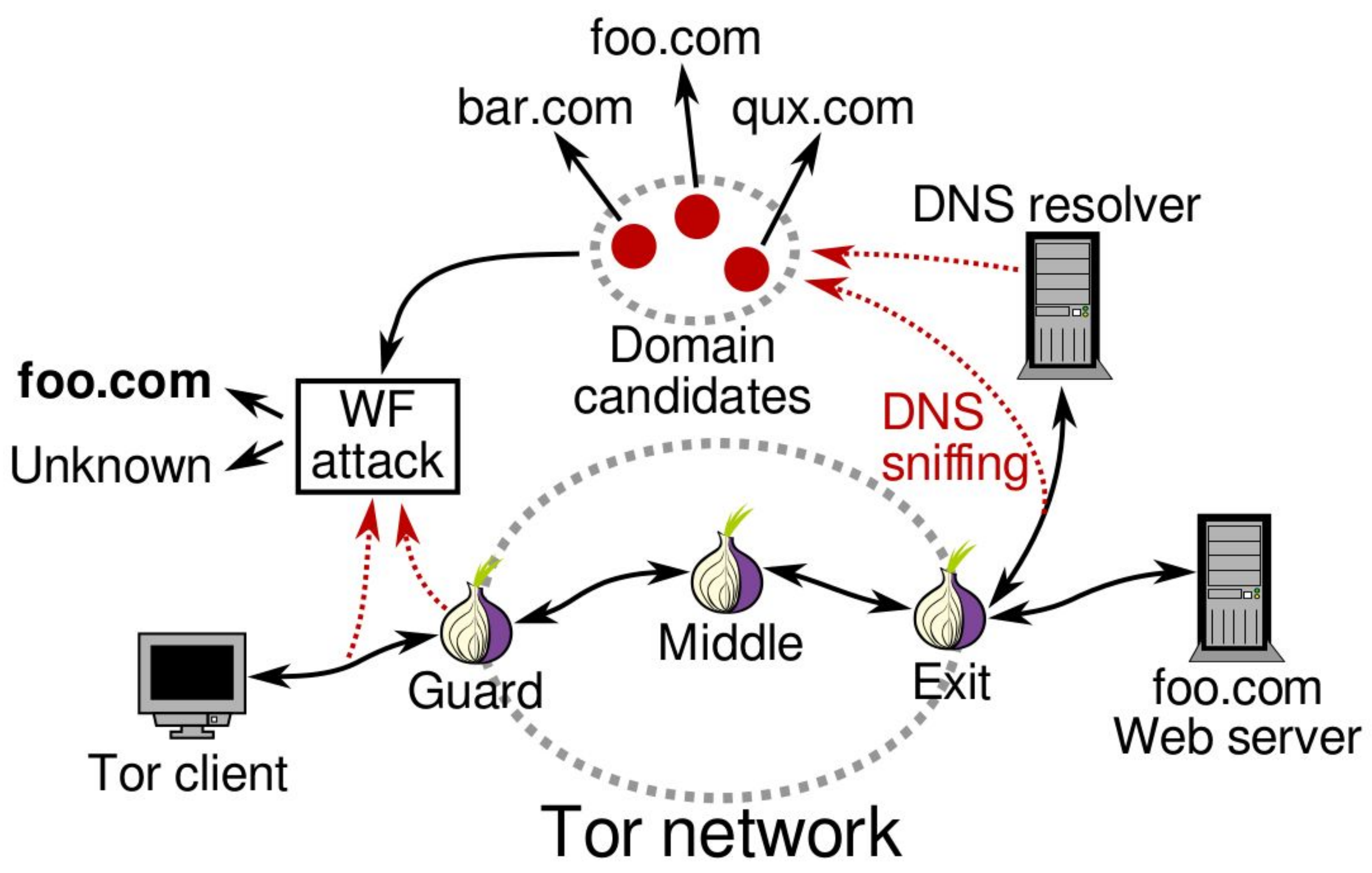
