Ting: Measuring and Exploiting Latencies Between All Tor Nodes

Frank Cangialosi   Dave Levin   Neil Spring
University of Maryland
Measuring latencies

Measurement Host

Ping

External Host

Limited to the nodes we control
Measuring latencies

Limited to the nodes we control

To gain broader insight, we can:

1. Control more nodes?
2. Estimate latencies?
Measuring latencies

**Ping**

Measurement Host  
External Host

**Goal**

Limited to the nodes we **control**

Latency between **arbitrary** nodes
King technique

[Gummadi et al, 2002]
King technique

[Gummadi et al, 2002]
King technique

[Gummadi et al, 2002]

Measure the RTT of recursive DNS queries
King technique

[Gummadi et al, 2002]

Measure the RTT of recursive DNS queries
King technique

[Gummadi et al, 2002]

Measure the RTT of recursive DNS queries
King technique

[Gummadi et al, 2002]

Measure the RTT of recursive DNS queries
King technique

[Gummadi et al, 2002]

Measure the RTT of recursive DNS queries

Subtract latencies
King technique
[Gummadi et al, 2002]

Measure the RTT of recursive DNS queries
Subtract latencies
King technique

[Gummadi et al, 2002]

Measure the RTT of recursive DNS queries

Subtract latencies

King does not directly measure the path between two hosts
King technique

[Gummadi et al, 2002]

Measure the RTT of recursive DNS queries

Subtract latencies

King does not directly measure the path between two hosts
King technique

[Gummadi et al, 2002]

Measure the RTT of recursive DNS queries
Subtract latencies

King does not *directly* measure the path between two hosts
King technique

[Parthasarathy et al, 2002]

Measure the RTT of recursive DNS queries

Subtract latencies

King does not directly measure the path between two hosts

Only 3% support recursive queries
TING  a tool for measuring latency between arbitrary Tor nodes

1. **Accurate** — measures the full path between end hosts

2. **Practical** — does not require modification of end hosts
What is Tor?

Anonymity-enabling overlay network

Packets routed through series of relays, called a circuit

Clients choose their own circuits
What is Tor?

Anonymity-enabling overlay network

Packets routed through series of relays, called a circuit

Clients choose their own circuits
Why Tor?

**Large**

6,536 relays in 77 countries (on October 29, 2015)

**Diverse**

5,520 /24s, ~61% residential networks

**Growing**

[https://metrics.torproject.org](https://metrics.torproject.org)
Tor-specific constraints
Tor-specific constraints
Tor-specific constraints

\[ x \rightarrow y \rightarrow x \]
Tor-specific constraints

Ping

x

y
Tor-specific constraints

Subtract latencies?
Tor-specific constraints

Tor traffic may be treated differently

Diagram:
- Node x
- Node y
- Arrows connecting x and y
Tor-specific constraints

Tor traffic may be treated differently

CDF of $\min(\text{ICMP}) - \min(\text{TCP})$ (msec)

- 10% ICMP slower
- 25% TCP slower

Diagram showing network topology with nodes x and y connected by arrows.
Tor-specific constraints

Tor traffic may be treated differently
Tor traffic may be treated differently
Tor-specific constraints

- Tor traffic may be treated differently
- Cannot create one-hop circuits
Tor-specific constraints

Tor traffic may be treated differently

Cannot create one-hop circuits

x

y

Tor
Tor-specific constraints

- Tor traffic may be treated differently
- Cannot create one-hop circuits
Tor-specific constraints

Tor traffic may be treated differently

Cannot create one-hop circuits
Tor-specific constraints

Tor traffic may be treated differently

Cannot create one-hop circuits

Must account for forwarding delays
Tor-specific constraints

Tor traffic may be treated differently

Cannot create one-hop circuits

Must account for forwarding delays

Queuing / Scheduling, Encryption & Decryption, Context Switches
Tor-specific constraints

- Tor traffic may be treated differently
- Cannot create one-hop circuits
- Must account for forwarding delays
- Queuing / Scheduling, Encryption & Decryption, Context Switches
Tor-specific constraints

