Is Bufferbloat a Privacy Issue?

Brian Trammell, ETH Zürich (w/ Mirja Kühlewind)

MAPRG IETF 102 — Montreal — Thursday 19 July 2018

mami

measurement and architecture for a middleboxed internet

measurement | architecture | experimentation

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688421.The opinions expressed and arguments employed reflect only the authors' view. The European Commission is not responsible for any use that may be made of that information.

Supported by the Swiss State Secretariat for Education, Research and Innovation under contract number 15.0268. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Swiss Government.

Executive summary

Yes*

bufferbloat has potential privacy impact

Executive summary

Yes*

bufferbloat has potential privacy impact

- *if a link has significant buffering

Executive summary

Yes*

bufferbloat has potential privacy impact

- *if a link has significant buffering
- *if the public IP address is associated only with that link

Executive summary

Yes*

bufferbloat has potential privacy impact

- *if a link has significant buffering
- *if the public IP address is associated only with that link
- *if the public IP address responds to ICMP Echo Request

Executive summary

Yes*

bufferbloat has potential privacy impact

- *if a link has significant buffering
- *if the public IP address is associated only with that link
- *if the public IP address responds to ICMP Echo Request
- *and if the Echo Request/Reply share the buffered queue

Executive summary

Yes*

bufferbloat has potential privacy impact

- *if a link has significant buffering
- *if the public IP address is associated only with that link
- *if the public IP address responds to ICMP Echo Request
- *and if the Echo Request/Reply share the buffered queue
- ****these conditions hold for one in seven networks we examined

Privacy and RTT-based geolocation

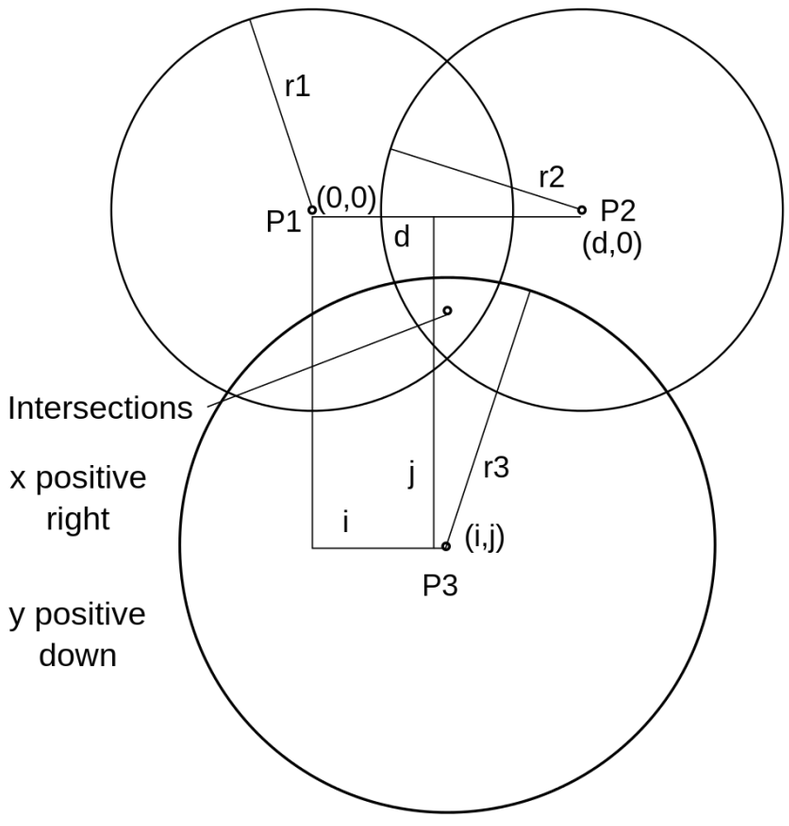
How did we get here?

measurement and architecture for a middleboxed internet

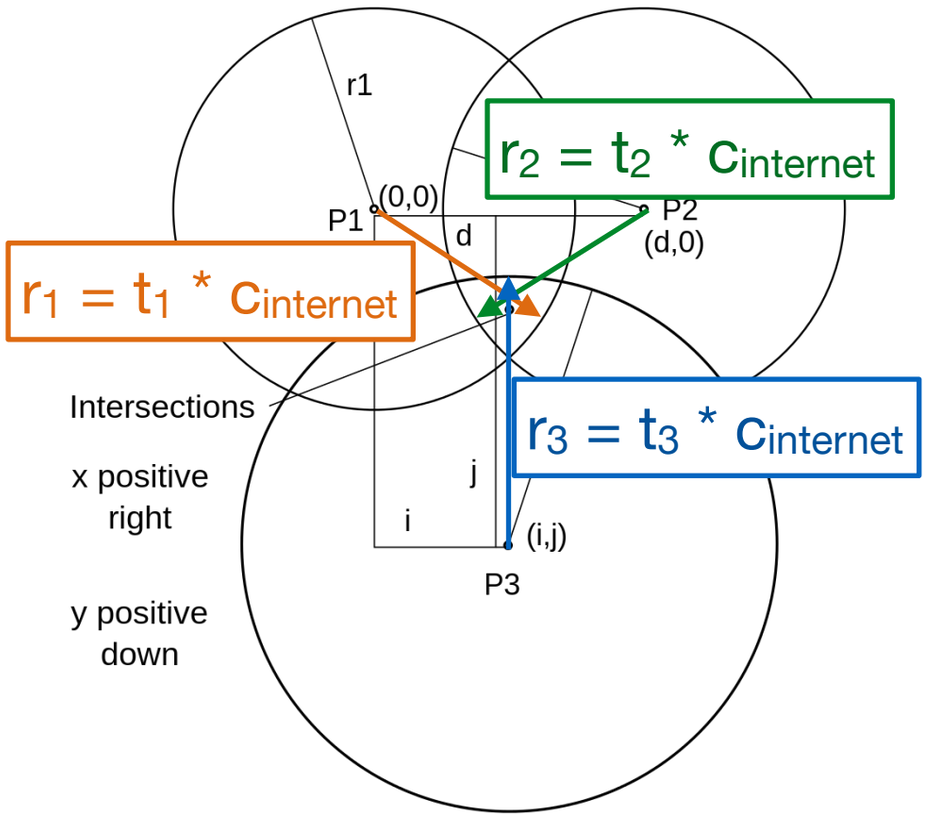
measurement | architecture | experimentation

"If I can ping you, I know where you are"

"If I can ping you, I know where you are"

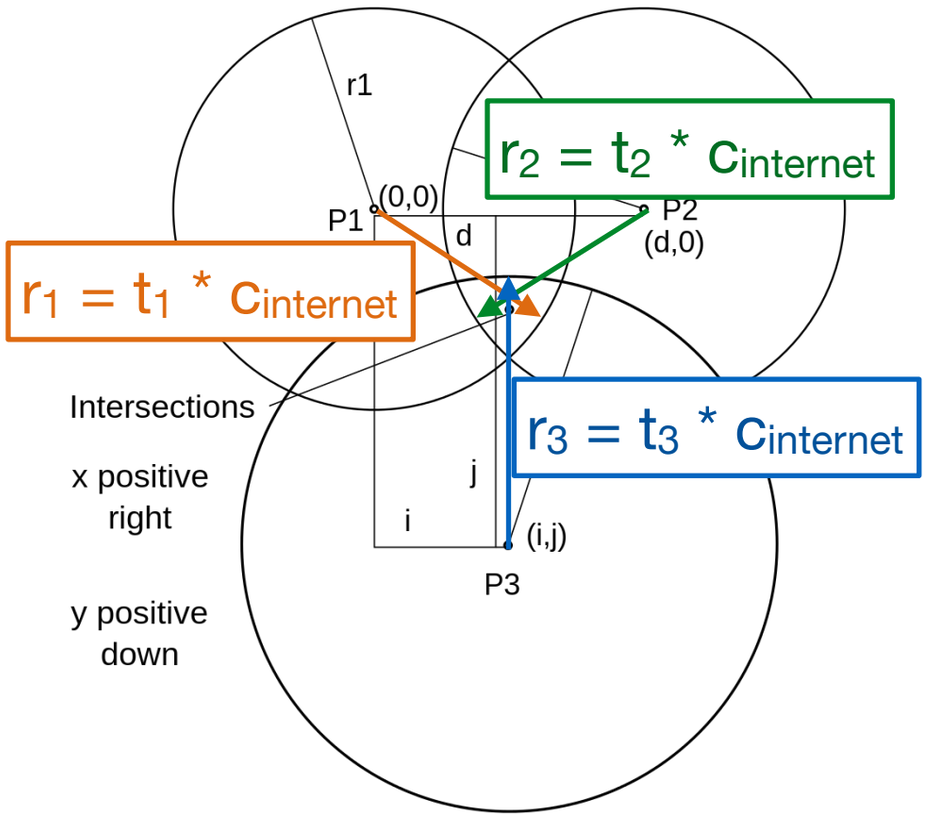


"If I can ping you, I know where you are"



CC-BY-SA-3.0 (wikipedia:Rhb100)

