U.S. NAVAL RESEARCH LABORATORY

On the Accuracy of Tor Bandwidth Estimation

Rob Jansen and Aaron Johnson
U.S. Naval Research Laboratory

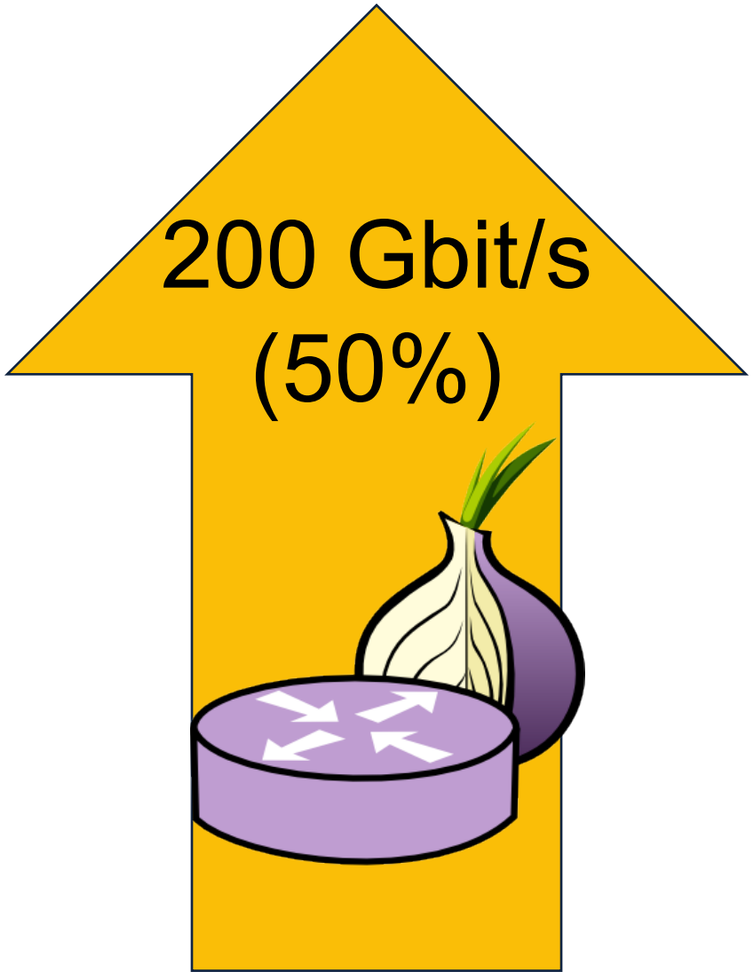
Rob Jansen
Center for High Assurance Computer Systems
U.S. Naval Research Laboratory

Passive and Active Measurement Conference 2021
Virtual Event
March 29th – 31st , 2021

U.S. NAVAL RESEARCH LABORATORY

Main Results

- **Tor underestimates its total network bandwidth capacity by about 200 Gbit/s (50%)**
  - ~20% of relays have > 50% variation in bandwidth estimates
- **We discovered significant error in bandwidth capacity estimation, with larger error for:**
  - High-capacity relays
  - Exit relays
  - Relays with lower uptimes
- **Inaccurate capacity measurements may lead to suboptimal load balancing**
  - Affects performance for all Tor users



U.S. NAVAL RESEARCH LABORATORY

Main Contributions

- **Analysis of passive relay measurements**
  - Understand variation in capacity estimates from historical data
  - Variation indicates inaccurate estimation

- **Active speed test experiment to measure relays**
  - Flood relays with traffic to drive up their observed bandwidth
  - Cause relays to learn their bandwidth limits and better estimate their capacity
  - Analyze change in bandwidth reports before/after speed test

U.S. NAVAL RESEARCH LABORATORY

Main Contributions

- **Analysis of passive relay measurements**
  - Understand variation in capacity estimates from historical data
  - Variation indicates inaccurate estimation

- **Active speed test experiment to measure relays**
  - Flood relays with traffic to drive up their observed bandwidth
  - Cause relays to learn their bandwidth limits and better estimate their capacity
  - Analyze change in bandwidth reports before/after speed test

