Characterizing Global Transnational Internet Performance and its Bottleneck

Pengxiong Zhu*, Keyu Man*, Zhongjie Wang*, Zhiyun Qian*, Roya Ensafi+, J. Alex Halderman+, Haixin Duan++

*University of California, Riverside
+University of Michigan, ++Tsinghua University

Accepted in ACM SIGMETRICS’ 20
Motivation

• Grew up in China, I rarely visit foreign websites
Motivation

• Visiting foreign website is slow

• Is this unique to China?

• Transnational network performance
  • Indicator of economic development, infrastructure investment, globalization of the countries

• How can we help improve the performance?
  • Characterization of bottleneck
Goal

• How well traffic flows across national borders?
  • Throughput?
  • How often slow speed happen
• What factors influence transnational performance?

World Submarine Cable Map

Global Transnational Throughput Measurement
Global Transnational Throughput Measurement: Experiment Setup

• Vantage Points
  • Cloud servers (VPS) [red]
  • Residential hosts [blue]
Global Transnational Throughput Measurement: Experiment Setup

- April 22, 2019 – April 27, 2019
Answers
How well traffic flows across national borders?

Slow and Unstable!

For the rest of this talk, we use “China” to refer specifically to mainland China.
Africa vs China*

*For the rest of this talk, we use “China” to refer specifically to mainland China.

China’s upstream is fast and stable

Boxplot of Throughput of global transnational links

*For the rest of this talk, we use “China” to refer specifically to mainland China.
Africa vs China

China’s transnational network performance has unique diurnal pattern.

(a) Lagos, Nigeria and Virginia, USA  (b) Cairo, Egypt and Virginia, USA  (c) Beijing, China and Virginia, USA

Throughput patterns from April 23 to April 27, 2019 (Beijing Time) for three receiver–sender pairs (a single connection is used in each pair).
China’s Transnational Throughput Measurement
Goals

• How widespread is the phenomenon throughout China?
• What are the performance characteristics?
• Where are the causes?
Goals

• How widespread is the phenomenon throughout China?
  • Locations specific? Time specific? Domestic traffic? Real-life web servers?

• What are the performance characteristics?

• Where are the causes?
Experiment Setup
Is it happening only in specific locations?

- Space Coverage: more vantage points in China
Is it happening only in specific time?

- Is it happening only in specific time?
  - Mar 27, 2019 - Apr 1, 2019
  - Apr 27, 2019 - May 2, 2019
  - Jul 28, 2019 - Aug 7, 2019
  - Sep 4, 2019 - Oct 3, 2019
  - 4 times, 53 days
Does a similar slowdown occur for domestic traffic?

- Add domestic traffic measurement
Does it happen to real-life web servers?

- Top Alexa Website Evaluation
  - May 31, 2019 – July 31, 2019 (61 days)

