Increasing Data Privacy with Self-Destructing Data

Roxana Geambasu and Hank Levy, UW

with:

- Yoshi Kohno, Amit Levy, Arvind Krishnamurthy (UW)
- Vinnie Mascaritolo (PGP Corporation)
- Paul Gardner (Vuze, Inc.)

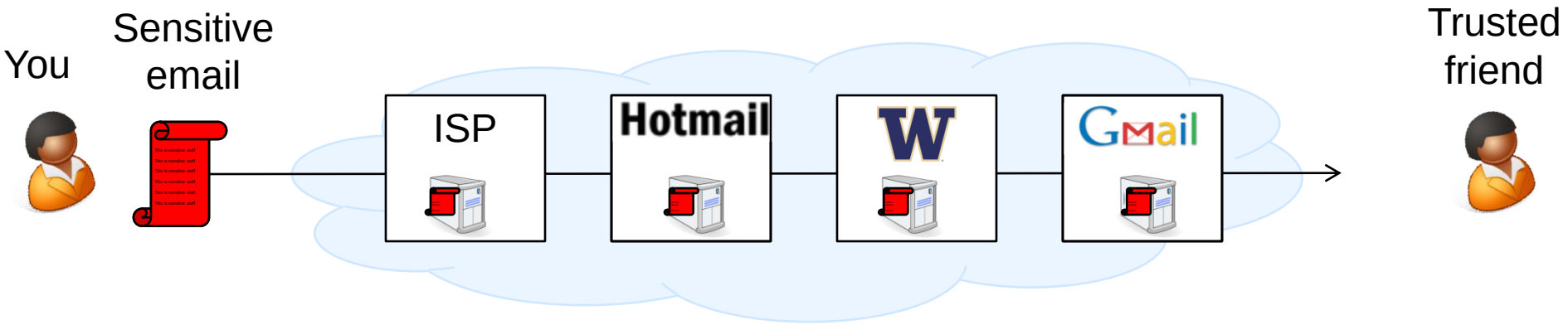
Outline

- Self-Destructing Data: Motivation and Goals
- Implementing a Decentralized Self-Destructing Data System

Data Lives Forever

- Huge disks have eliminated the need to ever delete data
  - Desktops store TBs of historical data
  - Phones, USB drives store GBs of personal data in your pocket
  - Data centers keep data forever
- The Web and cloud computing has made it impossible to delete data
  - Users have no direct control of their data
  - Web services are highly replicated, archival stores
  - Data has value, services want to mine that value