U.S. NAVAL RESEARCH LABORATORY

Definitions and Problem

- **Fowarding capacity (i.e., the *true* capacity)**
  - The max sustainable rate at which a relay can forward traffic
  - This value is unknown, so relays must estimate it

U.S. NAVAL RESEARCH LABORATORY

Definitions and Problem

- **Fowarding capacity (i.e., the *true* capacity)**
  - The max sustainable rate at which a relay can forward traffic
  - This value is unknown, so relays must estimate it

- **Observed bandwidth**
  - The max throughput a relay has sustained for any 10 second period over last 5 days
  - This value is reported to Tor metrics every 18 hours
  - Load balancing weights are derived from observed bandwidth

U.S. NAVAL RESEARCH LABORATORY

Definitions and Problem

- **Fowarding capacity (i.e., the *true* capacity)**
  - The max sustainable rate at which a relay can forward traffic
  - This value is unknown, so relays must estimate it
- **Observed bandwidth**
  - The max throughput a relay has sustained for any 10 second period over last 5 days
  - This value is reported to Tor metrics every 18 hours
  - Load balancing weights are derived from observed bandwidth
- **Problem: observed bandwidth != forwarding capacity**
  - Insufficient client traffic limits the observed bandwidth
  - Underutilized relay will never learn its true forwarding capacity
  - Weighting based on observed bandwidth will be inaccurate

U.S. NAVAL RESEARCH LABORATORY

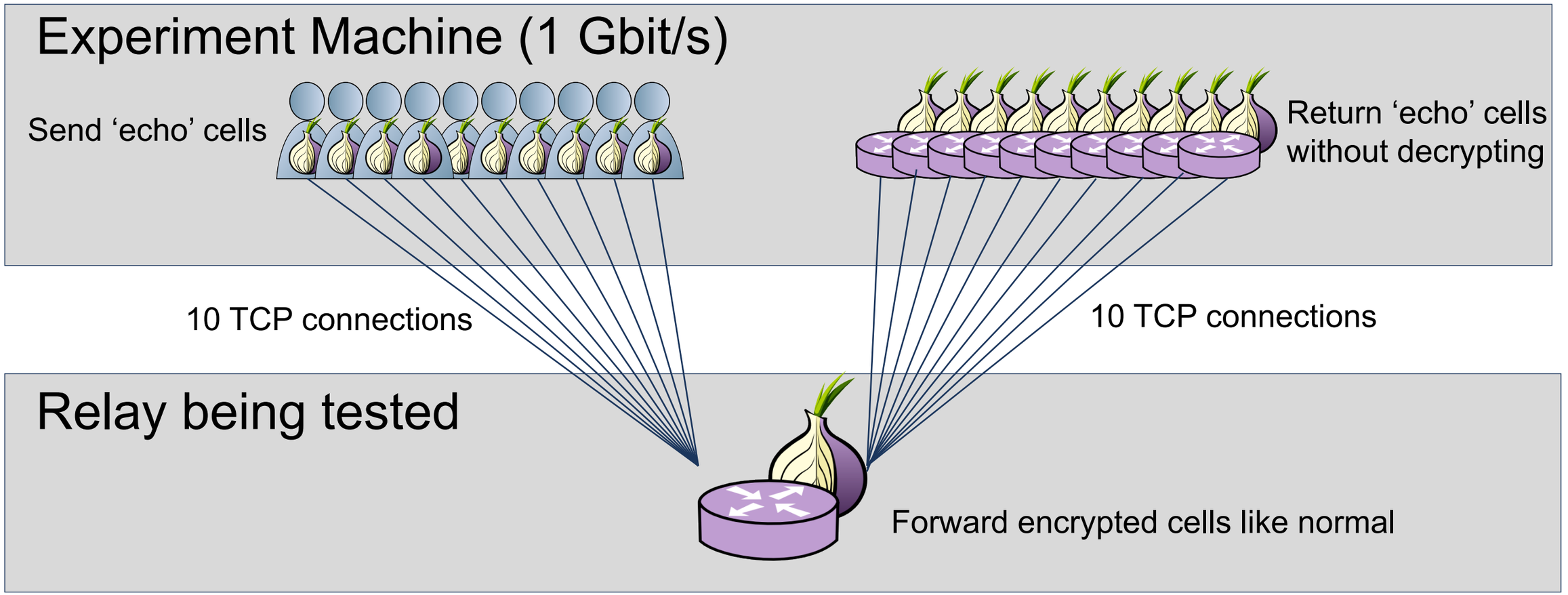
Active Speed Test Experiment

- **Hypothesis: the predominant error is to *underestimate* the true capacity of Tor relays**

