# MorphIT: Reconciling Anonymity with Internet Performance Transparency

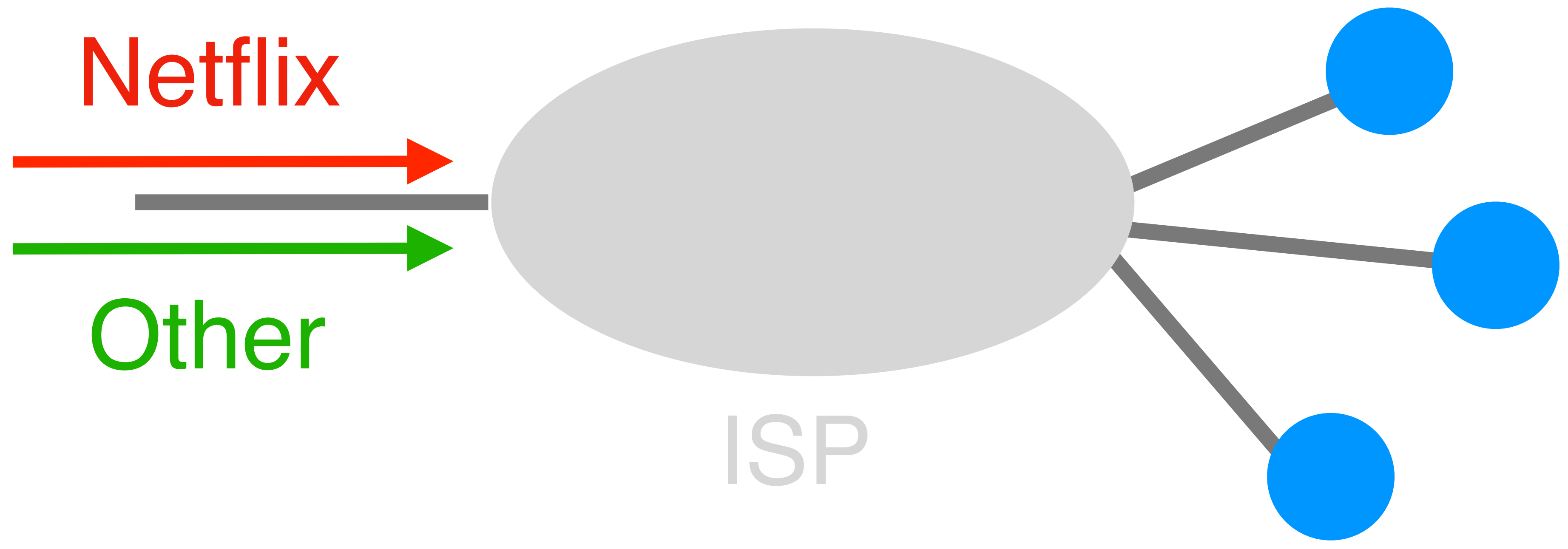
G. Fragkouli, K. Argyraki, B. Ford

IETF 109, 20/11/2020

EPFL

## Service Level Agreements

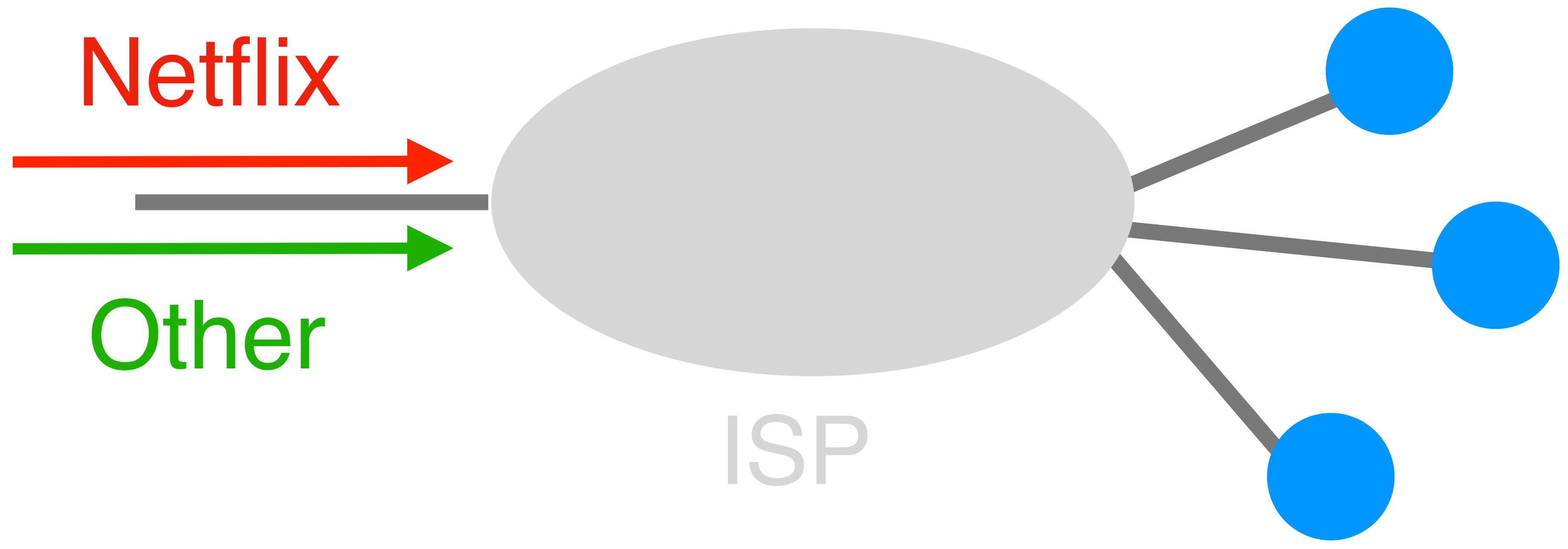
## Neutrality



# Transparency

## Service Level Agreements

## Neutrality



## Transparency

Service Level Agreements

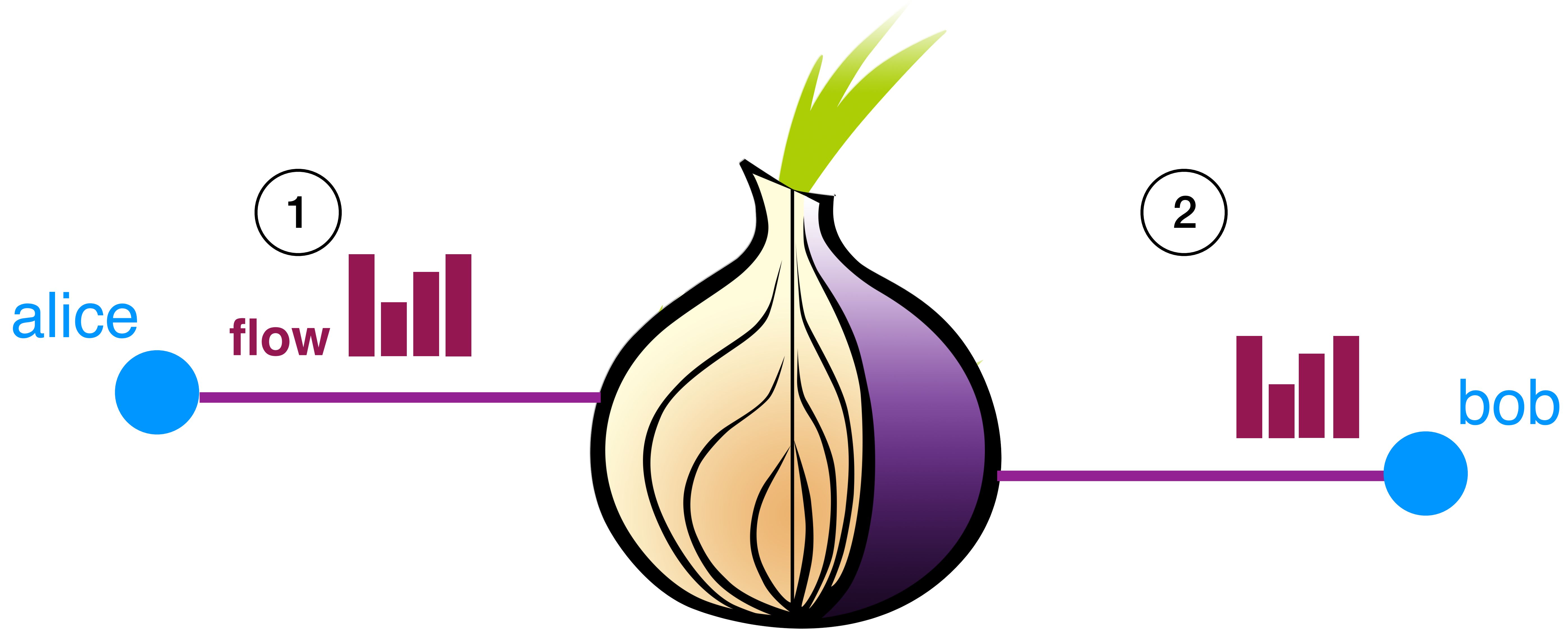
Neutrality

## Anonymity

Tor-like overlays

# Anonymity in Tor

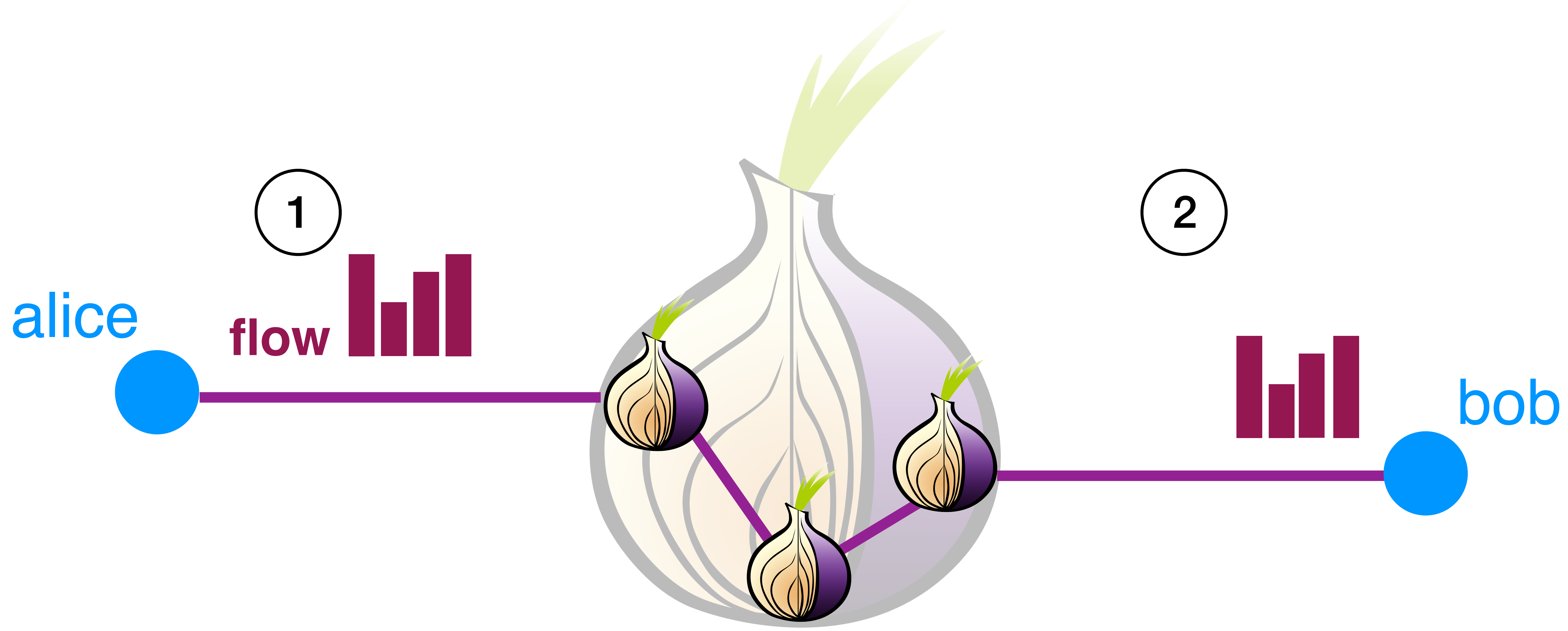
- Adversary cannot link sender to receiver



**Breaks against a global passive adversary**

# Anonymity in Tor

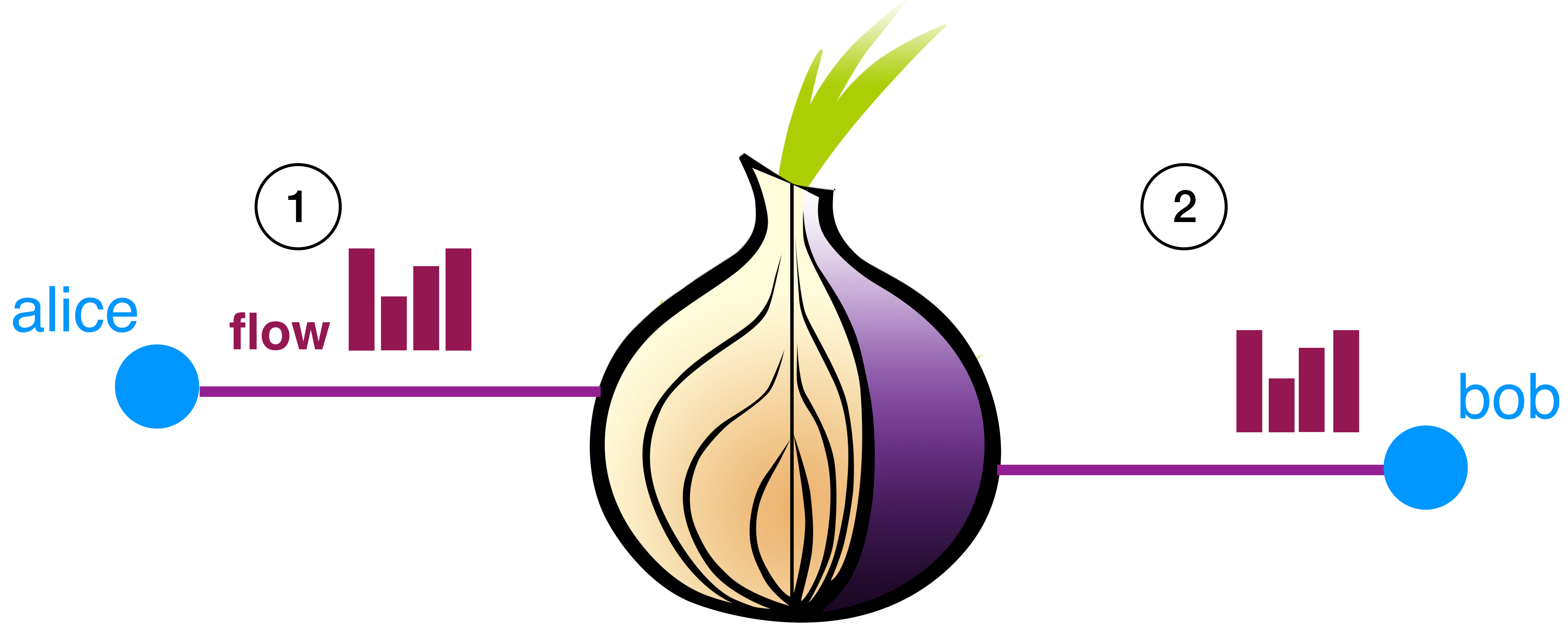
- Adversary cannot link sender to receiver



**Breaks against a global passive adversary**

# Anonymity in Tor

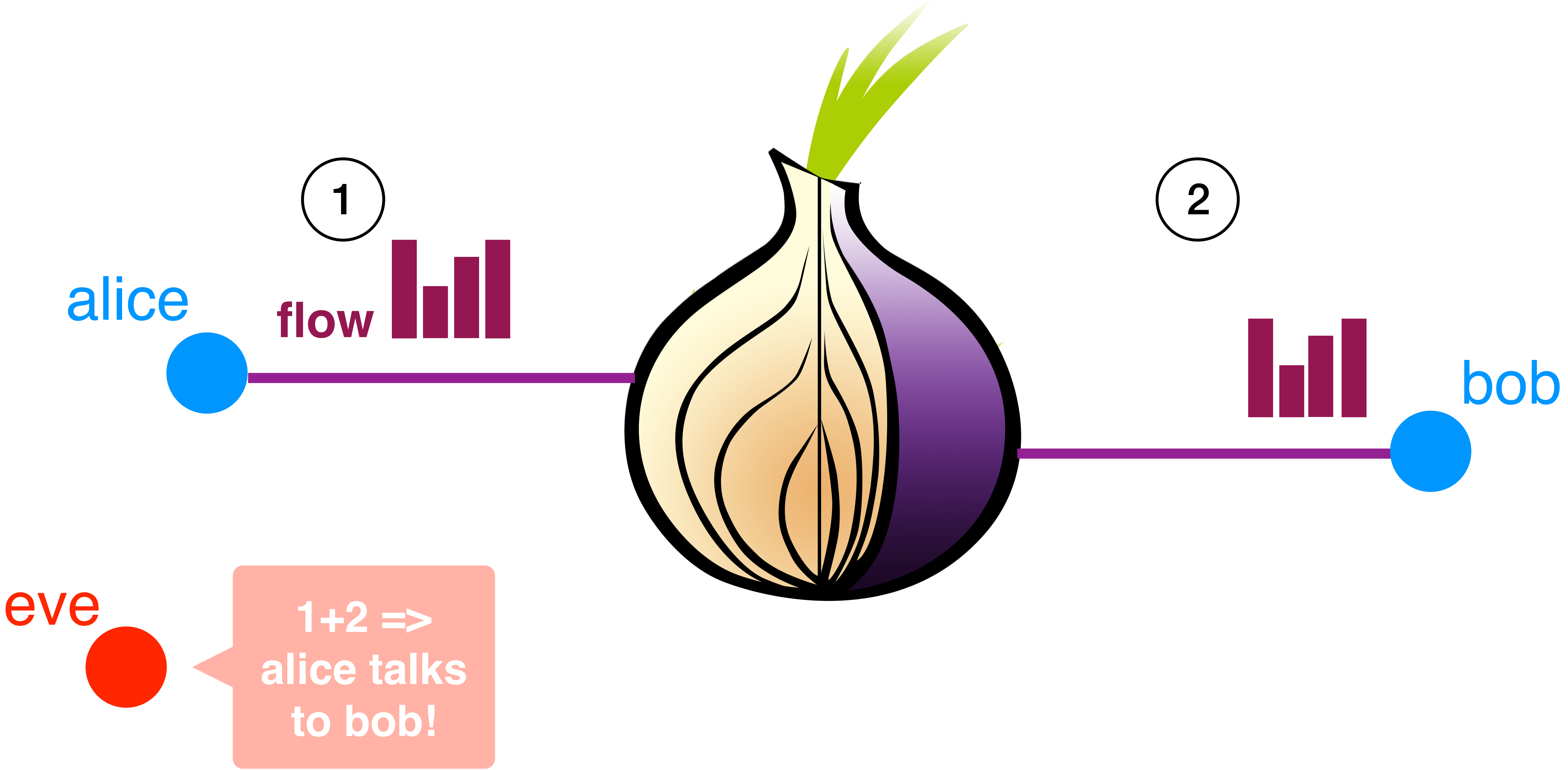
- Adversary cannot link sender to receiver



