Scalable, Accountable, Traffic Analysis Resistant Anonymity in Dissent

Bryan Ford
Yale University

project team:
David Isaac Wolinsky, Henry Corrigan-Gibbs, Joan Feigenbaum, Vitaly Shmatikov, Ewa Syta, Aaron Johnson, Ramakrishna Gummadi

Columbia University – September 25, 2012

Talk Outline

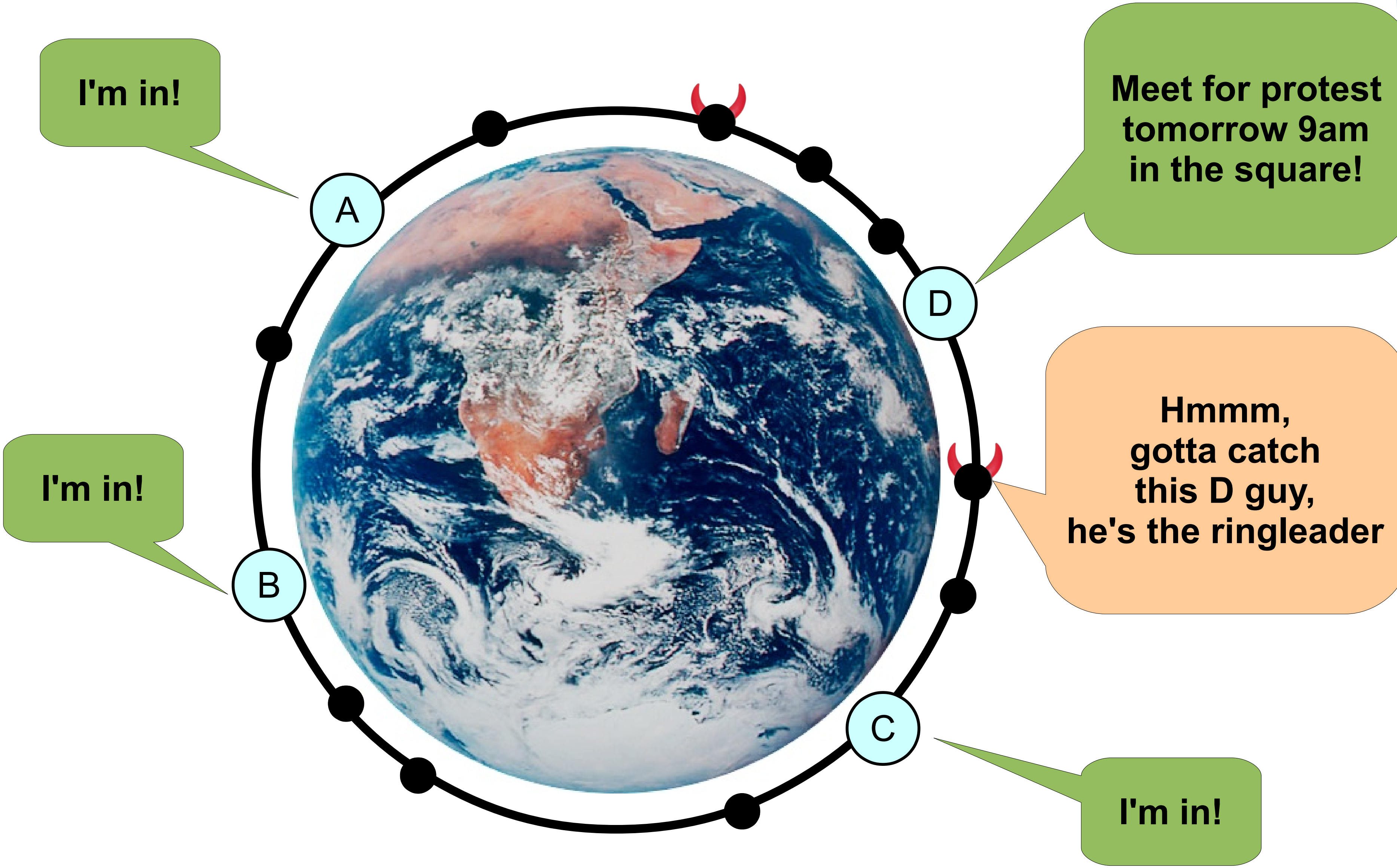
- Online anonymity: state-of-the-art, weaknesses
- Dining cryptographers: a cool, useless toy?
- Making DC-nets scale to “real” systems
- Accountability – in many flavors
- Anonymity scavenging and intersection attacks
- Conclusion

Why Anonymity?

Plays fundamental roles in democratic societies

- Discuss sensitive topics, freedom of speech
- Voting in elections or deliberative organizations
- Peer review processes
- Collaborative content creation, e.g., Wikipedia
- Protect dissidents in authoritarian states
- Whistleblowing
- Private bidding in auctions

A Protest in Repressistan



A Protest in Repressistan

Alice, Bob, Charlie, Dave, & friends

- Citizens of Repressistan
- Wish to connect, organize online safely

Government is powerful but not all-powerful

- Can't just “turn off Internet” indefinitely or throw *all* protesters in jail: cost is too high
- Must identify and make examples of the movement's outspoken “activist leaders”

Alice & friends need “strength in numbers”

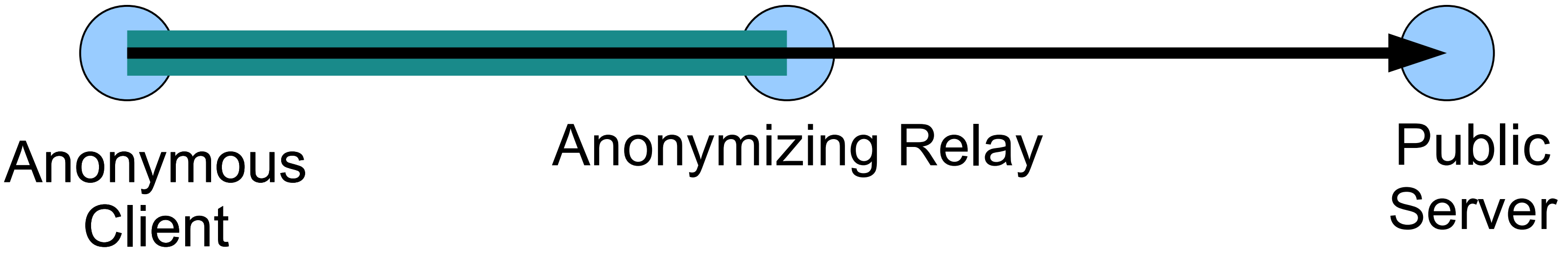
Being Anonymous: Naive Ways

Assume the Internet is “anonymous enough”

- IP addresses never provided real anonymity; many ways to track users, machines, browsers

Use centralized anonymizing relays/proxies

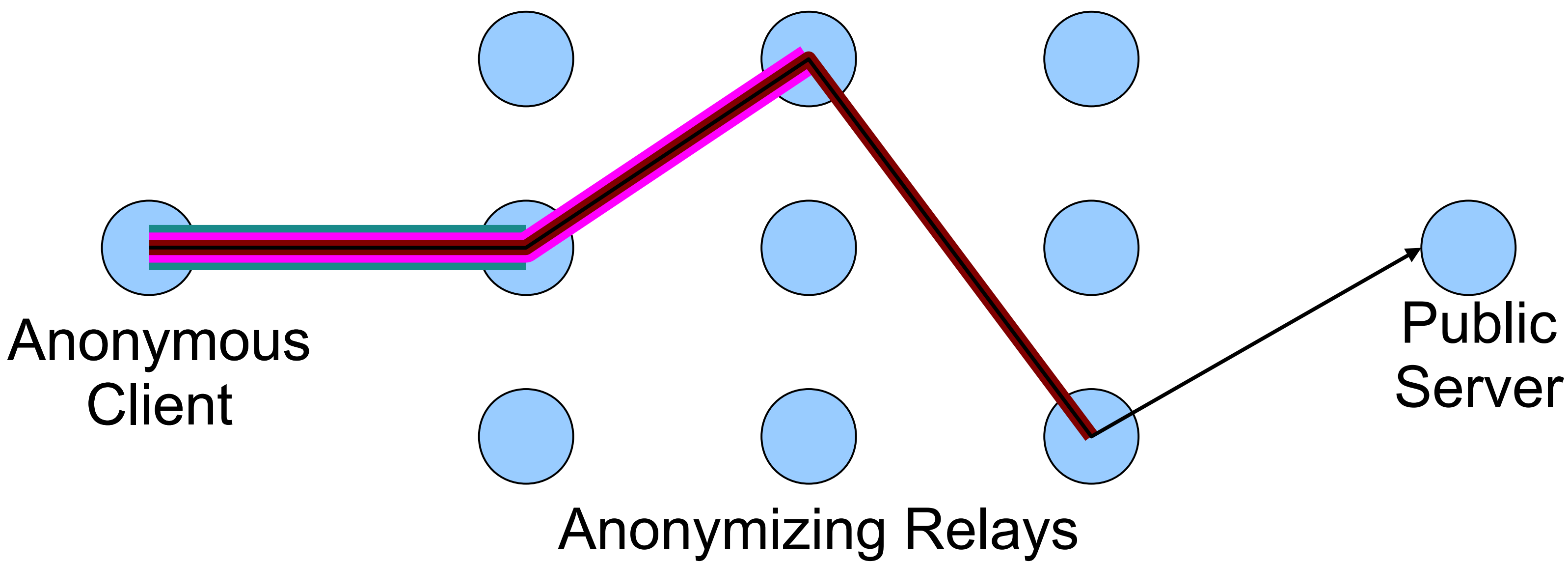
- Central point of failure, prime compromise target



Being Anonymous: Better Ways

MIX networks, onion routing systems: e.g., Tor

- Tunnel through a series of anonymizing relays
- Protects even if any one is malicious or hacked



Anonymity is Hard

Tor: The Onion Router [Dingledine'04]

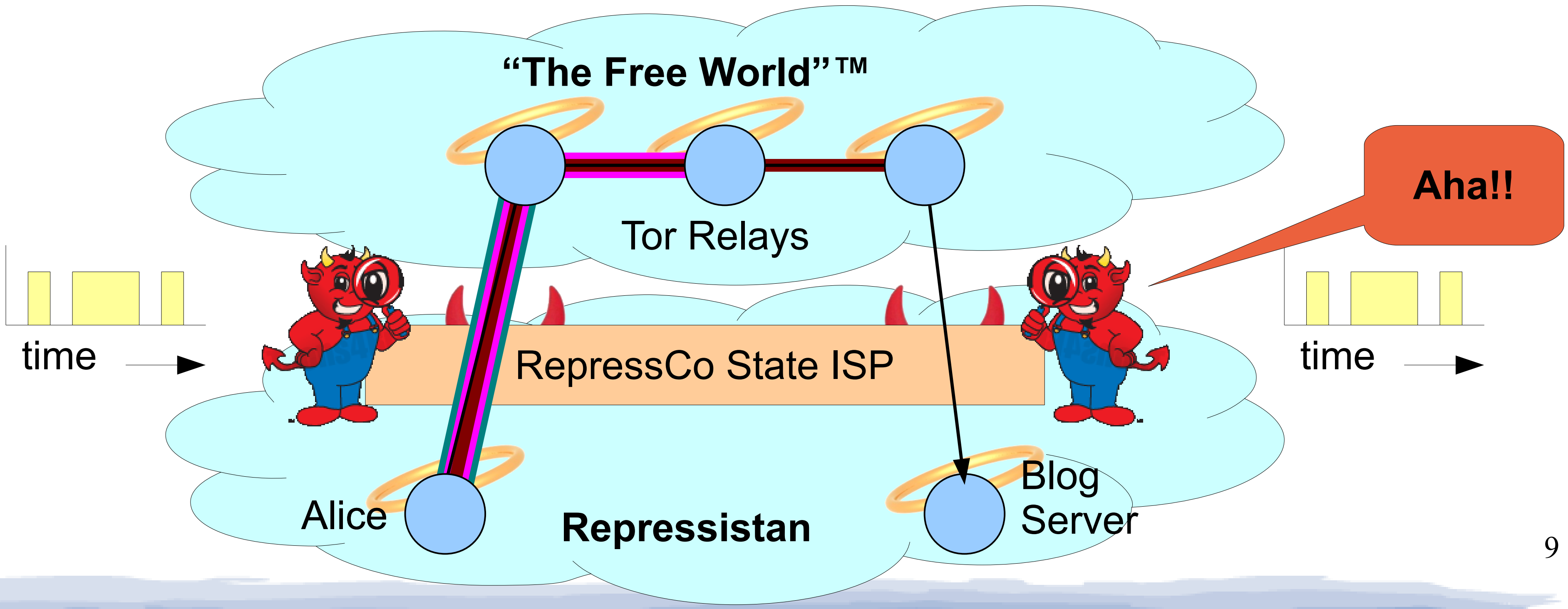
- Practical, scalable, convenient, widely deployed
- Likely best anonymity protection available now

