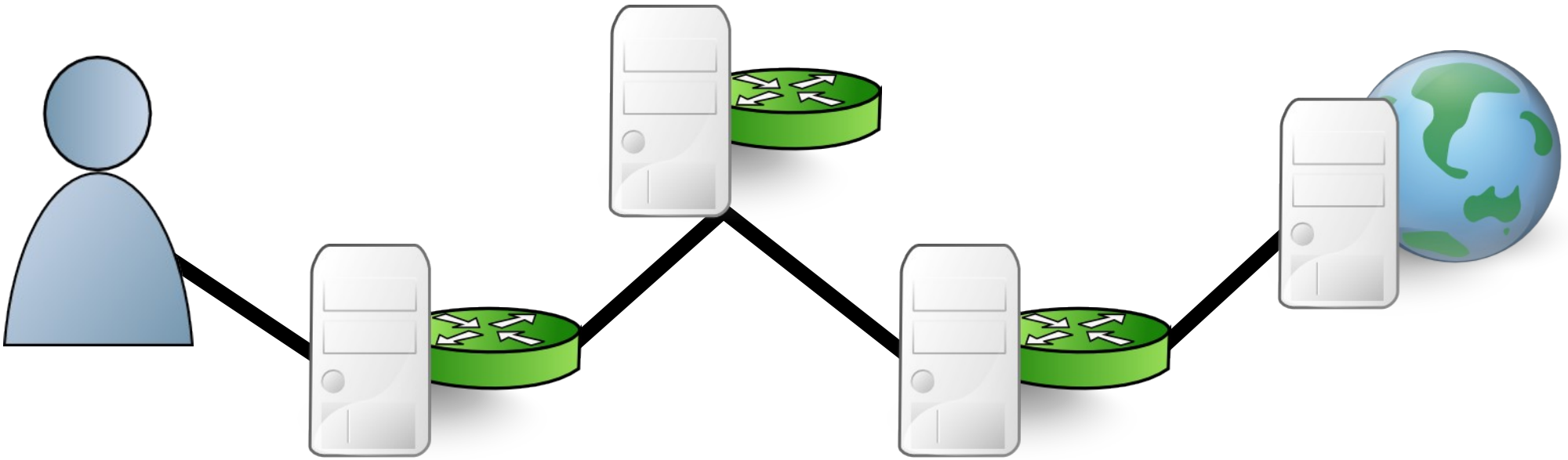
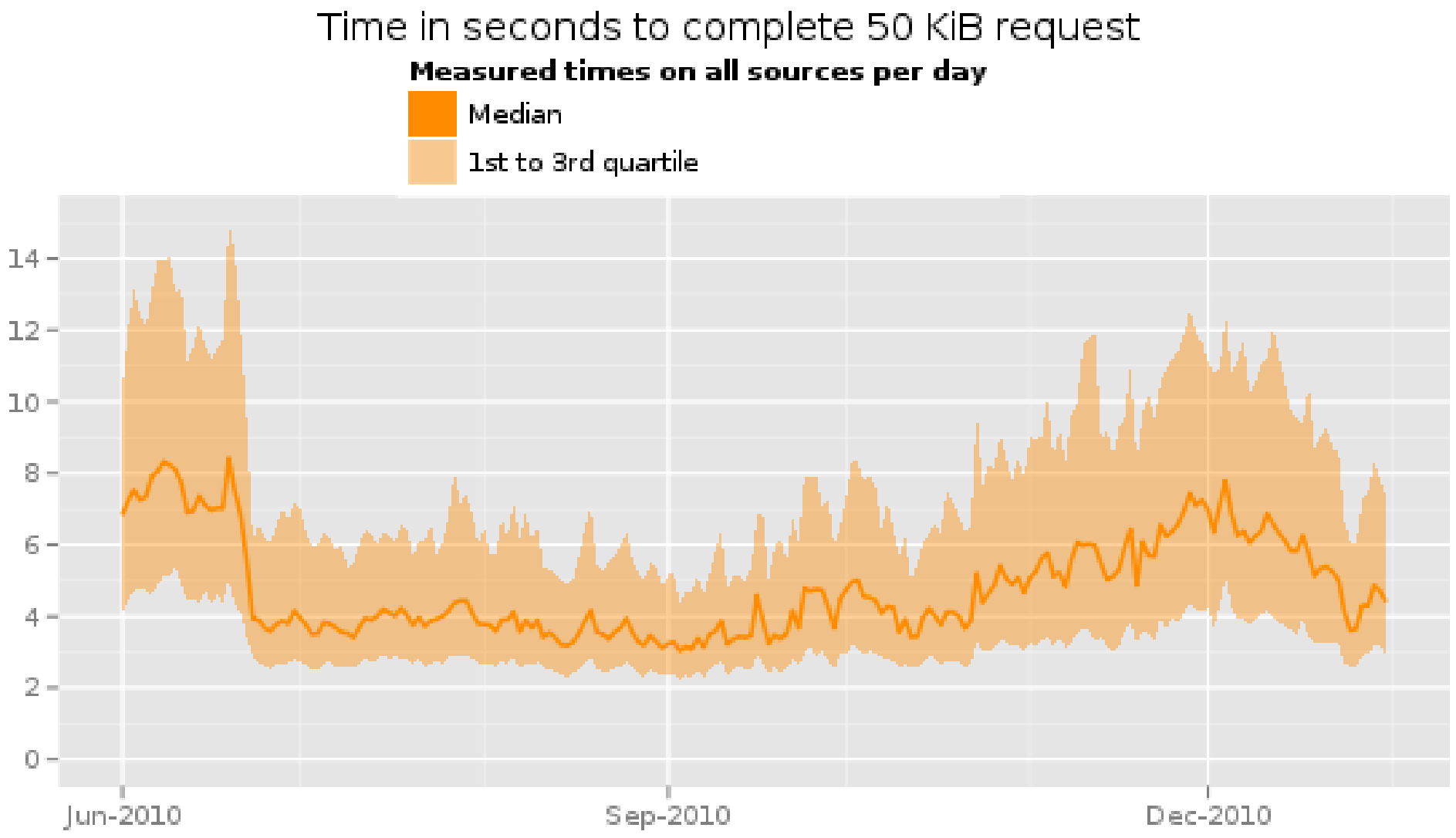
Shadow: Running Tor in a Box for Accurate and Efficient Experimentation

Rob Jansen and Nick Hopper
University of Minnesota
U.S. Naval Research Laboratory
rob.g.jansen@nrl.navy.mil

