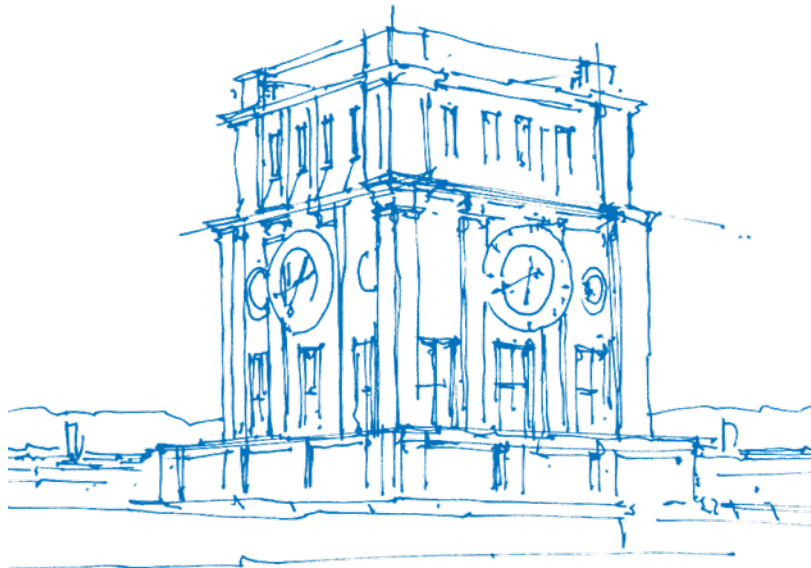


Analyzing Locality of Mobile Messaging Traffic Using the MATAdOR Framework

Quirin Scheitle, Matthias Wachs, Johannes Zirngibl, and Georg Carle
Heraklion, Greece, March 31, 2016



TUM Uhrenturm

Hypothesis: Mobile Messaging Services heavily distort traffic locality

Does the centrality of messaging services force traffic to leave region?

Mobile Messaging Services (MMS):

- Steady increase in Monthly Active Users (MAU)
- Taking market share from text messaging (SMS) and email

But:

- Proprietary services run by for-profit companies in few data centers
- Undisclosed protocols and applications

Violating traffic locality might impact user privacy:

- Content level access might be blocked by end-to-end encryption
- But meta data still accessible
- Centralized architectures attracts surveillance

Goal, Approach and Key Contributions

Our goals:

- Analyze centrality of mobile messaging services
- Analyze locality distortion caused by mobile messaging services
- Evaluate the impact of locality distortion on user privacy

Approach & outline:

- Automate largescale measurements between geographically distributed users
- Compare network path and application path between users
- Quantify amount of locality distortion

Key contributions:

- MATAdOR experimentation framework to conduct measurements with mobile applications
- Extensive dataset on mobile messaging service communication
- Quantification of locality distortion of mobile messaging service

Experimentation with Mobile Applications

Challenges with Apps:

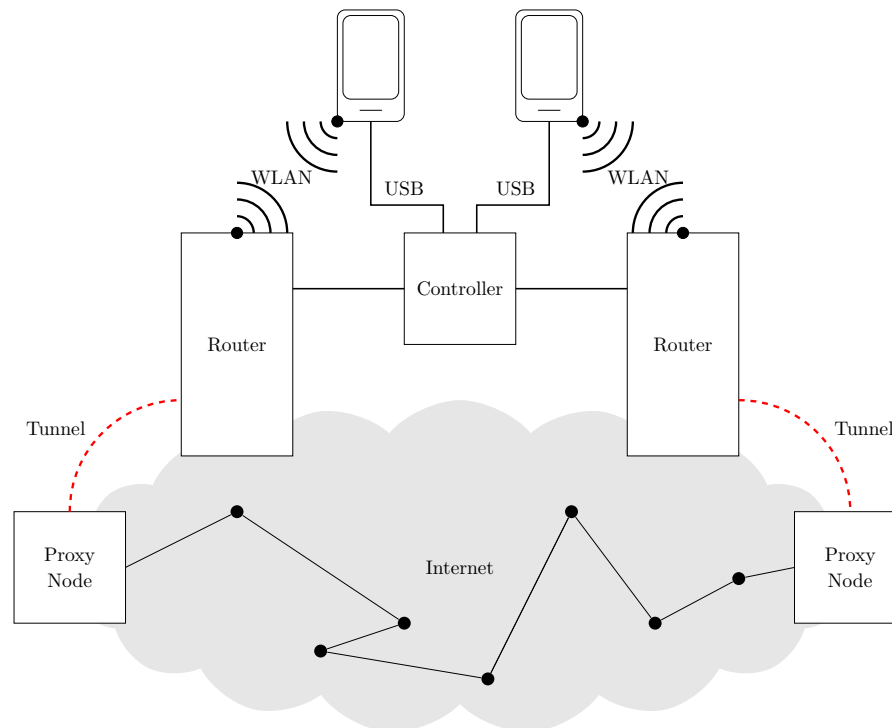
- Require cell phone number and SIM card
- Might behave differently if run in “strange” environments (emulation, proxy, VPN)
- Might behave differently based on their location
- Usually GUI-based without CLI, impeding anomaly handling (e.g., OS update dialog)

MATAdOR approach:

- Phone in flight mode, booked into private WLAN
- WLAN router transparently proxies traffic to arbitrary remote node
- XPrivacy reports remote destination to app under test
- Video screen record for validation of unusual captures

MATAdOR framework fully automates tests

- Phone and app automation: Android Debug Bridge
- Traffic routing automation: WLAN setup, remote tunnel setup, traffic capturing
- Path measurement automation: Target extraction from captures, in-protocol path measurements



Repository: <https://github.com/tumi8/matador>

Analyzing Traffic Locality by comparing Direct Path and Application Path

Redirect Mobile Messaging traffic through PlanetLab nodes:

- **Direct Path:** Forward path measurement between PlanetLab nodes
- **Application Path:** Forward path measurement to specific backend servers observed from that location

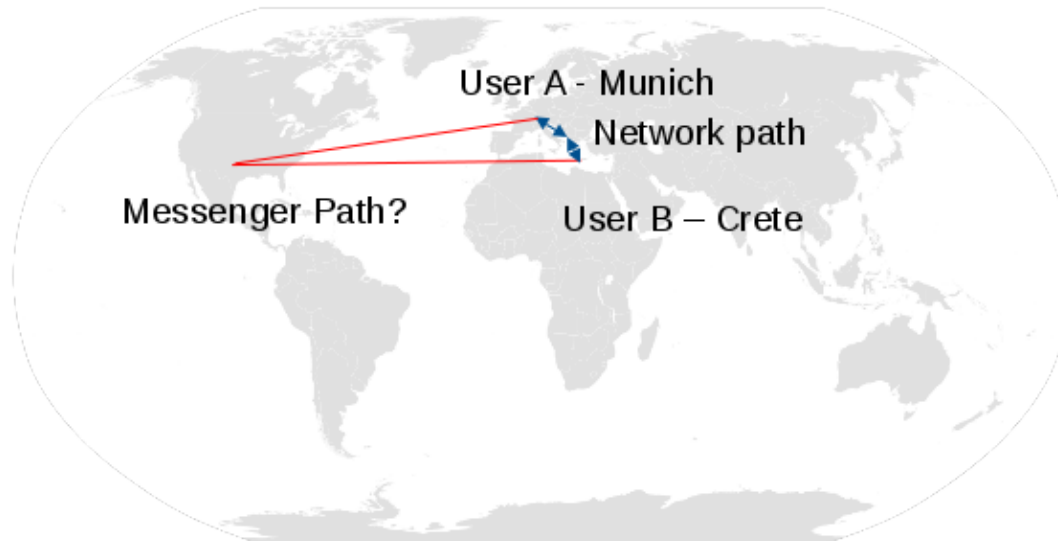


Image Source: <https://upload.wikimedia.org/wikipedia/commons/thumb/0/03/BlankMap-World6.svg>

Mobile Messaging Services: which to test?