But many known attacks, weaknesses

- Traffic analysis, traffic fingerprinting attacks
- Long-term intersection attacks [Kedogan'02]
- DoS attacks against anonymity [Borisov'07]
- Side-channel leaks/attacks [Abbott'07]

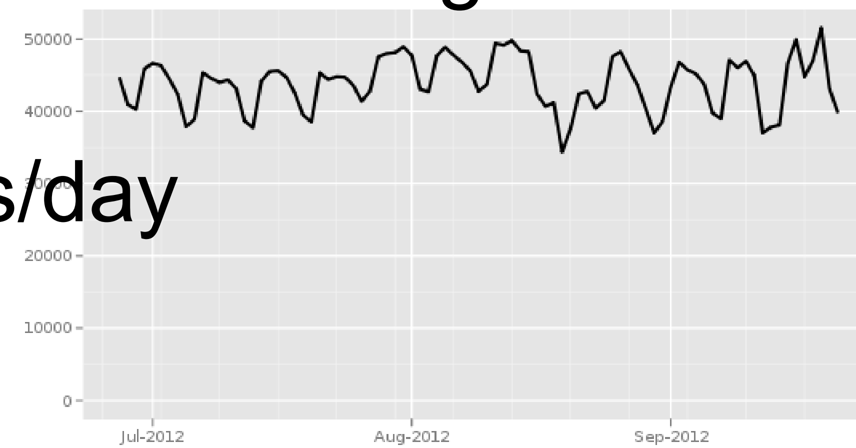
Traffic Analysis: Example 1

- Alice in Repressistan uses Tor to post on blog server hosted in Repressistan
- State ISP controls *both* entry and exit hops
- Fingerprint & correlate traffic to **deanonymize**



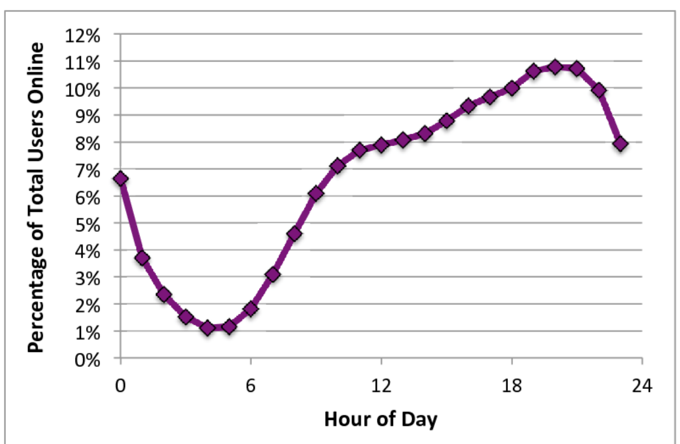
Traffic Analysis: Example 2

- Bob in Dictatopia posts via Tor to blog hosted in “The Free World”™
- Tor Metrics: 50,000 users/day connect from Dictatopia
  - Good anonymity, right?



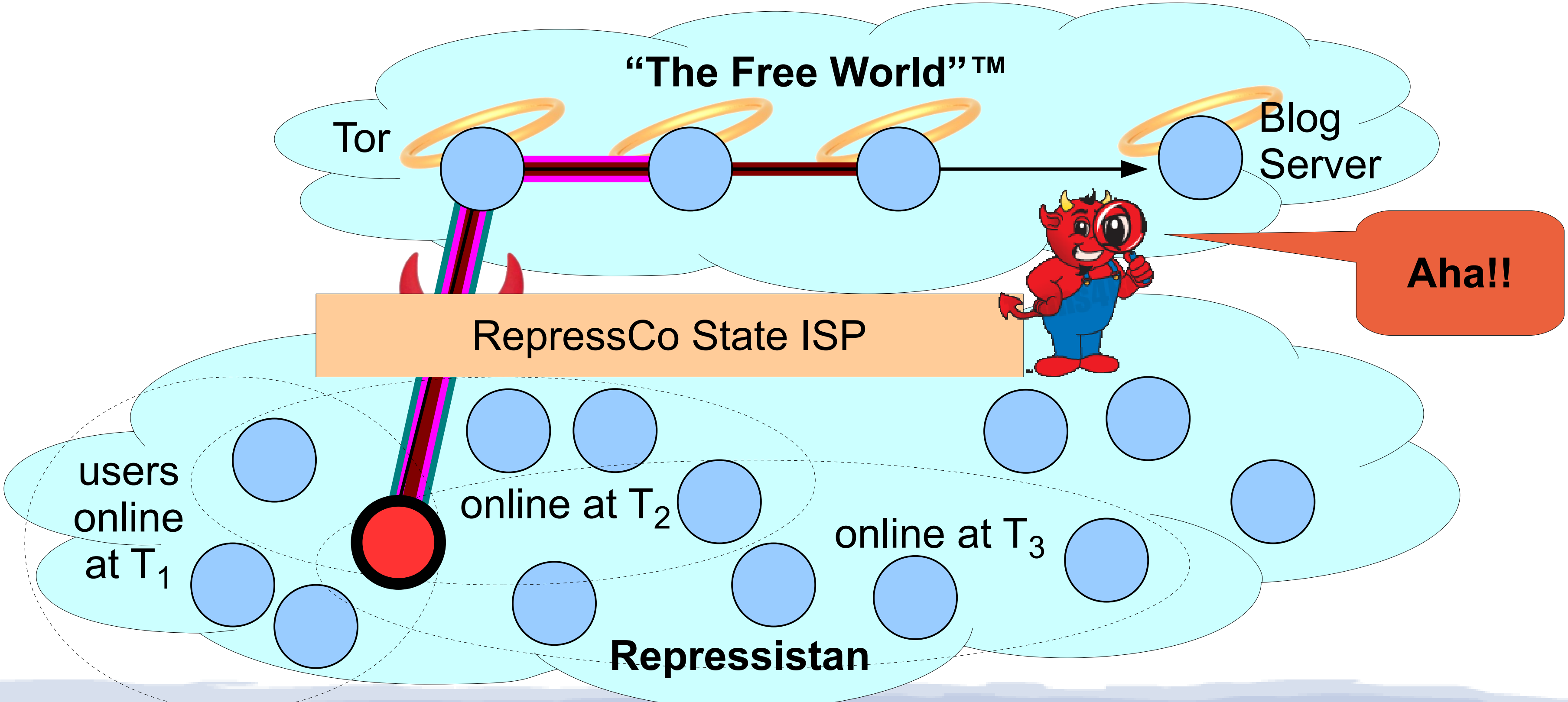
The Tor Project - https://metrics.torproject.org/

- But ISP logs tell police when users are online; blog post has timestamp
  - How many users are online ***at same time Bob posts?***
    - ~5,000 at 7PM?
      ~500 at 5AM?



Intersection Attack: Example

- Bob signs posts with pseudonym “AnoniBob”
  - Posts 3 signed messages at times T_1, T_2, T_3
  - Police find sets of users online each time, **intersect**



Maybe Anonymity is **Bad**?

Vulnerable to anonymous abuse by users, no **accountability** for misbehavior

- No one knows you're a dog
- So anybody can behave like one

Cause: unlimited supply of “free” pseudonyms

- Create sock-puppet “supporters” in online forums
- Vote many times in online polls, elections
- Get banned, respawn at new IP address
  - loser is *next* user of old IP address or Tor exit relay

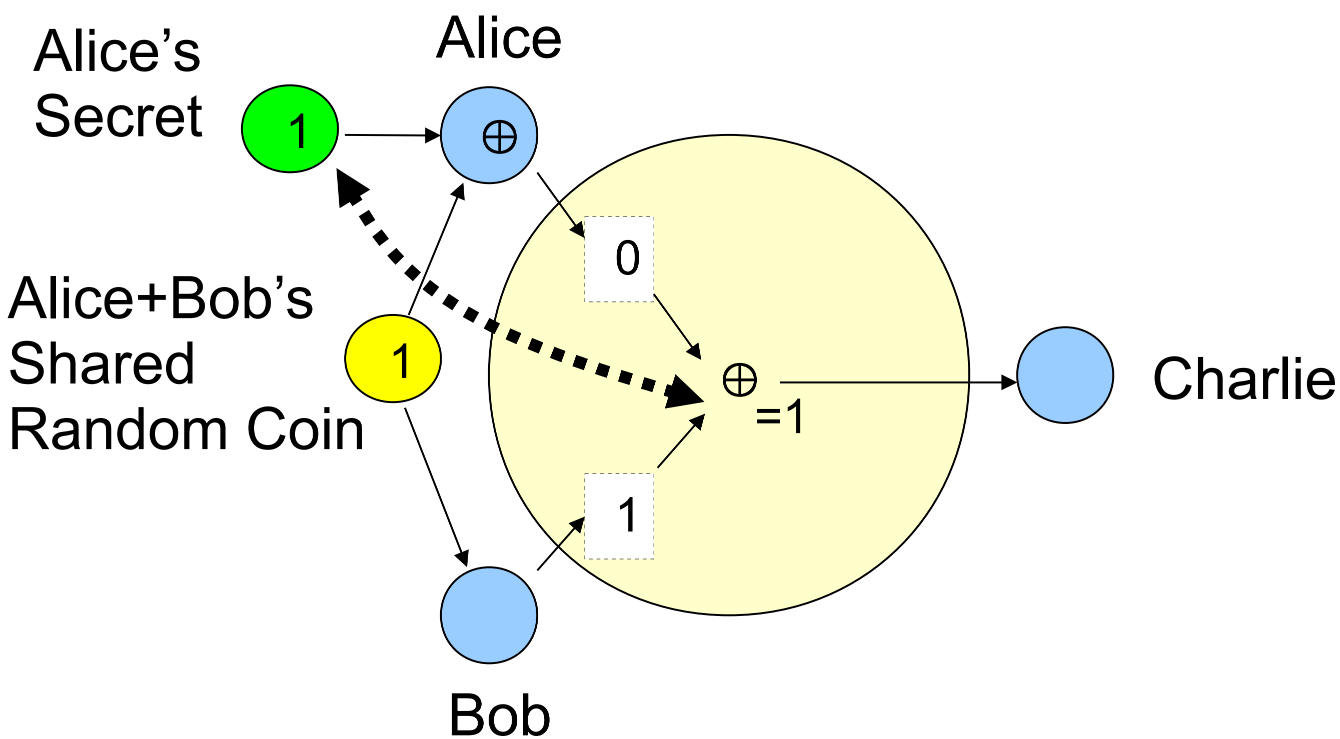
Talk Outline

- ✓ Online anonymity: state-of-the-art, weaknesses
- **Dining cryptographers: a cool, useless toy?**
- Making DC-nets scale to “real” systems
- Accountability – in many flavors
- Anonymity scavenging and intersection attacks
- Conclusion

Dining Cryptographers (DC-nets)

Another fundamental Chaum invention from the 80s...

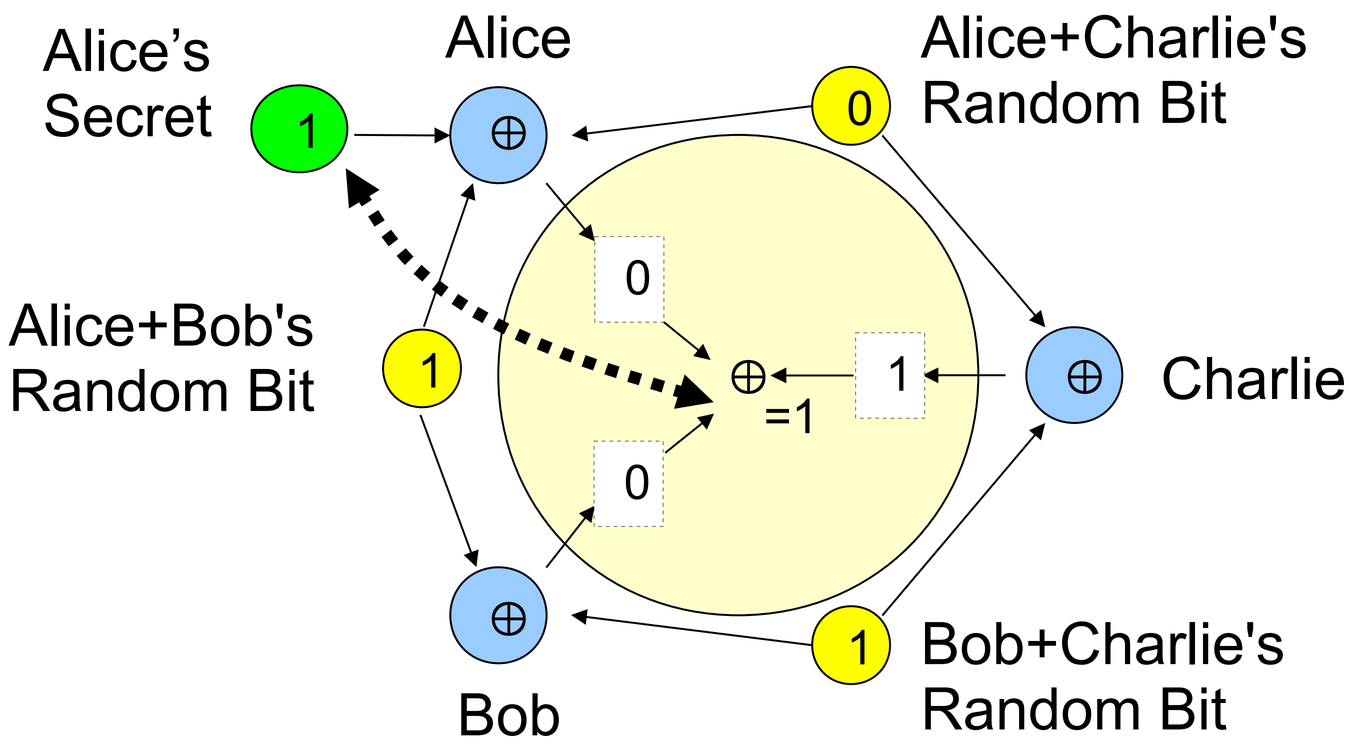
- Ex. 1: “Alice+Bob” sends a 1-bit secret to Charlie.