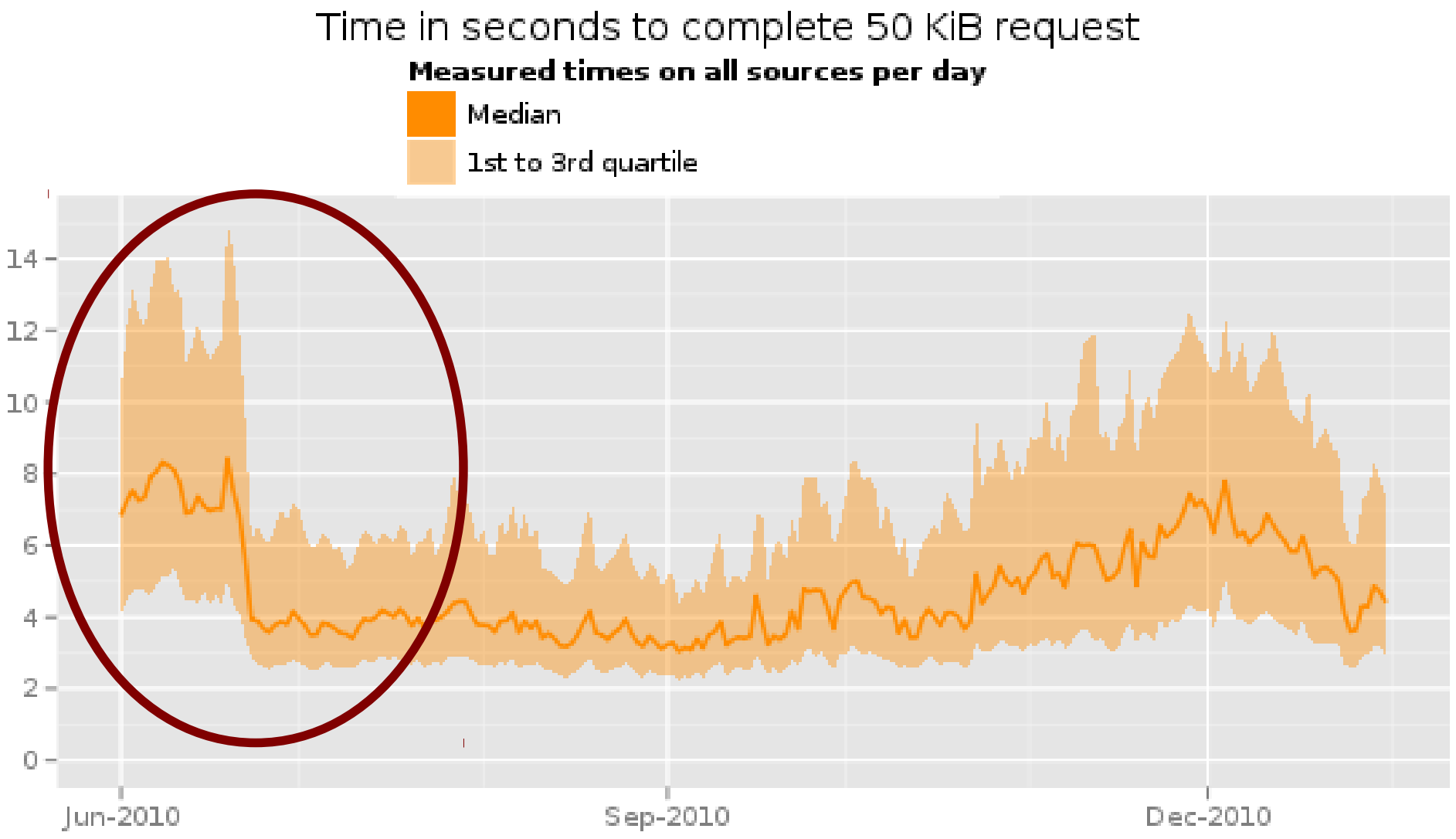
Anonymity with Tor



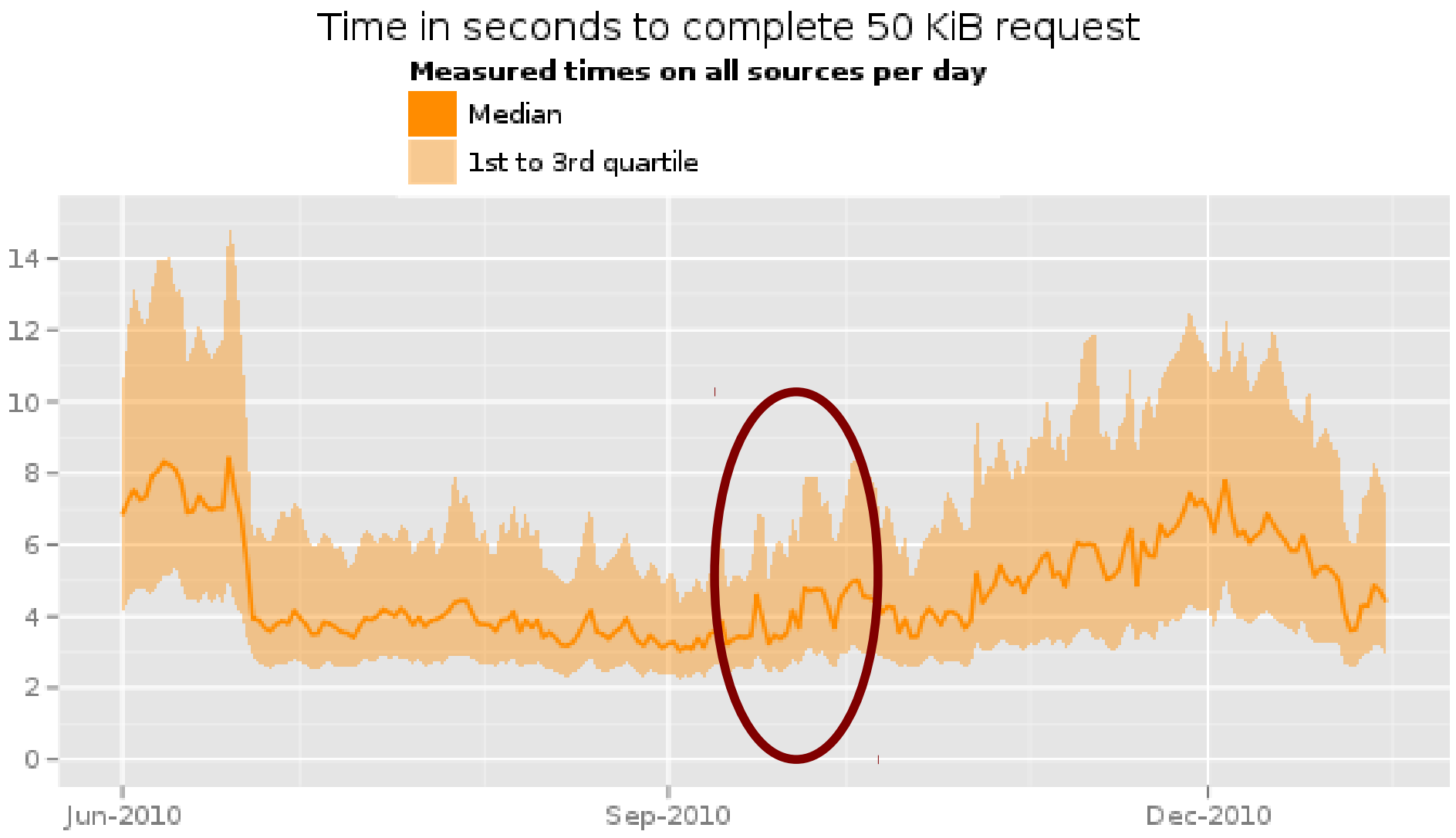
Client Relays Server



The Tor Project - https://metrics.torproject.org/



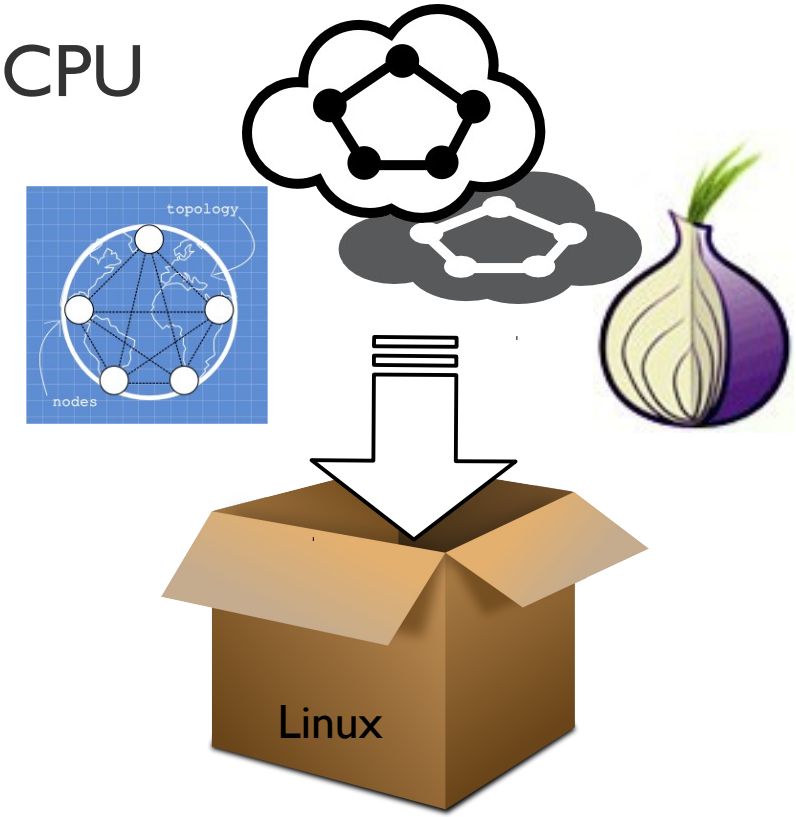
The Tor Project - https://metrics.torproject.org/



The Tor Project - https://metrics.torproject.org/

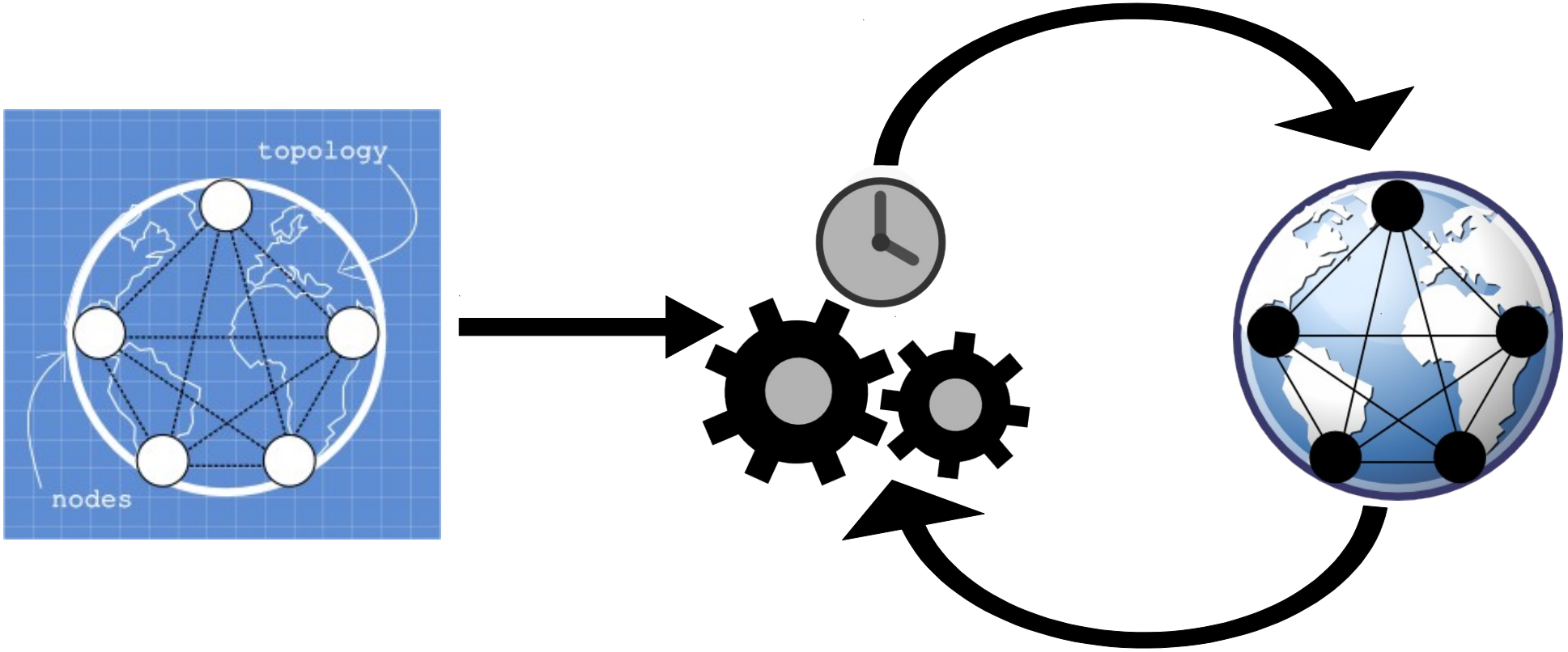
Tor in a Box with Shadow

- Discrete event network simulator
  - Natively executes real applications
  - Simulates time, network, crypto, CPU
  - Model latency and bandwidth
- Efficient, accurate, controlled
- Single Linux-box without root



