OneProbe: Measuring network path quality with TCP data-packet pairs

Rocky K. C. Chang

Internet Infrastructure and Security Group

The Hong Kong Polytechnic University

11 February 2011

ISMA 2011 AIMS-3

Our group

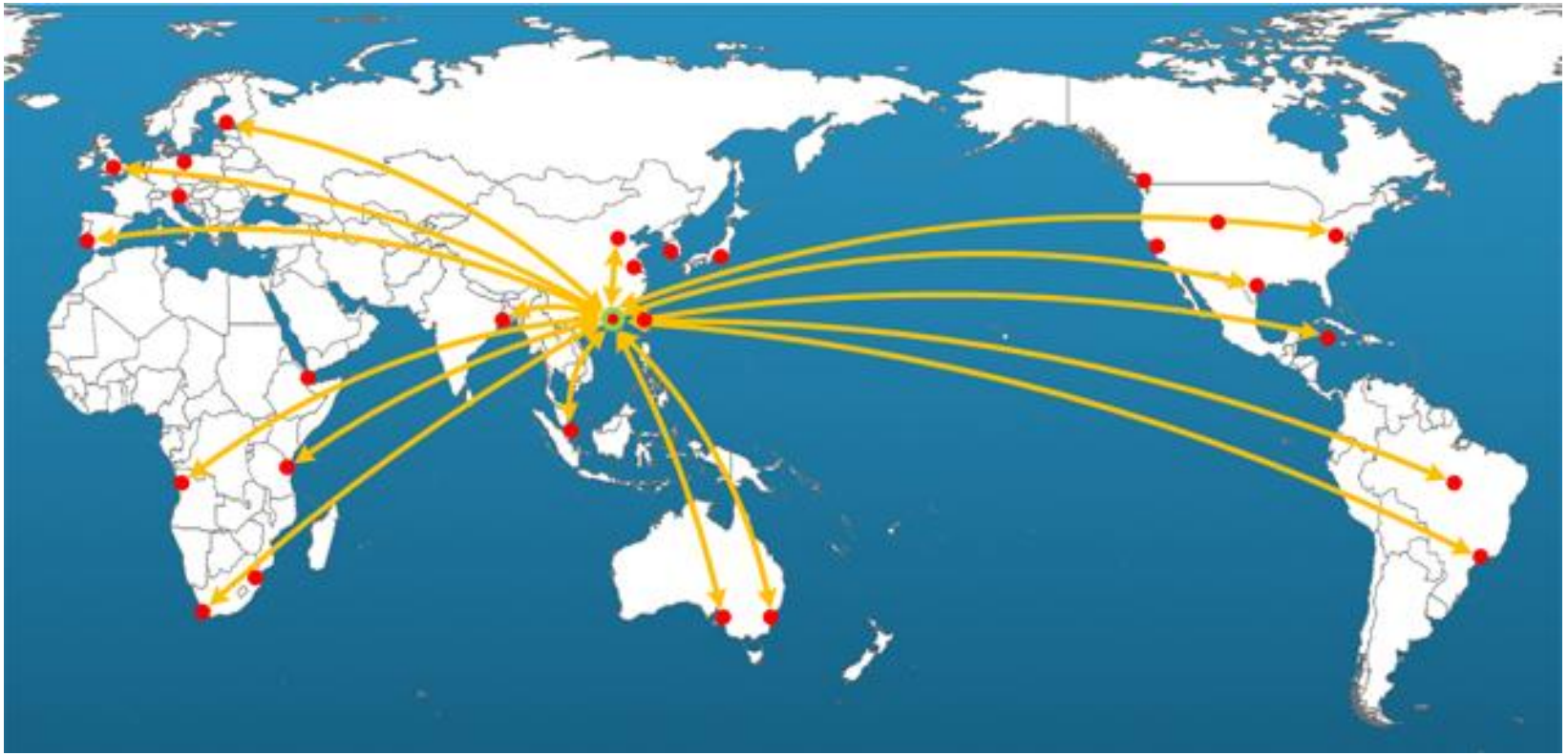
- Active measurement
  - **Non-cooperative** path-quality measurement methodologies
    - OneProbe (RTT, loss, reordering), capacity measurement, loss-pair measurement, traceroute analysis
  - Applications
    - Longitudinal analysis of network evolution, collaborative diagnosis of routing and performance problems, impact analysis of submarine cable faults, …
- Activities
  - Publications, research proposals, professional services
  - Work with HARNET, ISPs, data centers, ….
  - Plan to work with other groups, including CERNET in China

Outline

1. Path-quality measurement methodologies
2. Applications
   - Cooperative network measurement (a demo)
   - An impact analysis of a submarine cable fault
3. Conclusions and future works

1. Path-quality measurement

Measuring e2e network paths



Active measurement models

- Controlling both endpoints
  - E.g., one-way delay, OWAMP, TWAMP
- Controlling one endpoint (non-cooperative measurement)
  - Using/hacking existing protocols
  - E.g., ping, tulip, sting …
- Controlling zero endpoint
  - E.g., King

Active measurement models

- Controlling both endpoints
  - E.g., one-way delay, OWAMP, TWAMP
- Controlling one endpoint (non-cooperative measurement)
  - Using/hacking existing protocols
  - E.g., ping, tulip, sting …
- Controlling zero endpoint
  - E.g., King

(Invalid) assumptions

- Control-path quality = data-path quality
  - ICMP, TCP SYN, TCP RST
- Middleboxes not an issue
  - Dropping, rate-limiting, additional latency
- No changes in systems
  - Consecutive increment of IPID (e.g., tulip)
- Sampling rate and pattern not an issue

(Invalid) assumptions

- Control-path quality = data-path quality
  - ICMP, TCP SYN, TCP RST
- Middleboxes not an issue
  - Dropping, rate-limiting, additional latency
- No changes in systems
  - Consecutive increment of IPID (e.g., tulip)
- Sampling rate and pattern not an issue

Invalid assumptions beget unreliable measurement.

Other problems in practice

- Support only one or two metrics
- Round-trip measurement
- No control over packet sizes
- Not integrated with application protocols

Other problems in practice

- Support only one or two metrics
- Round-trip measurement
- No control over packet sizes
- Not integrated with application protocols

Practical issues stifle deployment.

Our design principles

- Use normal data packet to measure data-path quality.
- Use normal and basic data transmission mechanisms
- Integrated into normal application sessions.

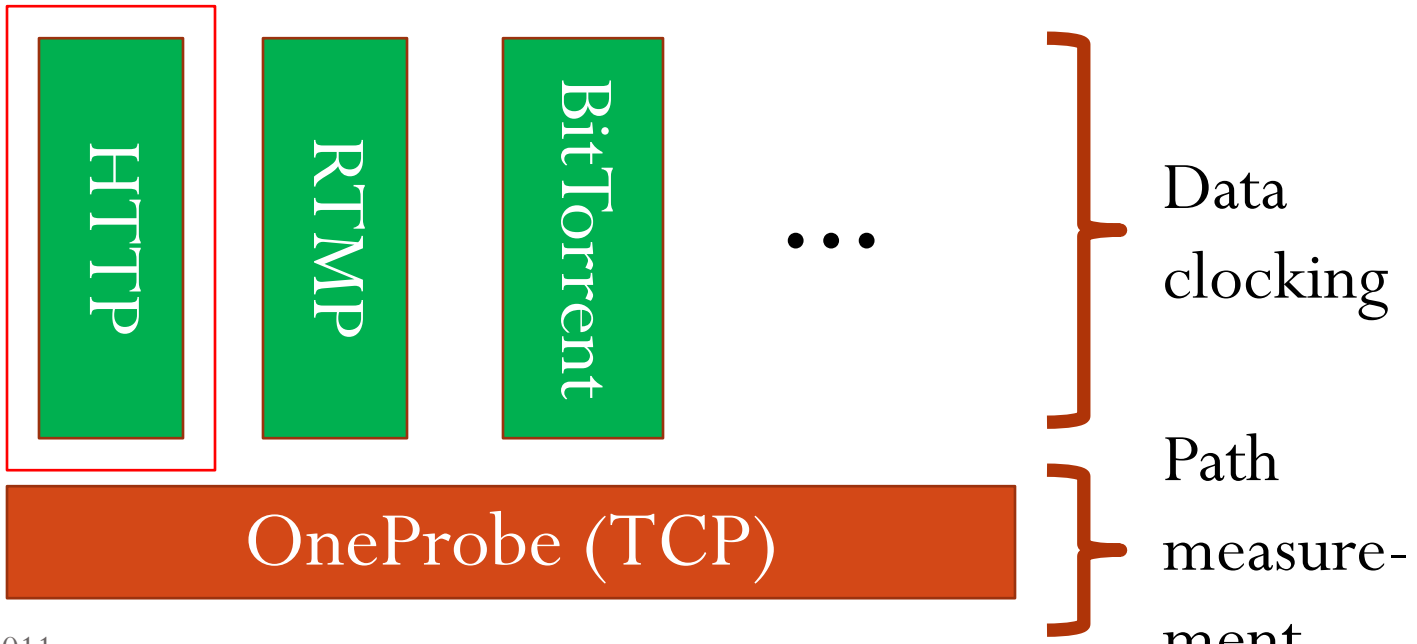
Our design principles

- Use normal data packet to measure data-path quality.
- Use normal and basic data transmission mechanisms
- Integrated into normal application sessions.

