



Vanish: Increasing Data Privacy with Self-Destructing Data

Roxana Geambasu

Yoshi Kohno

Amit Levy

Hank Levy

University of Washington



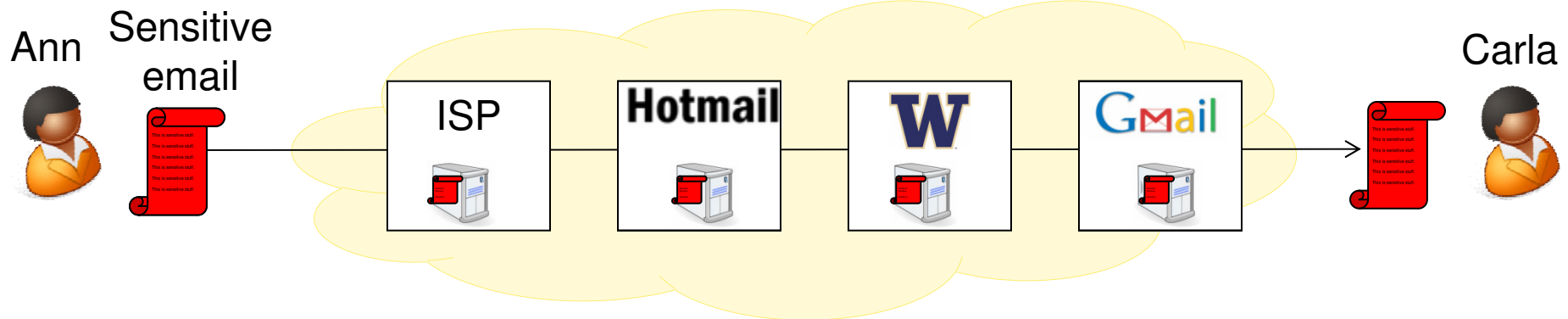
Outline

Part 1: Introducing Self-Destructing Data

Part 2: Vanish Architecture and Implementation

Part 3: Evaluation and Applications

Motivating Problem: Data Lives Forever



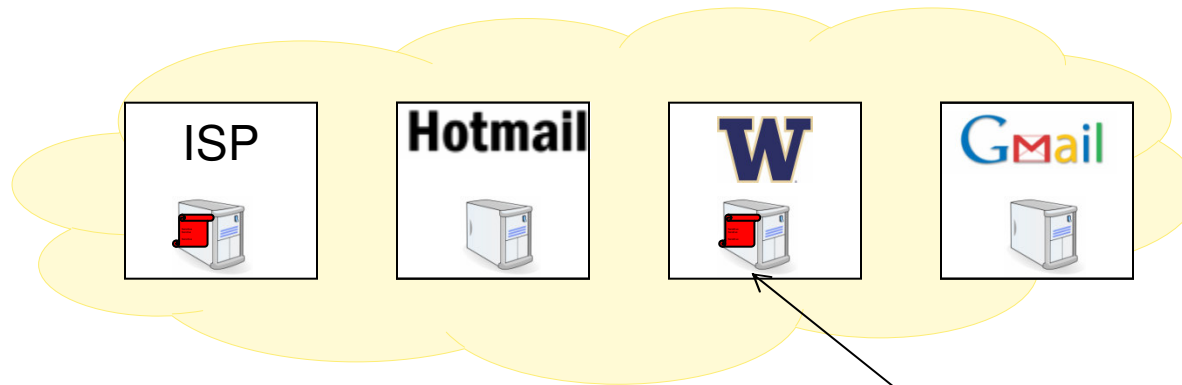
How can Ann delete her sensitive email?

- She doesn't know where all the copies are
- Services may retain data for long after user tries to delete

The screenshot shows the Ars Technica logo at the top left. Below it, the text "ars technica" is displayed. On the right side, it says "Last updated July 3, 2009". The main title of the article is "Are 'deleted' photos really gone from Facebook? Not always", with the word "Not" highlighted in yellow. The first sentence of the article reads: "When you delete embarrassing photos from sites like MySpace and Facebook, they don't disappear immediately."

Archived Copies Can Resurface Years Later

Ann



Carla



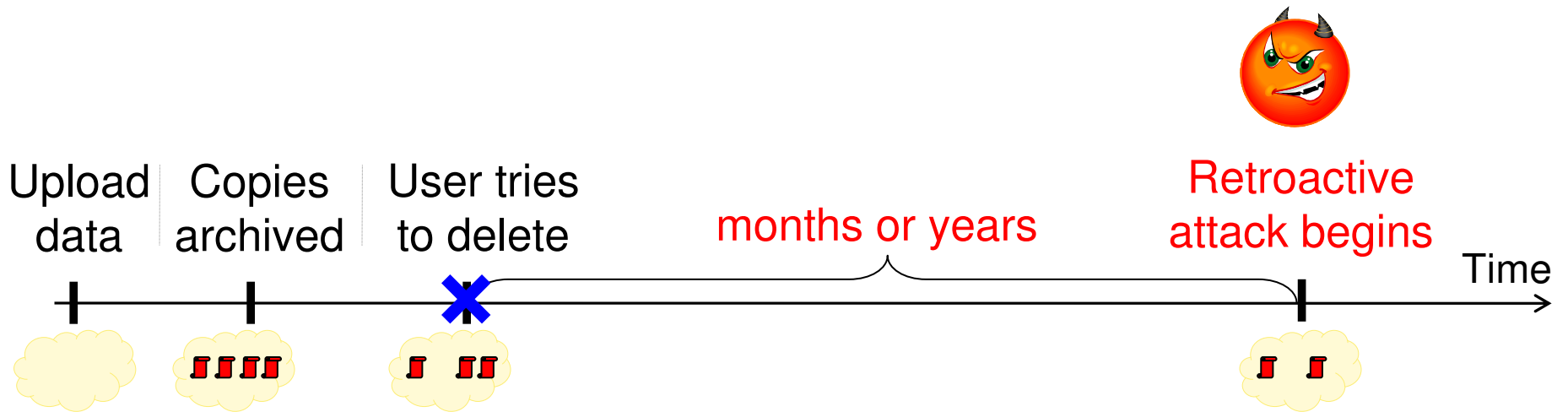
Some time later...

Subpoena,
hacking, ...

**Retroactive attack
on archived data**

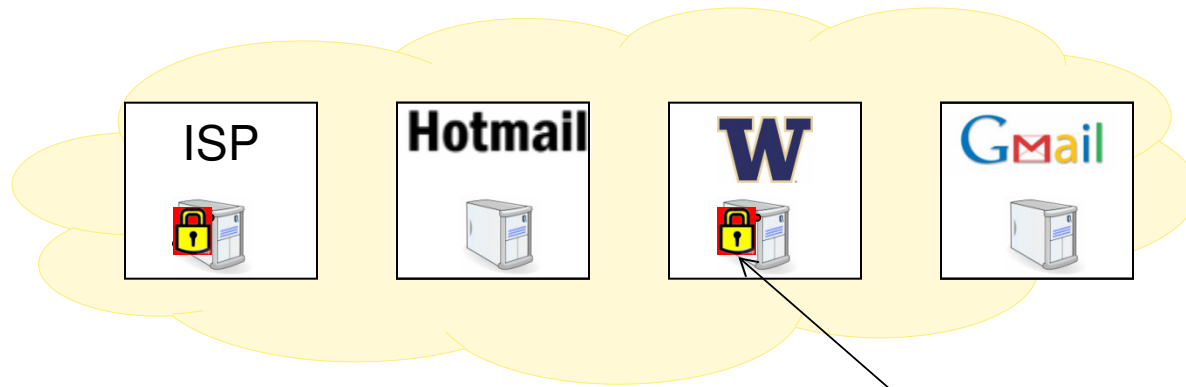


The Retroactive Attack



Why Not Use Encryption (e.g., PGP)?

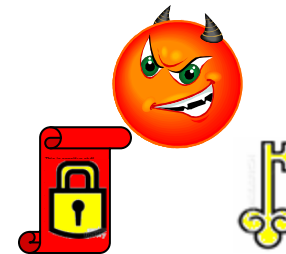
Ann



Carla

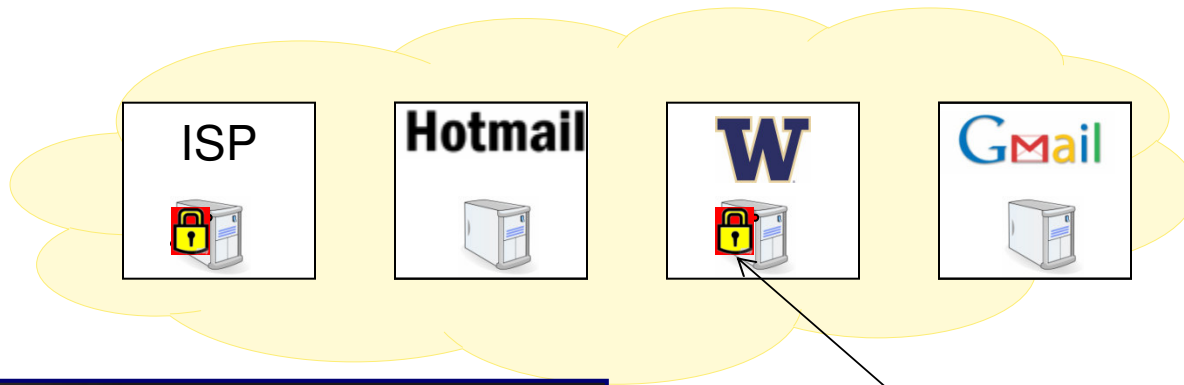


Subpoena,
hacking, ...



Why Not Use Encryption (e.g., PGP)?

