# Sparta: Practical Anonymity with Resistance to Traffic Analysis

Kyle Fredrickson, Ioannis Demertzis, Jim Hughes, Darrell D.E. Long

UC SANTA CRUZ
BaskinEngineering
CRSS
ssrc
STORAGE SYSTEMS RESEARCH CENTER
NSF

# Metadata and Why You Care

- **Goal:** Private messaging systems.
- Isn’t encryption enough?
  - Necessary — hides content
  - Not sufficient — leaks metadata
- Metadata is extremely valuable.
  - “With enough metadata you don’t really need content.”
    — Former NSA General Counsel
  - Statistical Relational Learning

# Metadata and Why You Care

**"We kill people based on metadata."**

Former NSA, CIA Director, Gen. Michael Hayden

# Existing Work

CRSS

**Atom: Horizontally Scaling Strong Anonymity**

Albert Kwon
MIT
Henry Corrigan-Gibbs
Stanford
Srinivas Devadas
MIT
Bryan Ford
EPFL

**Tor: The Second-Generation Onion Router**

*Roger Dingledine*
*The Free Haven Project*
*arma@freehaven.net*
*Nick Mathewson*
*The Free Haven Project*
*nickm@freehaven.net*
*Paul Syverson*
*Naval Research Lab*
*syverson@itd.nrl.navy.mil*

**Express: Lowering the Cost of Metadata-hiding Communication with Cryptographic Privacy**

Saba Eskandarian
*Stanford University*
Henry Corrigan-Gibbs
*MIT CSAIL*
Matei Zaharia
*Stanford University*
Dan Boneh
*Stanford University*

Clarion: Anonymous Communication from Multiparty Shuffling Protocols

Saba Eskandarian
UNC Chapel Hill
Dan Boneh
Stanford University

**XRD: Scalable Messaging System with Cryptographic Privacy**

Albert Kwon
*MIT*
David Lu
*MIT PRIMES*
Srinivas Devadas
*MIT*

Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms

David L. Chaum
University of California, Berkeley

**The Dining Cryptographers Problem:**
***Unconditional Sender and Recipient Untraceability***

David Chaum
Centre for Mathematics and Computer Science, Kruislan 413, 1098 SJ Amsterdam, The Netherlands

**Unobservable communication over fully untrusted infrastructure (extended version)***

Sebastian Angel
UT Austin and NYU
Srinath Setty
Microsoft Research

Vuvuzela: Scalable Private Messaging Resistant to Traffic Analysis

*Jelle van den Hooff, *David Lazar, Matei Zaharia, and Nickolai Zeldovich
*MIT CSAIL*

Groove: Flexible Metadata-Private Messaging

Ludovic Barman
EPFL
Moshe Kol
Hebrew University of Jerusalem
David Lazar
EPFL
Yossi Gilad
Hebrew University of Jerusalem
Nickolai Zeldovich
MIT CSAIL

Sabre: Sender-Anonymous Messaging with Fast Audits

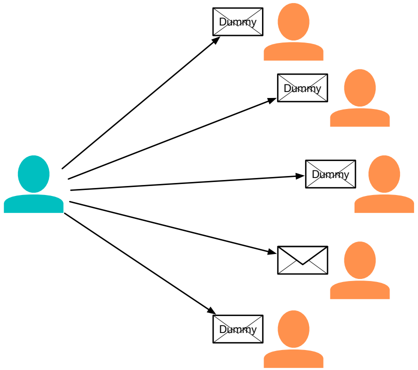
Adithya Vadapalli
University of Waterloo
adithya.vadapalli@uwaterloo.ca
Kyle Storrier
University of Calgary
kyle.storrier@ucalg
Ryan Henry
University of Calgary

**The Loopix Anonymity System**

Ania M. Piotrowska
*University College London*
Jamie Hayes
*University College London*
Tariq Elahi
*KU Leuven*
Sebastian Meiser
*University College London*
George Danezis
*University College London*

