Measuring the Mobile Internet

David Choffnes

Northeastern University

with U. Michigan, USC, Google, INRIA, and many others

Motivation

Most of the time people spend online is via a mobile device*

Motivation

Most of the time people spend online is via a mobile device*

When YouTube hangs, is it

- A slow device?
- Weak signal?
- Contention for bandwidth?
- Bad path to Google?
- ISP shaping?
- Bad server?

Motivation

Most of the time people spend online is via a mobile device*

When YouTube hangs, is it

- A slow device?
- Weak signal?
- Contention for bandwidth?
- Bad path to Google?
- ISP shaping?
- Bad server?

Key challenge: We need extensive network measurements to perform root cause analysis

There's an app for that

Solution: App-based measurement

- Speedometer
- Mobiperf

- MySpeedTest
- Netalyzr Mobile
- Namehelp Mobile
- …

How is this working out for us?

Lessons learned from mobile measurement

- What you measure may not be what you think

Lessons learned from mobile measurement

- What you measure may not be what you think

Example: Estimate page load time for google.com (low cost)

1. DNS lookup for google.com
2. Ping google.com
3. Estimate latency as DNS lookup time + N RTTs

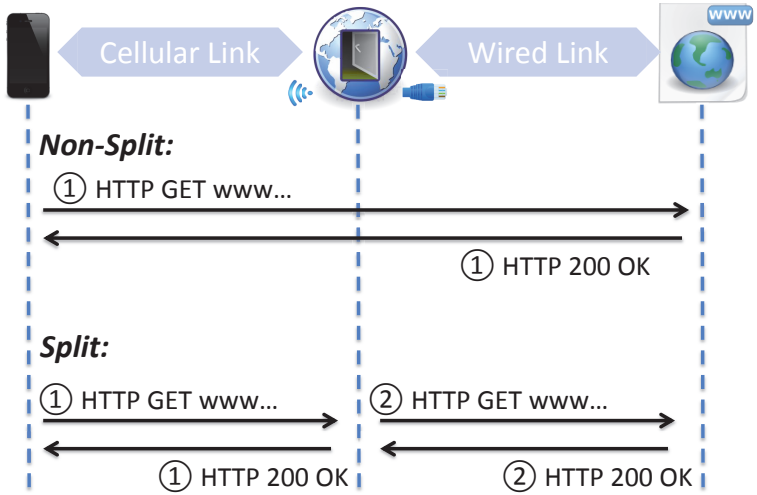
Lessons learned from mobile measurement

- What you measure may not be what you think

Example: Estimate page load time for google.com (low cost)

1. DNS lookup for google.com
2. Ping google.com
3. Estimate latency as DNS lookup time + N RTTs

- But when you do an HTTP GET for google.com…
  - Connection is split



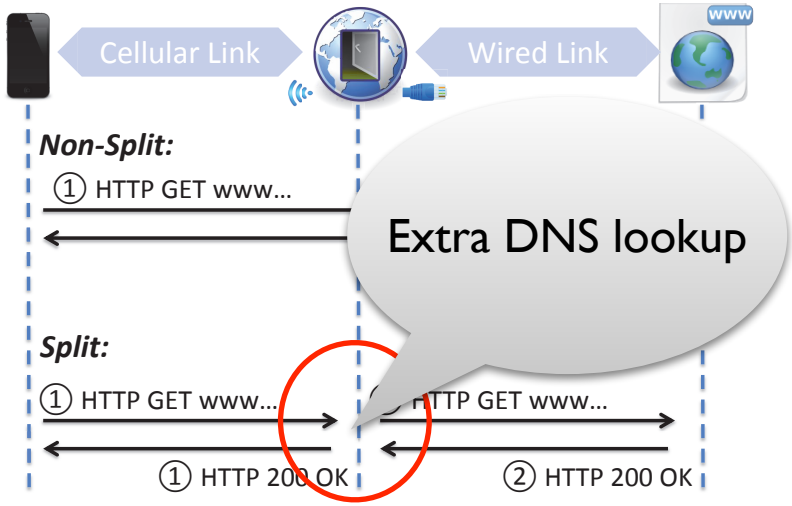
Lessons learned from mobile measurement

- What you measure may not be what you think

Example: Estimate page load time for google.com (low cost)