- **Experiment: perform a *speed test* on the live Tor network**
  - Actively attempt to send 1 Gbit/s of traffic through each relay
  - Extra traffic should increase relays' observed bandwidth
  - New observed bandwidths should better reflect forwarding capacity of relay

U.S. NAVAL RESEARCH LABORATORY

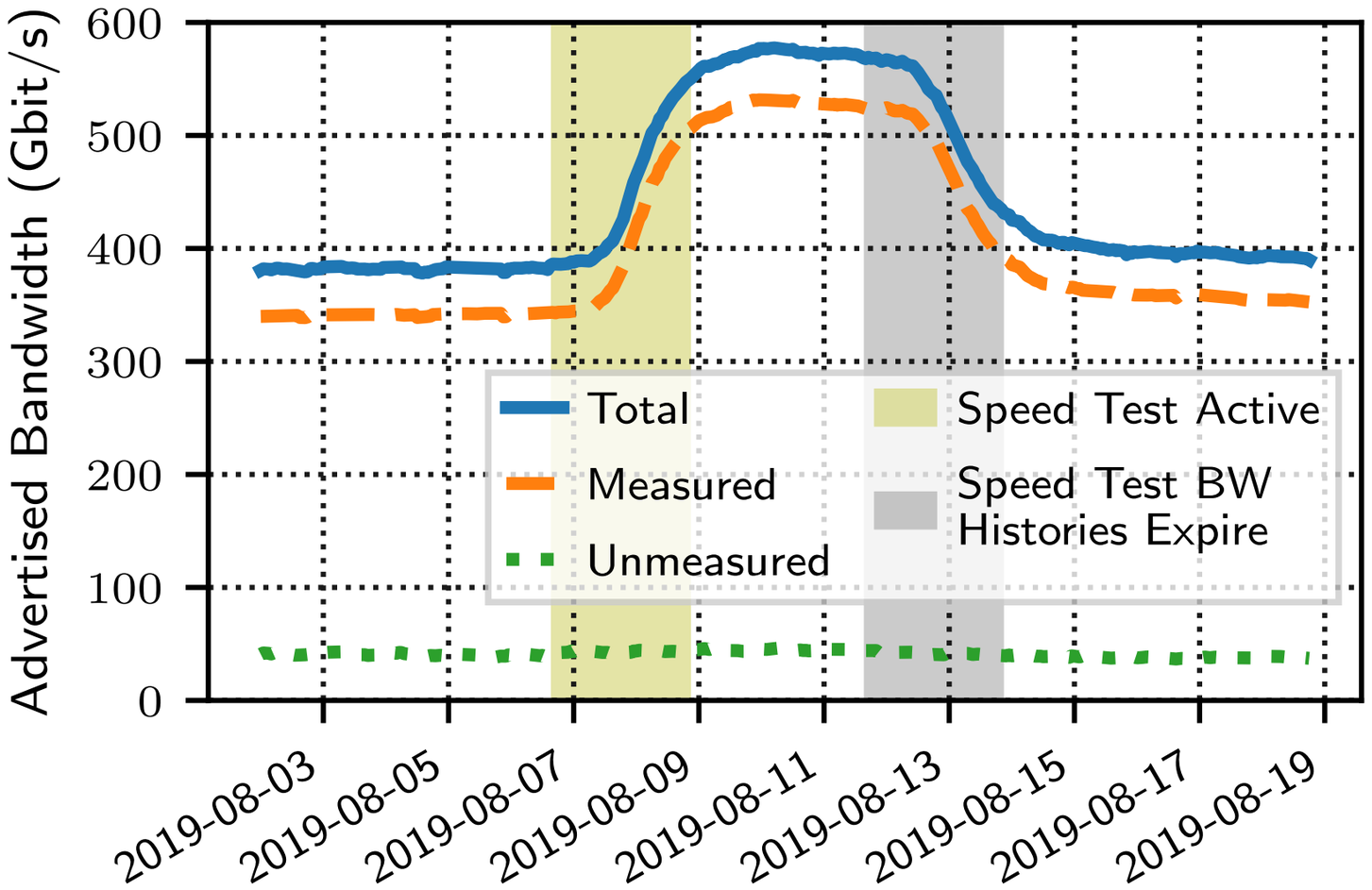
Speed Test Experiment



U.S. NAVAL RESEARCH LABORATORY

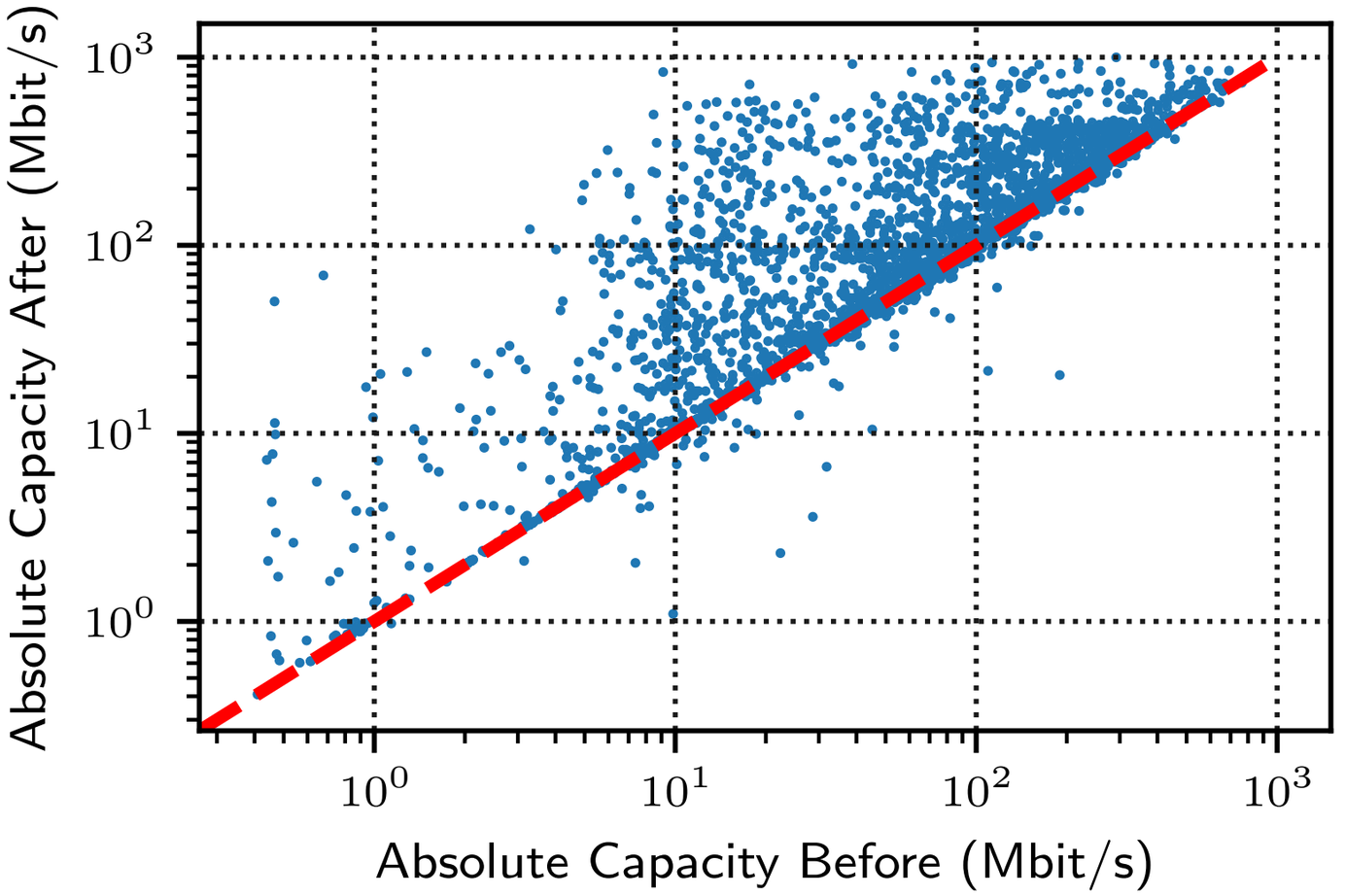
Speed Test Results

