

Percival: Secure Data Archival Using Secret Sharing

Thomas M Kroeger, Joel C Frank, Ethan L Miller

Sandia National Laboratories and University of California, Santa Cruz



Problems in Long Term Data Archive

- **Insider Threat:** Insiders can be put into one of three categories:
 1. *Users*--those with legitimate operational need for the data
 2. *System administrators*--system administrators with digital and usually physical access
 3. *Facilities personnel*--those with physical access (e.g. janitorial staff, security guards etc...)
- **Central Point for Attack:** The central design of most archives provide clear single target for attackers.
- **Planning for security events:** The rate of security events these days necessitates that any wise system designer needs to plan for parts of their system being compromise.



Secure Foundation: Shamir Secret Splitting

Uses points on a polynomial curve to provide information theoretically secure secret splitting.

Example: Any 3 of 5 Secret $S = 1234$

1. Create a polynomial of degree 2 by generating 2 random coefficient e.g

$$f(x) = S + 166x + 94x^2$$

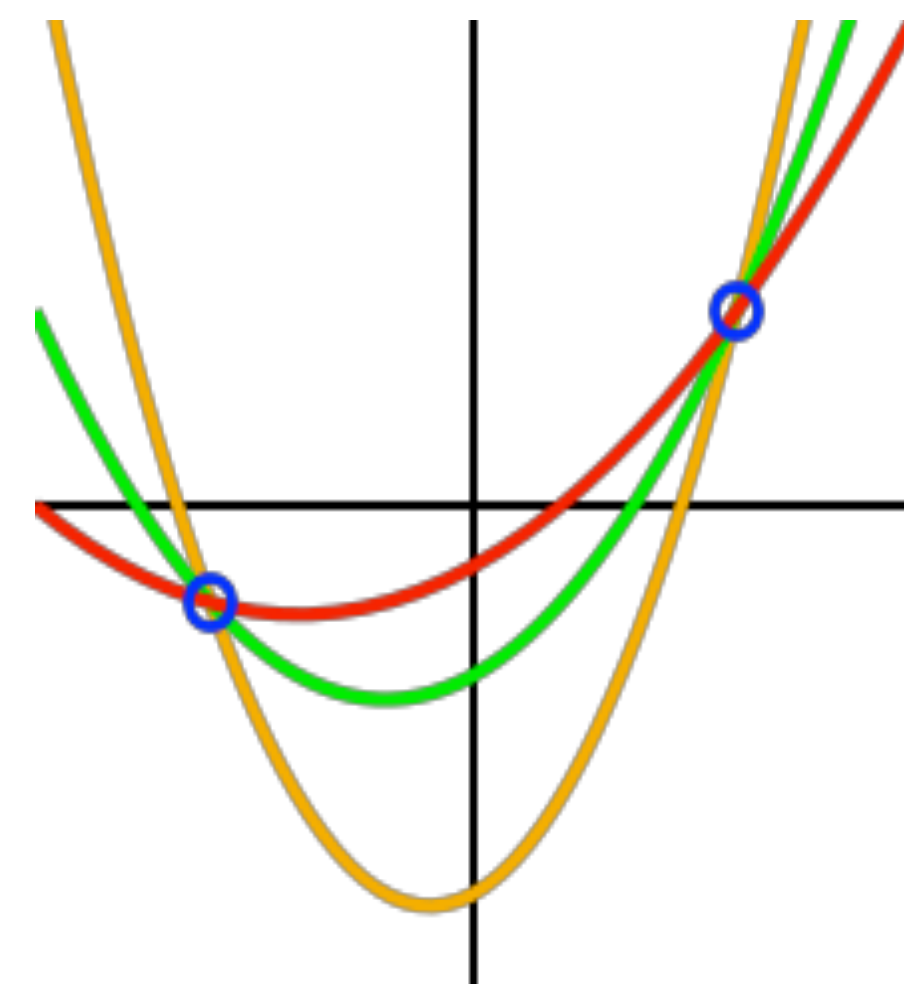
2. Generate 5 points along that curve:

(1,1494); (2,1942); (3,2578);
(4, 3402); (5,4414)

Any three enables me to solve for S.

All I know is S is a value on the Y-axis.

If I steal two points I don't learn anything new.



One can draw an infinite number of polynomials of degree 2 through 2 points.

