Applications for Measurement: Improving Anonymity Online
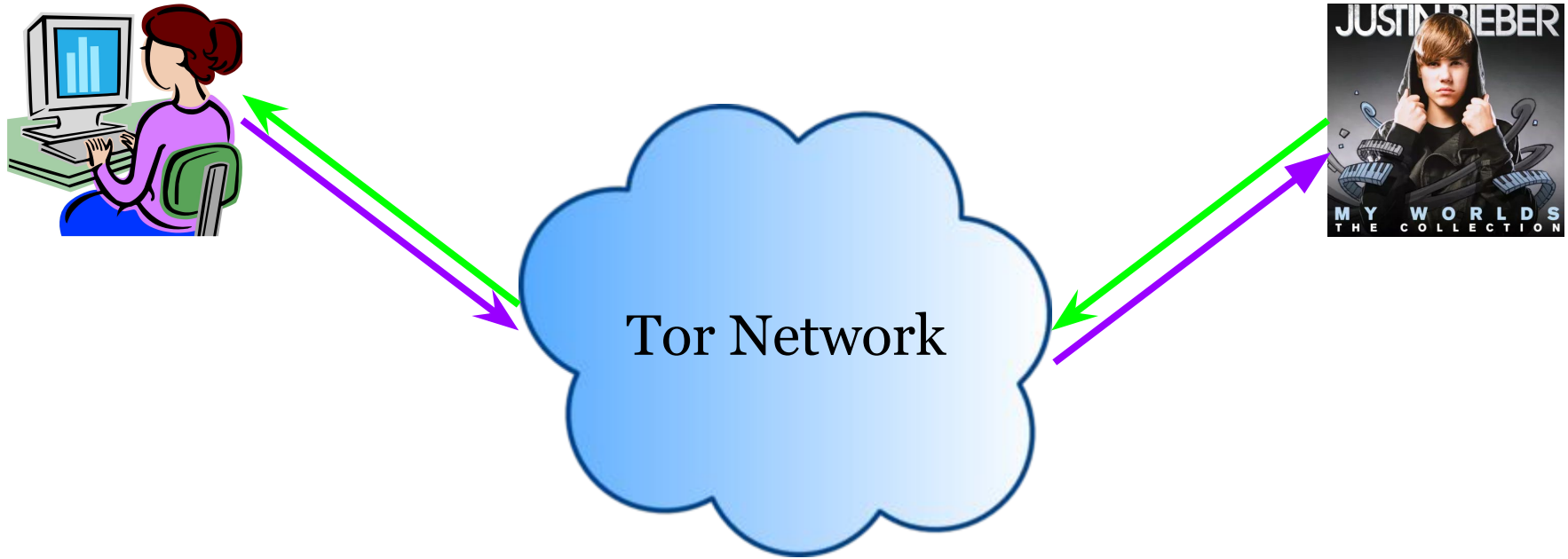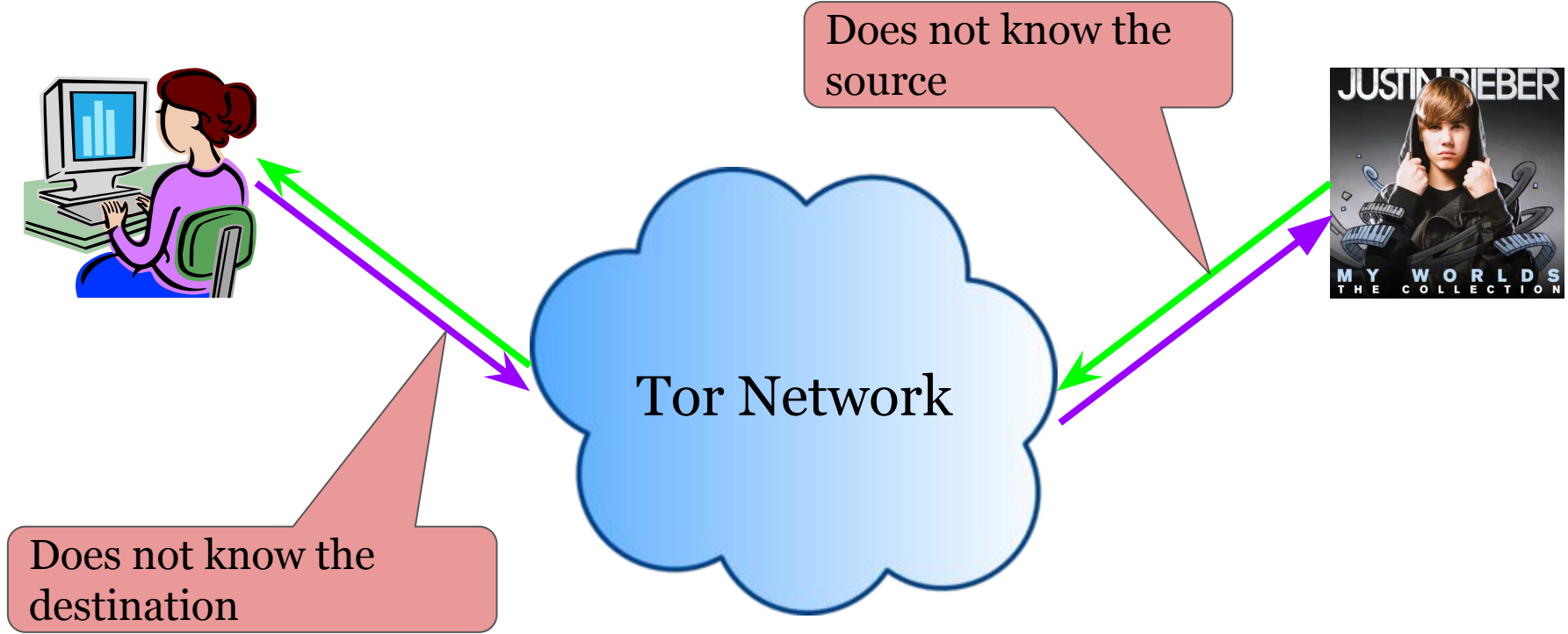