Dining Cryptographers (DC-nets)

Another fundamental Chaum invention from the 80s...

- Ex. 2: Homogeneous 3-member group anonymity



Dining Cryptographers (DC-nets)

Tantalizing theoretical properties

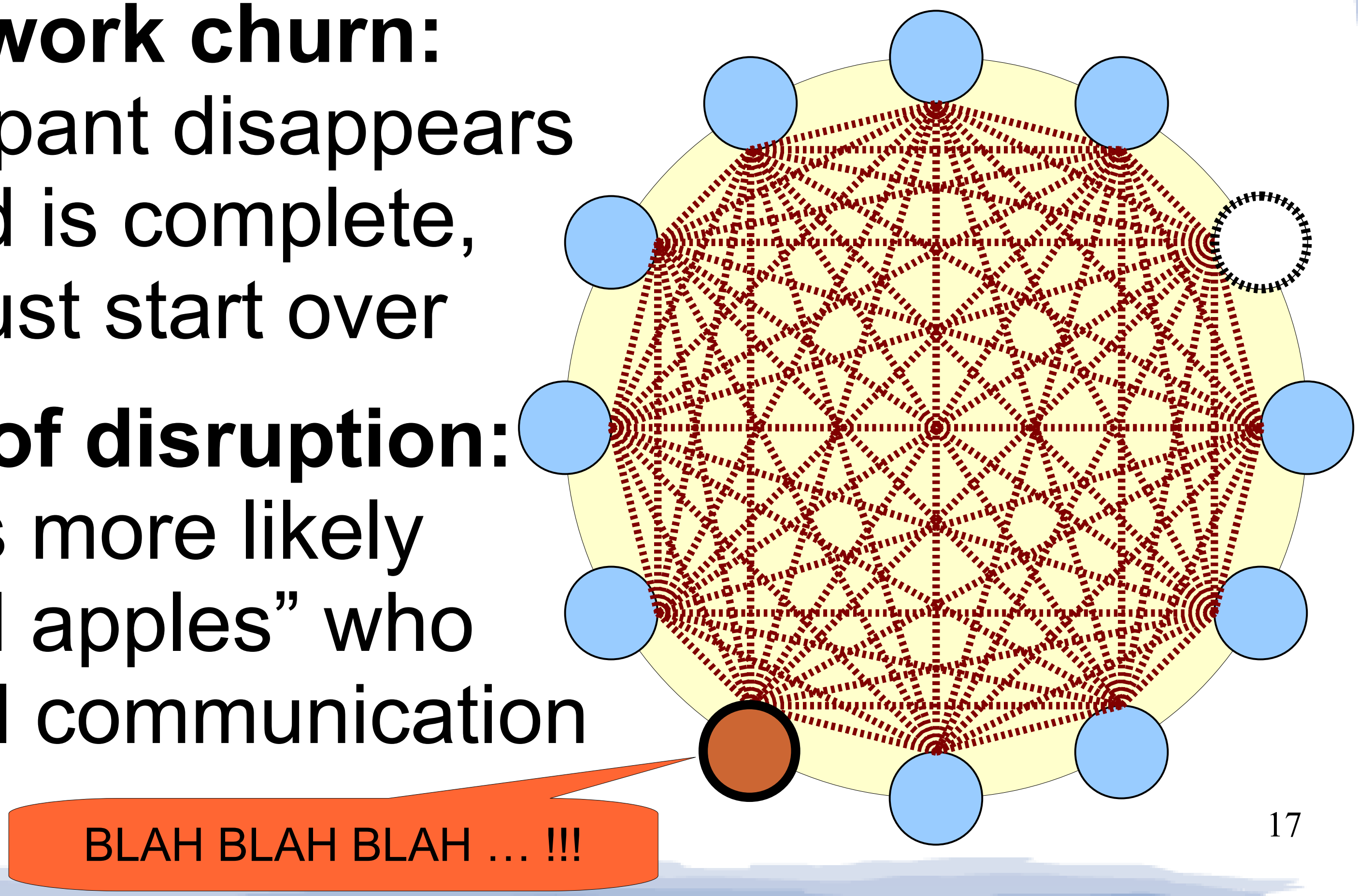
- Unconditional anonymity (if using “real” coins)
- Security against traffic analysis & collusion
    - Anonymity set = nodes *not* colluding against victim

Never successfully used in practical systems

- Easy to disrupt anonymously, no accountability
    - Malicious member can jam by sending random bits
- Not readily scalable to large groups
    - Especially with node failure, network churn

Why DC-nets Doesn't Scale

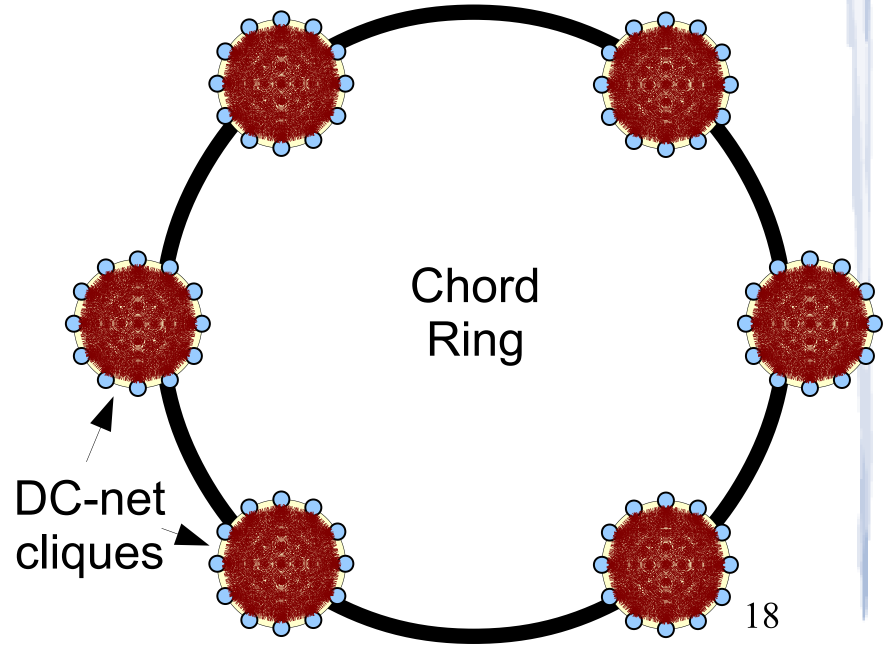
- **Computation cost:** *N* nodes each must flip, XOR together *N-1* shared coins per output bit
- **Typical network churn:** if *any* participant disappears before round is complete, *all* nodes must start over
- **Likelihood of disruption:** large groups more likely to have “bad apples” who jam some/all communication



Why Not Just Use **Small** Groups?

Exactly what **Herbivore** did [Sirer'04]

- Pioneering effort at making DC-nets practical
- Divides large network into many small *cliques*
  - If one gets jammed, join another
- Supports many users *total*, but guarantees *anonymity* only in *user's own clique*
  - Small anonymity sets, max 40 in experiments



The Dissent Project

(“DIning-cryptographers Shuffled-SEnd NeTwork”)

Fresh attempt to make DC-nets practical –
now 2nd year of 4-year DARPA-funded project

Goals:

- Scale to large *anonymity sets*, not just *networks*
- Add *accountability* to limit anonymous abuse
- Tolerate both *normal churn* and *disruption*
- *Quantifiable* security against *strong adversaries*

Selected Dissent Papers
(available at http://dedis.cs.yale.edu/2010/anon/)

Covered in part by this talk:

- “Dissent: Accountable Group Anonymity” [CCS'10]
- “Dissent in Numbers: Making Strong Anonymity Scale” [OSDI'12]
- “Dining in the Sunshine: Verifiable Anonymous Communication” (draft)
- “Scavenging for Anonymity with BlogDrop” (abstract) [ProvPriv'12]

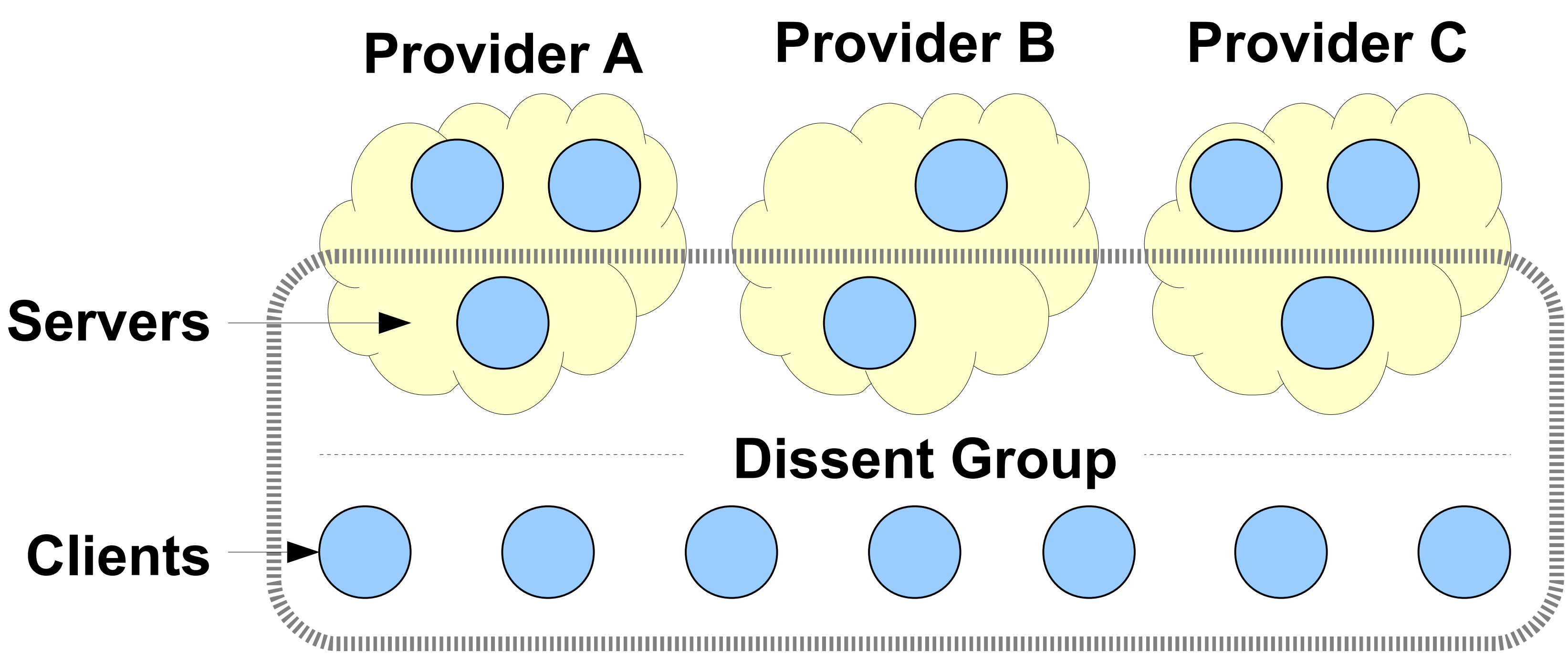
Talk Outline

- ✓ Online anonymity: state-of-the-art, weaknesses
- ✓ Dining cryptographers: a cool, useless toy?
- **Making DC-nets scale to “real” systems**
- Accountability – in many flavors
- Anonymity scavenging and intersection attacks
- Conclusion

