MLAB

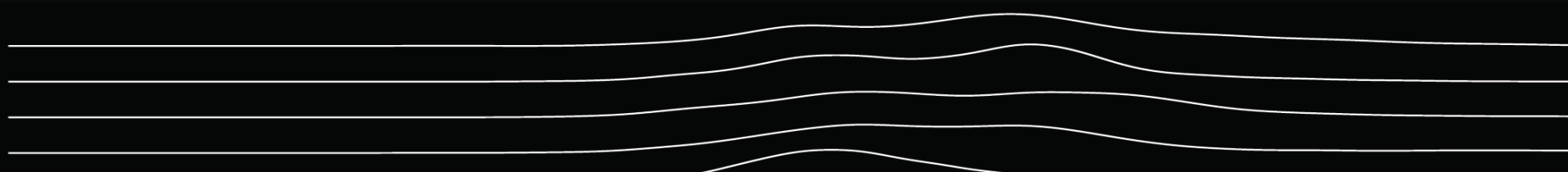
Measurement Lab

Supporting Open Internet Research

Lai Yi Ohlsen laiyi@measurementlab.net

Chris Ritzo critzo@measurementlab.net

Slides: https://bit.ly/2SwP9yl



What is M-Lab?

- People/Organization: A joint initiative between staff at Code for Science & Society, Google, and Princeton PlanetLab
- Data: An open repository of user-contributed, longitudinal, open-source derived Internet infrastructure data
- Infrastructure: A global infrastructure deployed, built, and run to aid in the creation of that data repository

M-Lab’s Mission

Measure the internet.
Save the data.
Make it universally accessible and useful.

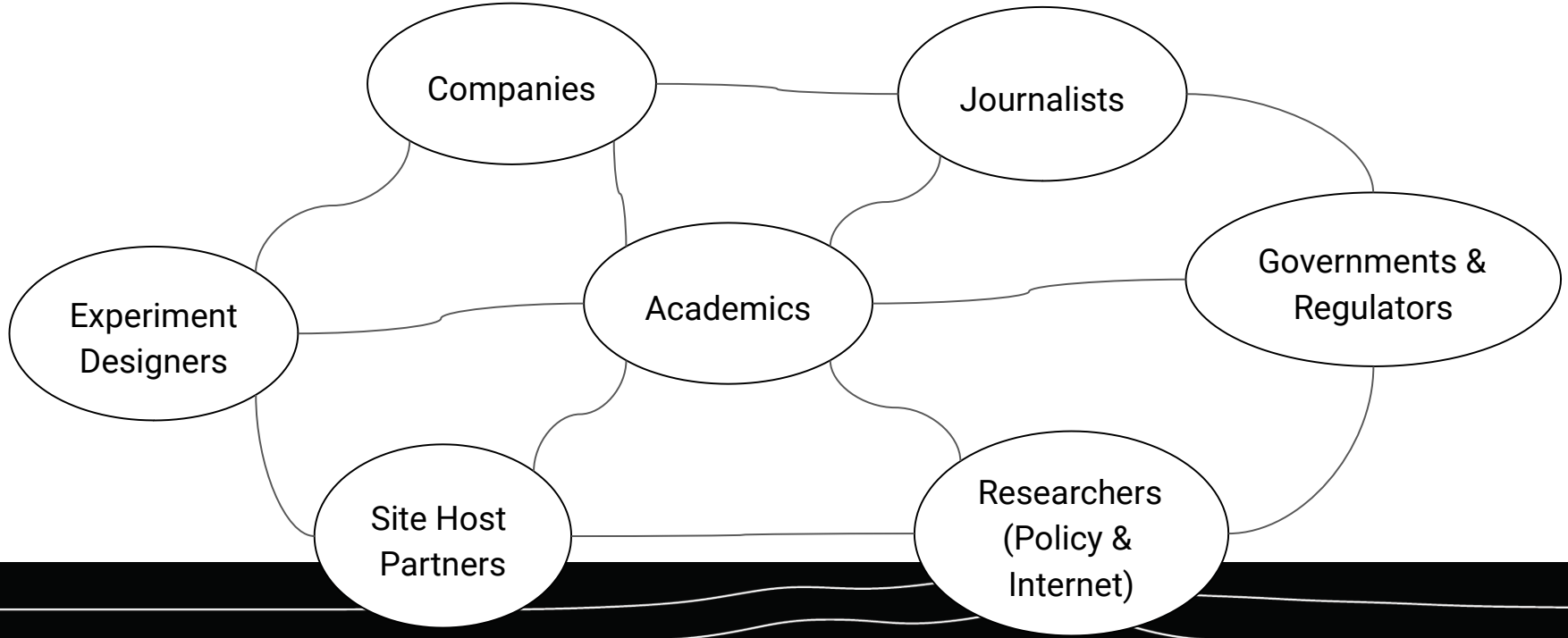
Note: we don’t measure the internet by ourselves -- people measure the Internet, using their own computers/phones and our servers, and we collect the data, and support them in their measurements

M-Lab Principles

- All measurements are active measurements
  - All the data is synthetic data, we take user privacy seriously.
  - Client initiated tests only, Servers do not start tests on their own.
- Experiments are curated and approved by a panel of reviewers, primarily developed by academics
- Edge clients come from the community.
  - Anyone can develop them.
  - Anyone can run tests against the platform.
- Openness
  - All of the data is released CC0.
  - All of the code is open source.