<table>
<thead>
<tr>
<th>Rank</th>
<th>Domain</th>
<th>File Domain</th>
<th>File Type</th>
<th>File Size(MB)</th>
<th>Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Baidu.com</td>
<td>downpack.baidu.com</td>
<td>apk</td>
<td>8.1</td>
<td>Chinese</td>
</tr>
<tr>
<td>3</td>
<td>Qq.com</td>
<td>dldir1.qq.com</td>
<td>apk</td>
<td>97</td>
<td>Chinese</td>
</tr>
<tr>
<td>4</td>
<td>Sohu.com</td>
<td>3g.k.sohu.com</td>
<td>apk</td>
<td>17</td>
<td>Chinese</td>
</tr>
<tr>
<td>5</td>
<td>Taobao.com</td>
<td>download.alicdn.com</td>
<td>apk</td>
<td>103</td>
<td>Chinese</td>
</tr>
<tr>
<td>8</td>
<td>Jd.com</td>
<td>storage.360buying.com</td>
<td>apk</td>
<td>81</td>
<td>Chinese</td>
</tr>
<tr>
<td>9</td>
<td>Sina.com.cn</td>
<td>downapp.sina.cn</td>
<td>apk</td>
<td>30</td>
<td>Chinese</td>
</tr>
<tr>
<td>31</td>
<td>Bing.com</td>
<td><a href="http://www.bing.com">www.bing.com</a></td>
<td>json</td>
<td>3.79</td>
<td>US</td>
</tr>
<tr>
<td>48</td>
<td>Yandex.ru</td>
<td>an.yandex.ru</td>
<td>js</td>
<td>1.1</td>
<td>Russian</td>
</tr>
<tr>
<td>50</td>
<td>Github.com</td>
<td>codeload.github.com</td>
<td>zip</td>
<td>20</td>
<td>US</td>
</tr>
<tr>
<td>68</td>
<td>Microsoft.com</td>
<td>download.microsoft.com</td>
<td>exe</td>
<td>15</td>
<td>US</td>
</tr>
<tr>
<td>76</td>
<td>Apple.com</td>
<td><a href="http://www.apple.com">www.apple.com</a></td>
<td>mp4</td>
<td>58</td>
<td>US</td>
</tr>
<tr>
<td>93</td>
<td>Sciencedirect.com</td>
<td>holdings sciencedirect.com</td>
<td>zip</td>
<td>11</td>
<td>US</td>
</tr>
<tr>
<td>123</td>
<td>Mail.ru</td>
<td>rfr.agent.mail.ru</td>
<td>exe</td>
<td>56</td>
<td>Russian</td>
</tr>
<tr>
<td>125</td>
<td>Nih.gov</td>
<td>obsrr.od.nih.gov</td>
<td>pdf</td>
<td>2.5</td>
<td>US</td>
</tr>
<tr>
<td>146</td>
<td>Ebay.com</td>
<td>developer.ebay.com</td>
<td>zip</td>
<td>35</td>
<td>US</td>
</tr>
<tr>
<td>190</td>
<td>Springer.com</td>
<td>link.springer.com</td>
<td>pdf</td>
<td>7.6</td>
<td>US</td>
</tr>
</tbody>
</table>

Top Alexa domains used to measure the impact of the Great Bottleneck of China on Chinese users.
Answers
Is it happening only in specific locations?

- No
Is it happening only in specific locations?

- No

Average number of slowdown hours per day by sender. Missing bars indicate the absence of the sender in the corresponding experiment.
Is it happening only in specific time?

• No

CDF of hours of slowdown per day for connections from 18 Chinese to 17 foreign vantage points.
Does it happen to real-life web servers?

• Yes

CDF of hours of slowdown per day for connections from 18 Chinese to 17 foreign vantage points.
Does it happen to real-life web servers?

- Yes
Does a similar slowdown occur for domestic traffic?

• No

![Diagram showing average slowdown hours for different websites](image-url)
Goals

• How widespread is the phenomenon throughout China?

• What are the performance characteristics?
  • What factors influence performance?
    • Traffic direction? Downstream or upstream?
    • Path? Time of day effect?
  • Does the slowdown look irregular?
    • Traffic Differentiation
    • Where is the bottleneck?

• Where are the causes?
What factors influence performance?

1. The slowdown follows varied diurnal patterns.
What factors influence performance?

2. Packets are lost from only the direction going into China
What factors influence performance?

• 3. In conformity with a normal congestion
Does the slowdown look irregular?

- Protocol
  - HTTP, HTTPS, Shadowsocks, VPN

- Packet type
  - TCP(SYN, SYN/ACK, ACK), UDP, ICMP

- Speed
  - Send at varying rate ignoring TCP congestion control

- No noticeable differences
Where is the bottleneck?

• The bottleneck almost always occurs within China
  • at a router managed by a Chinese ISP.
China’s unique network topology

- Oracle report[1]:
  - “Intra-net”
  - No foreign carrier allowed inside China
  - Network periphery for China exists outside of China.

China’s unique network topology

[Map showing China's network topology with connections to various countries and cities, including London, Amsterdam, Frankfurt, United States, Japan, South Korea, Kenya, Brazil, South Africa, Australia, Seattle, Bay Area, New York, DC, Los Angeles, Miami.]

Where is the bottleneck?

• The bottleneck mostly occurs within China.

Fig. 13. The location of the bottleneck hop.

Domestic infrastructure fails to match the expensive transnational infrastructure (e.g., submarine cable)?
Summary

1. How widespread throughout China?
   1) Prevalent nationwide everyday
   2) Top-Alexa websites too
   3) Not for domestic traffic

2. What are the performance characteristics?
   4) Packet loss in one direction
   5) Varied diurnal patterns.
   6) No irregular traffic throttling
   7) The bottleneck mostly within China.
Goals

• How widespread is the phenomenon throughout China?