Rishab Nithyanand | **Rachee Singh** | Shinyoung Cho | Phillipa Gill
Stony Brook University
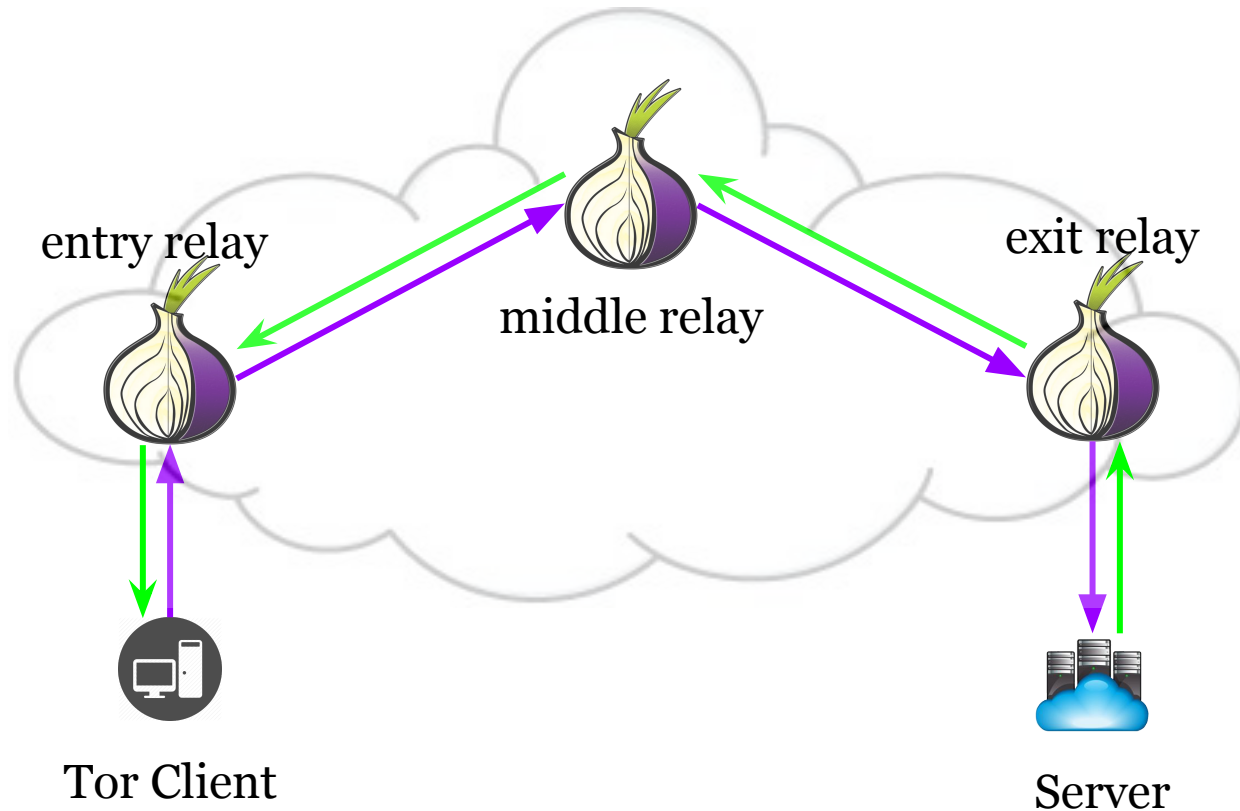
Anonymity on the Internet
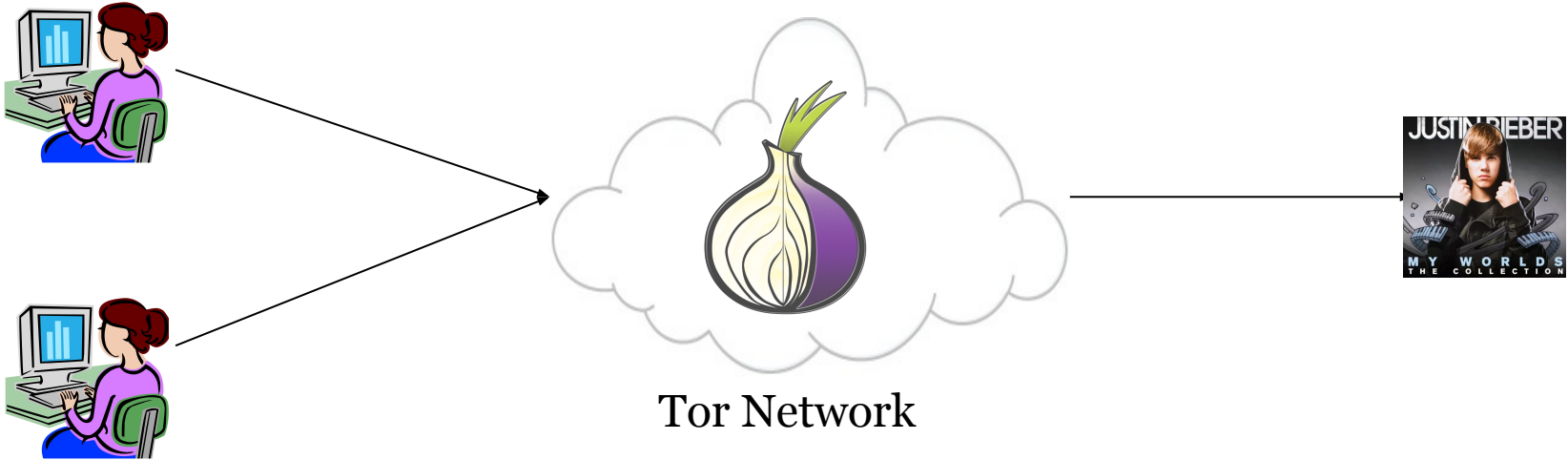
Anonymity on the Internet