Learn more about M-Lab: 10th Anniversary Convening blog post, videos: https://www.measurementlab.net/blog/mlab-10year-wrapup/

Our Community

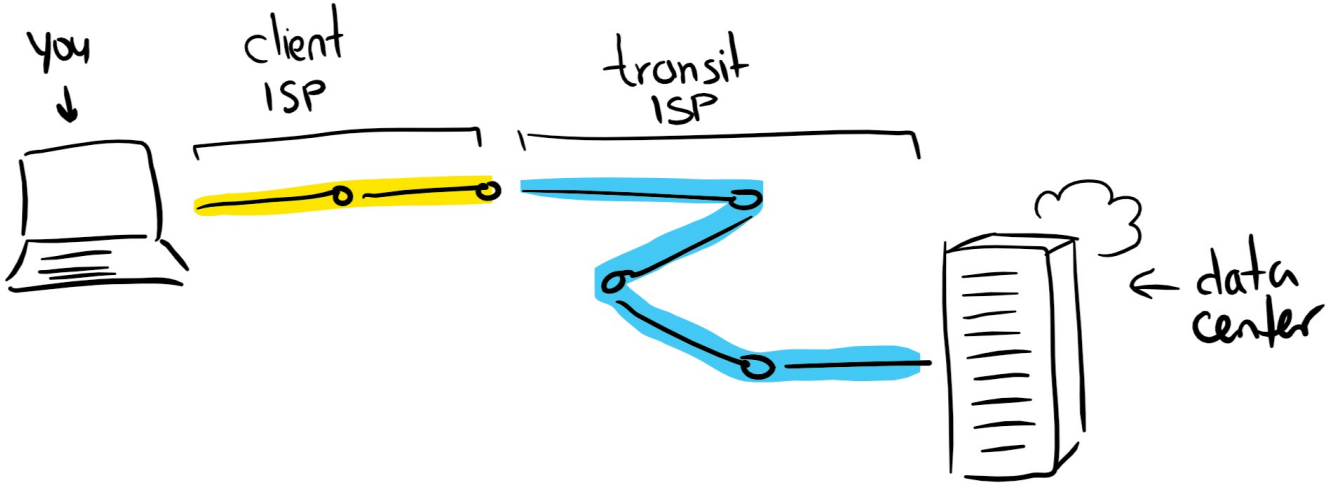


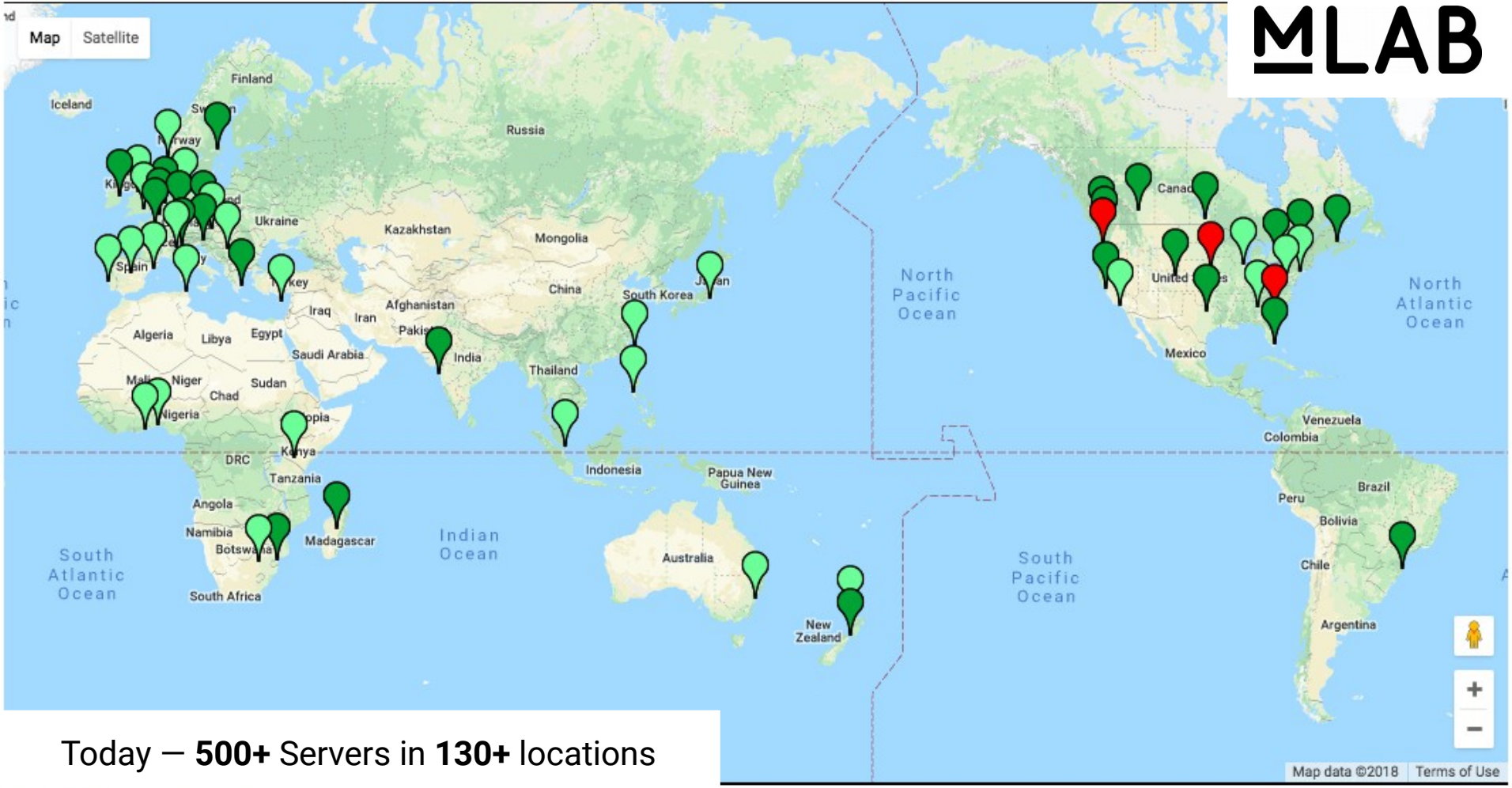
MLAB

M-Lab Platform Infrastructure

High capacity servers placed next to content

M-lab measures user experience of the **full route** from user to content





MLAB

Today — **500+** Servers in **130+** locations

MLAB

M-Lab Datasets

Experiments

MLAB

Glasnost
Max Planck Institute for Software Systems

Neubot
Nexa Center for Internet and Society, Politecnico di Torino

Project Bismark