**Breaks against a global passive adversary**

# Anonymity in Tor

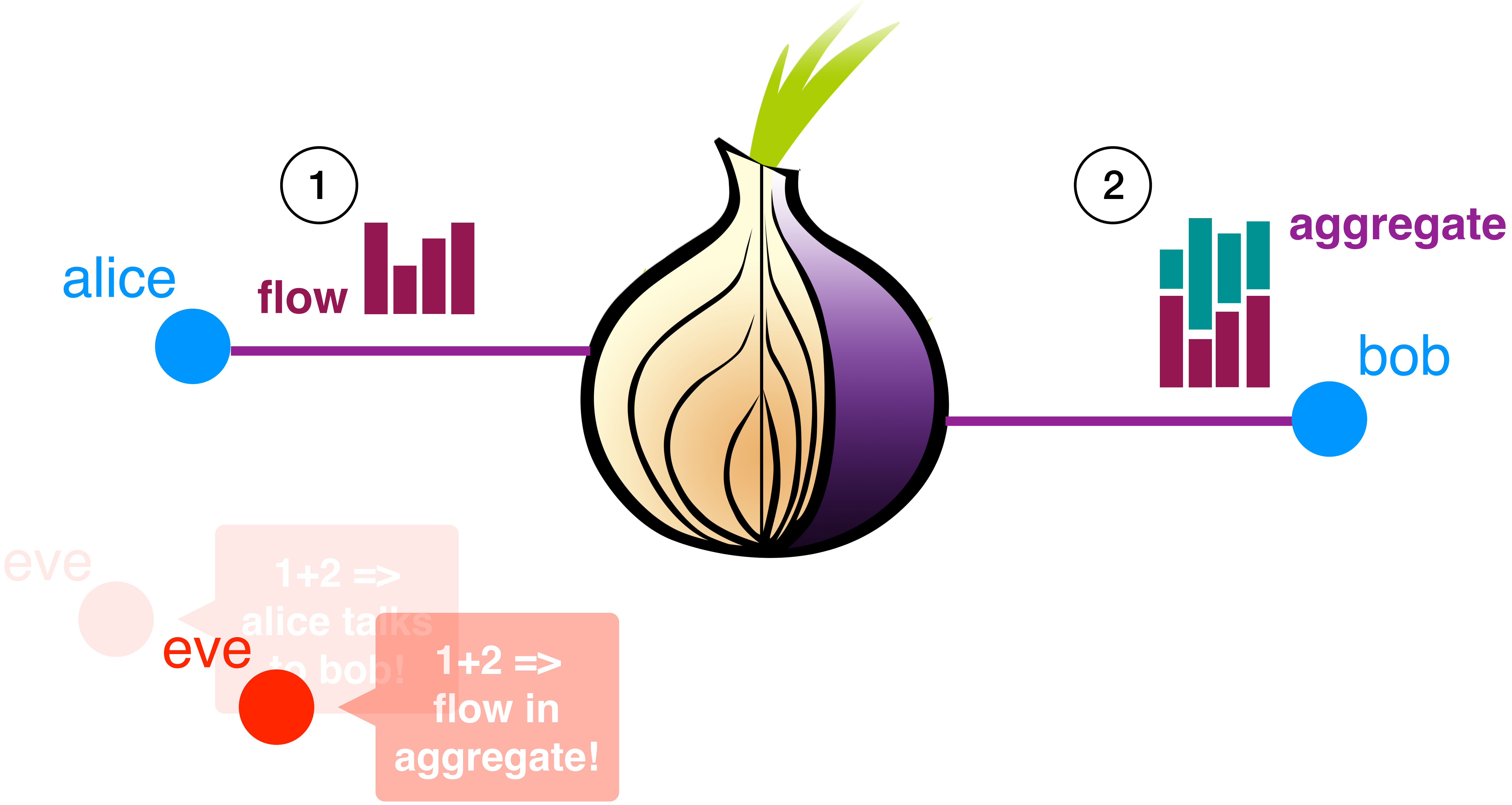
- Adversary cannot link sender to receiver



**Breaks against a global passive adversary**

# Anonymity in Tor

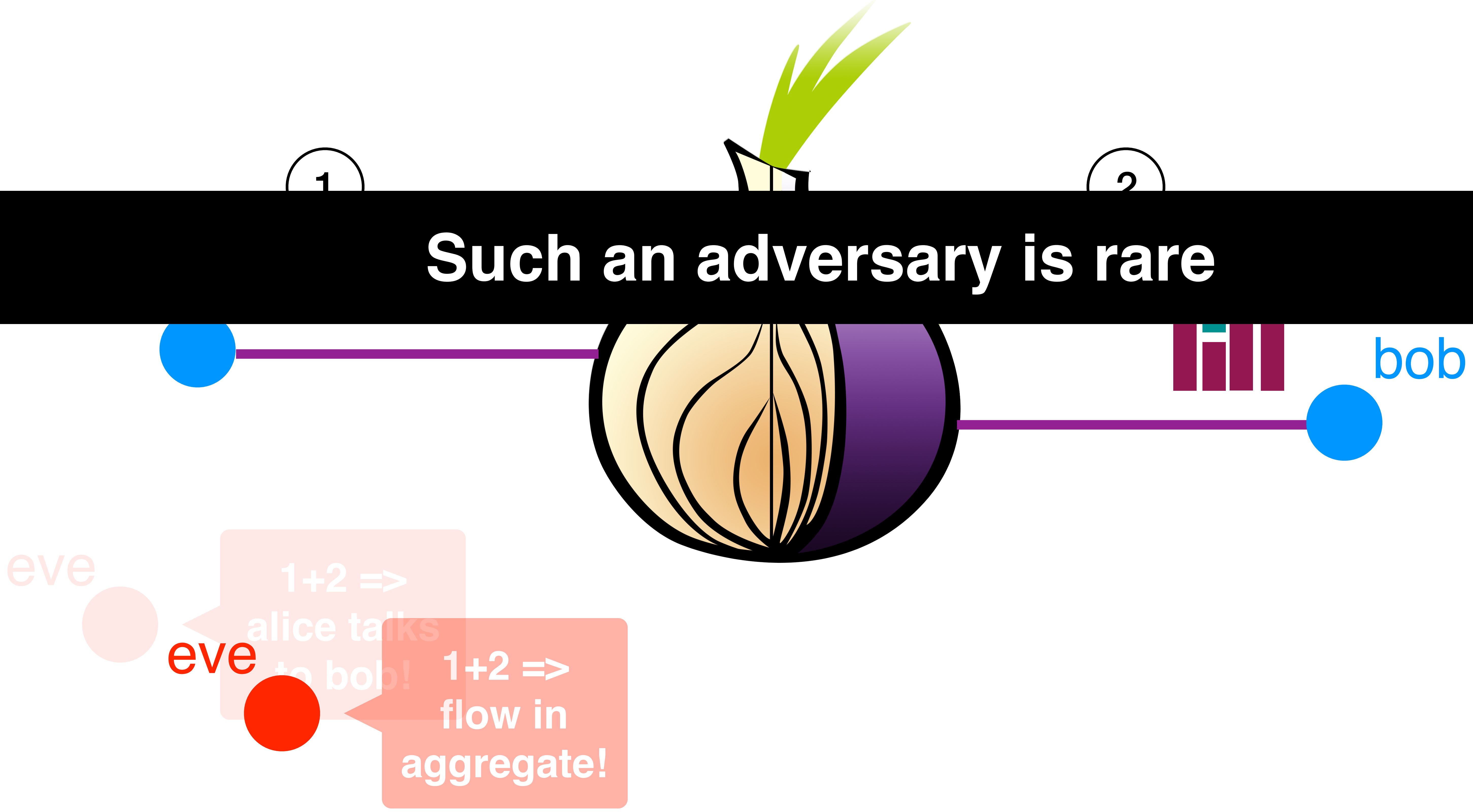
- Adversary cannot link sender to receiver



**Breaks against a global passive adversary**

# Anonymity in Tor

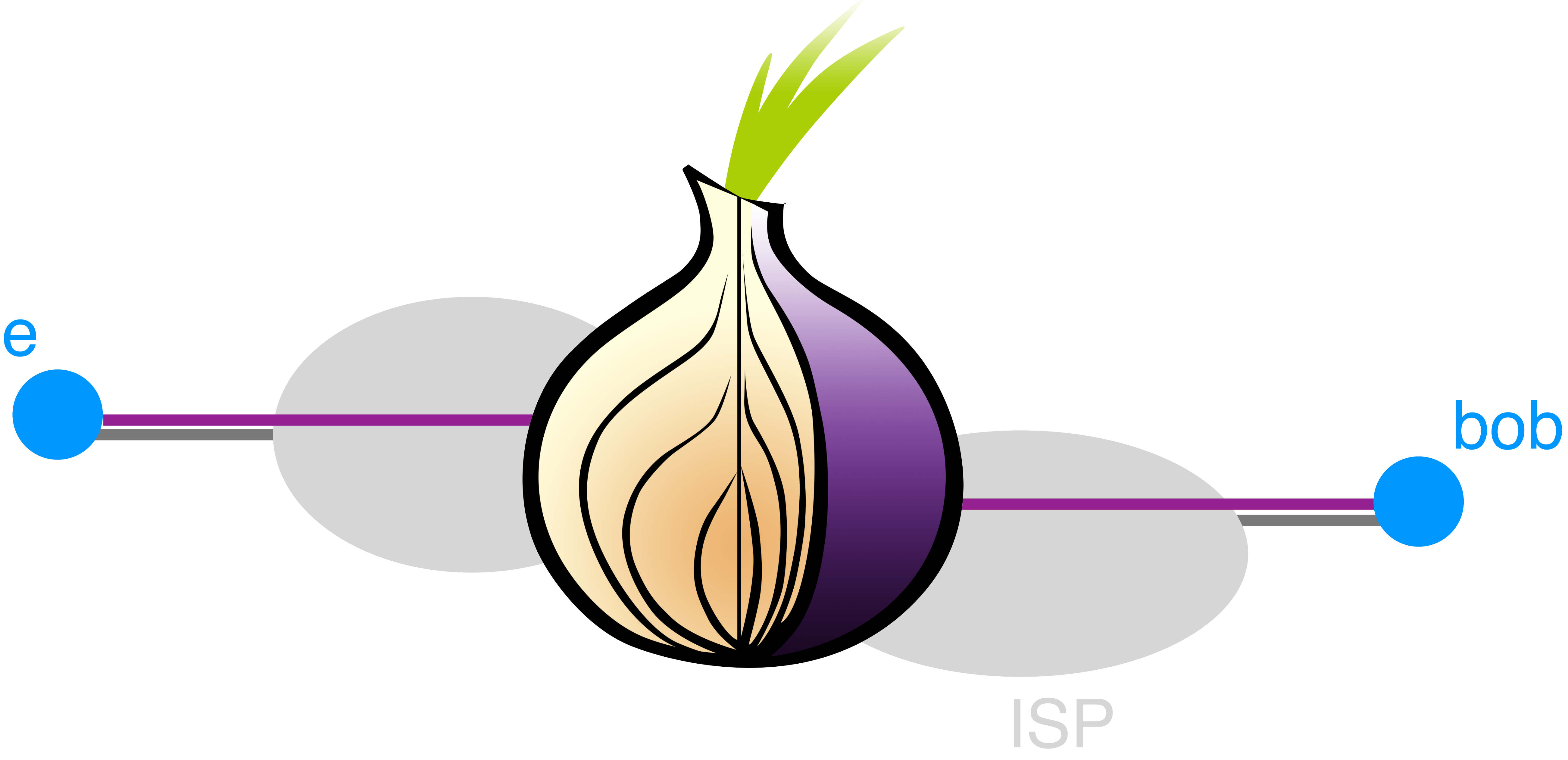
- Adversary cannot link sender to receiver



**Such an adversary is rare**

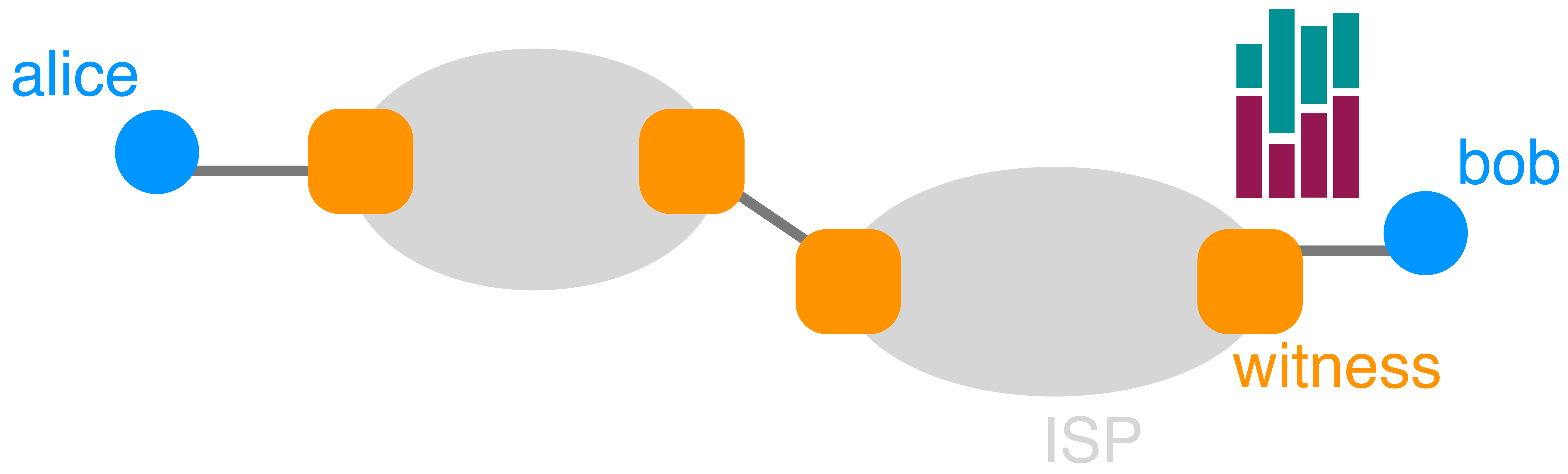
**Breaks against a global passive adversary**

# Internet performance transparency



# Internet performance transparency

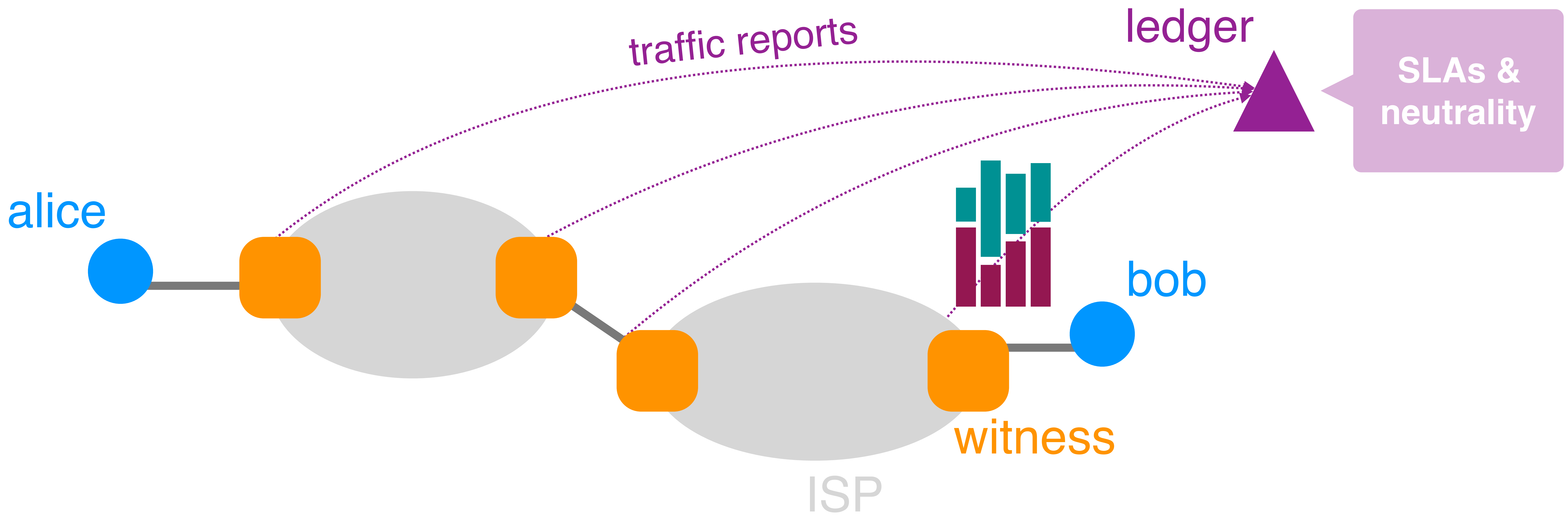
- Reports on traffic aggregates at ISP boundaries