"If I can ping you, I know where you are"

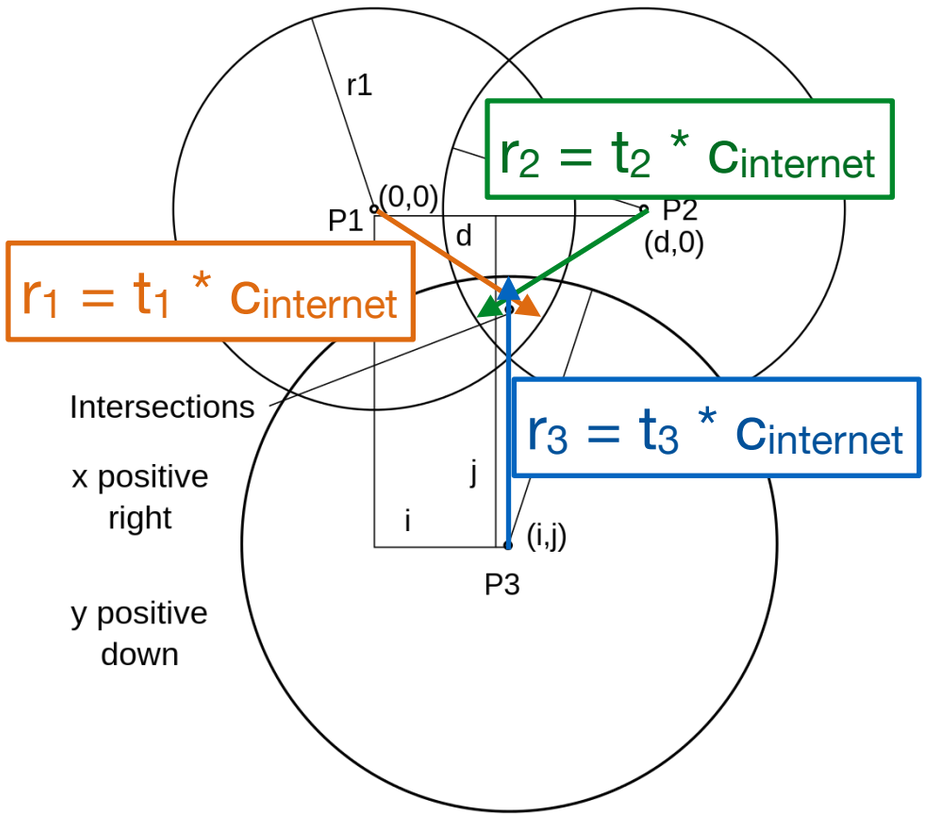


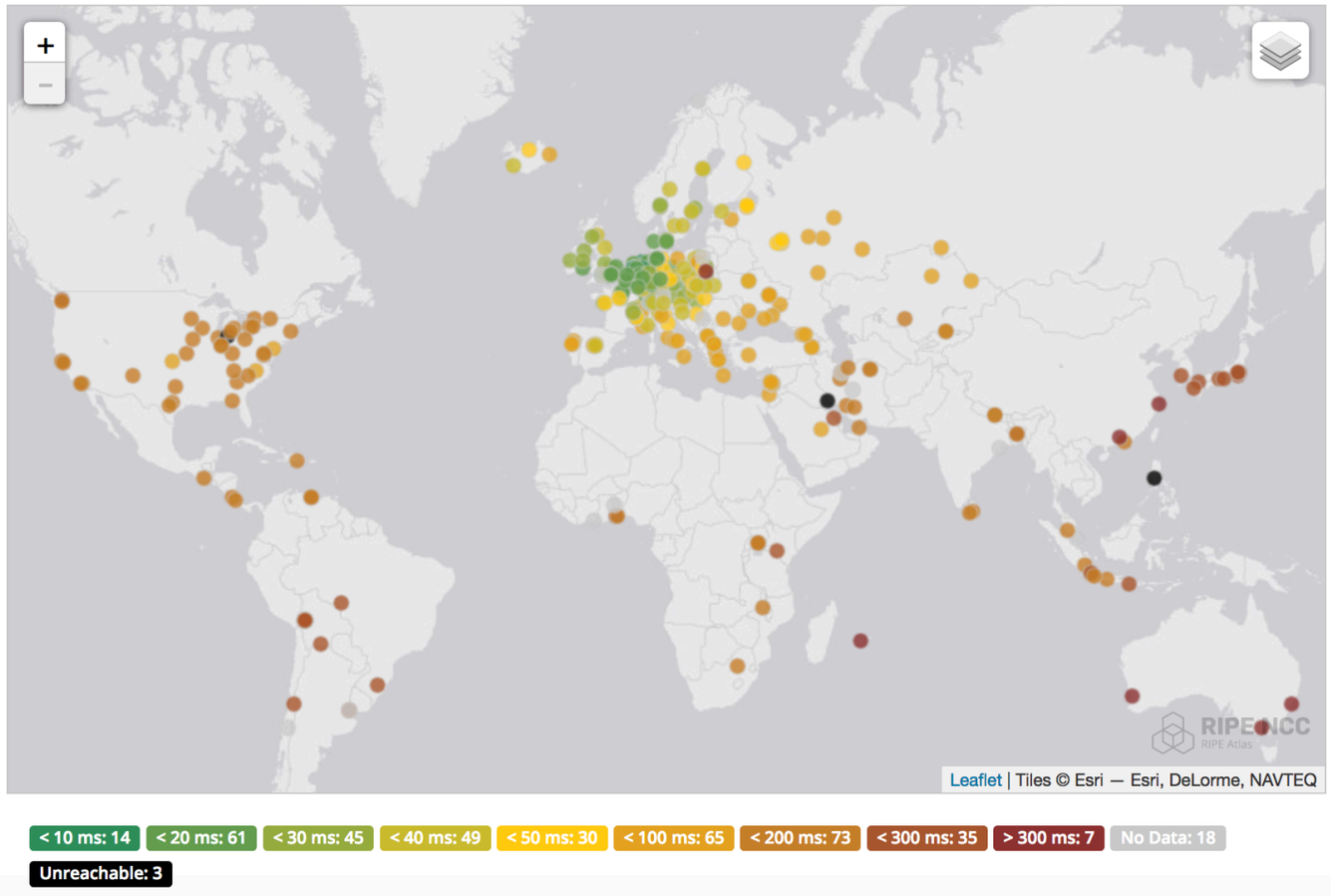
$$x = \frac{r_1^2 - r_2^2 + d^2}{2d}.$$

$$y = \frac{r_1^2 - r_3^2 - x^2 + (x-i)^2 + j^2}{2j} = \frac{r_1^2 - r_3^2 + i^2 + j^2}{2j} - \frac{i}{j}x.$$

$$z = \pm\sqrt{r_1^2 - x^2 - y^2}.$$

"If I can ping you, I know where you are"