Ann



Carla



Subpoena,
hacking, ...



cnet news
February 26, 2009 1:30 PM PST

Judge orders defendant to decrypt PGP-protected laptop

A federal judge has ordered a criminal defendant to decrypt his hard drive by typing in a ruling

v3.co.uk formerly vnunet.com

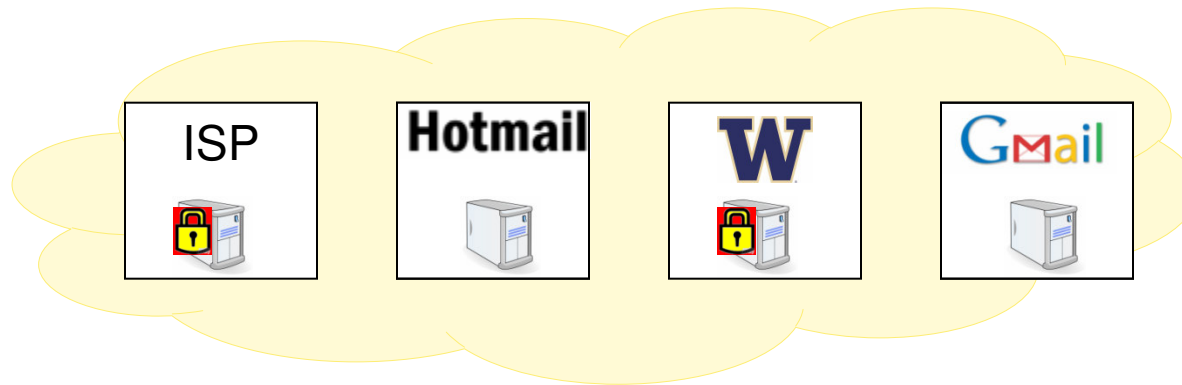
UK police can now demand encryption keys

vnunet.com, 03 Oct 2007

People in the UK who encrypt their data are now obliged by law to give up the encryption keys to law enforcement officials if requested under the Regulation of Investigatory Powers Act 2000 (RIP Act).


Why Not Use a Centralized Service?

Ann



Carla



Centralized Service	
	<p>“Trust us: we’ll help you delete your data on time.”</p>



Why Not Use a Centralized Service?

Ann



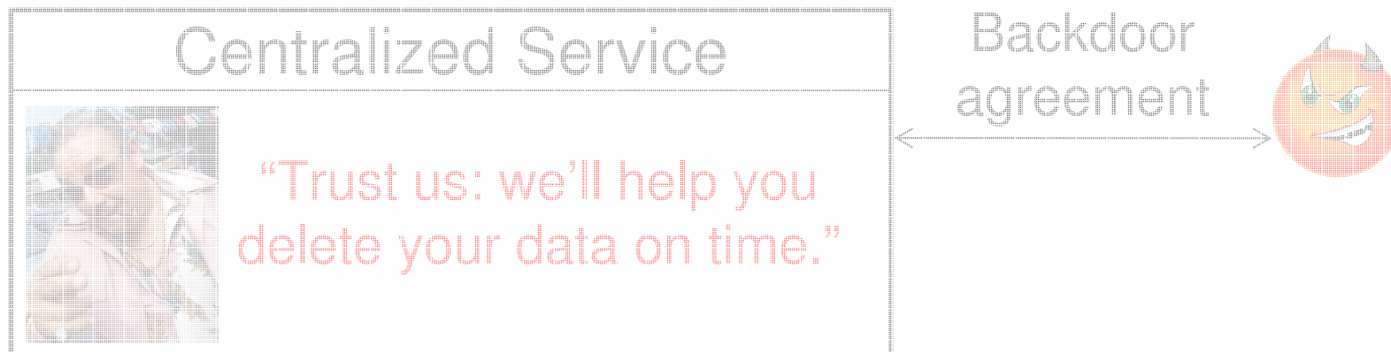
WIRED November 7, 2007 | 3:39 pm

Encrypted E-Mail Company Hushmail Spills to Feds

Hushmail, a longtime provider of encrypted web-based email, markets itself by saying that "not even a Hushmail employee with access to our servers can read your encrypted e-mail, since each message is uniquely encoded before it leaves your computer."

But it turns out that statement seems not to apply to individuals targeted by government agencies that are able to convince a Canadian court to serve a court order on the company.

Carla



The Problem: Two Huge Challenges for Privacy

1. Data lives forever

- On the web: emails, Facebook photos, Google Docs, blogs, ...
- In the home: disks are cheap, so no need to ever delete data
- In your pocket: phones and USB sticks have GBs of storage

2. Retroactive disclosure of both data and user keys has become commonplace

- Hackers
- Misconfigurations
- Legal actions
- Border seizing
- Theft
- Carelessness



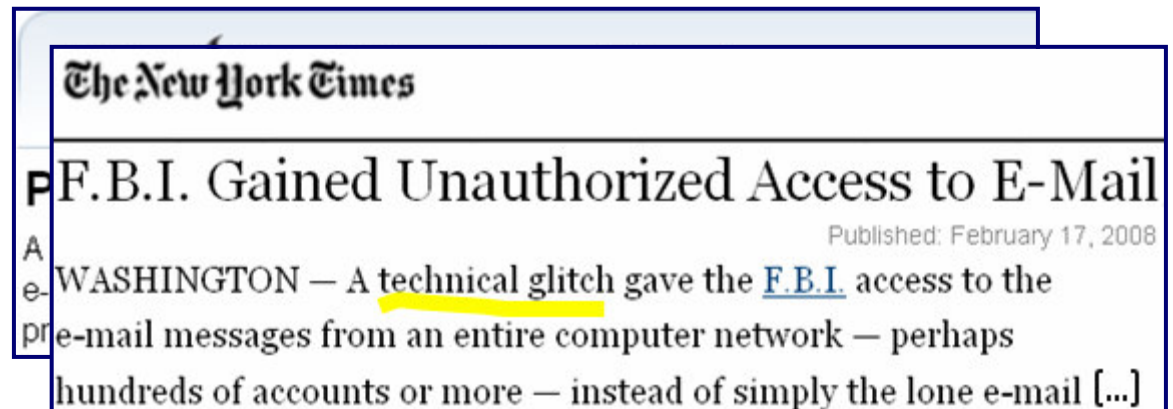
The Problem: Two Huge Challenges for Privacy

1. Data lives forever

- On the web: emails, Facebook photos, Google Docs, blogs, ...
- In the home: disks are cheap, so no need to ever delete data
- In your pocket: phones and USB sticks have GBs of storage

2. Retroactive disclosure of both data and user keys has become commonplace

- Hackers
- Misconfigurations
- Legal actions
- Border seizing
- Theft
- Carelessness



The Problem: Two Huge Challenges for Privacy

1. Data lives forever

- On the web: emails, Facebook photos, Google Docs, blogs, ...
- In the home: disks are cheap, so no need to ever delete data
- In your pocket: phones and USB sticks have GBs of storage

2. Retroactive disclosure of both data and user keys has become commonplace

- Hackers
- Misconfigurations
- Legal actions
- Border seizing
- Theft
- Carelessness



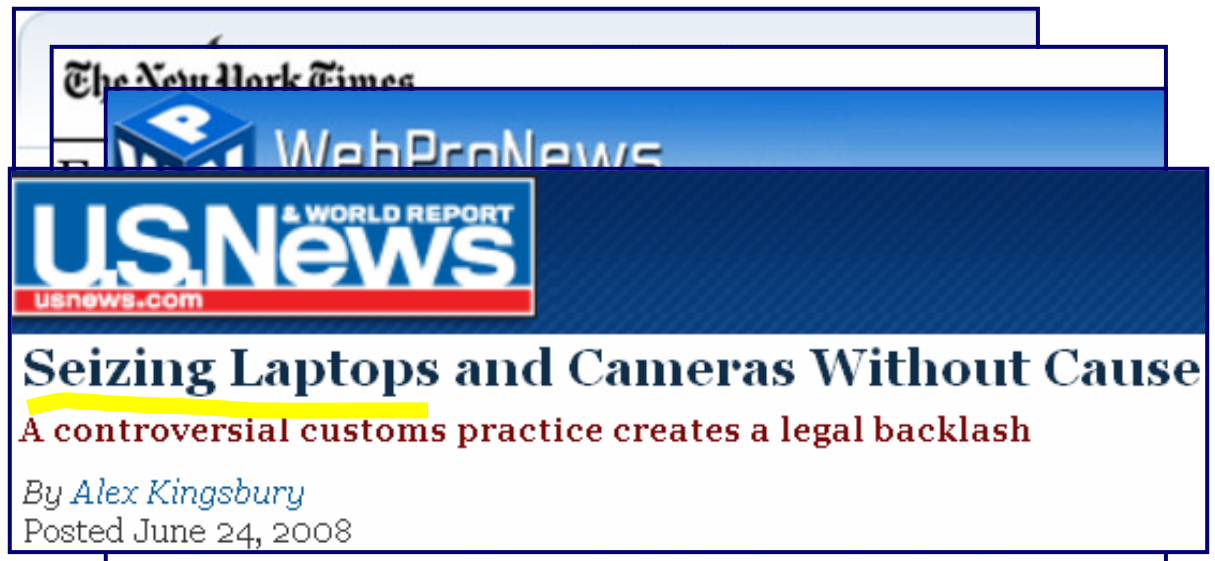
The Problem: Two Huge Challenges for Privacy

1. Data lives forever

- On the web: emails, Facebook photos, Google Docs, blogs, ...
- In the home: disks are cheap, so no need to ever delete data
- In your pocket: phones and USB sticks have GBs of storage

2. Retroactive disclosure of both data and user keys has become commonplace

- Hackers
- Misconfigurations
- Legal actions
- Border seizing
- Theft
- Carelessness