**Weakens Tor**

# Internet performance transparency

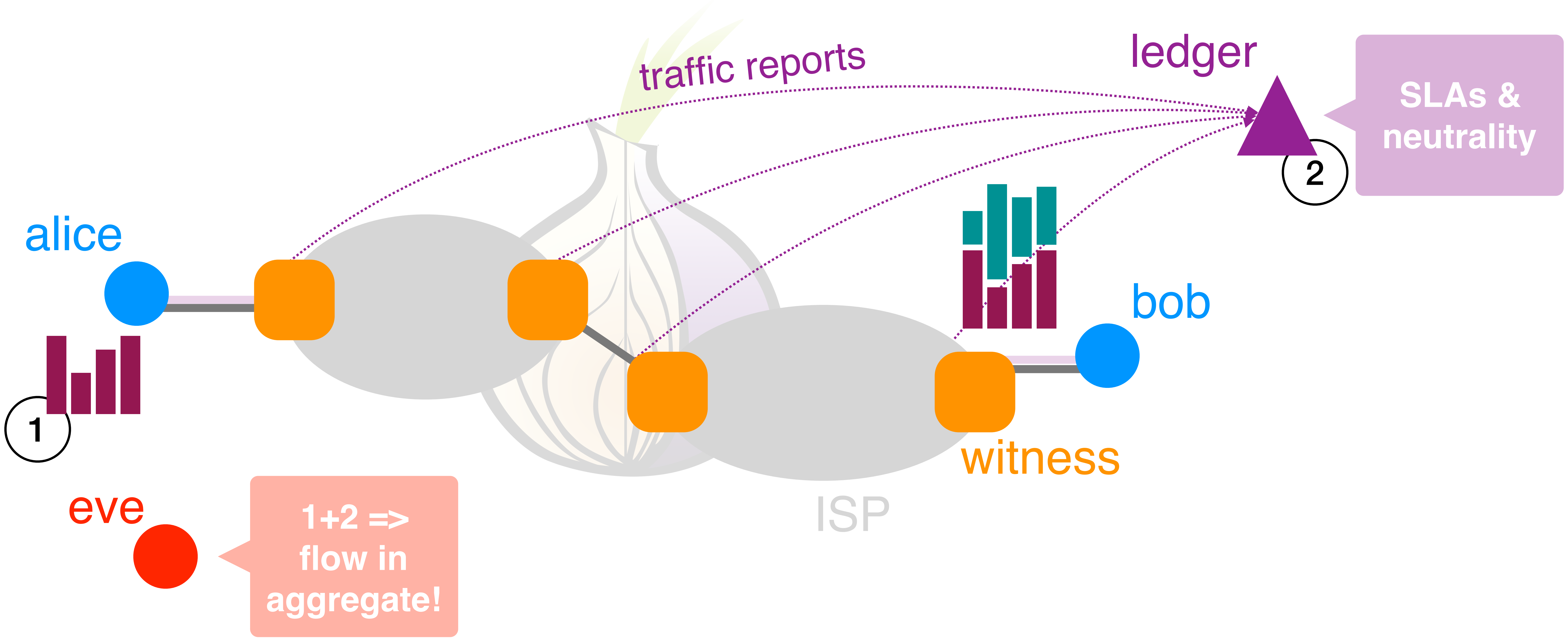
- Reports on traffic aggregates at ISP boundaries



**Weakens Tor**

# Internet performance transparency

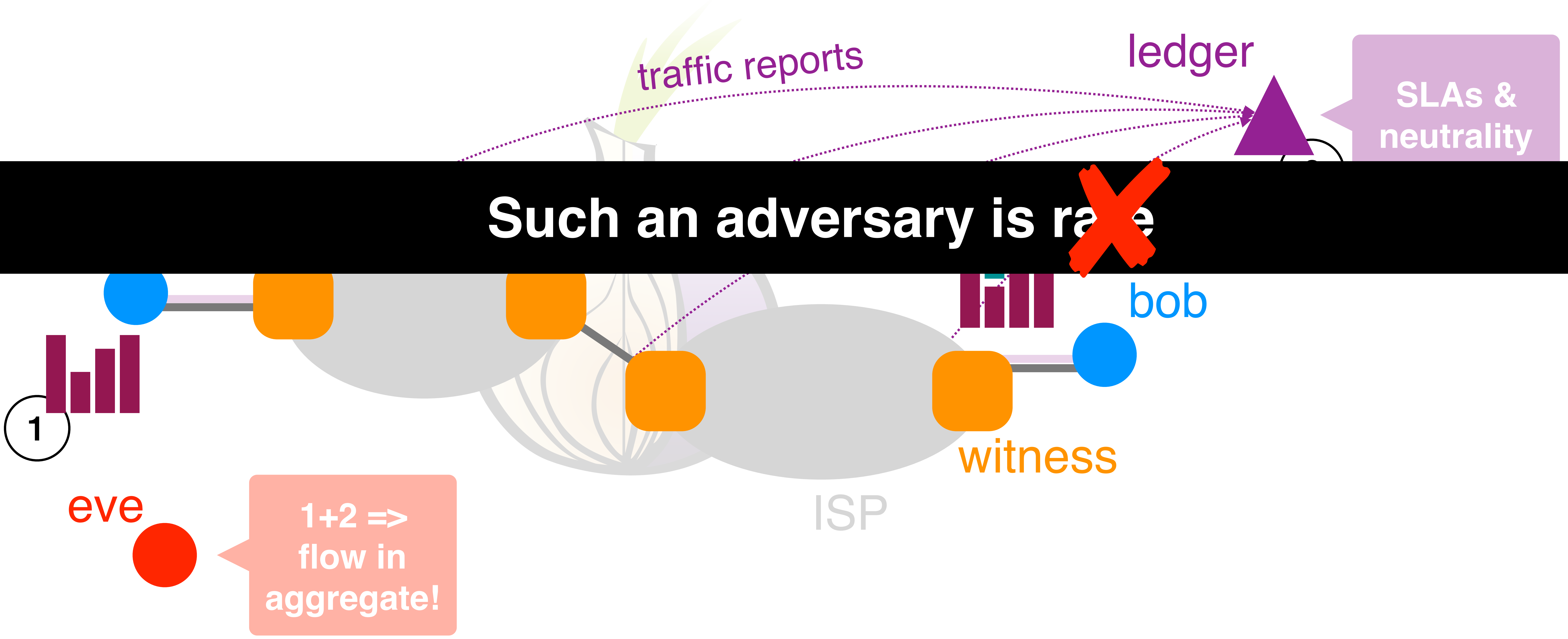
- Reports on traffic aggregates at ISP boundaries



**Weakens Tor**

# Internet performance transparency

- Reports on traffic aggregates at ISP boundaries



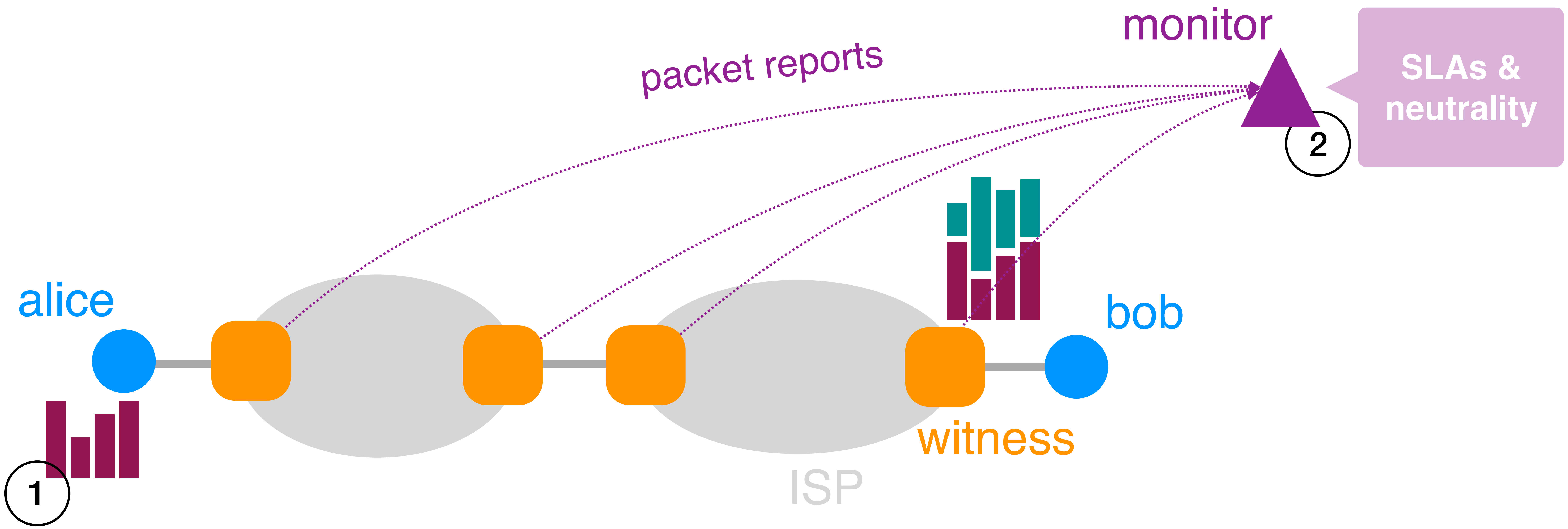
**Such an adversary is rare**

**Weakens Tor**

# Outline

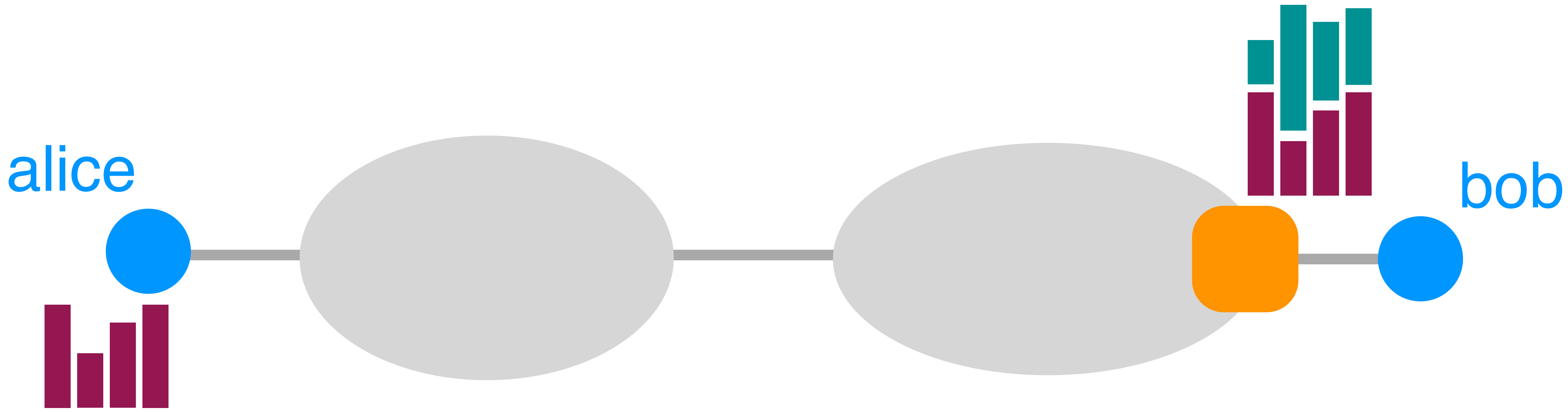
- **Measuring anonymity**
- Time granularity as noise
- Evaluation

# Measuring anonymity



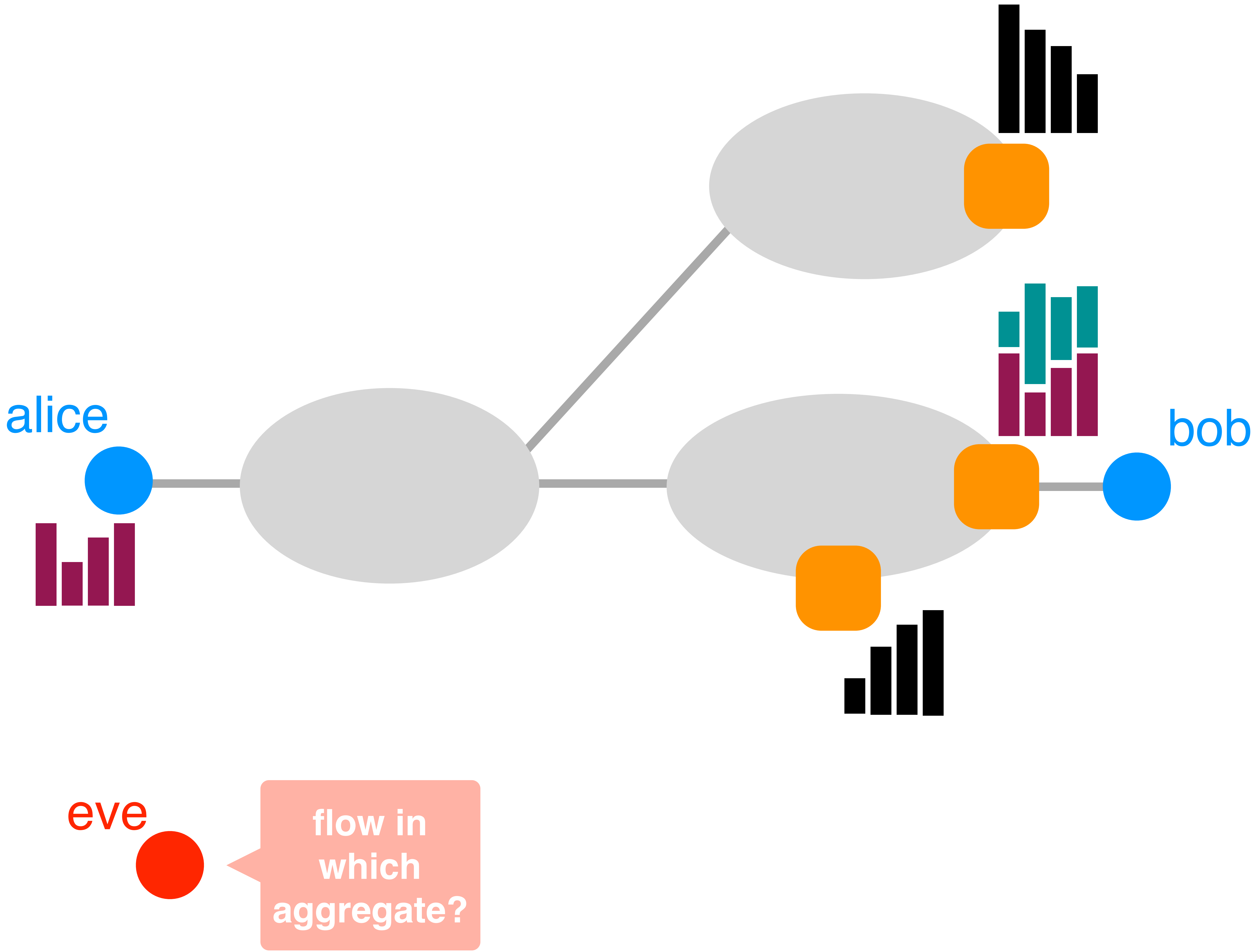
**T-anonymity set size captures deviation from ground truth**

# Measuring anonymity



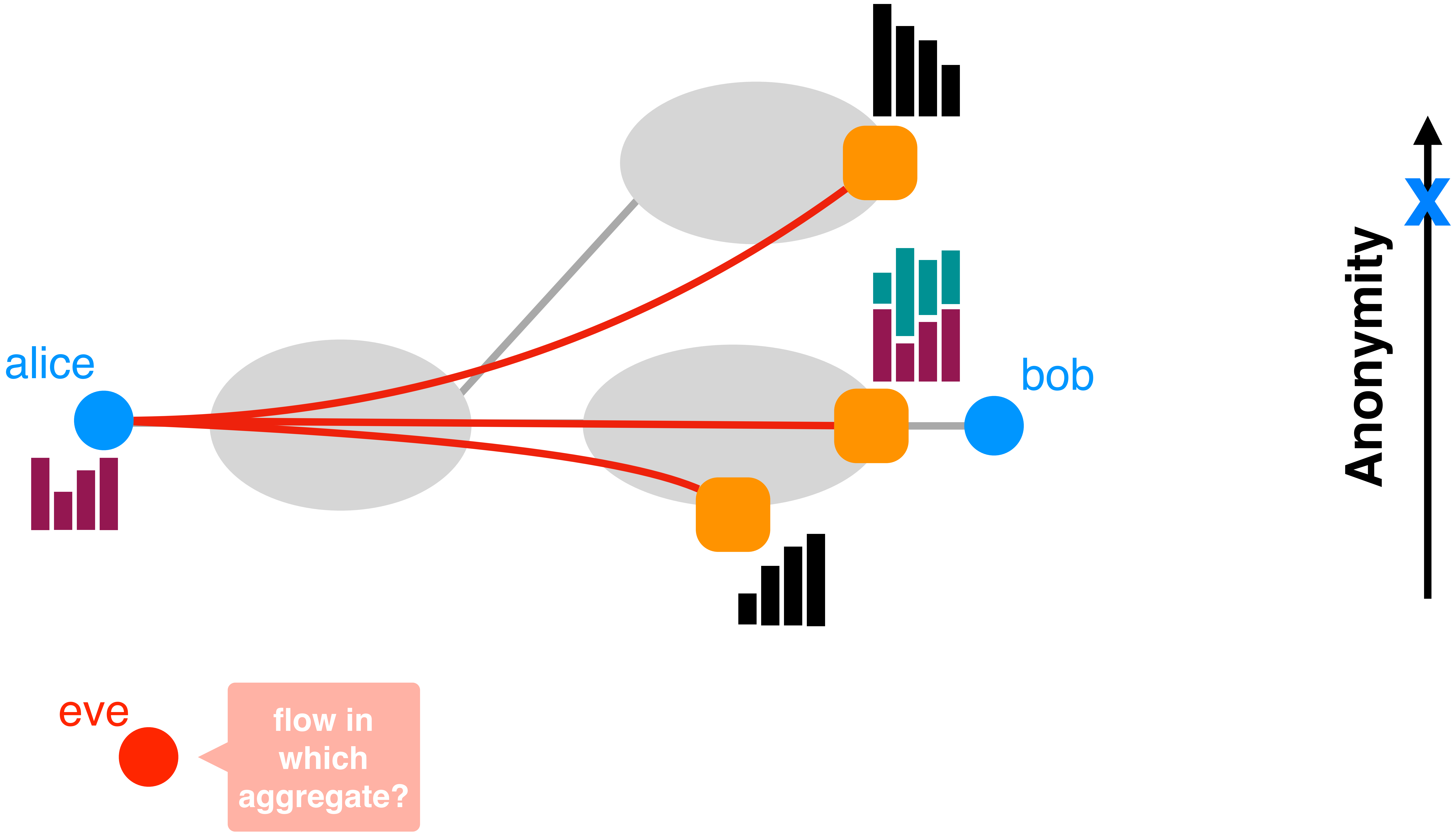
**T-anonymity set size captures deviation from ground truth**