Reliable measurement

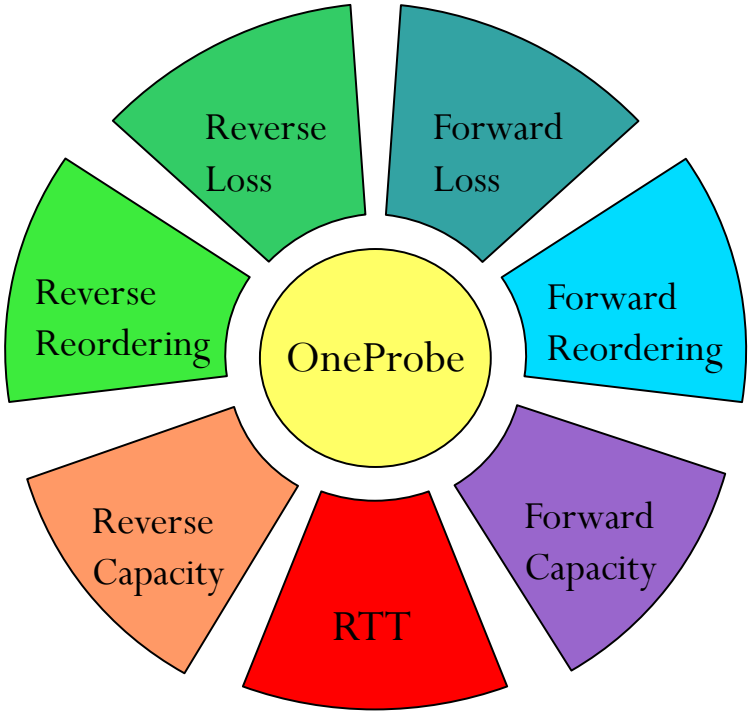
HTTP/OneProbe

- Use normal TCP data packet to measure data-path quality.
- Use normal and basic TCP data transmission mechanisms specified in RFC 793.
- Integrated into normal HTTP application sessions.

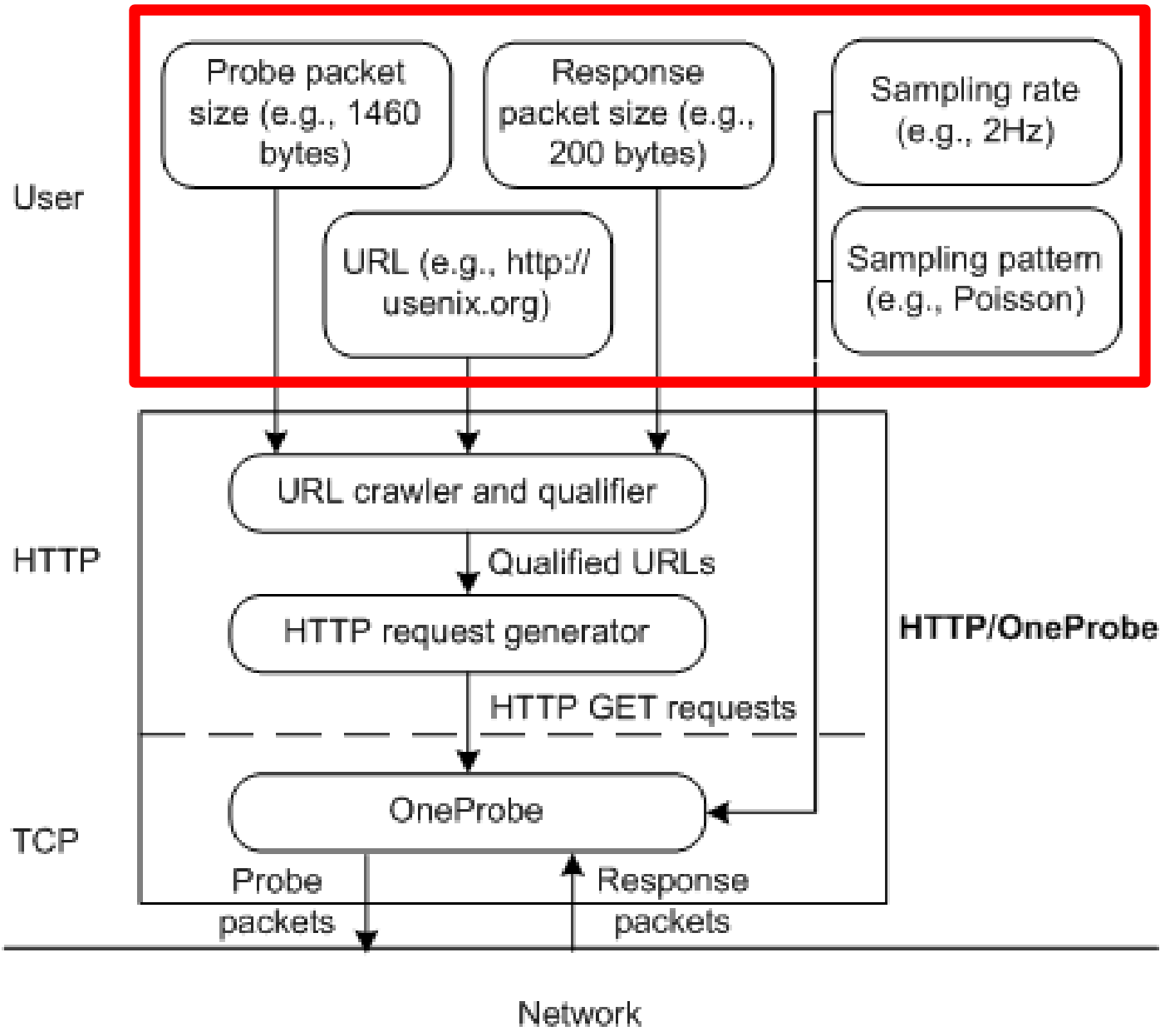


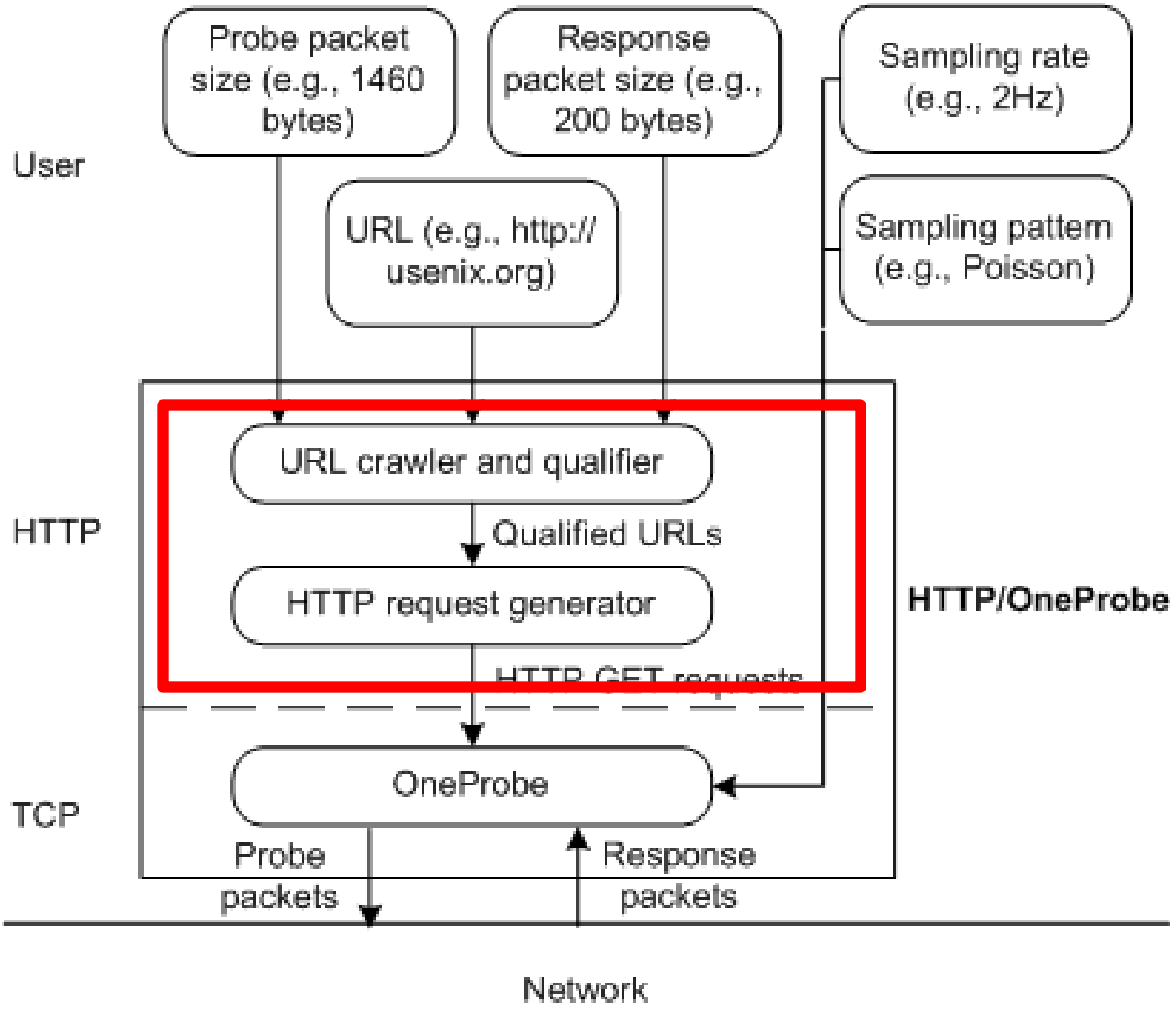
What does HTTP/OneProbe offer?