Tor underestimates its total capacity by about 50%



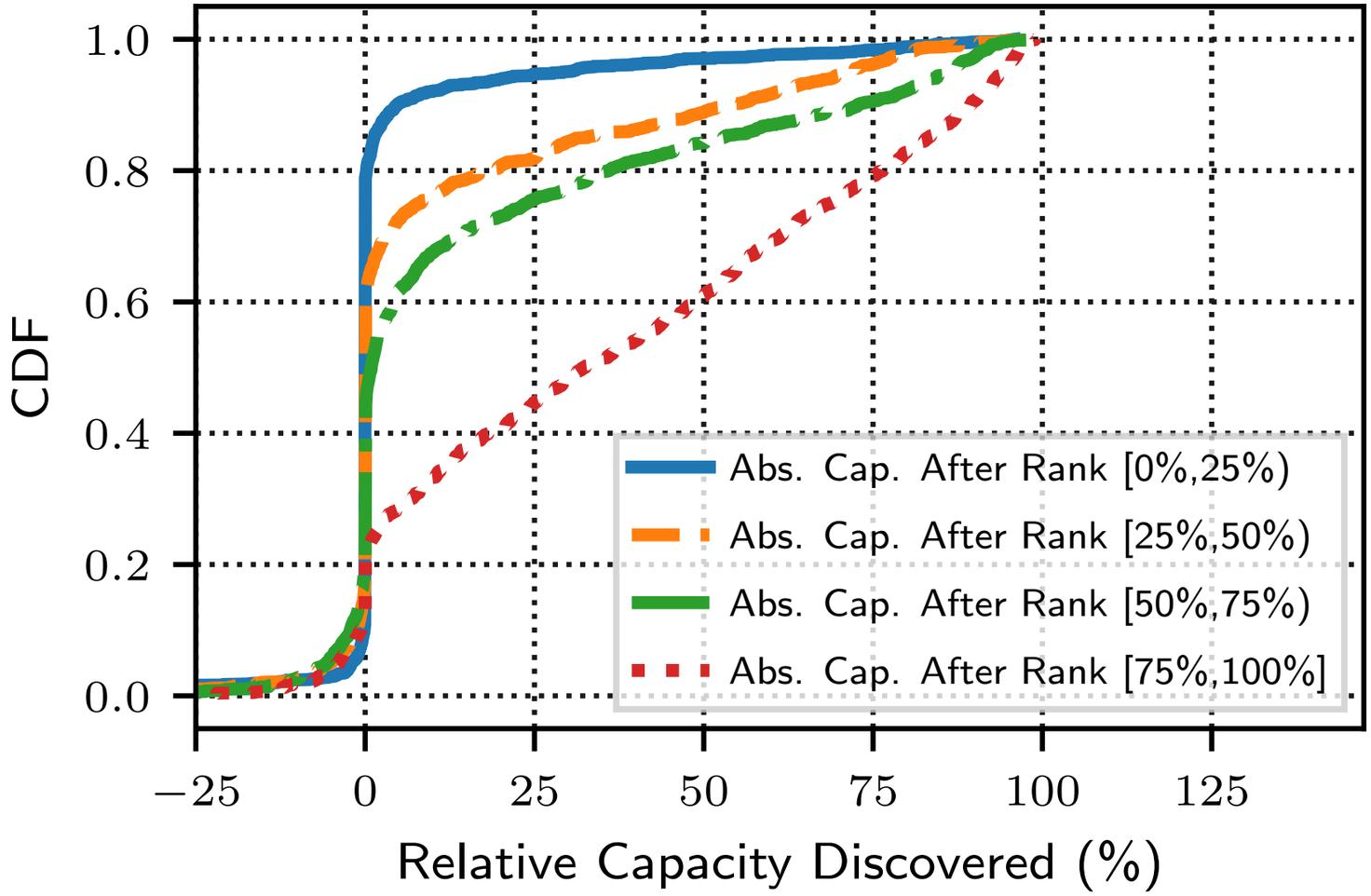
Speed Test Results

The estimated capacity increased after our experiment for most relays (some by a 10x or greater factor)