Percival Directions

Applied Research – Proto-type development & Testing

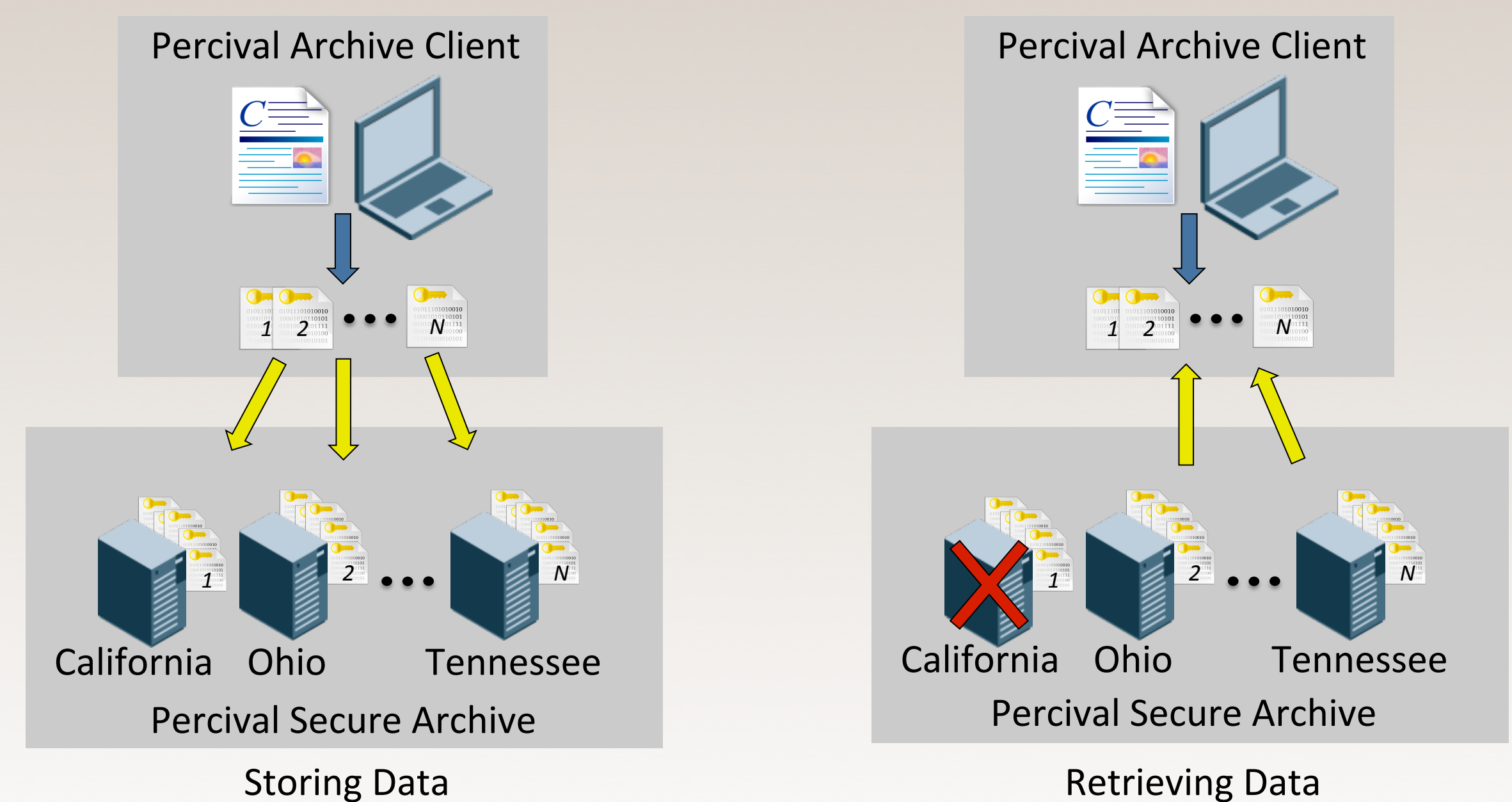
- Expand initial proof of concepts to pursue a detailed prototype
- Exploring and testing searching across the proto-type
- Explore adversarial models
 - Denial of service
 - Data integrity ECC type models to offer verification
 - Rebuilding after compromise
 - Detection of compromise
 - Man-in-the-middle threats

Open Research Questions

- Better search models
- Other transactions on split data
- More efficient transforms to split data
- Improved client side operations



