

# Paris Traceroute Data of M-Lab

[yachang@google.com](mailto:yachang@google.com), Google Inc.

# Paris Traceroute: History

Originally Proposed by

Brice Augustin, Xavier Cuvellier, Benjamin Orgogozo, Fabien Viger, Timur Friedman, Matthieu Latapy, Clémence Magnien, and Renata Teixeira, "Avoiding traceroute anomalies with Paris traceroute", in *Proc. Internet Measurement Conference*, October 2006



# Raw PT data on GCS

For Data Before 2017-06-30:

- <https://console.developers.google.com/storage/browser/m-lab/>

For Data After 2017-06-30:

- <https://console.developers.google.com/storage/browser/archive-mlab-oti/>



# Sample Raw Test

<https://github.com/m-lab/etl/blob/integration/parser/testdata/PT/20171208T00:00:04Z-35.188.101.1-40784-173.205.3.38-9090.paris>



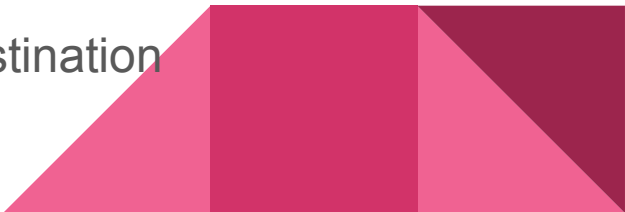
# PT data on BigQuery

- Parsing from raw data to BQ schema data is 100% open source now!

<https://github.com/m-lab/etl>

- Each hop is a separate row with hop src and destination in BQ
- Bundle of PT and NDT are made easier through Rollins which was launched from June 2016

Five-tuple test-id with triggering NDT server IP.

- For many metro >50% PT test did not reach final destination
  - Reconstruct of tree of hops is not obvious
- 

# ETL pipeline monitoring

<https://grafana.mlab-oti.measurementlab.net/d/7qq7W6Hmk/pipeline-pt?orgId=1&from=now-24h&to=now>



# Current schema

<https://bigquery.cloud.google.com/table/measurement-lab:public.traceroute?pli=1&tab=schema>

<https://bigquery.cloud.google.com/project/measurement-lab>

test_id	STRING
project	INTEGER
log_time	TIMESTAMP
type	INTEGER
connection_spec	RECORD
connection_spec.client_af	INTEGER
connection_spec.client_application	STRING
connection_spec.client_browser	STRING
connection_spec.client_hostname	STRING
connection_spec.client_ip	STRING
connection_spec.client_kernel_version	STRING
connection_spec.client_os	STRING
connection_spec.client_version	STRING
connection_spec.data_direction	INTEGER
connection_spec.server_af	INTEGER
connection_spec.server_hostname	STRING
connection_spec.server_ip	STRING
connection_spec.server_kernel_version	STRING
connection_spec.client_geolocation	RECORD
connection_spec.client_geolocation.area_code	INTEGER
connection_spec.client_geolocation.city	STRING

paris_traceroute_hop.protocol	
paris_traceroute_hop.src_ip	STRING
paris_traceroute_hop.src_af	INTEGER
paris_traceroute_hop.src_hostname	STRING
paris_traceroute_hop.src_geolocation	RECORD
paris_traceroute_hop.src_geolocation.area_code	INTEGER
paris_traceroute_hop.src_geolocation.city	STRING
paris_traceroute_hop.src_geolocation.continent_code	STRING
paris_traceroute_hop.src_geolocation.country_code	STRING
paris_traceroute_hop.src_geolocation.country_code3	STRING
paris_traceroute_hop.src_geolocation.country_name	STRING
paris_traceroute_hop.src_geolocation.latitude	FLOAT
paris_traceroute_hop.src_geolocation.longitude	FLOAT
paris_traceroute_hop.src_geolocation.metro_code	INTEGER
paris_traceroute_hop.src_geolocation.postal_code	STRING
paris_traceroute_hop.src_geolocation.region	STRING
paris_traceroute_hop.dest_ip	STRING
paris_traceroute_hop.dest_af	INTEGER
paris_traceroute_hop.dest_hostname	STRING
paris_traceroute_hop.dest_geolocation	RECORD
paris_traceroute_hop.dest_geolocation.area_code	INTEGER
paris_traceroute_hop.dest_geolocation.city	STRING

# Sample BQ search for PT

[https://bigquery.cloud.google.com/table/measurement-lab:base\\_tables.traceroute?pli=1](https://bigquery.cloud.google.com/table/measurement-lab:base_tables.traceroute?pli=1)





# PT data used by researchers

- Track congestion:

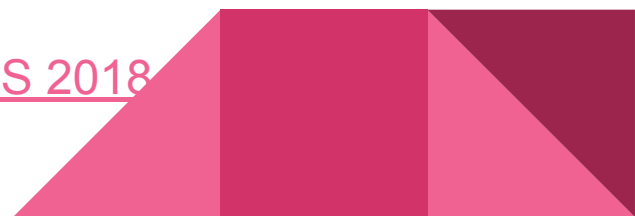
[Measurement and Analysis of Internet Congestion \(MANIC\)](#)

Amogh Dhamdhere

CAIDA, San Diego Supercomputer Center, University of California San Diego,  
2018

- Build mapping of AS

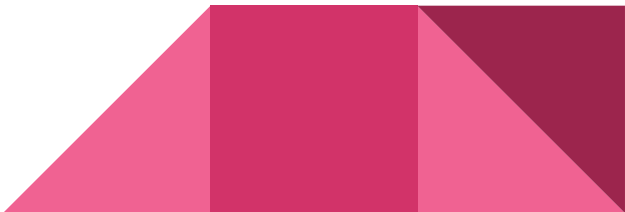
[Exhaustive Mapping of an Autonomous System AIMS 2018](#)



# Scamper: new binary for PT

- Originally Proposed in 2010

Scamper: a scalable and extensible packet prober for active measurement of the internet, by Matthew Luckie

- Open source: <https://www.caida.org/tools/measurement/scamper/>
  - Match the original Paris Traceroute implementation (tracelb)
  - Provide parallel probes and all kinds of options.
  - Json output with schema built-in
- 

# Sample Json output

One hop:

```
{"addr":"216.239.54.127", "q_ttl":1, "linkc":1, "links":[[{"addr":"172.217.7.138",  
"probes":[{"tx":{"sec":1533726748, "usec":648900}, "replyc":1, "ttl":8, "attempt":0,  
"flowid":1, "replies":[{"rx":{"sec":1533726748, "usec":650590}, "ttl":56, "rtt":1.690,  
"ipid":0, "icmp_type":0, "icmp_code":0, "icmp_q_tos":0}}]}]}]}  
{"type":"cycle-stop", "list_name":"default", "id":0,
```

The whole test:

```
"hostname":"mlab4.iad1t.measurement-lab.org", "stop_time":1533726748}
```

# Scamper on MLab v2.0

- Kubernetes cluster
- Master & pods
- Dockerize (<https://github.com/npad/sidestream/blob/scamper/Dockerfile>)
- Easy maintenance and monitoring

DEMO!