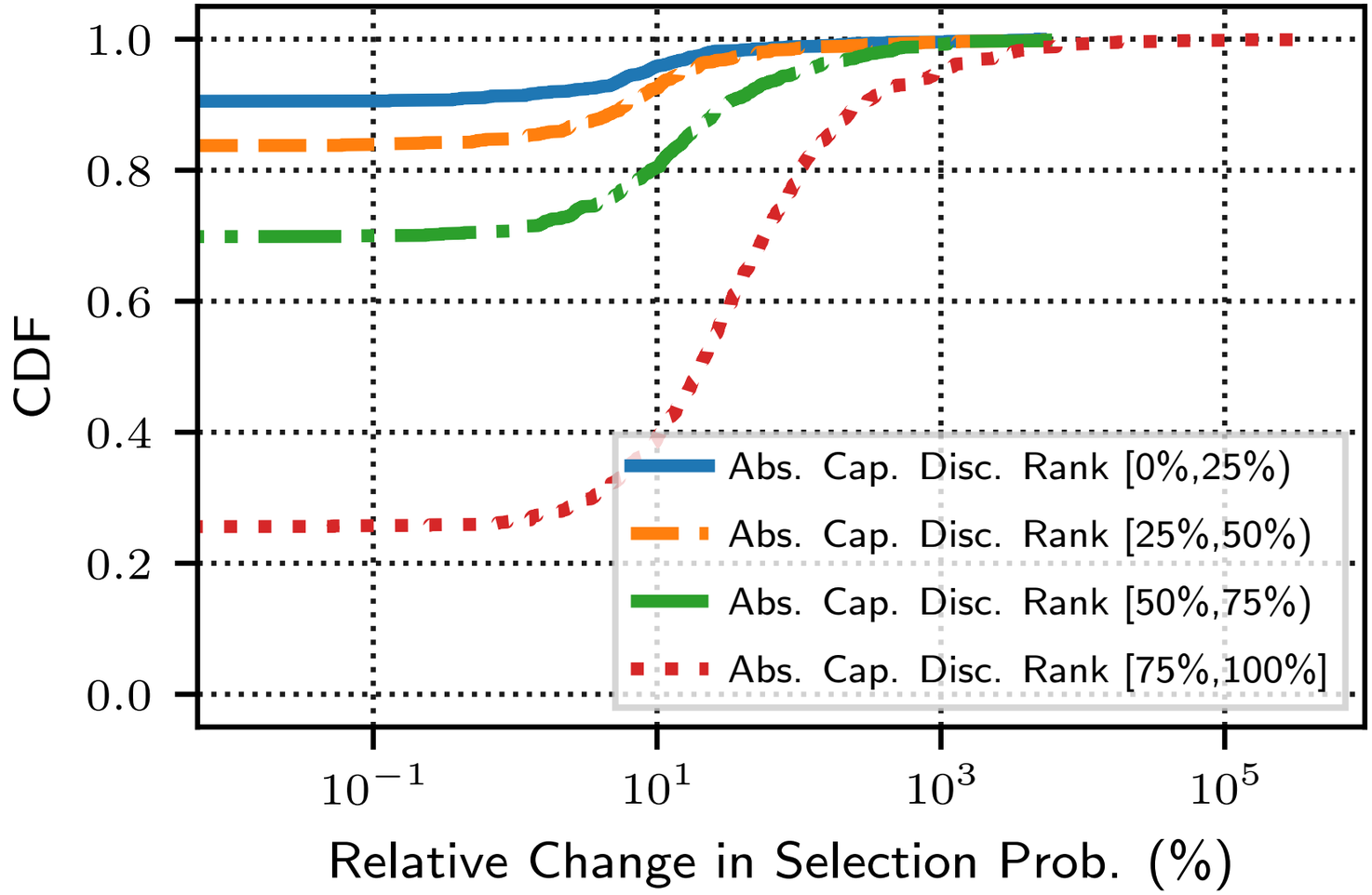
Speed Test Results

We discovered more capacity on higher-capacity relays



Speed Test Results

Load balancing weights changed significantly for many relays (some by a 10x or greater factor)