• What are the performance characteristics?

• Where are the causes?
What are the causes?

- Regulation
- Financial Motivation
Hypothesis 1: Regulation

- Traffic Differentiation:
  - HTTPS/VPN/Shadowsocks: No noticeable differences

- Great Firewall: On-path system[1]
  - Only copy packets
  - 34.45%: GFW hops match the bottleneck hops

- Great Canon: In-path system[2]
  - Only intercepting a subset of traffic
  - Might re-purposed to perform general traffic throttling.

Hypothesis 2: Network Resources Provisioned

• China Academy of Information and Communications Technology[1]
  • “China's international submarine cable development still lags the world's other major economies. The number of submarine cables in the U.S., Japan, the U.K., and Singapore is eight, two times that of China, respectively, and the per capita bandwidth is 20, 10, 73, and 265 times that of China, respectively”

![Table 2-3 Comparison of submarine cables between China and major countries in the world](image)

Hypothesis 2: Network Resources Provisioned

- In early years, peering with Chinese ISPs is hard
  - Grow their own transit business
  - Make to top tier ISPs

- Tiered services
  - All the three state-owned ISPs have set up a premium transnational network
**Hypothesis 2: Network Resources Provisioned**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Product</th>
<th>ChinaNet Paid-peer</th>
<th>China Access</th>
<th>Global Transit</th>
<th>Global Internet Access</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access Feature</td>
<td>BGP-4</td>
<td>BGP-4</td>
<td>BGP-4</td>
<td>BGP-4 or Static</td>
<td></td>
</tr>
<tr>
<td>Routes and Policy</td>
<td>CT Routes</td>
<td>ISP’s Routes in Mainland China</td>
<td>Global or Customized Routes</td>
<td>Global or customized Routes</td>
<td></td>
</tr>
<tr>
<td>Chinese Resources</td>
<td>Best effort</td>
<td>Best effort</td>
<td>Best effort</td>
<td>Premium</td>
<td></td>
</tr>
<tr>
<td>Bandwidth limitation</td>
<td>No limit</td>
<td>No limit</td>
<td>No limit</td>
<td>&lt;10G (To China&lt; 1G)</td>
<td></td>
</tr>
<tr>
<td>SLA Support</td>
<td>N/A</td>
<td>N/A</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Routes limitation</td>
<td>No limit</td>
<td>No limit</td>
<td>No limit</td>
<td>300</td>
<td></td>
</tr>
<tr>
<td>BGP Community</td>
<td>N/A</td>
<td>N/A</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Burstable Billing</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>MRTG Report</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Locking Glass</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>IP Address Service</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>✓</td>
</tr>
<tr>
<td>Anti-DDoS (Black-hole)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Anti-DDoS (Clean-pipe)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>✓</td>
</tr>
</tbody>
</table>
Hypothesis 2: Network Resources Provisioned

- Tiered services: Different loss rates

<table>
<thead>
<tr>
<th>Service</th>
<th>Average Loss Rate</th>
<th>Max Loss Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>GIA</td>
<td>3%</td>
<td>7%</td>
</tr>
<tr>
<td>GT</td>
<td>4%</td>
<td>14%</td>
</tr>
<tr>
<td>China Paid-Peer</td>
<td>5%</td>
<td>15%</td>
</tr>
<tr>
<td>China Access</td>
<td>5%</td>
<td>15%</td>
</tr>
</tbody>
</table>

Why are they offering tiered services?
Would tiered service benefit them?
Summary

1. How widespread throughout China?
   1) Prevalent nationwide everyday
   2) Top-Alexa websites too
   3) Not for domestic traffic

2. What are the performance characteristics?
   4) Packet loss in one direction
   5) Varied diurnal patterns.
   6) No irregular traffic throttling
   7) The bottleneck mostly within China.

3. What are the causes?
   • Require potentially insider knowledge
Future Work

• Longer-term trend at a global scale
• More countries to study, e.g., Russia and Iran
Thank you!

We hope our effort can help administrators or policy makers to locate the slowdown problem and eliminate the bottleneck.