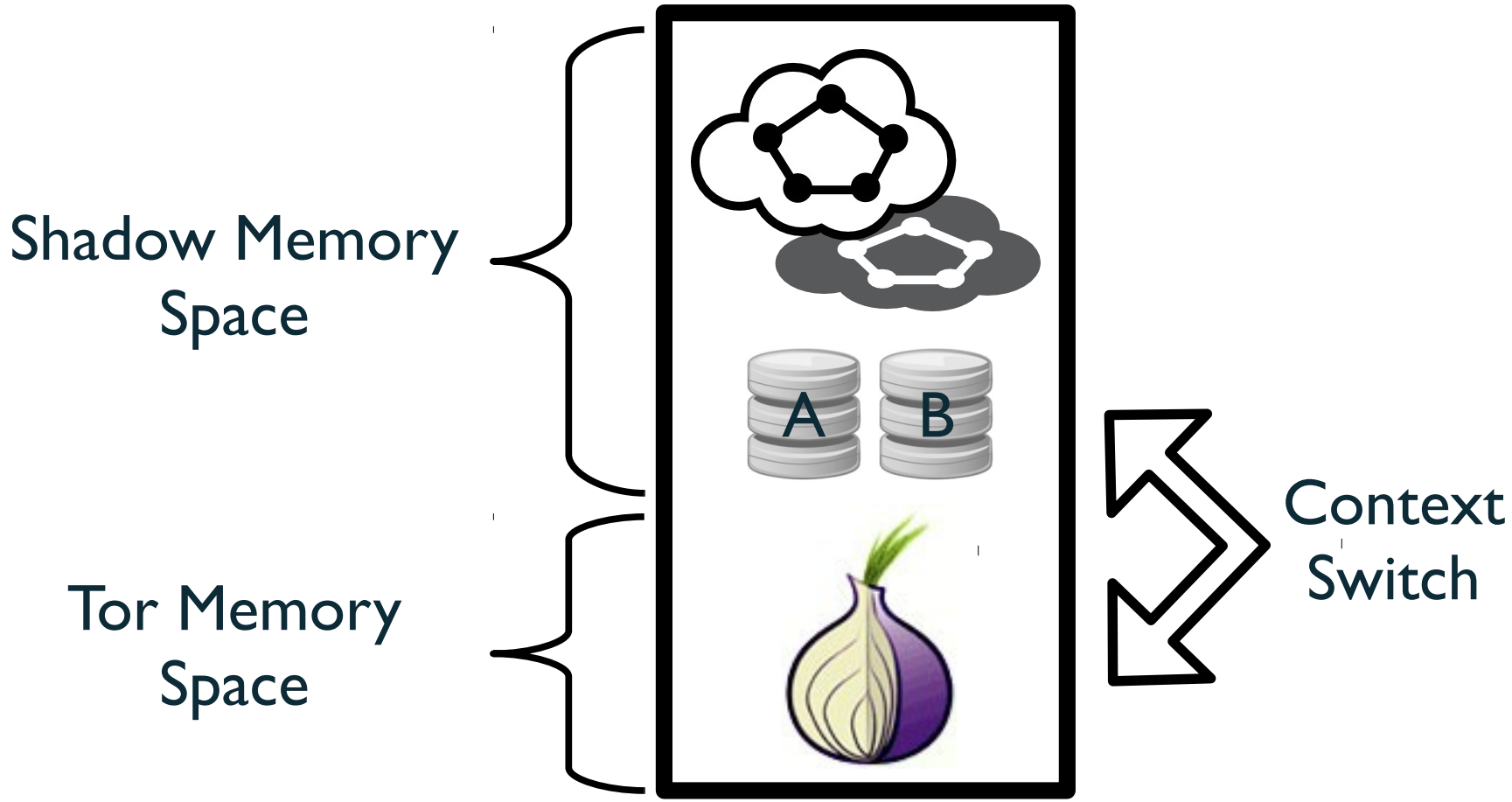
Shadow's Design I

- Simulation blueprint
- Discrete time events



Shadow's Design II

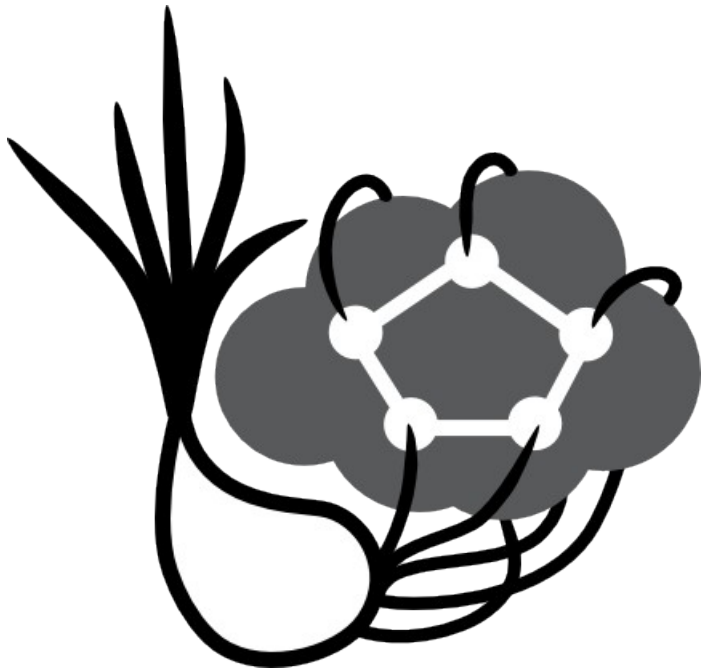
- Node management
- Function interposition



Scallion – A Plug-in Running Tor

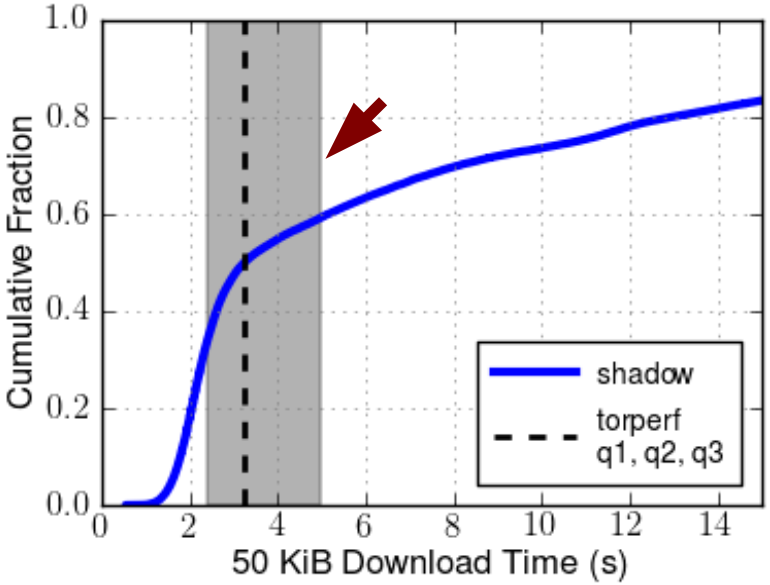
- Integrates Tor into Shadow

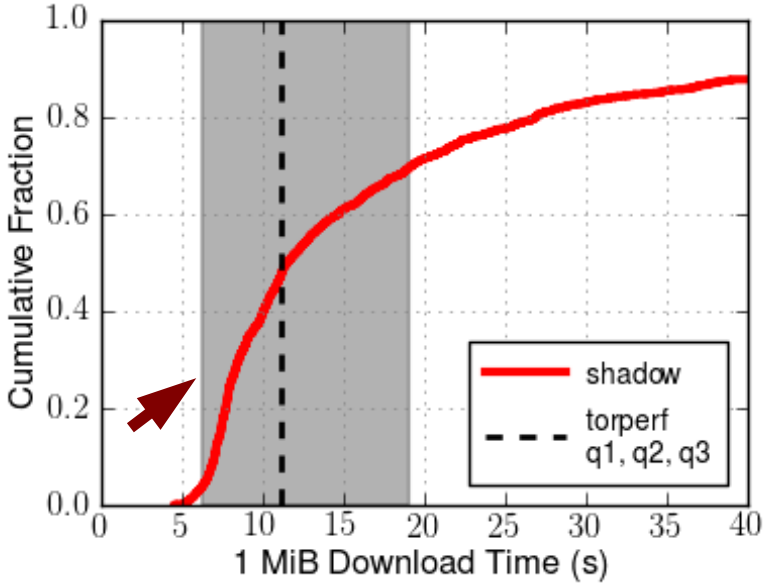
- Scalability
  - 1250 nodes in 10 GB RAM, 5x* - 10x** slowdown
  - 5750 nodes in 60 GB RAM, 40x** slowdown