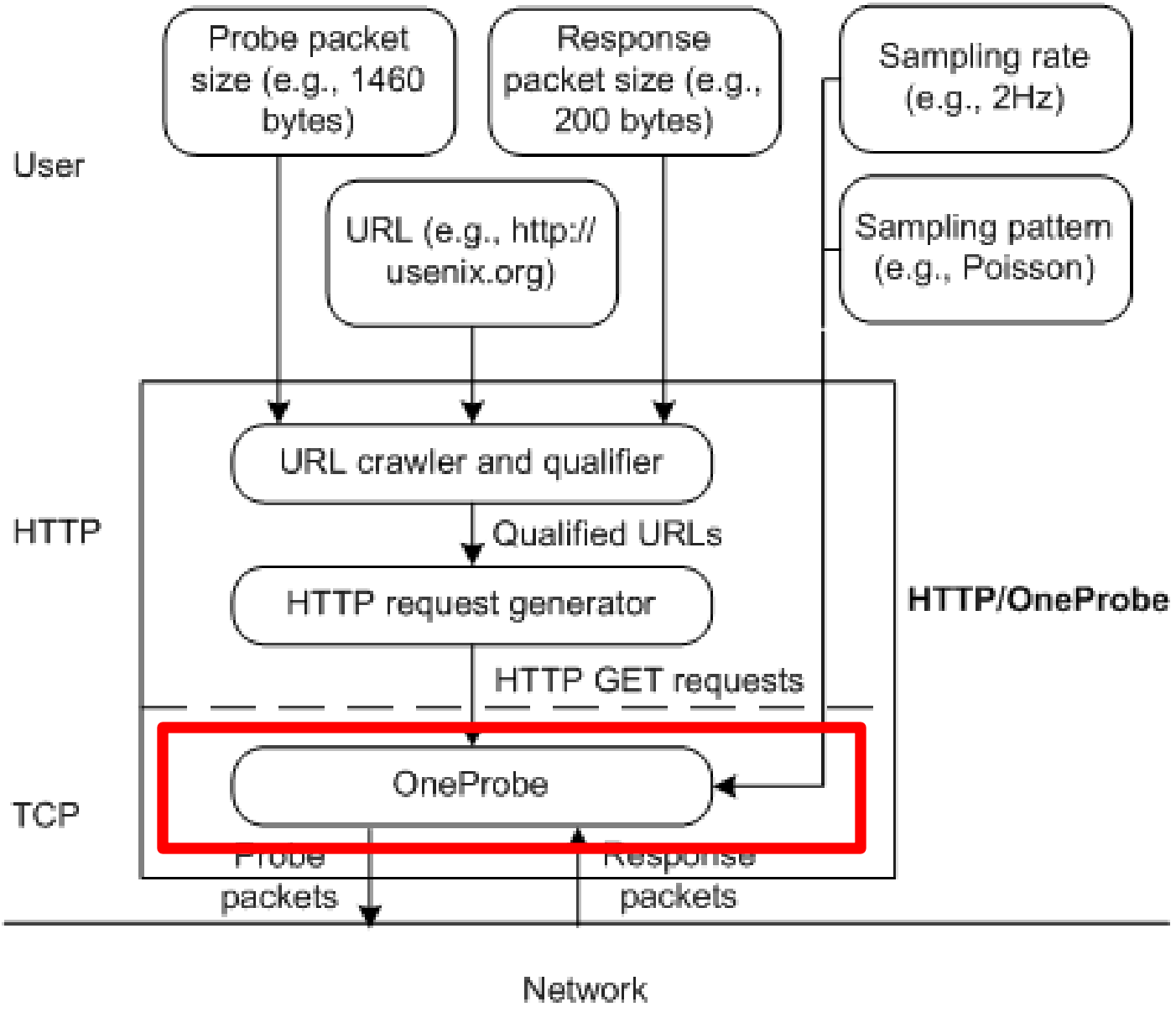
- Continuous path monitoring in an HTTP session (stateful measurement)
- All in one:
  - Round-trip time
  - Loss rate (uni-directional)
  - Reordering rate (uni-directional)
  - Capacity (uni-directional)
  - Loss-pair analysis
  - …



- "Design and Implementation of TCP Data Probes for Reliable and Metric-Rich Network Path Monitoring," *Proc. USENIX Annual Tech. Conf.*, June 2009.

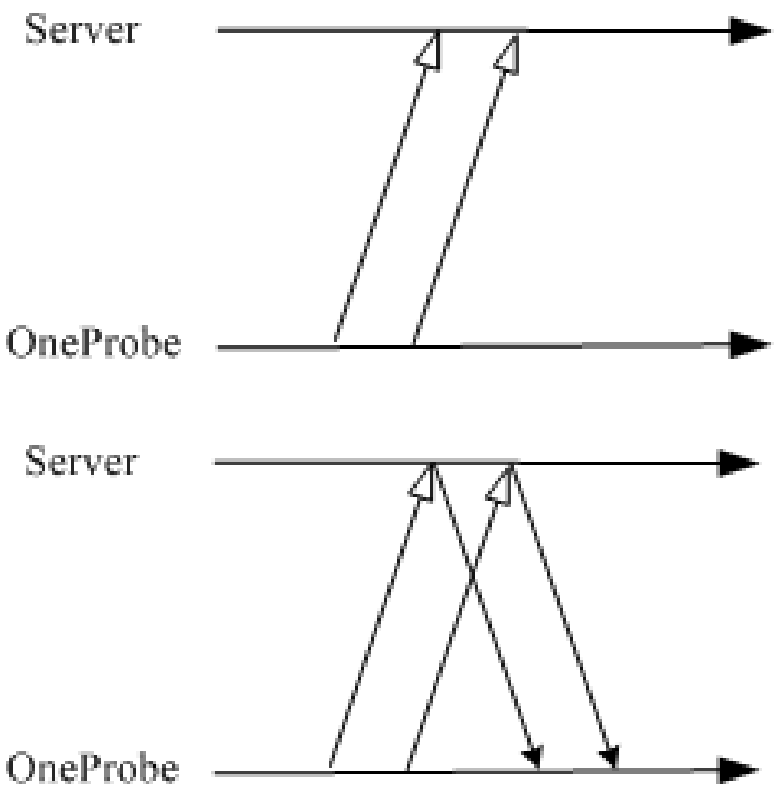

Probe packet size (e.g., 1460 bytes)
Response packet size (e.g., 200 bytes)
Sampling rate (e.g., 2Hz)
User
URL (e.g., http://usenix.org)
Sampling pattern (e.g., Poisson)
URL crawler and qualifier
HTTP
Qualified URLs
HTTP request generator
HTTP/OneProbe
HTTP GET requests
OneProbe
TCP
Probe packets
Response packets
Network


Probe packet size (e.g., 1460 bytes)
Response packet size (e.g., 200 bytes)
Sampling rate (e.g., 2Hz)
User
URL (e.g., http:// usenix.org)
Sampling pattern (e.g., Poisson)
URL crawler and qualifier
HTTP
Qualified URLs
HTTP request generator
HTTP/OneProbe
HTTP GET requests
OneProbe
TCP
Probe packets
Response packets
Network

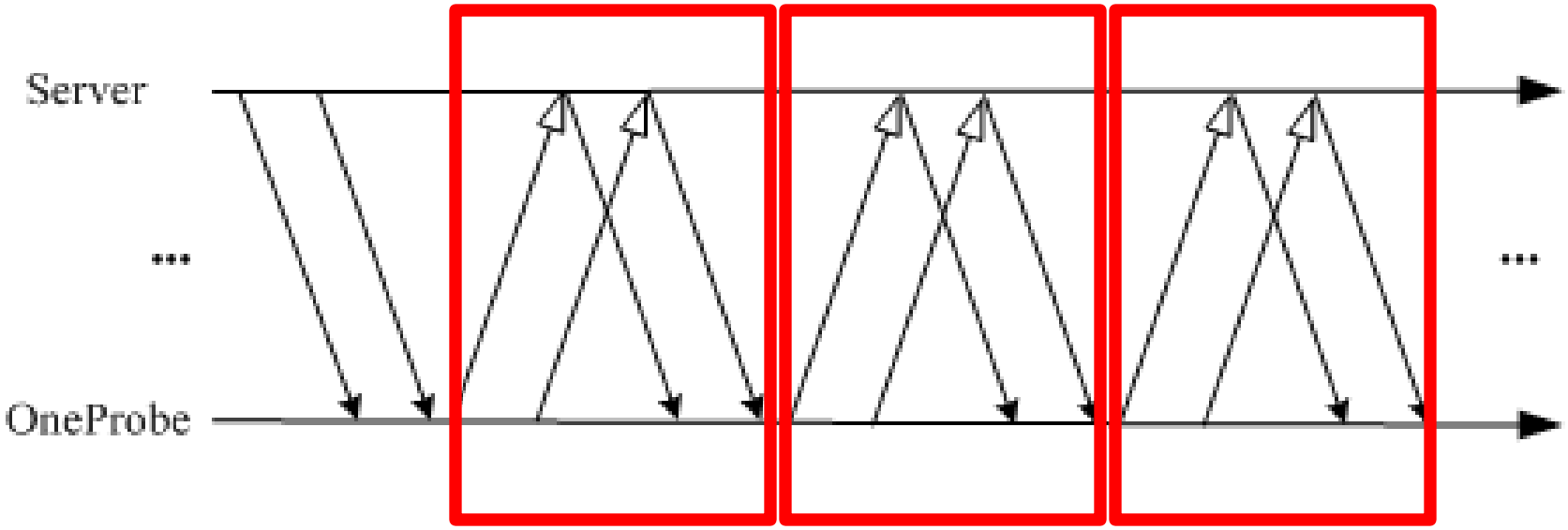

Probe packet size (e.g., 1460 bytes)
Response packet size (e.g., 200 bytes)
Sampling rate (e.g., 2Hz)
User
URL (e.g., http:// usenix.org)
Sampling pattern (e.g., Poisson)
URL crawler and qualifier
HTTP
Qualified URLs
HTTP request generator
HTTP/OneProbe
HTTP GET requests
OneProbe
TCP
packets
packets
Network

