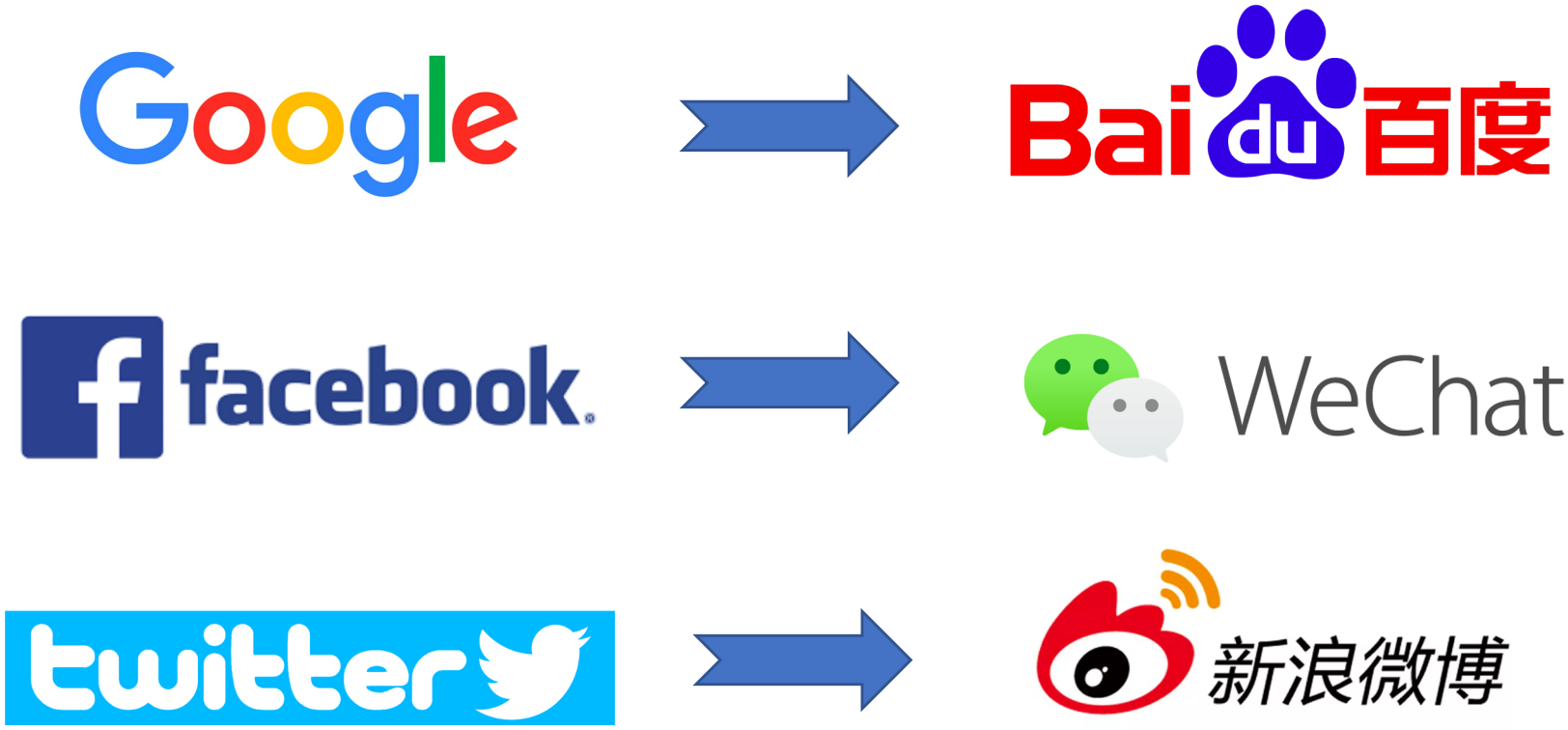
Characterizing Global Transnational Internet Performance and its Bottleneck

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



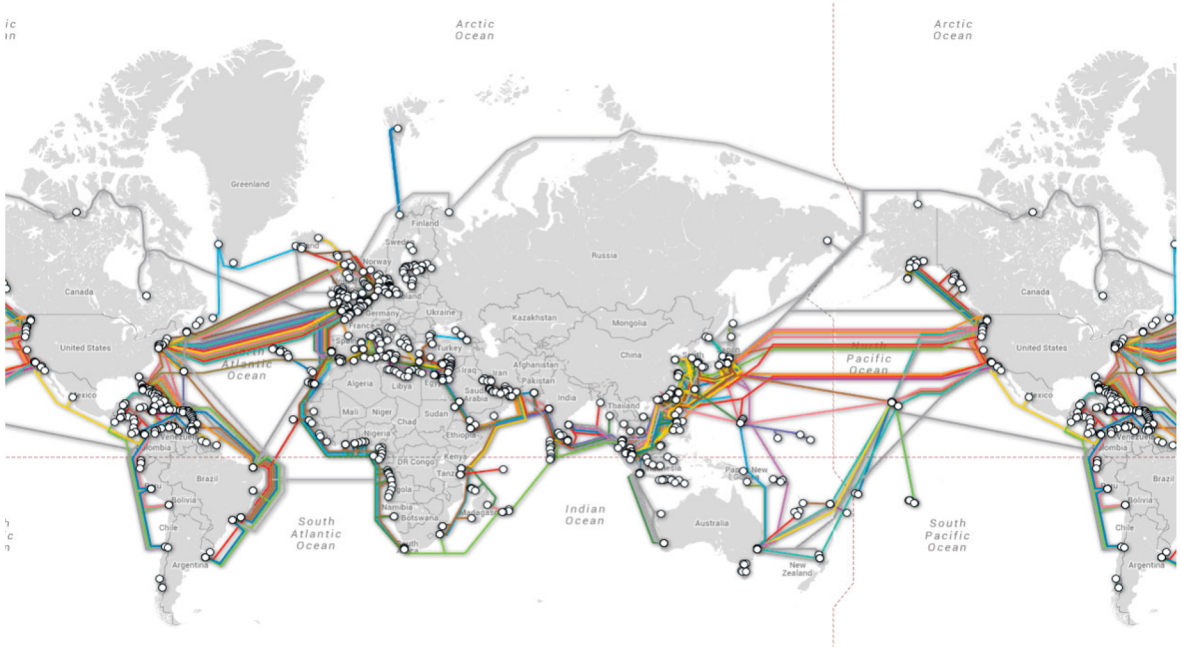
Motivation

- 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

http://www.wirelessinternetcapetown.co.za/this-is-what-south-africas-internet-actually-looks-like-mybroadband/