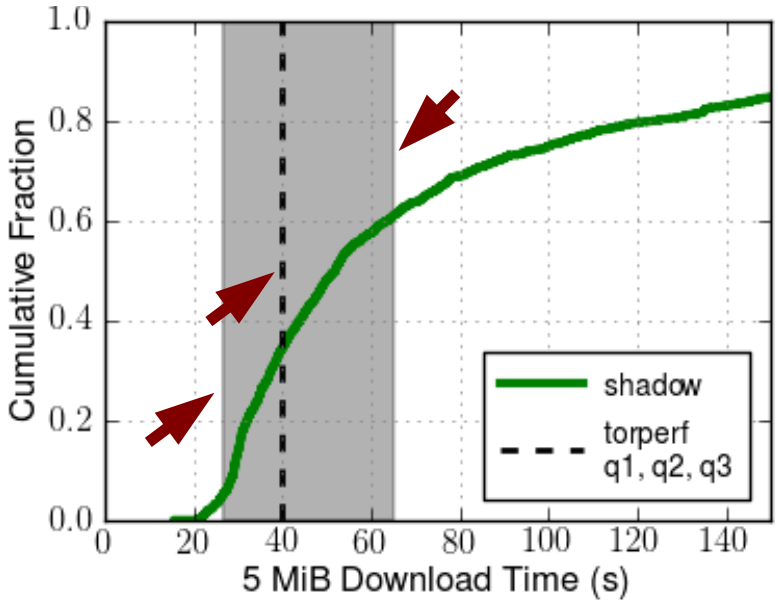


* 3.3 GHz AMD Phenom II X6 1100T ** 2.2 GHz AMD Opteron 6174

Accuracy Shadowing Tor

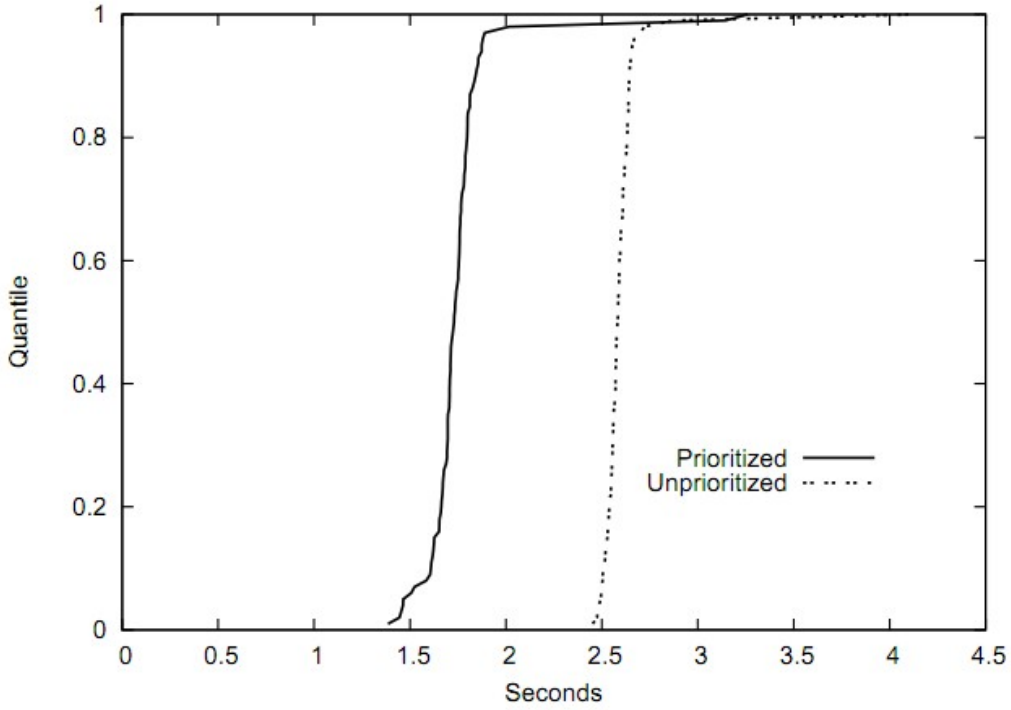




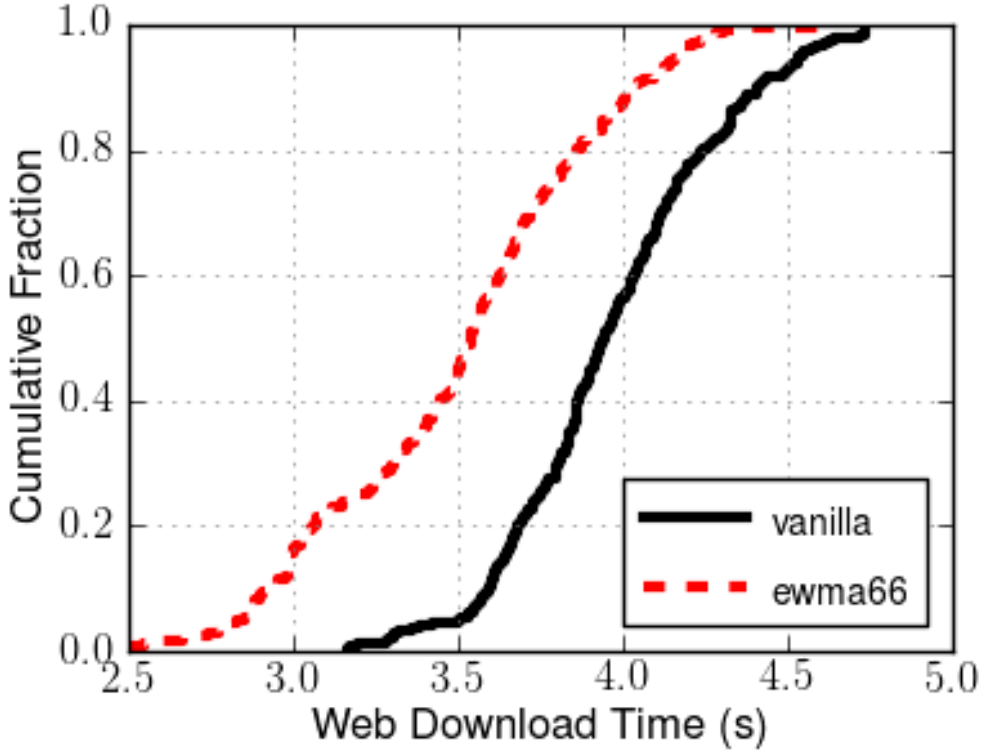


Demonstrating Shadow's Utility

Tang & Goldberg [CCS 10]



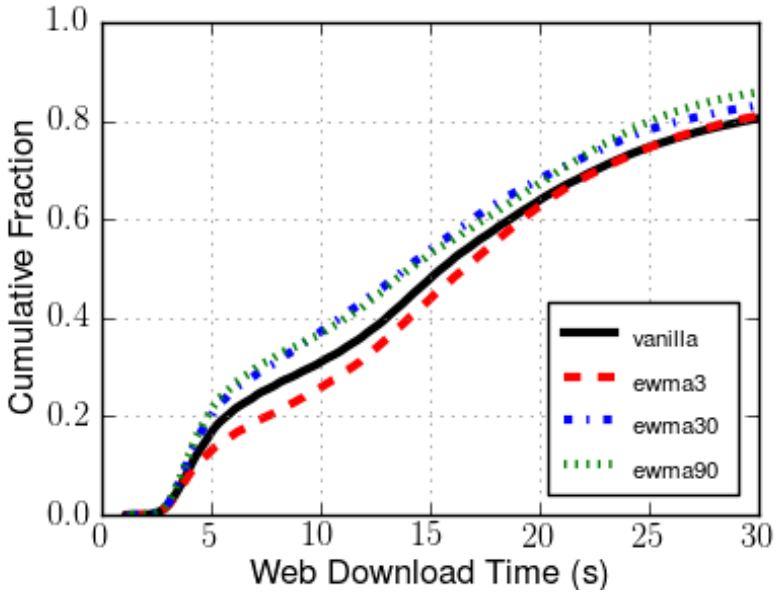
Shadow

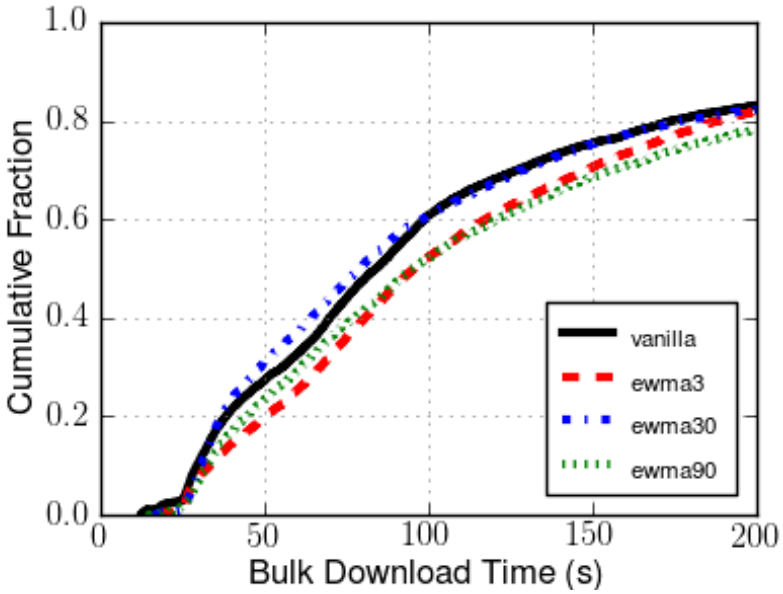


Web

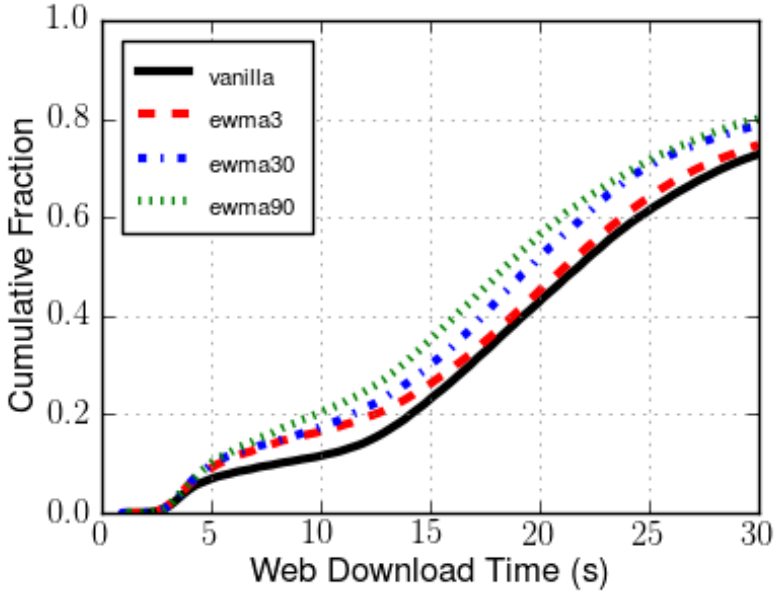
Bulk

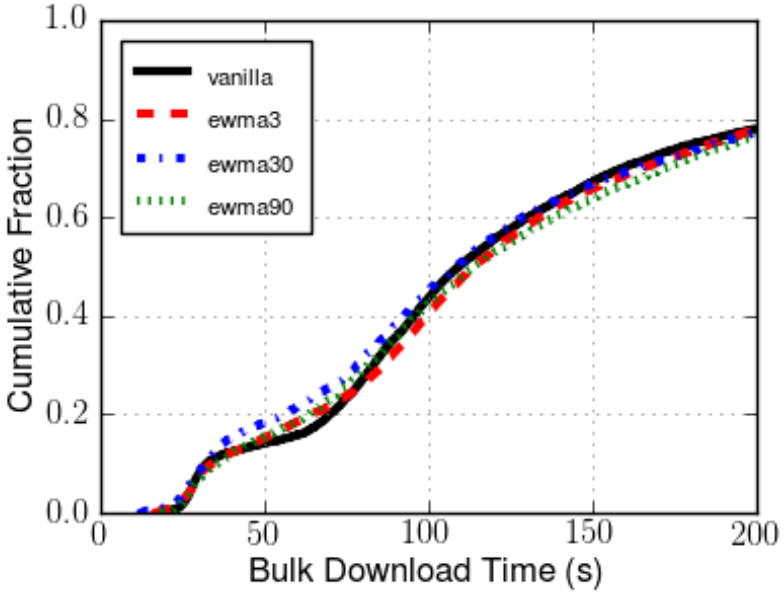
Lightly Loaded Tor





Heavily Loaded Tor





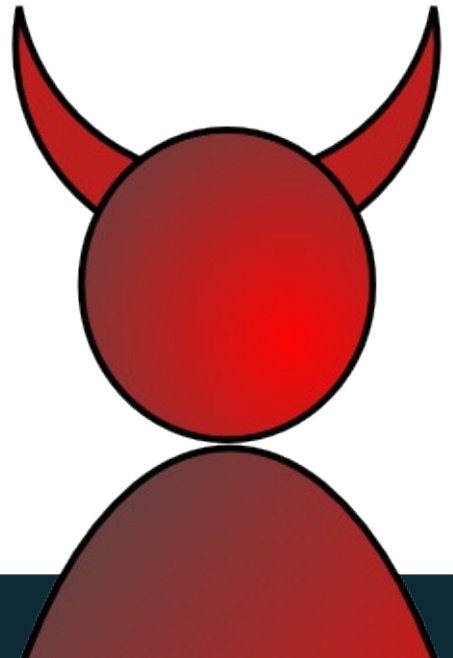
Conclusion

- Efficient, accurate, controllable, repeatable
- Tor experiments on one machine
  - Larger scale than previously possible
  - New results from new capabilities
- Able to run many applications
- Freely available and usable software

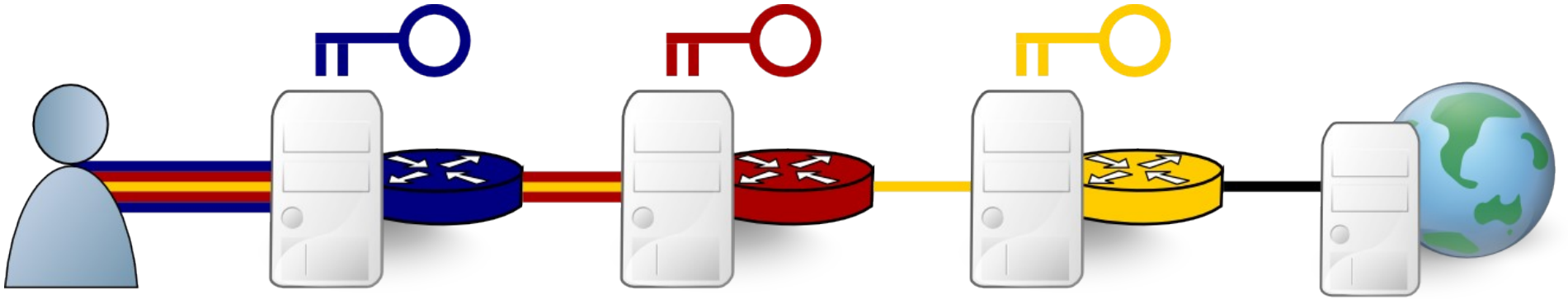
Questions?