Multi-Provider Cloud Model

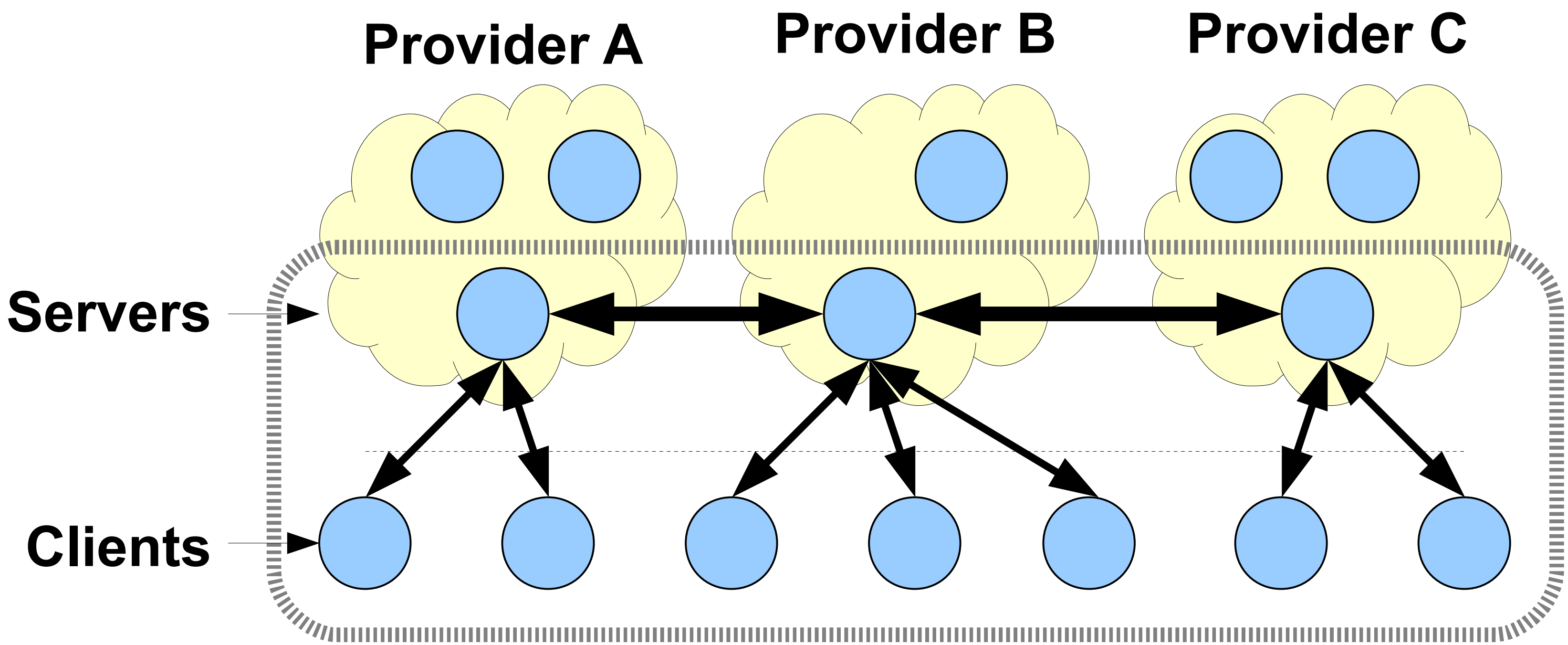
Dissent group (anonymity set) consists of:

- Large-*ish* number of unreliable *clients* (users)
- A few *servers*, each from a reputable *provider*



Communication Structure

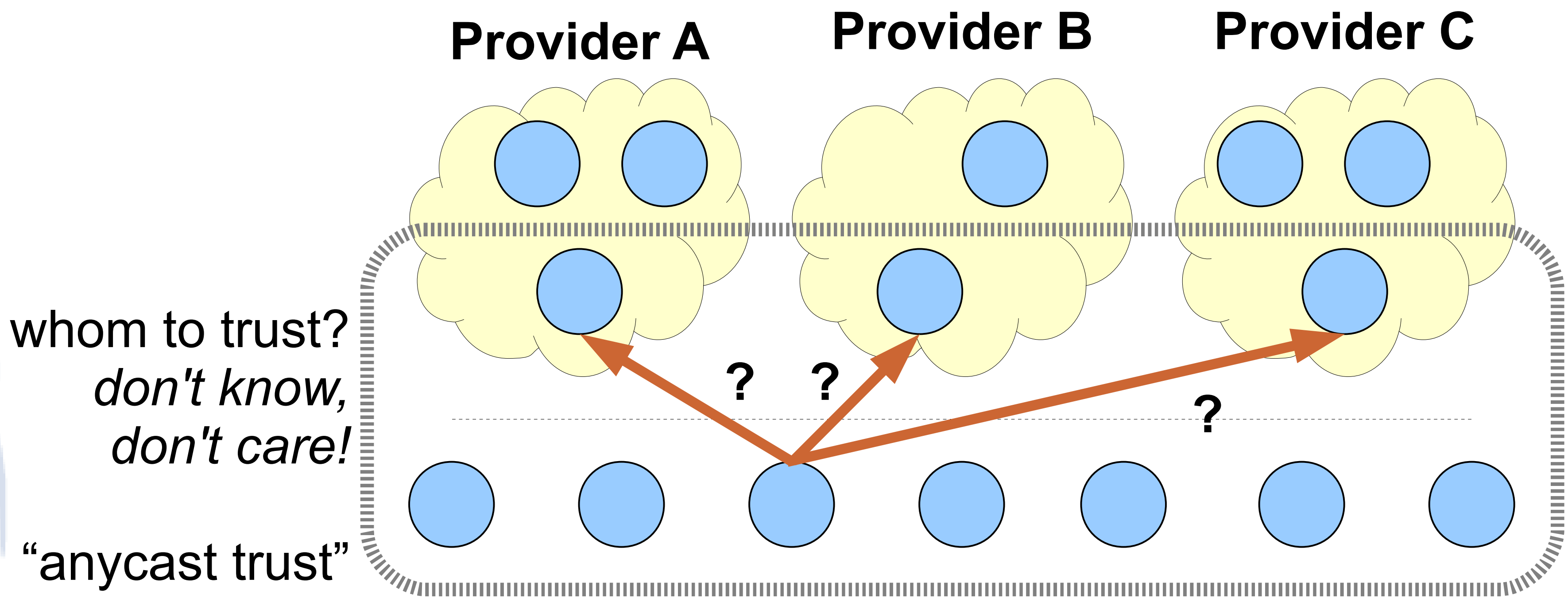
- Each client connects with *one* upstream server
- All servers coordinate directly with each other
  - Best if servers are “nearby” – low delay, high BW



"Anytrust" Assumption

Clients do *not* trust upstream (or any one) server

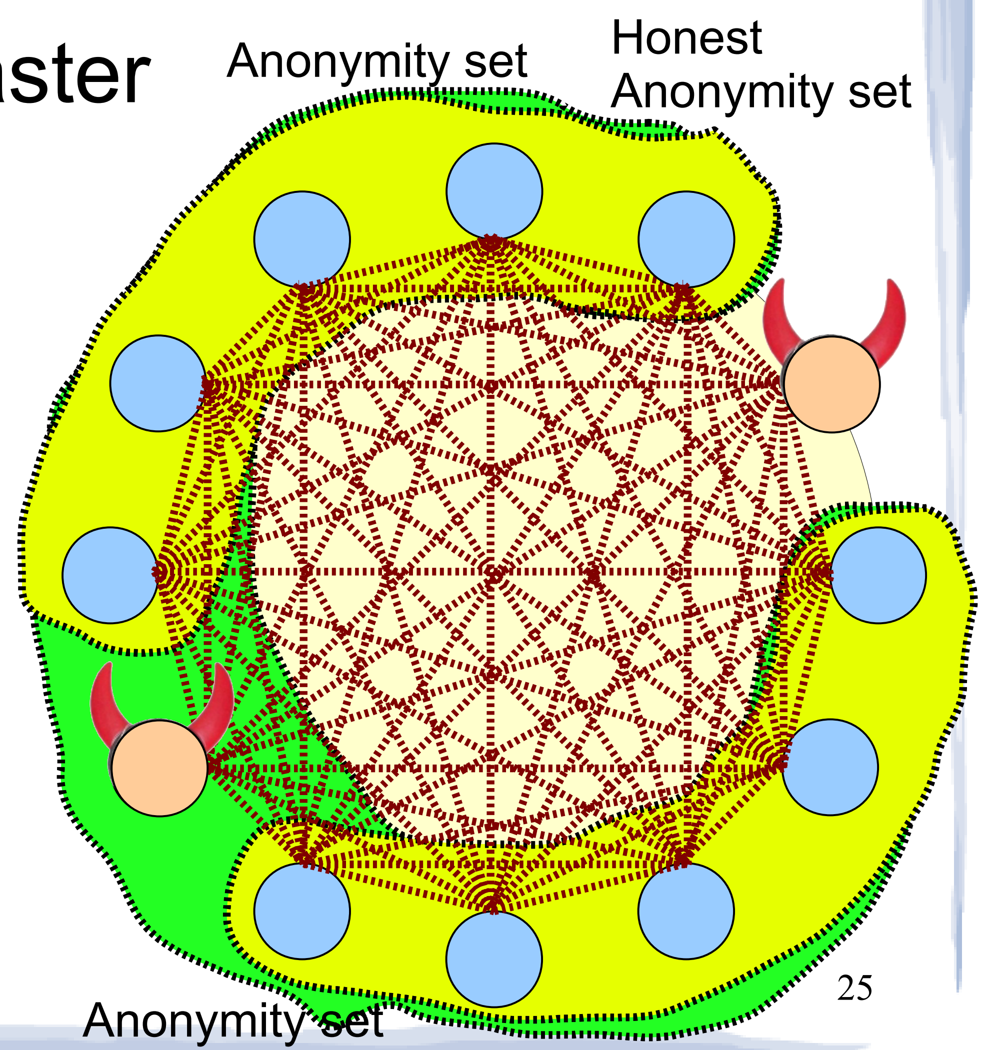
- Trust only that *some* server – *any* server – will not collude with *all* others against client



Sparse Coin-Sharing in DC-nets

Every pair of nodes needn't share coins/keys...

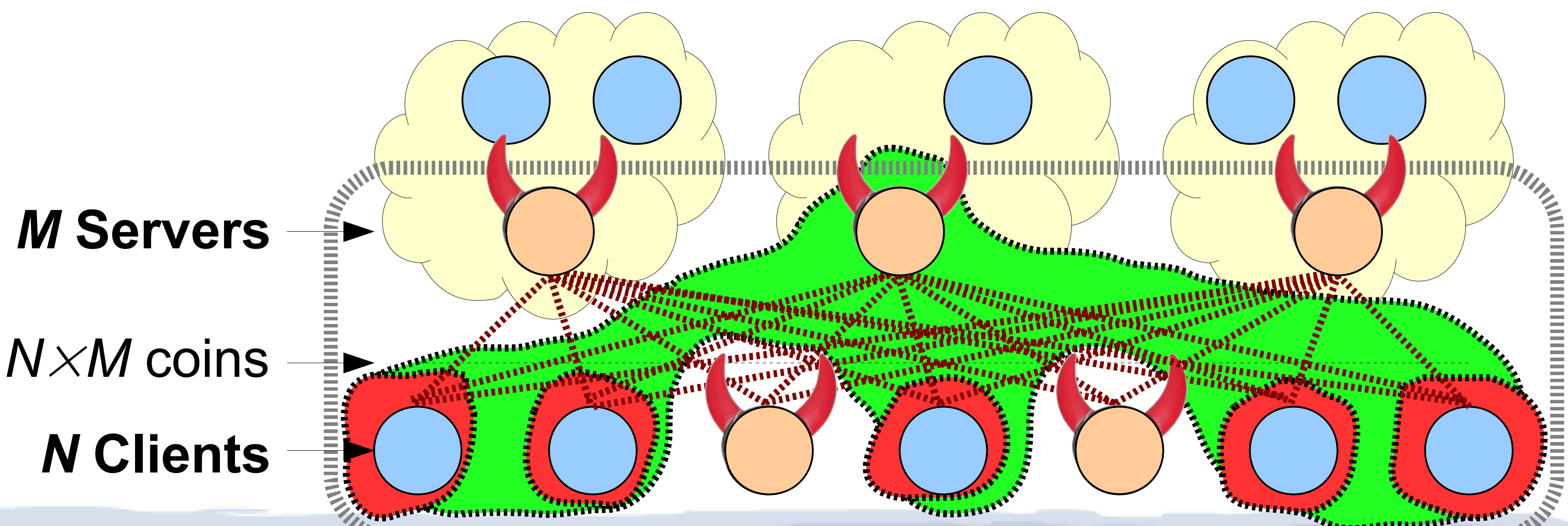
- Fewer shared coins → faster
- Reduces anonymity *if, and only if,* attack nodes split key-sharing graph
- Example: “ring” graph
  - OK if only 1 attack node
  - Bad if 2 or more collude



Dissent's Coin-Sharing Structure

Each client shares coins with *every server*

- Provided *there exists* one honest server, that server shares coins with *all* honest clients
  - *Optimal* anonymity – *if* assumption holds :)
  - *No* anonymity – if it doesn't :(



Why Client/Server Coin-Sharing?