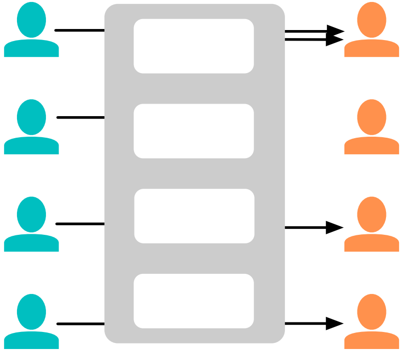
# Existing Work Doesn't Work

**Choose one**



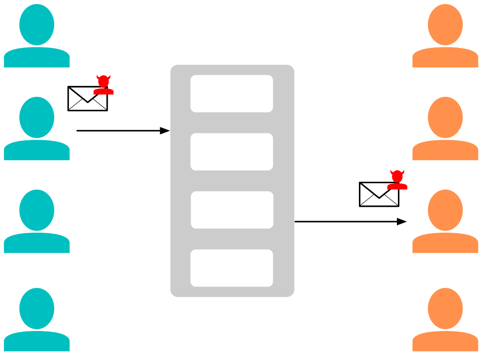

❌ High Costs

**OR**



❌ Unrealistic Assumptions

**OR**



❌ Insecure

# We Ask…

**Q1: Can systems provide long-term traffic analysis resistance practically?**

**Q2: Can they be securely and scalably implemented?**

# Our Contributions

- Precise definitions of traffic analysis resistance.

- New class of anonymity system.
  - Provably resists traffic analysis
  - Under weak assumptions
  - With low costs (**3400x** reduction in traffic)

- Sparta: securely implements this class using Intel SGX.
  - Scalable (**15x** faster)
  - Usable
  - Deployable

# Our Contributions

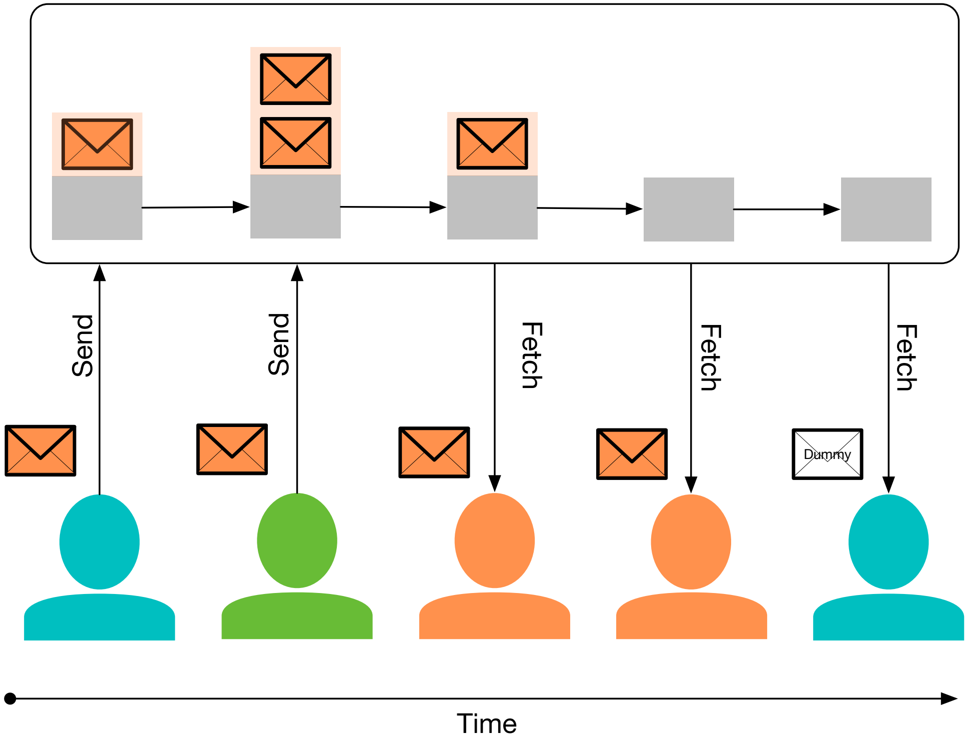
CRSS

- ~~Precise definitions of traffic analysis resistance.~~

- **New class of anonymity system.**
  - Provably resists traffic analysis
  - Under weak assumptions
  - With low costs (**3400x** reduction in traffic)

- Sparta: securely implements this class using Intel SGX.
  - Scalable (**15x** faster)
  - Usable
  - Deployable

# Oblivious MultiQueues (OMQs)




OMQs are a set of queues.