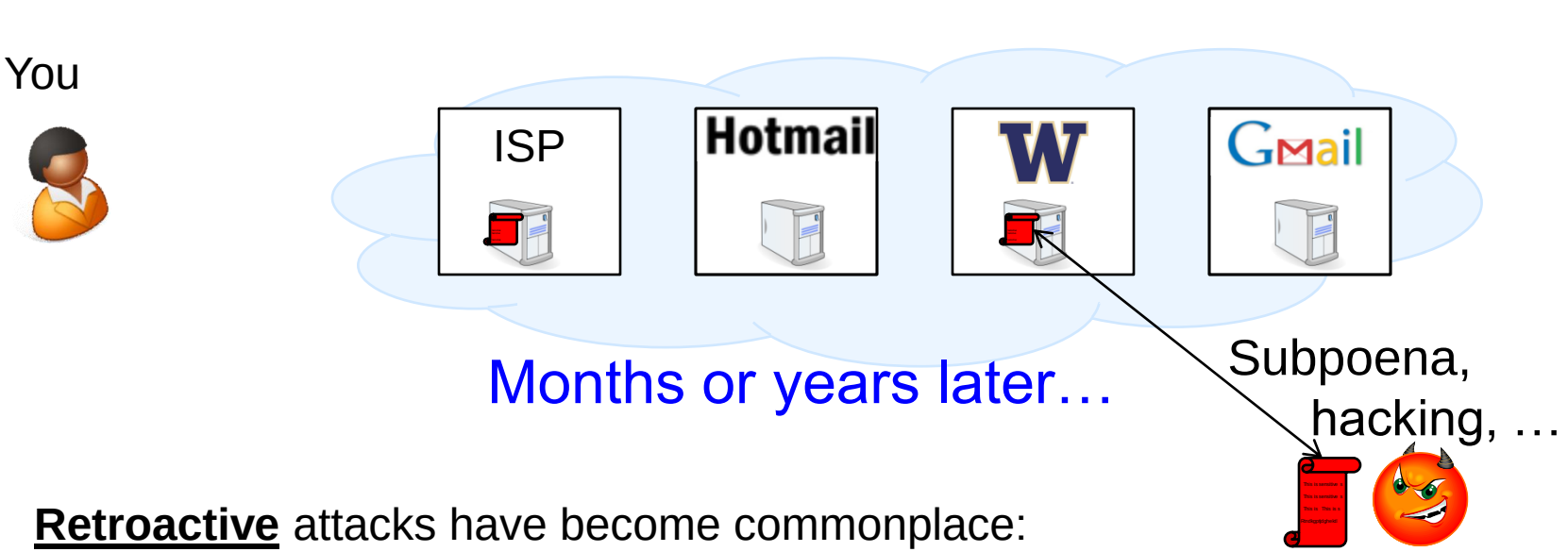
Data Lives Forever: Example, Email



A few days later…

- You want to delete the email, but:
  - You don't know where all the copies are
  - You can't be sure that all services will delete all copies (e.g., from backups and replicas)
  - Even deleting your account doesn't necessarily delete the data (e.g., Facebook)

Archived Copies Can Resurface Years Later



Retroactive attacks have become commonplace:

- Hackers
- Subpoenas
- Misconfiguration
- Laptops seized a
- Device theft
- Carelessness
- …

Telegraph.co
Chinese hum
Google em
Published: 2:33PM GMT

WebProNews
Breaking eBusiness and Search News
Email Being Used More In Divorce Cases

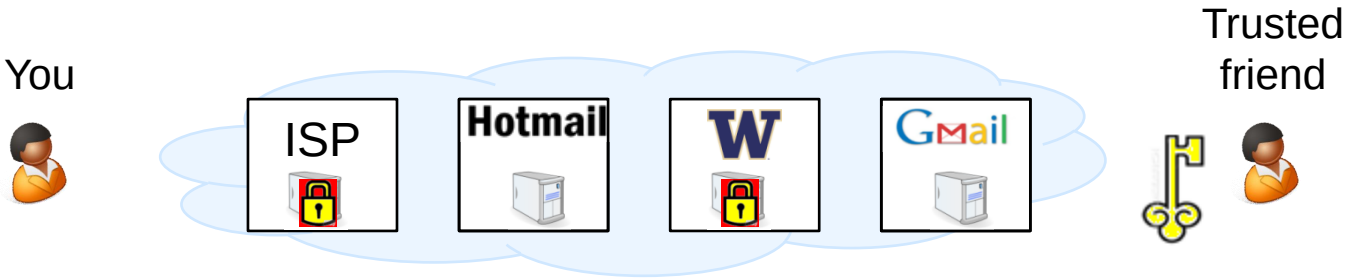
The New York Times

F.B.I. Gained Unauthorized Access to E-Mail

Published: February 17, 2008

WASHINGTON — A technical glitch gave the F.B.I. access to the e-mail messages from an entire computer network — perhaps hundreds of accounts or more — instead of simply the lone e-mail [...]

Why Not Rely On Encryption (e.g., PGP)?



- It's possible for an attacker to get both encrypted data and decryption key
  - PGP keys are long-lived (stored on disks, backed up)

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 e
law enforcement officials if requested under the Regulation of Investigatory Pow
Act).

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 his PGP passphrase so prosecutors can view the unencrypted files, a ruling that raises serious concerns about self-incrimination in an electronic

Why Not Rely On A Centralized Service?

DeleteMyData.com

Trust us: we'll help you delete your data!

- Huge trust concerns for relying on a centralized service

WIRED
BLOG NETWORK

Encrypted E-Mail Company Hushmail Spills to Feds

By Ryan Singel ✉ November 07, 2007 | 7:39:41 PM

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."