rob.g.jansen@nrl.navy.mil
cs.umn.edu/~jansen

shadow.cs.umn.edu
github.com/shadow



How Tor Works



Client Relays Server

Testing Tor Improvements

- Most popular anonymous communication system
  - 500K – 1M users
- New algorithms/protocols need testing
- No standard experimentation approach

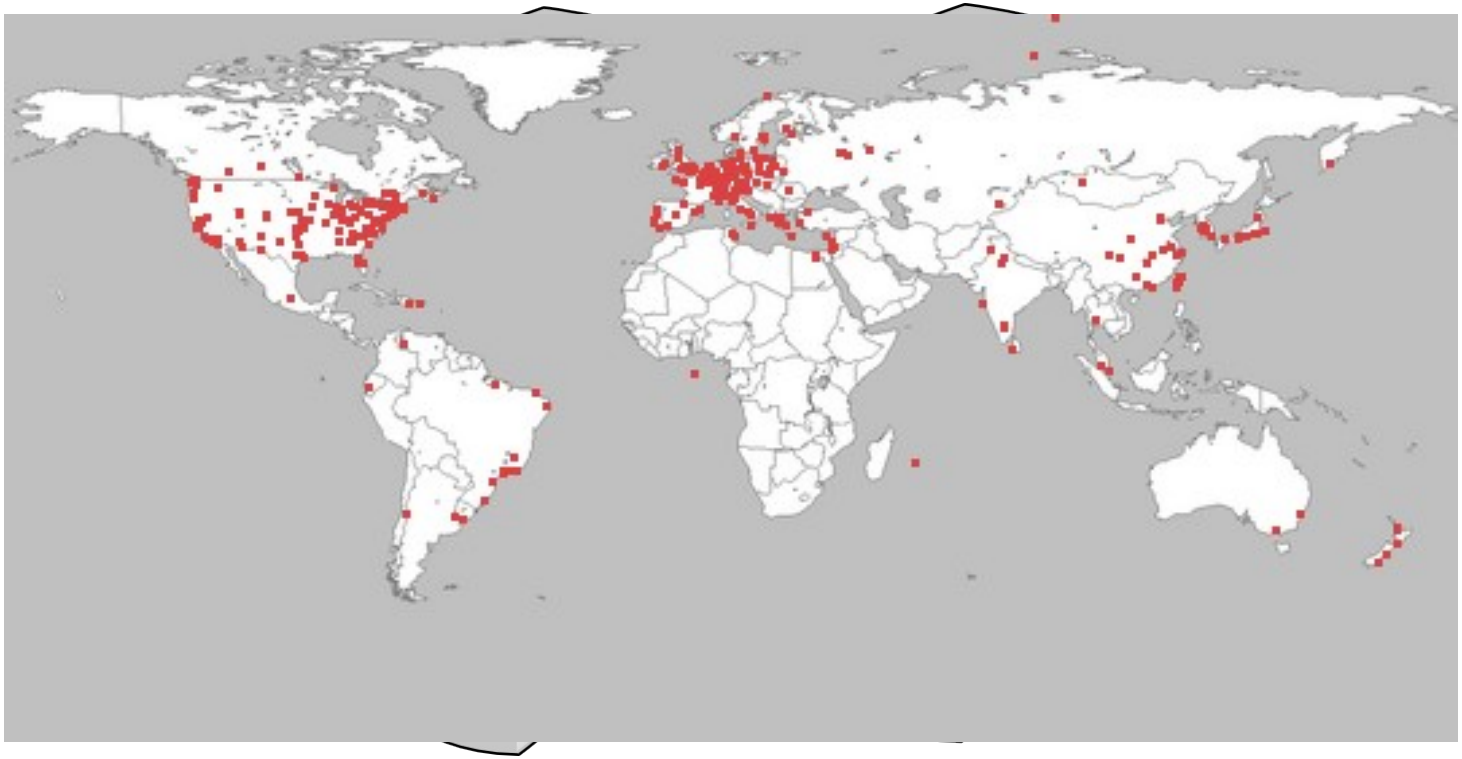
Recent Tor Experimentation*

Live Tor and PlanetLab	Bauer et al. [WPES 07], Hopper et al. [CCS 07], Tang and Goldberg [WPES 07], McCoy et al. [PETS 08], Snader and Borisov [NDSS 08], McLachlan and Hopper [WPES 09], McLachlan et al. [CCS 09], Chaabane et al. [NSS 10], Mulazzani et al. [CMS 10], Tang and Goldberg [CCS 10], Luo et al. [ACSAC 11]
Emulation	Chakravarty et al. [ESORICS 10], AlSabah et al. [PETS 11], Moore et al. [ACSAC 11]
Simulation and Modeling	Borisov et al. [CCS 07], O'Gorman and Blott [ASIAN 2007], Murdoch and Watson [PETS 08], Ngan et al. [FC 10], Jansen et al. [CCS 10]

* Not a comprehensive list

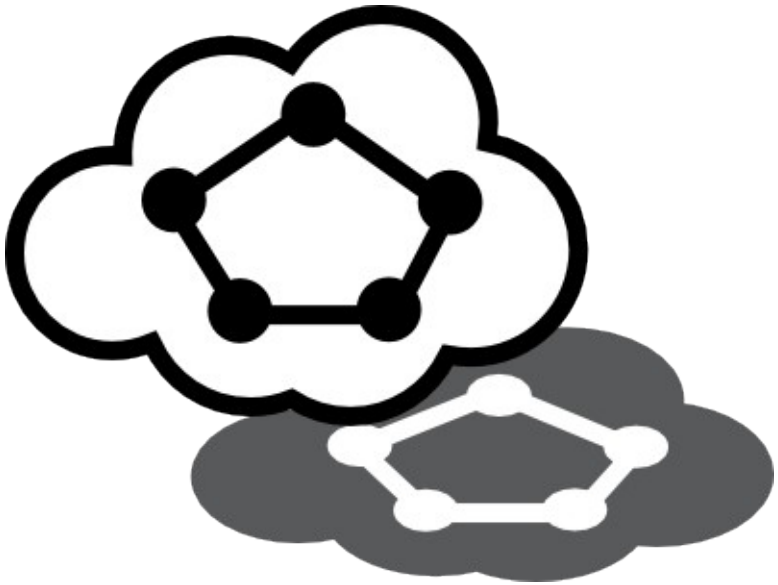
Network Experimentation

Approach	Disadvantages
Simulation	Not generalizable, inaccurate
Emulation	Large overhead, kernel complexities
PlanetLab	Hard to manage, bad at modeling

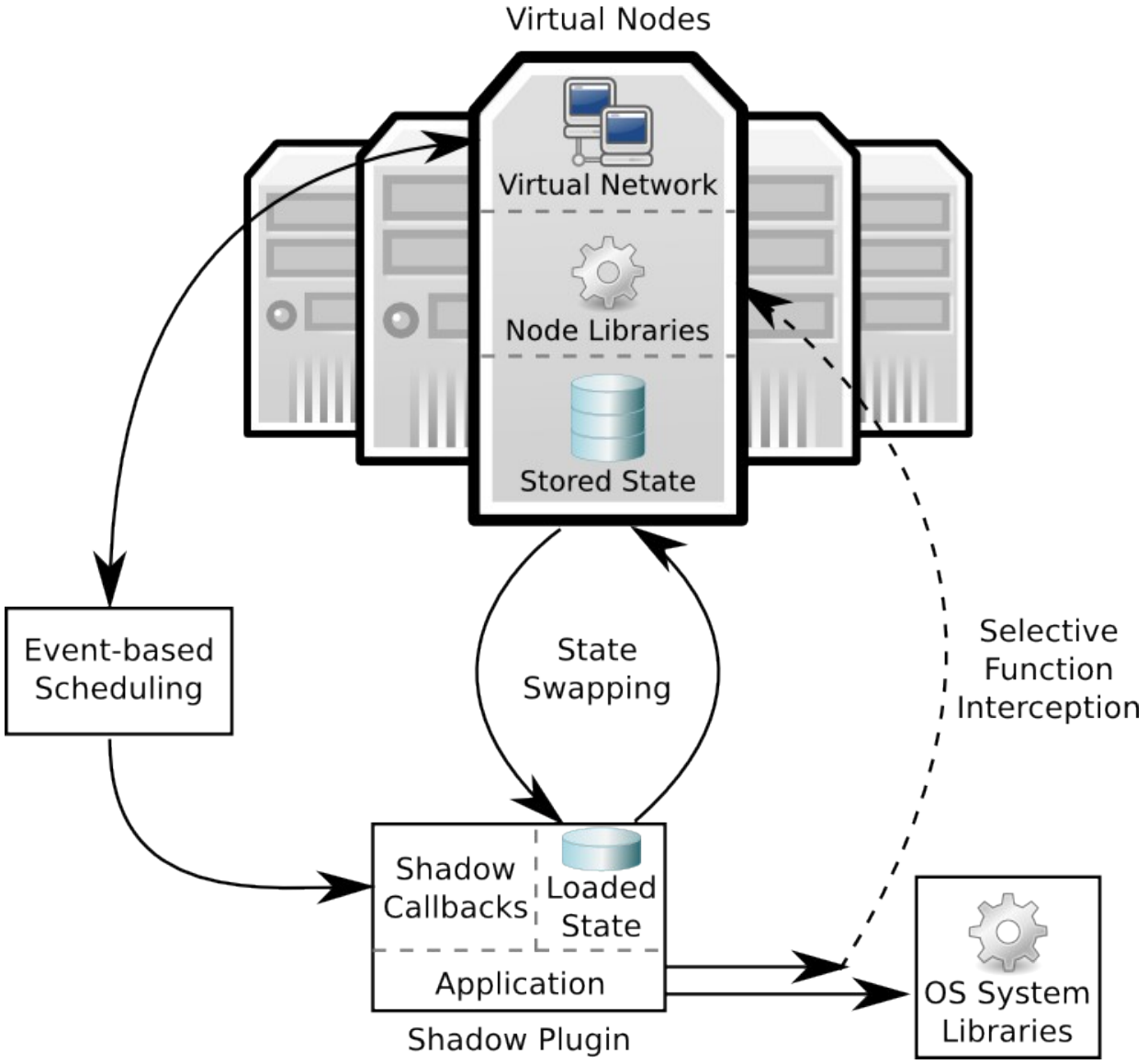


Tor in a Box with Shadow

- Discrete event network simulator
- Runs real application without modification
- Accurate, efficient, scalable
- Runs on Linux without root privileges