Global Transnational Throughput Measurement

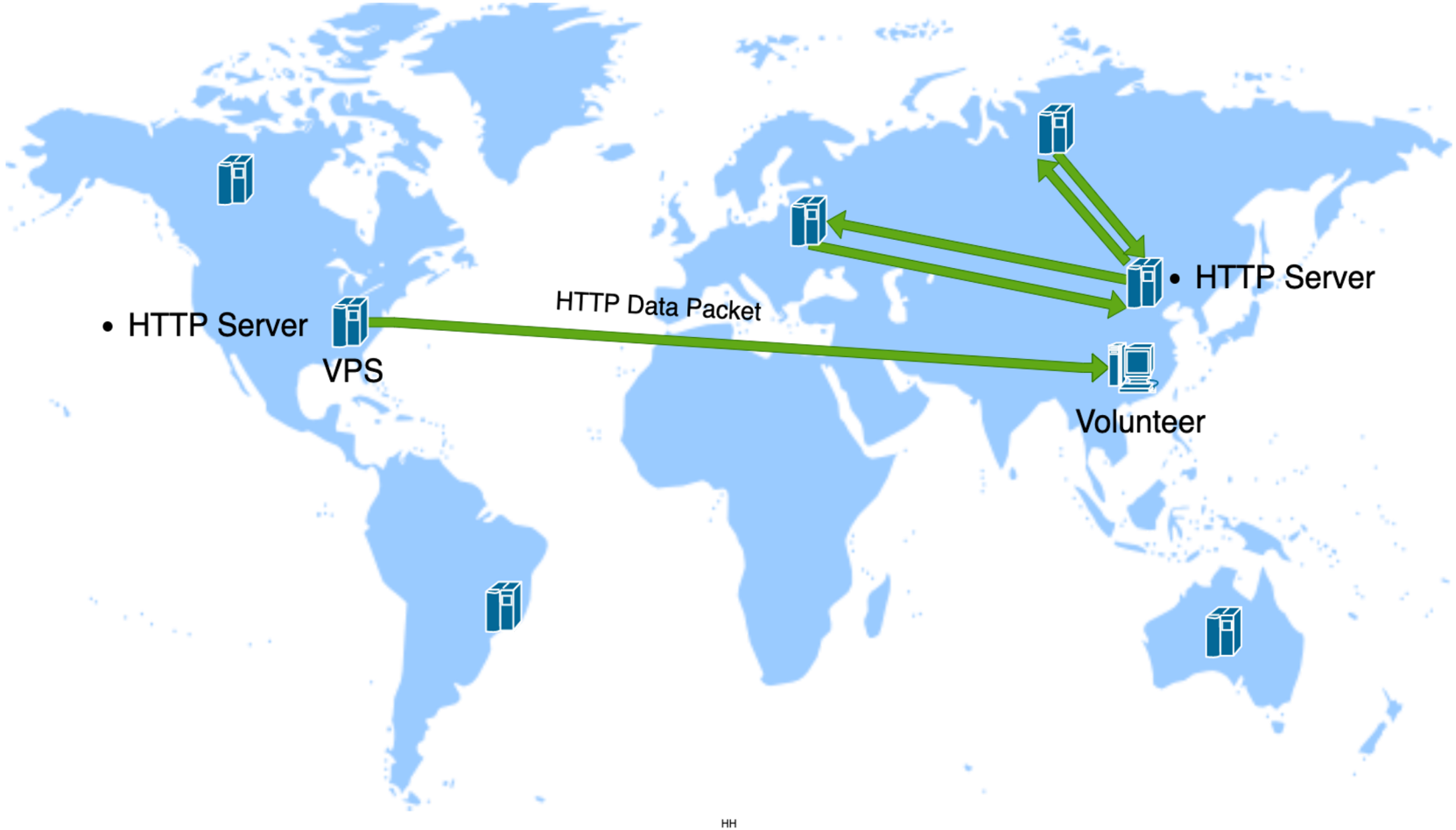
Global Transnational Throughput Measurement: Experiment Setup

- Vantage Points
  - Cloud servers(VPS) [red]
  - Residential hosts [blue]