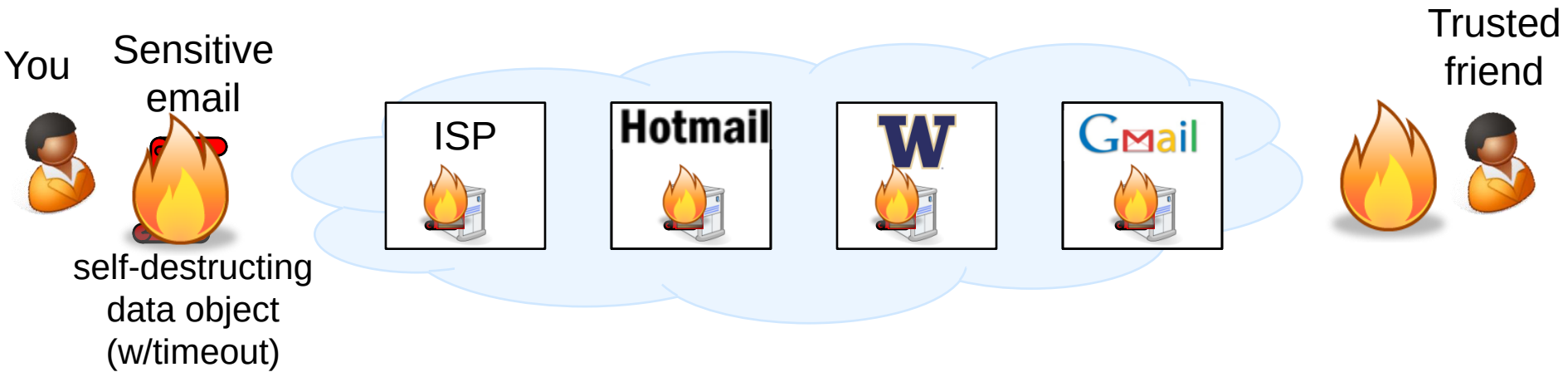
Question:

Can we empower users with control of data lifetime?

Answer:

Self-destructing data

Self-Destructing Data Model



Goals

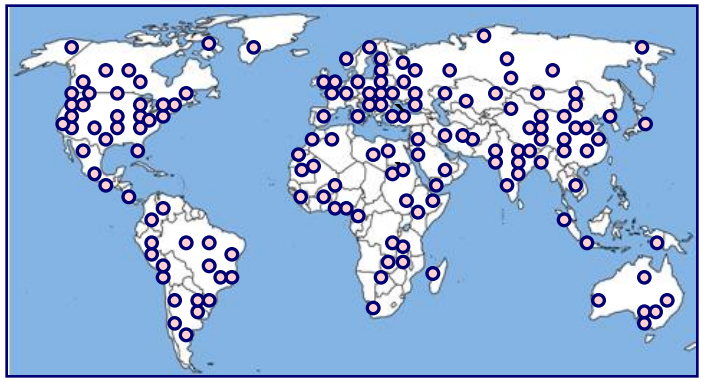
1. Until timeout, users can read original data
2. After timeout, all copies become permanently unreadable
   - 2.1 both online and offline copies
   - 2.2 even for attackers who obtain an archived copy & user keys
   - 2.3 without requiring any explicit action
   - 2.4 without having to trust any centralized services

Outline

- Self-Destructing Data: Motivation and Goals
- Implementing a Decentralized Self-Destructing Data System