OneProbe: the probe design

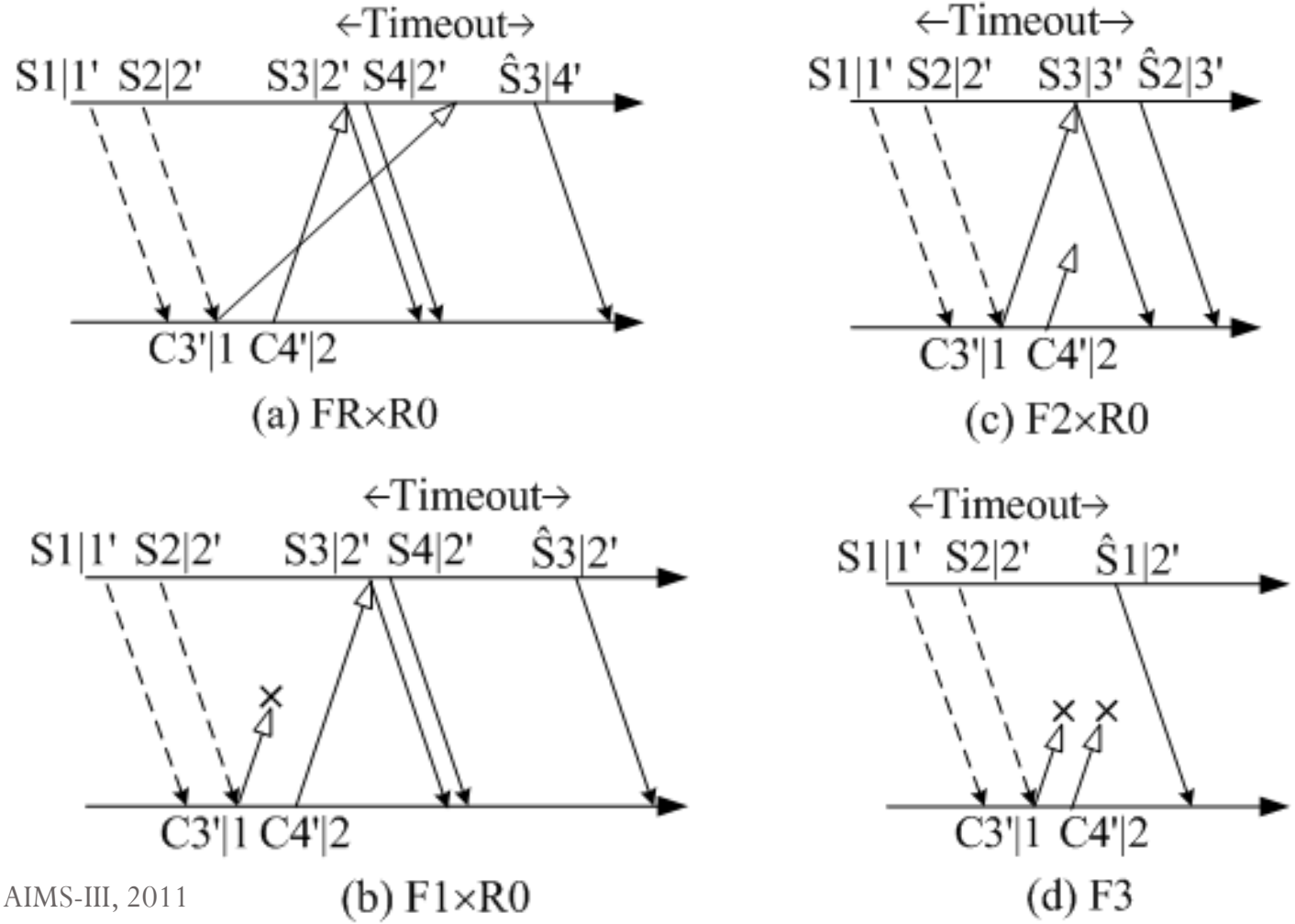
- Send two back-to-back probe data packets.
  - Capacity measurement
  - Packet reordering
  - Determine which packet is lost.
- Similarly for the response packets
  - Each probe packet elicits a response packet



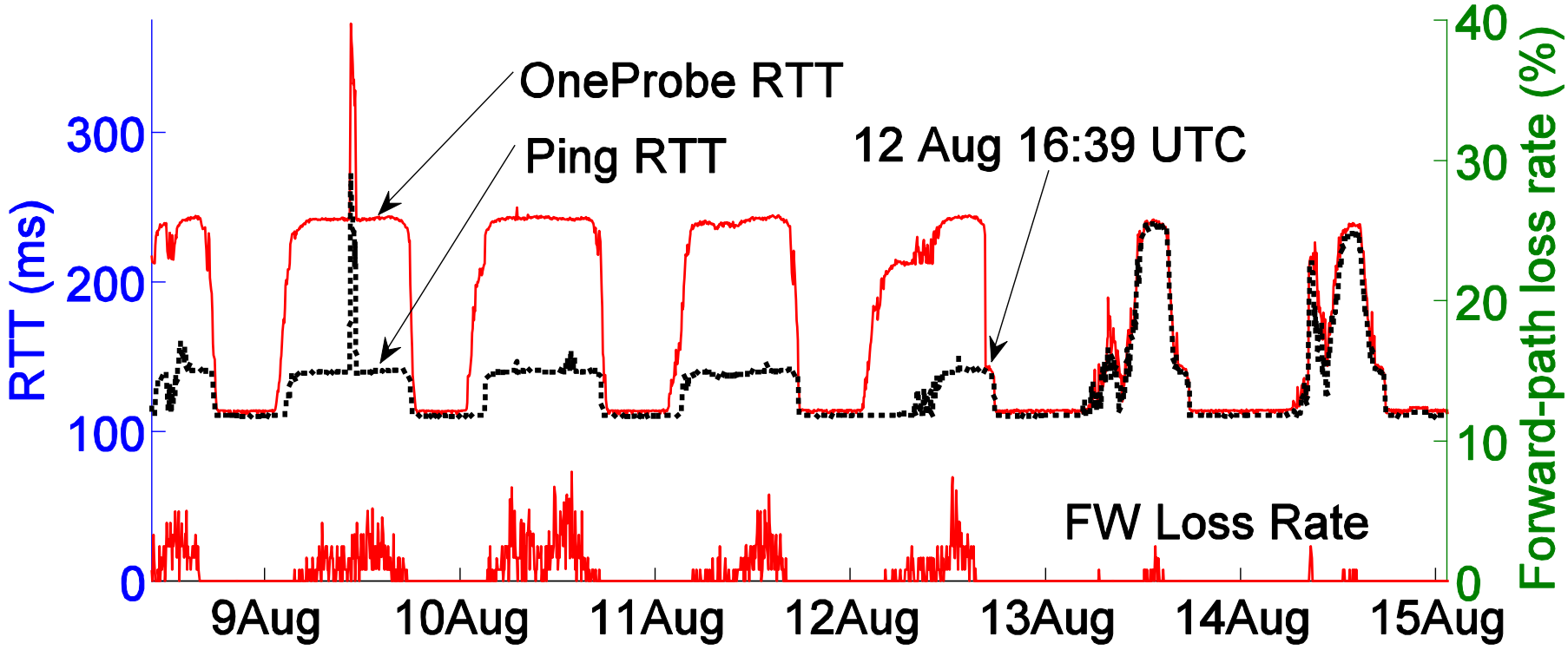
OneProbe: Bootstrapping and continuous monitoring



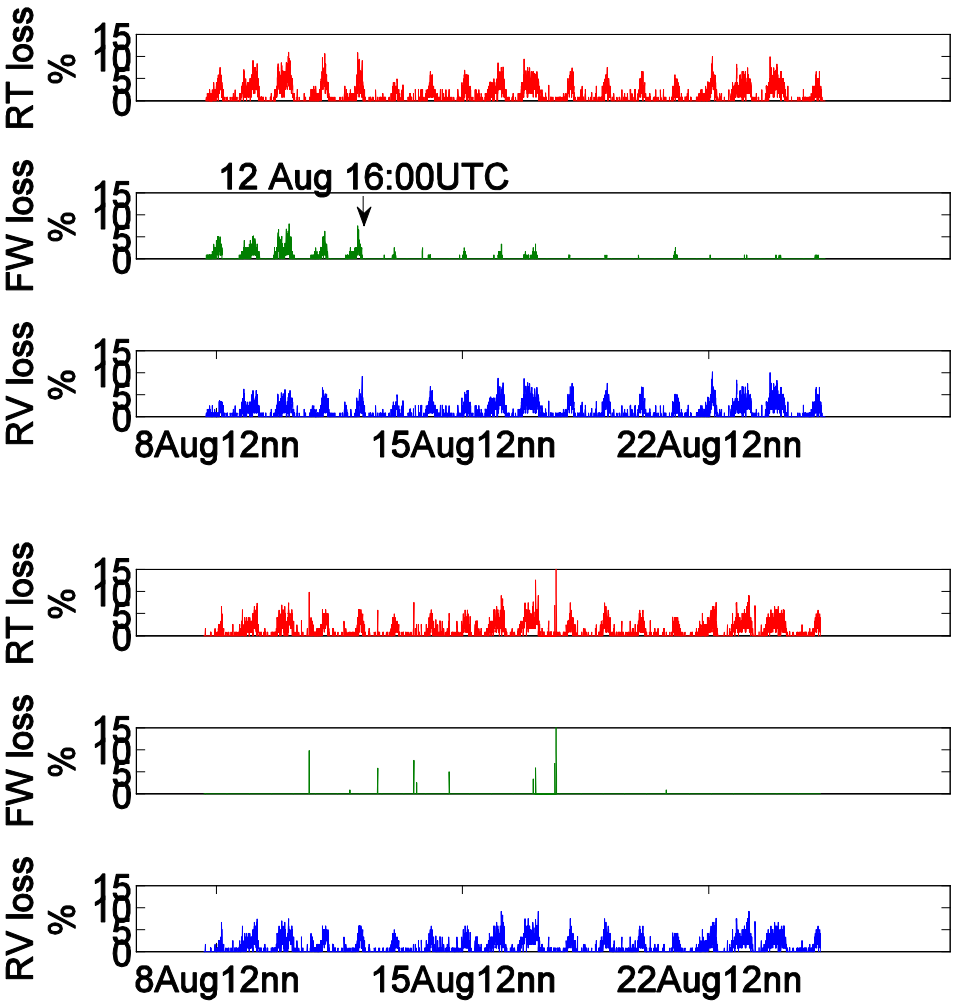
OneProbe: Loss and reordering measurement via response diversity

