Fathom: A Browser-based Network Measurement Platform

Mohan Dhawan, Justin Samuel, Renata Teixeira, Christian Kreibich, Mark Allman, Nicholas Weaver & Vern Paxson

Rutgers University, UC Berkeley, CNRS & UPMC and ICSI
User/edge-centric measurement
User-centric measurement

- Dasu [W-MUST’11]
- HomeNet Profiler [HomeNets’10, PAM’12]
- Hostview [HotMetrics’10]
- Netalyzr [IMC’10]
- Shaperprobe [IMC’11]
- Speedtests
User-centric measurement

- Dasu [W-MUST’11]
- HomeNet Profiler [HomeNets’10, PAM’12]
- Hostview [HotMetrics’10]
- Netalyzr [IMC’10]
- Shaperprobe [IMC’11]
- Speedtests

Vuze plugin, Java

Java

Native App

Native App

Flash
User-centric measurement
User-centric measurement

No general purpose measurement platform
Is there a better platform?
Is there a better platform?

+ Ubiquitous
+ All the time

− No proper API
− Security model

November 14, 2012
Contributions

Fathom is a practical browser-based network measurement platform

– Provides a programmable interface for writing and launching measurements from web pages
– Enables unparalleled access to the edge via the ubiquity of browsers
– Supports novel analyses via passive and active measurements
– Combines existing security primitives to safely expand capabilities of in-page JavaScript
Outline

• Concrete Example
• Design
• Implementation
• Evaluation
• Conclusion
Google Maps: End-user problems

What is up with Google Maps? It is either incredibly slow or pages just don't appear unless I click on the 'Still loading ...Slow? Try Basic HTML' button that has started to appear at the of of the page. This has happening for at least a week, maybe much longer. My connection runs at about 6Mb/sec and is fine for other sites so this is a google map issue. I notice many other posts over the last year where people have had problems at times when I haven't so is just my turn to be frustrated?
Hiya,

To make progress with this, we'll need some more information from you. Ideally, answers to these questions from everyone:

1) When did you first notice this problem?

2) Can you quantify "incredibly slow", eg, "It used to take around 4 seconds to load Maps and now it takes around 20" or "Maps loads fast but search results take at least 15 seconds to come back" etc.

3) Which browsers, operating systems, and (if possible) security patchlevels do you have?

4) What other software is installed on your computer, eg, do you have any toolbars, chat clients, etc installed?

5) If you know how to use the CPU monitoring facility of your operating system, is the CPU maxed out during page load, or does it seem to be waiting on the network?

6) Which area of the world are you connecting from? Metro-area granularity is ideal but I'll take country-level too :-)

Any information you can provide will be very helpful.
Google Maps: Inefficient solution

Hiya,

To make sure I understand your concerns, let's go through each of your questions:

1) When did you start noticing the issues?
2) Can you describe how slow it has become? Examples?
3) What happens to other Google services when Maps is not working?
4) What are your system specifications (CPU, memory, etc.)?
5) If you have a location history, could you please share it?
6) What is your country/region of residence?

Since when?
How slow?
CPU & N/W load?
Location?

Any information you can provide will be very helpful.

Hiya,

To make sure I understand your concerns, let's go through each of your questions:

1) When did you start noticing the issues?
2) Can you describe how slow it has become? Examples?
3) What happens to other Google services when Maps is not working?
4) What are your system specifications (CPU, memory, etc.)?
5) If you have a location history, could you please share it?
6) What is your country/region of residence?

Since when?
How slow?
CPU & N/W load?
Location?