U.S. NAVAL RESEARCH LABORATORY

Summary

- **Contributions**
  - Historical measurements: ~20% of relays have >50% variation in bandwidth estimates
  - Active speed test experiment:
    - Tor underestimates total capacity by ~50%
    - Larger error associated with high-capacity, exit, and lower uptime relays

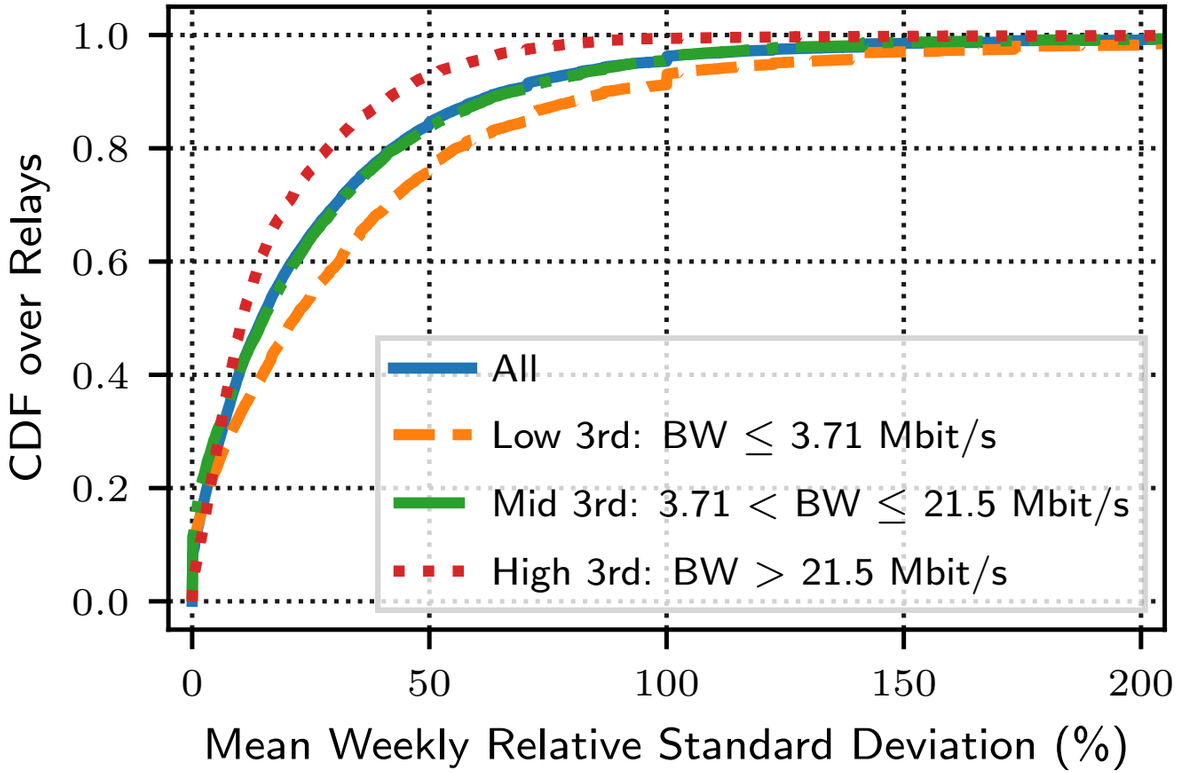
- **Research artifacts available at:**
  - https://torbwest-pam2021.github.io

- **Contact**
  - <rob.g.jansen@nrl.navy.mil>, robgjansen.com, @robgjansen

U.S. NAVAL RESEARCH LABORATORY

Analysis of Historical Bandwidth Data

- **Use relative standard deviation to understand variation**
  - A(r,w) = adv. bws for relay r during week w
  - RSD(A(r,w)) = stdev(A(r,w)) / mean(A(r,w))
- **We find significant variation in relays' bandwidths**
  - the capacity estimates of 25% of relays vary by 41% or more
  - the capacity estimates of 10% of relays vary by 71% or more
  - some relays' capacity estimates vary by more than 200%



Speed Test Results

We discovered more capacity on lower-uptime relays