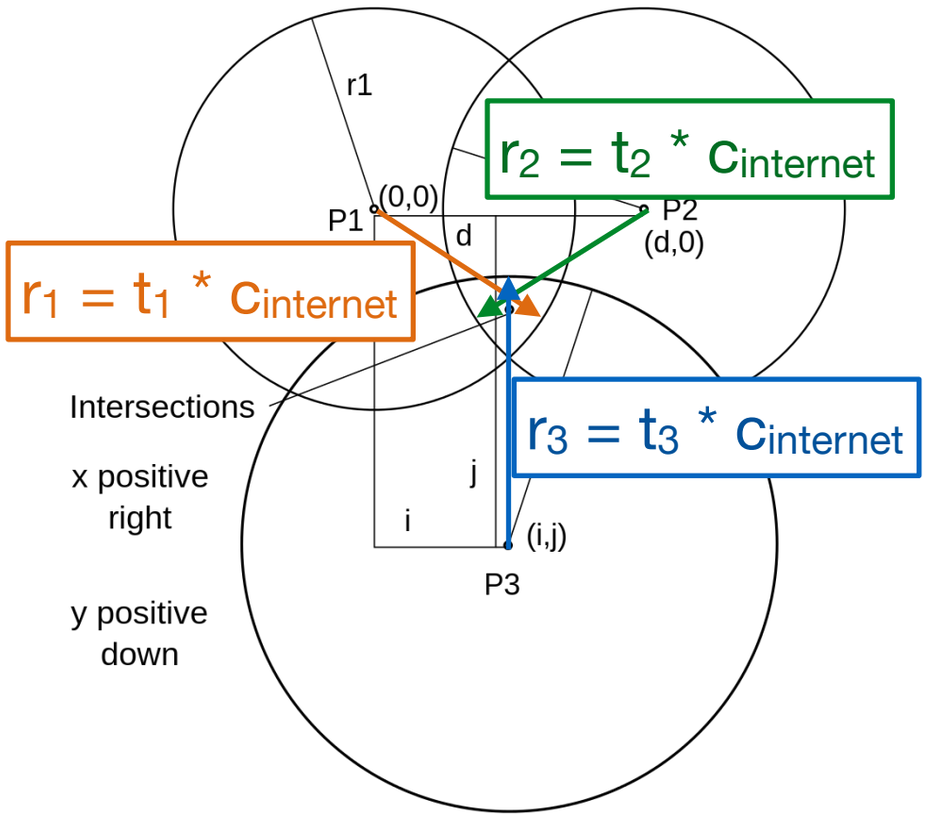


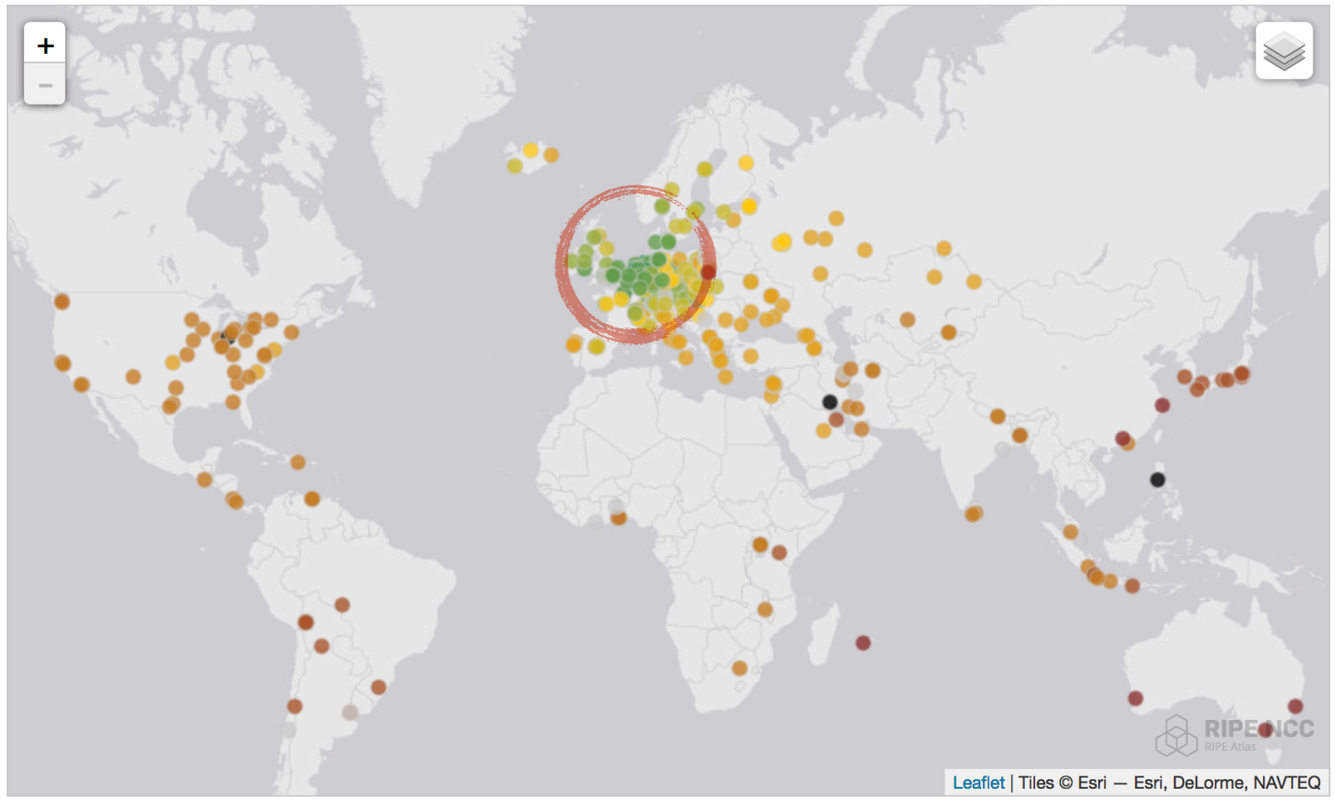


CC-BY-SA-3.0 (wikipedia:Rhb100)

https://atlas.ripe.net/measurements/11583536/#!map

"If I can ping you, I know where you are"





https://atlas.ripe.net/measurements/11583536/#!map

...actually not so much.

...actually not so much.

- Internet RTT is the sum of delays at each hop, some terms of which are variable:

$$RTT_{obs} = \sum_{n=0}^{f}(D_{prop_{n\to n+1}} + D_{queue_n} + D_{proc_n})+$$
$$\sum^{r}(D_{prop_{m\to m+1}} + D_{queue_m} + D_{proc_m})+$$
$$D_{stack} + D_{app}$$

- Distance can be derived only when queueing, stack, and application delay are held to zero:

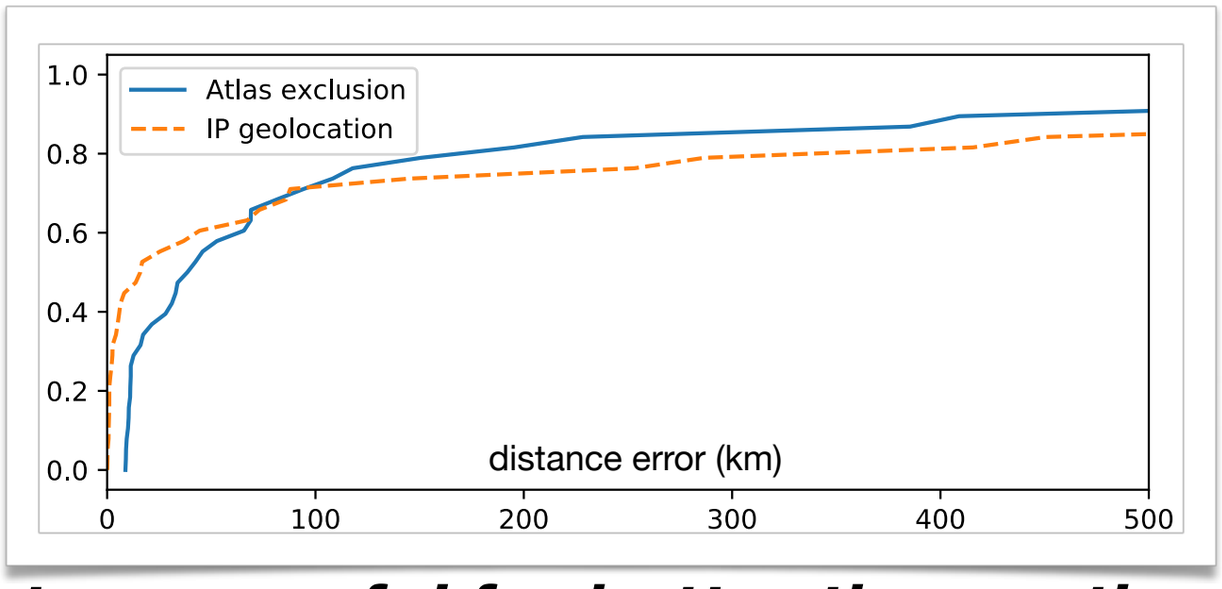
$$dist < \frac{\sum_{n=0}^{f} D_{prop_{n\to n+1}} + \sum_{m=0}^{r} D_{prop_{m\to m+1}}}{2} \times c_{internet}$$

- When target address is redacted, the risk is entirely dependent on how close the known address is to the unknown address:
  - 1ms RTT → <100km distance

...actually not so much.

- Internet RTT is the sum of delays at each hop, some terms of which are variable:

$$RTT_{obs} = \sum_{n=0}^{f}(D_{prop_{n\to n+1}} + D_{queue_n} + D_{proc_n})+$$
$$\sum^{r}(D_{prop_{m\to m+1}} + D_{queue_m} + D_{proc_m})+$$
$$D_{stack} + D_{app}$$



- ***RTT > 10ms not very useful for better than national location.***

Sometimes the answer is another question....

- We were concerned about the geoprivacy implications of *passive* observation of RTT
  - (which turns out not to be all that scary)

- But does *active* observation of RTT pose a problem?
  - What else can we extract from RTT data?

RTT-based load telemetry

measurement and architecture for a middleboxed internet

measurement | architecture | experimentation

Remote Load Telemetry

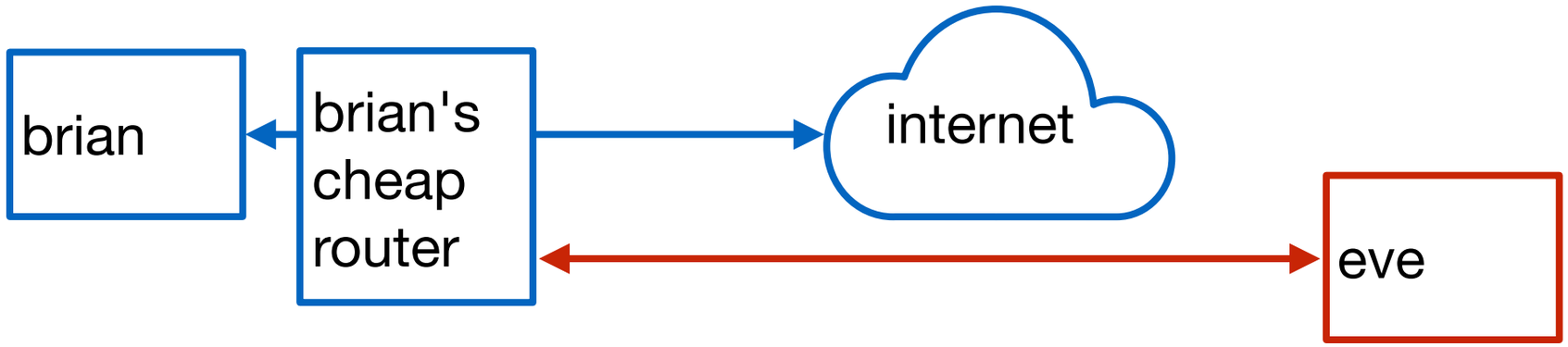
- Can a remote entity armed only with *ping* extract information about the operation of machines on my network?