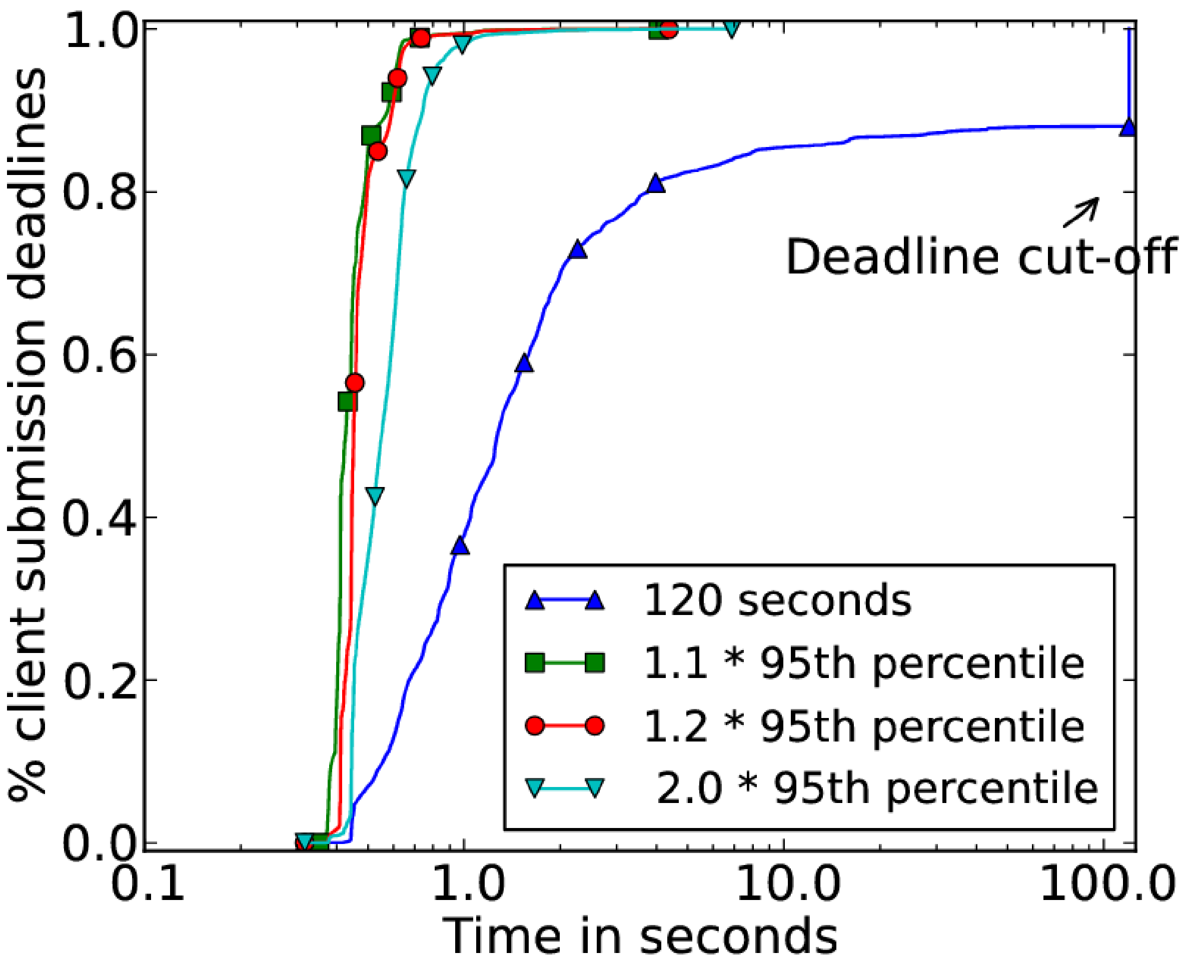
Two key benefits:

1. **Reduce computation load on clients**
   - Compute only $M \ll N$ pseudo-random coins per bit of anonymous transmission bandwidth
2. **Servers can adapt to slow or offline clients**
   - Client ciphertexts depend *only* on servers, *not* on which other clients are online in this round
   - Servers collect client ciphertexts until a *deadline*, then compute *their* ciphertexts based on results
   - *No* wait for slowest client, or restart on disconnect

Group of 500 PlanetLab Clients

Without deadline, 50% of rounds take over 1 sec, 20% over 5 sec, 15% timeout

With deadline, 90% of rounds take < 0.4 sec, *no timeouts*

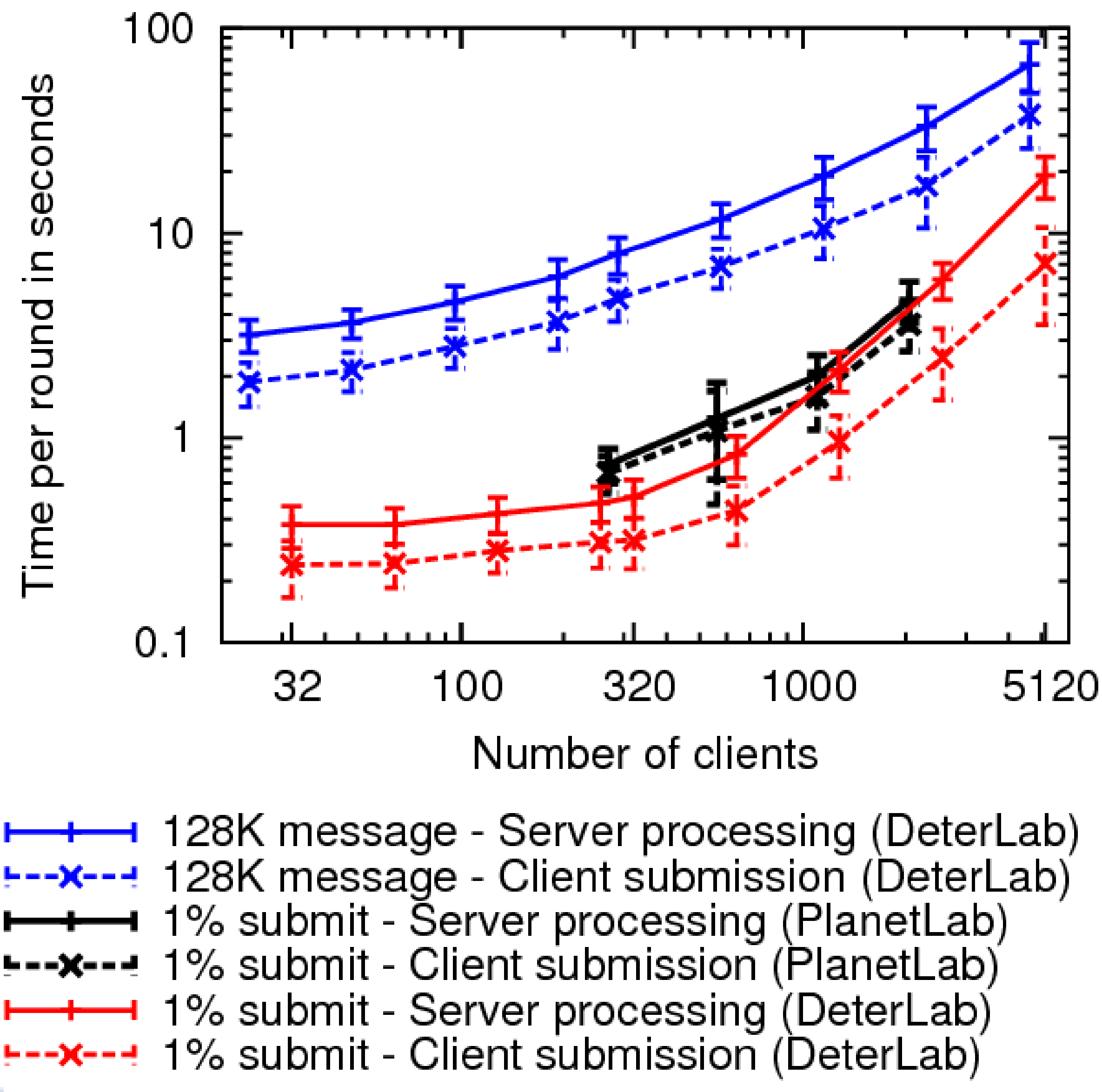


Scaling to Thousands of Clients

Anonymity sets **100× larger** than previously demonstrated

- Herbivore, Dissent v1: ~40 clients

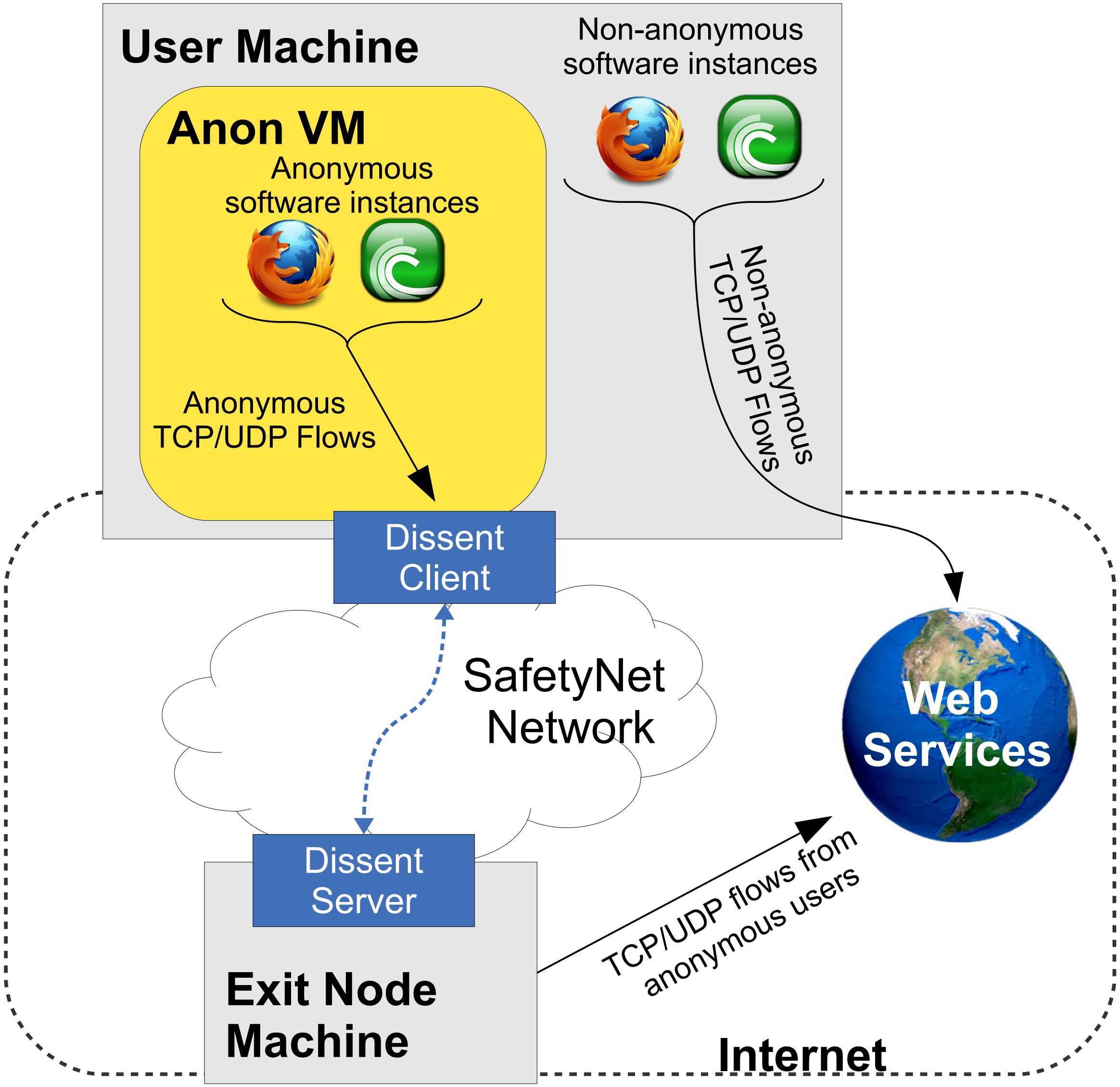
Sub-second latencies in 1000-client groups