Any information you can provide will be very helpful.
Outline

• Concrete Example
• Design
• Implementation
• Evaluation
• Conclusion
Fathom APIs

• Four main `fathom` namespaces
  – `fathom.socket.*`
    • Low-level socket management
Fathom APIs

• Four main `fathom` namespaces
  – `fathom.socket.*`
    • Low-level socket management
  – `fathom.proto.*`
    • Implement protocols like DNS, mDNS, HTTP and UPnP
Fathom APIs

- Four main `fathom` namespaces
  - `fathom.socket.*`
    - Low-level socket management
  - `fathom.proto.*`
    - Implement protocols like DNS, mDNS, HTTP and UPnP
  - `fathom.system.*`
    - Controlled access to system utilities
Fathom APIs

• Four main **fathom** namespaces
  – **fathom.socket.***
    • Low-level socket management
  – **fathom.proto.***
    • Implement protocols like DNS, mDNS, HTTP and UPnP
  – **fathom.system.***
    • Controlled access to system utilities
  – **fathom.utils.***
    • Miscellaneous browser APIs, timer, passive metrics, etc.
Fathom API use

```html
<html>
<body>
<script>
  function tr_callback(results) {...}
  fathom.system.traceroute(
    "maps.google.com", tr_callback);
</script>
<body>
<html>
```
Security

• Fathom APIs are powerful
  – Malicious scripts can misuse client resources
Security

• Fathom APIs are powerful
  – Malicious scripts can misuse client resources

• Five-fold defense mechanism
  – Client policy
  – User confirmation
  – Script manifest
  – Code signing
  – Server manifest
Security

• Fathom APIs are powerful
  – Malicious scripts can misuse client resources

• Five-fold defense mechanism
  – Client policy
  – User confirmation
  – Script manifest
  – Code signing
  – Server manifest
Security: User confirmation

HTTP Request

http://www.foo.com

Alice
Security: User confirmation

HTTP Response

http://www.foo.com

Measurement

JS

Alice
Security: User confirmation

http://www.foo.com

Security Dialog

Alice
Security: User confirmation

This security dialog is not informative!

Security Dialog

http://www.foo.com

Alice
Security: Script manifest

HTTP Request

http://www.foo.com

Alice
Security: Script manifest

HTTP Response

http://www.foo.com

Alice
Security: Script manifest

```javascript
var manifest = {
  'api': [
    'proto.*',
    'socket.*',
    'system.*',
    'util.*'
  ],
  'destinations': [
    'http://*.bar.com:*'
  ]
};
```
Security: Script manifest

A web page is requesting Fathom privileges.

http://www.foo.com

The page wants to use these APIs:
- proto:* APIs allow DNS, mDNS, HTTP and UPnP protocol access.
- socket:* APIs allow low-level socket communication.
- system:* APIs allow access to system utilities.
- util:* APIs allow access to timer, browser APIs and metrics.

The page wants to communicate with:
- http://*.bar.com:*

Security Dialog

Alice
Security: Script manifest

But still difficult for average users
Security: Server manifest

Untrusted script wants to open TCP connection to www.bar.com
Fathom requests the server’s manifest file

http://www.bar.com
Security: Server manifest

Fathom parses the server manifest

```xml
<allow-access-from domain="*.google.com"/>
<allow-access-from ip="198.168.150.25"/>
<deny-access-from domain="*.untrusted.com"/>
<deny-access-from ip="221.23.42.30"/>
```

http://www.bar.com
Security: Server manifest

```xml
<allow-access-from domain="*.google.com"/>
<allow-access-from ip="198.168.150.25"/>
<deny-access-from domain="*.untrusted.com"/>
<deny-access-from ip="221.23.42.30"/>
```

http://www.bar.com

Fathom prevents untrusted script from opening the TCP connection

Alice
Outline

• Concrete Example
• Design
  • Implementation
• Evaluation
• Conclusion
Fathom: A measurement platform for Firefox

- Firefox is a popular browser
- JavaScript-only extension (open source)
  - Rich extension API
  - Portable (across Firefox’s platforms)
Fathom: A measurement platform for Firefox

- Firefox is a popular browser
- JavaScript-only extension (open source)
  - Rich extension API
  - Portable (across Firefox’s platforms)

Technical details about the implementation are mentioned in the paper.
Outline

• Concrete Example
• Design
• Implementation
• Evaluation
• Conclusion
Overhead of page load time