Remote Load Telemetry

- Can a remote entity armed only with *ping* extract information about the operation of machines on my network?

$$load_{net} \propto \sum_{n=0}^{f} D_{queue_n} + \sum_{m=0}^{r} D_{queue_m}$$

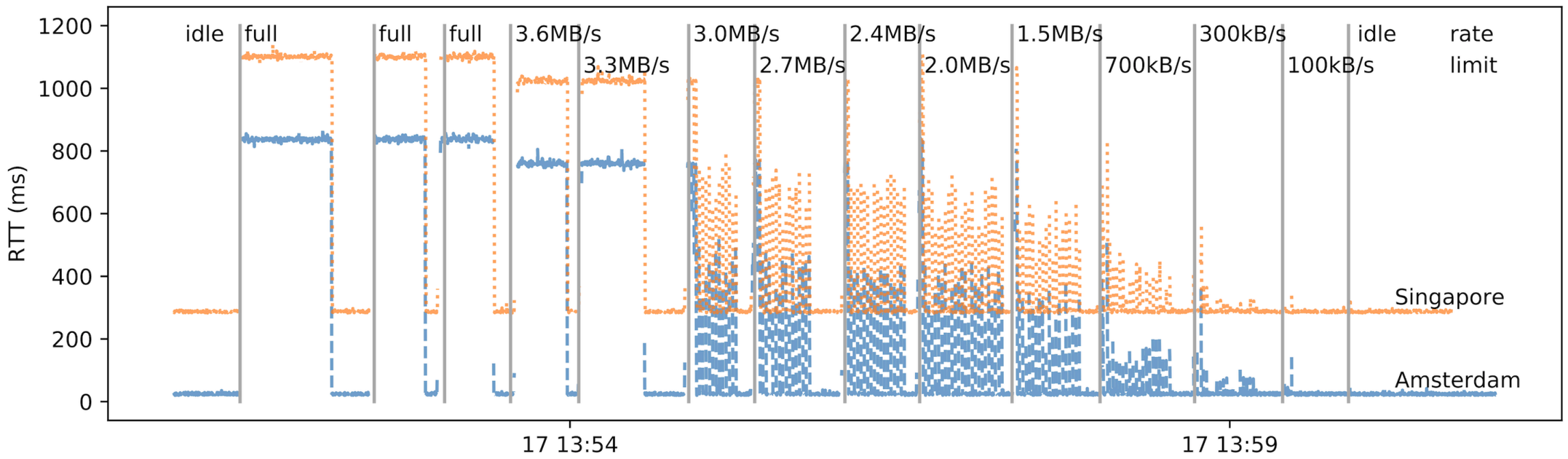
Remote Load Telemetry



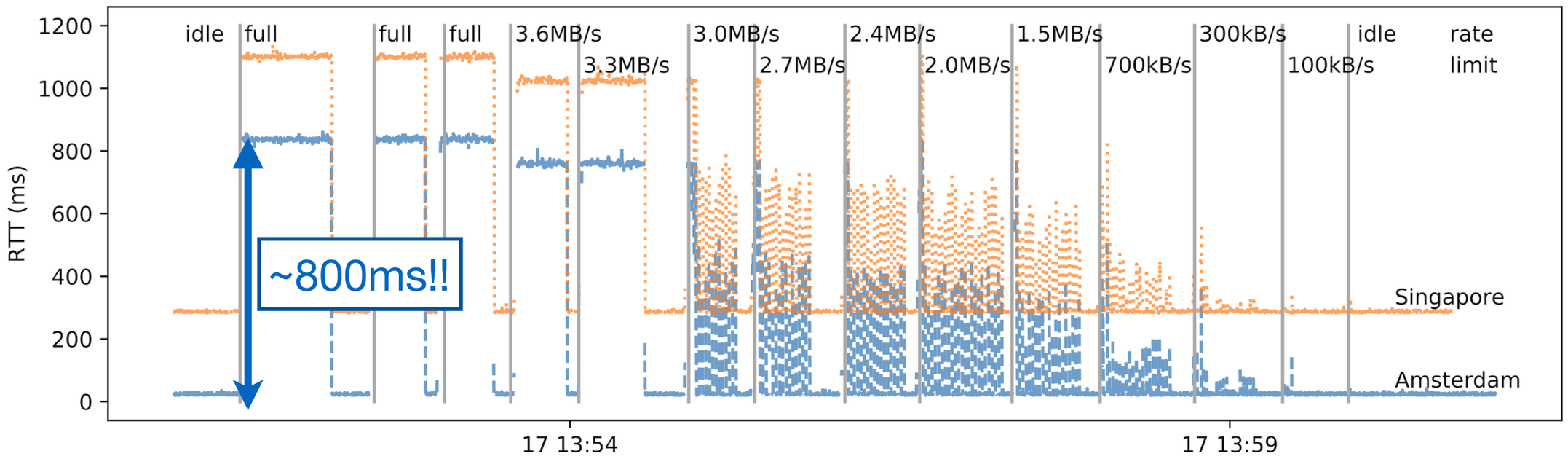
- Can a remote entity armed only with *ping* extract information about the operation of machines on my network?