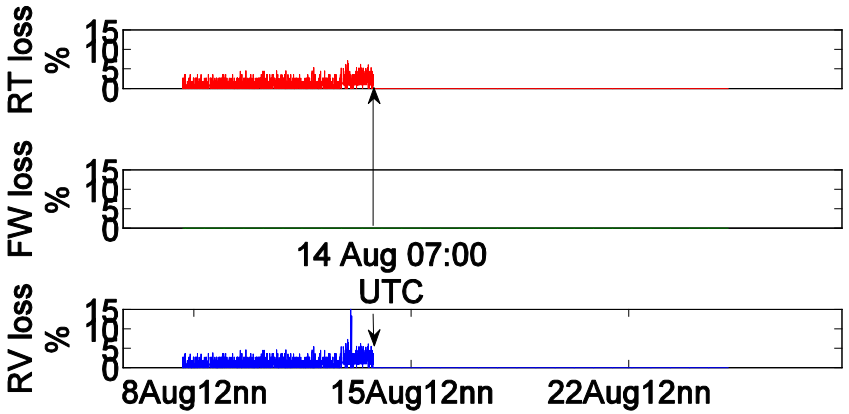


Discrepancy between ping RTT and OneProbe RTT

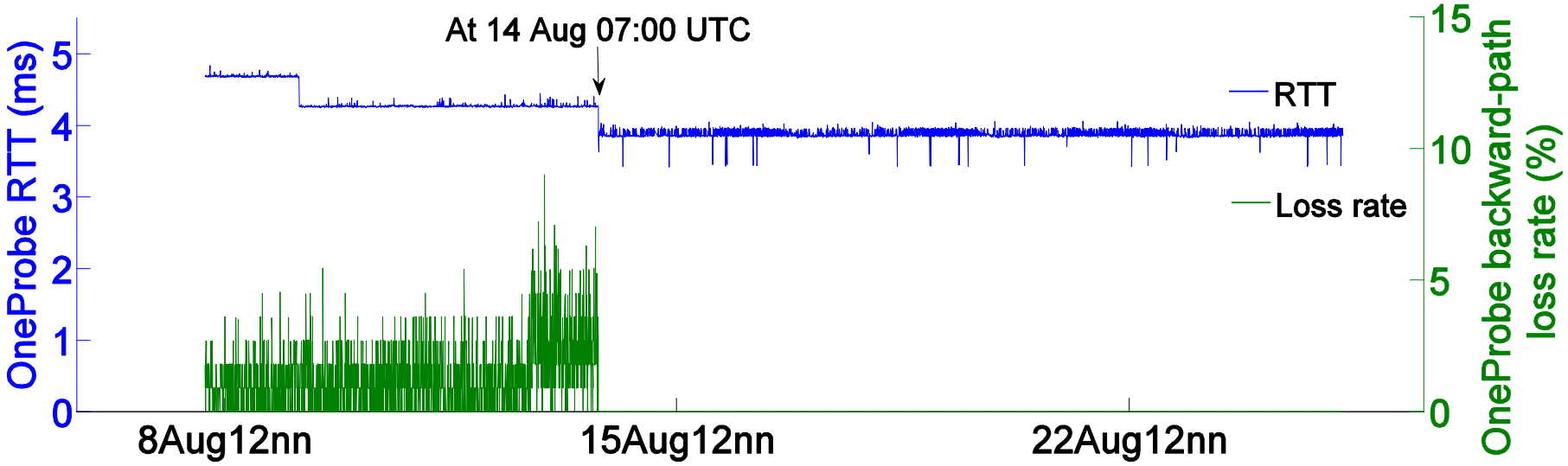


Highly asymmetric loss rates



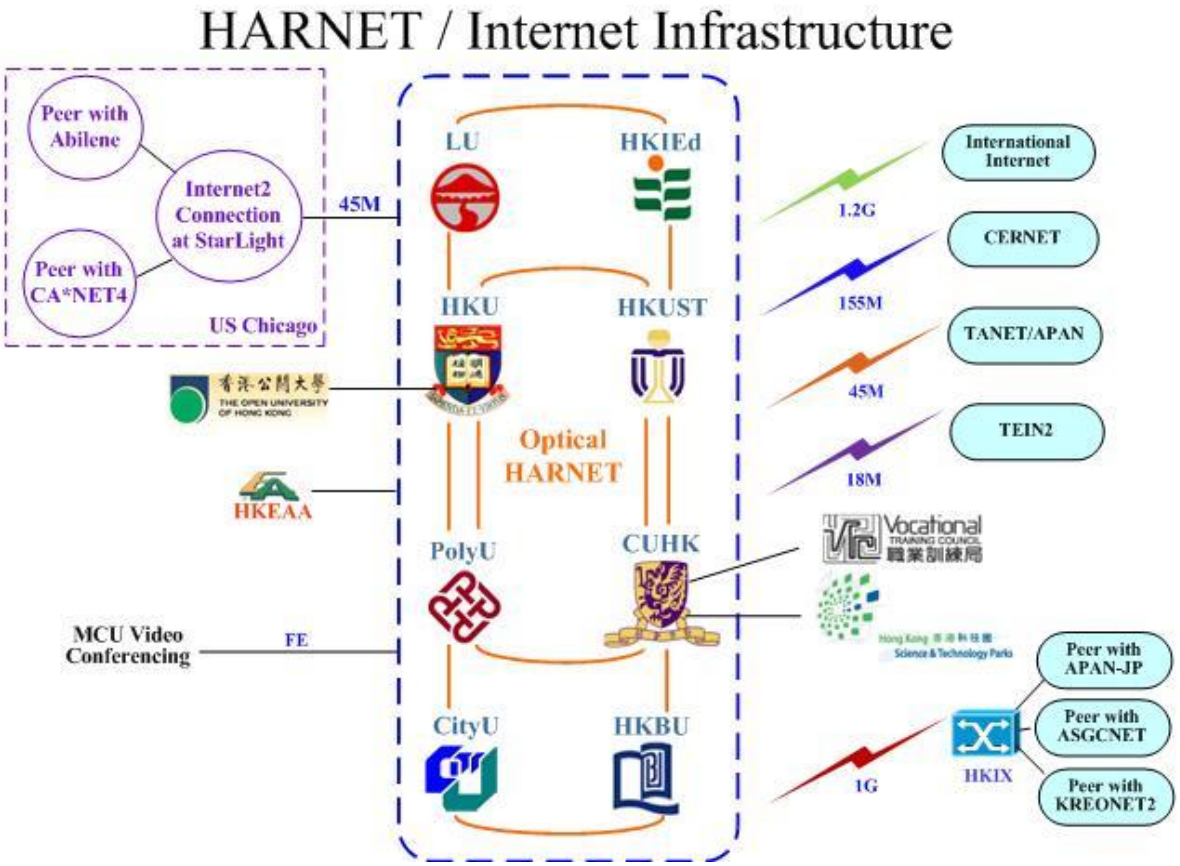


Impact of configuration changes



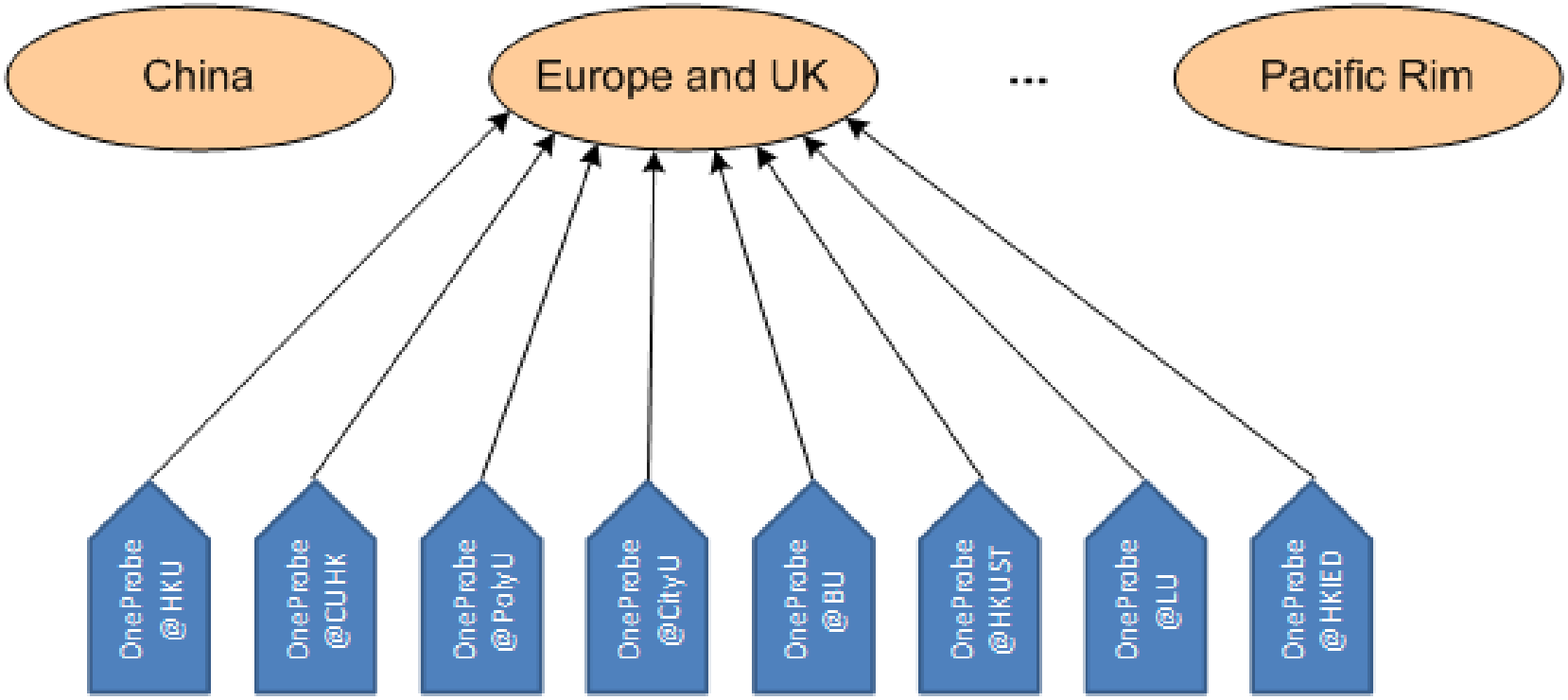
2.1 Application: Collaborative path-quality measurement