Data-driven approach to select 4 of 19 services, full details in paper:

- Popularity: WhatsApp and WeChat due to large user base
- Security: Signal (then TextSecure) due to high rank in EFF scorecard
- Location: Threema due to Switzerland-based infrastructure
- Architecture: ~~Bleep due to P2P Architecture~~

Table 1. Properties of mobile messaging services and applications.

Application (Version)	Monthly active users ¹ [22]	EFF Scorecard ² Points [4]	Architecture	Server Distribution	Mobile First
WhatsApp (2.12.176)	800-900mn [12, 23] [27, p.23]	2	client-server	n/a	✓
WeChat (6.2.4)	400-600mn [27, p.22] [26, p.4]	n/a	client-server	n/a	✓
Facebook ³	350-600mn [5], [27, p.22]	2	client-server	n/a	✗
Skype	300mn [15]	1	client-server	n/a	✗
QQ International	843mn [26, p.4]	2	client-server	n/a	✗
Viber	249mn [21]	1	client-server	n/a	✓
LINE	211mn [13]	n/a	client-server	n/a	✓
Kik	300mn ⁴ [27]	1	client-server	n/a	✓
Tango	270mn [24]	n/a	client-server	n/a	✓
KakaoTalk	48mn [2]	n/a	client-server	n/a	✓
Yahoo Messenger	n/a	1	client-server	n/a	✗
TextSecure (2.24.1)	> 10mn ⁴ [17]	7	client-server	global	✓
Silent Text	n/a	7	client-server	n/a	✓
Telegram	30-50mn [27, p.22] [28]	4 ⁵	client-server	global	✓
Wickr	4mn ⁶ [20]	5	client-server	global	✓
Bleep (1.0.616)	n/a	n/a	P2P	n/a	✓
FireChat	n/a	n/a	mesh P2P	n/a	✓
Threema (2.41)	3mn ⁴ [29]	5	client-server	Switzerland	✓
SIMSme	1 mn ⁶	n/a	client-server	Germany	✓

1: Around July 30, 2015, for exact date see app-specific source 2: EFF Secure Messaging Scorecard [4]
 3: Stand-alone Facebook Messenger 4: Registered users 5: Score of 7 in secure chats 6: App Store Downloads

MATAdOR sent 6,496 automated messages between 28 countries and 4 applications

- One measurement [5-10 minutes]:
 - Tunnel setup, Location setting through XPrivacy
 - 1 Mobile messaging application
 - 2 phones, traffic proxied through 2 remote locations
 - 2 messages sent in each direction
 - Path measurements to IP addresses active in measurement
- Release of full, unaltered data set on our website, unique in its kind

Manual post-processing of traces required to identify messaging backend servers

Classification of IPs

- Lots of Android system background noise (identification aided by empty measurements)
- Correlation of send/receive timestamps to identify messaging servers
- IPs usually resolved through DNS, very few hardwired IPs

Insights:

- WeChat complements DNS by a custom DNS-over-HTTP protocols
- WeChat is the only app to use different servers based on location
- Other messengers do not use regional differentiation at all, all traffic directed to same IP (or subnet)

Mapping paths to countries and regions required some manual post-processing

- Path IP addresses mapped to countries using the ip2location.com database
- Countries mapped to regions according to United Nations GeoScheme
- Countries mapped to interest groups (5 Eyes, EU, ...)
- Manual validation of high-impact IP mappings (based on round-trip time, reverse DNS, neighbouring hops, active measurements)

Mobile Messaging Services frequently direct traffic out of region

Table 2. Mobile messaging services in almost all cases direct traffic out of region.

Region	# Measurements	Traffic leaving region					
		Network Path		Application Path			
		#	%	TextSecure	Threema	WeChat	WhatsApp
Europe	120	0	0%	100%	0%	100%	100%
Oceania	3	0	0%	100%	100%	100%	100%
Asia	28	6	21%	100%	100%	50%	100%
Americas	10	0	0%	0%	100%	100%	0%
South America	3	1	33%	100%	100%	100%	100%
Northern America	3	0	0%	0%	100%	100%	0%

Legend: > Network Path

Mobile Messaging Services generally make traffic more accessible to Interest Groups

Table 3. Mobile messaging services in most cases increase traffic accessibility for interest groups.