$$load_{net} \propto \sum_{n=0}^{f} D_{queue_n} + \sum_{m=0}^{r} D_{queue_m}$$

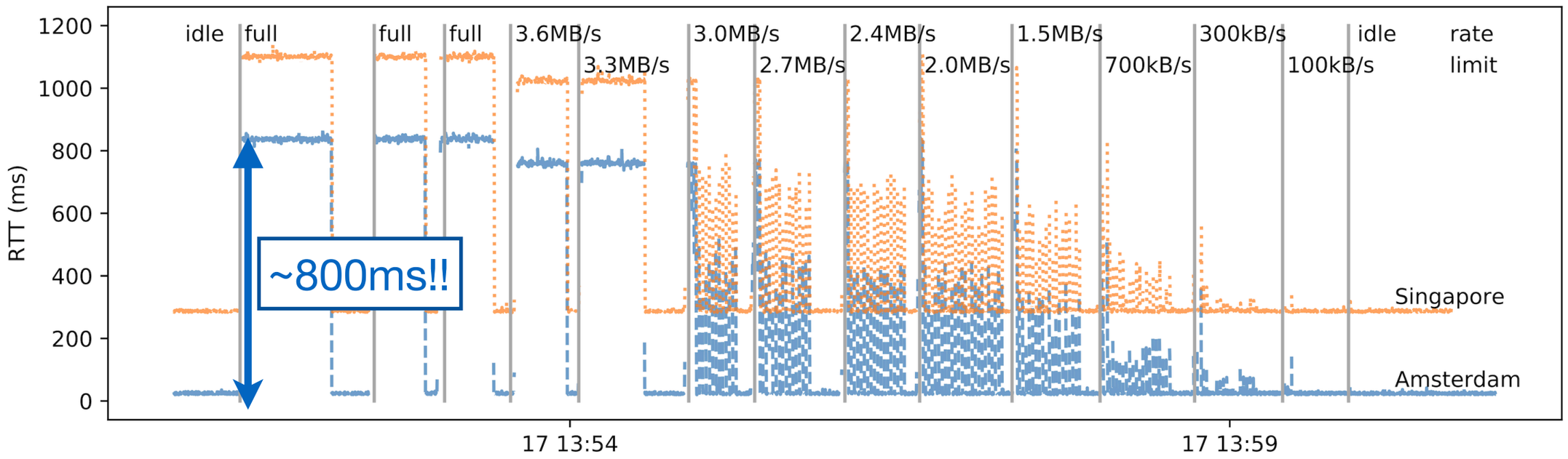
Remote Load Telemetry



Remote Load Telemetry

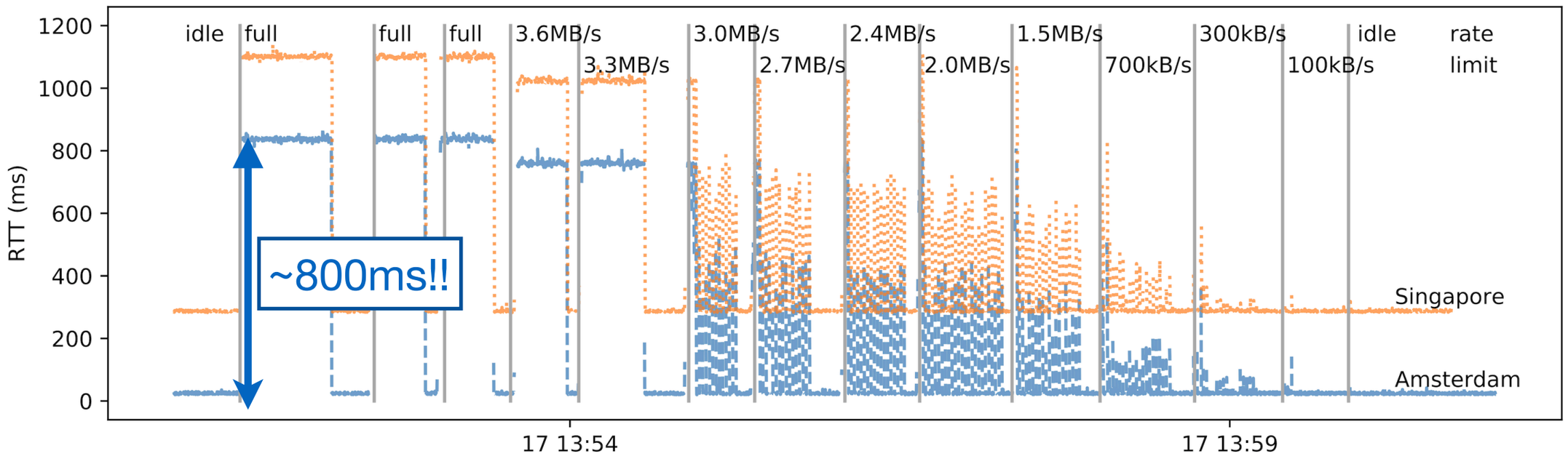


Remote Load Telemetry



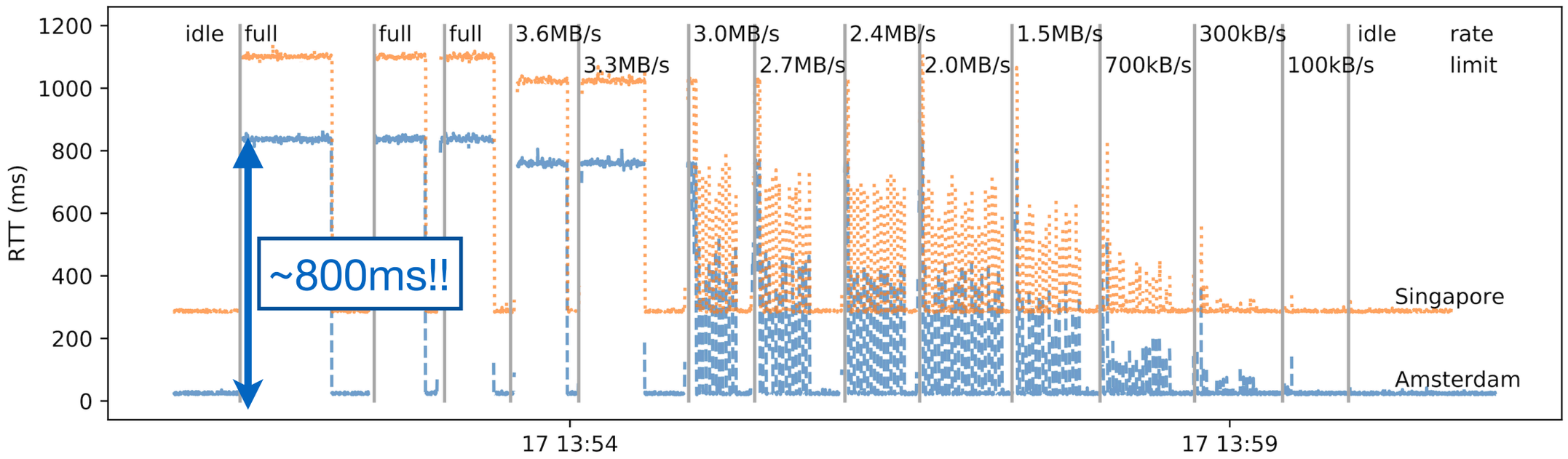
- Okay, so we know my connection sucks.

Remote Load Telemetry



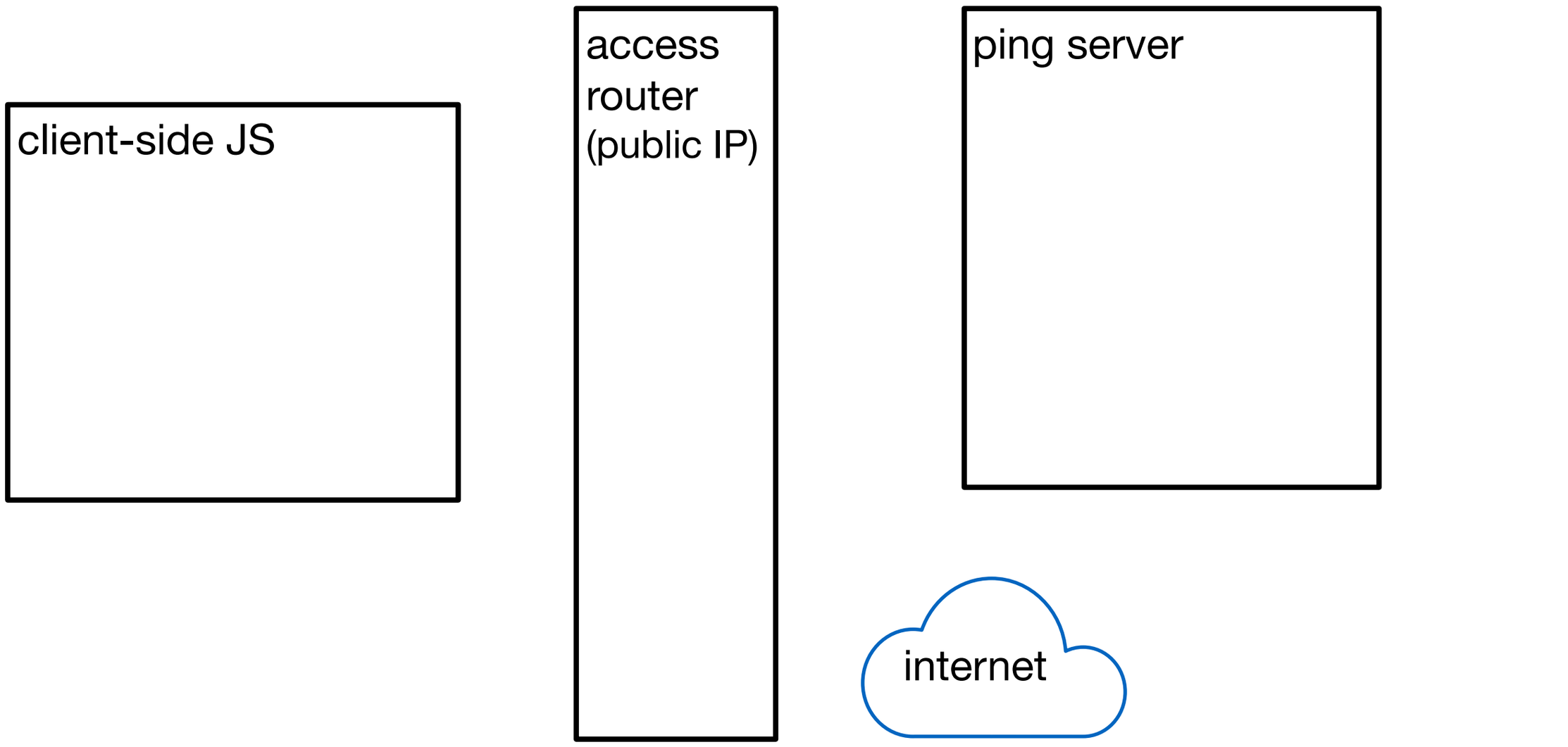
- Okay, so we know my connection sucks.
- How widespread is this phenomenon?