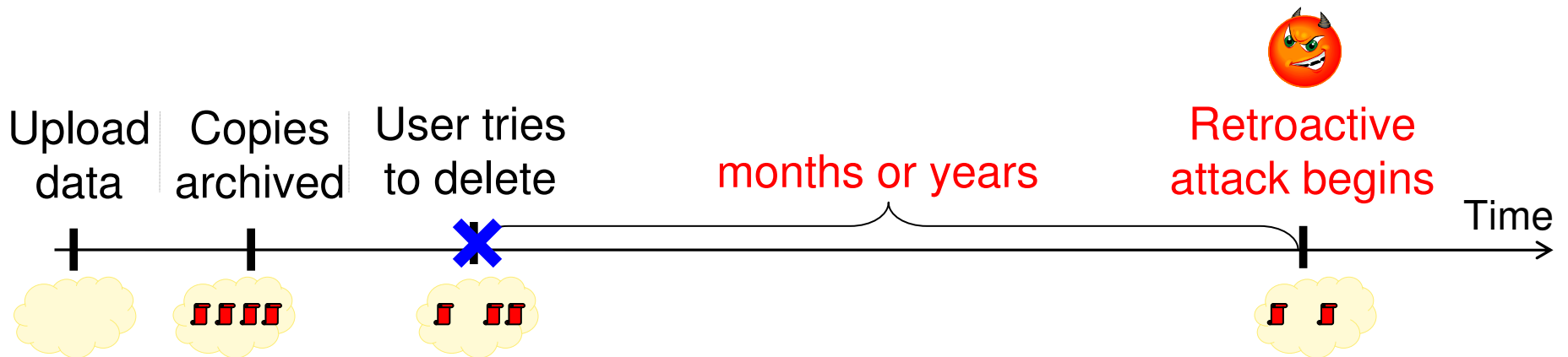


Question:

Can we empower users with control of data lifetime?

Answer:

Self-destructing data

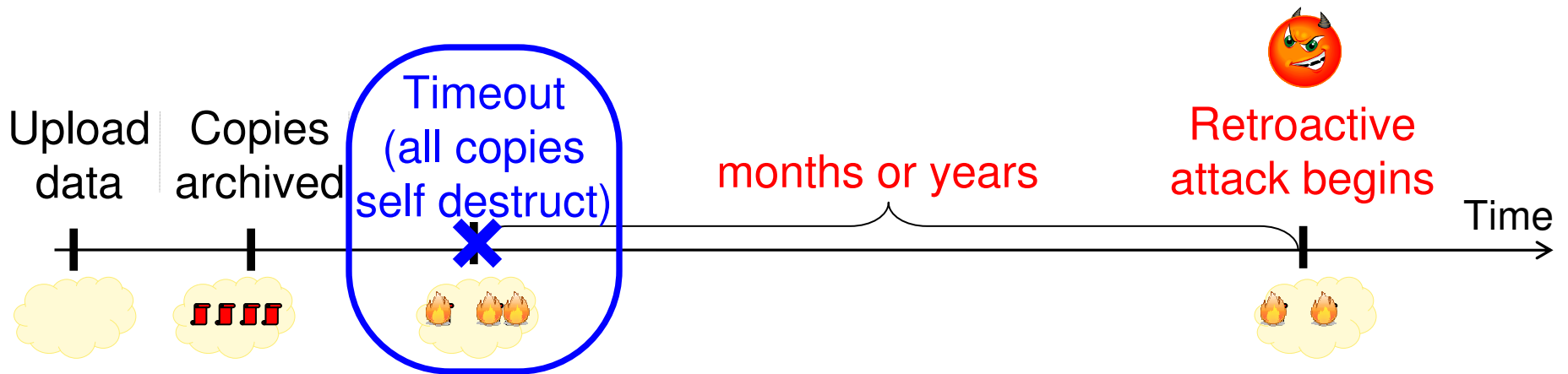


Question:

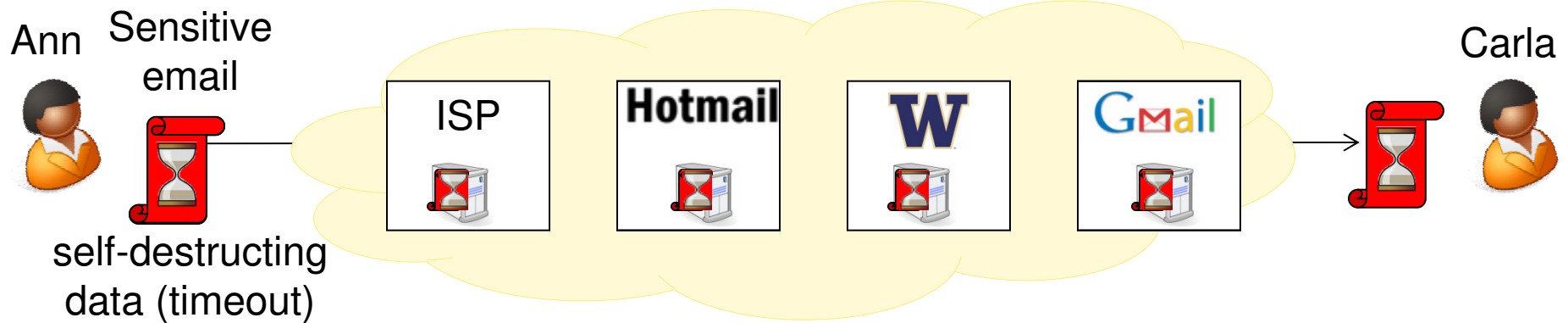
Can we empower users with control of data lifetime?

Answer:

Self-destructing data

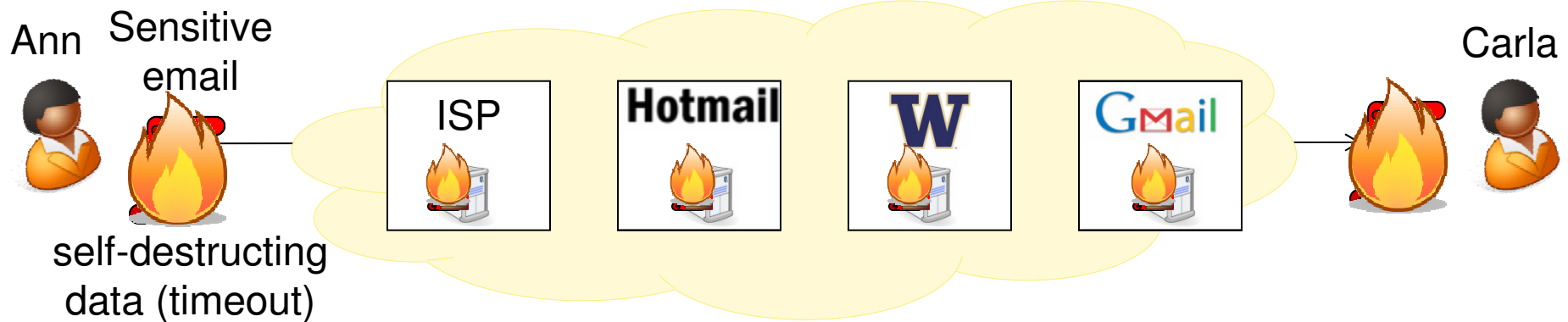


Self-Destructing Data Model



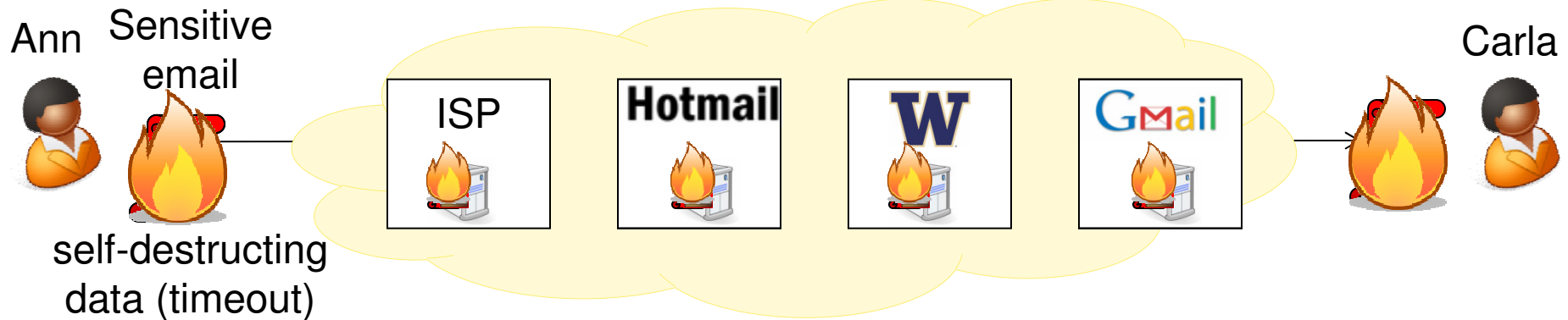
1. Until timeout, users can read original message

Self-Destructing Data Model



1. Until timeout, users can read original message
2. After timeout, **all copies** become **permanently unreadable**
 - 2.1. even for attackers who obtain an **archived copy** & **user keys**
 - 2.2. without requiring **explicit delete action** by user/services
 - 2.3. without having to trust **any centralized services**

Self-Destructing Data Model



Goals of Self-Destructing Data

1. Until timeout, users can read original message
2. After timeout, **all copies** become **permanently unreadable**
 - 2.1. even for attackers who obtain an **archived copy** & **user keys**
 - 2.2. without requiring **explicit delete action** by user/services
 - 2.3. without having to trust **any centralized services**



Outline

Part 1: Introducing Self-Destructing Data

Part 2: [Vanish Architecture and Implementation](#)

Part 3: Evaluation and Applications



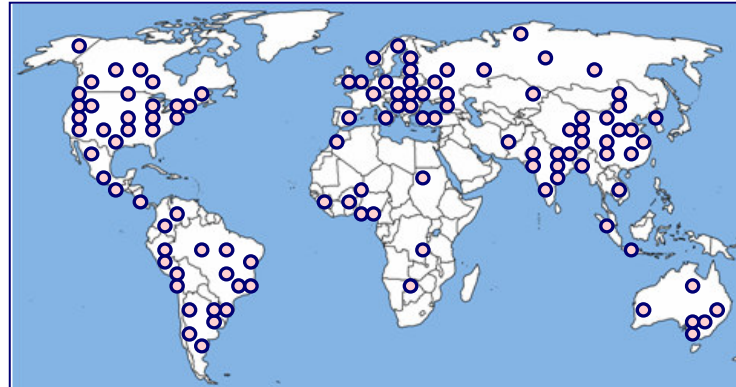
Vanish: Self-Destructing Data System

- Traditional solutions are not sufficient for self-destructing data goals:
 - PGP
 - Centralized data management services
 - Forward-secure encryption
 - ...
- Let's try something completely new!