Does not know the source

Tor Network

Does not know the destination

Online Anonymity via Tor
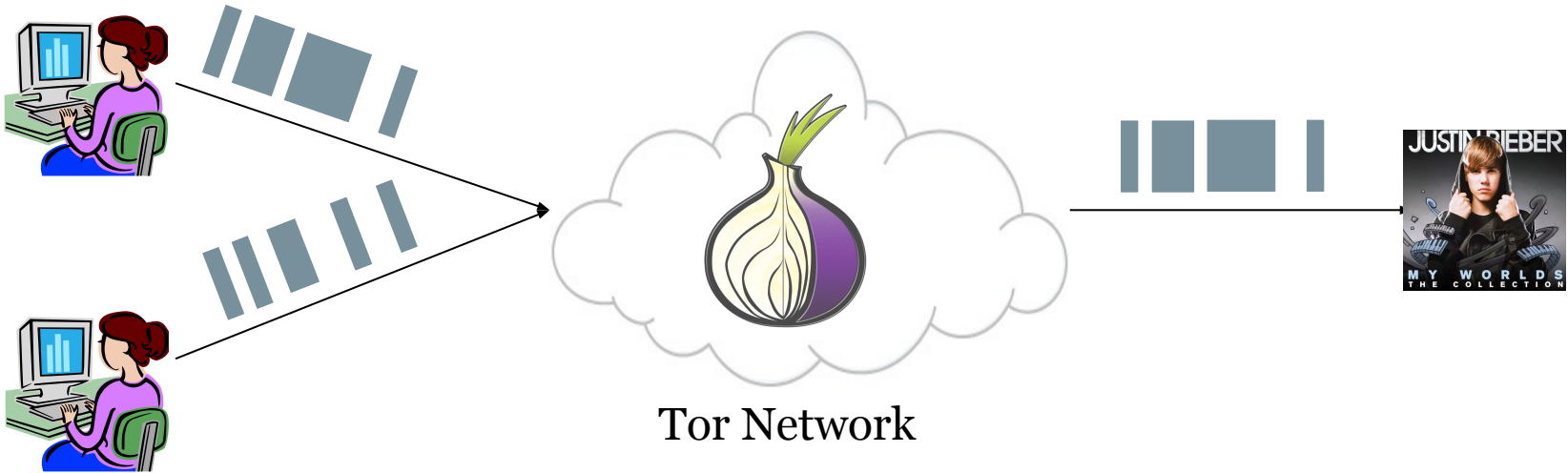
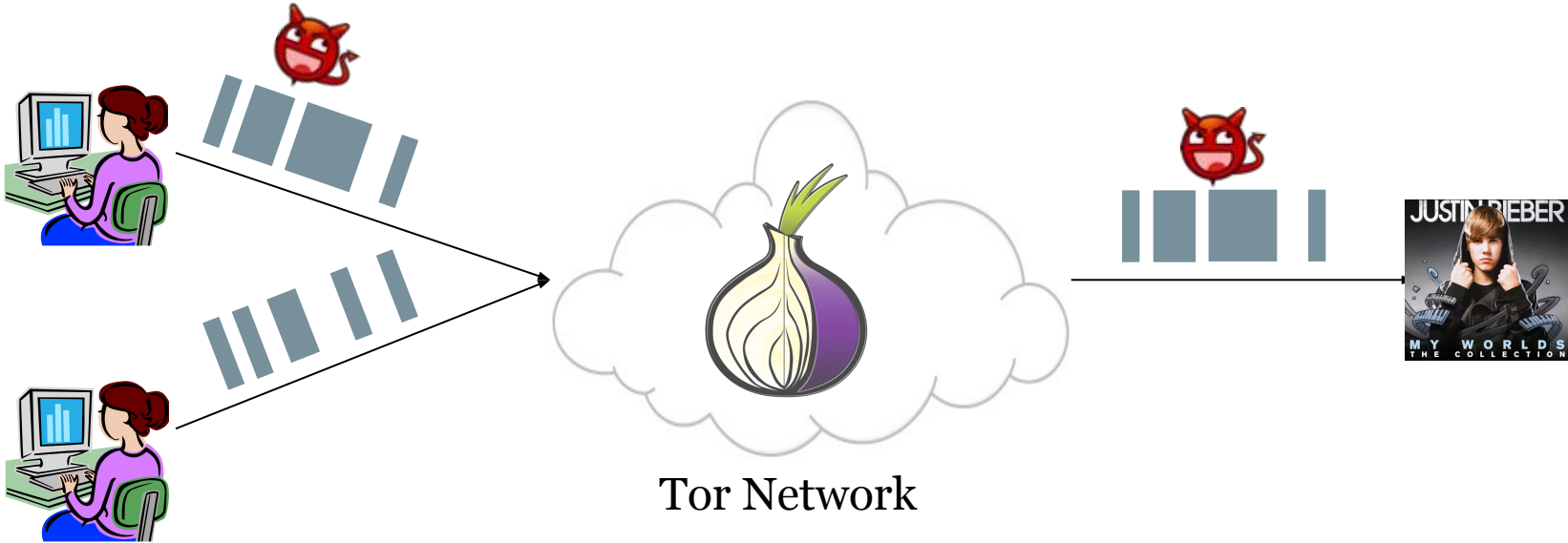entry relay