**Security Properties**

- Send$(q_i, m)$ should **not** leak which queue is written to
- Fetch$(q_i, k)$ should **only** leak $k$

# Deferred Retrieval

**Traffic Analysis Resistance**: No correlation between sender and receiver traffic.

**Assumptions for OMQs**

- Users fetch independently of received messages.
  - ✅ Users can go offline
  - ✅ Users can have different rates
  - ✅ Users can change their rate, e.g. at night, while on cellular networks
  - ❌ Users cannot change their rate based on received traffic (inherent)

**Assumptions for Prior Work**

- **All** users send **one message** during **every** interval $R$.
  - ❌ Users cannot go offline.
  - ❌ Users cannot have different rates
  - ❌ Users cannot change their rate, e.g. at night, while on cellular networks.
  - ❌ Users cannot change their rate based on received traffic (inherent)

# What We Did

- ~~Leakage Analysis~~
- ~~New class of anonymity system.~~
  - ✅ Provably resists traffic analysis
  - ✅ Under weak assumptions
  - ✅ With low costs
- **Sparta: securely implements this deferred retrieval using Intel SGX.**
  - Scalable
  - Usable
  - Deployable

# Intel SGX

- Hardware-based trusted execution environment.
- Guarantees
  - Isolation — establishes region of memory accessible only by an enclave.
  - Attestation — enclave is running expected code.

**Side Channels**

**Access Patterns**

```
let x = data[secret];
```

**Control Flow**

```
if secret {
    stuff();
} else {
    other_stuff();
}
```

```
for x in 0..secret {
    stuff();
}
```

**Solution: SGX + Oblivious Algorithms**

# Sparta

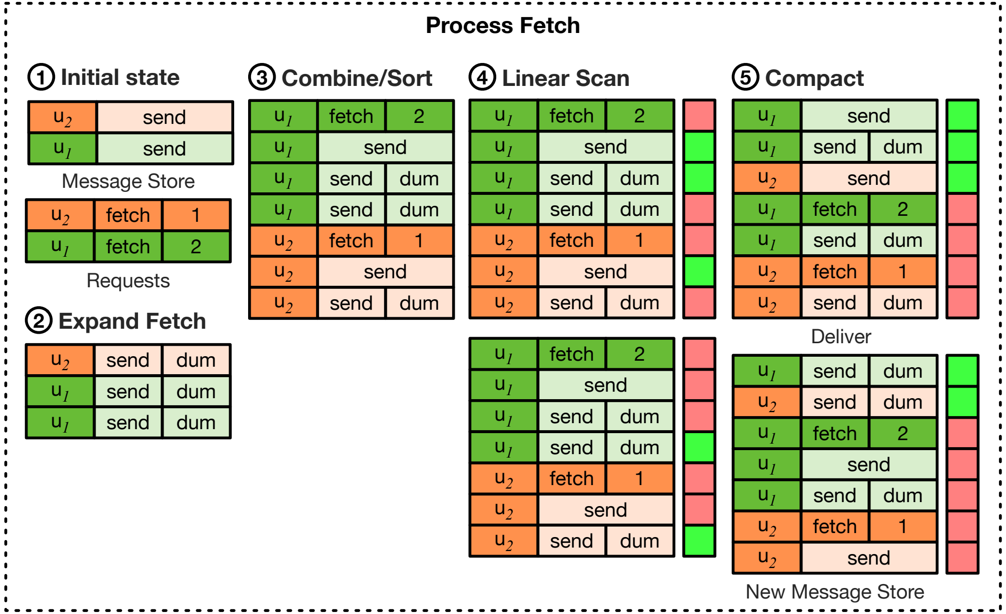
- A family of solutions implementing OMQS.
  - Sparta-LL — optimized for low-latency
  - Sparta-SB — optimized for high throughput
  - Sparta-D — optimized for high throughput in a distributed environment

# Sparta

- A family of solutions implementing OMQS.
  - Sparta-LL — optimized for low-latency
  - **Sparta-SB — optimized for high throughput**
  - Sparta-D — optimized for high throughput in a distributed environment

# Sparta-SB

- Based on oblivious sort and oblivious compaction.
- Send$(u_i, m)$: appends a message into the state
- Fetch$(\{u_i, k_i\})$:



# Experiments

**Experiment 1:**

- How do our relaxed assumptions affect performance under real workloads?