- Tor traffic may be treated differently
- Cannot create one-hop circuits
- Must account for forwarding delays
Ting technique

\[ x \quad y \]
Measurement Host

Measure full path between x and y

Summary
Measure full path between x and y

Summary
Measure full path between x and y

Summary
Measure full path between x and y

Summary
Measurement Host

Measure **full** path between x and y

Summary
Measurement Host

Isolate **RTT** between client and x

Summary
Measurement Host

Isolate RTT between client and x

Summary
Measurement Host

Isolate **RTT** between client and x

**Summary**

Isolate RTT between client and x
Isolate **RTT** between client and x

Summary
Measurement Host

Isolate RTT between client and x

Summary
Isolate **RTT** between client and `x`
Measurement Host

Isolate **RTT** between client and x

Subtract latencies

Summary

1 - 2

F F
Isolate **RTT** between client and $x$
Isolate **RTT** between client and \( y \)
Measurement Host

Isolate **RTT** between client and **y**

Summary
Isolate RTT between client and y

Measurement Host

Summary
Isolate RTT between client and y

Summary
Isolate **RTT** between client and \( y \)
Isolate **RTT** between client and $y$
Isolate **RTT** between client and \( y \)

**Summary**

1. Subtract latencies
2. (Green operations)
3. (Red operations)
Isolate RTT between client and $y$
1 - 2 - 3 = RTT(x,y) + F_x + F_y
Minimum of multiple, independent samples of each circuit
Ting evaluation

Implemented Ting using the Stem Tor controller
No modifications to the Tor client

How well does Ting work?

- How accurate?
- How many samples?
- How consistent?
How well does Ting work?

- How accurate?
- How many samples?
- How consistent?

31 Tor relays

930 Total pairs
How accurate is Ting?

Fraction of PlanetLab Pairs

Measured / Real

All Pairs

28
How accurate is Ting?

Fraction of PlanetLab Pairs

Measured / Real

Ideal
How accurate is Ting?

Fraction of PlanetLab Pairs

Measured / Real

All Pairs

Ideal
How accurate is Ting?

Ting estimates typically within 10% of “real”
How accurate is Ting?

Ting estimates typically within 10% of “real”
How accurate is Ting?

Ting estimates typically within 10% of “real”
How accurate is Ting?

Ideal

Relative error lower at higher latencies

Forwarding delays create an additive error

RTT(x,y) + Fx + Fy

Ting estimates typically within 10% of “real”
How many samples does Ting need?
How many samples does Ting need?

Cumulative Fraction of Pairs vs. Cumulative Samples Per Circuit

- Measured Min
- Within 1ms
- Within 1%
- Within 5%
- Within 10%

Fraction of Pairs

Cumulative Samples Per Circuit

PLANETLAB
How many samples does Ting need?

Only ~15 samples needed to reach values within 5% of min

Ting remains accurate with few samples
How consistent are Ting measurements?

**30 pairs** of real Tor relays, measured *once an hour* over a week.
How consistent are Ting measurements?

30 pairs of real Tor relays, measured once an hour over a week

50% vary by < 5ms over a week

Ting measurements need not be updated often
Evaluation summary

How well does Ting work?

- Typically within 10% of real latency
- Remains accurate with few samples
- Vary by only a few ms over time
Evaluation summary

How well does Ting work?

- Typically within 10% of real latency
- Remains accurate with few samples
- Vary by only a few ms over time

Ting’s **stability and accuracy** permit collection of an **all-pairs RTT** dataset
All-pairs RTT dataset

50 Tor relays outside of our control

1,225 pairs in all-pairs RTT dataset

Geographic distribution

Latency distribution

CDF

Latency (msec)
Applications

1. Speeding up deanonymization of Tor circuits
2. Improving Tor’s path selection algorithm
3. Gain insight into non-Tor nodes
Applications

1. Speeding up deanonymization of Tor circuits
Classic traffic-analysis attack
[Murdoch and Danezis, 2005]
Classic traffic-analysis attack

[Murdoch and Danezis, 2005]
Classic traffic-analysis attack

[Murdoch and Danezis, 2005]
Classic traffic-analysis attack

[Murdoch and Danezis, 2005]

**Attacker’s Goal:** find all nodes in the **circuit**
Classic traffic-analysis attack

[Murdoch and Danezis, 2005]

**Attacker’s Goal:** find all nodes in the **circuit**
Classic traffic-analysis attack
[Murdoch and Danezis, 2005]

**Attacker’s Goal**: find all nodes in the **circuit**

![Diagram showing a client, Tor network, and a malicious destination with a probe.]
Classic traffic-analysis attack