middle relay

exit relay

Tor Client

Server

Threat Model: Network Based Attacks

Tor Network

Threat Model: Network Based Attacks

Tor Network

Threat Model: Network Based Attacks

Tor Network

Internet routing and timing attacks

Tor Network

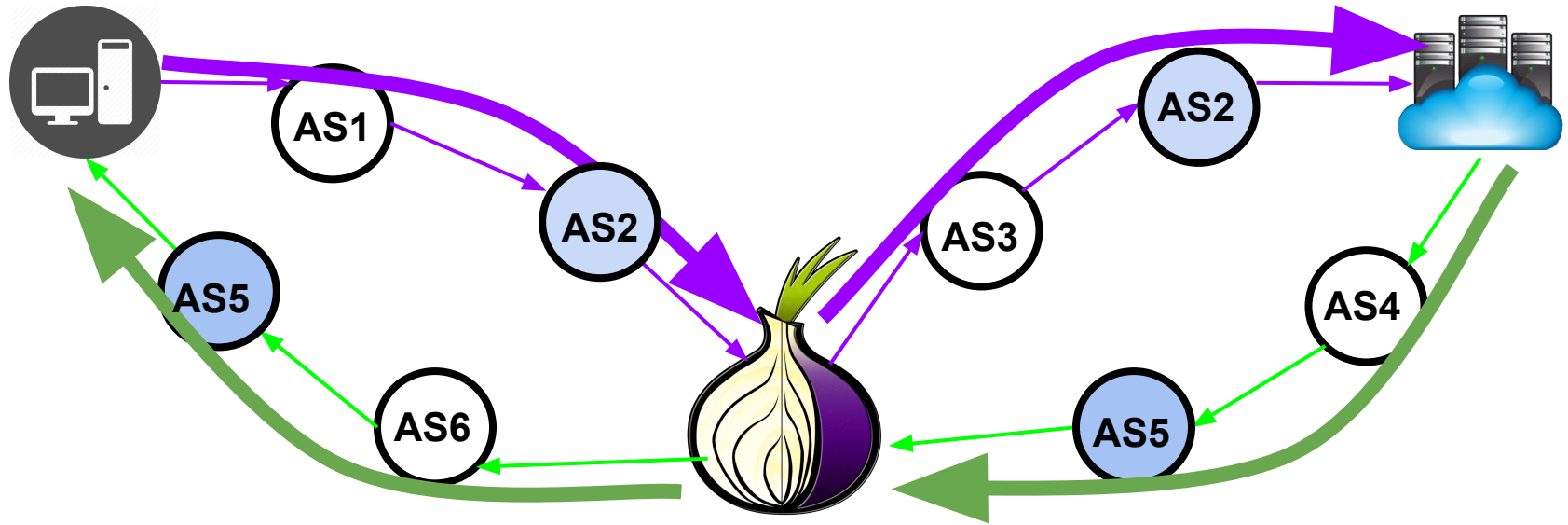# Internet routing and timing attacks