Percival: Using Secret Splitting for Resilient Data Archive



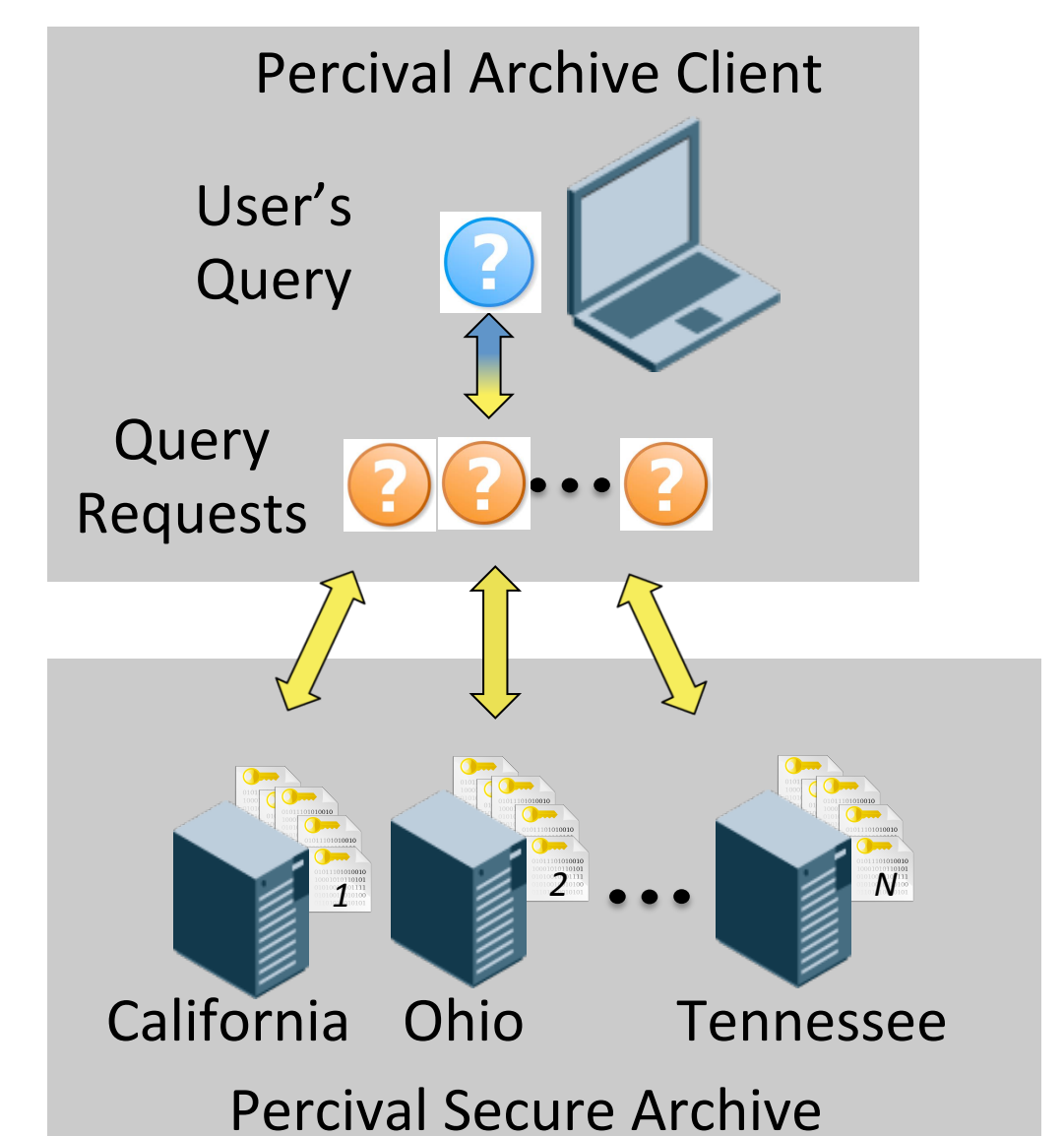
- Splits a 1K document into N shares of 1K that are stored by data custodians creating a distributed archive.
- To retrieve a document the client gets and combines T shares.
- Resilient to insider threat. T different site admins would have to collude.
- Able to operate with integrity and data protection in the presence of compromise.
- Resilient to crypto analytical attack, $T-1$ shares provides no info. All values equally likely.
- No key needed to protect the data simplifies long term archive management.

Searching a Percival Archive

- Data is vulnerable when on the client.
- Create ability to operate on the data without needing to reassemble the shares.

Keyword Search

- Recent research has explored protocols that enable keyword search across our archive.
- We use pre-indexing, key hashes and bloom filters to provide a blinded search.
- The custodians do the searching but are blinded to the terms searched and the results of the query.
- Initial results with data from the Gutenberg project and Wikipedia have shown promise.



Data custodians blindly perform the search

Authentication within Percival

Percival's design enables a broad range of capabilities for authentication:

- **Multiple users required**--For sensitive environments where user burden is acceptable, one could implement a configuration that requires two separate people to authenticate in order to retrieve a document. In the same way retail stores require two people to verify a refund this multi-user access requirement can help mitigate abuses among those with operational need to access sensitive data.
- **Key-chain**--when ease of use is paramount, one could use a key-chain with credentials to all the necessary repositories. The user could then unlock this keychain and gain access to the archive. Such a configuration would appear to the user as one simple authentication while in the background his client would have a diverse set of credentials across different domains hopefully using different protocols.