1. DNS lookup for google.com
2. Ping google.com
3. Estimate latency as DNS lookup time + N RTTs

- But when you do an HTTP GET for google.com…
  - Connection is split
  - New DNS lookup based on `Host:` field



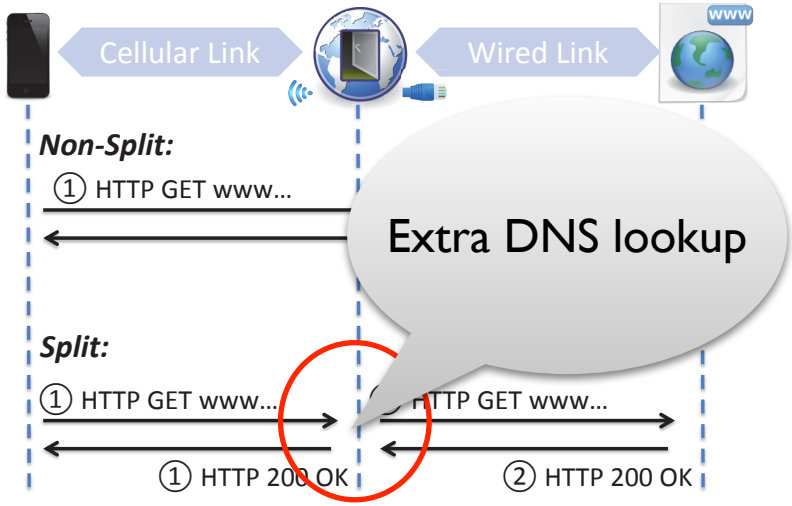
Lessons learned from mobile measurement

- What you measure may not be what you think

Example: Estimate page load time for google.com (low cost)

1. DNS lookup for google.com
2. Ping google.com
3. Estimate latency as DNS lookup time + N RTTs

- But when you do an HTTP GET for google.com…
  - Connection is split
  - New DNS lookup based on `Host:` field
  - See Xing's poster for more info



Lessons learned from mobile measurement

- Spray and pray measurement can get you breadth…

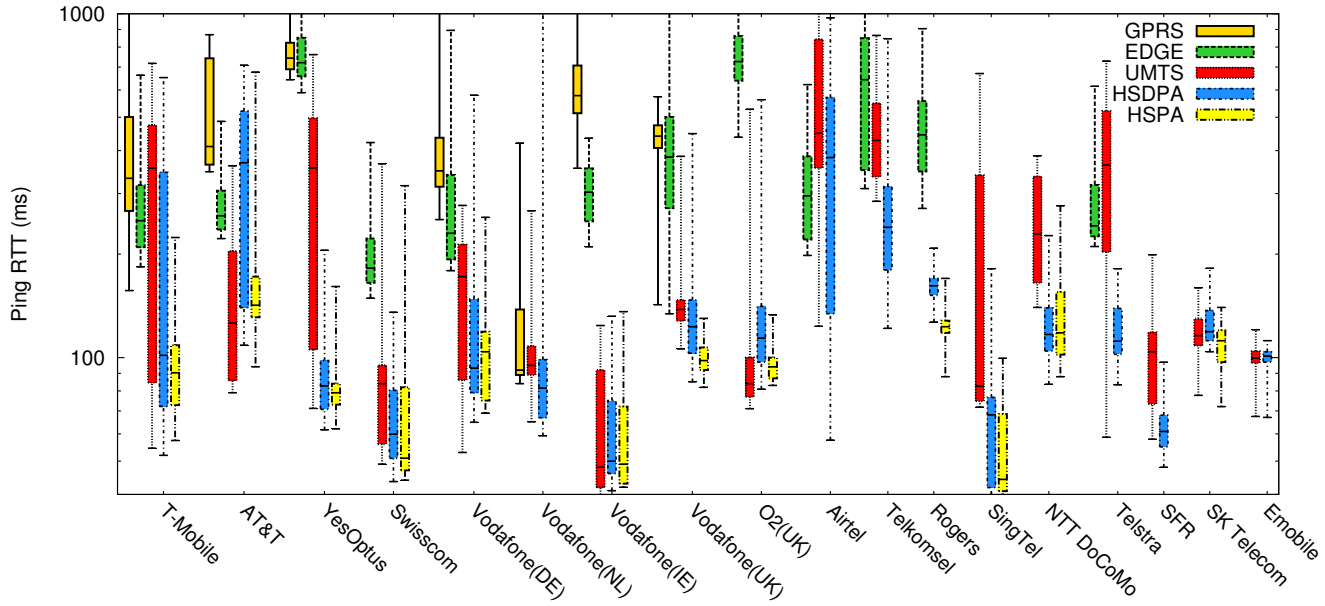
	HSPA	HSDPA	UMTS	EDGE	GPRS	LTE	EVDO	eHRPD	1xRTT
# of Measurements	439K	2326K	563K	506K	58K	1460K	2183K	301K	68K
# of Carriers	50	111	96	85	48	7	8	2	3