# Measuring anonymity



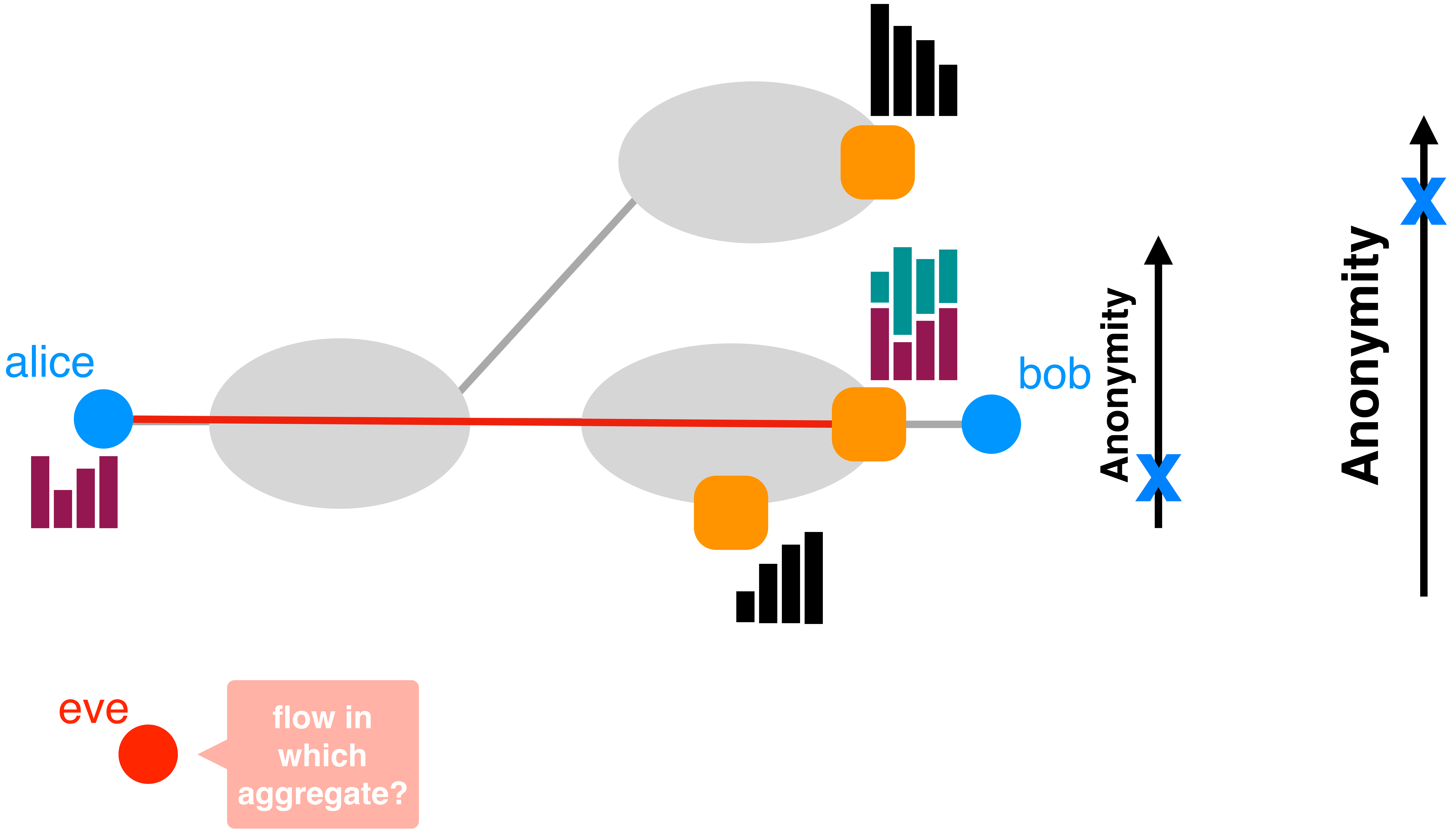
T-anonymity set size captures deviation from ground truth

# Measuring anonymity



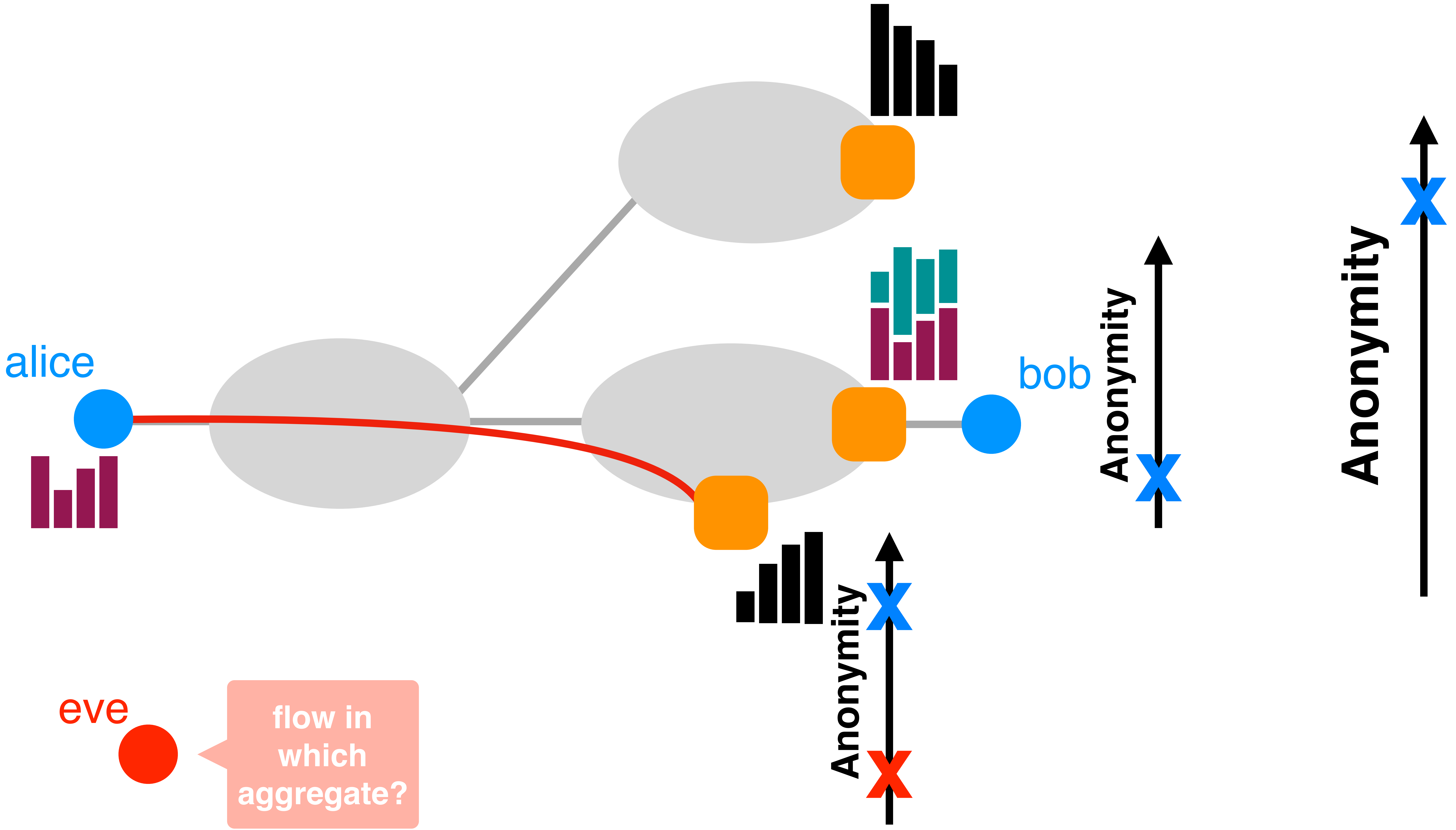
T-anonymity set size captures deviation from ground truth

# Measuring anonymity



**T-anonymity set size captures deviation from ground truth**

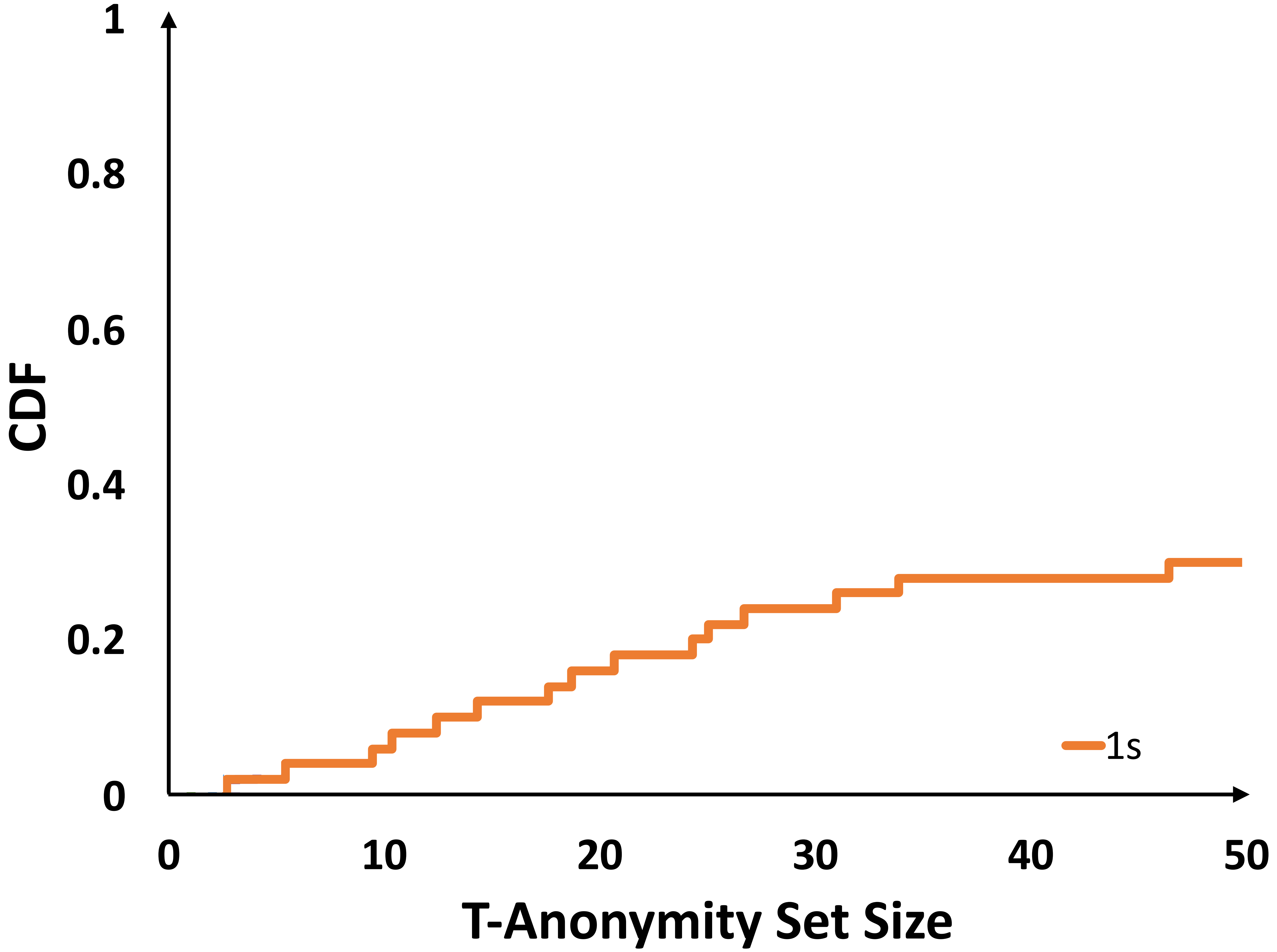
# Measuring anonymity



T-anonymity set size captures deviation from ground truth

# Dense aggregates are NOT anonymous

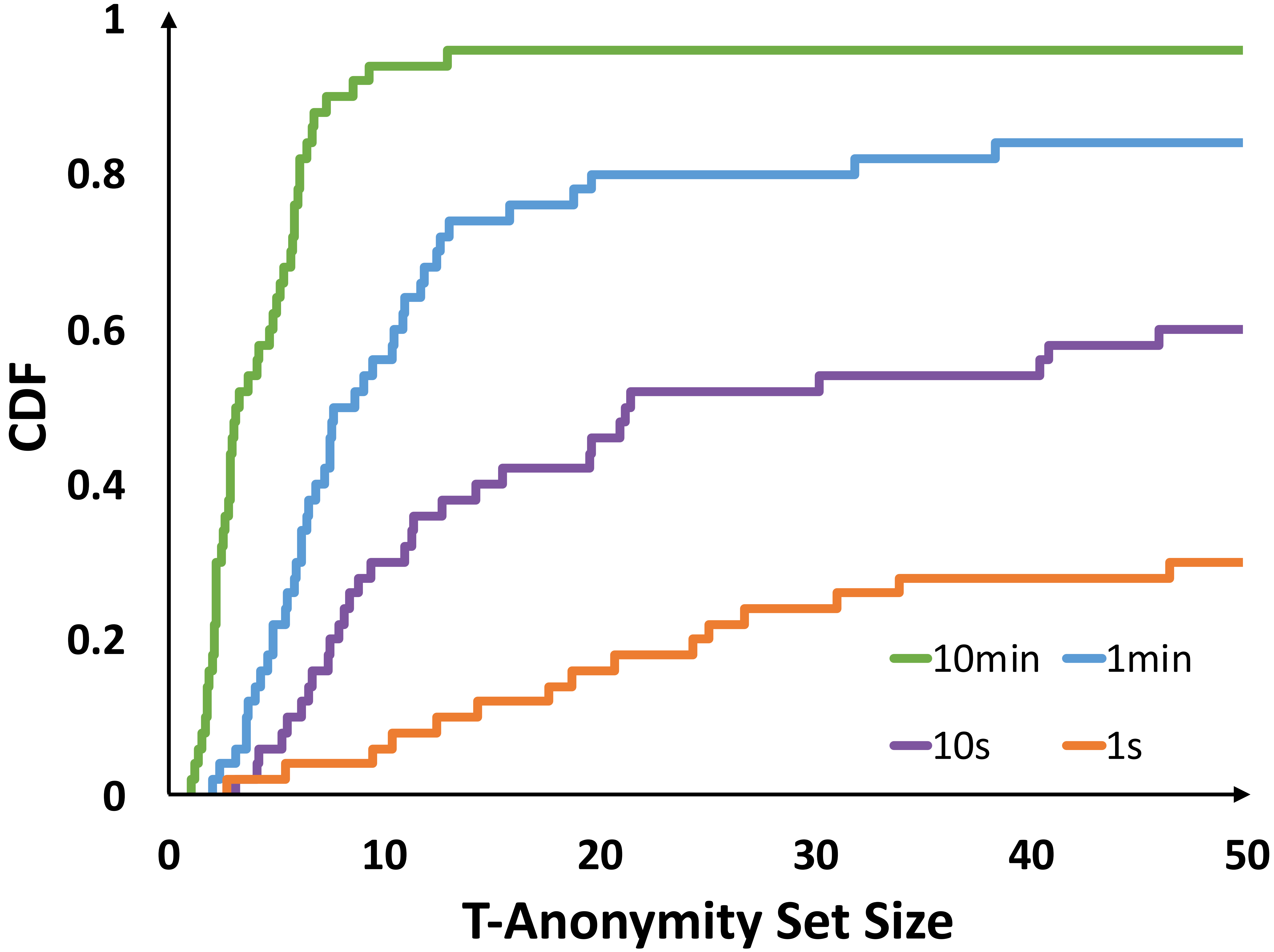
2018 CAIDA Internet traces
50 target flows/aggregates
512 flows per aggregate
reports per 1ms