Tor Network
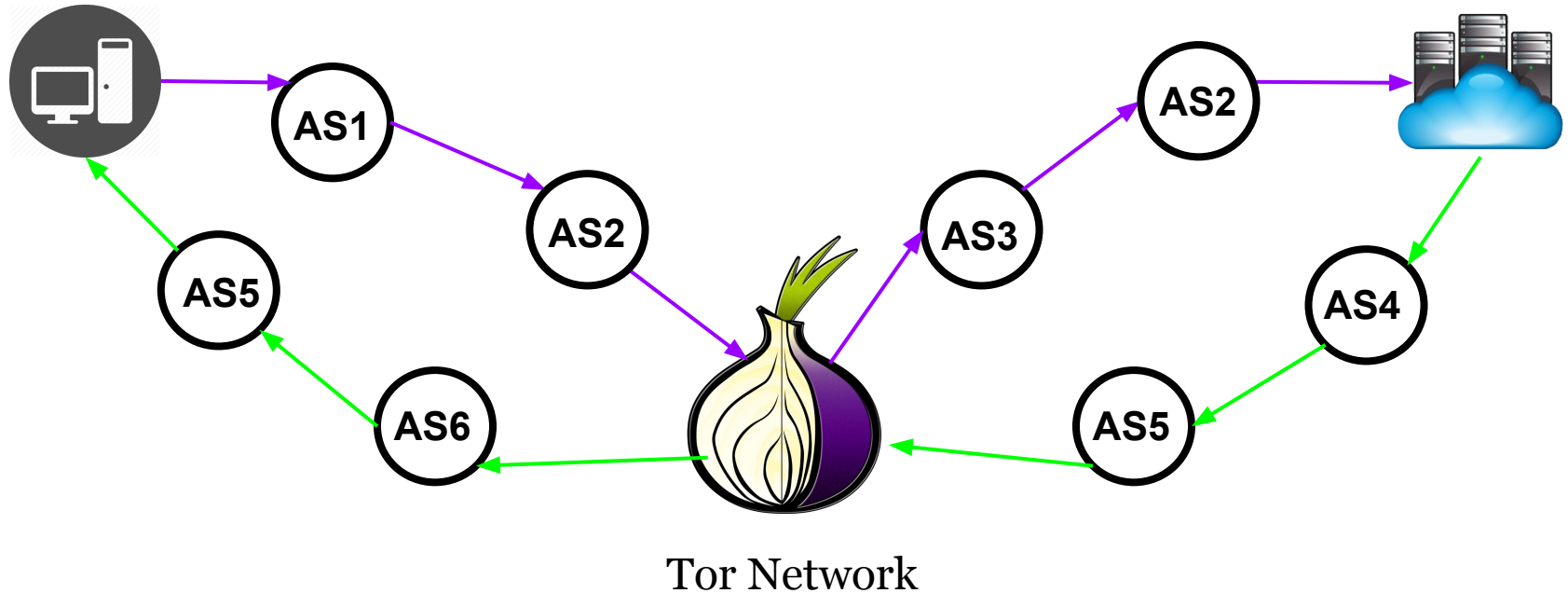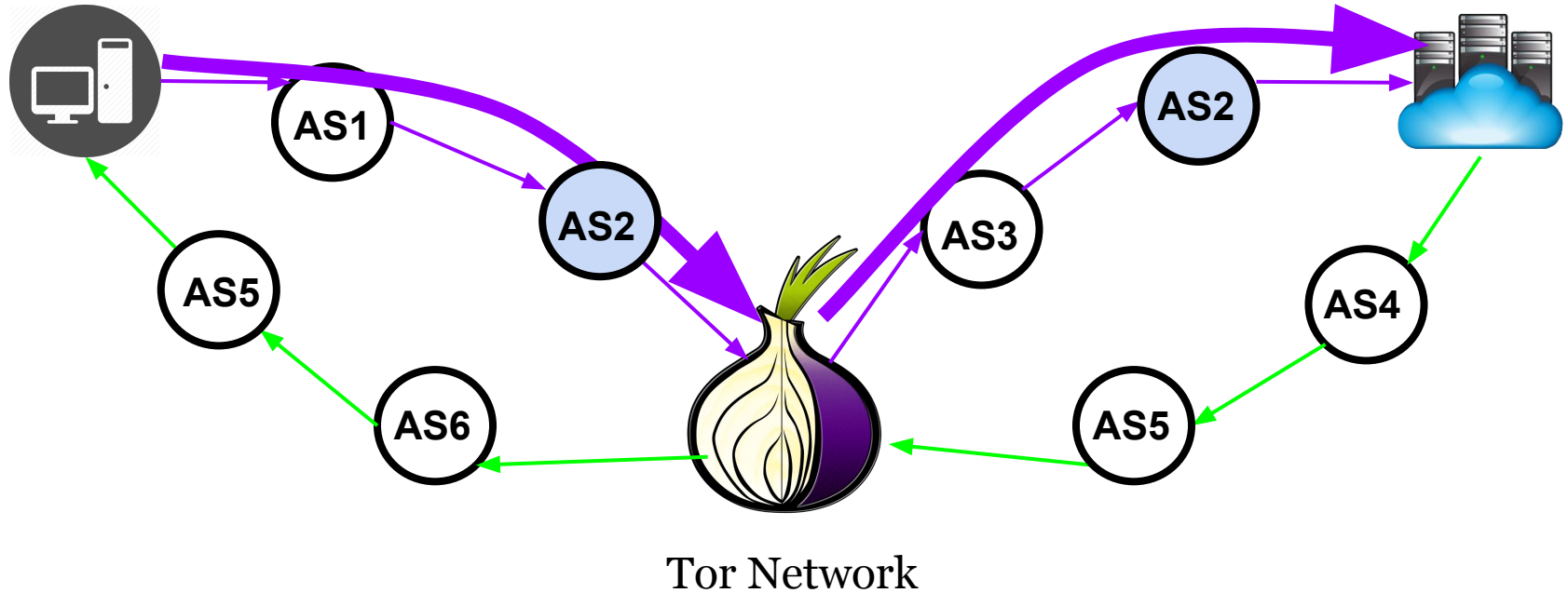
Internet routing and timing attacks

Path asymmetry => **Increases the attack surface** [RAPTOR, USENIX 2015]
TCP ACK numbers leak timing and size info on reverse path!