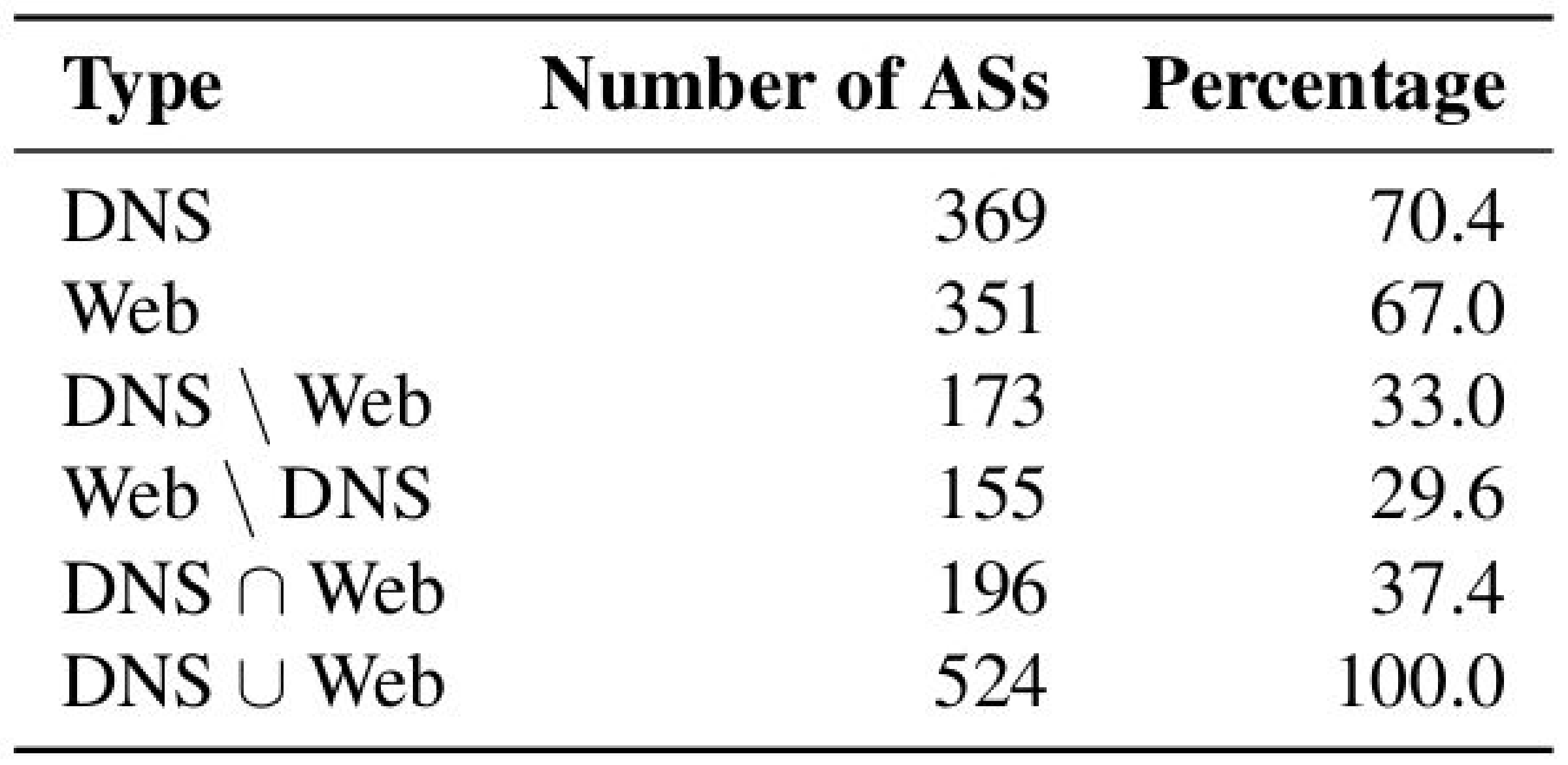
DNS-based Traffic Correlation Attacks