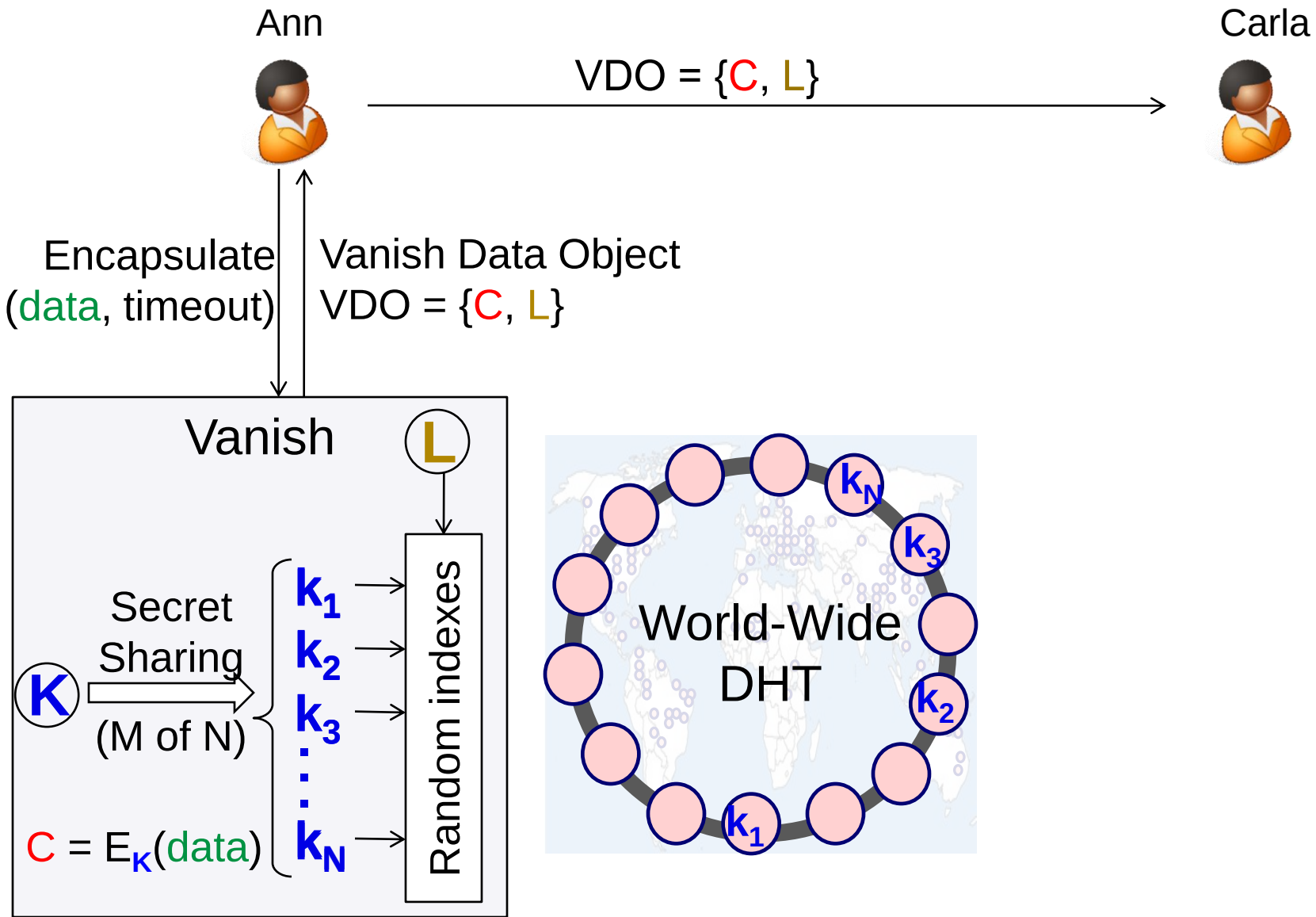
Our Idea

- Suppose we had access to millions of public “places” all around the world, where:
  - we could “hide” some information (*needle in a haystack*)
  - it would be impossible to find those locations later
  - the places would “lose” or time out our data over time (*churn*)