**Given enough time, adversary de-anonymizes ~60% of cases**

# Dense aggregates are NOT anonymous

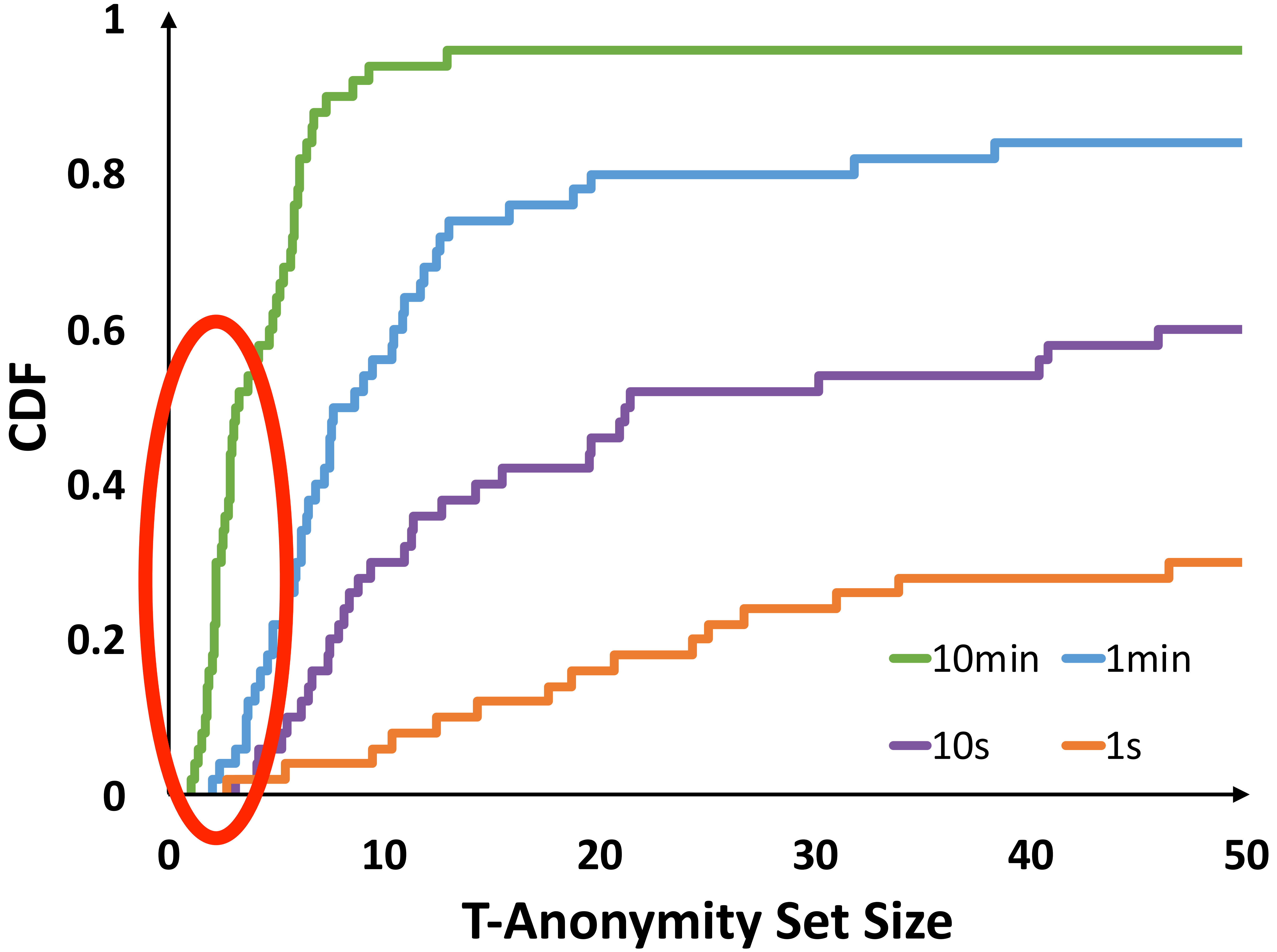
2018 CAIDA Internet traces
50 target flows/aggregates
512 flows per aggregate
reports per 1ms