Vantage Point Map

Global Transnational Throughput Measurement: Experiment Setup

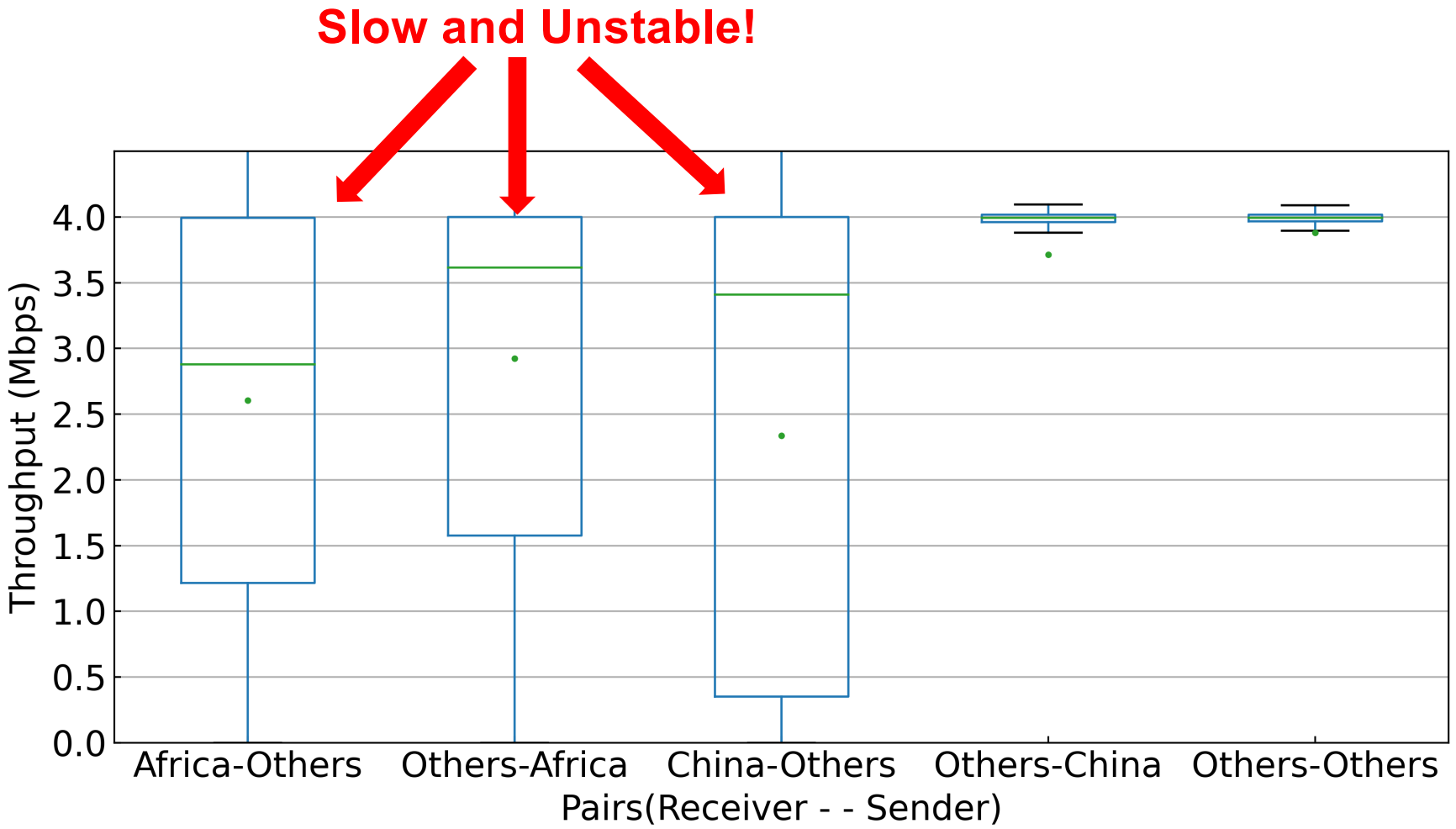
- April 22, 2019 – April 27, 2019



Experiment Setup

Answers

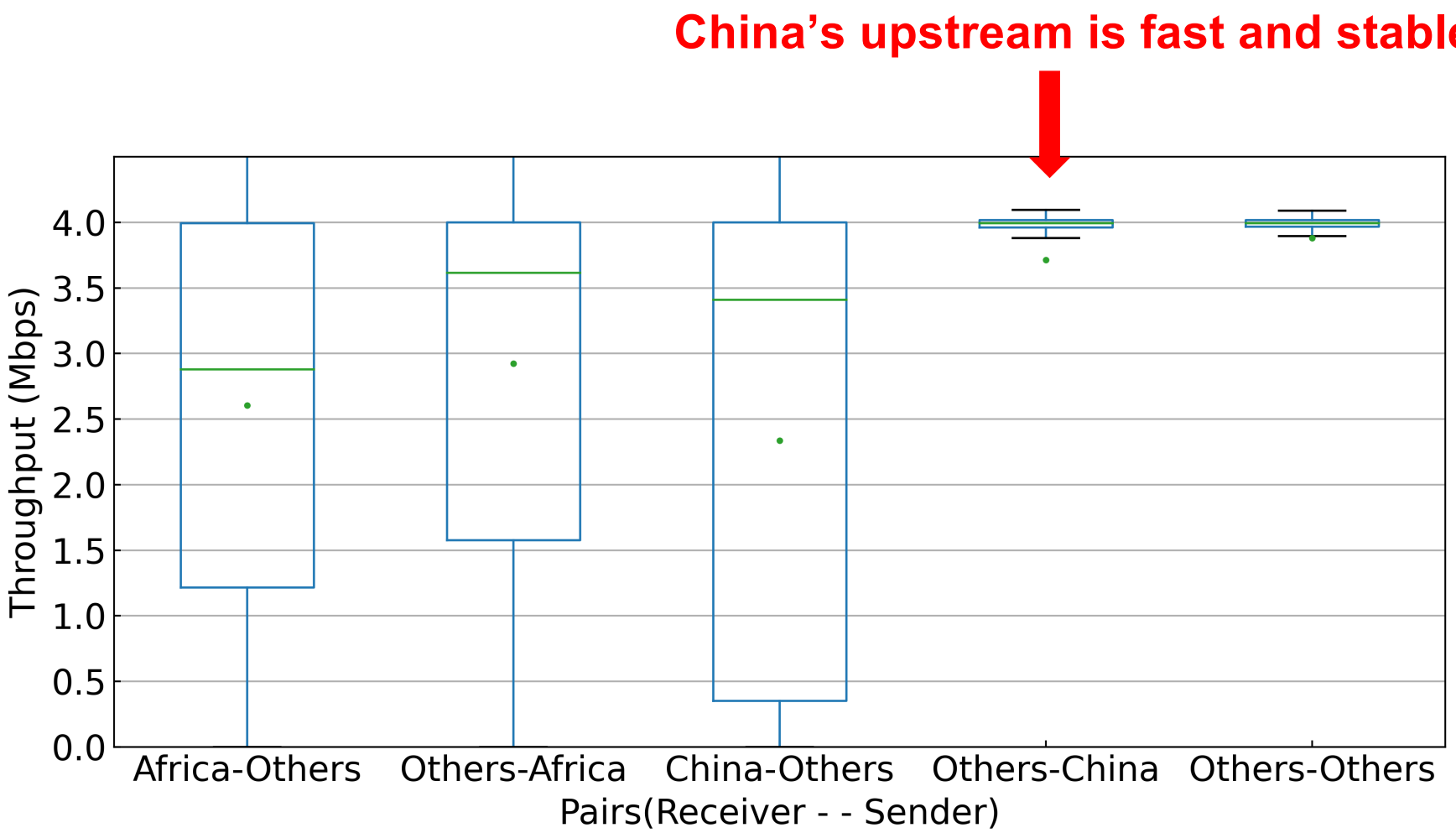
How well traffic flows across national borders?



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*

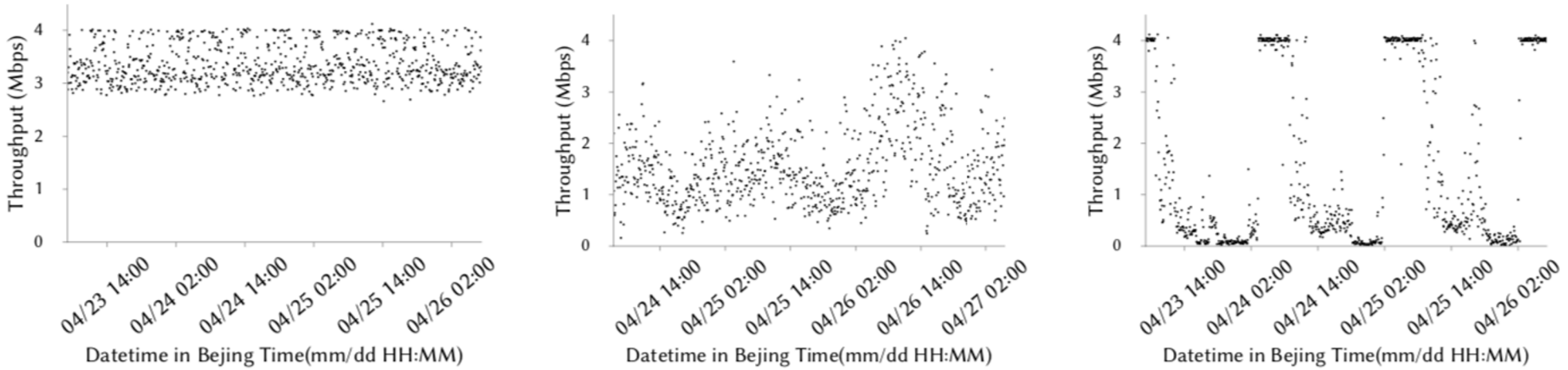


Boxplot of Throughput of global transnational links

*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

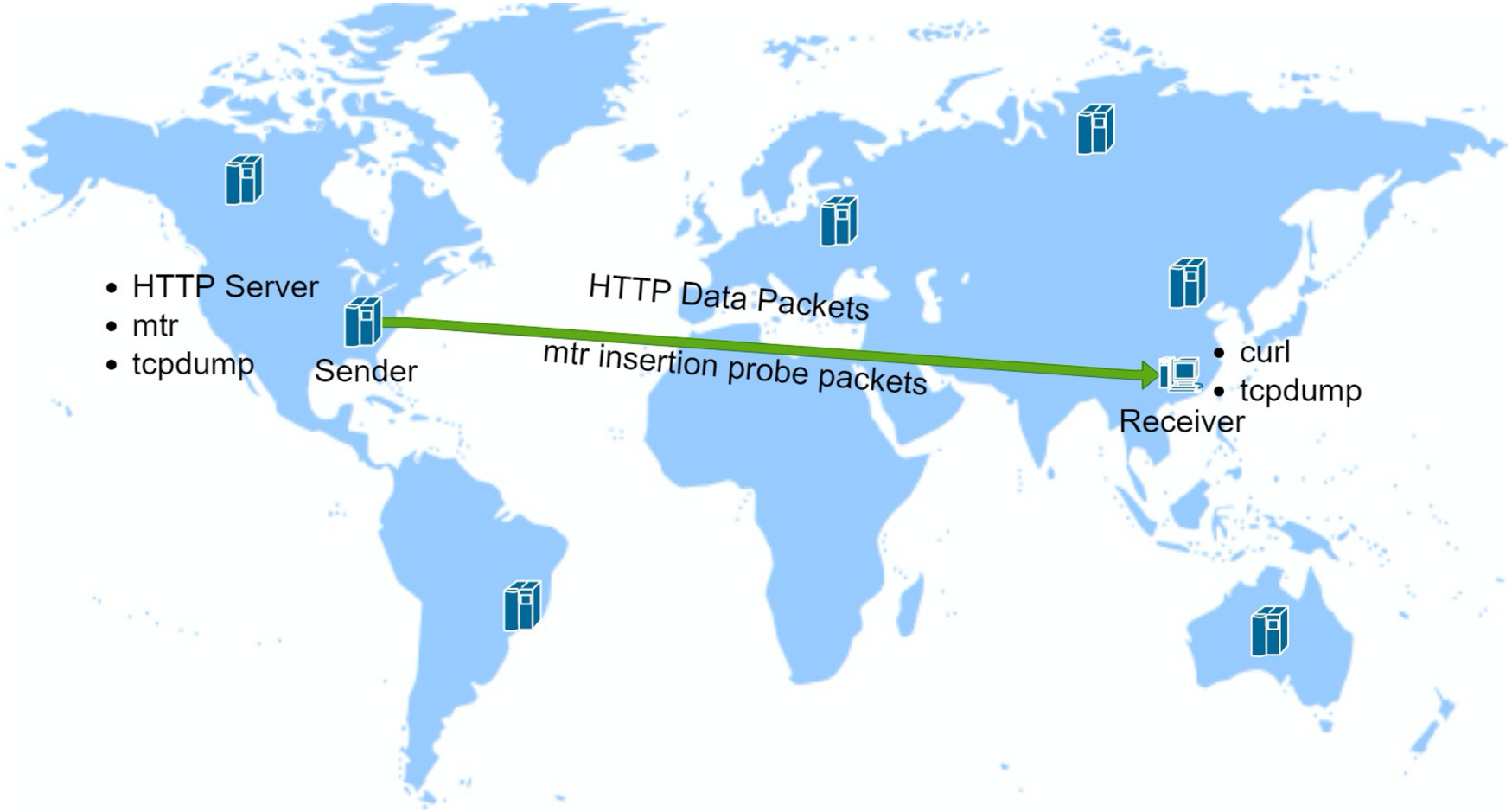
Geolocations of vantage points within China.