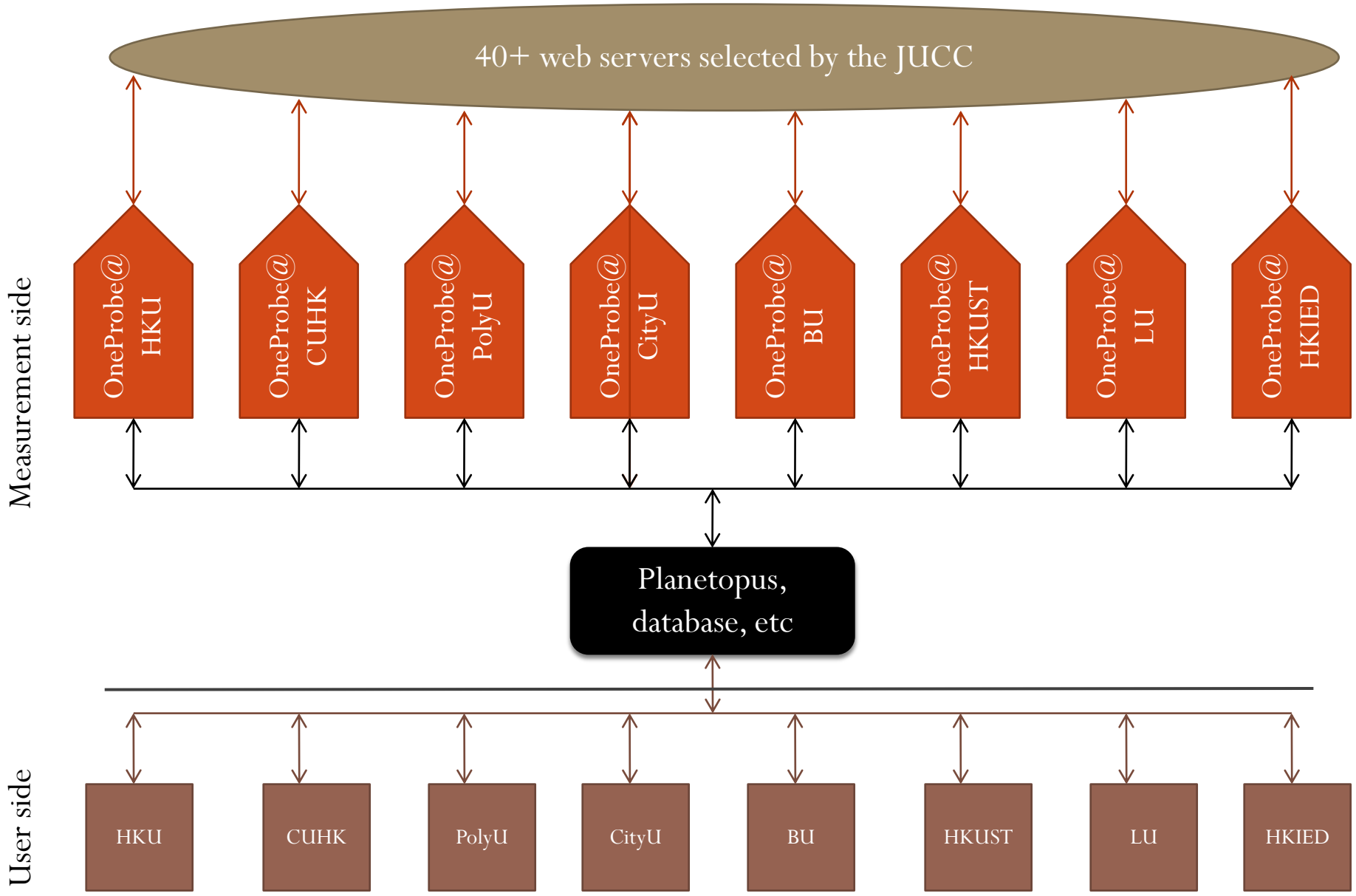
HARNET measurement (since 1 Jan 2009)