Princeton University

Reverse Traceroute
University of Washington

MobiPerf
University of Michigan

NPAD
Pittsburgh Supercomputing Center

Sharperprobe
Georgia Tech College of Computing

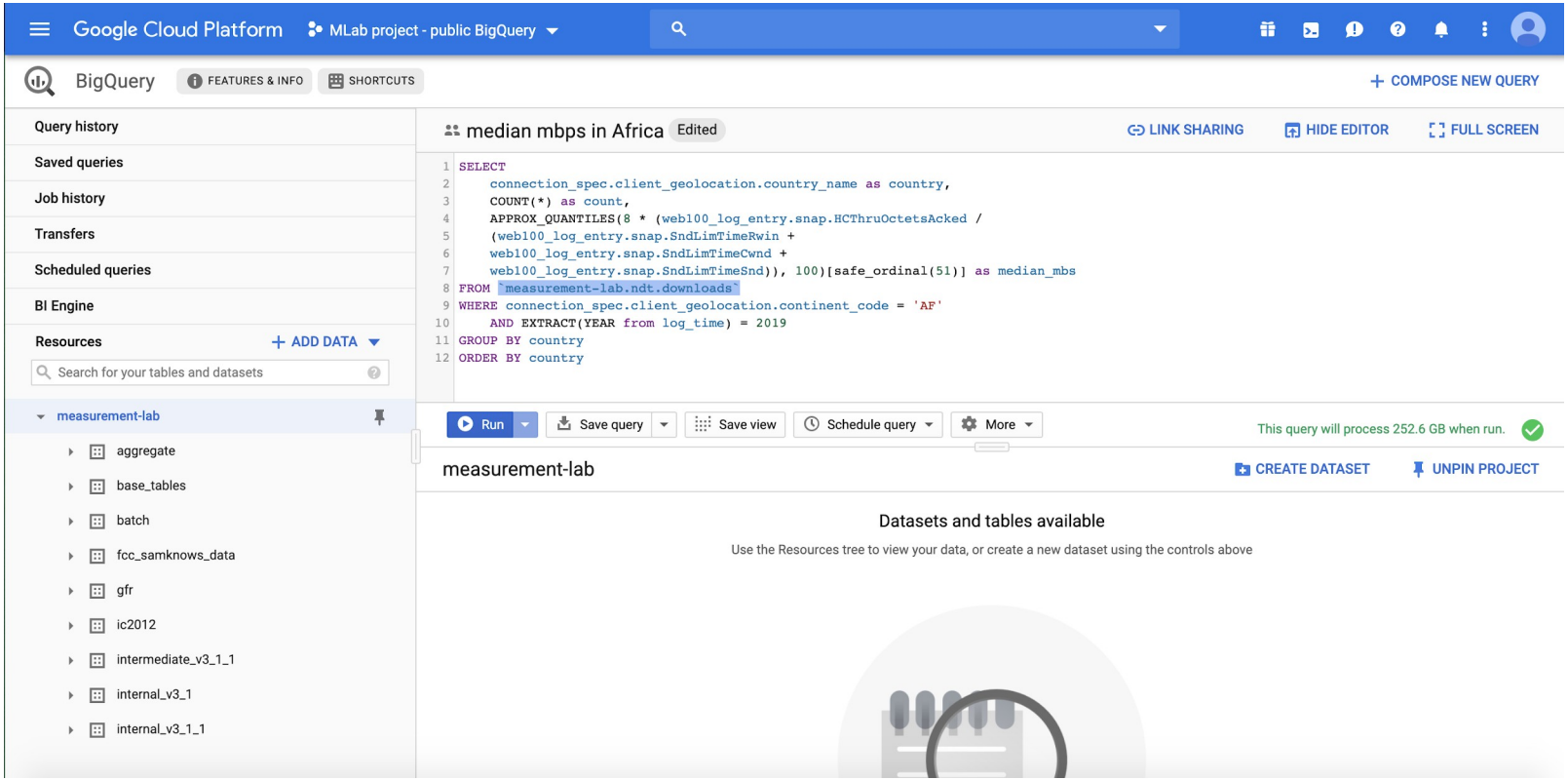
Network Diagnostic Tool
Internet2

Paris Traceroute
University Pierre et Marie Curie

Windrider
Northwestern University

Datasets in BigQuery

MLAB



Datasets

Network Diagnostic Tool
Internet2

- ndt
  - downloads
  - recommended
  - uploads
  - web100
- sidestream
  - web100

Paris Traceroute
University Pierre et Marie Curie

- node
  - traceroute

New schema coming soon!