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



End-to-End Experiment Setup

Does it happen to real-life web servers?

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

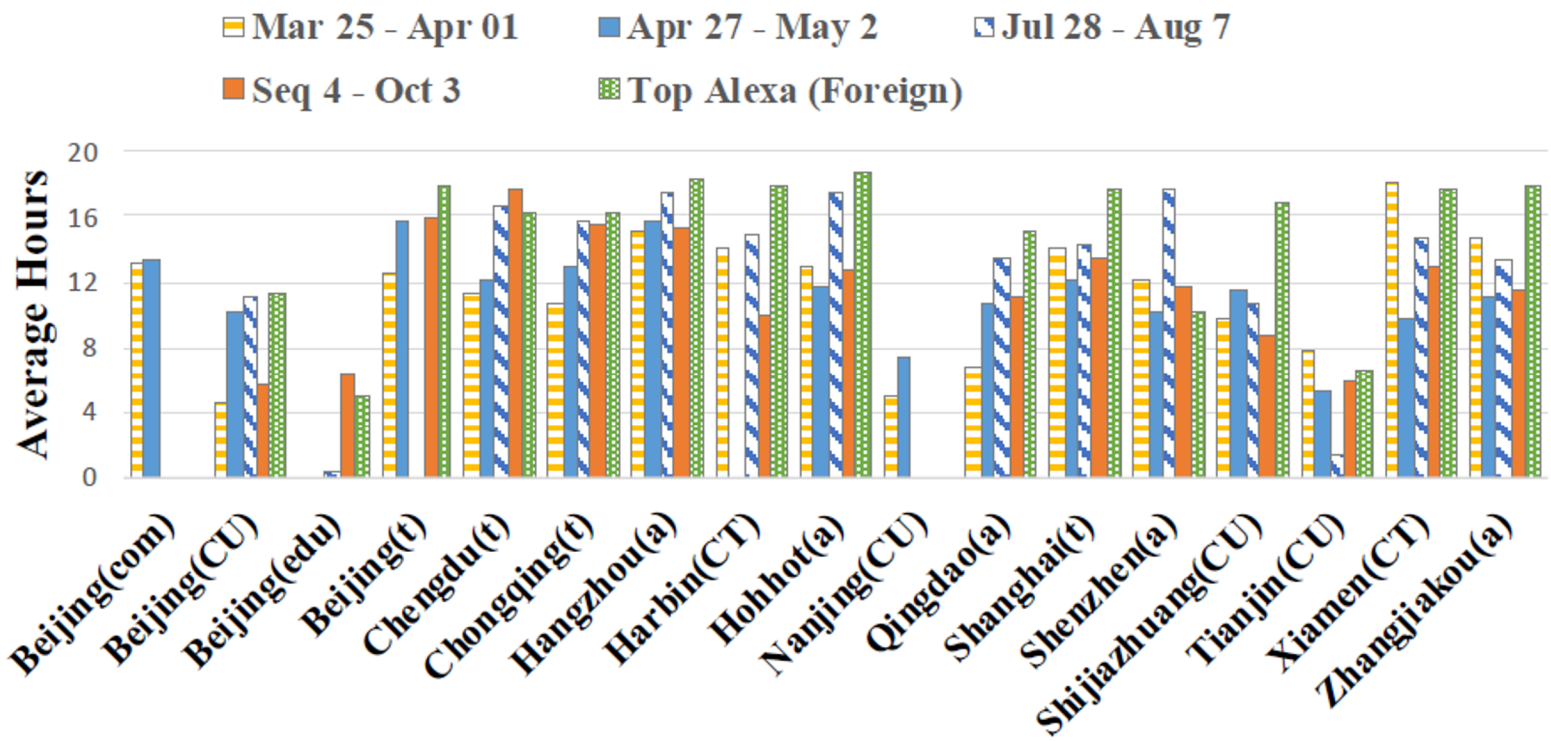
Rank	Domain	File Domain	File Type	File Size(MB)	Origin
1	Baidu.com	downpack.baidu.com	apk	8.1	Chinese
3	Qq.com	dldir1.qq.com	apk	97	Chinese
4	Sohu.com	3g.k.sohu.com	apk	17	Chinese
5	Taobao.com	download.alicdn.com	apk	103	Chinese
8	Jd.com	storage.360buyimg.com	apk	81	Chinese
9	Sina.com.cn	downapp.sina.cn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

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?