Lessons learned from mobile measurement

- Spray and pray measurement can get you breadth…

	HSPA	HSDPA	UMTS	EDGE	GPRS	LTE	EVDO	eHRPD	1xRTT
# of Measurements	439K	2326K	563K	506K	58K	1460K	2183K	301K	68K
# of Carriers	50	111	96	85	48	7	8	2	3

- …but noise in the data frustrates characterization



Lessons learned from mobile measurement

How do we isolate cases of bad performance? [PAM '14]

Lessons learned from mobile measurement

How do we isolate cases of bad performance? [PAM '14]

Example: Path inflation evolution in mobile networks

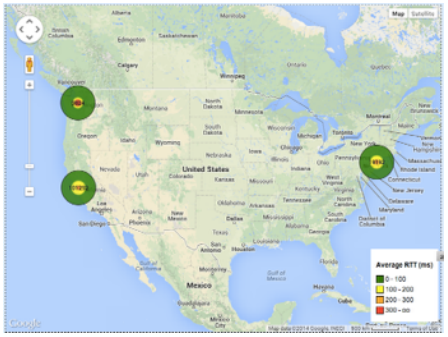
- Combination of traceroute and DNS lookups
- Knowledge of Google's infrastructure
- Continuous measurements

Lessons learned from mobile measurement

How do we isolate cases of bad performance? [PAM '14]

Example: Path inflation evolution in mobile networks

- Combination of traceroute and DNS lookups
- Knowledge of Google's infrastructure
- Continuous measurements

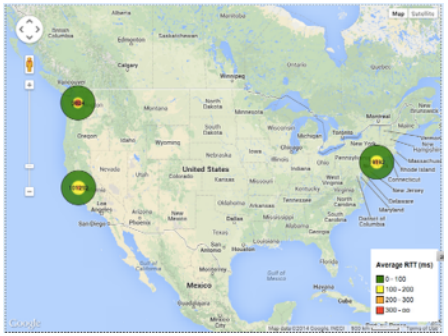


Lessons learned from mobile measurement

How do we isolate cases of bad performance? [PAM '14]

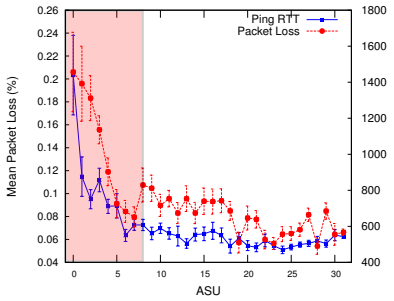
Example: Path inflation evolution in mobile networks

- Combination of traceroute and DNS lookups
- Knowledge of Google's infrastructure
- Continuous measurements



Example: Shotgun correlations

- Vary signal strength only, observe impact on performance
- Doesn't really generalize to factors such as location



Next generation of mobile measurement

Our 1st-generation apps have been useful!