Given enough time, adversary de-anonymizes ~60% of cases

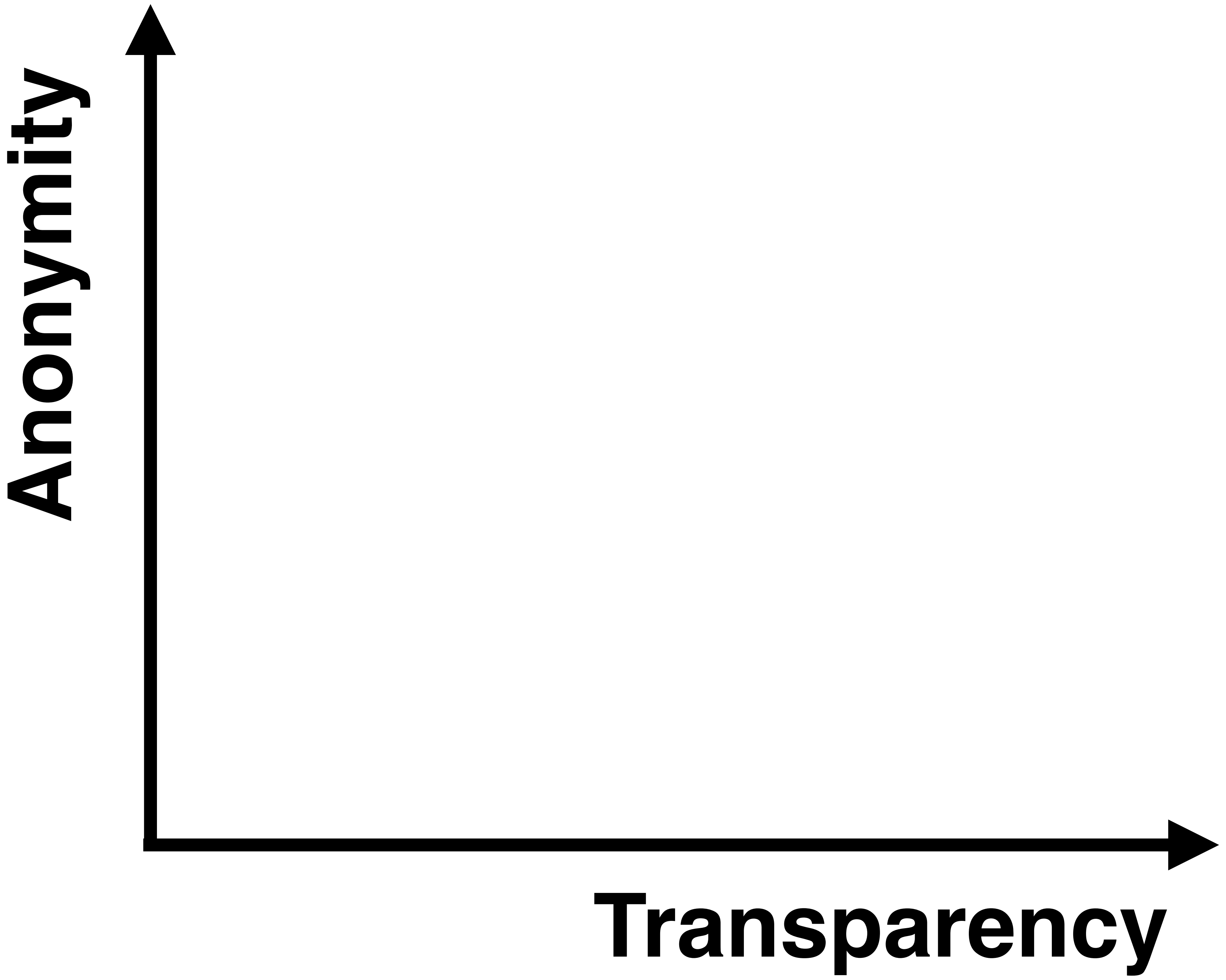
# Dense aggregates are NOT anonymous

2018 CAIDA Internet traces
50 target flows/aggregates
512 flows per aggregate
reports per 1ms

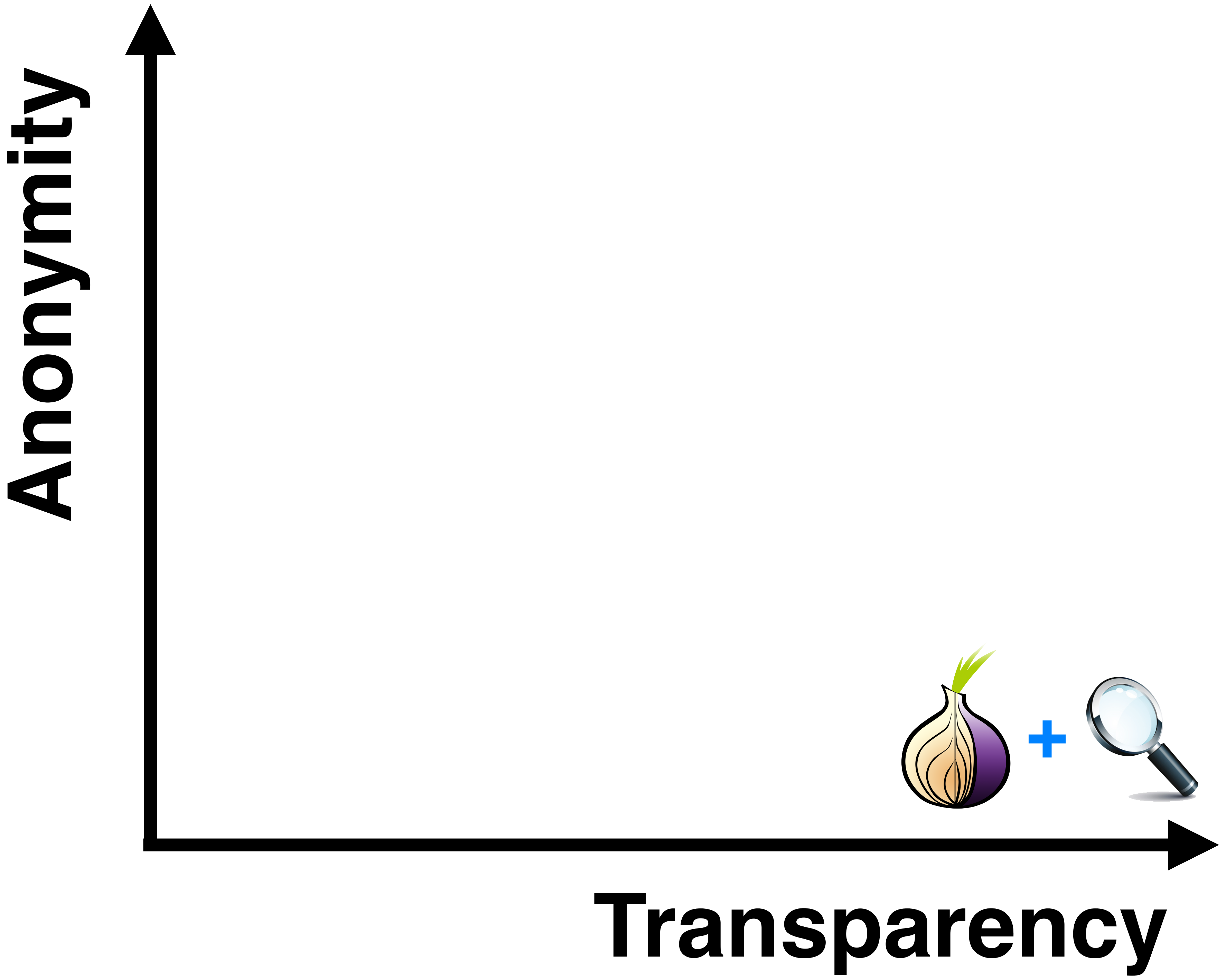


**Given enough time, adversary de-anonymizes ~60% of cases**

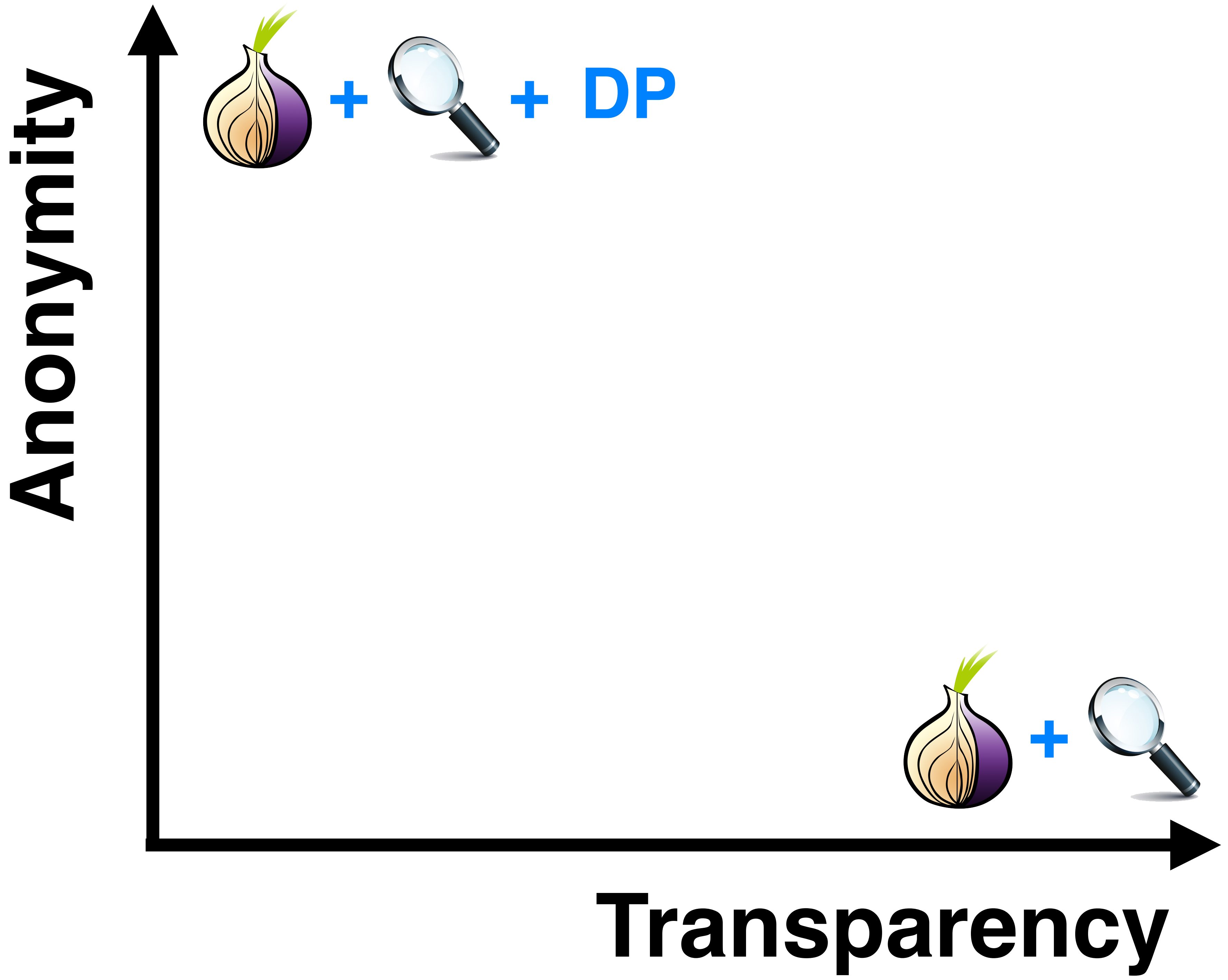
# The anonymity-transparency trade-off



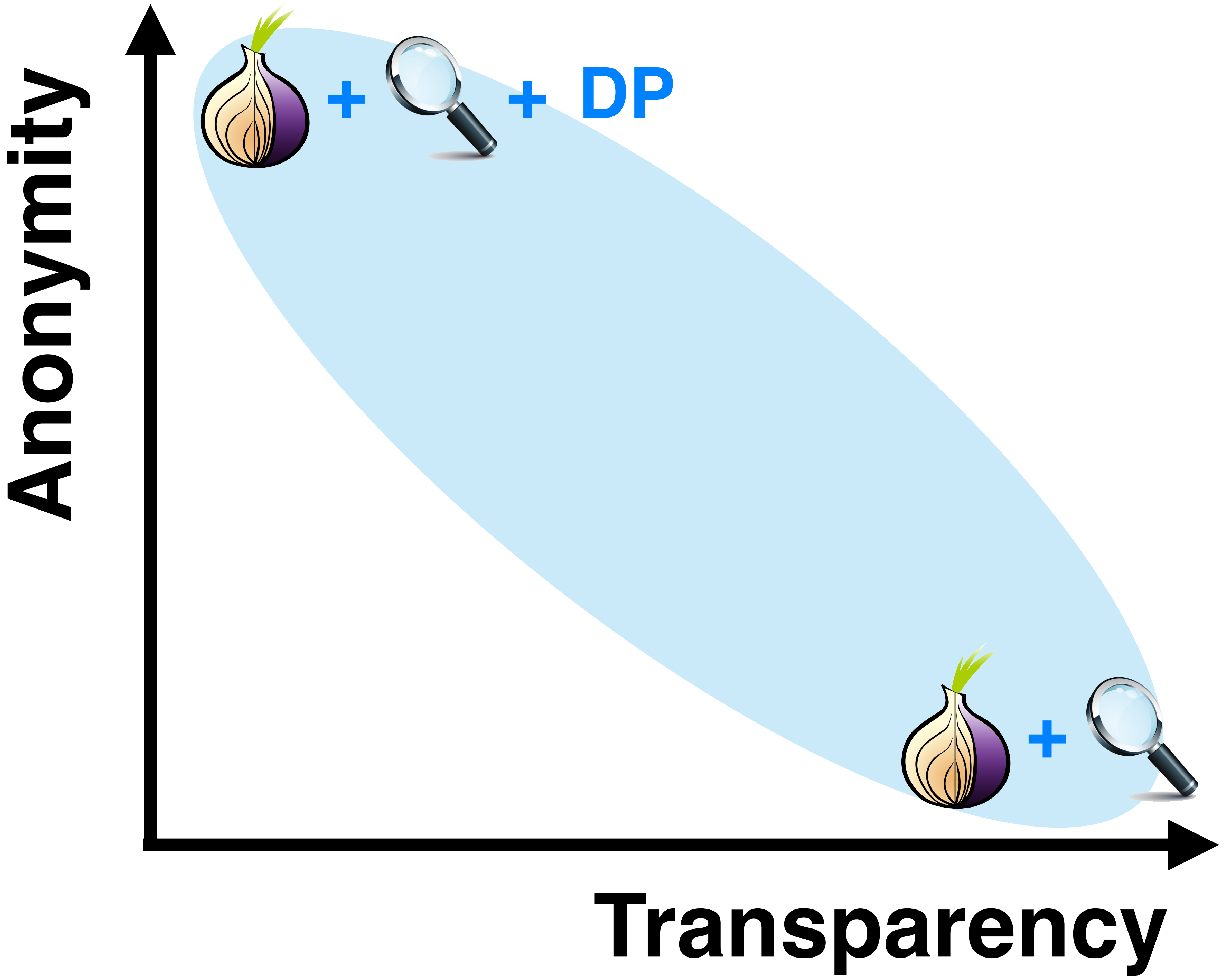
# The anonymity-transparency trade-off



# The anonymity-transparency trade-off

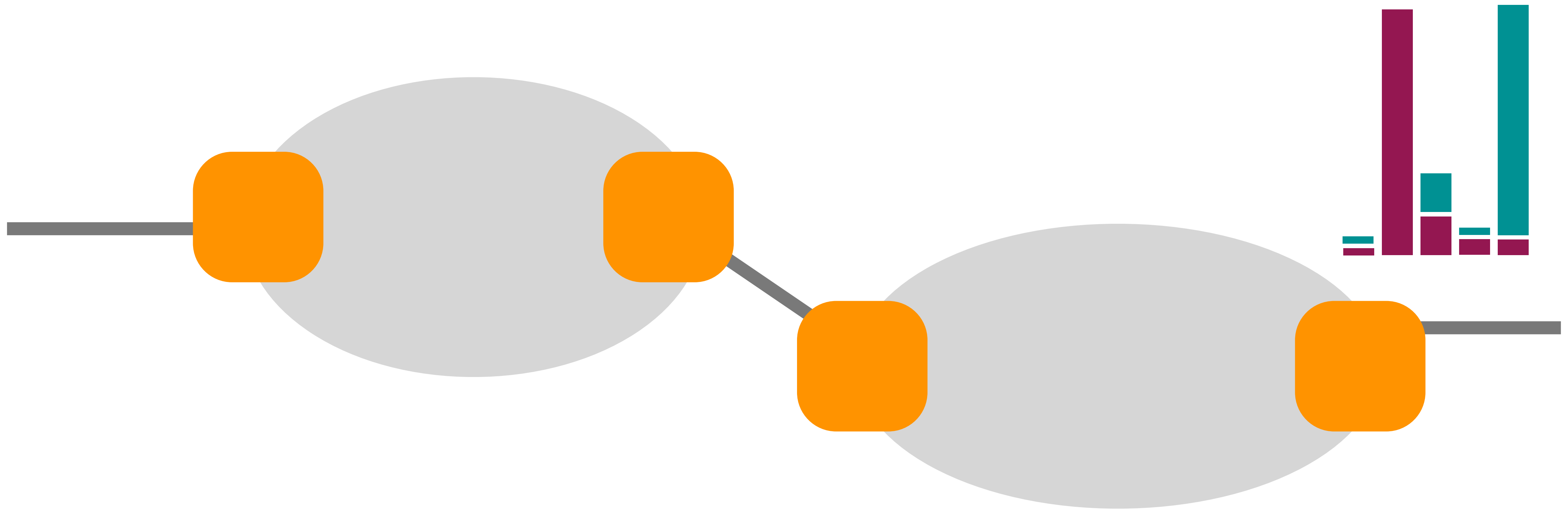


# The anonymity-transparency trade-off



# Improving anonymity is not easy

- Any flow could be a target
- No network coordination

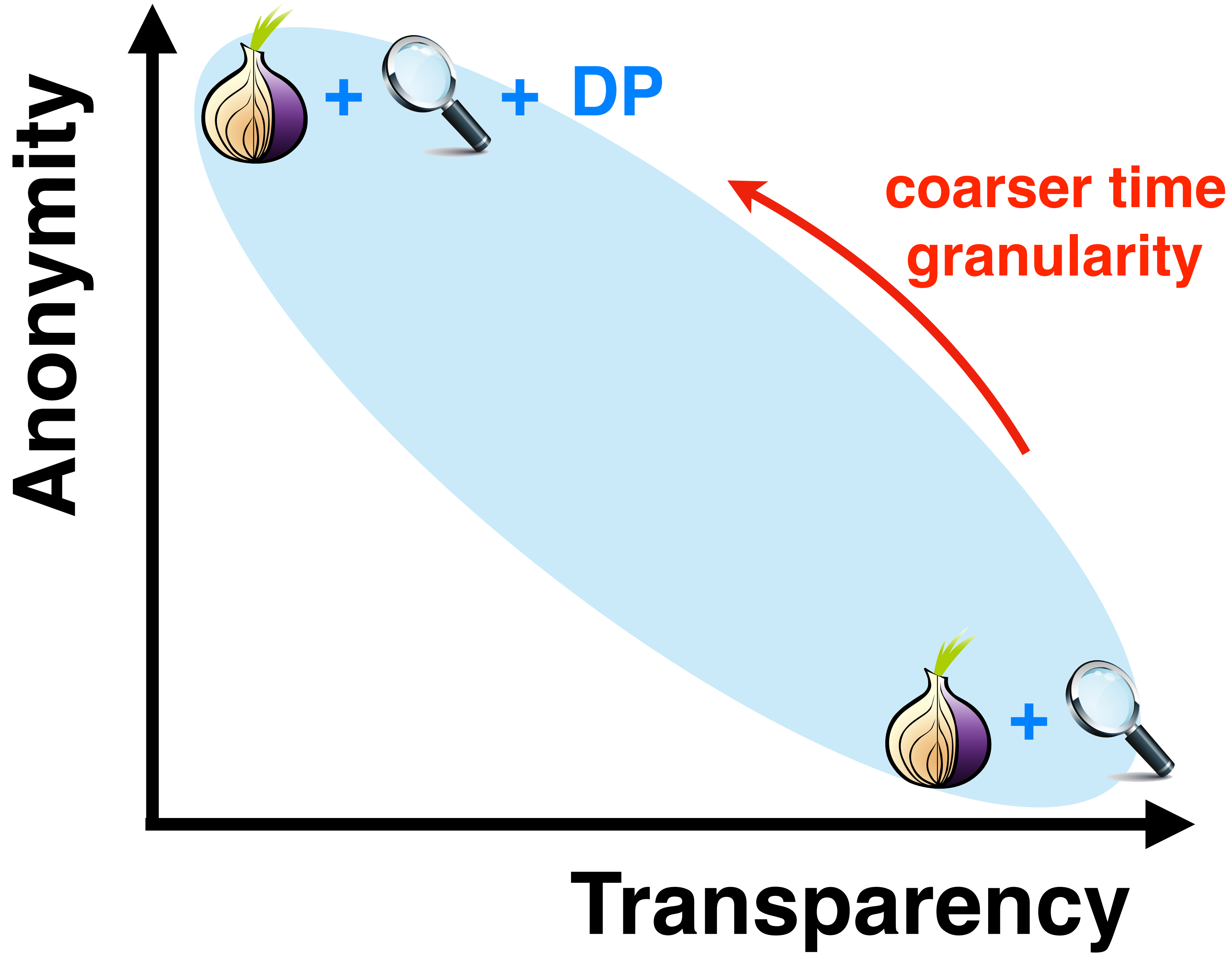


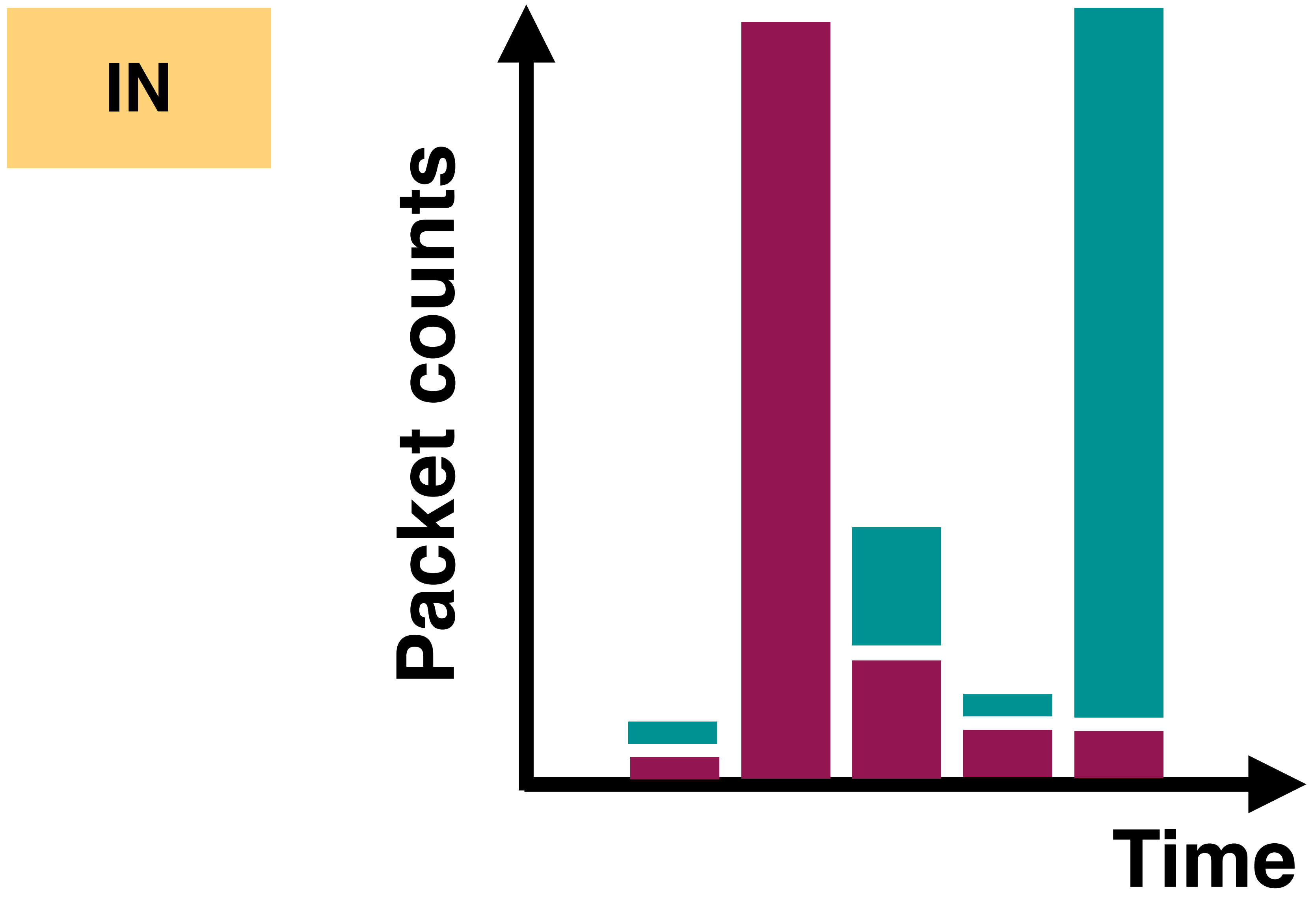
**Strike a good balance for all flows with ISP-local decisions**