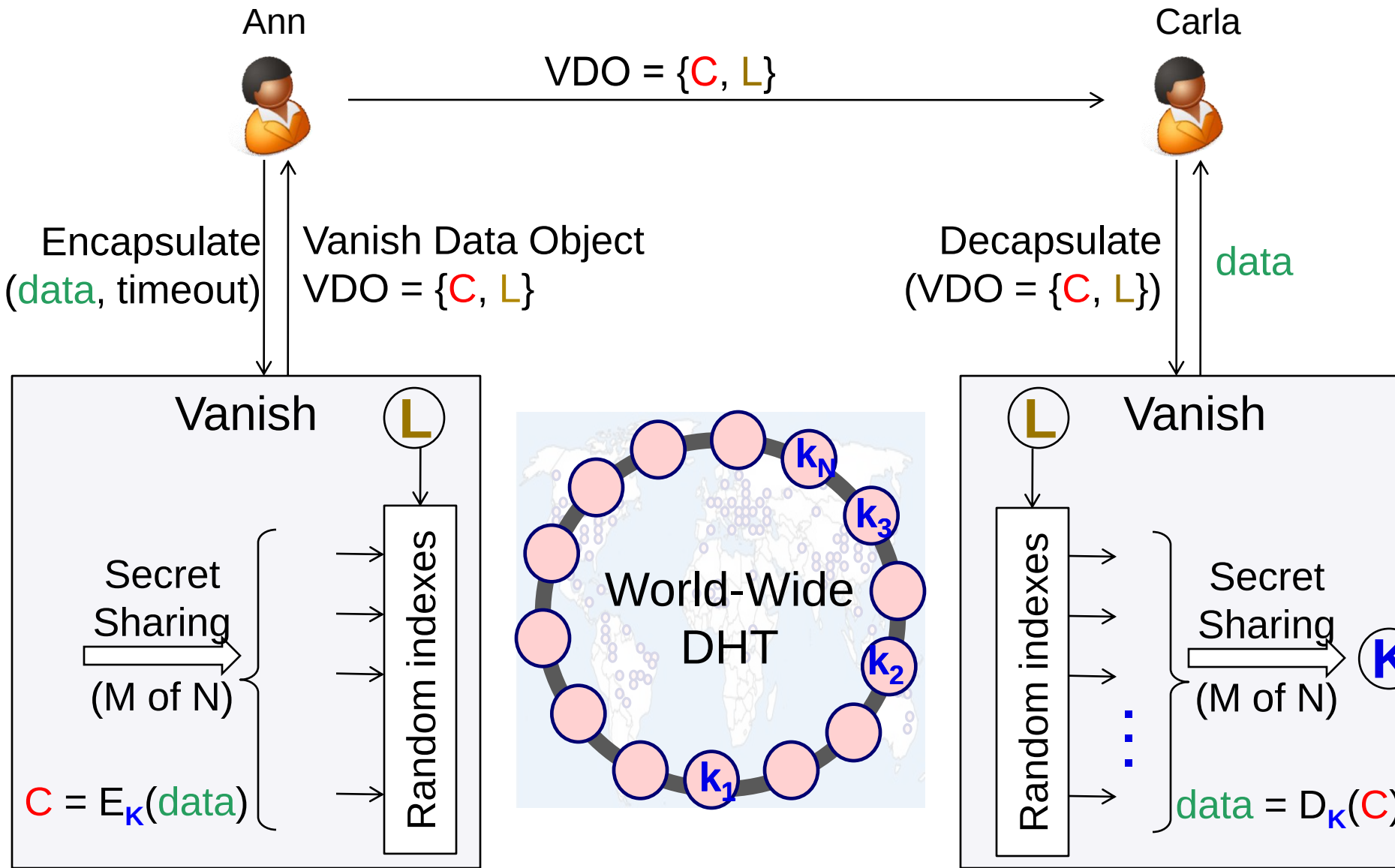


- How could we use this to create self-destructing data?

One example: DHTs (Vanish)