Making further strides in this domain

- Needs a large(r) set of participating devices
- Real-time monitoring of performance
- On-demand (adaptive) measurements for RCA

Great, let's all go build new apps!

Pitfalls of App Proliferation

Coverage

- Who has a billion-install idea? Million-install?
  - If so, what the heck are you doing here?

Consistency

- How do we join datasets?
- What measurement implementations?

Cost

- How many times do we need to run the same ping?

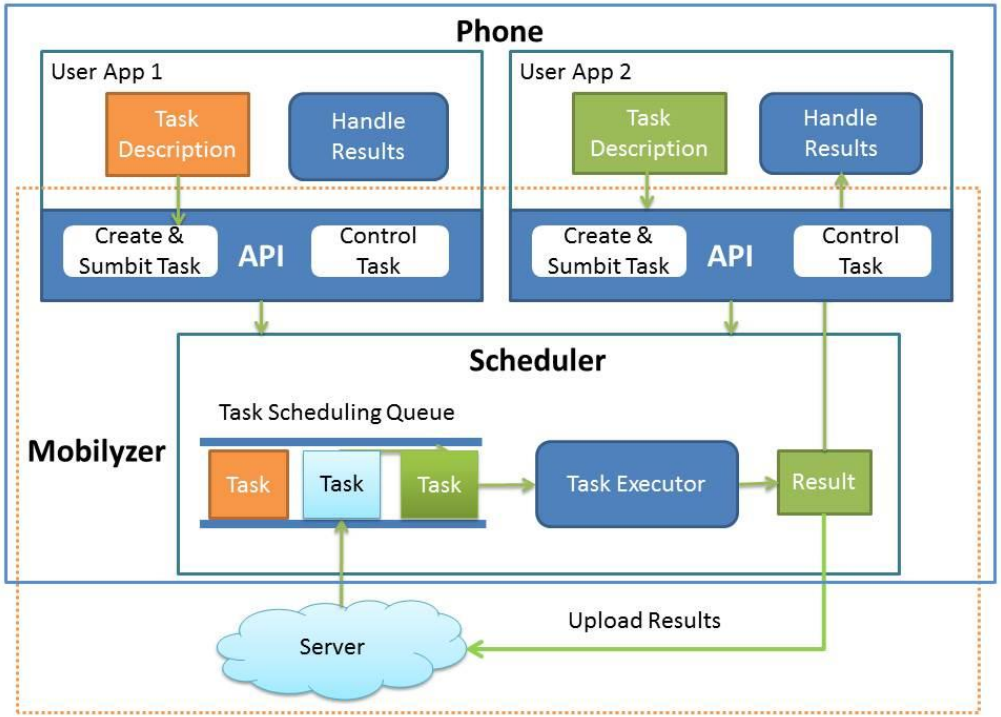
Interference

- What if MST runs during Mobiperf's ping test?