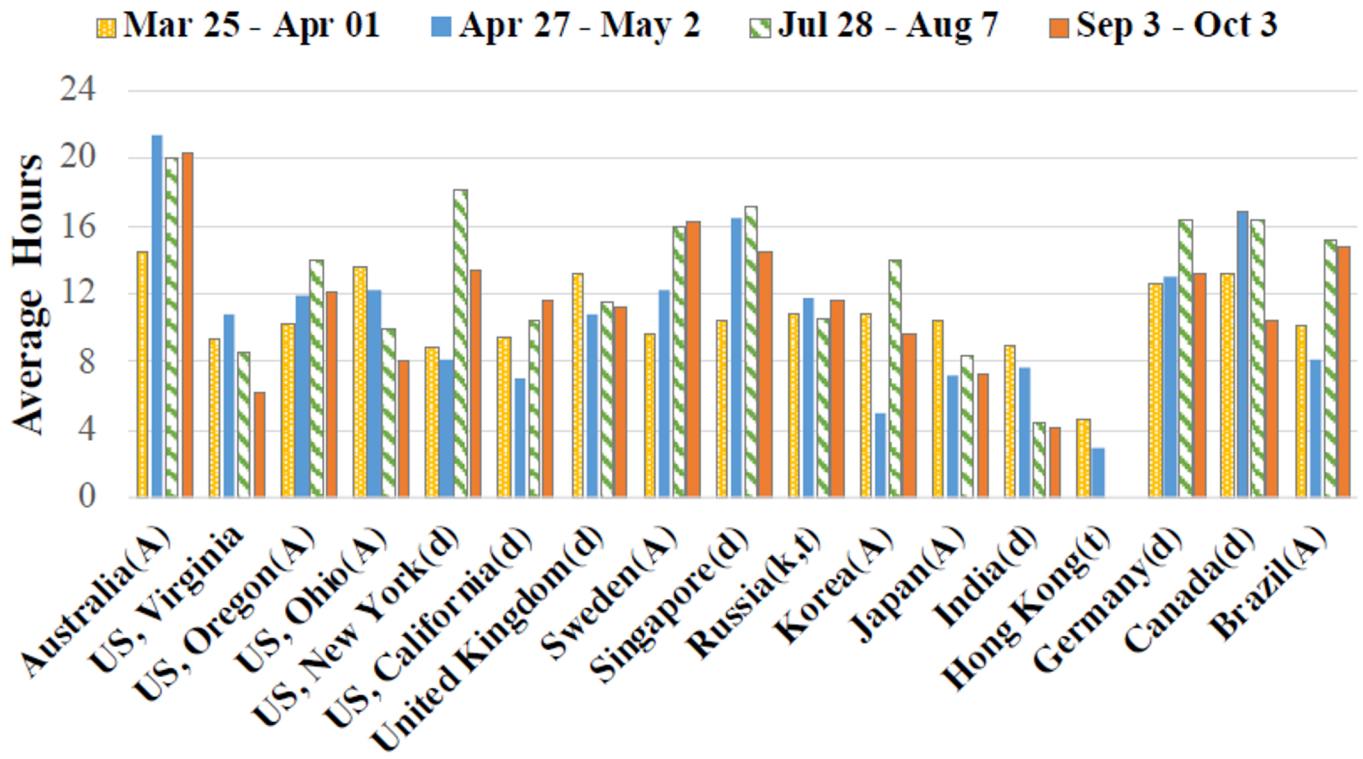
- No



Average number of slowdown hours per day by receiver. Missing bars indicate the absence of the receiver in the corresponding experiment. The ISPs labeled in the parentheses follow the convention in Table 2

Is it happening only in specific locations?

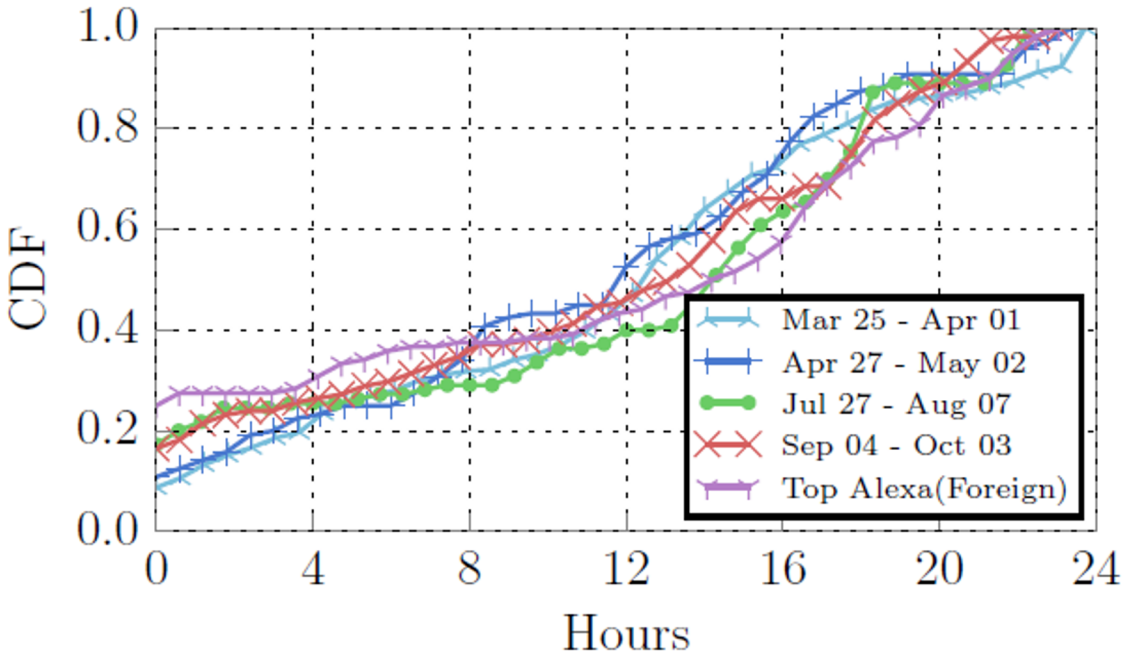
- No



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?

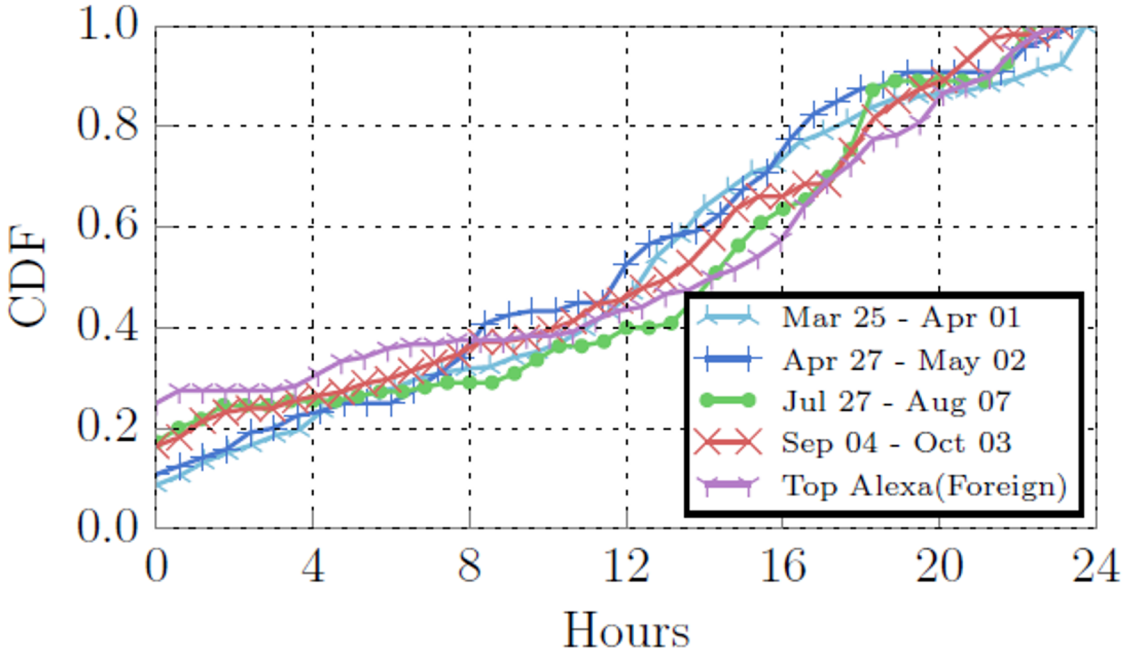
- No



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?