WiNon: Web Browsing via Dissent

Fast enough for interactive use in small local-area groups, e.g., WiFi

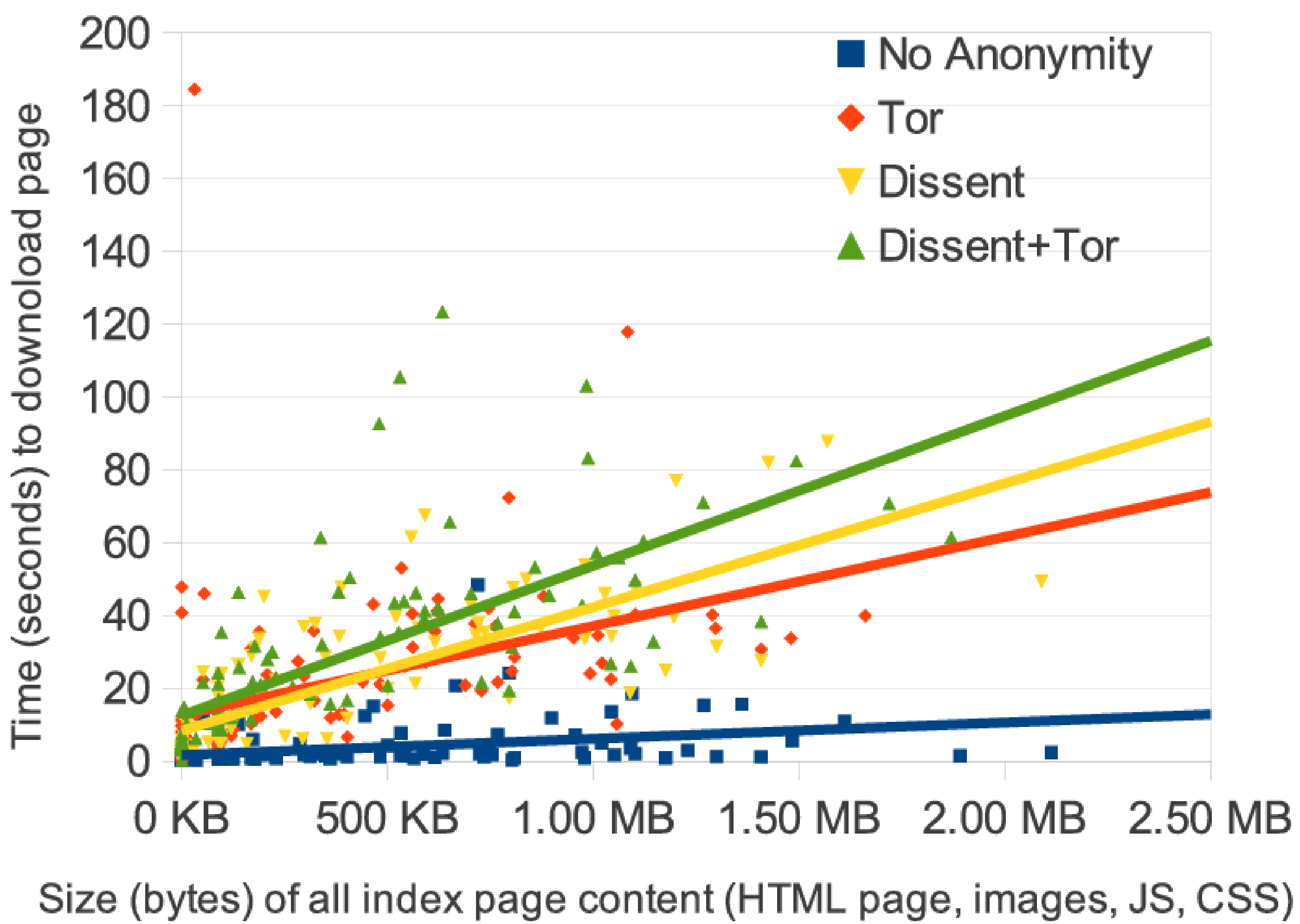
"Strong, small" anonymity sets complementing **"large, weak"** sets Tor offers



WiNon Browsing Latency

5 servers,
24 clients,
WiFi LAN
→ usability
comparable
to Tor

Illustrative only –
“apples-to-
oranges”

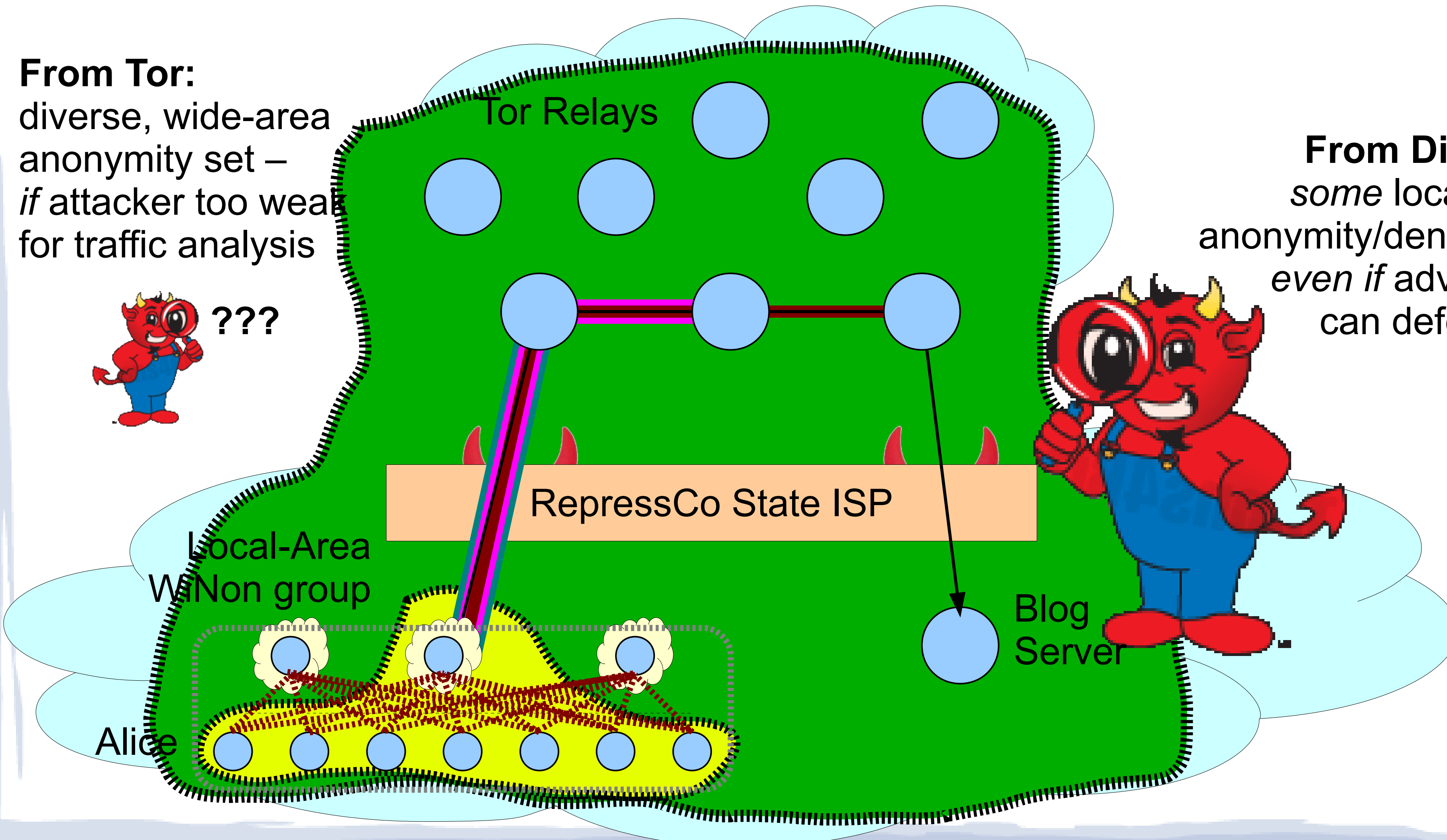


Why is Dissent+Tor Interesting?

Defend against “Little Brother” *and* “Big Brother”

From Tor:
diverse, wide-area anonymity set –
if attacker too weak for traffic analysis

From Dissent:
some local-area anonymity/deniability,
even if adversary can defeat Tor

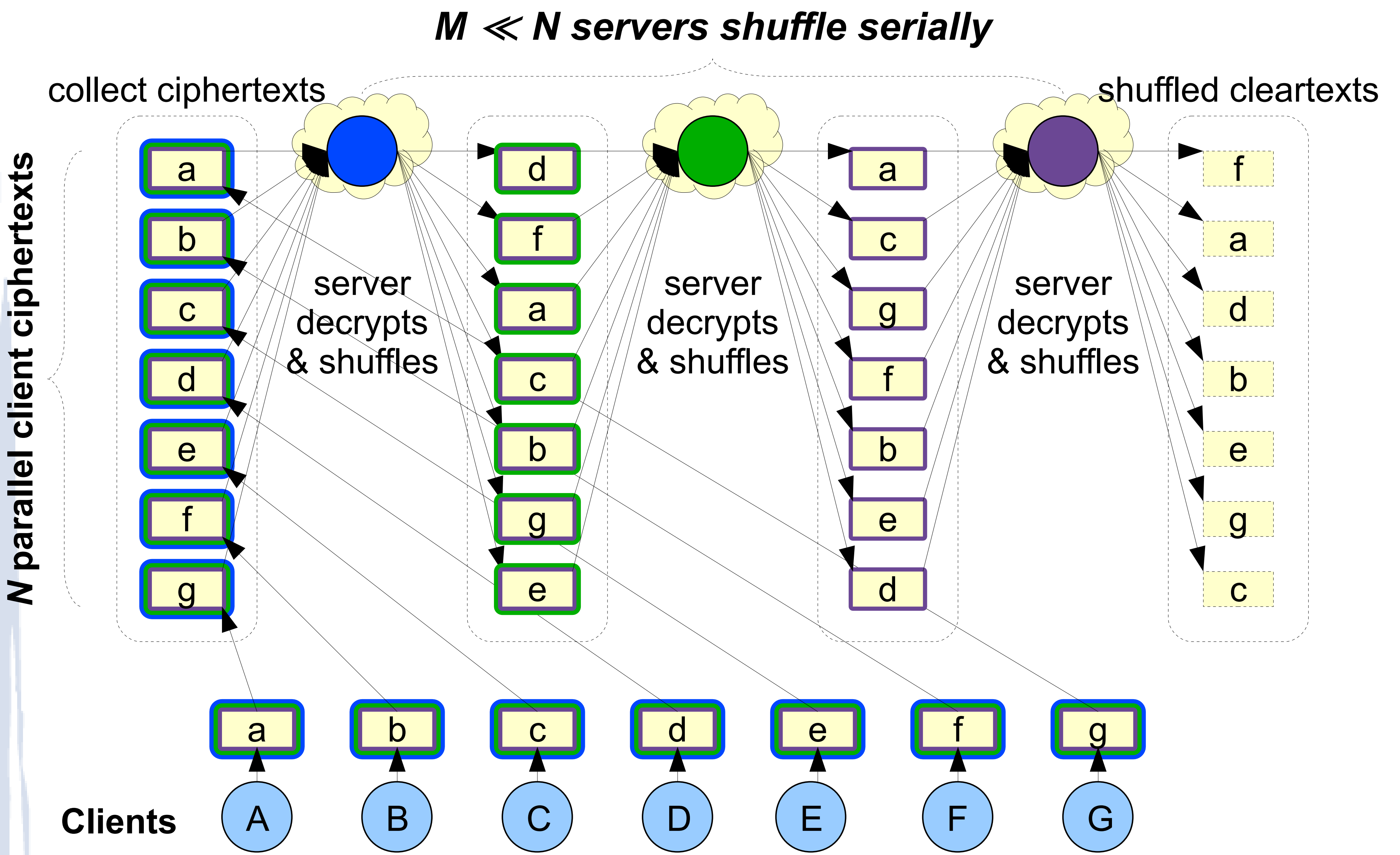


Scheduling DC-net Transmissions

How does each client know *when* to transmit?