Mobilyzer: Mobile Measurement Library

Our proposal: One measurement library, many apps

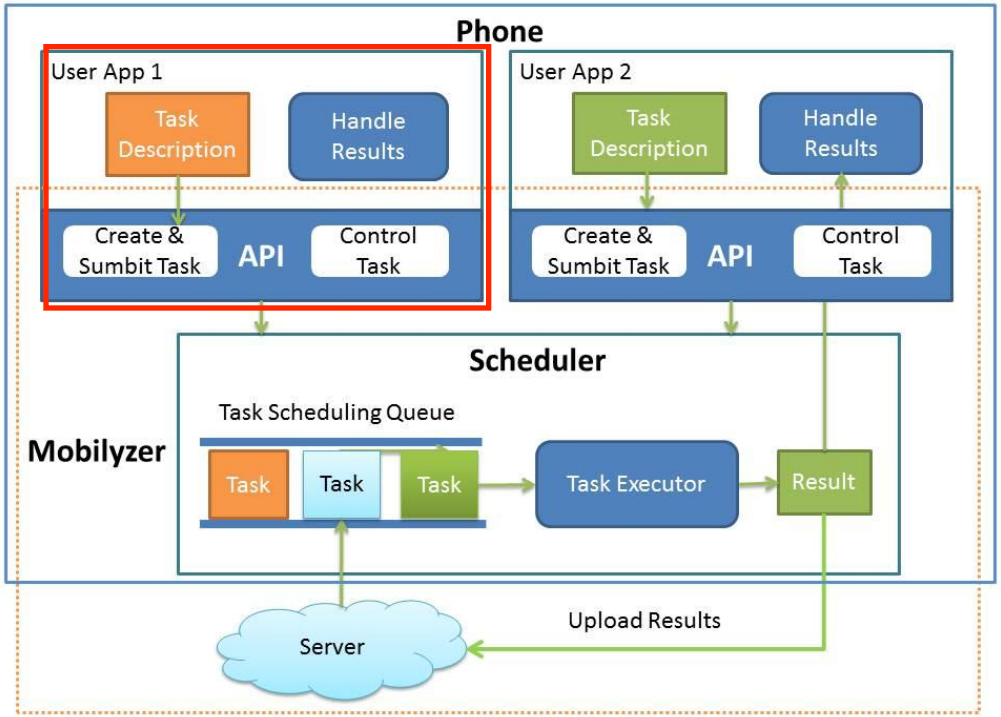
- API for measurements simplifies app dev



Mobilyzer: Mobile Measurement Library

Our proposal: One measurement library, many apps

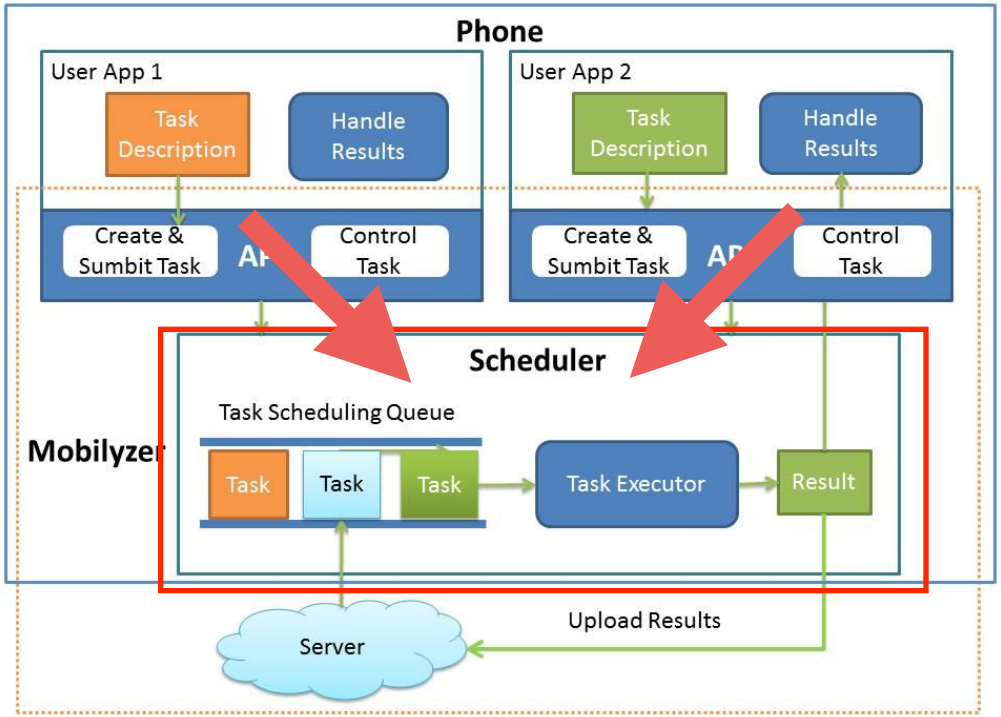
- API for measurements simplifies app dev