Astoria [NDSS2016] + Cipollino[*]

- What if the Tor client could pick relays to avoid timing attacks?
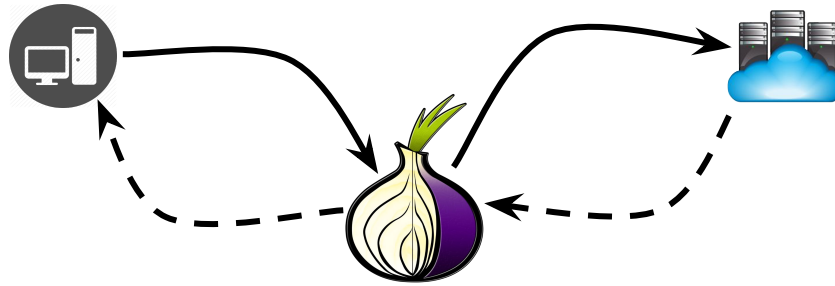
- We show that there usually is a safe option [NDSS2016]
- **Challenge:**
  - How can the Tor client learn network paths?
- Astoria: Policy-based simulations on empirically derived AS graphs
- Cipollino: **Based on measured paths**

* https://en.wikipedia.org/wiki/Cipollino

Tor client measurement-plane requirements

- Data needs to be current
 - E.g., if a path changes to go through a new AS we need to know!
- Path computations need to be **local**
  - The client can't ask a third party about paths to the destination!
- Data needs to be **compact** + **accurate**
- Trade-off between:
  - measured data as **relevant** as possible (**near real-time**).
  - Cipollino Tor client **low-latency** (not on-demand measurements)

Our solution: PathCache

- **Basic idea:** Reuse measurements already being made!
- Combine publicly accessible traceroute measurements to learn new paths
  - Currently using RIPE Atlas + iPlane data
  - Augmenting with control-plane data RIPE NCC, Routeviews
  - **Longer term:** Efficient use of new measurements to increase coverage
- http://pathcache.cs.stonybrook.edu
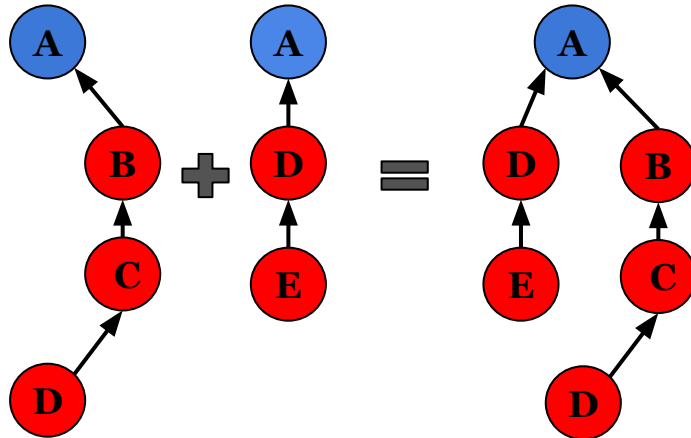
Why is PathCache Useful?

- Everyone needs traceroutes!
- But why run **redundant traceroutes** at the expense of a **constrained** measurement **budget**?
- Measurement hardware can be more effectively utilised.
- Standard and compact graphs as JSONs, easy to work with.
- Try it out here: http://pathcache.cs.stonybrook.edu/api/v1/174 (AS174's dest based graph)
- http://pathcache.cs.stonybrook.edu/api/v1/path?src=2119&dst=174

Future Work?

- Need more data
  - Huge benefit of periodic measurements from RIPE Atlas!
- Path prediction as a service
- A platform to maintain AS-level paths over time (benefit of retrospective measurements)
- Provide a generic interface to upload user-run measurements.
- As a community, increase our coverage via measurement reuse