# Outline

- Measuring anonymity
- **Time granularity as noise**
- Evaluation

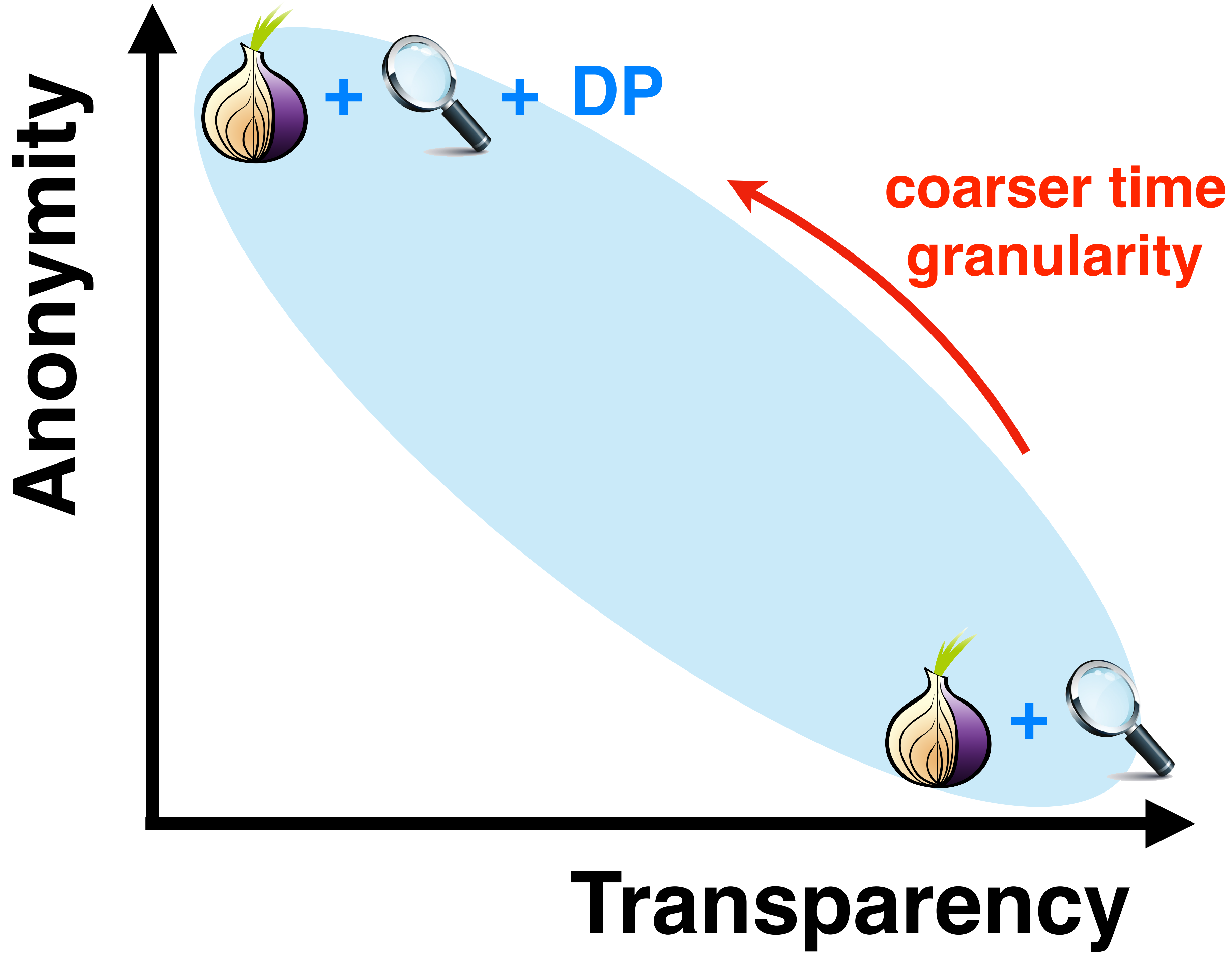
# Time granularity as noise

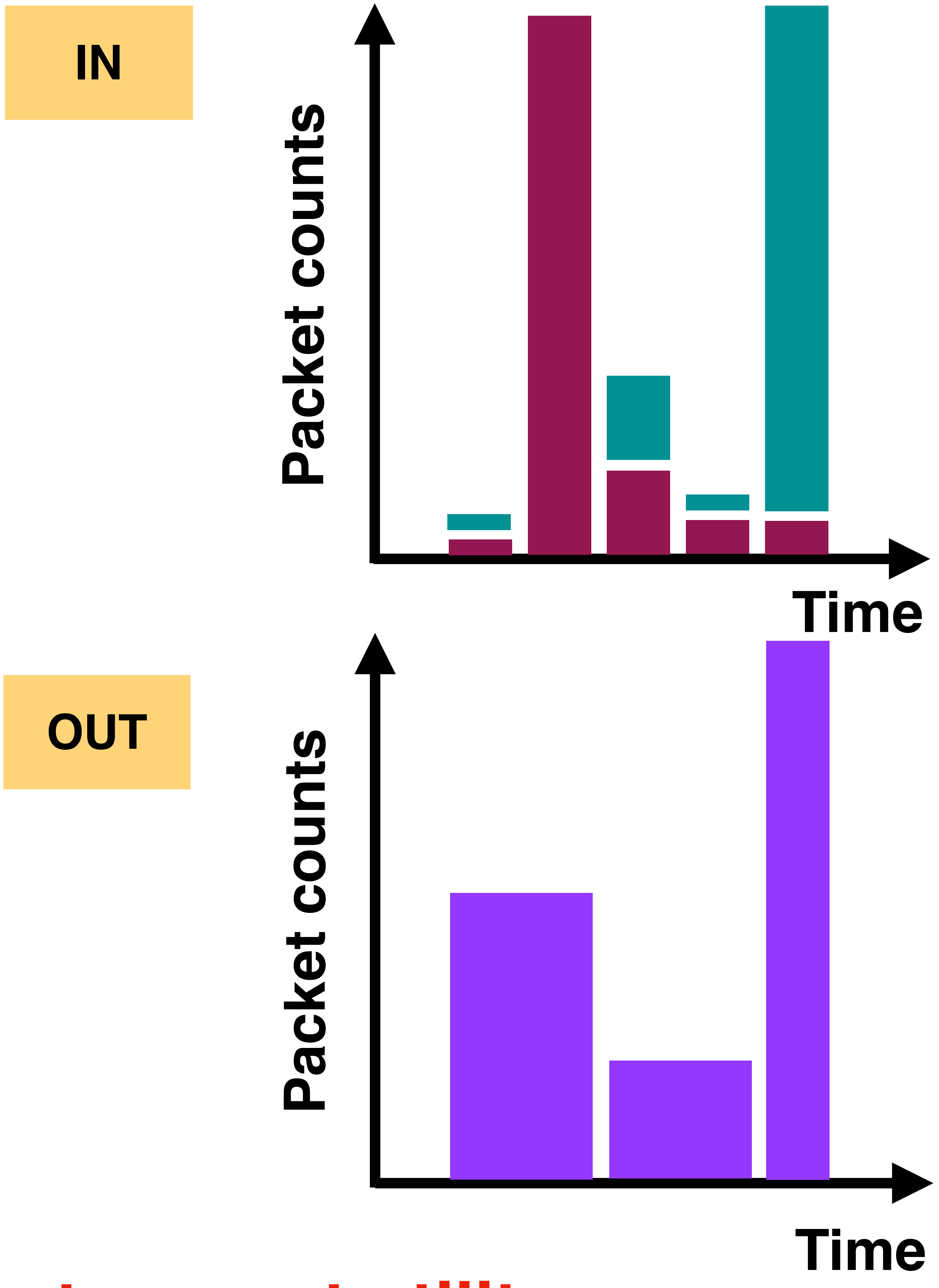




**Hides flow patterns, but impacts report utility**

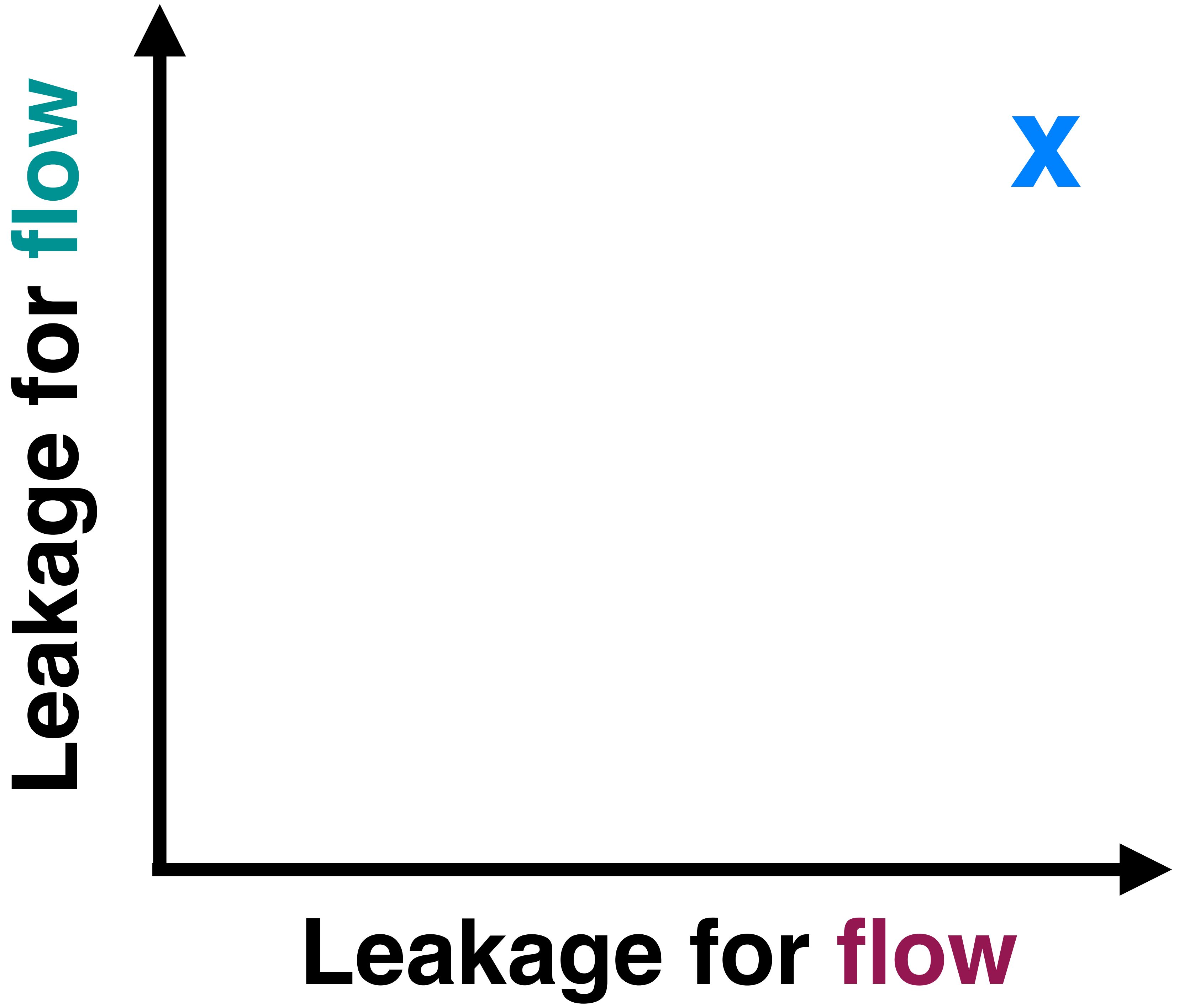
# Time granularity as noise



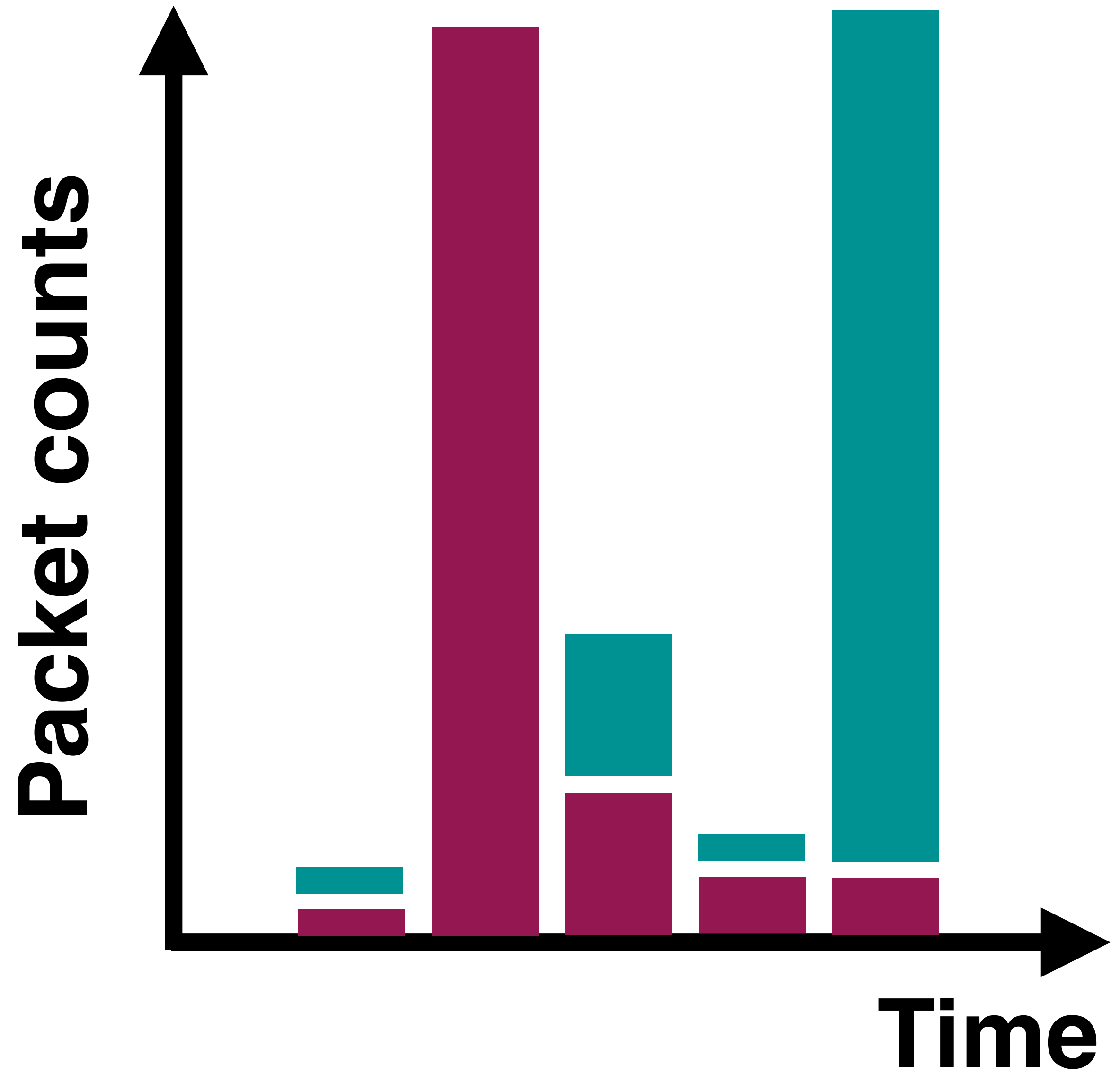


Hides flow patterns, but impacts report utility

# Adaptive binning

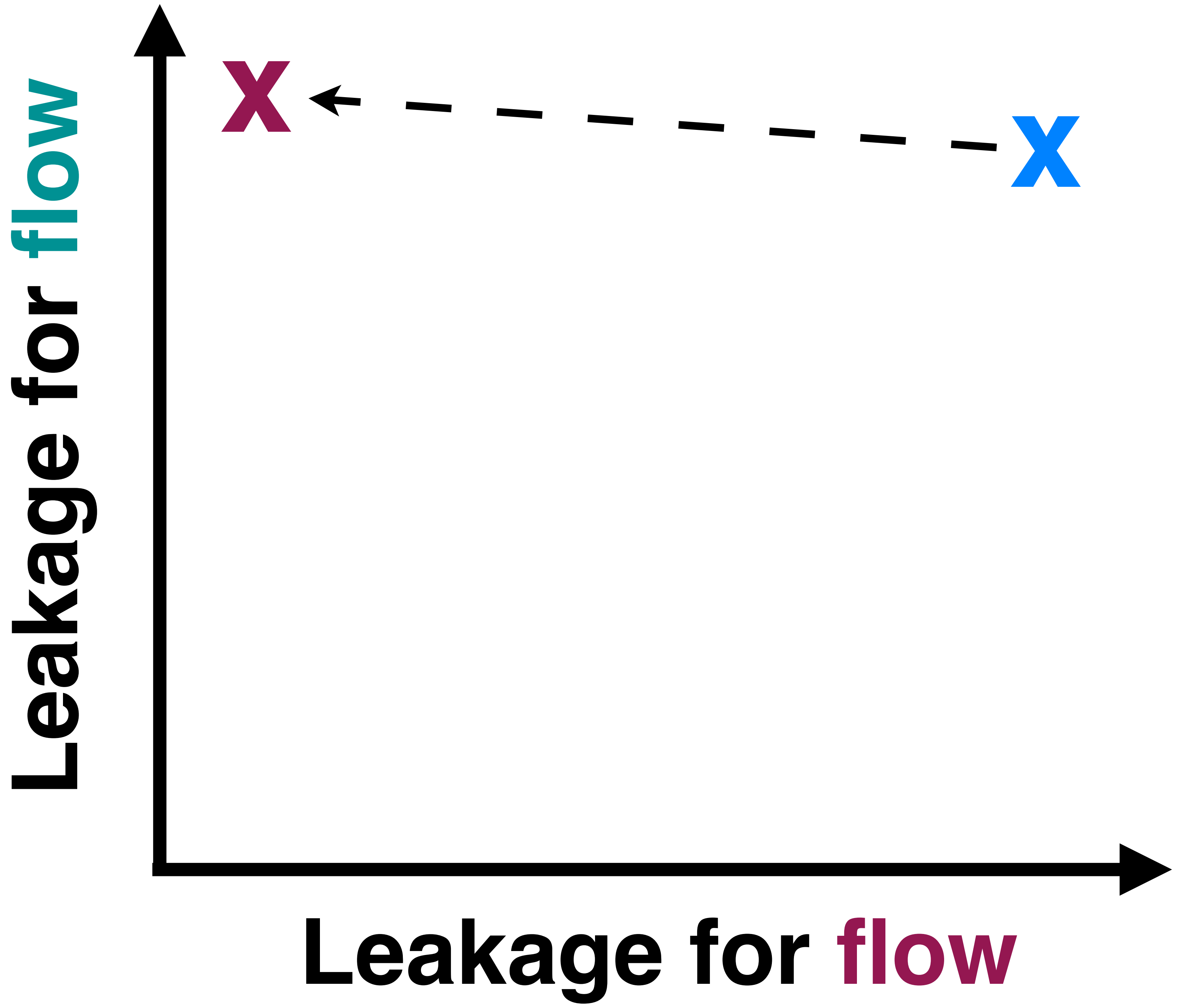


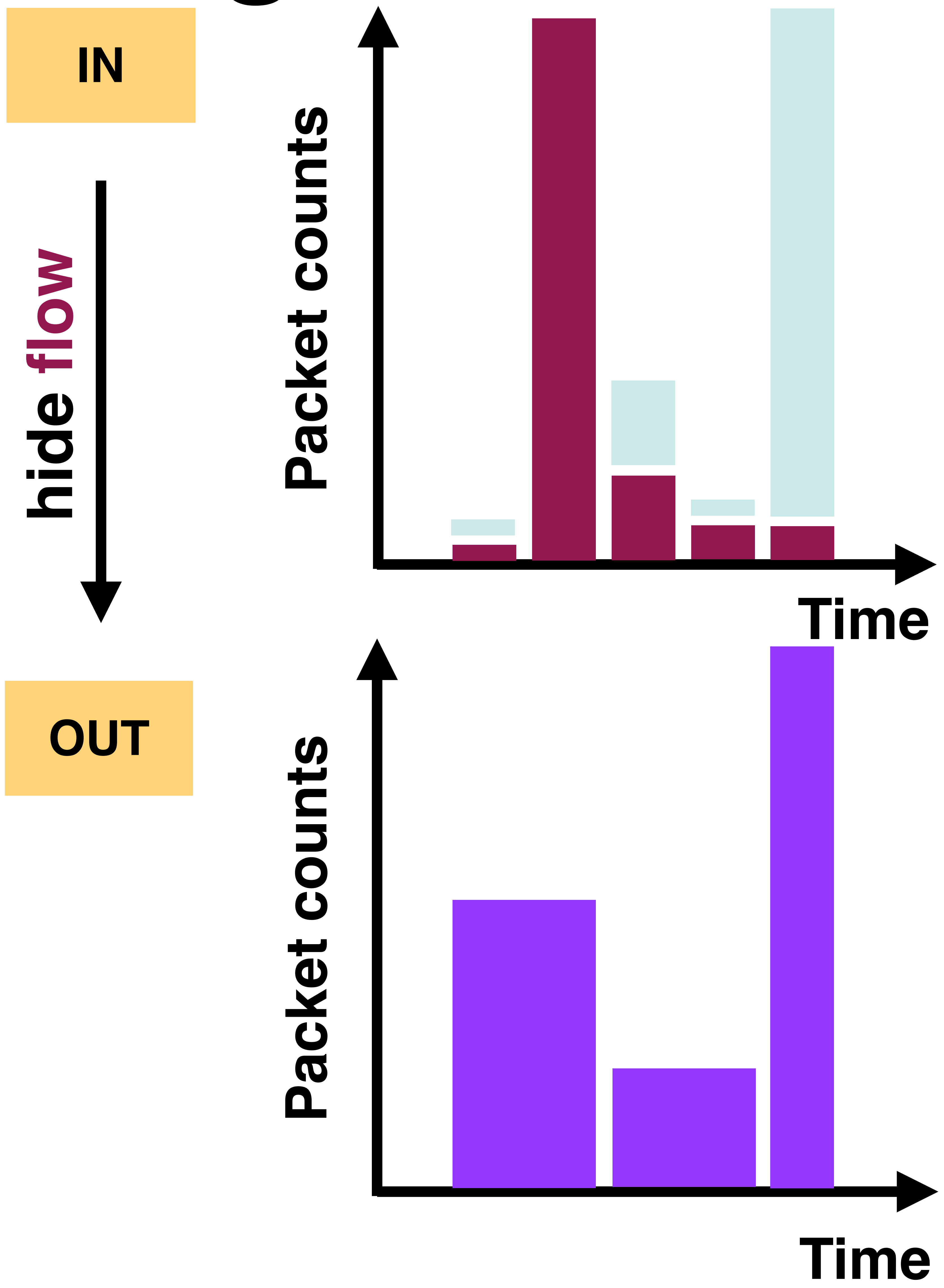
IN



Hide "worst-off" flow, subject to an upper time granularity

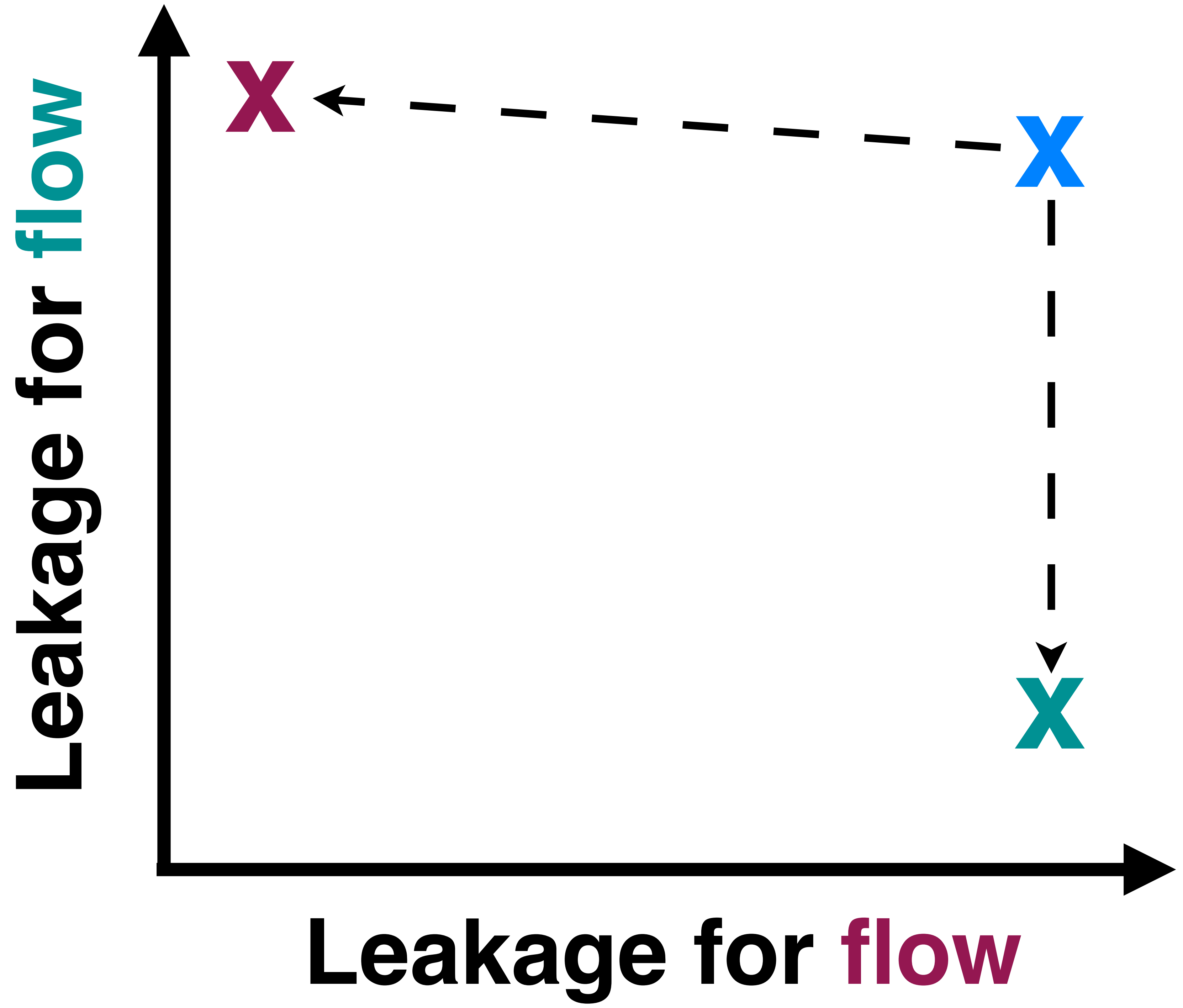
# Adaptive binning

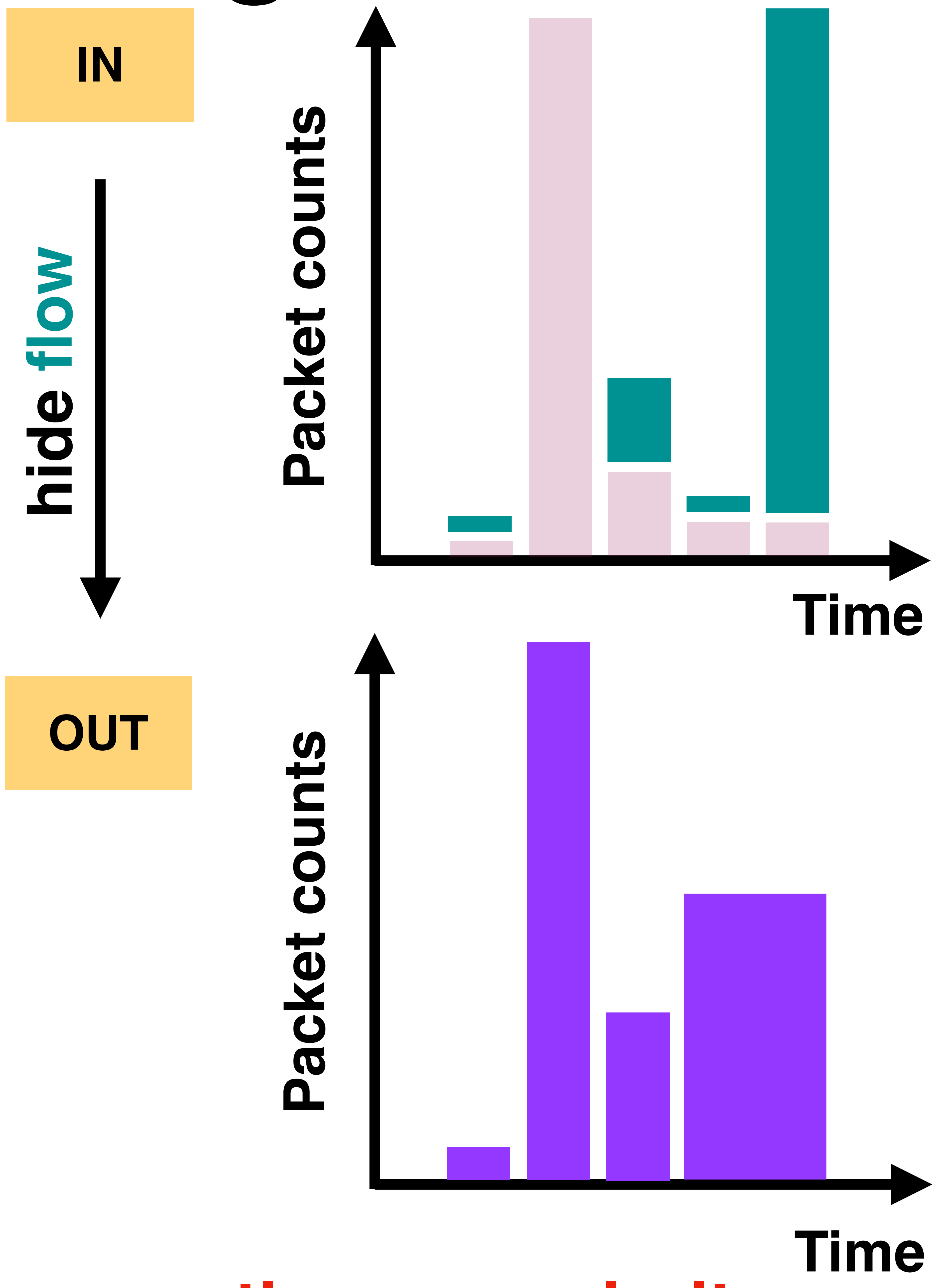




Hide “worst-off” flow, subject to an upper time granularity

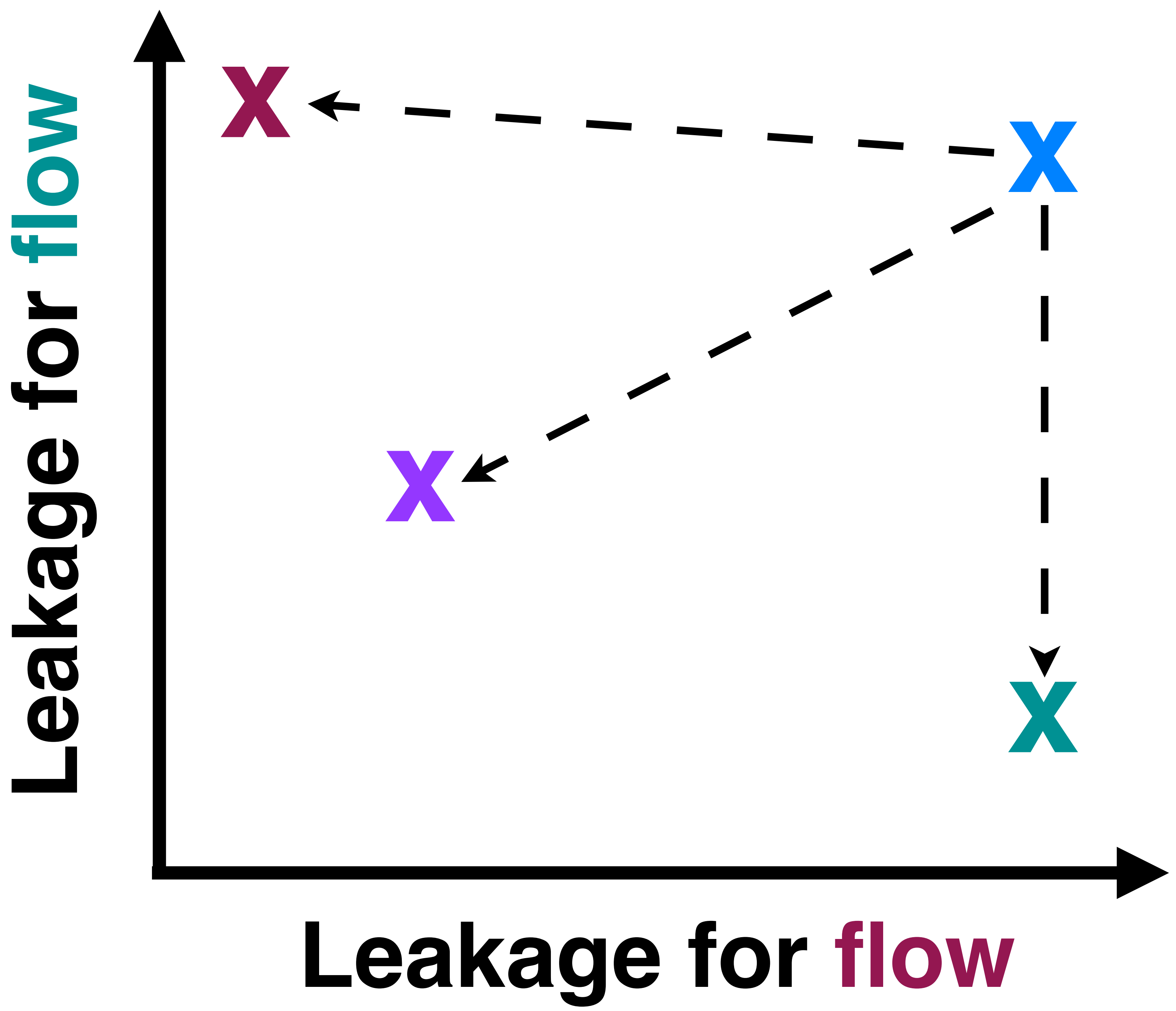
# Adaptive binning

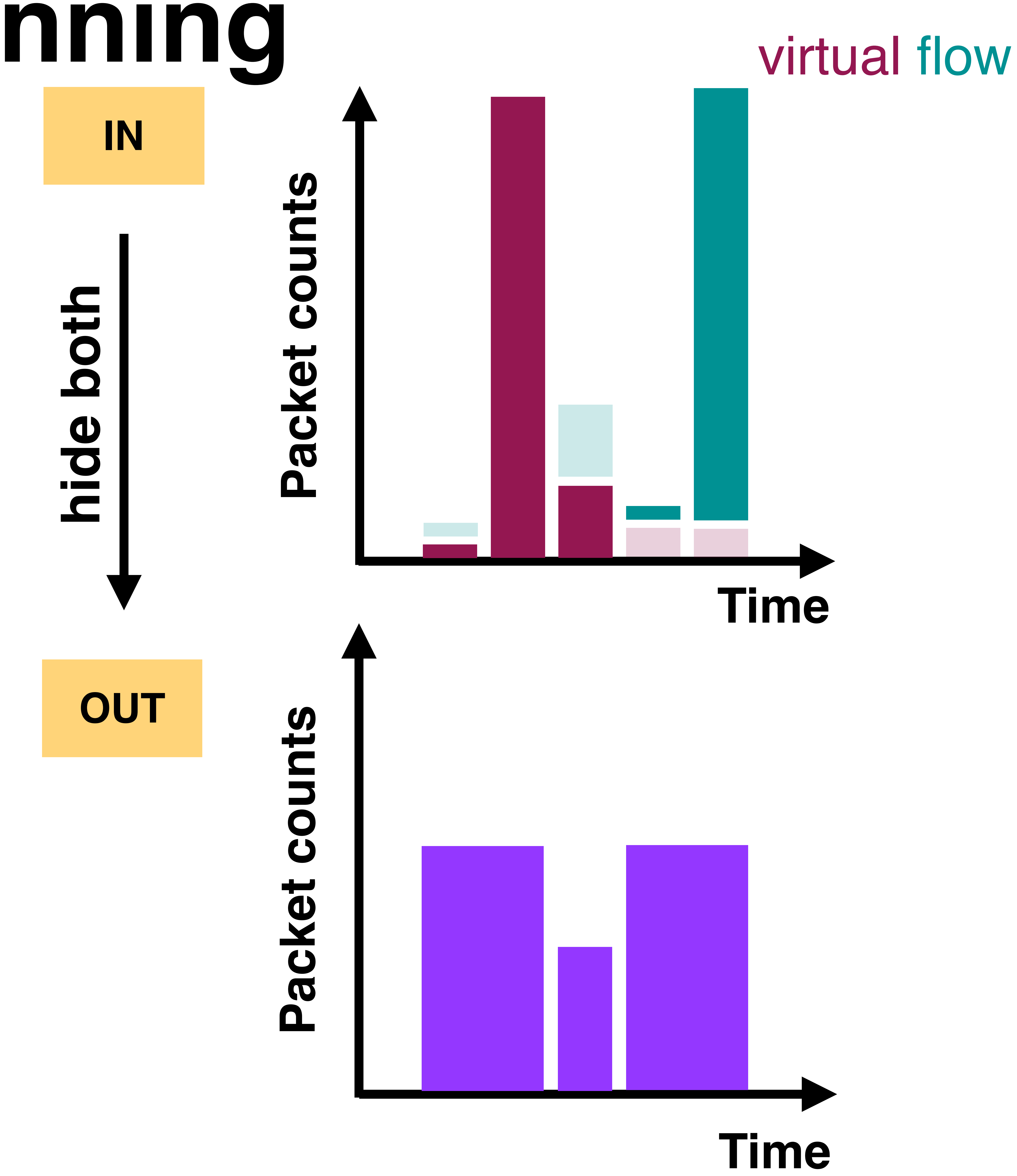