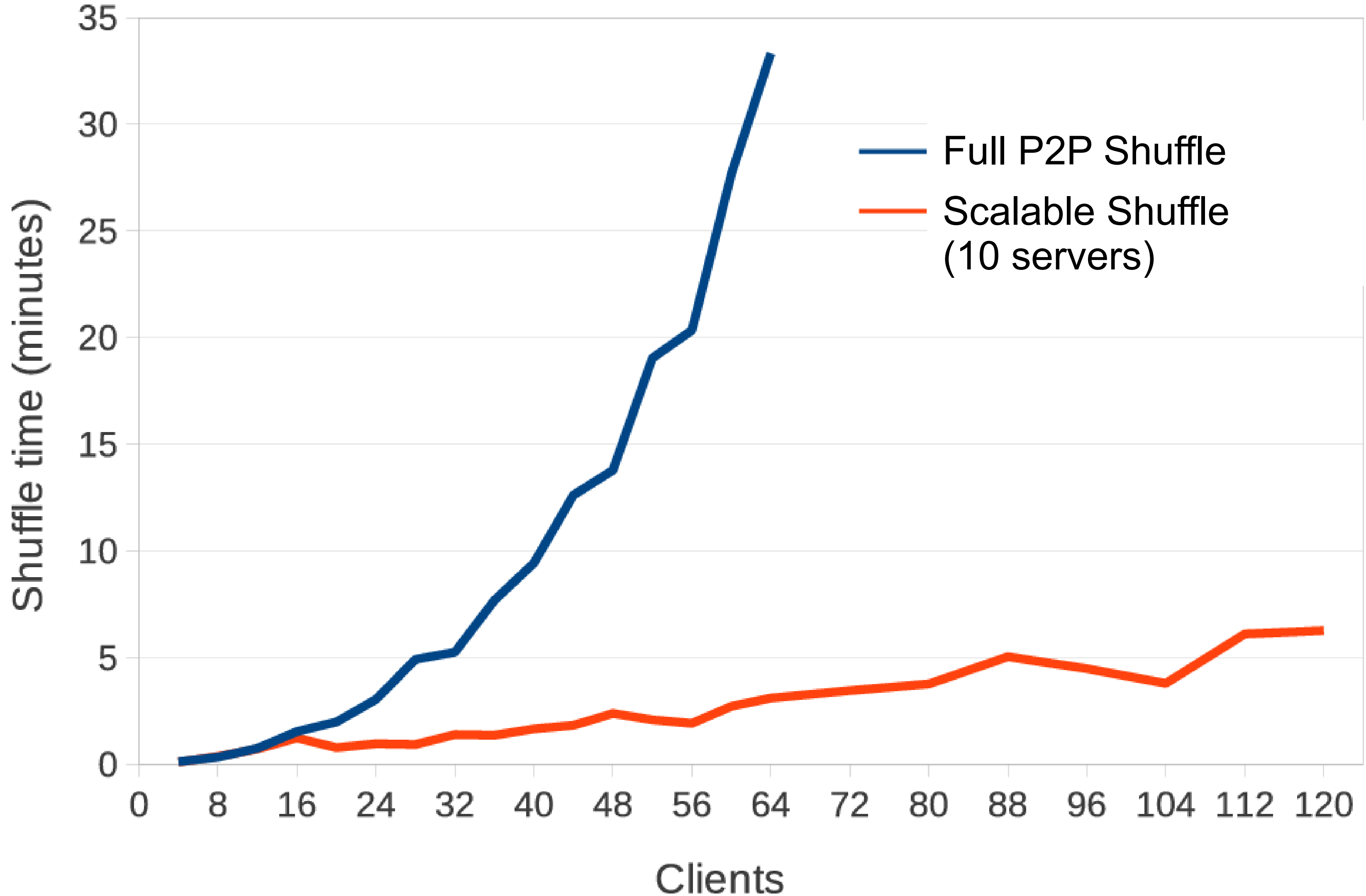
- Like airwaves, DC-nets messages get garbled if more than one client transmits at once
- Dissent uses **verifiable shuffles** [Neff'01] to form schedule of anonymous transmission slots
  - See papers for details
- Scalable shuffling in Dissent *also* relies on multi-provider cloud model, anytrust assumption

Scalable Shuffling at a Glance



clients choose *pseudonym keys*, onion-encrypt for *all servers*

Scalable Shuffle Comparison



Talk Outline

- ✓ Online anonymity: state-of-the-art, weaknesses
- ✓ Dining cryptographers: a cool, useless toy?
- ✓ Making DC-nets scale to “real” systems
- **Accountability – in many flavors**
- Anonymity scavenging and intersection attacks
- Conclusion

Accountable Anonymity

Accountability can mean many things

- “Accountability & Deterrence” [Feigenbaum'11]

In Dissent, accountability means:

- **Disruption-resistance:** group can trace, expel any member attempting to jam communication
- **Proportionality:** each member gets *exactly* 1 bandwidth share, 1 vote, 1 pseudonym, etc.

In Dissent, “accountability” does **not** mean “de-anonymize people who say things I don't like”

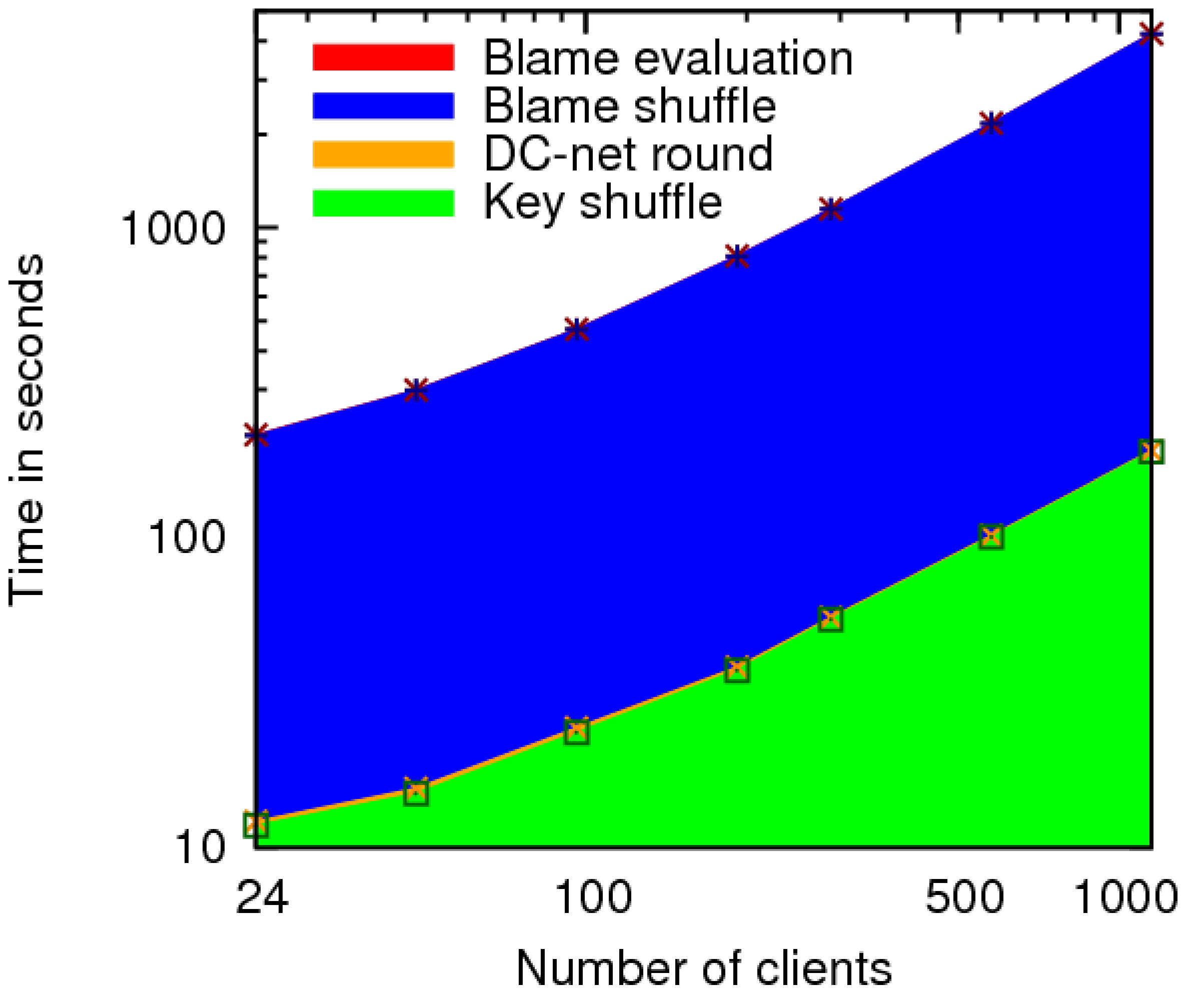
Jam-Proofing DC-nets: 4 Ways

1. **Herbivore:** flee to new group if jammed
   - Must keep groups small to minimize jamming risk
   - Could land in a group that's not jammed because it's *completely* owned by adversary! [Borisov'07]
2. **Dissent v1** [CCS'10]:
   use verifiable shuffle to distribute *assignments* with ciphertext hashes before each round
   - Makes jamming easy to identify and trace
   - Requires slow, expensive shuffle for *every round*

Jam-Proofing DC-nets: 4 Ways

3. **Dissent** [OSDI '12]:
   retroactive disruption-tracing "blame" protocol
   - Victim finds a "witness a bit" attacker flipped 0→1; broadcasts pointer to witness bit in "blame shuffle"
   - Nodes reveal all coins contributing to witness bit, find source of "odd-one-out" that flipped it to 1
   - **Upsides:** minimal overhead when no jamming, $\geq \frac{1}{2}$ chance of catching jammer in each round
   - **Downsides:** complex, slow due to blame shuffle; attacker with *f* nodes can stop progress for *f* rounds