- utilization
  - switch

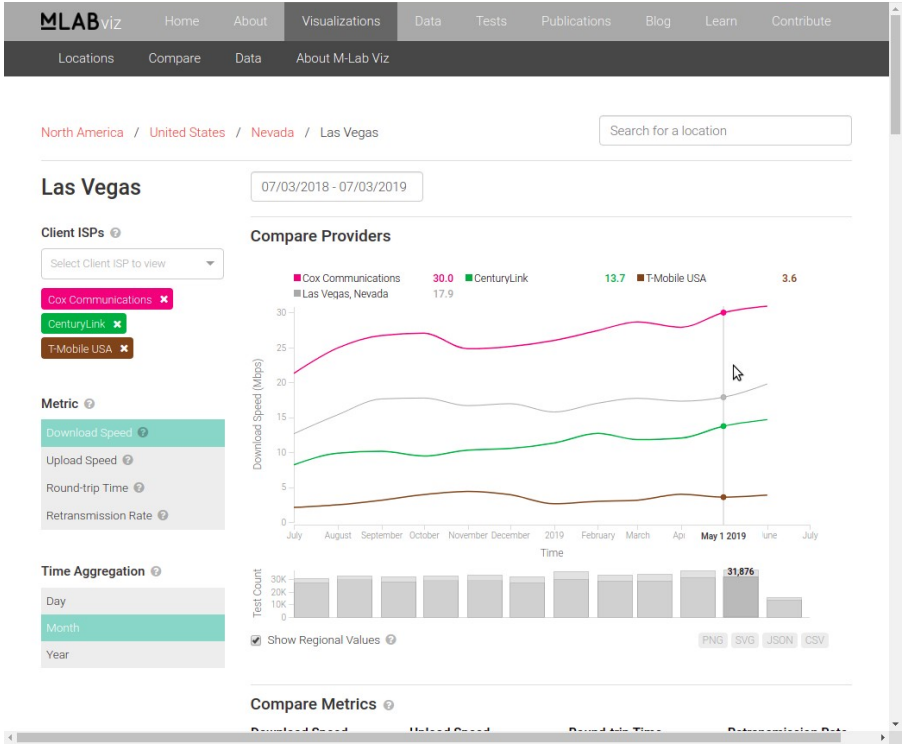
Access data schemas in Bigquery.

Schema | Details

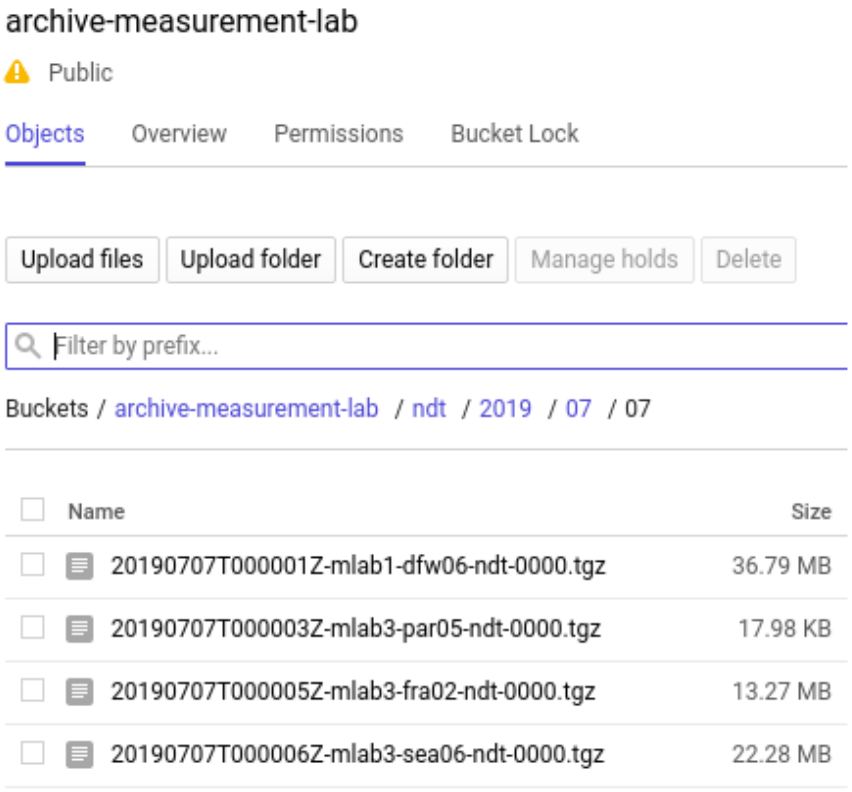
Field name	Type	Mode	Descri
partition_date	DATE	NULLABLE	
test_id	STRING	NULLABLE	
task_filename	STRING	NULLABLE	
parse_time	TIMESTAMP	NULLABLE	
parser_version	STRING	NULLABLE	
log_time	TIMESTAMP	NULLABLE	
blacklist_flags	INTEGER	NULLABLE	
anomalies	RECORD	NULLABLE	
anomalies. no_meta	BOOLEAN	NULLABLE	
anomalies. snaplog_error	BOOLEAN	NULLABLE	
anomalies. num_snaps	INTEGER	NULLABLE	
anomalies. blacklist_flags	INTEGER	NULLABLE	
connection_spec	RECORD	NULLABLE	
connection_spec. client_af	INTEGER	NULLABLE	

Accessing M-Lab Data

- Source data (raw packet traces & metadata)
- Visualization website
- BigQuery (parsed data)
- Other M-Lab & community-developed applications