Shadow Architecture



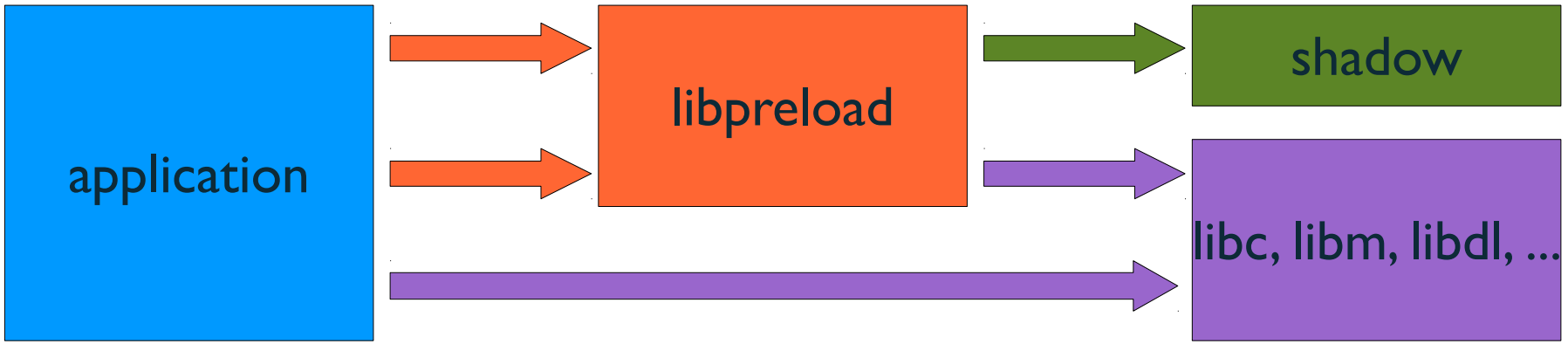
Function Interposition

- Intercept, redirect function calls
- $ readelf -s shadow
  - 0 FUNC GLOBAL **UND** socket@@GLIBC_2.2.5
  - 210 FUNC GLOBAL 13 vsocket_socket
- $ ldd shadow
  - libm.so.6 => /lib64/libm.so.6
  - libdl.so.2 => /lib64/libdl.so.2
  - libc.so.6 => /lib64/libc.so.6