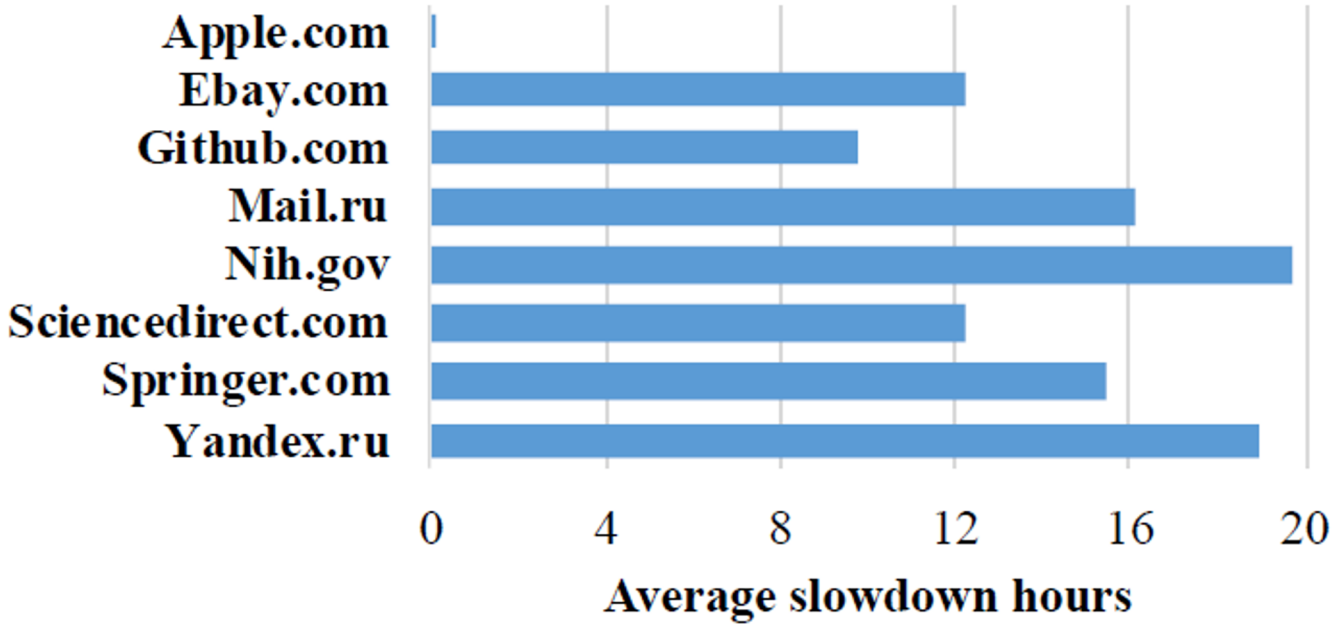
- Yes



CDF of hours of slowdown per day for connections from 18 Chinese to 17 foreign vantage points.

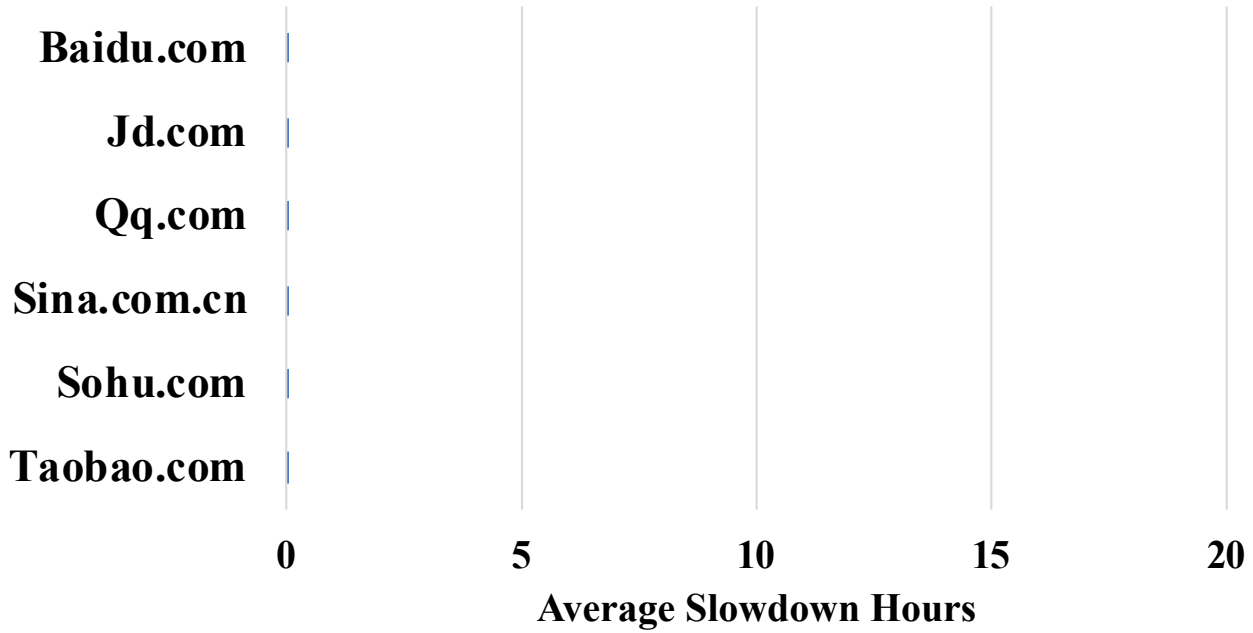
Does it happen to real-life web servers?

- Yes



Does a similar slowdown occur for domestic traffic?

- No

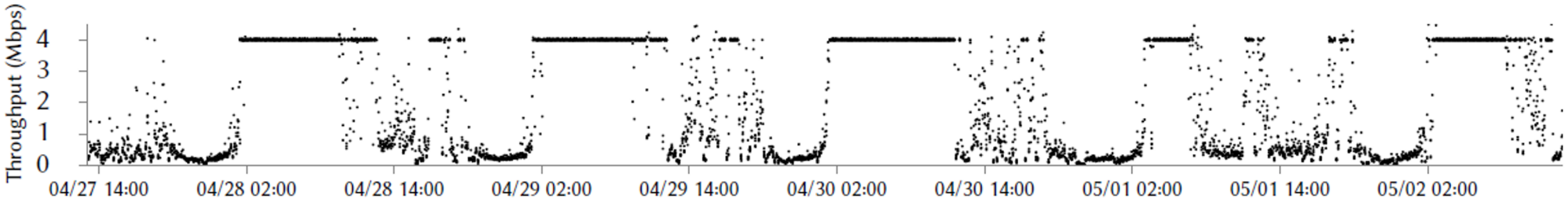


Goals

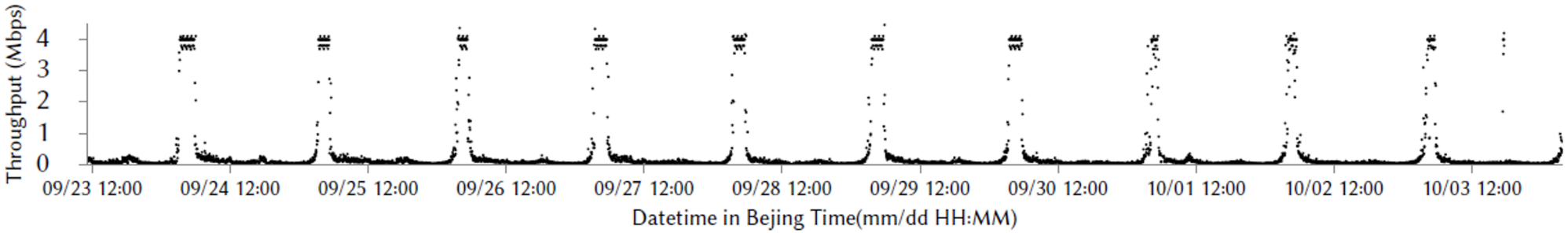
- How widespread is the phenomenon throughout China?
- What are the performance characteristics?
  - What factors influence performance?
    - Traffic direction? Downstream or upstream?
    - Path? Time of day effect?
  - Does the slowdown look irregular?
    - Traffic Differentiation
  - Where is the bottleneck?
- Where are the causes?