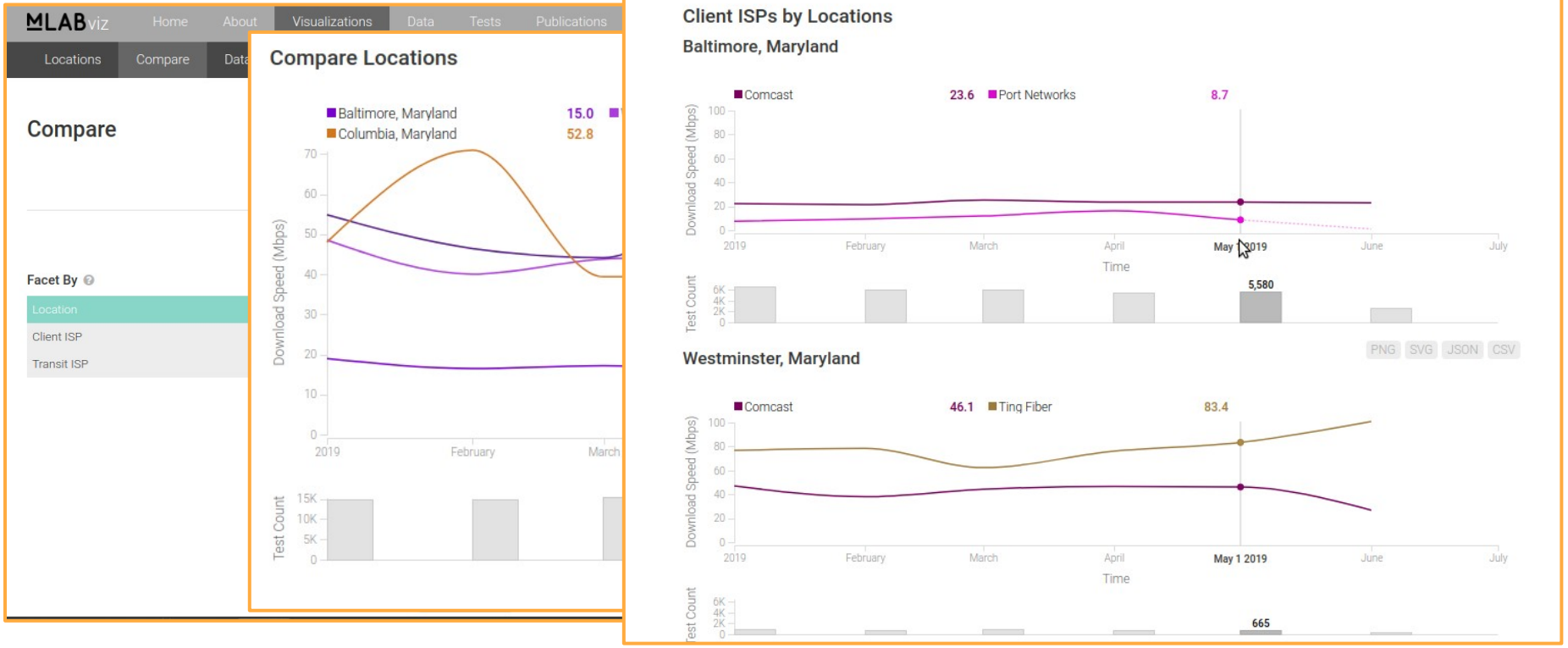


Accessing M-Lab Data: *Source Data*



- Openly available, raw packet traces, metadata for every test
  - Stored in Google Cloud Storage: https://console.cloud.google.com/storage/browser/archive-measurement-lab
  - Used primarily by academic researchers or in some cases by M-Lab partners with research resources
  - https://measurementlab.net/data/docs/gcs/

Accessing M-Lab Data:
M-Lab Visualization Website



M-Lab Data in BigQuery - Free!

- Sign up for the M-Lab Discuss List to get open & free access to the data:
  - Quick Start - www.measurementlab.net/quickstart/
  - Data Docs - www.measurementlab.net/data/docs/
  - Get help - support@measurementlab.net

MLAB

How Researchers Use M-Lab

How Researchers Use M-Lab

MLAB

Data Analysis

Selected Publications: https://www.measurementlab.net/publications/

- Estimating Residential Broadband Capacity using Big Data from M-Lab
  - University of New South Wales, Sydney, Australia
- Access is more than cost: Measuring the quality of mobile broadband service
  - Web Foundation, Alliance for Affordable Internet
- Can Competition-Enhancing Regulation Bridge the Quality Divide in Internet Provision?
  - National University of Singapore
- NETPerfTrace – Predicting Internet Path Dynamics and Performance with Machine Learning
  - Université de Liège - ULiège

How Researchers Use M-Lab

Hosting a Measurement Experiment

- M-Lab will soon use Docker/Kubernetes to manage the platform
- Opportunities for new experiments to be hosted on M-Lab
- Researchers designing server/client measurement experiments can apply to host with M-Lab
  - WeHe - Net Neutrality focused test https://dd.meddle.mobi/
- support@measurementlab.net

MLAB

How M-Lab Supports Researchers

M-Lab Supports Researchers

MLAB

Support, Documentation, Workshops, Training

- Email support@measurementlab.net
  - How to access & query data free of charge
  - Schema descriptions, sample queries, documentation
- Presenting workshops and training
  - African Internet Summit - Kampala, Uganda - June 2019
  - Univ. of Guelph, Ontario, Canada - June 2019
  - Africom / Afrinic 2018
  - SIGCOMM 2018 Hackathon - Budapest 2018
- Sponsoring Research Events
  - TMA School
  - Internet Measurement Conference

M-Lab Supports Researchers

MLAB

Partner in research program to build structured data collection tools

Simmons UNIVERSITY

INTERNET2®

CS&S Code for Science & Society

- IMLS Measuring Library Broadband Networks
  - Partners: Internet2, Simmons University
  - Year 1: 10 Libraries / Year 2: +50 Libraries
  - NDT, speedtest.net, and more
  - https://slis.simmons.edu/blogs/mlbn/
- Goal: Open source code, automated testing via on-premise devices, data visualization, local/regional data comparison
- Alexandria, Virginia Pilot - https://tinyurl.com/y5m3rmek
  - SBCs running NDT tests on a randomized schedule
  - Devices in classrooms (on wired & wireless)