End-to-end correlation attacks

- Adversary seeks to control **both ends** of low-latency anonymity network, e.g., Tor
- Then, simple **packet counting** techniques allow **deanonymization**
- Past work focused on client-to-server TCP stream, **ignoring** DNS's distributed nature



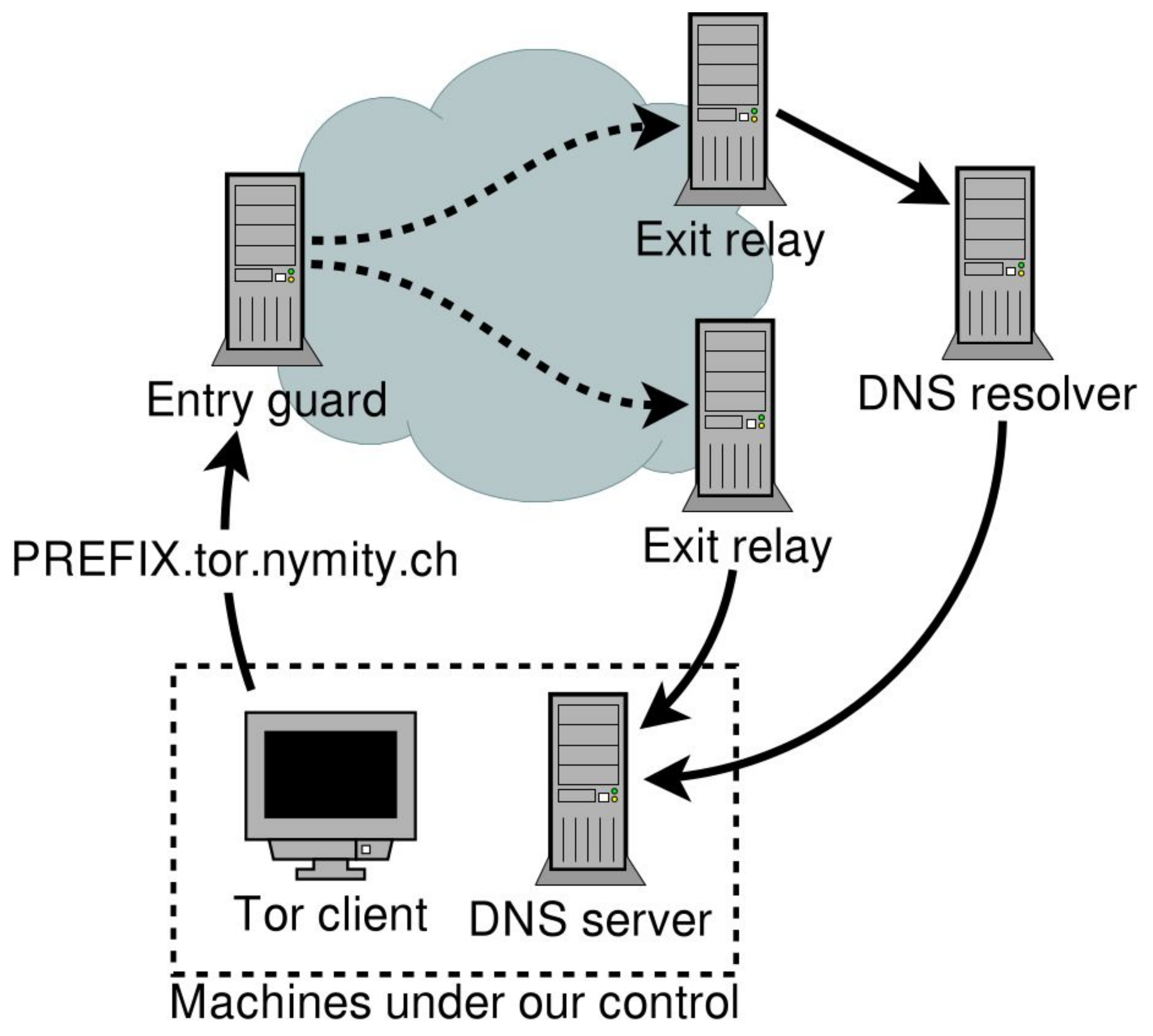


Type	Number of ASs	Percentage
DNS	369	70.4
Web	351	67.0
DNS \ Web	173	33.0
Web \ DNS	155	29.6
DNS ∩ Web	196	37.4
DNS ∪ Web	524	100.0

Table 2: *The set relations between unique traversed ASs for DNS and unique traversed ASs for Web.*

Why is DNS an issue?

- Iterative queries **traverse many paths** in addition to point-to-point TCP connection
- Third-party resolvers (e.g., 8.8.8.8) shouldn't learn **what Tor users do**
- Tor's DNS resolution is entirely up to exit relays. Here be dragons.