Region	Interest Group	#Total	Accessible for Interest Group									
			Network Path		TextSecure		Threema		WeChat		WhatsApp	
			#	%	#	%	#	%	#	%	#	%
Europe	5 Eyes	120	86	72%	120	100%	68	57%	119	99%	120	100%
Europe	EU	120	118	98%	119	99%	119	99%	120	100%	120	100%
Europe	China	120	0	0%	0	0%	0	0%	120	100%	0	0%
Oceania	5 Eyes	3	3	100%	3	100%	3	100%	3	100%	3	100%
Oceania	EU	3	0	0%	0	0%	3	100%	0	0%	0	0%
Oceania	China	3	0	0%	0	0%	0	0%	3	100%	0	0%
Asia	5 Eyes	28	6	21%	28	100%	21	75%	14	50%	28	100%
Asia	EU	28	6	21%	7	25%	18	64%	7	25%	7	25%
Asia	China	28	10	36%	7	25%	7	25%	28	100%	7	25%
South America	5 Eyes	3	1	33%	3	100%	3	100%	3	100%	3	100%
South America	EU	3	0	0%	0	0%	2	67%	0	0%	0	0%
South America	China	3	0	0%	0	0%	0	0%	3	100%	0	0%
North America	5 Eyes	3	3	100%	3	100%	3	100%	3	100%	3	100%
North America	EU	3	0	0%	0	0%	2	67%	0	0%	0	0%
North America	China	3	0	0%	0	0%	0	0%	3	100%	0	0%

Legend:

< Network Path

> Network Path

Key Findings

- Peer-to-peer services difficult to implement on mobile devices
 - Silent fall-back may contradict user expectations
- Mobile messaging services generally centralized and location-independent
- WeChat with interesting DNS-over-HTTP technique and regional optimization
- Mobile messaging services usually distort traffic locality
- Mobile messaging services can improve traffic locality in few cases

→ No single “best” messenger, choose based on your usage pattern.

Key Contributions

- Insights into mobile messaging traffic locality
- Ready-to-use, publicly available MATAdOR framework
- Full, unaltered dataset of bidirectional communication of 4 mobile messaging applications

Key Contributions

- Insights into mobile messaging traffic locality
- Ready-to-use, publicly available MATAdOR framework
- Full, unaltered dataset of bidirectional communication of 4 mobile messaging applications

Questions?

Quirin Scheitle <scheitle@net.in.tum.de>

Matthias Wachs <wachs@net.in.tum.de>

More information, framework and data set available under:

<https://net.in.tum.de/pub/mobmes/>



Further Material

[Table 1: Mobile Messaging Services selection](#)

[Definition of Interest Groups](#)

[Table 2: Traffic Leaving Region](#)

[Table 3: Accessibility for Interest Groups](#)

[Anomaly: US to Switzerland without EU](#)

[Changes to ip2location database](#)

[Hidden parts of service infrastructure](#)

[WeChat DNS-over-HTTP](#)

[PlanetLab nodes used](#)

[Full DNS table](#)

Table 1: Mobile Messaging Services selection

Table 1. Properties of mobile messaging services and applications.