MLAB

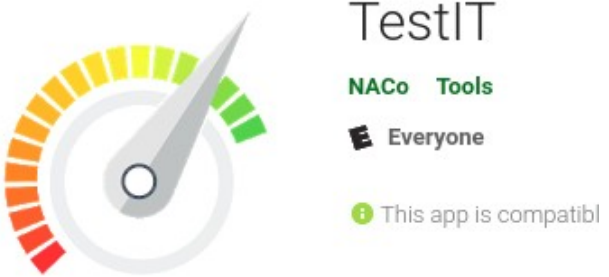
M-Lab Supports Community Based Research

Community Based Research

App / Client for Data Collection

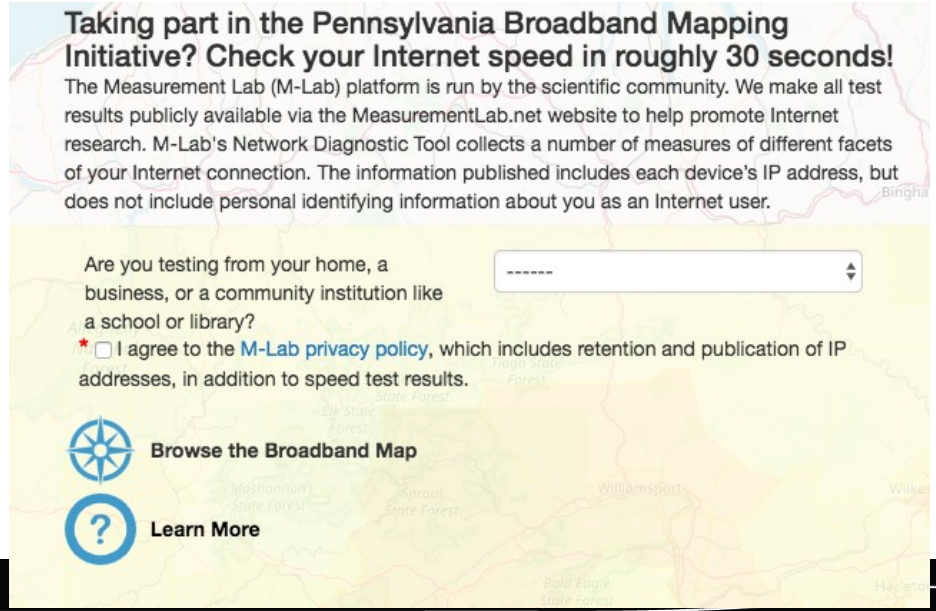
- National Association of Counties (NACo) using a branded app integration of NDT to collect data for rural areas
- Uses the M-Lab platform as a measurement service
- NDT test results are sent to M-Lab

- ***Enhanced location data*** is saved in NACo's database, with a copy of the standard test results



Community Based Research

Survey & Mapping Initiatives



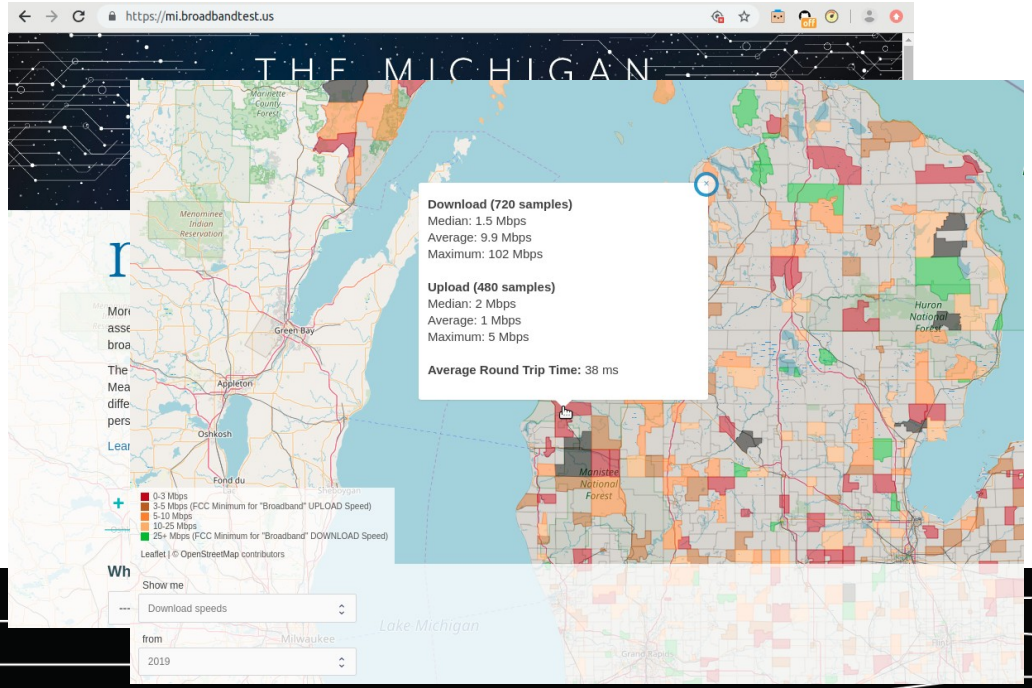
- M-Lab prototyped community tool used to conduct broadband survey, speed test, & aggregation: *Piecewise*

- City of Seattle, WA
- Stevens County & Ferry County, WA
- Clearwater County, ID