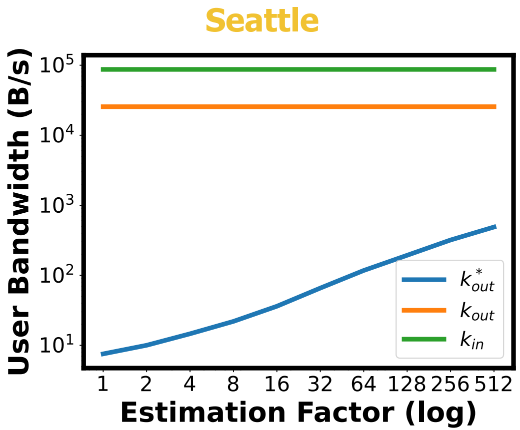
**Experiment 2:**

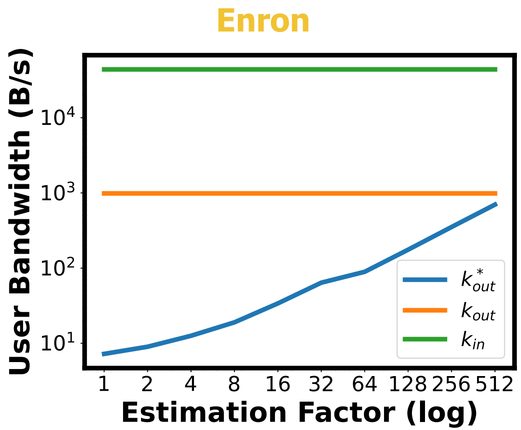
- How does Sparta perform as the database and compute scale?

# Experiment 1 Results

- Our systems rely on estimations of a user's download rate.
    - Existing work — network overhead under optimal download rates?
    - Sparta only — how do imperfect estimations affect network overhead?

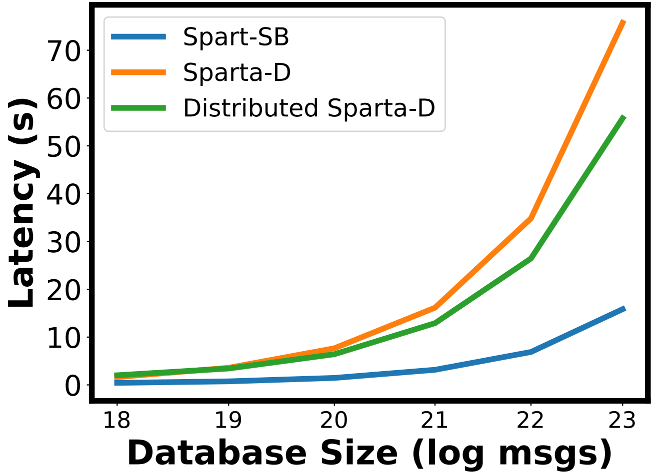
- **3400x reduction in overhead for the same latency (optimal)**
- **710x reduction in overhead for the same latency (estimate OOM)**

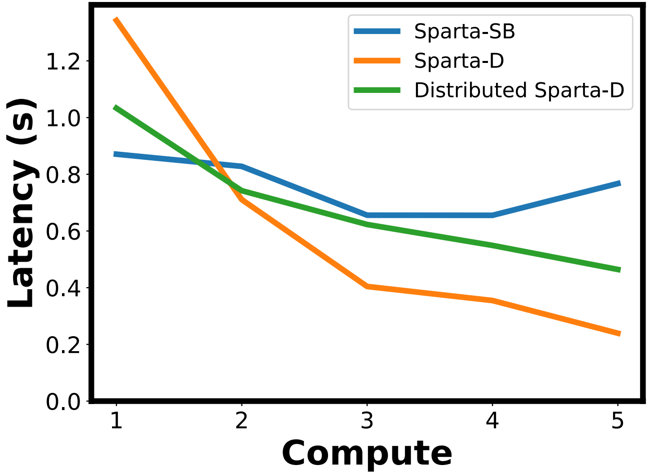


# Experiment 2 Results

- **15x improvement over prior fastest work.**
- Experiment 2.1 — scaling up the size of the database state.
- Experiment 2.2 — scaling up the amount of compute allocated to the systems.