Application (Version)	Monthly active users ¹ [22]	EFF Scorecard ² Points [4]	Architecture	Server Distribution	Mobile First
WhatsApp (2.12.176)	800-900mn [12, 23] [27, p.23]	2	client-server	n/a	✓
WeChat (6.2.4)	400-600mn [27, p.22] [26, p.4]	n/a	client-server	n/a	✓
Facebook ³	350-600mn [5], [27, p.22]	2	client-server	n/a	✗
Skype	300mn [15]	1	client-server	n/a	✗
QQ International	843mn [26, p.4]	2	client-server	n/a	✗
Viber	249mn [21]	1	client-server	n/a	✓
LINE	211mn [13]	n/a	client-server	n/a	✓
Kik	200mn ⁴ [25]	1	client-server	n/a	✓
Tango	70mn [24]	n/a	client-server	n/a	✓
KakaoTalk	48mn [2]	n/a	client-server	n/a	✓
Yahoo Messenger	n/a	1	client-server	n/a	✗
TextSecure (2.24.1)	>10mn ⁴ [17]	7	client-server	global	✓
Silent Text	n/a	7	client-server	n/a	✓
Telegram	30-50mn [27, p.22] [28]	4 ⁵	client-server	global	✓
Wickr	4mn ⁶ [20]	5	client-server	global	✓
Bleep (1.0.616)	n/a	n/a	P2P	n/a	✓
FireChat	n/a	n/a	mesh P2P	n/a	✓
Threema (2.41)	3mn ⁴ [29]	5	client-server	Switzerland	✓
SIMSme	1 mn ⁶	n/a	client-server	Germany	✓

1: Around July 30, 2015, for exact date see app-specific source 2: EFF Secure Messaging Scorecard [4]
3: Stand-alone Facebook Messenger 4: Registered users 5: Score of 7 in secure chats 6: App Store Downloads

Definition of Interest Groups

- *5 Eyes* consisting of: Great Britain, United States, New Zealand, Canada
- *European Union* consisting of: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, United Kingdom
- *Arab League* consisting of: Algeria, Bahrain, Comoros, Djibouti, Egypt, Iraq, Jordan, Kuwait, Lebanon, Libya, Mauritania, Morocco, Oman, Palestine, Qatar, Saudi Arabia, Somalia, Sudan
- *Russia* with the only member Russia
- *China* with the only member China.

Table 2: Traffic Leaving Region

Table 2. Mobile messaging services in almost all cases direct traffic out of region.

Region	# Measurements	Traffic leaving region						
		Network Path		Application Path				
		#	%	TextSecure	Threema	WeChat	WhatsApp	
Europe	120	0	0%	100%	0%	100%	100%	
Oceania	3	0	0%	100%	100%	100%	100%	
Asia	28	6	21%	100%	100%	50%	100%	
Americas	10	0	0%	0%	100%	100%	0%	
South America	3	1	33%	100%	100%	100%	100%	
Northern America	3	0	0%	0%	100%	100%	0%	

Legend:

> Network Path

Table 3: Accessibility for Interest Groups

Table 3. Mobile messaging services in most cases increase traffic accessibility for interest groups.

Region	Interest Group	#Total	Accessible for Interest Group									
			Network Path		TextSecure		Threema		WeChat		WhatsApp	
			#	%	#	%	#	%	#	%	#	%
Europe	5 Eyes	120	86	72%	120	100%	68	57%	119	99%	120	100%
Europe	EU	120	118	98%	119	99%	119	99%	120	100%	120	100%
Europe	China	120	0	0%	0	0%	0	0%	120	100%	0	0%
Oceania	5 Eyes	3	3	100%	3	100%	3	100%	3	100%	3	100%
Oceania	EU	3	0	0%	0	0%	3	100%	0	0%	0	0%
Oceania	China	3	0	0%	0	0%	0	0%	3	100%	0	0%
Asia	5 Eyes	28	6	21%	28	100%	21	75%	14	50%	28	100%
Asia	EU	28	6	21%	7	25%	18	64%	7	25%	7	25%
Asia	China	28	10	36%	7	25%	7	25%	28	100%	7	25%
South America	5 Eyes	3	1	33%	3	100%	3	100%	3	100%	3	100%
South America	EU	3	0	0%	0	0%	2	67%	0	0%	0	0%
South America	China	3	0	0%	0	0%	0	0%	3	100%	0	0%
North America	5 Eyes	3	3	100%	3	100%	3	100%	3	100%	3	100%
North America	EU	3	0	0%	0	0%	2	67%	0	0%	0	0%
North America	China	3	0	0%	0	0%	0	0%	3	100%	0	0%

Legend:

< Network Path

> Network Path

Anomaly: US to Switzerland without EU

```

traceroute to 5.148.175.201 (5.148.175.201) , 30 hops max, 60 byte packets
 1 host129.190-227-163.telecom.net.ar (190.227.163.129)  1.395 ms 1.378 ms 1.454 ms
  :
 8 xe-11-0-0.edge2.miami1.Level3.net (63.209.150.165)  130.908 ms 137.789 ms 138.165
   ms
 9 ae-2-26.bar1.Zurich1.Level3.net (4.69.142.133)  278.270 ms 278.336 ms 278.079 ms
10 ae-2-26.bar1.Zurich1.Level3.net (4.69.142.133)  248.524 ms 272.676 ms 272.715 ms
11 NINE-INTERN.bar1.Zurich1.Level3.net (213.242.67.94)  263.285 ms 266.851 ms *
12 * * *
  :
16 * ps1.threema.ch (5.148.175.201)  259.566 ms 255.001 ms

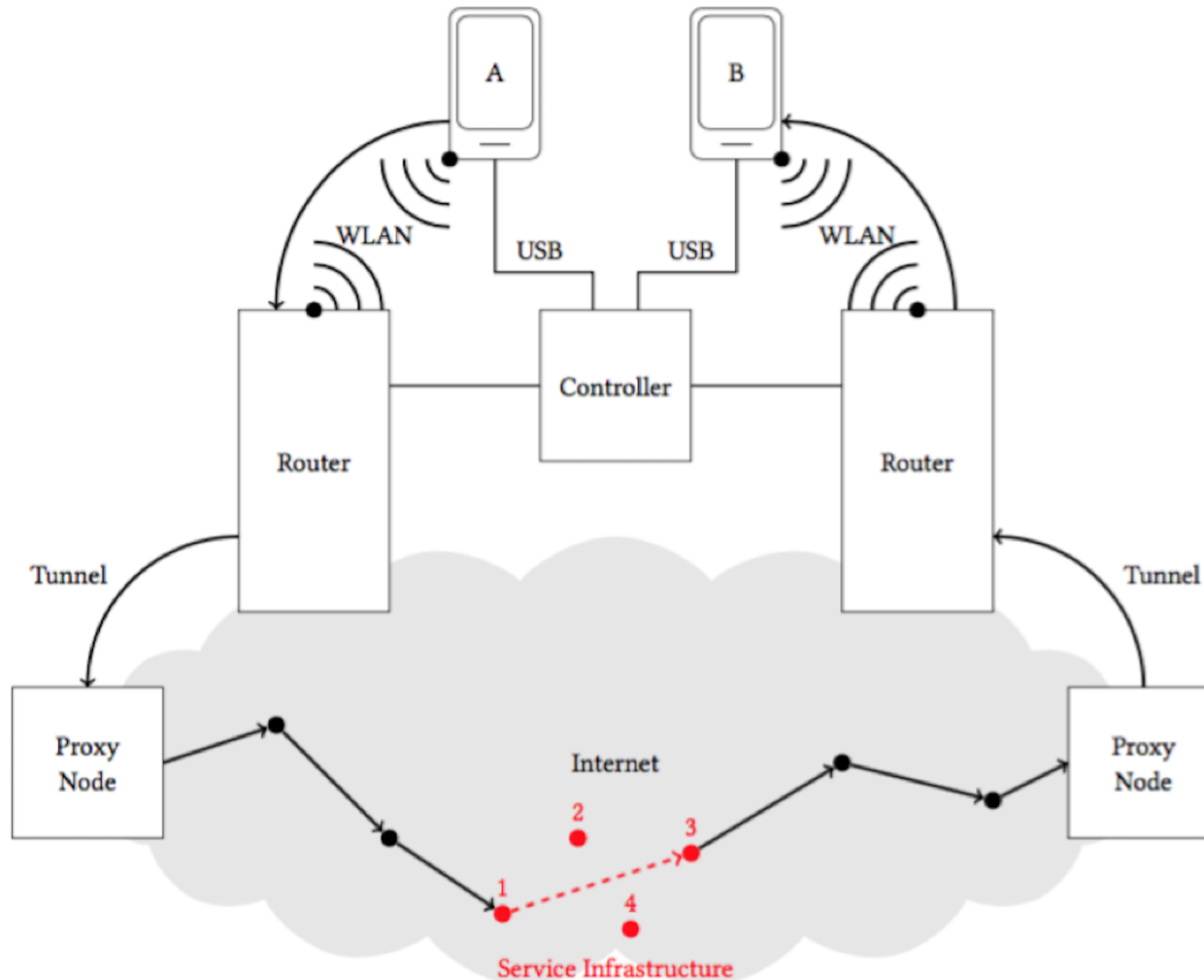
```