Remote Load Telemetry

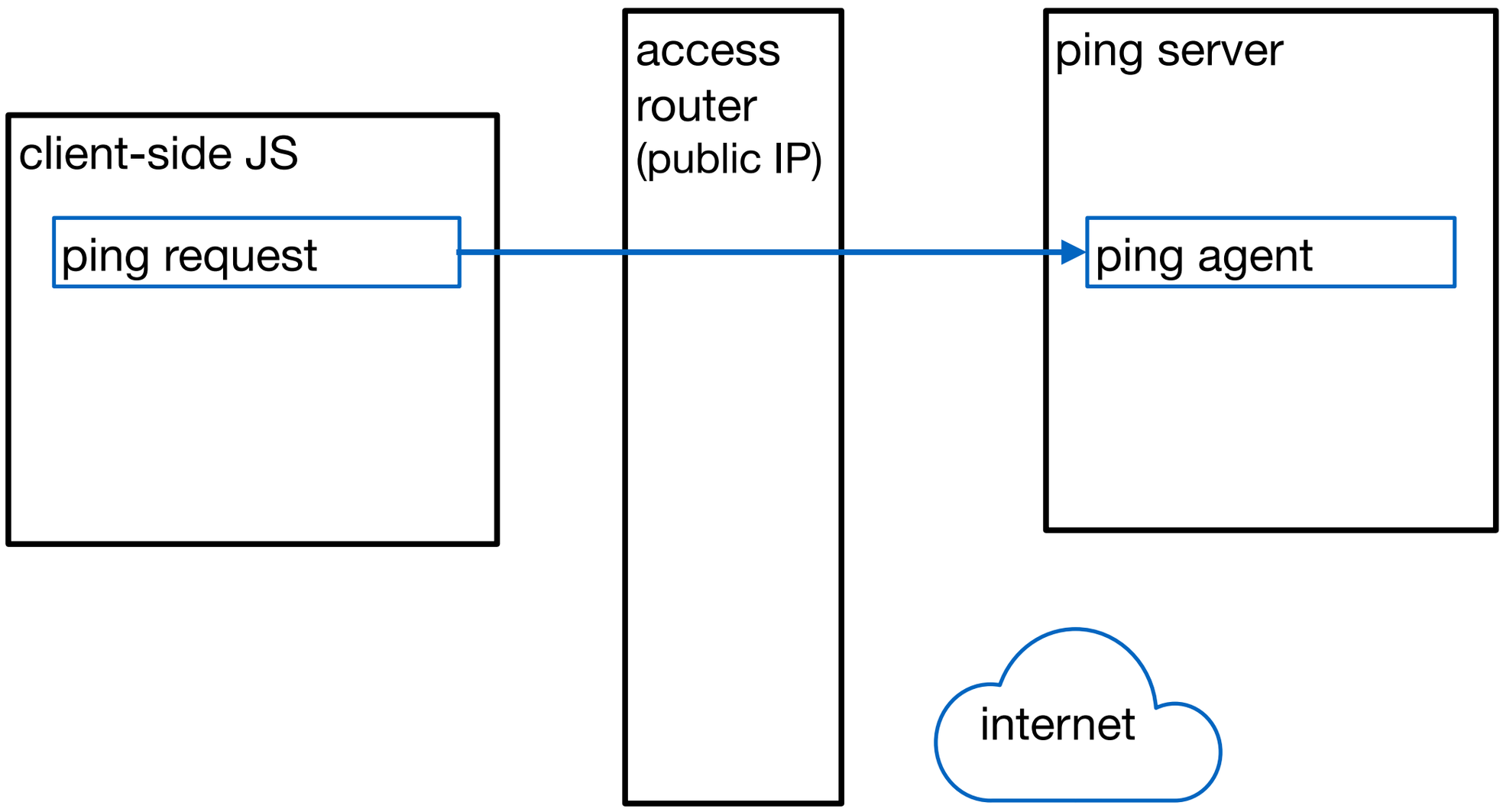


- Okay, so we know my connection sucks.
- How widespread is this phenomenon?
  - https://pingme.pto.mami-project.eu