- Fathom continuously monitors system & browser

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>Overhead (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Craigslist</td>
<td>1.1</td>
</tr>
<tr>
<td>CNN</td>
<td>3.1</td>
</tr>
<tr>
<td>ESPN</td>
<td>2.5</td>
</tr>
<tr>
<td>Google Maps</td>
<td>2.9</td>
</tr>
<tr>
<td>NY Times</td>
<td>1.3</td>
</tr>
<tr>
<td>Slashdot</td>
<td>1.4</td>
</tr>
<tr>
<td>Yahoo</td>
<td>1.0</td>
</tr>
<tr>
<td>YouTube</td>
<td>3.1</td>
</tr>
</tbody>
</table>
Overhead of page load time

- Fathom continuously monitors system & browser

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>Overhead (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Craigslist</td>
<td>1.1</td>
</tr>
<tr>
<td>CNN</td>
<td>3.1</td>
</tr>
<tr>
<td>ESPN</td>
<td>2.5</td>
</tr>
<tr>
<td>NY Times</td>
<td>1.3</td>
</tr>
<tr>
<td>Slashdot</td>
<td>1.4</td>
</tr>
<tr>
<td>Yahoo</td>
<td>1.0</td>
</tr>
<tr>
<td>YouTube</td>
<td>3.1</td>
</tr>
</tbody>
</table>

Overhead < 3.2%
Accuracy: Methodology

• Timestamp and Timer
Accuracy: Methodology

• Timestamp and Timer
• Experiment scenarios
  – No browsing

No competing browsing activity
Accuracy: Methodology

• Timestamp and Timer
• Experiment scenarios
  – No browsing
  – Passive browsing

Only asynchronous HTTP requests
Accuracy: Methodology

• Timestamp and Timer
• Experiment scenarios
  – No browsing
  – Passive browsing
  – Active browsing

Simultaneous web page loads in progress
Accuracy: Methodology

• Timestamp and Timer
• Experiment scenarios
  – No browsing
  – Passive browsing
  – Active browsing
  – Cross-traffic on the host

Saturate the local N/W
Accuracy: Results

• Timestamp

Accuracy of 1ms under normal operating conditions
Accuracy: Results

• Timestamp
  – Accuracy of 1ms under normal operating conditions
  – Degrades with high cross-traffic on host
    • Use `fathom.utils.metrics` to check host traffic
Accuracy: Results

• Timestamp
  – Accuracy of 1ms under normal operating conditions
  – Degrades with high cross-traffic on host
    • Use `fathom.utils.metrics` to check host traffic

• Timer
  – Degrades with concurrent browsing activity

Browser tabs share the same runtime
Accuracy: Results

• Timestamp
  – Accuracy of 1ms under normal operating conditions
  – Degrades with high cross-traffic on host
    • Use `fathom.utils.metrics` to check host traffic

• Timer
  – Degrades with concurrent browsing activity
    • Check wall-clock time using JavaScript’s `Date` API
Case Studies

• Netalyzr (IMC’10)
  – Researchers can use Fathom to build complex toolkits
  – Fathom-powered Netalyzr confirms its versatility
Case Studies

• Netalyzr (IMC’10)
  – **Researchers** can use Fathom to build complex toolkits
  – Fathom-powered Netalyzr confirms its versatility

• Web access failure
  – “Debug my connection” application
  – Incentive for **end-users** to install Fathom
Case Studies

• Netalyzr (IMC’10)
  – Researchers can use Fathom to build complex toolkits
  – Fathom-powered Netalyzr confirms its versatility

• Web access failure
  – “Debug my connection” application
  – Incentive for end-users to install Fathom

• Web services debugging (Google Maps)
  – Web developers can use Fathom APIs to assist users
Outline

• Concrete Example
• Design
• Implementation
• Evaluation
• Conclusion
Conclusion

Fathom is a practical browser-based network measurement platform

– Programmable interface for writing and launching measurements from web pages
– Prototype available for Firefox web browser
– Low runtime overhead and acceptable accuracy
– Built-in checks to ensure user security and privacy
Thank You.

Download Fathom at http://fathom.icsi.berkeley.edu/

Contact: fathom@icsi.berkeley.edu