Idea:
Leverage P2P systems

P2P 101: Intro to Peer-To-Peer Systems

- A system composed of individually-owned computers that make a portion of their resources available directly to their peers without intermediary managed hosts or servers. [~wikipedia]

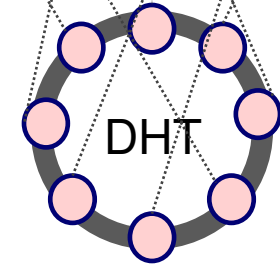
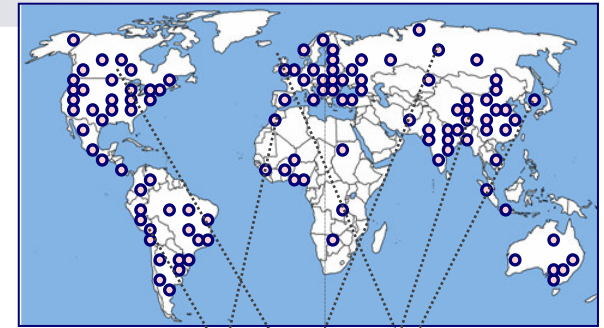


Important P2P properties (for Vanish):

- **Huge scale** – millions of nodes
- **Geographic distribution** – hundreds of countries
- **Decentralization** – individually-owned, no single point of trust
- **Constant evolution** – nodes constantly join and leave

Distributed Hashtables (DHTs)

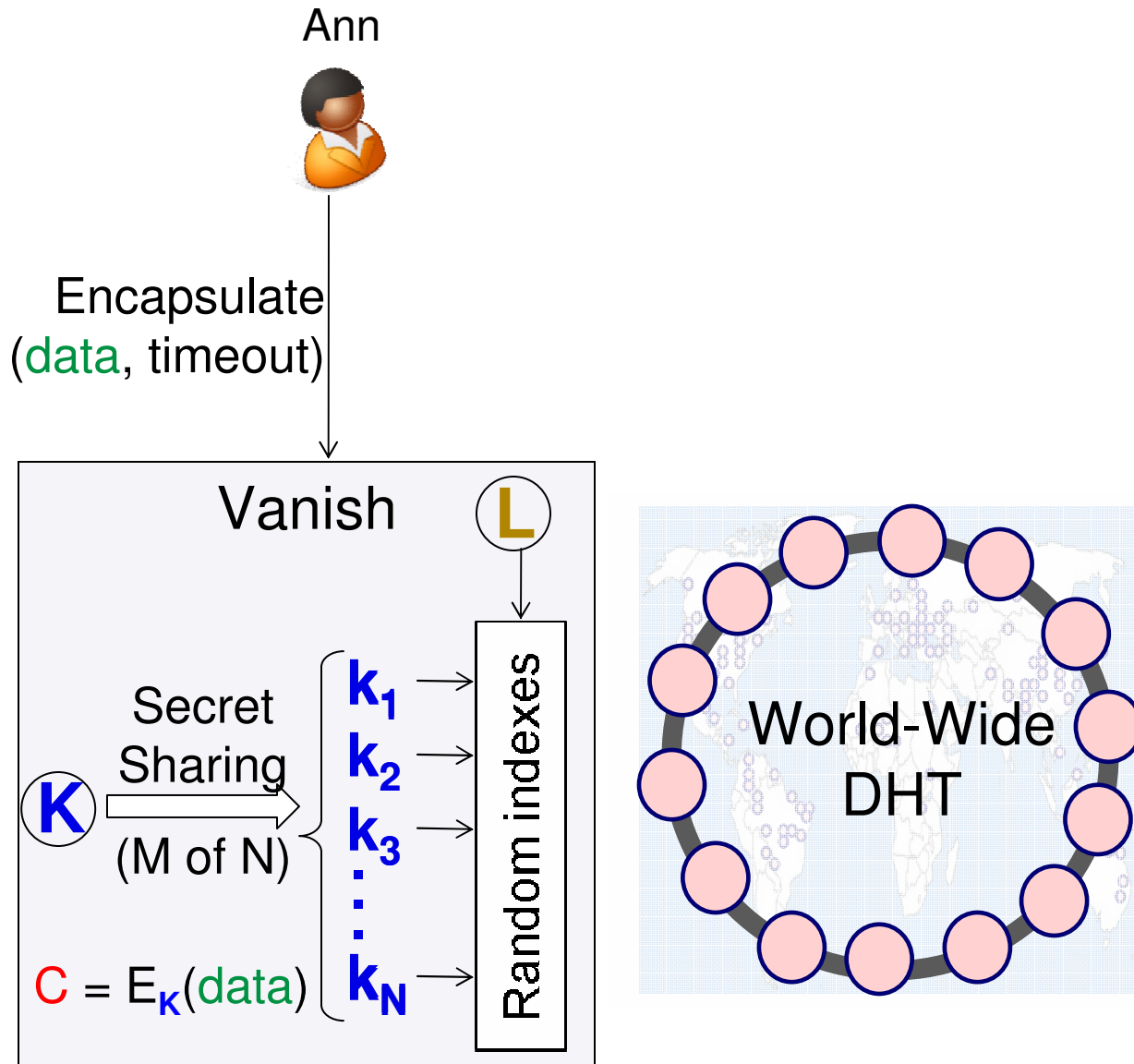
- Hashtable data structure implemented on a P2P network
 - Get and put (index, value) pairs
 - Each node stores part of the index space



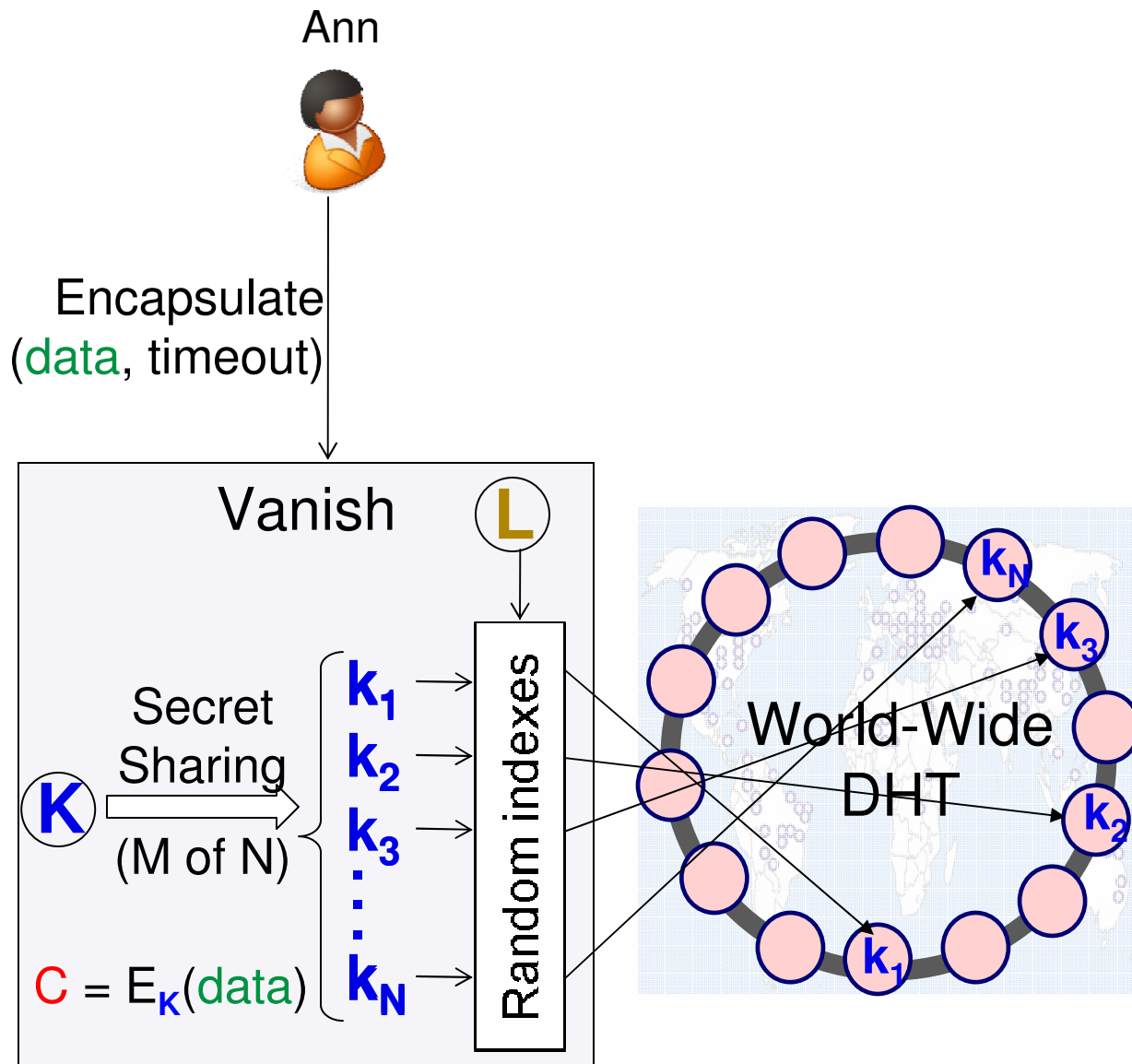
Logical structure

- DHTs are part of many file sharing systems:
 - Vuze, Mainline, KAD
 - Vuze has ~1.5M simultaneous nodes in ~190 countries
- **Vanish leverages DHTs to provide self-destructing data**
 - One of few applications of DHTs outside of file sharing