Listing 7.5: Traceroute from planet-lab2.itba.edu.ar to a Threema Server

Changes to ip2location database

Geolocation			Geolocation			Geolocation		
IP Address	Old	New	IP Address	Old	New	IP Address	Old	New
4.68.111.178	US	NL	176.32.125.145	IE	US	176.32.125.229	IE	US
4.69.137.81	US	CH	176.32.125.147	IE	US	176.32.125.231	IE	US
61.8.59.37	AU	JP	176.32.125.148	IE	US	178.236.3.24	IE	US
61.8.59.38	AU	JP	176.32.125.152	IE	US	178.236.3.96	IE	US
62.115.140.203	US	FR	176.32.125.155	IE	US	178.236.3.98	IE	US
62.216.145.74	GB	HK	176.32.125.159	IE	US	185.70.203.37	IT	AR
63.216.156.106	US	HK	176.32.125.160	IE	US	185.70.203.61	IT	AR
63.216.156.86	US	HK	176.32.125.162	IE	US	185.70.203.63	IT	AR
63.218.230.73	US	DE	176.32.125.164	IE	US	195.22.199.177	IT	US
63.223.29.10	US	HK	176.32.125.167	IE	US	195.22.219.145	IT	BR
63.223.29.14	US	HK	176.32.125.169	IE	US	195.22.219.147	IT	BR
63.223.29.18	US	HK	176.32.125.171	IE	US	195.22.219.171	IT	BR
77.19.128.210	NO	CH	176.32.125.172	IE	US	195.22.219.175	IT	BR
80.77.0.181	EG	HK	176.32.125.174	IE	US	195.22.219.177	IT	BR
80.77.0.182	EG	HK	176.32.125.176	IE	US	195.22.219.179	IT	BR
82.197.168.121	HR	CH	176.32.125.179	IE	US	198.32.141.11	US	SG
85.95.25.105	US	HK	176.32.125.181	IE	US	198.32.141.146	US	SG
85.95.25.90	US	HK	176.32.125.183	IE	US	202.147.40.114	AU	US
85.95.26.102	AE	HK	176.32.125.186	IE	US	202.147.58.146	AU	US
89.221.41.175	IT	US	176.32.125.191	IE	US	202.147.58.147	AU	GB
89.221.41.177	IT	US	176.32.125.193	IE	US	202.147.58.150	AU	US
109.105.97.73	US	DE	176.32.125.196	IE	US	202.147.58.151	AU	GB
129.250.6.162	JP	US	176.32.125.198	IE	US	202.84.140.234	AU	SG
149.3.181.101	IT	US	176.32.125.200	IE	US	202.84.143.253	AU	HK
149.3.181.65	IT	US	176.32.125.203	IE	US	202.84.221.25	AU	SG
149.3.181.87	IT	US	176.32.125.207	IE	US	202.84.221.26	AU	KR
149.3.181.91	IT	US	176.32.125.208	IE	US	202.84.249.161	AU	HK
149.3.181.95	IT	US	176.32.125.210	IE	US	203.233.2.234	KR	US
149.3.181.97	IT	US	176.32.125.212	IE	US	212.162.10.81	US	DE
149.3.181.99	IT	US	176.32.125.215	IE	US	212.162.10.82	US	DE
150.99.188.201	JP	US	176.32.125.217	IE	US	213.242.73.74	US	CH
176.32.125.136	IE	US	176.32.125.219	IE	US	217.239.45.6	DE	US
176.32.125.138	IE	US	176.32.125.220	IE	US	217.6.51.246	DE	US
176.32.125.140	IE	US	176.32.125.224	IE	US	-	-	-
176.32.125.143	IE	US	176.32.125.227	IE	US	-	-	-