Mobilyzer: Mobile Measurement Library

Our proposal: One measurement library, many apps

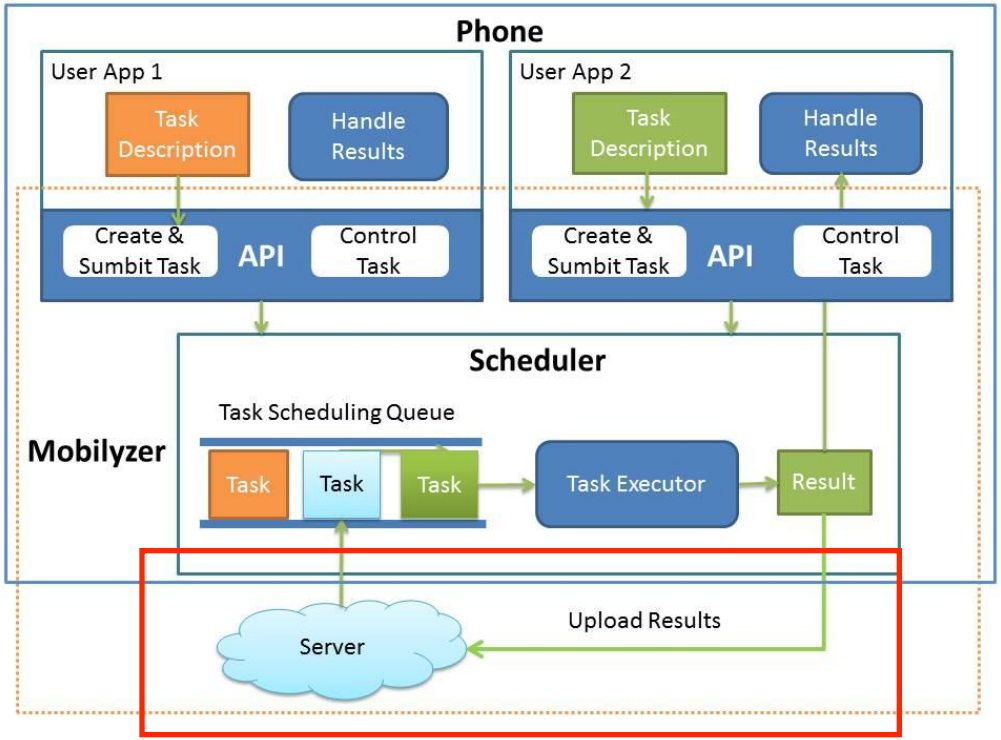
- API for measurements simplifies app dev
- Single scheduler per device allows apps to coexist



Mobilyzer: Mobile Measurement Library

Our proposal: One measurement library, many apps

- API for measurements simplifies app dev
- Single scheduler per device allows apps to coexist
- Data collection



Why Mobilyzer?

Designed with researcher incentives in mind

- Simplifies app development
  - Just ask Nick's student
- PlanetLab-like participation model
  - Bring one Mobilyzed app, run measurements on all devices
- Dynamic server-side control of measurements

See Hongyi's and Ashkan's posters for more

One last thing

What about detecting traffic differentiation in mobile?

One last thing

What about detecting traffic differentiation in mobile?

- Challenges
  - What traffic triggers differentiation?
  - How do we tell that the ISP is doing anything (vs. noise)?
- Our approach
  - Record and replay traces of real traffic from mobile apps
  - Use VPN tunnels (**Meddle**) as controlled experiments

See Abbas' poster for more info!

Thanks!

Obligatory project URLS

http://mobilyzer-project.mobi

http://mobiperf.com

http://meddle.mobi

http://openmobiledata.appspot.com