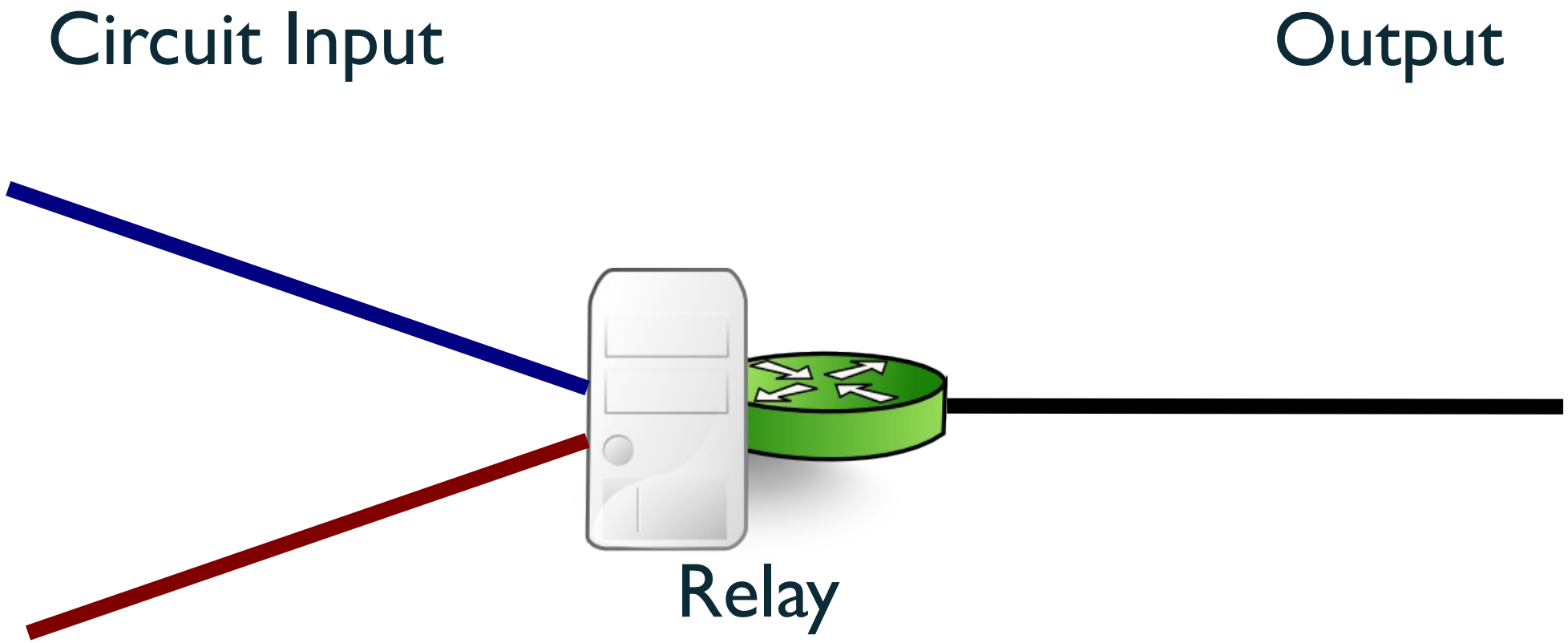
Function Interposition

- LD_PRELOAD=/home/rob/libpreload.so

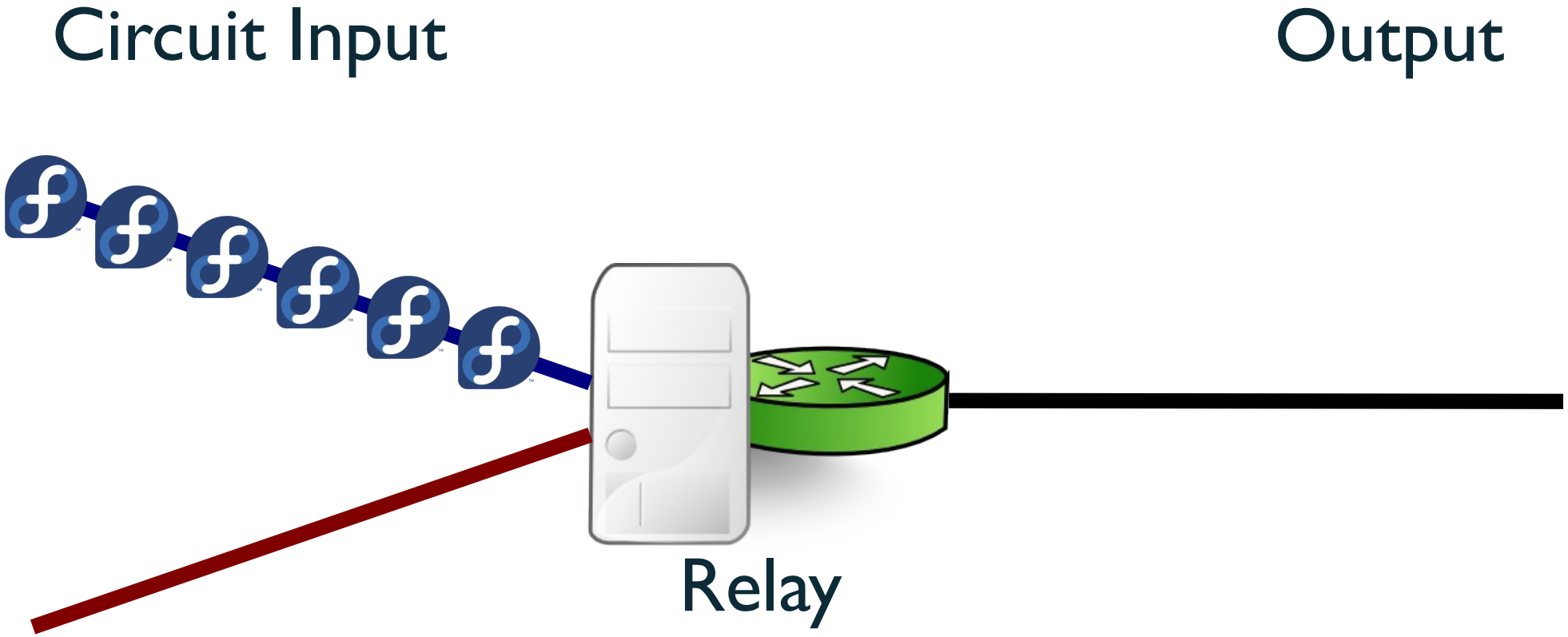
- Search my library first



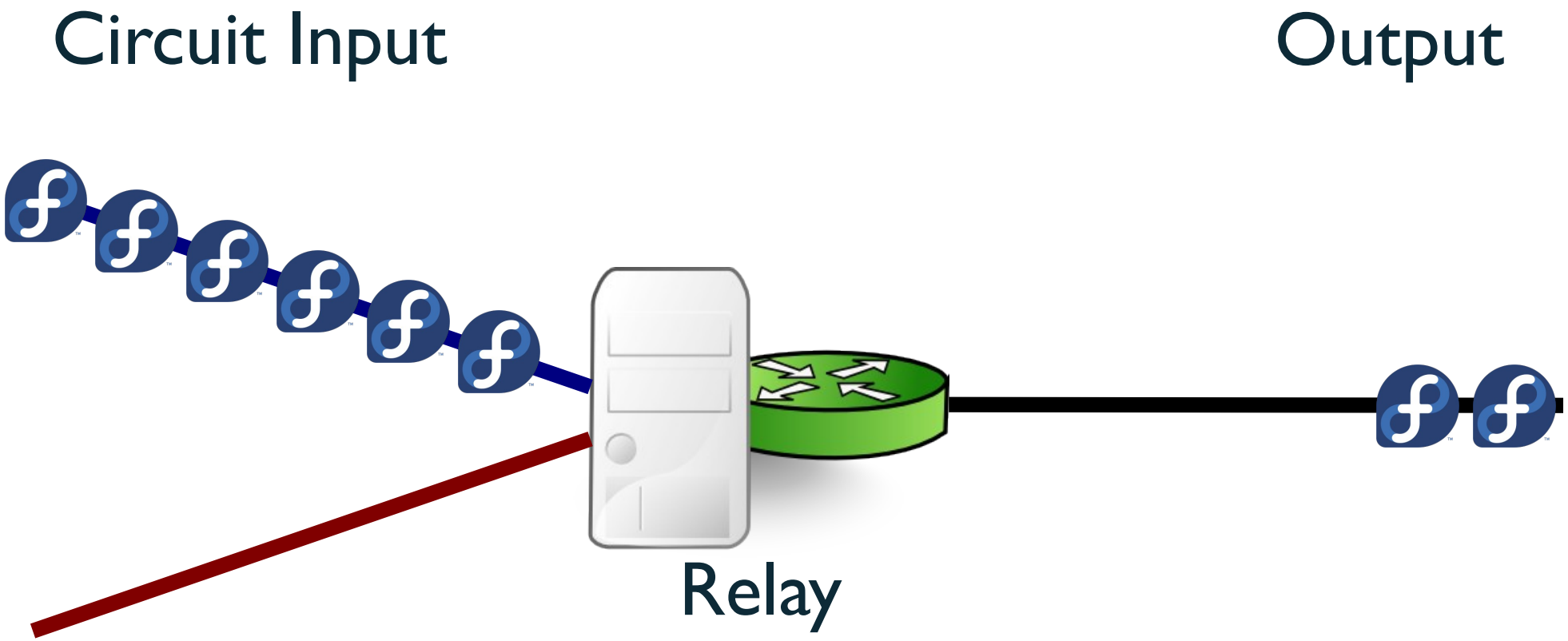
Tor Circuit Scheduling



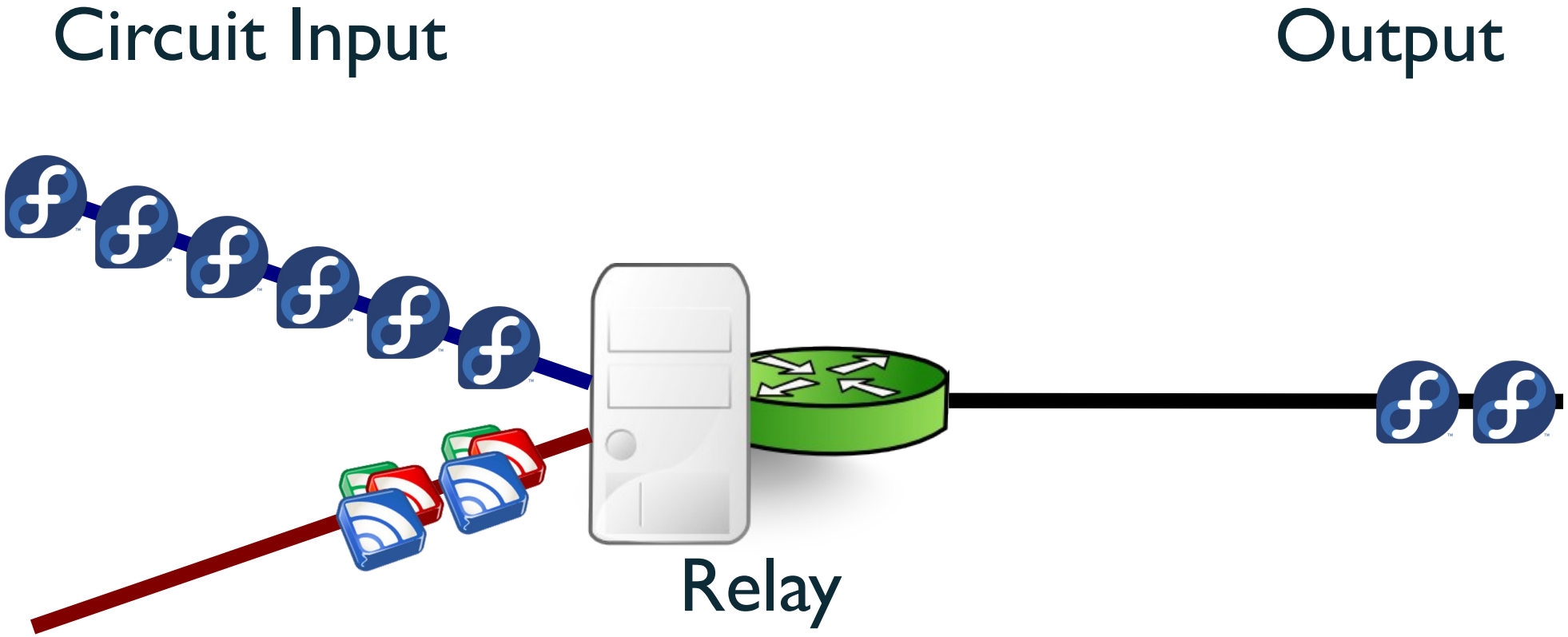
Tor Circuit Scheduling



Tor Circuit Scheduling



Tor Circuit Scheduling

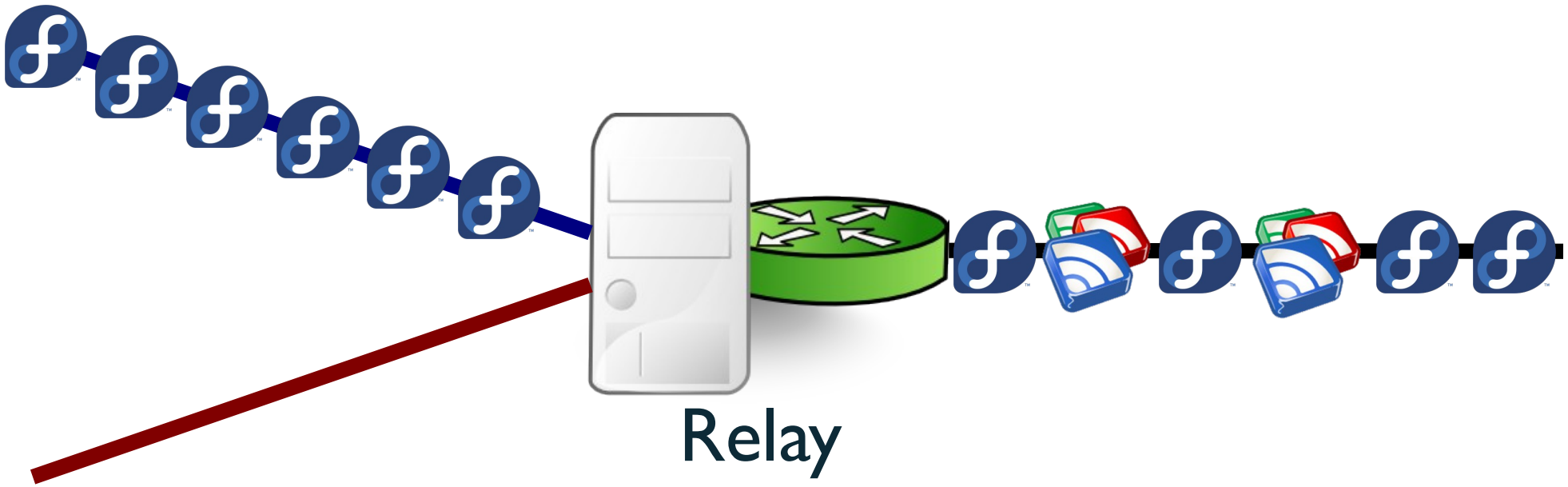


Tor Circuit Scheduling

Round Robin

Circuit Input

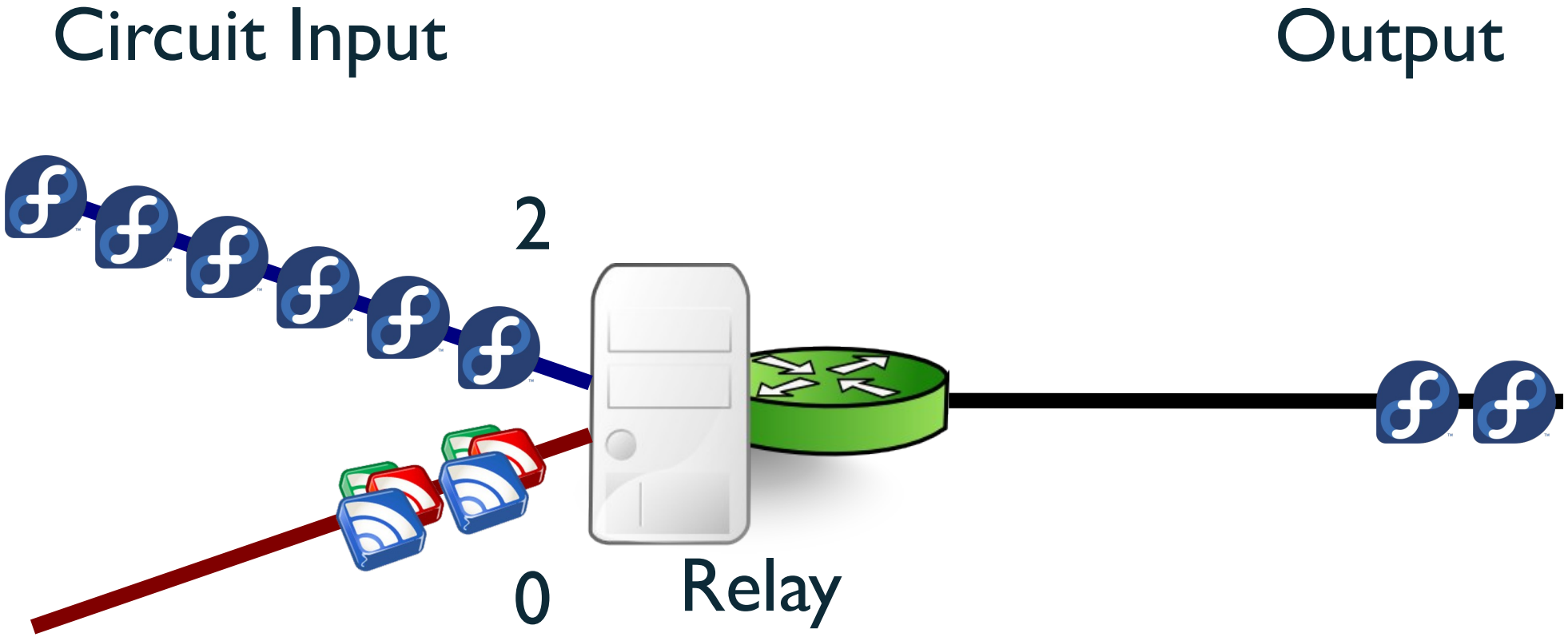
Output