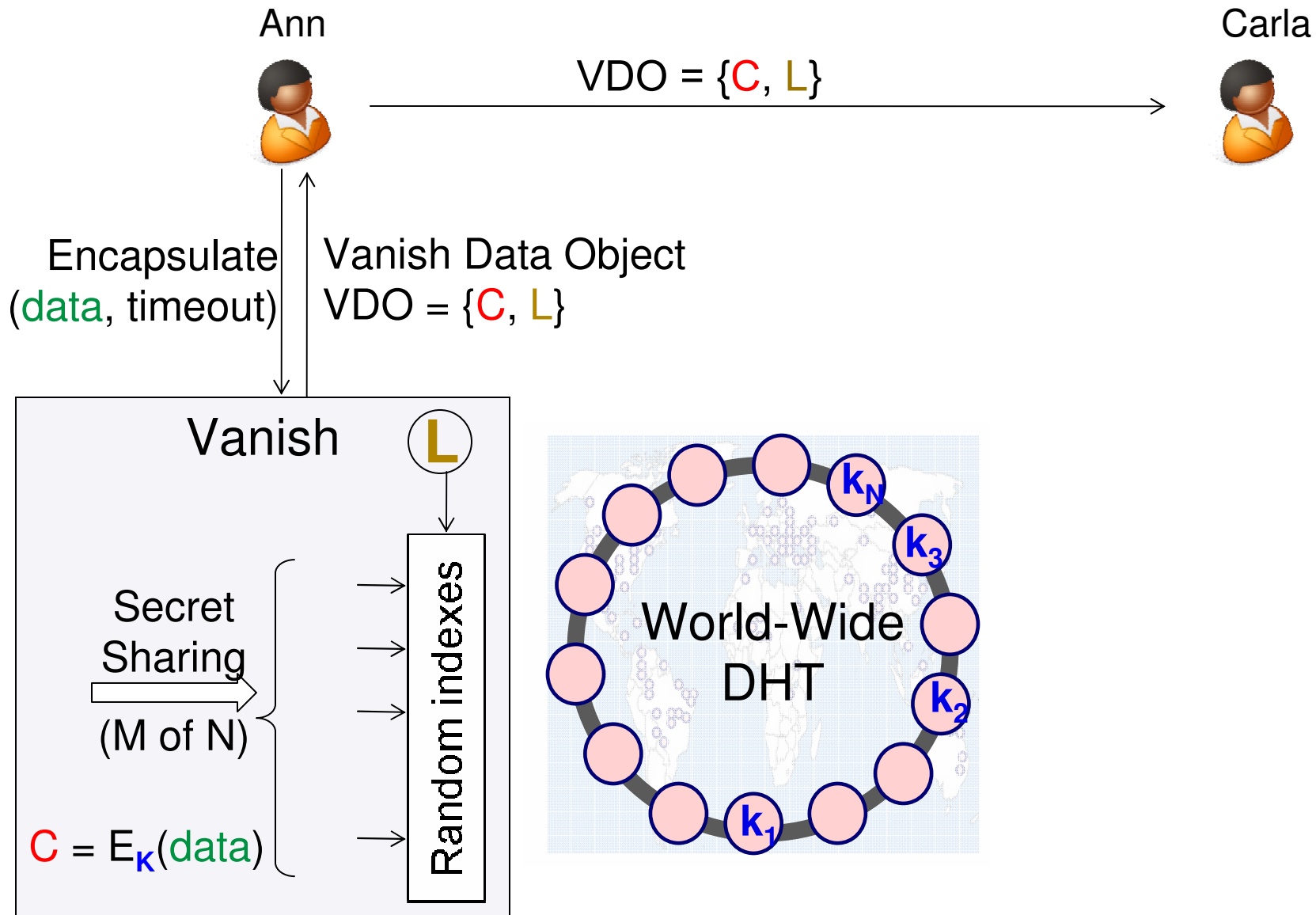
How Vanish Works: Data Encapsulation



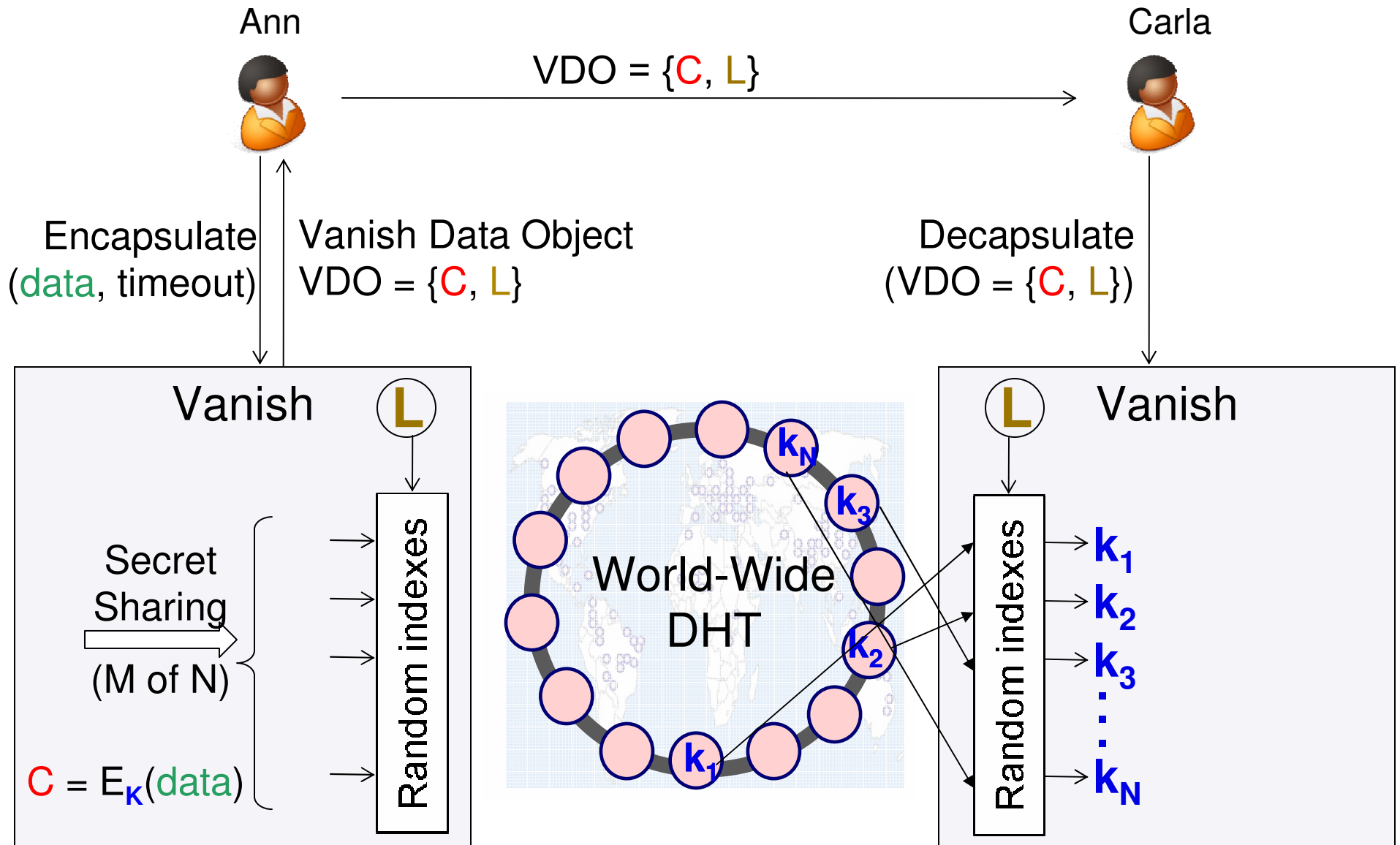
How Vanish Works: Data Encapsulation



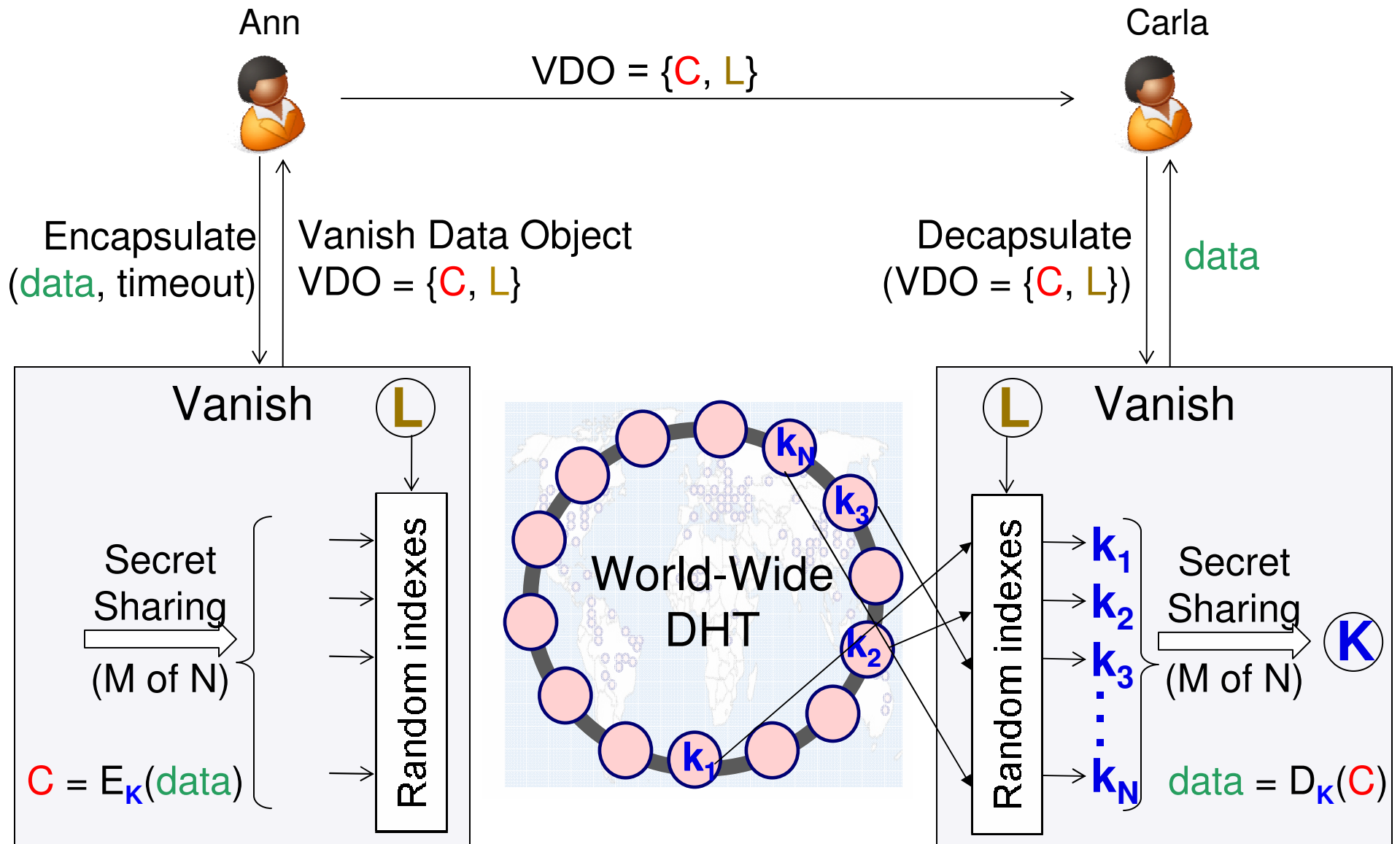
How Vanish Works: Data Encapsulation



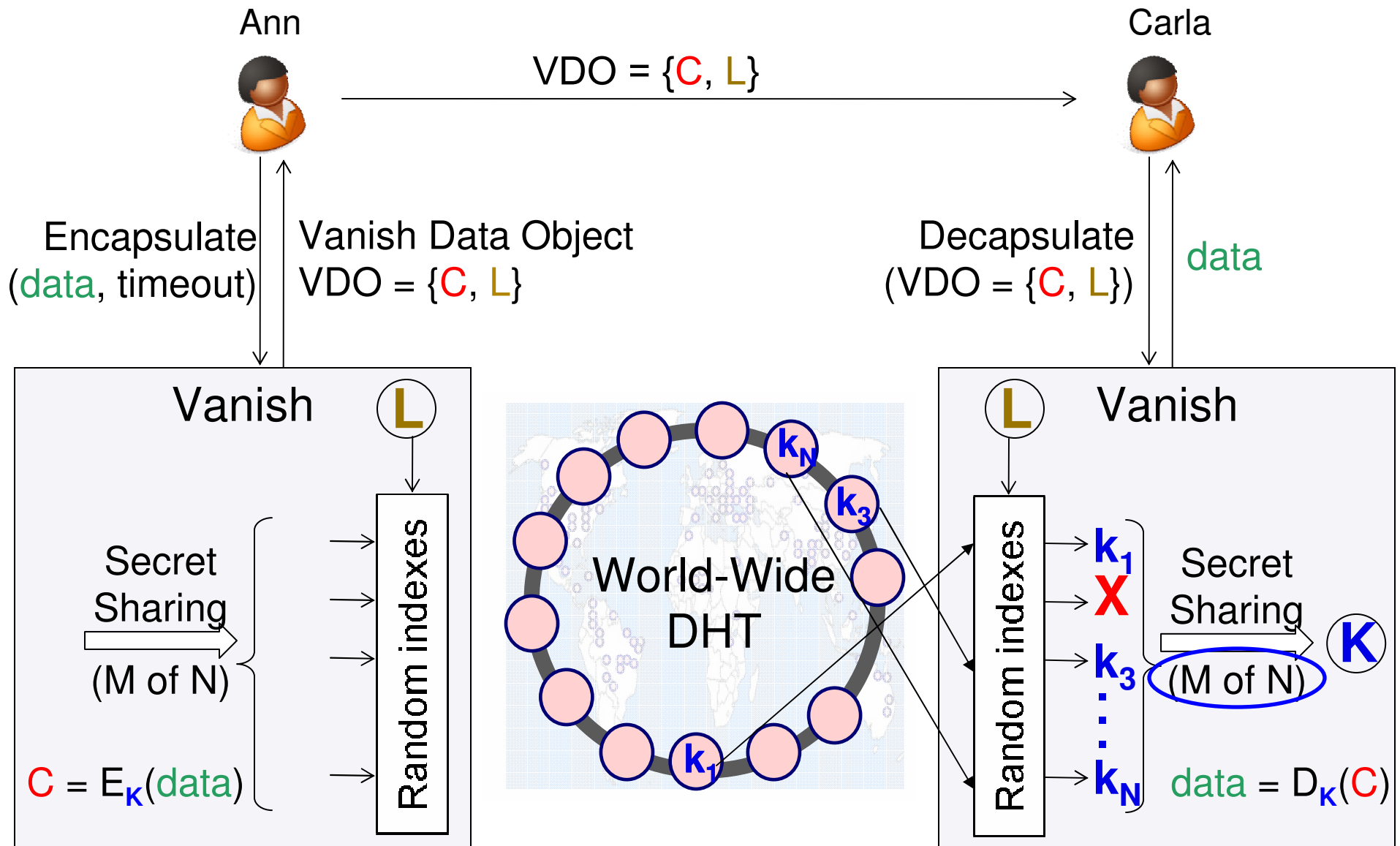
How Vanish Works: Data Decapsulation



How Vanish Works: Data Decapsulation

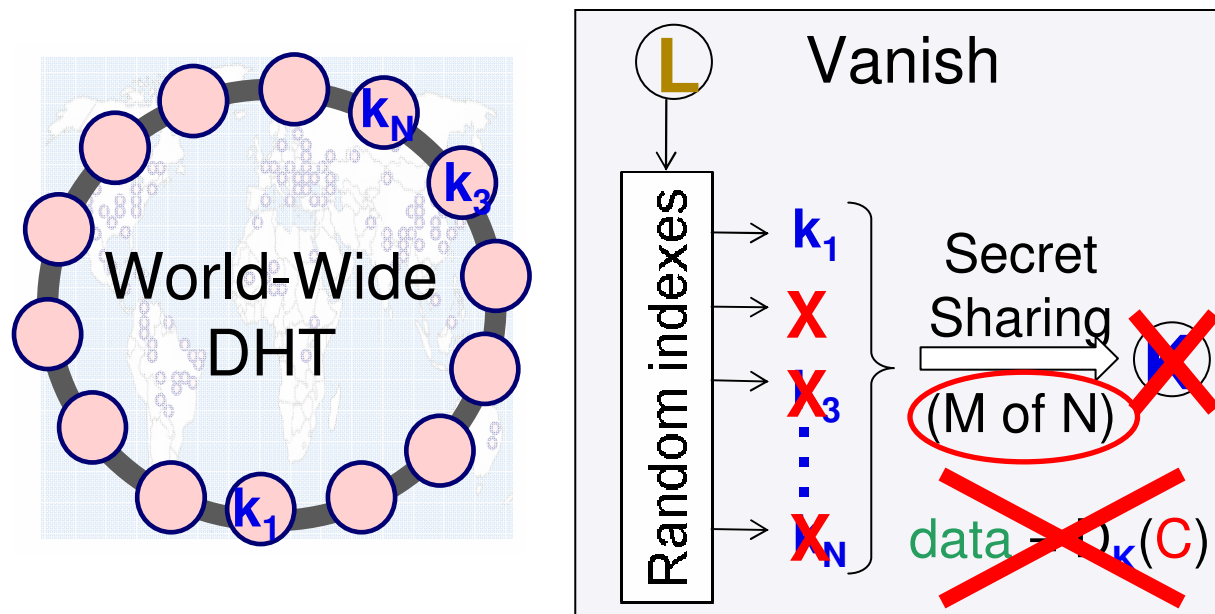


How Vanish Works: Data Decapsulation



How Vanish Works: Data Timeout

- The DHT **loses key pieces** over time
 - Natural churn: nodes crash or leave the DHT
 - Built-in timeout: DHT nodes purge data periodically



- Key loss makes all data copies **permanently unreadable**



Outline

Part 1: Introducing Self-Destructing Data

Part 2: Vanish Architecture and Implementation

Part 3: Evaluation and Applications

Evaluation

- Experiments to understand and improve:

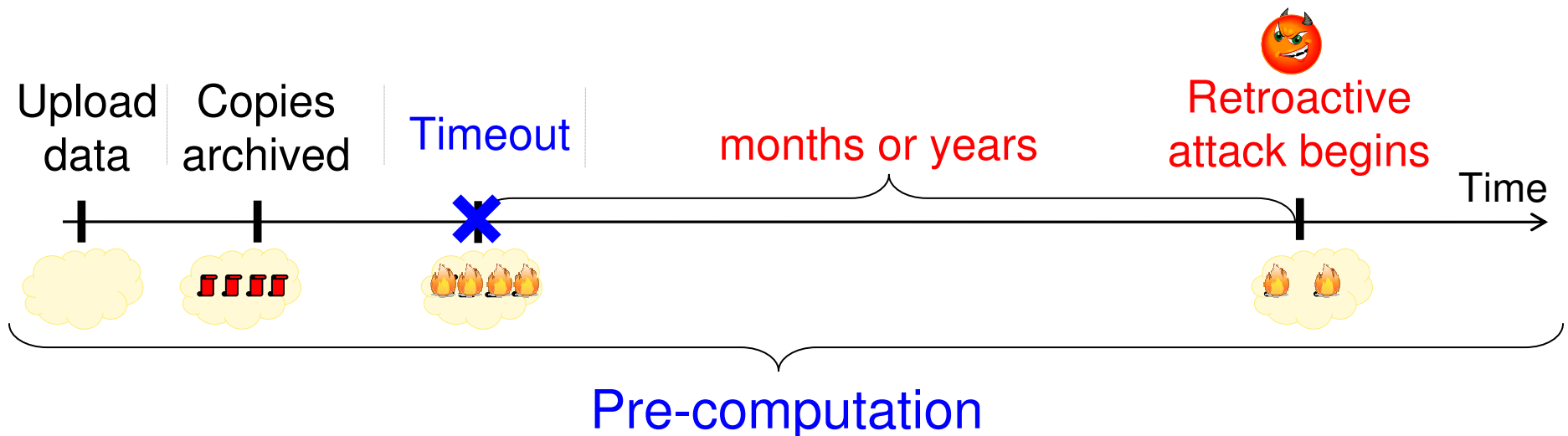
1. data availability before timeout
 2. data unavailability after timeout
 3. performance
 4. security
- } In the paper
- } Discussed next

- Highest-level results:

- **Secret sharing** parameters (N and M) affect availability, timeout, performance, and security
- **Tradeoffs** are necessary

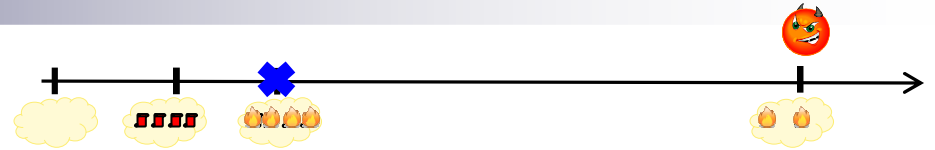
Threat Model

- Goal: protect against **retroactive attacks** on old copies
 - Attackers don't know their target until after timeout
 - Attackers may do non-targeted “**pre-computations**” at any time



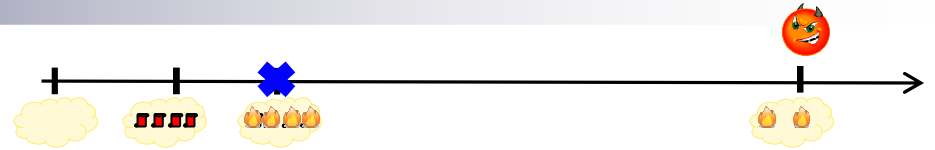
- Communicating parties trust each other
 - E.g., Ann trusts Carla not to keep a plain-text copy

Attack Analysis

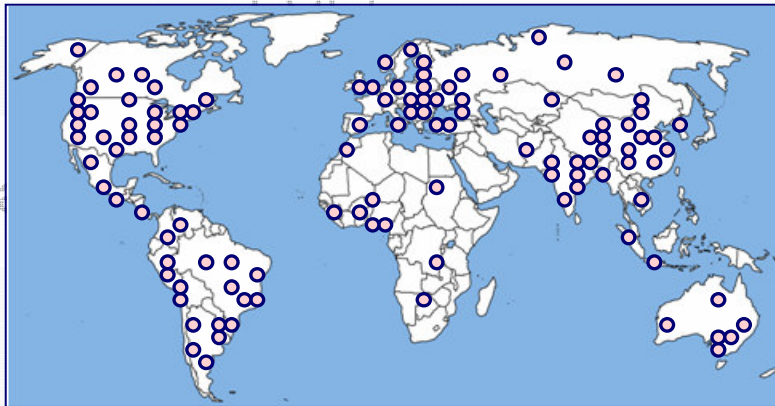


Retroactive Attack	Defense
Obtain data by legal means (e.g., subpoenas)	P2P properties: constant evolution , geographic distribution , decentralization
Gmail decapsulates all VDO emails	Compose with traditional encryption (e.g., PGP)
ISP sniffs traffic	Anonymity systems (e.g., Tor)
DHT eclipse, routing attack	Defenses in DHT literature (e.g., constraints on routing table)
DHT Sybil attack	Defenses in DHT literature; Vuze offers some basic protection
Intercept DHT “get” requests & save results	Vanish obfuscates key share lookups
Capture key pieces from the DHT (pre-computation)	P2P property: huge scale
More (see paper)	

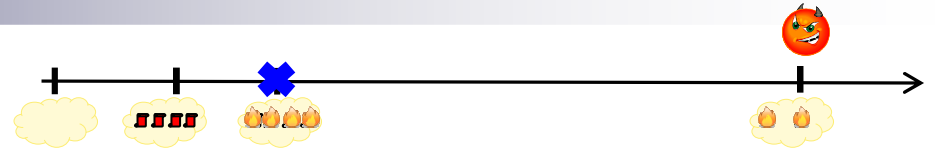
Attack Analysis



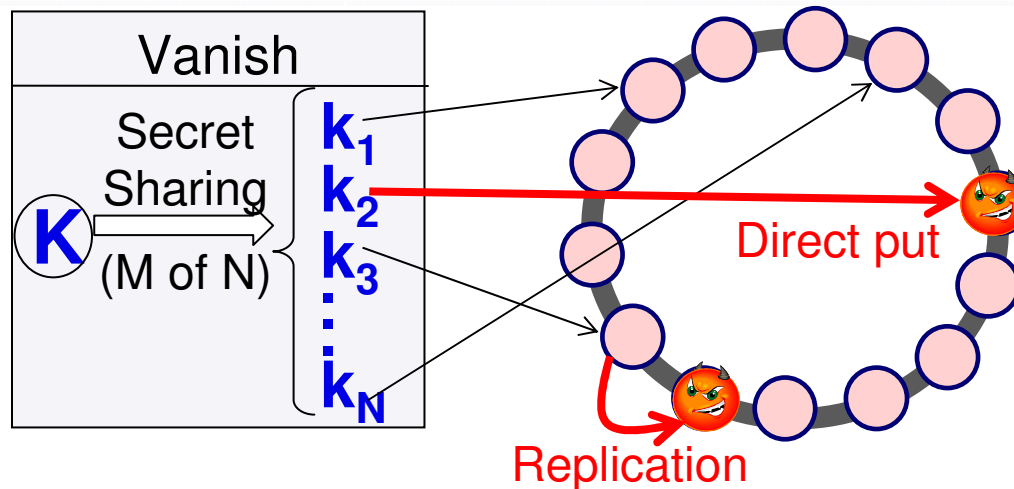
Retroactive Attack	Defense
Obtain data by legal means (e.g., subpoenas)	P2P properties: constant evolution , geographic distribution , decentralization
Gmail decapsulates all VDO emails	Compose with traditional encryption (e.g., PGP)
ISP sniffs traffic	g., Tor)
DHT eclipse, routing	ture (e.g., constraints
DHT Sybil attack	ture; Vuze offers
Intercept DHT "get" requests & save results	Vanish obfuscates key share lookups
Capture key pieces from the DHT (pre-computation)	P2P property: huge scale
More (see paper)	



Retroactive Attacks



Attack	Defense
Capture any key pieces from the DHT (pre-computation)	P2P property: huge scale



- Given the huge DHT scale, how many nodes does the attacker need to be effective?
- Current estimate:
 - Attacker must join with ~8% of DHT size, for 25% capture
 - There may be other attacks (and defenses)