# Conclusion

- Contributions
    - We formalized traffic analysis resistance.
    - Deferred retrieval leads to orders of magnitude cheaper systems.
    - Our implementations are an order of magnitude faster.

- Sparta is deployable.
- Sparta is usable.

# News & Upcoming Work

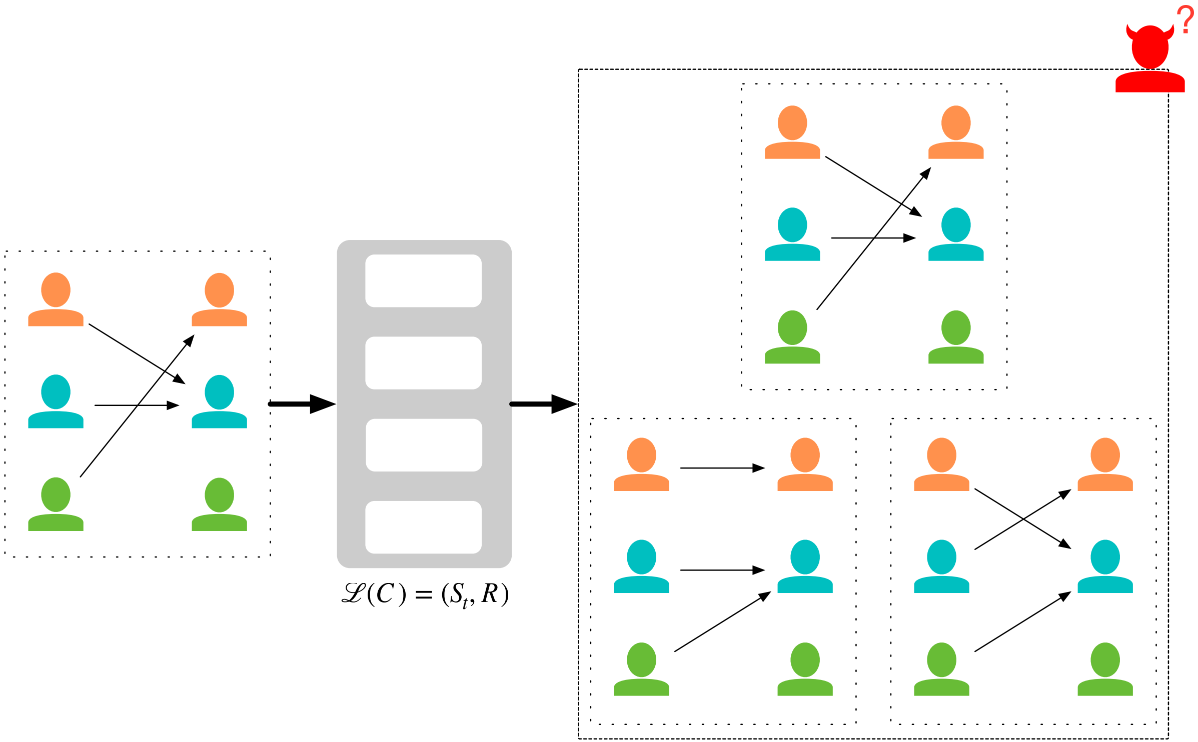
- Sparta was accepted!
  - IEEE Security & Privacy (Oakland) 2025
- SoK: The Traffic Analysis and Performance of Anonymous Communication Systems
  - Submitting to Oakland tomorrow

- Under Construction
  - Synchronous Systems are Dead; Long Live the Asynchronous
  - Raptor: Recipient Adjustable Padding for Traffic Analysis Resistance
  - Graduating in Spring 2025

# Thanks for Listening!

**Kyle Fredrickson**
kyfredri@ucsc.edu

# Leakage Analysis




$\mathscr{L}(C) = (S_t, R)$