Observatory web UI:

  - https://observatory.mami-project.eu/#/observatory

- All sources on GitHub!

  - data upload: https://github.com/mami-project/observatory-upload

  - core: https://github.com/mami-project/pto-core

  - web interface: https://github.com/mami-project/pto-web