Community Based Research

MLAB

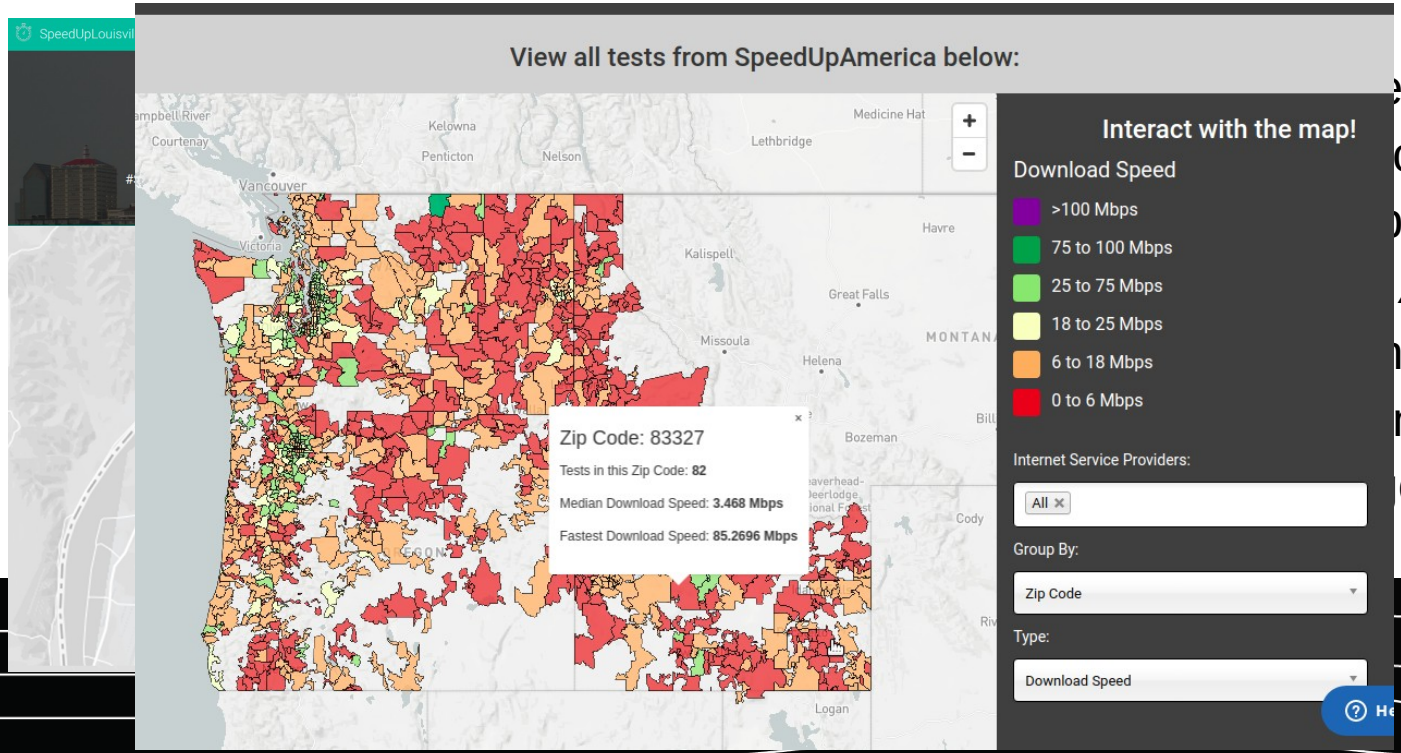
Survey & Mapping Initiatives



- Michigan / MERIT - broadband testing & mapping with a homework gap hook
  - R&E network, K12 statewide systems
  - Homework assignment to run a test
- https://mi.broadbandtest.us/

Community Based Research

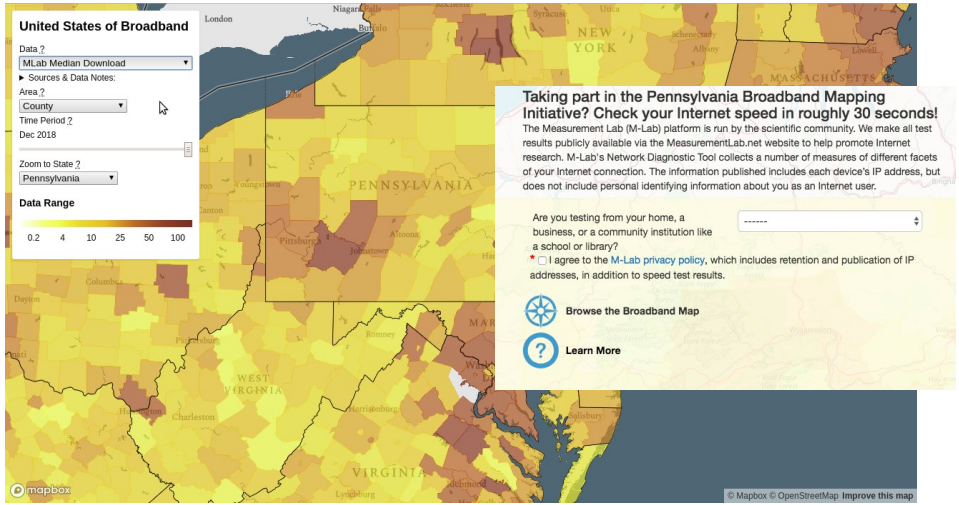
Survey & Mapping Initiatives



Louisville
eloped by Louisville,
c/private partnership
ported by M-Lab
America
ner development at
nology Assoc. of
on

Community Based Research

Survey & Mapping Initiatives



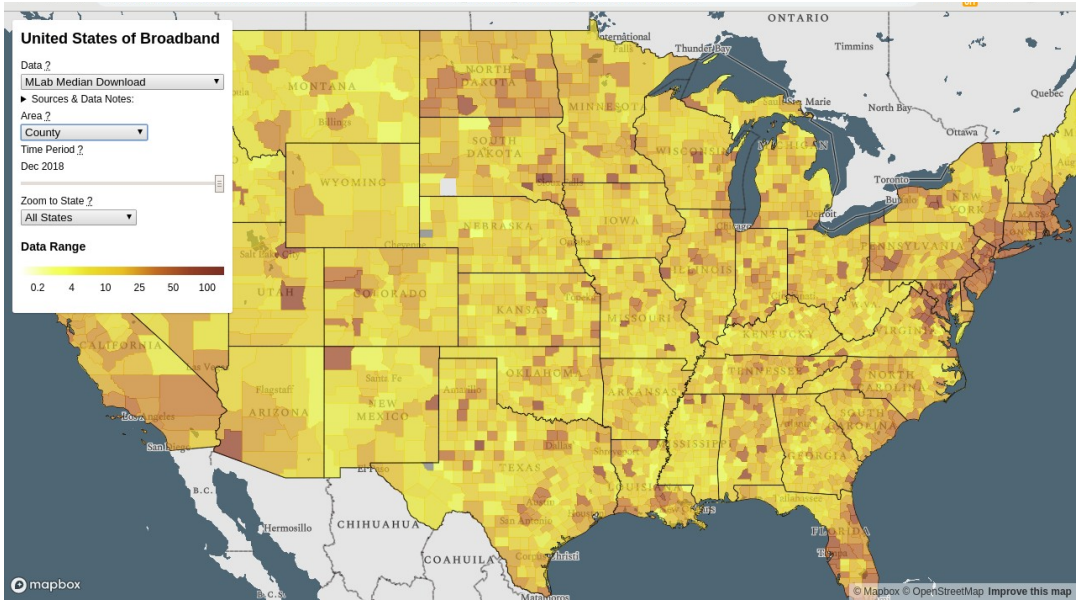
- Study conducted by Penn State, compares measured speeds with each FCC 477 data release
- Select M-Lab speeds & metrics, FCC, or +/- difference layers
- Aggregate by county, state house/senate, zipcode, census tract
- https://pa.broadbandtest.us

