Where is the Internet?

Answers from Yahoo! Answers:

(G. Whilikers) Out there.

(Mike) the way I see it, the "internet" has to be somewhere. a router collects the internet to my house sure. but somewhere on earth there HAS to be some where that the internet originates by sending out the first signal so others can collect it. where is that place? also is it like... a box or something? I'm so confused. one more for those who want to answer, if some maintenance guy were to spill coffee on it would the worlds internet crash? like... everywhere?

(tp5com) Everywhere. It's not really a physical being.

(evaohell) The internet is stored in a little black box with a blinking light:
And what does it look like?
The Network Geography of the Internet*

Péter Hága
Eötvös Loránd University
Budapest, Hungary

* Spotter: A Model Based Active Geolocation Service, INFOCOM 2011
* On the Network Geography of the Internet, INFOCOM 2011
* On the Spatial Properties of Internet Routes, submitted to Journal on Selected Areas of Communications
Geolocation in general
Geolocation in general

• passive geolocation
  o extracting location information from domain names (DNS)
  o registries, Whois databases, commercial databases
  o large and geographically dispersed IP blocks are allocated to a single entity prohibiting the study of several geographic aspects

• active geolocation
  o active probing (delay, topology)
  o measurement nodes with known location
  o constraint based techniques
Whois based location estimation
example for passive geoloc

<table>
<thead>
<tr>
<th>Registered location</th>
<th>number of distinct IPv4 address</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mountain View, CA, US</td>
<td>222340</td>
</tr>
<tr>
<td>Plano, TX, US</td>
<td>331</td>
</tr>
<tr>
<td>Chicago, IL, US</td>
<td>46</td>
</tr>
<tr>
<td>Irvine, CA, US</td>
<td>30</td>
</tr>
<tr>
<td>Waterloo, ON, CA</td>
<td>22</td>
</tr>
<tr>
<td>San Francisco, CA, US</td>
<td>21</td>
</tr>
<tr>
<td>Atlanta, GA, US</td>
<td>15</td>
</tr>
<tr>
<td>Phoenix, AZ, US</td>
<td>15</td>
</tr>
<tr>
<td>Southfield, MI, US</td>
<td>15</td>
</tr>
<tr>
<td>Mequon, WI, US</td>
<td>14</td>
</tr>
<tr>
<td>Pittsburgh, PA, US</td>
<td>14</td>
</tr>
<tr>
<td>Gladwyne, PA, US</td>
<td>7</td>
</tr>
<tr>
<td>Richardson, TX, US</td>
<td>7</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>222877</strong></td>
</tr>
</tbody>
</table>

Cumulative distribution of the maximal distances from Pamplona, Spain to 4000 Google IPs. The maximal distances are calculated from the network delays assuming 200000 km/sec signal propagation speed. The vertical line represents the real geographical distance between Pamplona and Mountain View, CA, showing that 47% of the nodes must be closer to Pamplona than Mountain View.
Spotter – theoretical background
Spotter – theoretical background

- Roundtrip travel time measured between the landmark (L) and the target (T) node by ping
- RTT is transformed to geographic distance
- "Triangulation" from multiple landmarks
Travel time – distance relation

- reference dataset (nodes with known location)
- distance between the source and destination
- measured RTT
Standardized travel time – distance distribution

- Standardized values, fitted normal distribution
The distances are normally distributed for a given RTT. Each RTT point out a radial profile with $\mu(d)$ and $\sigma(d) \Rightarrow$ which defines the spatial probability distribution of the target.
"Triangulation"

The target’s location must satisfy all the individual spatial constraints.
Spotter - online geolocation service

- online: free, easy to use
- offline: batch mode
  ~15k addresses/day
- to fix: DNS resolve problem

visit and use!

http://spotter.etomic.org
Where is the WikiLeaks site (46.59.1.2)?

hostip.info
Where is the WikiLeaks site (46.59.1.2)?
Where is the WikiLeaks site (46.59.1.2)?

MaxMind

MaxMind GeoIP Demo

<table>
<thead>
<tr>
<th>Hostname</th>
<th>Country Code</th>
<th>Country Name</th>
<th>Region Name</th>
<th>City</th>
<th>Postal Code</th>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>wikileaks.ch</td>
<td>AU</td>
<td>Australia</td>
<td></td>
<td></td>
<td>-27.0000</td>
<td>133.0000</td>
<td></td>
</tr>
</tbody>
</table>

These results were generated with the Perl API and the commercial GeoIP City, GeoIP ISP, and GeoIP Organization databases.

To find countries and cities, enter IP addresses/hostnames, separated by spaces: (To get a demo for your IP address, click here.)

Lookup IP Addresses

There is a limit of 25 demo lookups per day. Request a test account for more lookups.

MaxMind, GeoIP and related marks are registered trademarks of MaxMind, Inc.

Copyright © 2010 MaxMind, Inc. All Rights Reserved. Terms of use.
Where is the WikiLeaks site (46.59.1.2)?

Results:

wikileaks.ch
IP Address: 46.59.1.2
Status: Host is available.
Expected Coordinates: 59.35 17.94
City-level Location: Stockholm, Sweden

Map:
Where is the WikiLeaks site (46.59.1.2)?

http://www.youtube.com/watch?v=qwlATf9xse4
Spotter’s accuracy

- Estimation accuracy for reference node set
  (thanks to Bradley@CAIDA)
- Spotter – “pure probabilistic method”
- Spotter city – population density as extra constraint
Some basic geographic properties of the Internet

- Are there typical link lengths?
- How circuitous are the network paths?
- What is the extent of route asymmetry?
Router density visualization

- routers of the full-mesh PlanetLab topology (700x700)
- discovered by traceroute
- both academic and commercial ISPs
- IP level network
- approx. 16,000 addresses
Link length frequency & distribution
Each unique link is represented once.

Links are weighted up with their prevalence in the traceroute data set collected at one time.

# of identified IP level links: approx. 44,000
Link length distribution

- power law, where \( P(d) \propto d^{-\delta} \)?
- logarithmic relation, where \( P(d) \propto 1/d \)?
- model behind the observation?
Circuitousness of network paths
Circuitousness of network paths

relative circuitousness = \frac{\text{aggregated link length}}{\text{spherical distance of the endpoints}}
Circuitousness of network paths

Intra-continental routes: significant circuitousness

Inter-continental routes: determined by the intercontinental gateways
Symmetry of network routes
Symmetry of network routes

Ratio of the route lengths of the forward and the backward directions.

- A: United Kingdom – Hong Kong
- B: California, USA – Hong Kong
- C: California, USA – Singapore
Around the world in 300 ms

from the United Kingdom to Hong Kong:
• forward direction: eastward through Europe and Asia
• backward direction: eastward through the USA
Where is the Internet?*
And what does it look like?*

*as our dataset shows
Thank you for your attention!

for more info &
for the online service please visit:
http://spotter.etomic.org