What factors influence performance?

1. The slowdown follows varied diurnal patterns.



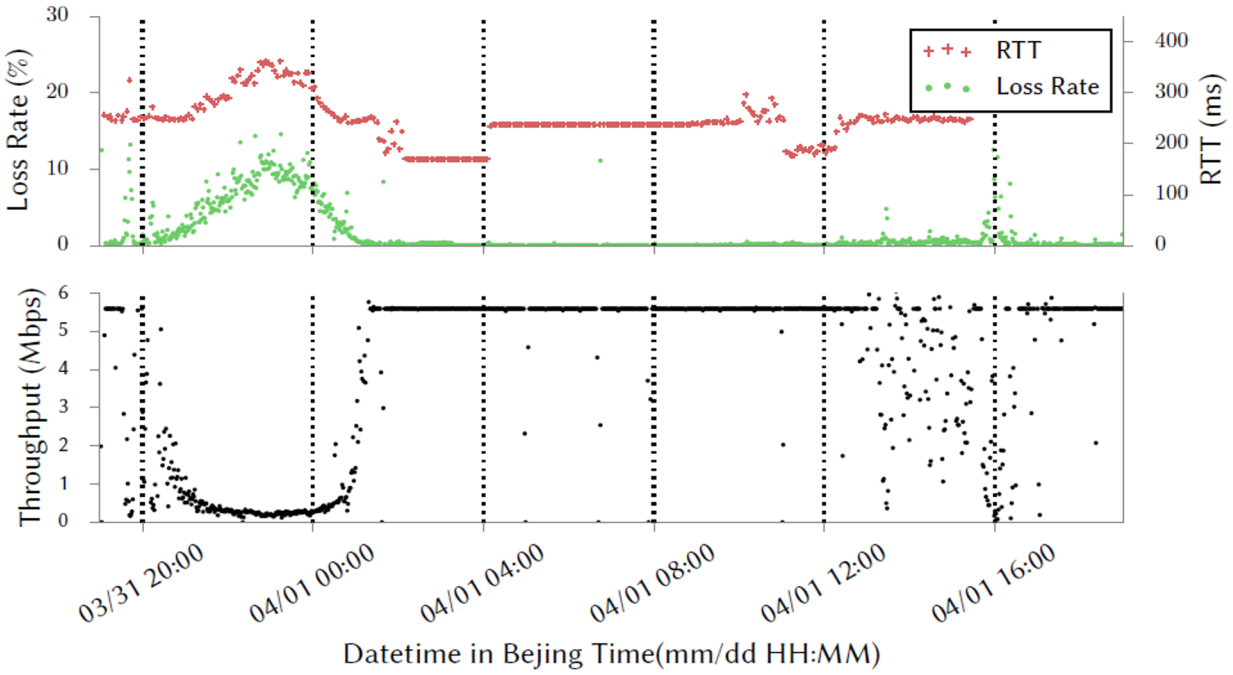
(a) Shanghai, China (VPS) from Oregon, USA, April 27 to May 2



(b) Beijing, China (residential) from Stockholm, Sweden, Sep 23 to Oct 03

What factors influence performance?

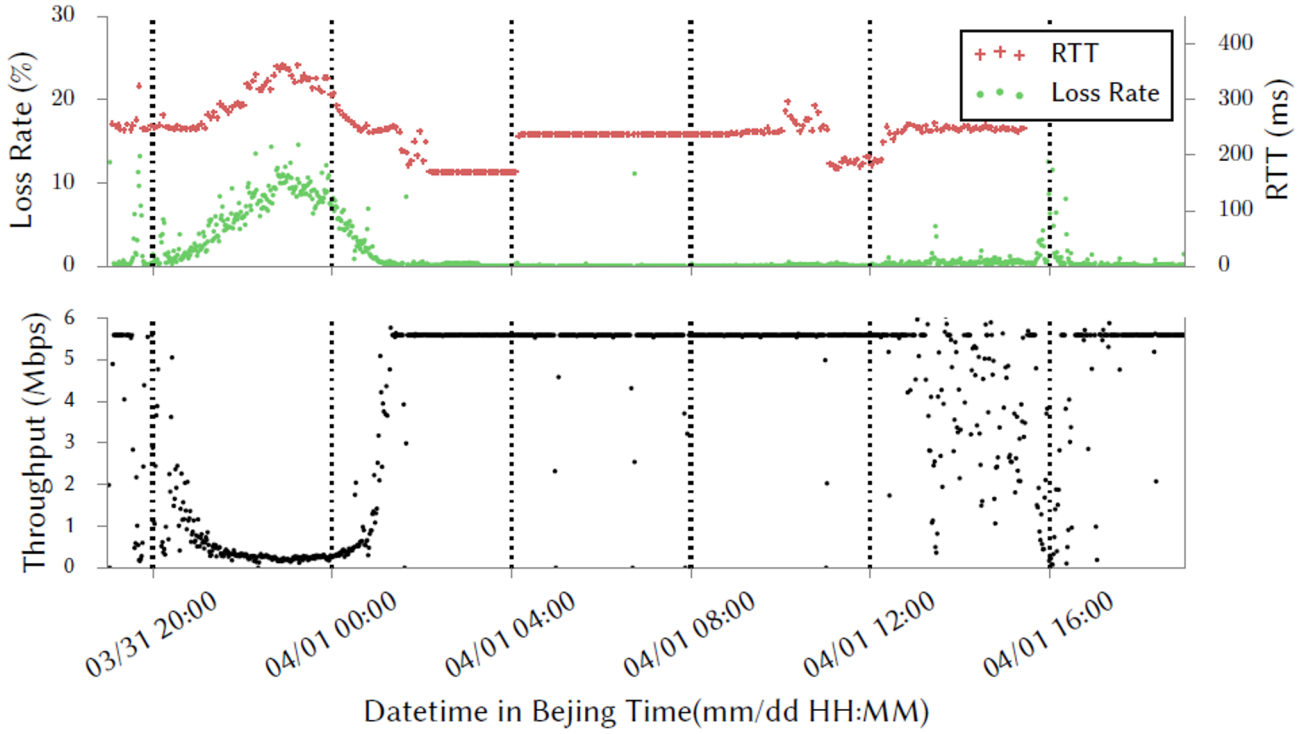
2. Packets are lost from only the direction going into China



Zhangjiakou, China (VPS) from London, United Kingdom, March 31

What factors influence performance?