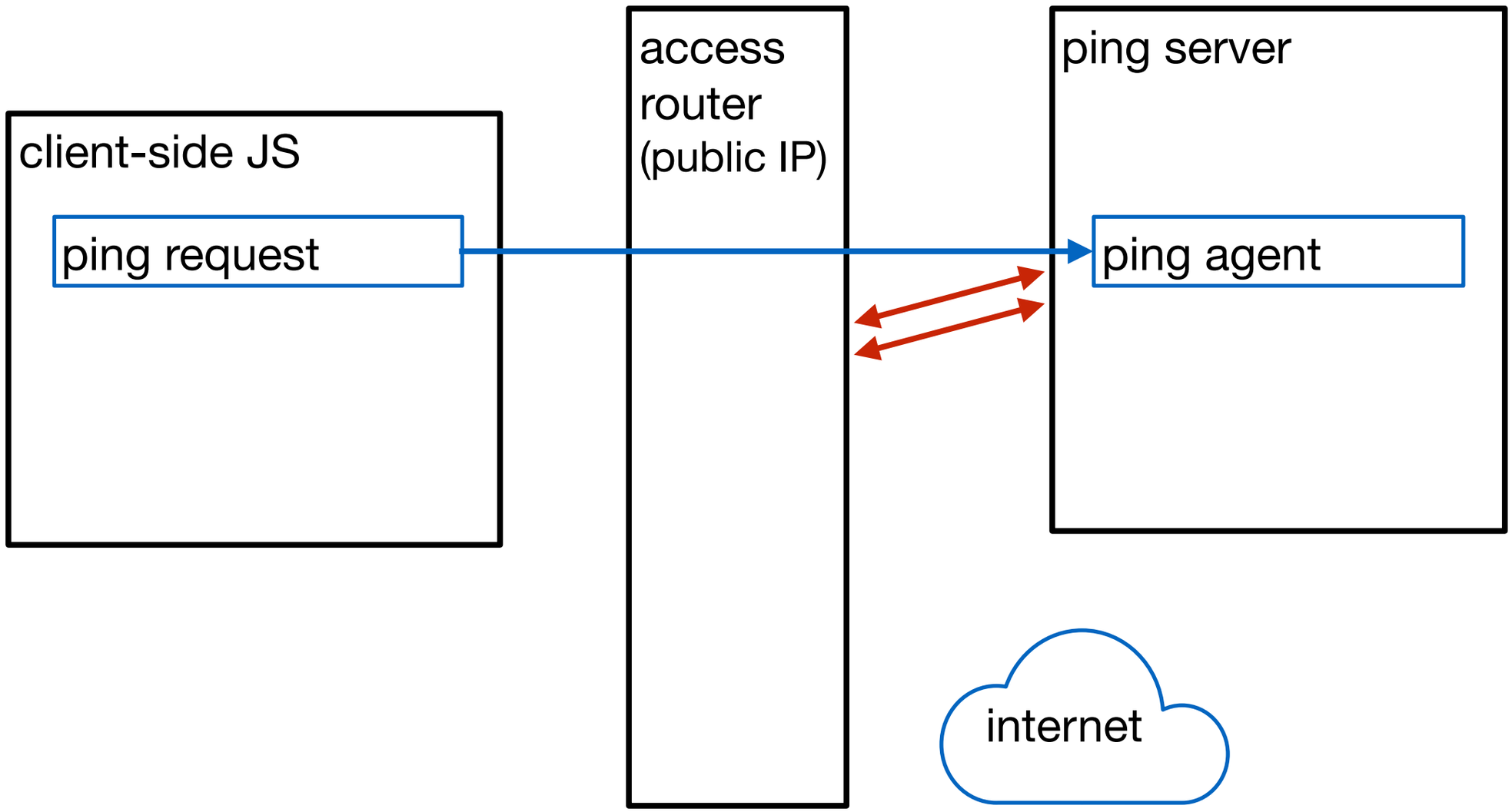
github.com/mami-project/pingme



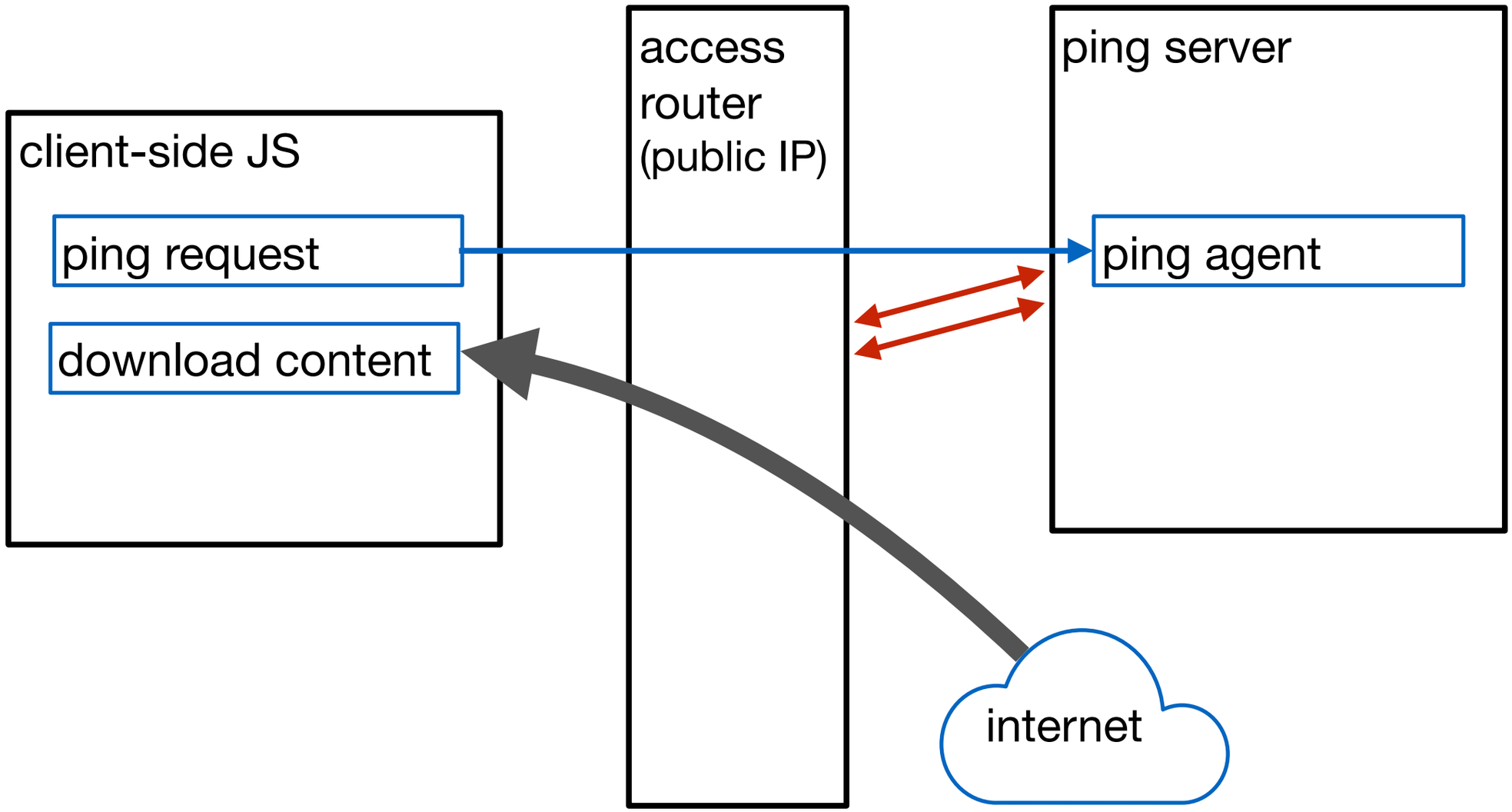
github.com/mami-project/pingme



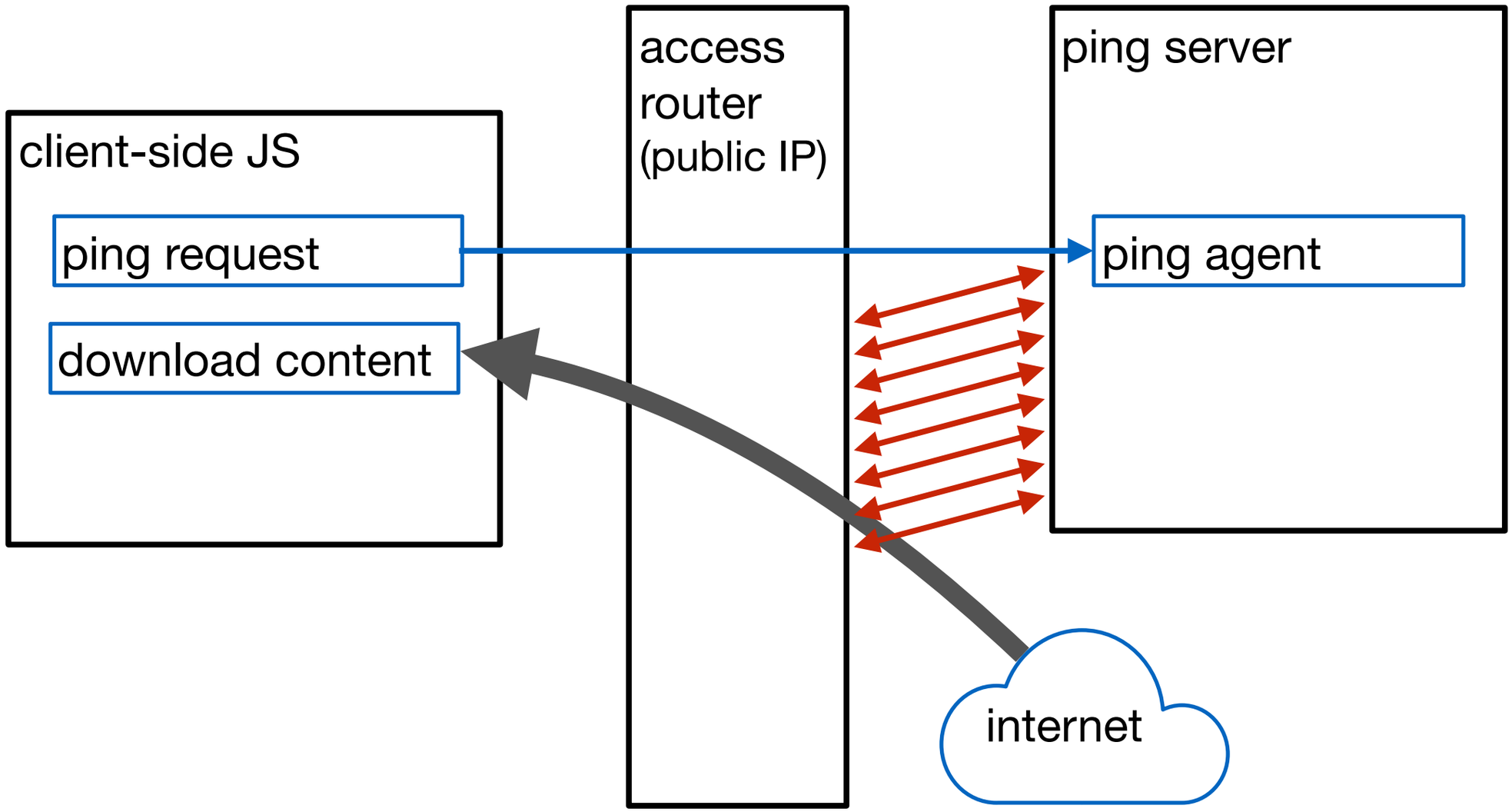
github.com/mami-project/pingme



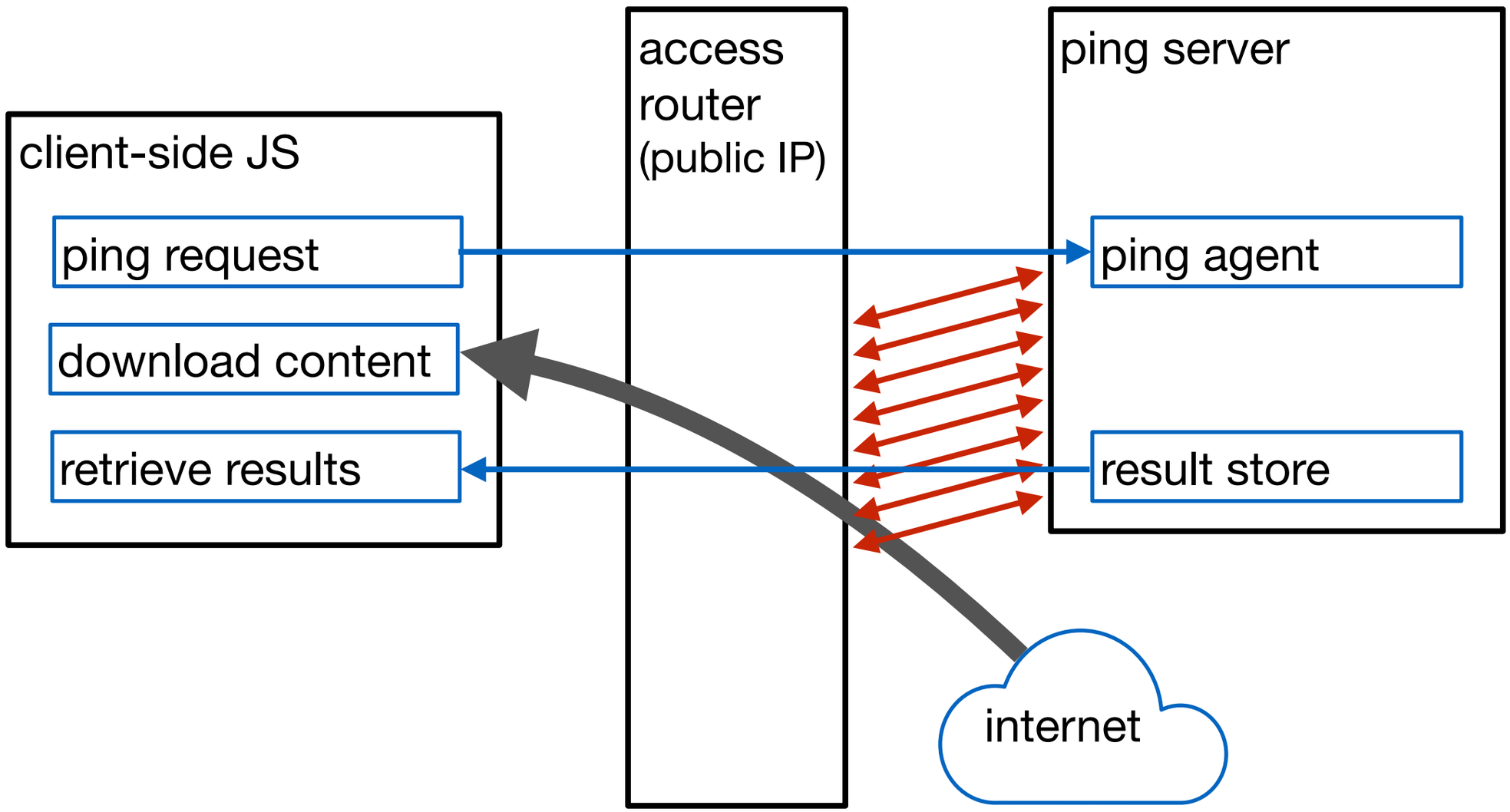
github.com/mami-project/pingme



github.com/mami-project/pingme



github.com/mami-project/pingme

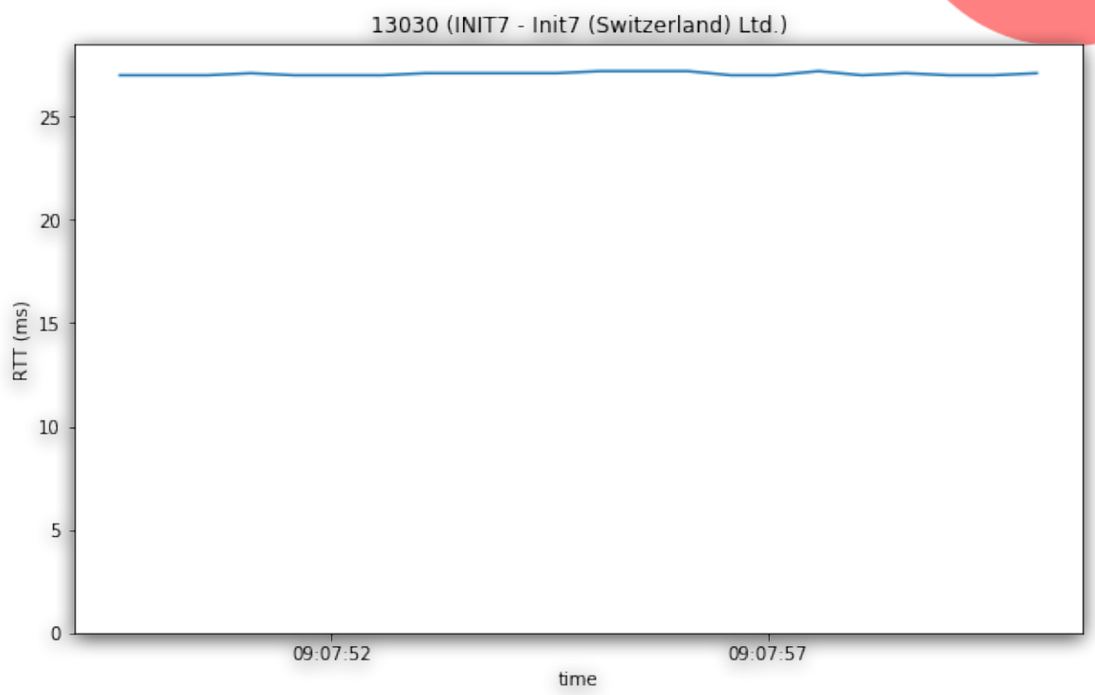


Results

- 106 measurements from 66 networks
  - 33 (50%) networks always block ICMP
    - (7/8 definitely-mobile networks block ICMP)
- On 24 (33%) networks, no indication of load-dependent RTT
- Remote load telemetry might work on 9 (14%) networks

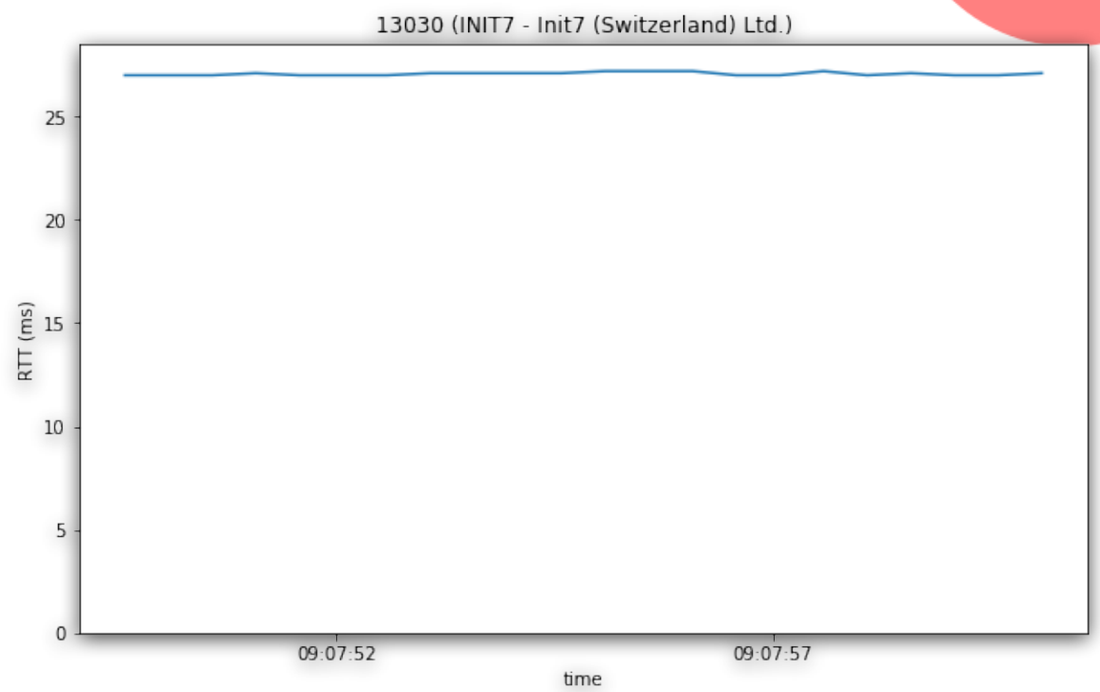
Results

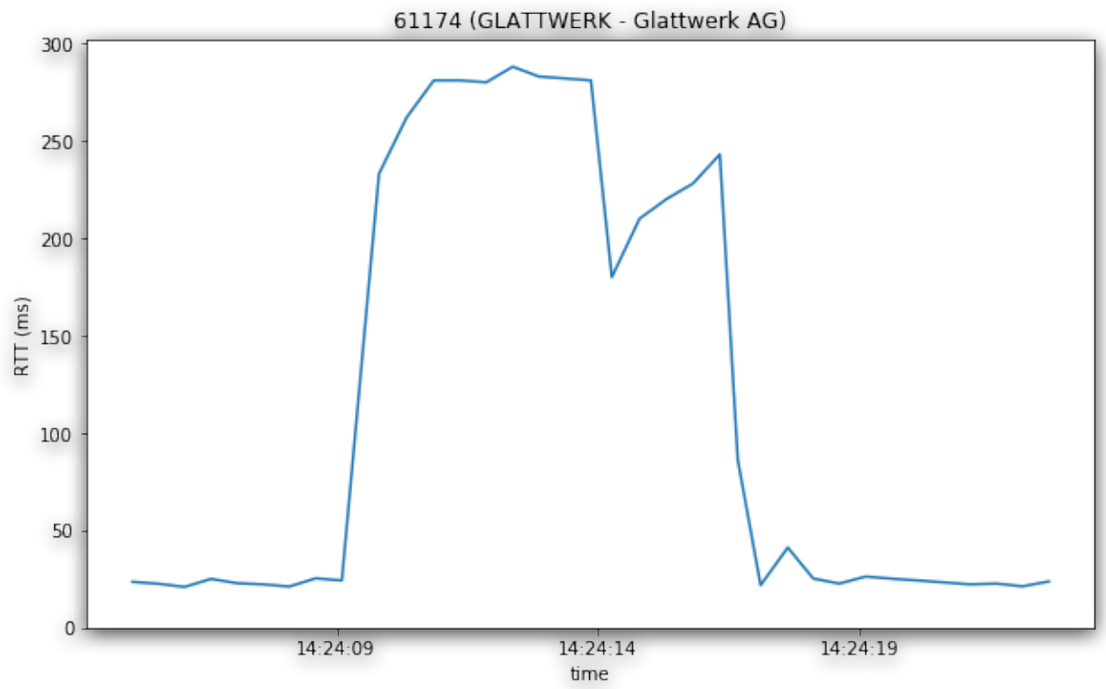
- 106 measurements from 66 networks
  - 33 (50%) networks always block ICMP
    - (7/8 definitely-mobile networks block ICMP)
- On 24 (33%) networks, no indication of load-dependent RTT
- Remote load telemetry might work on 9 (14%) networks



Results

- 106 measurements from 66 networks
  - 33 (50%) networks always block ICMP
    - (7/8 definitely-mobile networks block ICMP)
- On 24 (33%) networks, no indication of load-dependent RTT
- Remote load telemetry might work on 9 (14%) networks





Recommendations

- Remote load telemetry allows anyone who can ping you to measure your network activity.
  - Why this is bad is left as an exercise to the audience.

- Good advice: de-bloat ***all*** the buffers, deploy AQM/ECN.

- Bad advice: roll out CGN everywhere, block ICMP.