Relay

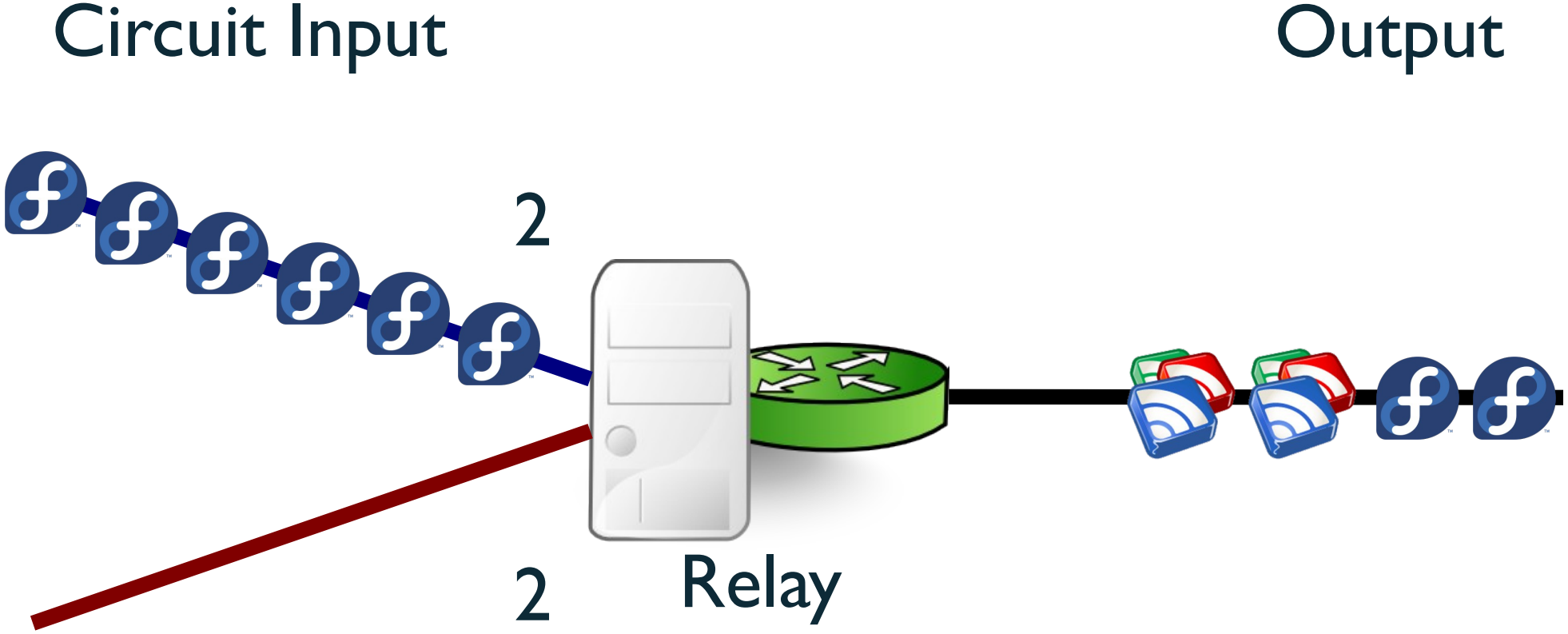
Tor Circuit Scheduling

EWMA [Tang and Goldberg CCS 2010]

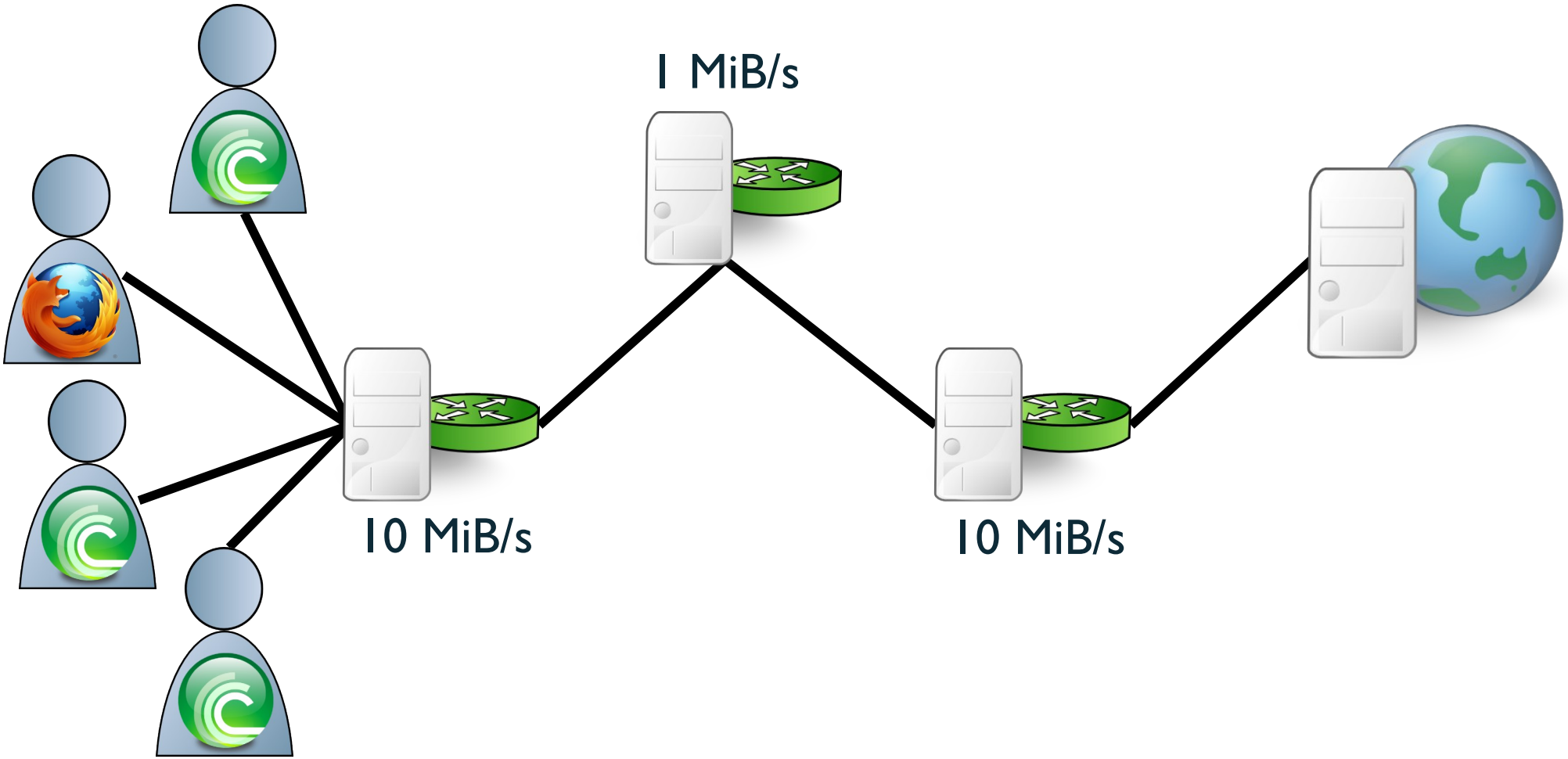


Tor Circuit Scheduling

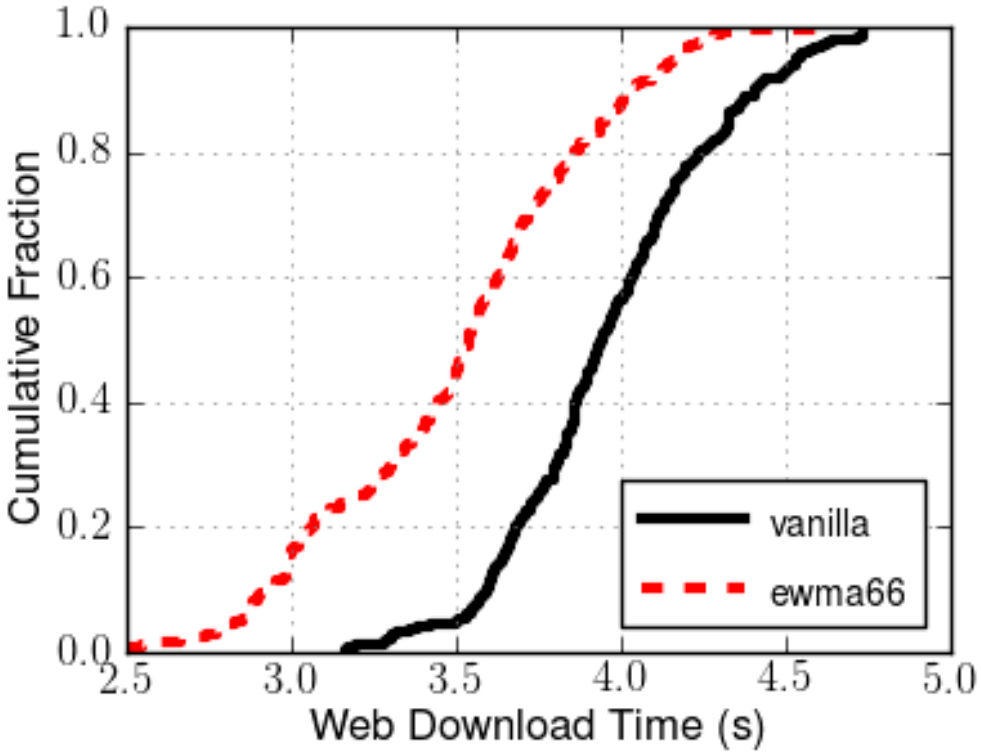
EWMA [Tang and Goldberg CCS 2010]

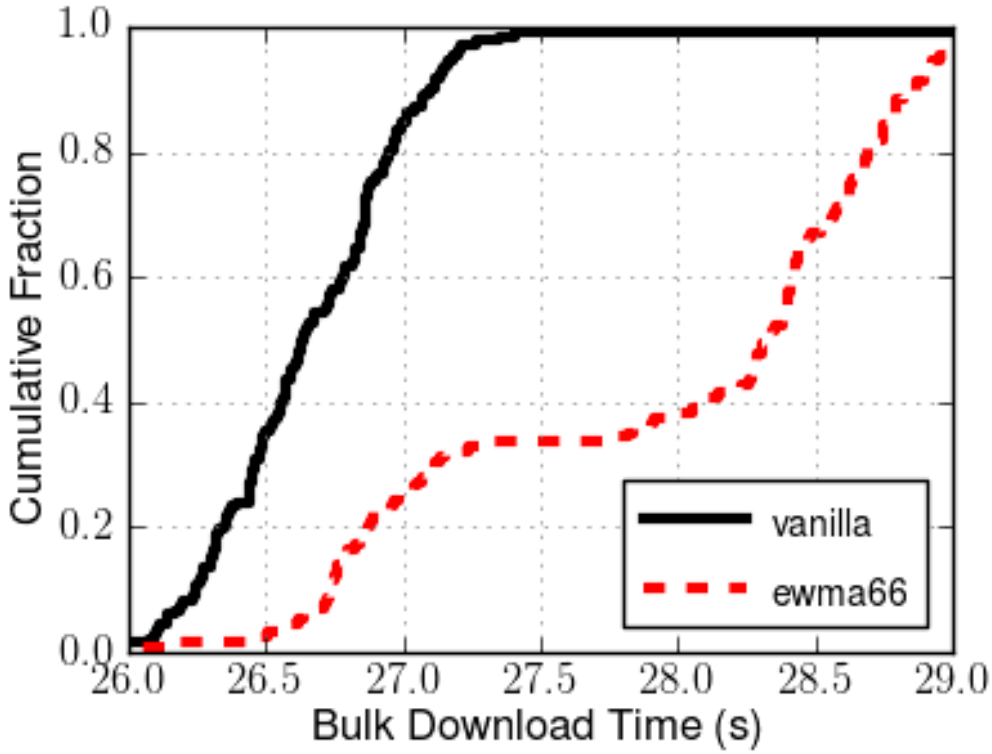


EWMA: Bottleneck



EWMA: Bottleneck





Summary

- Simulate time, network stack, crypto ciphers
- Model network latency and node bandwidth from real measurements
- Natively executes real application code