Hidden parts of service infrastructure



WeChat DNS-over-HTTP

```

1  <?xml version="1.0" encoding="utf-8"?>
2  <dns>
3    <retcode>0</retcode>
4    <domainlist>
5      <domain name="extshort.weixin.qq.com" timeout="1800">
6        <ip>103.7.31.152</ip>
7      </domain>
8      <domain name="localhost" timeout="1800">
9        <ip>127.0.0.1</ip>
10     </domain>
11     <domain name="long.weixin.qq.com" timeout="1800">
12       <ip>103.7.31.151</ip>
13     </domain>
14     <domain name="minorshort.weixin.qq.com" timeout="1800">
15       <ip>103.7.31.152</ip>
16     </domain>
17     :
101    <domain name="hkshort.weixin.qq.com" timeout="1800">
102      <ip>203.205.151.160</ip>
103      <ip>203.205.147.168</ip>
104      <ip>203.205.129.101</ip>
105    </domain>
106  </domainlist>
107  <builtiniplist>
108    <ip>203.205.151.164</ip>
109    <ip>203.205.143.141</ip>
110    <ip>203.205.129.102</ip>
111  </builtiniplist>
112  <clientip>193.190.168.49</clientip>
113  <clientspid>0</clientspid>
114  <timestamp>1443628800</timestamp>
115  <signature>MDwCHFqMSx/
      PdY6OtMi59uAjQ0JAqe3ZsJ8IU7Fii0CHDuAtxbwX/XV094cGvj00r83+
      iHY2fFwYKZqmqqs=</signature>
116 </dns>

```

PlanetLab nodes used

Region	Country	# ¹	Node Used
(15)	Europe		
		Belgium	1 planetlab1.extern.kuleuven.be
		Czech Republic	4 planetlab1.cesnet.cz ²
		Denmark	2 planetlab1.diku.dk
		Germany	14 iraplab1.iralab.uni-karlsruhe.de
		Finland	1 planetlab4.hiit.fi
		France	13 planetlab1.jcp-consult.net
		Greece	2 planetlab3.cslab.ece.ntua.gr
		Ireland	2 planetlab-coffee.ait.ie
		Italy	3 planet-lab-node1.netgroup.uniroma2.it
		Norway	2 planetlab1.ifi.uio.no
		Poland	6 planetlab1.mini.pw.edu.pl
		Portugal	3 planet2.servers.ua.pt
		Spain	5 planetlab1.um.es
		Sweden	5 planetlab-1.ida.liu.se
	Switzerland	4 planetlab2.inf.ethz.ch	
(7)	Asia		
		China	9 pl1.6test.edu.cn
		Hong Kong	3 planetlab1.ie.cuhk.edu.hk
		Israel	3 planetlab1.mta.ac.il
		Japan	7 pl1.sos.info.hiroshima-cu.ac.jp
		Korea, Republic of	1 netapp6.cs.kookmin.ac.kr
		Singapore	3 planetlab1.comp.nus.edu.sg
	Thailand	2 ple1.ait.ac.th	
(2)	Oceania		
		Australia	2 pl1.eng.monash.edu.au
	New Zealand	4 planetlab1.cs.otago.ac.nz	
(2)	North America		
		Canada	9 cs-planetlab3.cs.surrey.sfu.ca
	U.S.	71 planetlab01.cs.washington.edu	
(2)	South America		
		Argentina	2 planet-lab2.itba.edu.ar
	Brazil	3 planetlab1.pop-mg.rnp.br	

1: Number of accessible nodes during the measurement

2: The Czech PlanetLab node was not accessible for a period of time. Some measurements use planetlab1.fit.vutbr.cz