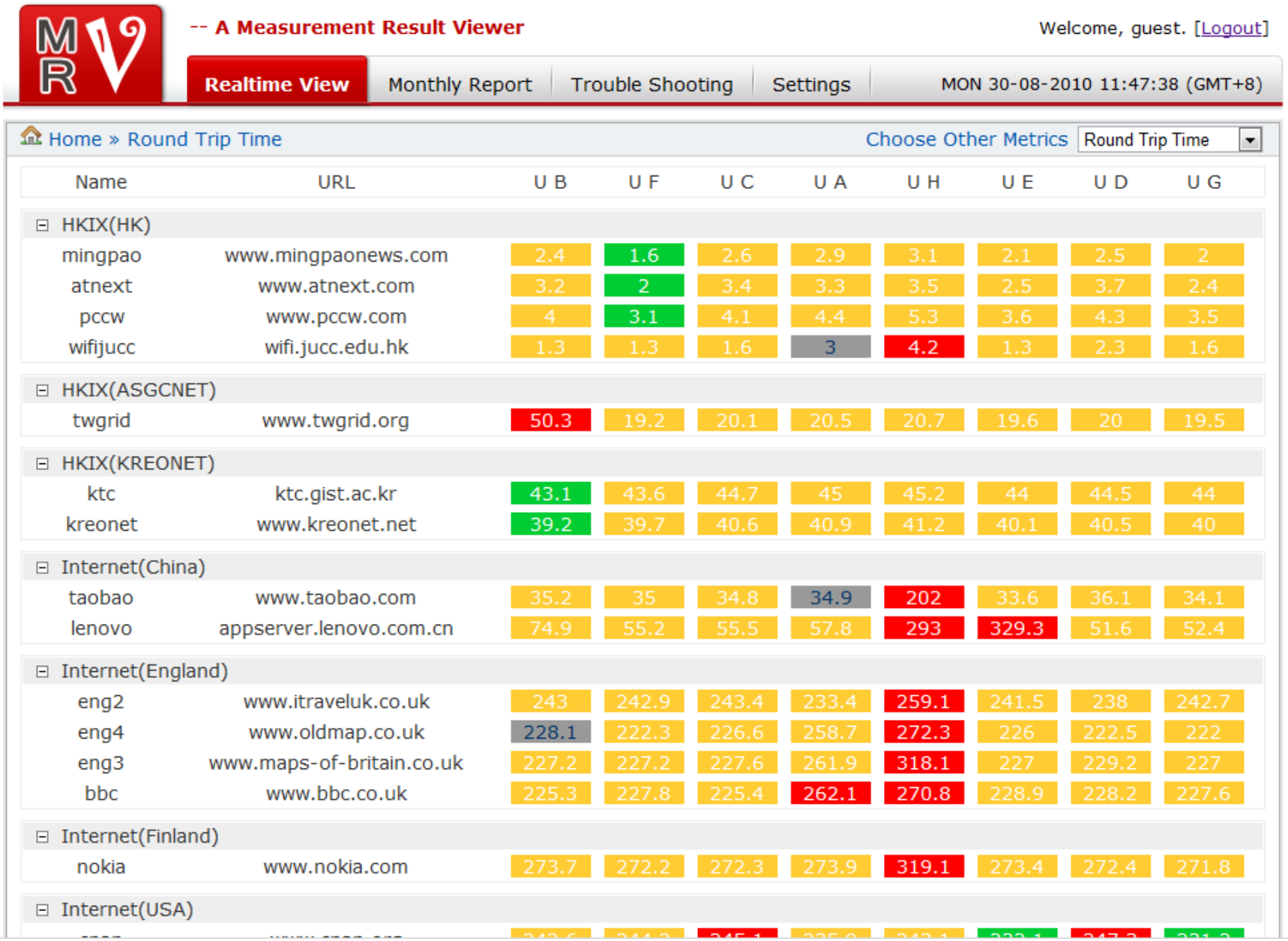


Running OneProbe at the 8 Us

- 24x365 probing of the paths to 40+ websites



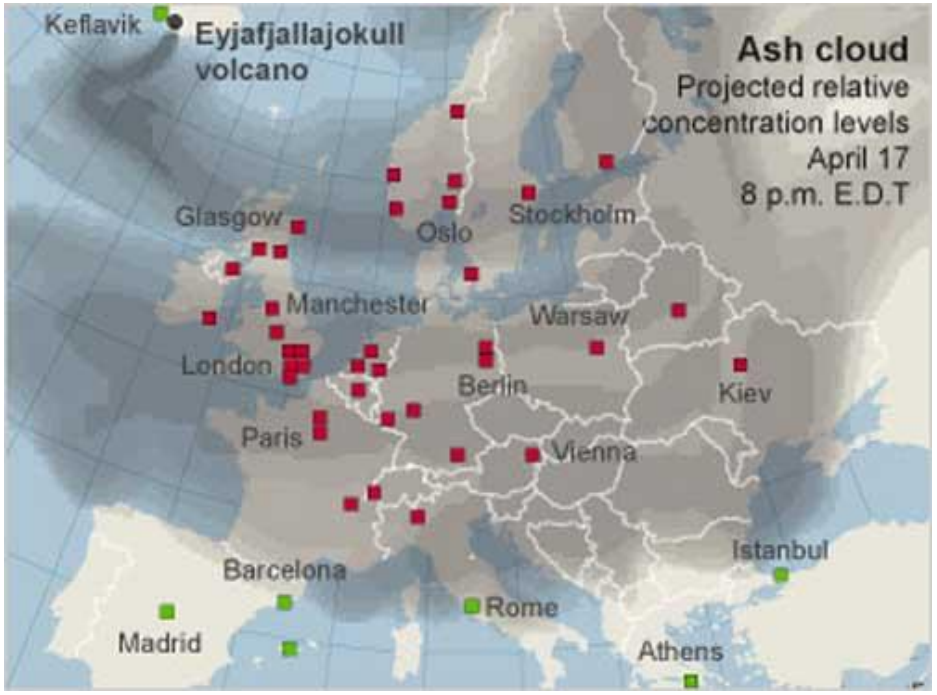

40+ web servers selected by the JUCC
Measurement side
OneProbe@ HKU
OneProbe@ CUHK
OneProbe@ PolyU
OneProbe@ CityU
OneProbe@ BU
OneProbe@ HKUST
OneProbe@ LU
OneProbe@ HKIED
Planetopus, database, etc
User side
HKU
CUHK
PolyU
CityU
BU
HKUST
LU
HKIED



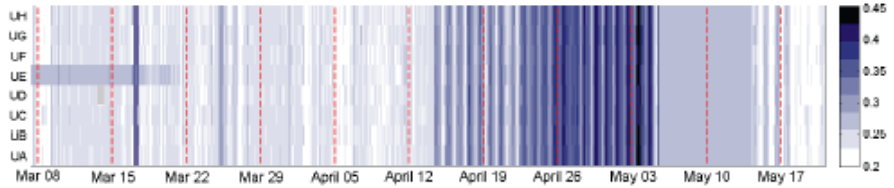
Name	URL	U B	U F	U C	U A	U H	U E	U D	U G
HKIX(HK)									
mingpao	www.mingpaonews.com	2.4	1.6	2.6	2.9	3.1	2.1	2.5	2
atnext	www.atnext.com	3.2	2	3.4	3.3	3.5	2.5	3.7	2.4
pccw	www.pccw.com	4	3.1	4.1	4.4	5.3	3.6	4.3	3.5
wifijucc	wifi.jucc.edu.hk	1.3	1.3	1.6	3	4.2	1.3	2.3	1.6
HKIX(ASGCNET)									
twgrid	www.twgrid.org	50.3	19.2	20.1	20.5	20.7	19.6	20	19.5
HKIX(KREONET)									
ktc	ktc.gist.ac.kr	43.1	43.6	44.7	45	45.2	44	44.5	44
kreonet	www.kreonet.net	39.2	39.7	40.6	40.9	41.2	40.1	40.5	40
Internet(China)									
taobao	www.taobao.com	35.2	35	34.8	34.9	202	33.6	36.1	34.1
lenovo	appserver.lenovo.com.cn	74.9	55.2	55.5	57.8	293	329.3	51.6	52.4
Internet(England)									
eng2	www.itraveluk.co.uk	243	242.9	243.4	233.4	259.1	241.5	238	242.7
eng4	www.oldmap.co.uk	228.1	222.3	226.6	258.7	272.3	226	222.5	222
eng3	www.maps-of-britain.co.uk	227.2	227.2	227.6	261.9	318.1	227	229.2	227
bbc	www.bbc.co.uk	225.3	227.8	225.4	262.1	270.8	228.9	228.2	227.6
Internet(Finland)									
nokia	www.nokia.com	273.7	272.2	272.3	273.9	319.1	273.4	272.4	271.8
Internet(USA)									

2.2 Application: Impact analysis of submarine cable faults

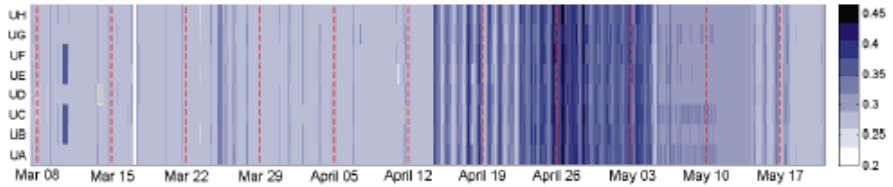
Eyjafjallajöekull volcano eruption



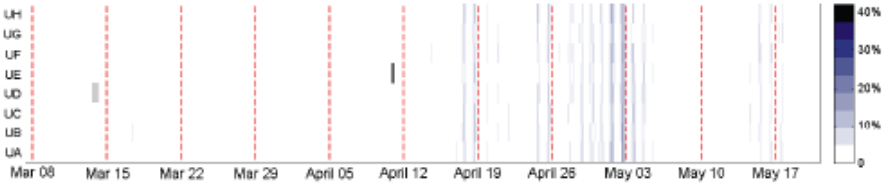
Path-quality degradation for NOK (Finland) and ENG (in UK)



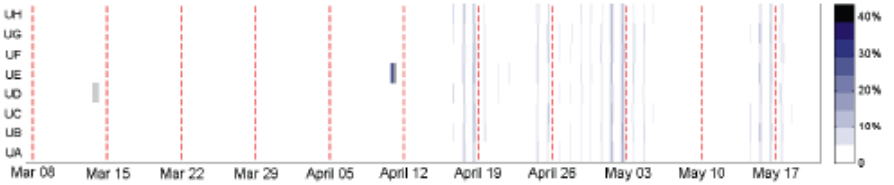
(a) The ENG path's RTT.



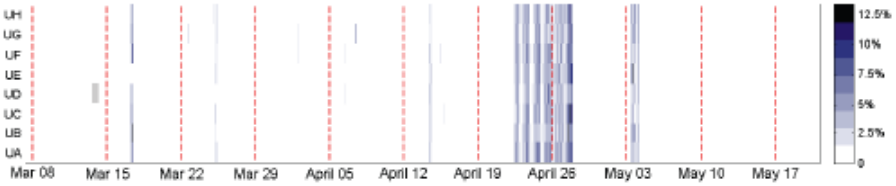
(b) The NOK path's RTT.



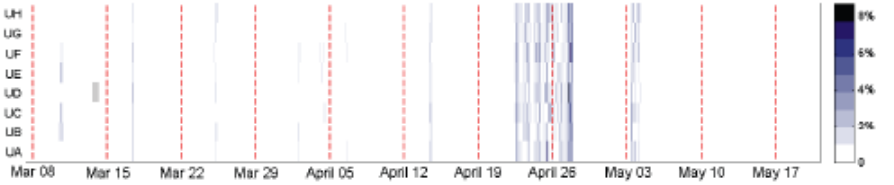
(c) The ENG path's forward-path loss.



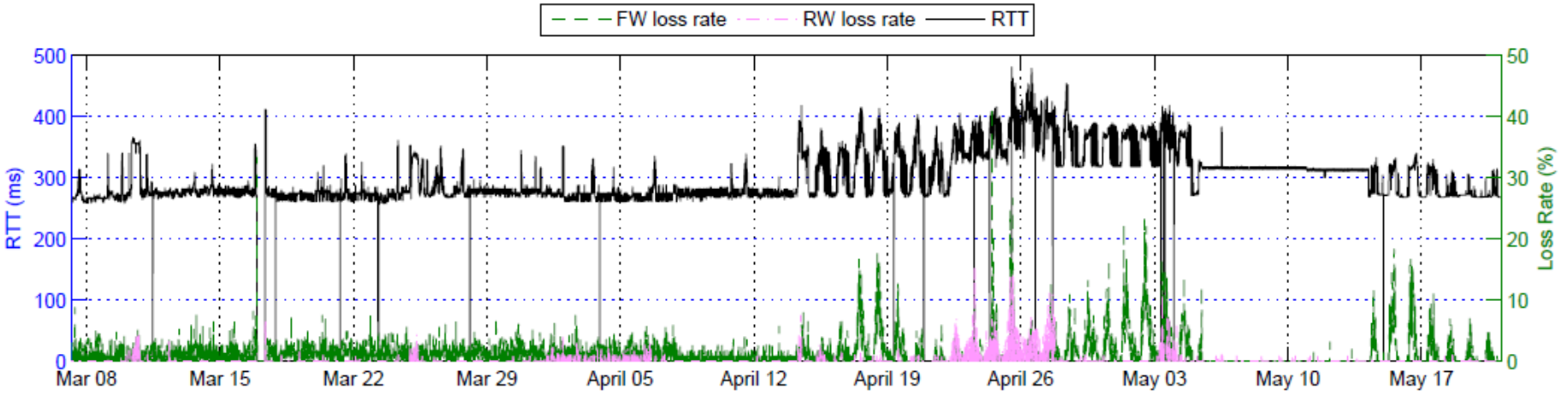
(d) The NOK path's forward-path loss.