- 3.In conformity with a normal congestion



Zhangjiakou, China (VPS) from London, United Kingdom, March 31

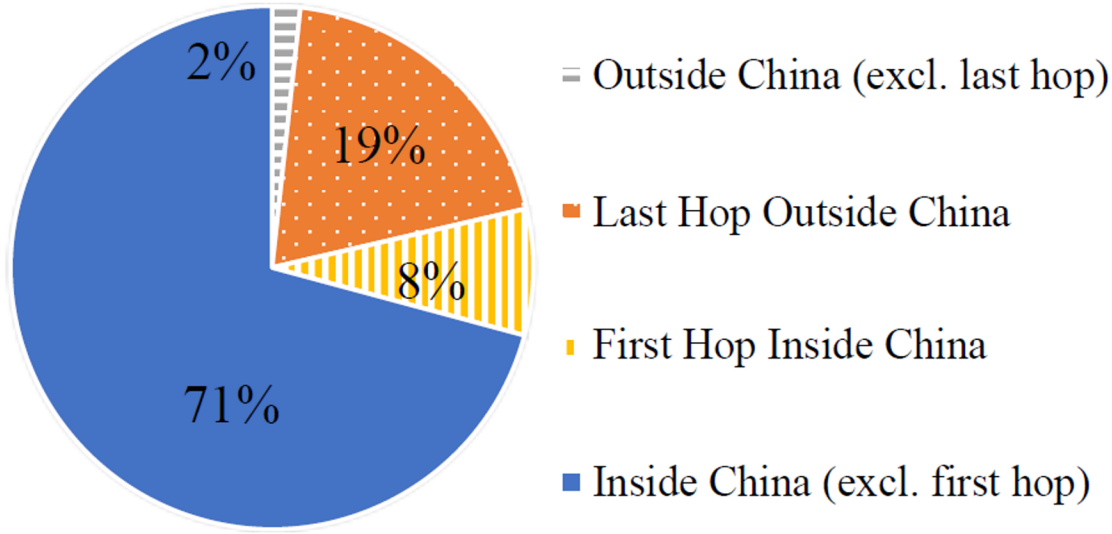
Does the slowdown look irregular?

- 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.



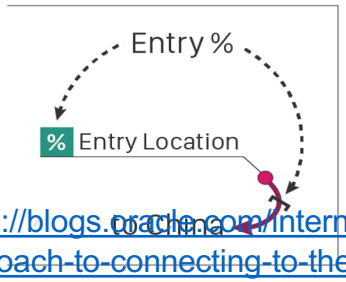
The location of the bottleneck hop.

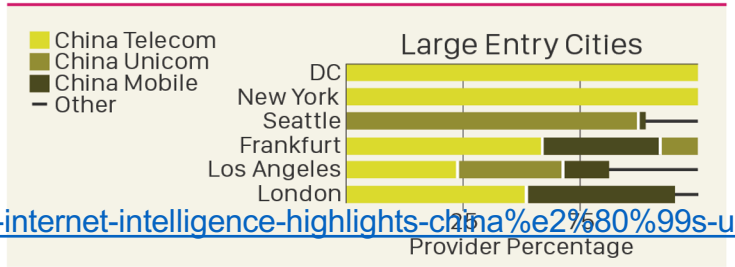
China’s unique network topology

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

[1] https://blogs.oracle.com/internetintelligence/analysis-by-oracle-internet-intelligence-highlights-china%e2%80%99s-unique-approach-to-connecting-to-the-global-internet

China's unique network topology





https://blogs.oracle.com/internetintelligence/analysis-by-oracle-internet-intelligence-highlights-china%e2%80%99s-unique-approach-to-connecting-to-the-global-internet

Where is the bottleneck?

- The bottleneck mostly occurs within China.

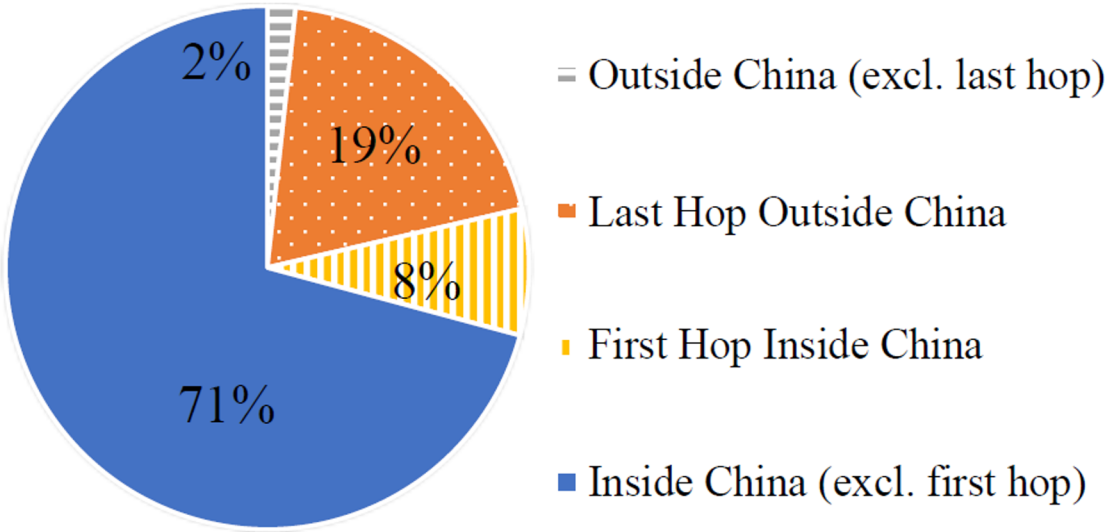


Fig. 13. The location of the bottleneck hop.

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

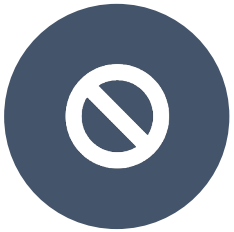
Summary

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.

[1] Zhongjie Wang, Yue Cao, Zhiyun Qian, Chengyu Song, and Srikanth V Krishnamurthy. 2017. Your state is not mine: a closer look at evading stateful internet censorship. In ACM SIGCOMM Conference on Internet Measurement Conference (IMC). ACM, 114–127.
[2] Bill Marczak, Nicholas Weaver, Jakub Dalek, Roya Ensafi, David Fifield, Sarah McKune, Arn Rey, John Scott-Railton, Ron Deibert, and Vern Paxson. 2015. An analysis of china's "great cannon". In USENIX Workshop on Free and Open Communications on the Internet.

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

	China	US	Japan	UK	Singapore
Number of submarine cables[4]	10	80	23	53	24
Total International bandwidth in 2017 (Gbps)	43445	201527	38799	151066	46544
Per capita international bandwidth (Mbps)	0.031	0.618	0.306	2.289	8.297

Source: TeleGeography and MIIT

[1] China Academy of Information and Communications Technology. 2018. White Paper on China International Optical Cable Interconnection. http://www.caict.ac.cn/english/yjcg/bps/201808/P020180829385778461678.pdf.

Hypothesis 2: Network Resources Provisioned

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

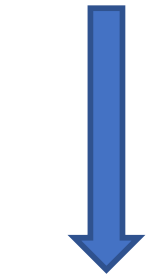
- Tiered services
  - All the three state-own ISPs have set up a premium transnational network

Hypothesis 2: Network Resources Provisioned

Feature \ Product	ChinaNet Paid-peer	China Access	Global Transit	Global Internet Access
Access Network	ChinaNet (AS 4134)	ChinaNet (AS 4134)	ChinaNet (AS 4134)	CN2 (AS 4809)
Access Feature	BGP-4	BGP-4	BGP-4	BGP-4 or Static
Routes and Policy	CT Routes	ISP's Routes in Mainland China	Global or Customized Routes	Global or customized Routes
Chinese Resources	Best effort	Best effort	Best effort	Premium
Bandwidth limitation	No limit	No limit	No limit	<10G (To China< 1G)
SLA Support	N/A	N/A	√	√
Routes limitation	No limit	No limit	No limit	300
BGP Community	N/A	N/A	√	√
Burstable Billing	√	√	√	√
MRTG Report	√	√	√	√
Looking Glass	√	√	√	√
IP Address Service	N/A	N/A	N/A	√
Anti-DDoS (Black-hole)	√	√	√	√
Anti-DDoS (Clean-pipe)	N/A	N/A	N/A	√

Hypothesis 2: Network Resources Provisioned

- Tiered services: Different loss rates



Service	Average Loss Rate	Max Loss Rate
GIA	3%	7%
GT	4%	14%
China Paid-Peer	5%	15%
China Access	5%	15%

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.