Vanish Applications

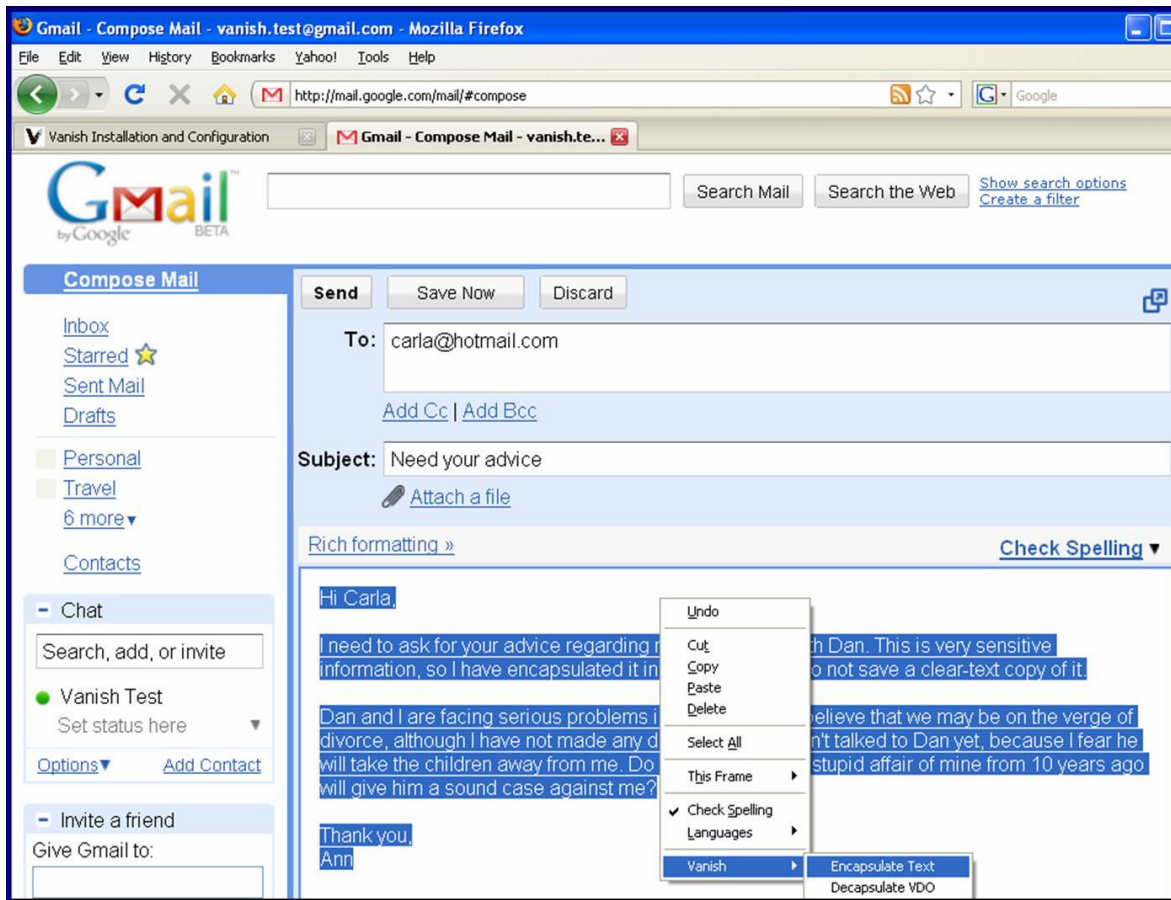
- Self-destructing data & Vanish support many applications

Example applications:

- **Firefox plugin**
 - Included in our release of Vanish
- Thunderbird plugin
 - Developed by the community two weeks after release 😊
- Self-destructing files
- Self-destructing trash-bin
- ...

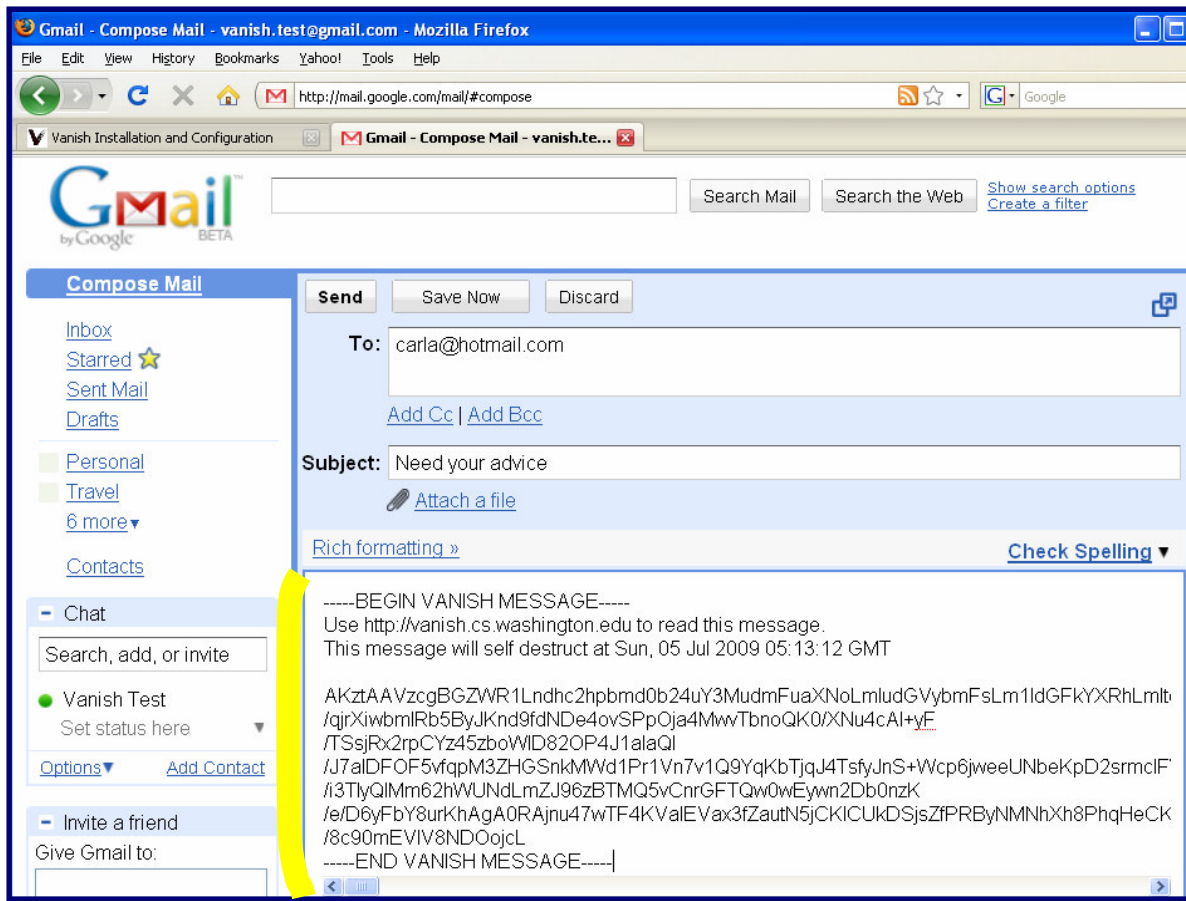
Firefox Plugin For Vanishing Web Data

- Encapsulate text in **any text area** in self-destructing VDOs



Firefox Plugin For Vanishing Web Data

- Encapsulate text in **any text area** in self-destructing VDOs



Firefox Plugin For Vanishing Web Data

- Encapsulate text in **any text area** in self-destructing VDOs

The image displays four overlapping Firefox browser windows, each showing the 'Vanish' plugin interface. The windows are:

- Gmail - Compose Mail**: Shows the 'Compose Mail' sidebar with a 'Vanish Test' button highlighted by a yellow arrow.
- Windows Live Hotmail**: Shows the 'Hotmail' interface with a 'Vanish Concepts and Architecture' sidebar and a yellow arrow pointing to the 'Vanish Test' button.
- Divorce Document - Google Docs**: Shows a document titled 'Divorce Document' with a 'Vanish Concepts and Architecture' sidebar and a yellow arrow pointing to the 'Vanish Test' button.
- Facebook | Message: Need your advice**: Shows a Facebook message from 'Ann Gerobo' with a 'Vanish Concepts and Architecture' sidebar and a yellow arrow pointing to the 'Vanish Test' button.

The message content in the Facebook window is as follows:

-----BEGIN VANISH MESSAGE-----
Use <http://vanish.cs.washington.edu> to read this message.
This message will self destruct at Sun, 05 Jul 2009 06:21:18 GMT

AKztAAVzcgBGZWR1Lndhc2hpbm
d0b24uY3MudmbAgACsGAMZXBvY
2hWR1Lndhc2hpbmd0b24uY3Mud
mFuaXNoLmludGvYbmfLm1ldGF
kYXRhLmltcGwusW5kaXJlY3RlZ
XlNZXRhZGF0YUlteGw6bmc16f5
f7QIAAlsAEmVuY3J35c-RlZF9kY
XRhX2tleXQAAltCTAAIbWV0YWR
hdGFxAH4AAAXhwchNyeAEFZHJud
2FzaGluz3Rvbi5jcy52YW5pc2g
uaW50ZXJuuYwubWV0YWRhdGEua
W1wbC5CYXNpY01ldGFkYXRhSW1
wbNgVQUjt/E3XAgACsGANbG9Y
XRpb25fc2VlZEwABnBhcmFtc3Q
ANkoZlHvd2FzaGluz3Rvbi9jc
y92YW5pc2graW50ZXJuuYwubWV
DYWRhdGEvVkrPUGFyYW1zO3hw
sCcB1ldGFkYXRhLlZET1BhcmFt
c7292Mmle6MAGAlSgALY3JlYX
Rpb25fdHJABVlbnNyeXB0aW9u
X2tleV9sZW5ndGhJAApudW1fc2

Firefox Plugin For Vanishing Web Data

- Encapsulate text in **any text area** in self-destructing VDOs

Effect:

Vanish empowers users with seamless control over the lifetime of their Web data

The screenshot shows a Gmail 'Compose Mail' window in Mozilla Firefox. The 'Vanish' plugin interface is visible, including a 'Vanish Test' button and a 'Vanish' message being composed. The message content includes a header 'Need your advice Between You', a sender 'Ann Gerold', and a body with a URL and a long alphanumeric string. A blue callout box with white text is overlaid on the screenshot, stating 'Effect: Vanish empowers users with seamless control over the lifetime of their Web data'. Yellow arrows point from the callout box to the 'Vanish' message content in the screenshot.

Conclusions

- Two formidable challenges to privacy:
 - Data lives forever
 - Disclosures of data and keys have become commonplace
- **Self-destructing data** empowers users with lifetime control
- Vanish:
 - Combines global-scale **DHTs** with **secret sharing** to provide self-destructing data
 - Firefox plugin allows users to set timeouts on text data **anywhere on the web**
- Vanish ≠ Vuze-based Vanish
 - Customized DHTs, hybrid approach, other P2P systems
 - Further extensions for security in the paper