Towards Fine-Grained, High-Coverage Internet Monitoring at Scale

Hongyu Wu\textsuperscript{1}, Qi Ling\textsuperscript{2}, Penghui Mi\textsuperscript{3}, Chaoyang Ji\textsuperscript{3}, Yinliang Hu\textsuperscript{3}, Yibo Pi\textsuperscript{1}

\textsuperscript{1}Shanghai Jiao Tong University
\textsuperscript{2}University of Michigan
\textsuperscript{3}Huawei Cloud Computing Technologies Co., Ltd.
Public Internet: Bottleneck of Cloud Services

- Data Center
- Routers
- End User
- Bottleneck
- Public Internet
Monitoring public Internet is crucial
Challenge of Internet Monitoring

Large overhead

/8 network
Challenge of Internet Monitoring

How to efficiently achieve large-scale Internet monitoring?
- Two basic assumptions

1. **Similarity assumption**: Clients in the same /24 have similar paths
   - Only representative in each /24 needs to probed, reducing probing overhead

2. **Coverage assumption**: Tracking the performance to each /24 suffices for full-coverage monitoring

Both assumptions are challenged by the increasing prevalence of load balancing
Paths to addresses in the same /24 could be very different due to load balancing
Probing path to only representatives of /24s would leave many links uncovered
Methodology

Ground truth: all visible links

Use D-Miner [NSDI’20] to find all visible links

Flow paths

Modify Zmap to find the flow path from vantage points to each end user

Simulation

- Simulate real-world downstream traffic from DCs to clients
- Evaluate current practice against ground truth

Evaluate current practice against ground truth
VP and Target Selection

BABA: Beijing
AMZN: Tokyo
BABA&AMZN: Sydney
BABA&AMZN: Frankfurt
BABA&AMZN: Silicon Valley
AMZN: São Paulo
AMZN: Cape Town
BABA: Beijing
AMZN: Cape Town
BABA&AMZN: Sydney
VP and Target Selection

/8 network covers most IP addresses in the country where DC is located.
Similarity assumption: Clients in the same /24 have similar paths

Path difference = \frac{(|s_A \cup s_B| - |s_A \cap s_B|)}{(|s_A| + |s_B|)}

Path difference = \frac{(3 - 1)}{(2 + 2)} = 0.5
(discard last-hop link)
Evaluating Similarity Assumption of Current Practices

I24 Network

Amazon

<table>
<thead>
<tr>
<th>CDF</th>
<th>Path difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
</tr>
</tbody>
</table>

Alibaba

<table>
<thead>
<tr>
<th>CDF</th>
<th>Path difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
</tr>
</tbody>
</table>

x.x.x.1  x.x.x.2  x.x.x.(1+d)
Evaluating Similarity Assumption of Current Practices

Similarity assumption fails: The representative fails to indicate performance of entire /24
Evaluating Coverage Assumption of Current Practices

Current Practice: Selecting .1 addr of every /24 as representative

Link coverage = \( \frac{\text{Covered links}}{\text{Total links}} \)
Evaluating Coverage Assumption of Current Practices

Current Practice: Selecting .1 addr of every /24 as representative

Coverage assumption fails: Current practices leave 70% of links unwatched

Link coverage = \frac{#\text{Covered links}}{#\text{Total links}}
Evaluating Coverage Assumption of Current Practices

Current Practice: Selecting .1 addr of every /24 as representative

Events occurring to 70% of links will not be observed by current practices

Link coverage = \frac{\#Total\ links}{\text{Total links}}

Then, how to improve link coverage?
Traditional wisdom to boost link coverage

Traditional wisdom: Increase the granularity of monitoring.
Traditional wisdom to boost link coverage

Traditional wisdom: Increase the granularity of monitoring.

Boosting link coverage by increasing granularity is not scalable.
By monitoring end-hosts A, C, E, we can coverage all visible links.
Why Our End-to-End Approach is More Scalable?

Normalized link count = \( \frac{\text{# Total links}}{\text{#/24s}} \)

Links scale much slower than network size
Why Our End-to-End Approach is More Scalable?

Normalized link count = \( \frac{\# \text{ Total links}}{\#/24s} \)

Can we achieve high link coverage by carefully selecting targets?
A Greedy End-to-End Approach

Always choose targets contribute most new links

Selected Targets

Targets

Link Set

\[\text{max new links}\]

\[\text{discard duplications}\]
Evaluating the Greedy End-to-End Approach

- More difficult to achieve full coverage for larger network
- For /8, monitoring $x$ links only requires probing $0.6^*x$ targets
Trade-off between Coverage & Overhead

- Only the first 30% of overhead can discover >1 new links

- 1/3 overhead for 80% coverage
Takeaways

- Current practices fail to monitor the changes of a majority of links in the Internet, leaving critical links unwatched.

- High link coverage can be achieved by carefully selecting probing targets with reasonable overhead.

- Our dataset is published at https://github.com/SJTU-NMS-Lab/APNet23
Thank You!

Q & A
Future Directions

- IPv6 exploration

- Fast start without long-time data collection

- Real-time detection on link failure/congestion
Internet Monitoring: Passive vs Active

Passive Measurements

Active Measurements

Our Work
Our Contributions

- Evaluate the link coverage of two rule-of-thumb practices for scalable Internet measurement from a cloud-centric view

- Evaluate the predictability of performance for client flows to the same /24s

- Propose to achieve high-coverage monitoring with an end-to-end approach

- Estimate the overhead for high-coverage monitoring
- Tool: D-Miner [Vermeulen et al, NSDI’ 20]

- Divide /8 into /16s

- Send two back-to-back scans at 100,000pps

- Goal: Find ALL visible links at confidence level of 99%(95% for one scan)
## Dataset

<table>
<thead>
<tr>
<th>Ground Truth</th>
<th>Random Flow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Divide /8 into /16s</td>
<td>Setup</td>
</tr>
<tr>
<td>Send two back-to-back scans at 100,000pps</td>
<td>- src port: 80</td>
</tr>
<tr>
<td>Goal: Find ALL visible links at confidence level</td>
<td>- dest port: Random</td>
</tr>
<tr>
<td>of 99%(95% for one scan)</td>
<td>- TTL: 1~32</td>
</tr>
<tr>
<td></td>
<td>Goal: Simulate the downstream traffic flows from DCs to clients</td>
</tr>
<tr>
<td>Ground Truth</td>
<td>Random Flow</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Divide /8 into /16s</td>
<td>Setup</td>
</tr>
<tr>
<td>Send two back-to-back scans at 100,000pps</td>
<td>- src port: 80</td>
</tr>
<tr>
<td>Goal: Find ALL visible links at confidence level of 99% (95% for one scan)</td>
<td>- dest port: Random</td>
</tr>
<tr>
<td></td>
<td>- TTL: 1~32</td>
</tr>
<tr>
<td></td>
<td>Goal: Simulate the downstream traffic flows from DCs to clients</td>
</tr>
</tbody>
</table>
Uncovered links cause poor visibility to the Internet

What is the link coverage of current practices under load balancing?

- Cloud
- /24 Network
  - End User A
  - End User B
  - End User C

Missing Events

Congestion

这个图可以用，再加一些说明文字表述出标题的这个意思
Load Balancing Challenges Coverage Assumption

Cloud

/24 Network
End User A
End User B (representative)

/24 Network
End User C (representative)
End User D