[Murdoch and Danezis, 2005]

**Attacker’s Goal**: find all nodes in the **circuit**

*Time- and bandwidth-intensive*
Classic traffic-analysis attack

[Murdoch and Danezis, 2005]

**Attacker’s Goal**: find all nodes in the circuit

*Time- and bandwidth-intensive*

*Done in an unguided fashion*
Classic traffic-analysis attack

[Murdoch and Danezis, 2005]

Attacker’s Goal: find all nodes in the circuit

Client → Tor → Malicious Destination

Time- and bandwidth-intensive
Done in an unguided fashion
Classic traffic-analysis attack

CDF

Fraction of nodes tested

RTT-unaware

Ignore too-large RTTs + informed target selection
Classic traffic-analysis attack

CDF

Fraction of nodes tested

RTT-unaware
Faster deanonymization with Ting

Apply what the attacker knows about latencies
Faster deanonymization with Ting

Apply what the attacker knows about latencies

e2e RTT
Faster deanonymization with Ting

Apply what the attacker knows about latencies

e2e RTT

Directly measured
Faster deanonymization with Ting

Apply what the attacker knows about latencies

e2e RTT

Pre-measured with Ting

Directly measured
Faster deanonymization with Ting

Apply what the attacker knows about latencies

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured
Faster deanonymization with Ting

Apply what the attacker knows about latencies

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured
Faster deanonymization with Ting

Reason about what the client → entry RTT would have to be

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured

40
Client’s RTT to the entry node is unknown. Pre-measured with Ting. Directly measured.
Faster deanonymization with Ting

Reason about what the client → entry RTT would have to be

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured
Ruling out nodes without probing them

Reason about what the client → entry RTT would have to be

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured
Ruling out nodes without probing them

Reason about what the client $\rightarrow$ entry RTT would have to be

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured
Ruling out nodes without probing them

Reason about what the client → entry RTT would have to be

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured
Ruling out nodes without probing them

Reason about what the client → entry RTT would have to be

Client’s RTT to the entry node is unknown

Client’s RTT would have to be negative

Directly measured

Pre-measured with Ting
Ruling out nodes without probing them

Reason about what the client → entry RTT would have to be

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured
Ruling out too-large RTTs

CDF

RTT-unaware
Ignore too-large RTTs

Fraction of nodes tested
Ruling out too-large RTTs

CDF

Fraction of nodes tested

RTT-unaware
Ignore too-large RTTs
Informed target selection

Probe nodes according to probability that they are on the circuit

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured

Too-Large RTT

Client’s RTT would have to be negative
Informed target selection

Probe the more likely circuits first

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured
Informed target selection

Probe the more likely circuits first

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured
Informed target selection

Probe the more likely circuits first

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured
Informed target selection

Probe the more likely circuits first

Average Ting RTT

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured
Informed target selection
Probe the more likely circuits first

Client’s RTT to the entry node is unknown
Pre-measured with Ting
Directly measured
Informed target selection

Probe the more likely circuits first

Score

Lower score implies higher likelihood

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured

46
Informed target selection

Probe the more likely circuits first

Score
Lower score implies higher likelihood

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured
Informed target selection

Probe the more likely circuits first

Score

Lower score implies higher likelihood

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured
Informed target selection

Probe the more likely circuits first

Score

Lower score implies higher likelihood

Client’s RTT to the entry node is unknown

Pre-measured with Ting

Directly measured
Informed target selection decreases search time by a median of 1.5x
Informed target selection decreases search time by a median of **1.5x**
Faster deanonymization with Ting

Informed target selection \textbf{decreases} search time by a median of \textbf{1.5x}
Summary

We lack a **practical** tool for measuring the latency between two arbitrary hosts.

**TING** measures the latency between **Tor** nodes is fast, accurate, and practical.

Source code and data available at:

Implementation

Ting Client
- Language: Python
- Tor Controller: Stem
- Tor-0.2.3.25-patched
- SLOC: ~400

Test Relays
- Tor-0.2.4.22 (latest)
- PublishDescriptors 0
- Restricted Exit Policy
- Uptime: > 1 month