Hide “worst-off” flow, subject to an upper time granularity

# Adaptive binning





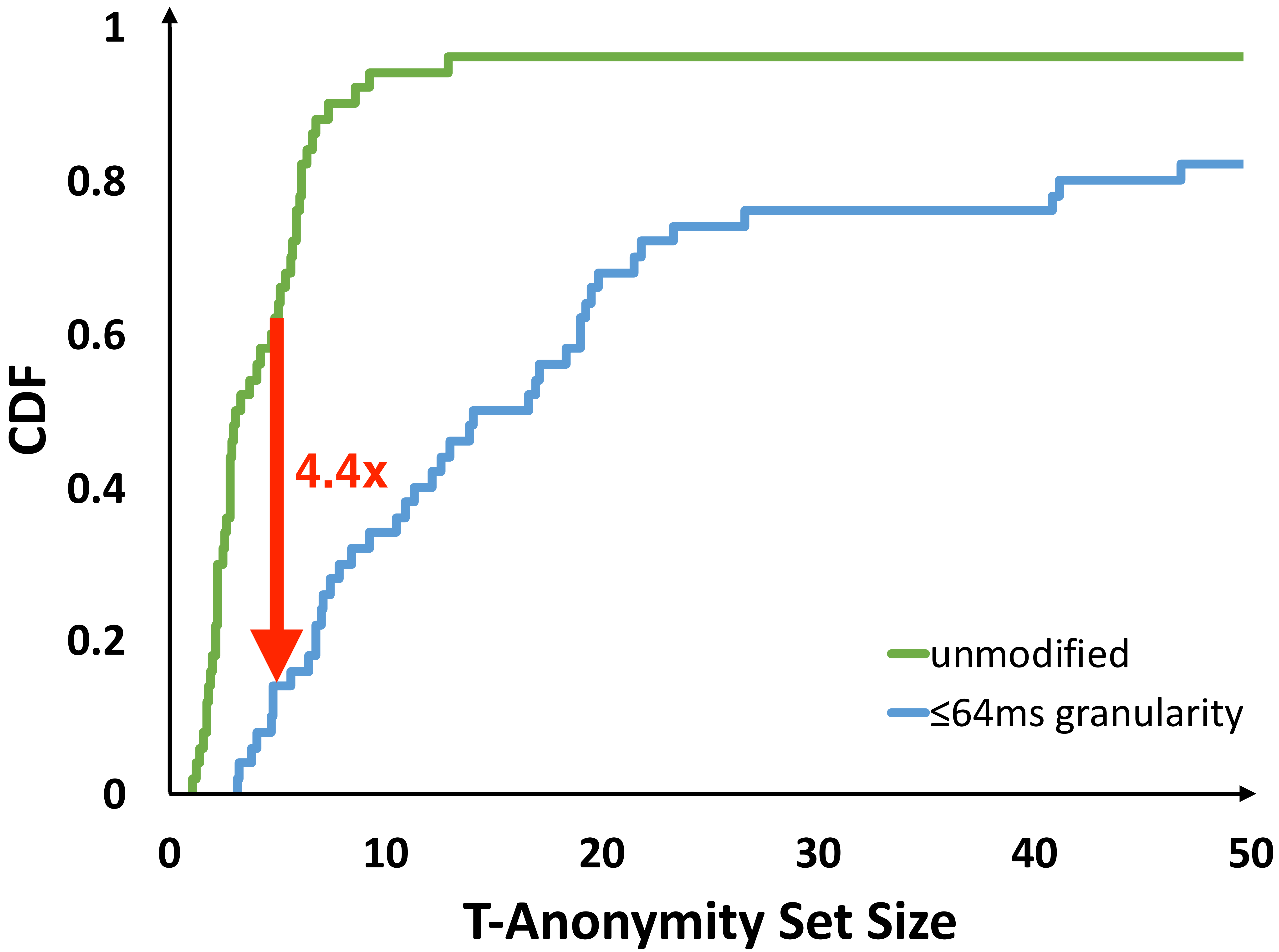
**Hide “worst-off” flow, subject to an upper time granularity**

# Outline

- Measuring anonymity
- Time granularity as noise
- **Evaluation**

# Anonymity re-assessed

2018 CAIDA Internet traces
50 target flows/aggregates
512 flows per aggregate
10min observation



**4.4x improvement at sub-second granularity**

# In the paper

- Other experimental setups
  - Sparse aggregates, Poisson traffic, “on-off” traffic
- The cost of differential privacy to transparency
- Scalability of the algorithm

# Conclusion

- Rethink transparency, as it can greatly damage Tor anonymity
- Time granularity as noise

- Trustworthy performance metrics over untrusted networks
- Reconcile transparency with privacy of network topology

**Thank you**