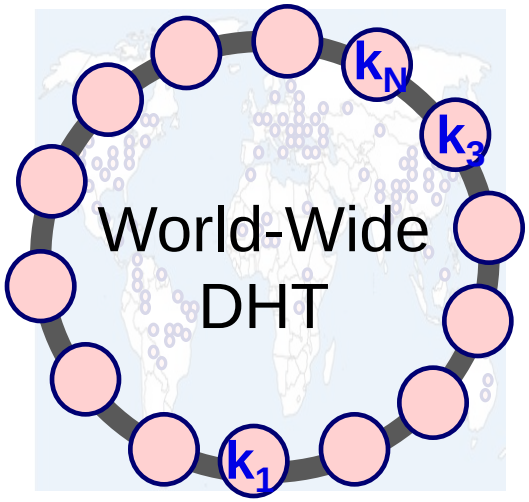


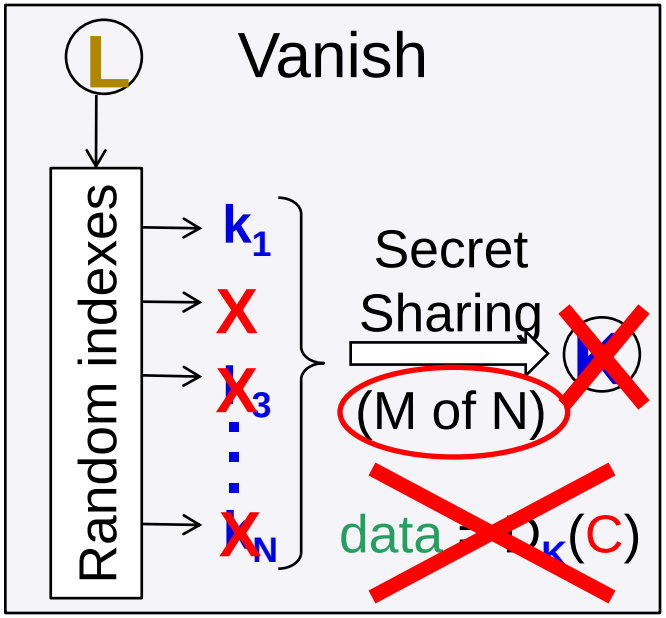
How Vanish Works: Data Decapsulation



How data times out in the DHT

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

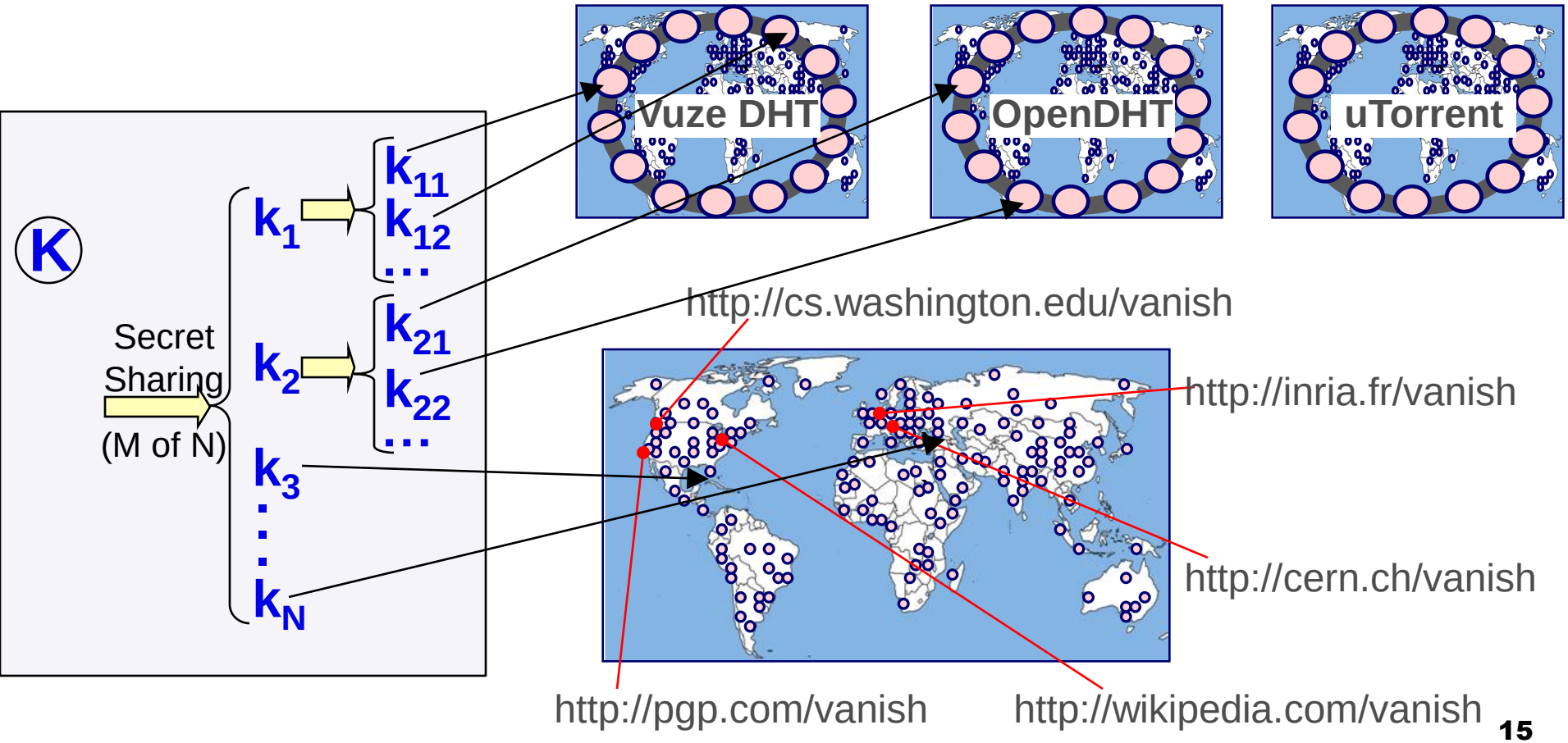




- Key loss makes all data copies permanently unreadable
- Random indexes / node IDs are useless in the future

Extending the trick: hierarchical secret sharing

- Keys are spread over multiple *key storage systems*
- No single system has enough keys to decrypt the data



History and Current State

Jul 09: We released Vanish (based on Vuze DHT)

- Description & source code available at: http://vanish.cs.washington.edu/

Aug 09: We presented Vanish at 2009 USENIX Security

- Won Outstanding Student Paper award

Sep 09: Others showed data crawling weaknesses in Vuze DHT

Oct 09: We designed, evaluated, and deployed at scale DHT defenses that strengthen Vuze against data crawling

- We raised attack bar by two orders of magnitude

Nov 09-now: Designed and evaluated hierarchical schemes (new paper in process)

Summary

- Formidable challenges to privacy in the Web:
  - Data lives forever
  - Disclosures of data and keys have become commonplace
- Self-destructing data empowers users with lifetime control
- Our approach:
  - Combines secret sharing with global-scale, distributed-trust, decentralized systems to achieve data destruction
  - Can combine the best security properties of multiple systems to raise the bar against attack