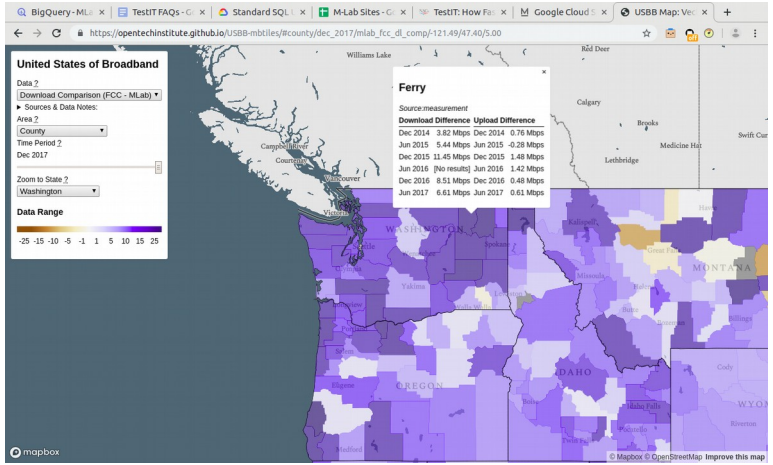
Community Based Research

MLAB

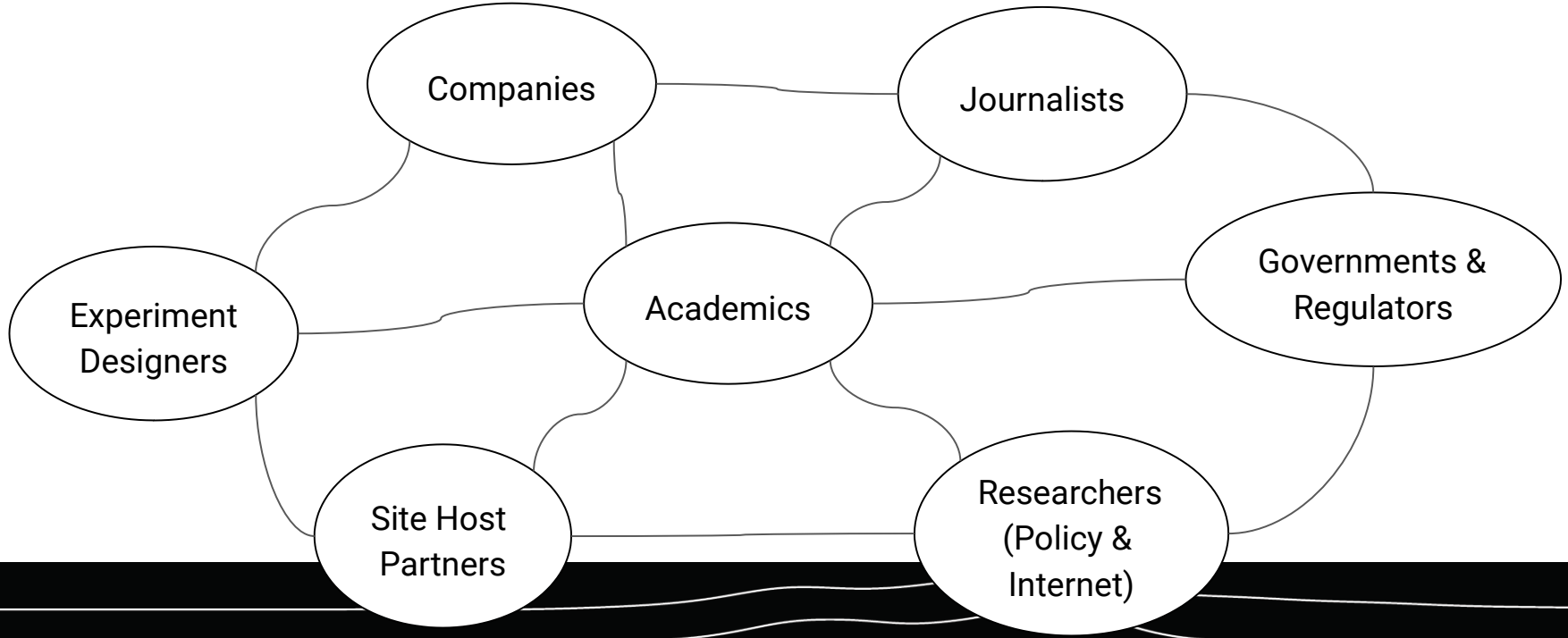
United States of Broadband

https://www.newamerica.org/oti/reports/united-states-broadband-map/





Our Community



MLAB

Measurement Lab

Supporting Open Internet Research

Lai Yi Ohlsen laiyi@measurementlab.net

Chris Ritzo critzo@measurementlab.net