Round Latency Breakdown

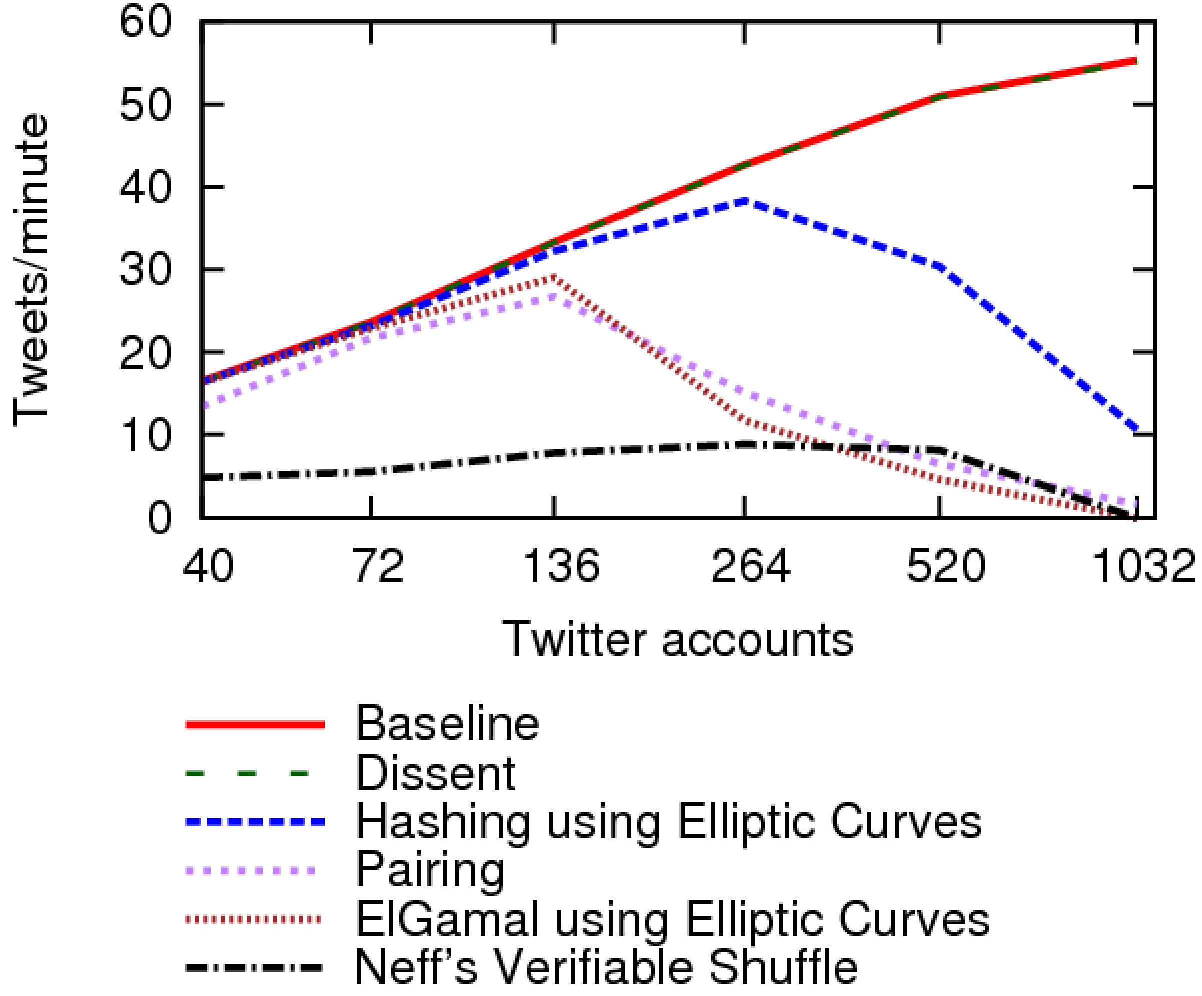


Jam-Proofing DC-nets: 4 Ways

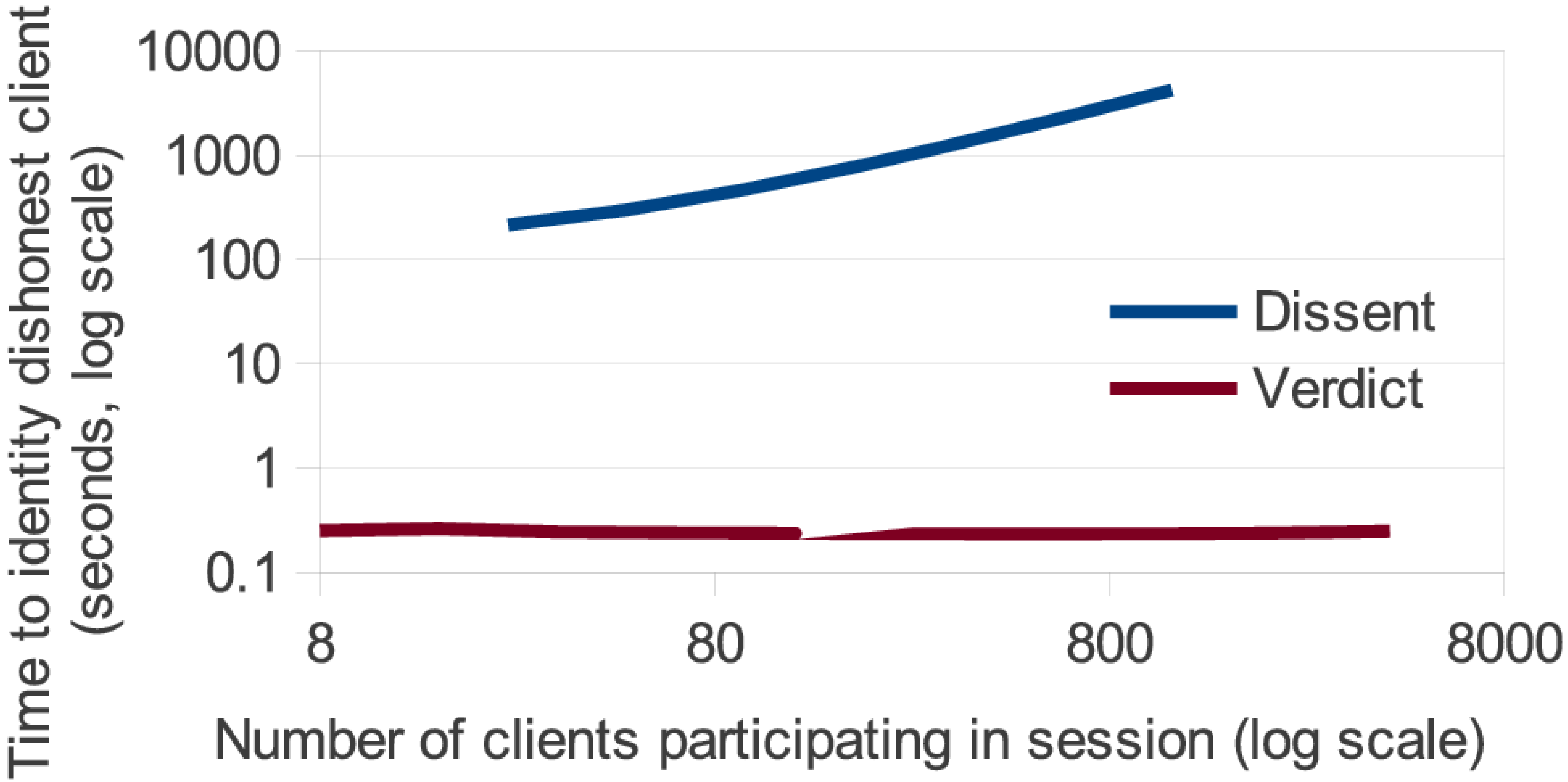
4. **“Dining in the Sunshine”** [see Dissent page]: *proactive* verifiability via cryptographic proofs

- Clients encode messages in algebraic groups, show correct construction via discrete log proofs
- 3 schemes: pairing-based [Golle/Juels'04], plus faster schemes usable with Schnorr or EC groups
- **Upsides:** disruptors cannot jam communication; ciphertexts can be build offline and “dropped off”; potential asymptotic benefits in large groups
- **Downsides:** complex, slow and CPU-intensive, especially in small groups, due to group arithmetic

Retroactive vs Proactive: The Bad News



Retroactive vs Proactive: The Good News



Talk Outline

- ✔ Online anonymity: state-of-the-art, weaknesses
- ✔ Dining cryptographers: a cool, useless toy?
- ✔ Making DC-nets scale to “real” systems
- ✔ Accountability – in many flavors
- **Anonymity scavenging, intersection attacks**
- Conclusion

Anonymity Scavenging

Bob in Dictatopia posts to dissident blog each day

- Tolerates latency, needs large anonymity set, *even under traffic analysis & intersection attack:* risks jail time if identity discovered

Alice wants to microblog casually on blocked sites

- Needs low latency, but low security sensitivity: “everyone does it” → unlikely to be prosecuted

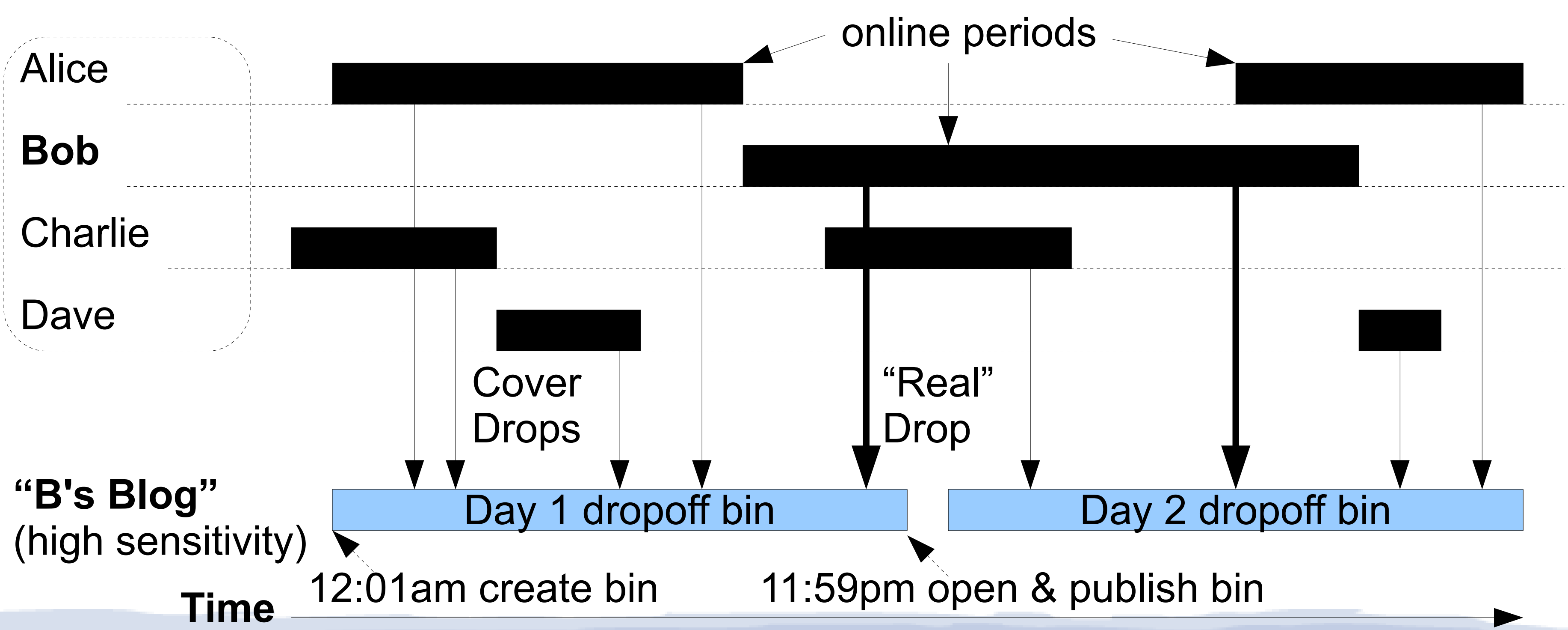
How can we meet both Alice's and Bob's needs?

Better, how can Alice (unwittingly) help Bob?

Dropoff Communication Model

Many users come online per day *at different times*

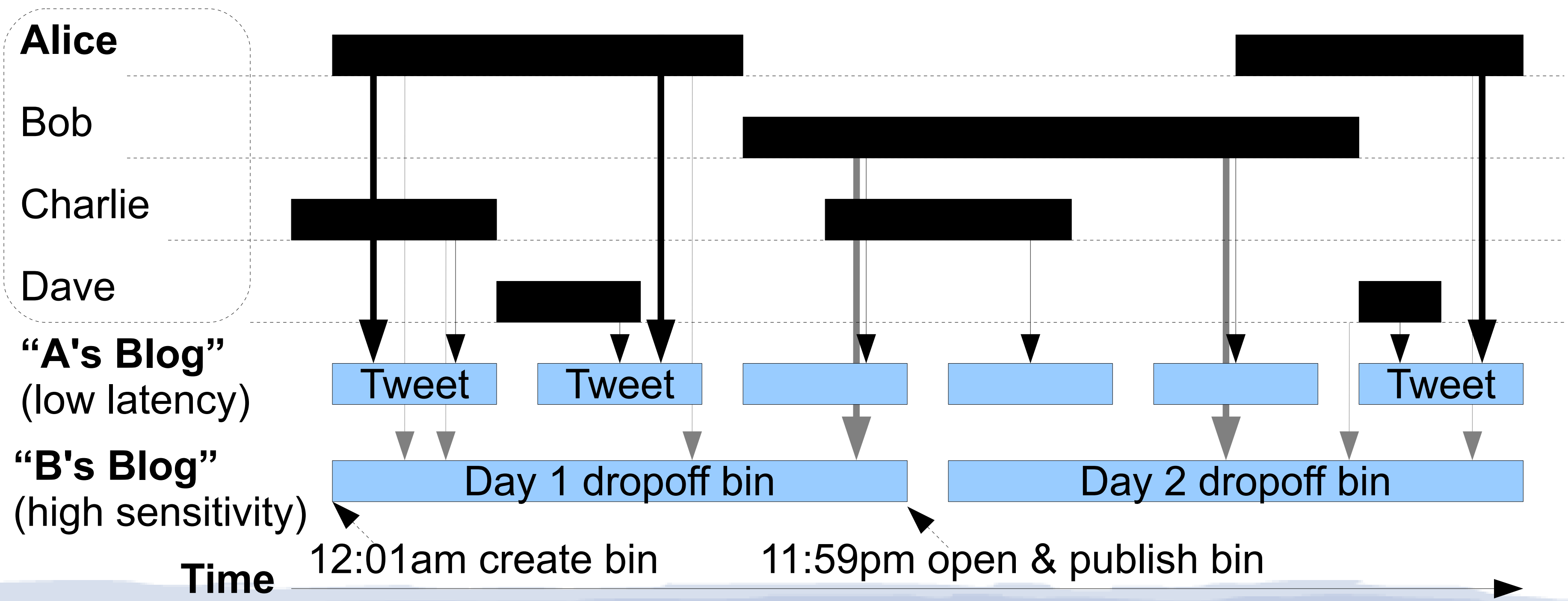
- Drop DC-nets ciphertext into Bob's dropoff bin
- Servers “open” bin, publish contents at midnight



Scavenging from Diverse Users

Alice frequently microblogs low-sensitivity chitchat

- Gets lower anonymity against traffic analysis
- But contributes to Bob's large anonymity set



Work-in-Progress

Builds on, depends on verifiable DC-nets

- Dropped-off ciphertexts *must* be verifiable

Extend DC-net traffic analysis security “over time”

- Can we get 50,000-user anonymity in a day?

Under *some* conditions we *think* we can address long-term intersection attacks this way too

- Becomes “real-time” system for sensitive users
- Bob can avoid leaking identity even long-term – *if* he (and others) show up *at least once per day*

Talk Outline

- ✓ Online anonymity: state-of-the-art, weaknesses
- ✓ Dining cryptographers: a cool, useless toy?
- ✓ Making DC-nets scale to “real” systems
- ✓ Accountability – in many flavors
- ✓ Anonymity scavenging, intersection attacks
- **Conclusion**

Summary and Current Status

What we've done so far:

- Made DC-nets scale to 5000+ node groups
- Wide-area microblogging, local-area browsing uses
- Developed 3 new approaches to accountability

In-progress:

- Proactively verifiable DC-nets (mostly done)
- Scavenging large anonymity sets across time
- Protection against long-term intersection attacks

- Very experimental code available on GitHub

Conclusion

The Dissent project asks:
can we use dining cryptographers as a foundation to get stronger, quantifiable anonymity in practice?

- Anonymity: even against traffic analysis
- Accountability: resistant to sybil attacks, disruption
- Eventually: resistance to intersection attacks??

We're optimistic, but many open questions!

http://dedis.cs.yale.edu/2010/anon/