Improving Web Availability for Clients with MONET

David G. Andersen
Carnegie Mellon University
dga@cs.cmu.edu

Hari Balakrishnan, M. Frans Kaashoek,
Rohit N. Rao
MIT Computer Science and Artificial
Intelligence Lab
fhari,kaashoek,rohitg@csail.mit.edu

http://nms.csail.mit.edu/ron/ronweb
Contents

• Problem Description
• Multi-Homed Overlay Network (MONET)
• MONET Test Network
• General Failures
• The ICP+ Protocol Probing a Path Through a Peer
• Waypoint Selection Algorithm
• MONET Data Collection
• Performance
• Conclusion
• Remarks & Discussion
Problem Description

• Clients experience outage rates of a few percent when accessing Web site

• To address the problems of availability in the Internet the following has been done
  • multi-homing
  • path and data redundancy,
  • Increase in capacity

• Results
  • Availability 95-99%. Outages are still observed in the network

• Solution
  • MONET improves client availability to web sites
  • MONET reduces periods of downtime and exceptional delays that lead to a poor user experience.
Multi-Homed Overlay Network (MONET)

- Set of Web proxies deployed across the Internet
- One site might have one or a few proxies

MONET addresses two questions:
- 1. How to obtain multiple paths from a client to a site?
- 2. Which path(s) to use, and at what times?

MONET finds these paths in three ways:
- Using of a combination of link multi-homing and cooperative overlay network of peer proxies
- Forwarding requests and responses via a small overlay network of peer MONET proxies
- Contacting multiple server replicas.
MONET Test Network

Clients

CSAIL Proxy

Cogent

DSL

Internet

UUNET

ELI

Aros

MIT

Internet2

Genuity

UTAH

Towerstream

NYU

100 Mbps

100 Mbps

1.5Mbps

Measurement Time
6 Dec-27 Jan 2004

1.5Mbps

100 Mbps

Measurement Time
24 Jan-04 Feb 2004

NYU Proxy

Aros Proxy

Utah Proxy

Mazu Proxy

Saved Traces

Cogent

MIT

NYU

“Internet”

“Internet2”

Genuity

Towerstream

Aros

UTAH

UUNET

ELI

INTERNET

INTERNET2

Genuity

UUNET

ELI

Aros

UTAH

1.5Mbps

100 Mbps

100 Mbps

1.5Mbps
General Failures

- Failures generally can prevent a client from accessing the Web site
  - Client's access link may be down
  - Name System (DNS) may not respond or may have incorrect information
  - Misconfigurations
  - Congestion,
  - Routing pathologies
  - Server itself or its access network may be down.
The ICP+ Protocol Probing a Path Through a Peer

Client Proxy                  Peer Proxy                  Web Server

ICP QUERY

ICP MISS

RTT=...

GET x

DATA

DATA

SYN

SYN/ACK

GET x (+ACK)
Parallel Client Proxy Queries

1. Request starts
2. Local DNS Resolution
3. Peer Proxy Probe
4. Local TCP Connections
5. Fetch via first responder
Waypoint Selection Algorithm

- A waypoint selection algorithm,
  - Picks a good small subset of all available paths
  - Uses probes that check the availability of multiple underlying components
  - Dynamically decides the order in which the many possible paths between client and server should be used,
  - What time to use any given path

- Returns a static list of (path, delay) pairs that it chooses based upon the name of the destination Web site
- Maintains statistics about path success rates and connection times through different interfaces and peers

- Reducing overhead
  - After opening one connection successfully, MONET closes the remaining probe connections.
MONET Data Collection

• MONET proxies recorded the following:
  
  • Cached objects
    • Request Time
  
  • Uncached objects
    • DNS Resolution duration
    • TCP connect duration
    • ICP+ duration
### Performance (1/3)

<table>
<thead>
<tr>
<th>Failure Type</th>
<th>CSAI L</th>
<th>MAZU</th>
<th>MAZU</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MIT</td>
<td>Cog</td>
<td>DSL</td>
</tr>
<tr>
<td>DNS</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Site RST</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Site Unreach</td>
<td>173</td>
<td>173</td>
<td>173</td>
</tr>
<tr>
<td>Client Access</td>
<td>152</td>
<td>14</td>
<td>2016</td>
</tr>
<tr>
<td>Wide-area</td>
<td>201</td>
<td>238</td>
<td>1828</td>
</tr>
<tr>
<td><strong>Availability</strong></td>
<td><strong>99.6%</strong></td>
<td><strong>99.7%</strong></td>
<td><strong>97%</strong></td>
</tr>
</tbody>
</table>
Performance (2/3)

The diagram illustrates the fraction of successful connections over time for different setups. The x-axis represents the time (in seconds) taken for DNS and connect operations, while the y-axis shows the fraction of successful connections.

Key points on the diagram include:
- All concurrently
- MONET
- MIT+Cogent+DSL
- MIT + ICP peers
- MIT
- Cogent
- DSL

The curves indicate how well each setup handles the connection process under varying time conditions.
Performance (3/3)

Fraction successful connects

<table>
<thead>
<tr>
<th>Route</th>
<th>All paths</th>
<th>Cogent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cog+MIT+DSL</td>
<td>solid</td>
<td></td>
</tr>
<tr>
<td>Cog+DSL</td>
<td>dashed</td>
<td></td>
</tr>
<tr>
<td>Cog+ICP</td>
<td>dotted</td>
<td></td>
</tr>
<tr>
<td>MIT+ICP</td>
<td>dash-dot</td>
<td></td>
</tr>
<tr>
<td>MIT</td>
<td>dash-dot</td>
<td></td>
</tr>
</tbody>
</table>

- MONET best – multihomed servers
- MONET best – all servers
- Baseline performance

All servers

Multi-homed servers
Conclusion

- MONET, a Web proxy system to improve the end-to-end client-perceived availability of accesses to Web sites.
- MONET masks several kinds of failures that prevent clients from connecting to Web sites, including access link failures, Internet routing failures, DNS failures, and subset of server-side failures.
Remarks & Discussion

• Failures are maskable, except for server failures themselves.

• If the different paths available between a proxy and server all share a single point of failure (e.g., a particular network link, a misconfigured DNS database, etc.), MONET will not mask the failure of that element.

• The current MONET implementation does not mask mid-stream failures that might occur in the middle of a TCP connection; such failures may be recovered from by issuing appropriate HTTP range requests or using transport-layer techniques.

• Access Authorization and update in time?

• Link capacity seems to play a role in availability. Link capacities vary in the internet and middle boxes