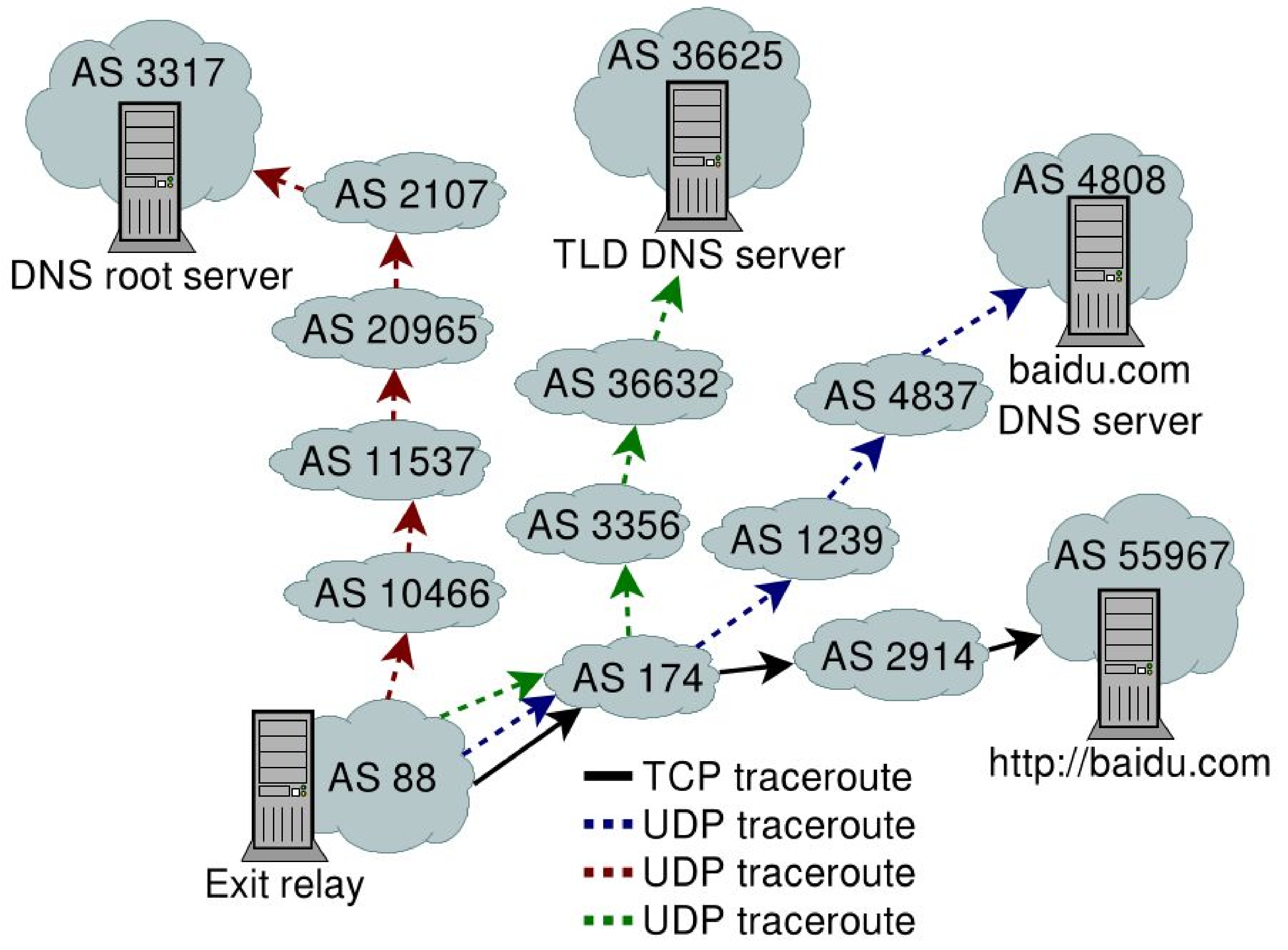


Who we are

At Princeton	At Karlstad	At KTH
Nick Feamster	Tobias Pulls	Benjamin Greschbach
Jon Metzman		
Laura Roberts		
Philipp Winter		

More information:

https://nymity.ch/dns-traffic-correlation/



Preliminary results

- Google gets to see **~25%** of DNS requests exiting Tor (bad)
- 12% of DNS requests come from self-hosted resolvers (good...?)
- Most exit relay resolvers use **0x20 encoding** and **random source ports** (good)
- DNS traffic traverses **more ASs** than Web traffic (bad)
- ~33% of exit resolvers don't validate DNSSEC (bad)