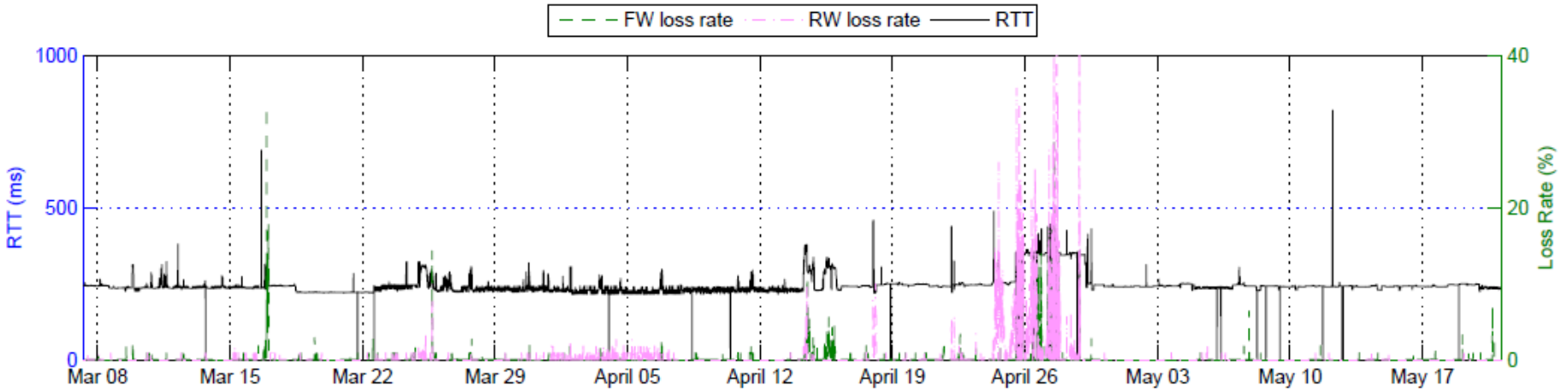
(e) The ENG path's reverse-path loss.



(f) The NOK path's reverse-path loss.



(a) UB→NOK's RTT and loss rates.



(b) UB→BBC's RTT and loss rates.

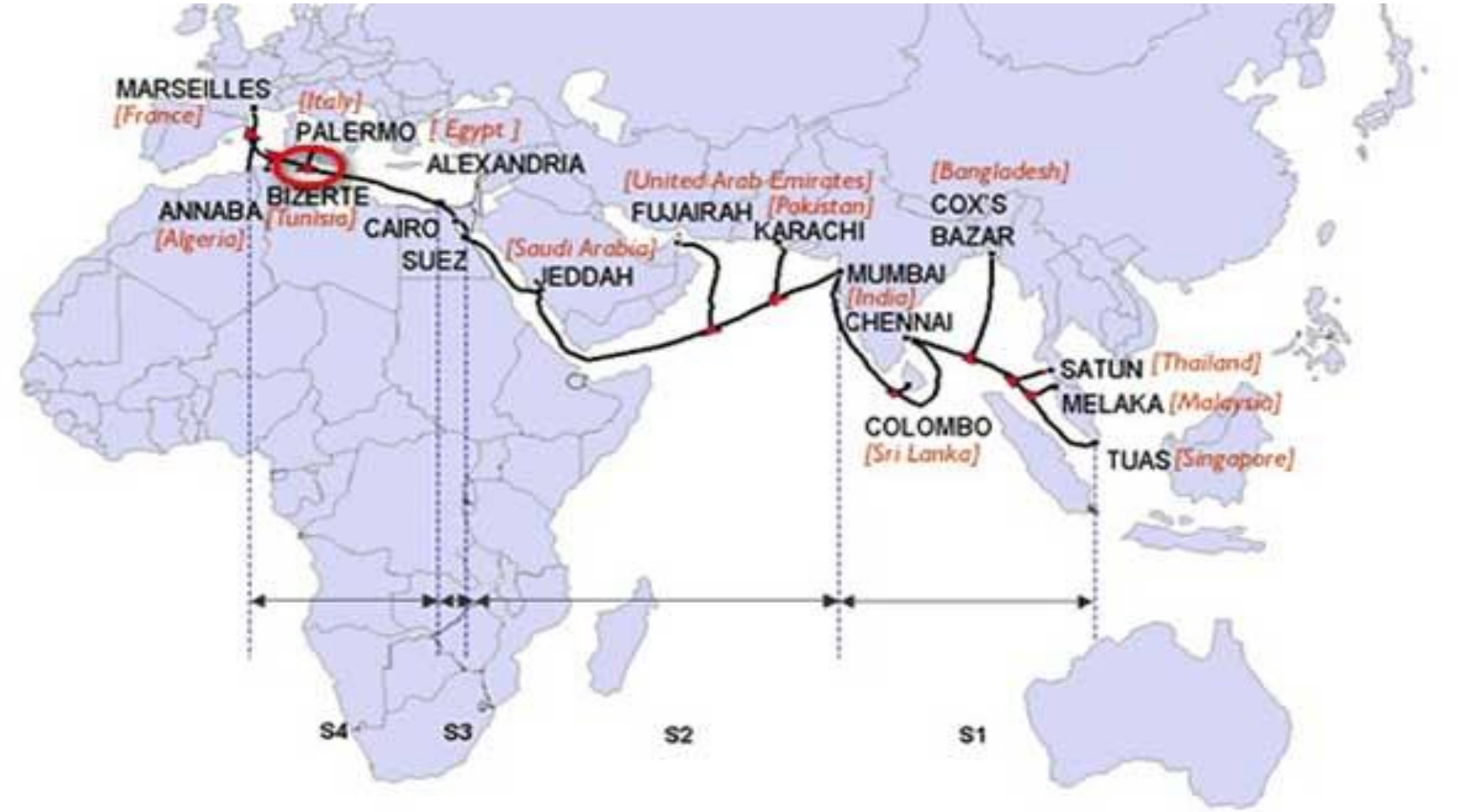
Network congestion caused by the volcano ashes?

- The surges on packet loss and RTT occurred on 14 April 2009.
- But
  - The onsets of the path congestion and air traffic disruption do not entirely match.
  - Some of the peak loss rate and RTT occurred on weekends.
  - Path congestion can still be observed at the end of the measurement period.

A SEA-ME-WE 4 cable fault

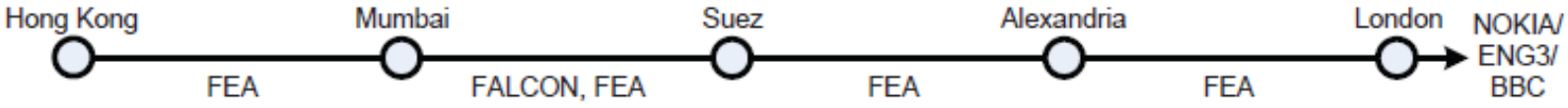
- The SEA-ME-WE 4 cable encountered a shunt fault on the segment between Alexandria and Marseille on 14 April 2010.
- The repair was started on 25 April 2010, and it took four days to complete.
- During the repair, the service for the westbound traffic to Europe was not available.
- "Non-cooperative Diagnosis of Submarine Cable Faults," *Proc. PAM 2011*, March 2011.

The SEA-ME-WE 4 cable

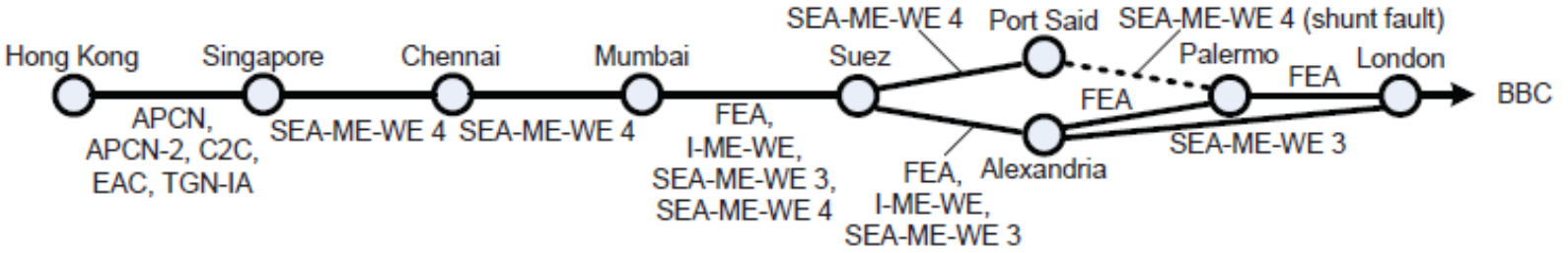


A plausible explanation for the network congestion

- The congestion in the FLAG network was caused by taking on rerouted traffic from the faulty SEA-ME-WE 4 cable.
  - FLAG does not use the SEA-ME-WE 4 cable for Hong Kong → NOKIA, ENG3, and BBC.
  - FLAG uses FEA for Hong Kong → NOKIA, ENG3, and BBC
  - TATA uses different cables between Mumbai and London.



(a) FLAG (in phases (a)–(c)).



(b) TATA (in phase (d)).

Conclusions and current works

- Turning a network protocol into a measurement protocol.
- Coming up a novel measurement method is just half a story.
- Making it work in the non-cooperative Internet is hard.
- Current works
  - Expanding OneProbe's capability (e.g., asymmetric available bandwidth)
  - Applications: fault localizations, SLA measurement, speed test, net measurement neutrality, correlating with QoE, …

oneprobe.org