Thesis Proposal

Background

Currently, there are many commercial cloud services that allow both individuals and organizations to do three main tasks: 1) have multiple copies of their data, 2) access/share their data from multiple geographical locations, and 3) modify their data through synchronous or asynchronous mechanisms. Although these services are incredibly useful to thousands of users, there is a sector that has yet to take full advantage of cloud storage: nonprofits.

Nonprofits, particularly those located outside the United States and some of the wealthiest European states, face unique challenges that inhibit them from using commercial cloud storage services. First, the data that nonprofits deal with is highly sensitive, ranging from databases of interviews and reports of violations that amounted to genocide\(^1\) to government corruption information that has gotten journalists killed by their government’s authorities\(^2\). The information is too sensitive to be placed in some company’s server abroad. Even in the United States, data stored in cloud services can be accessed without requiring the data owner’s consent if there is sufficient evidence for criminal activity or for matters of national security.\(^3\) If the sensitive data of these nonprofits is to be stored in a remote location, it needs to be protected by encryption systems that will ensure the privacy of both the users and the data itself.

The other major issue that nonprofits must deal with are technological and economic constraints. From my experience working with nonprofits, many of them lack the technological expertise to develop or even use systems that would allow them to both send and store their highly sensitive data safely. This exposes them to hackings from their governments or other organizations, thus endangering their sources and the staff of the nonprofits themselves. Lastly, nonprofits rarely have the economic resources to pay for proprietary services. Even if the price of $10 for cloud storage is reasonable in the United States, such amount in pesos (191.24) or Rwandan francs (8985.13) is simply not affordable.

Problem Statement

Nonprofits need a distributed storage system that could help them replicate, share and modify their sensitive data. The system must offer data privacy protection and secure communication channels for the upload and download of data. In addition, the system needs to provide a friendly, easy-to-use user interface that could be integrated in as many workspaces as possible.

---

Challenges

1. All types of data.
   a. Nonprofits can handle everything from databases to word documents to photographs. The system must operate under the assumption that there is no standard type of data.
2. Highly unstructured data.
   a. Even across documents that have the same format, the data in each of them is often unorganized and must be worked on in order to be analyzable.
3. Low cost and efficient.
   a. The system must be free and open-source so that developers can contribute to the system, and it must operate in low-resource settings (low bandwidths, high latency, few cores).
4. Available to all users, easy to use.
   a. The system must be highly intuitive and must integrate easily into the user’s common workflow.

Deliverables

- **Architecture**
  The architecture will contain all the specifications with regards to:
  o How data is uploaded/downloaded to/from multiple servers.
  o How data is distributed and replicated among multiple servers.
  o How data privacy could be achieved through multiple encryption techniques such as public-key encryption and/or homomorphic encryption.
  o Any other architectural issues that may arise throughout the development of this project.

- **Algorithms**
  All the algorithms developed or modified to achieve data consistency across multiple servers and to ensure data privacy will be included as well.

- **Implementation prototype**
  o Client implementation
  o Server implementation
    - The server will be implemented through a Raspberry Pi so that host nonprofits spend as little money as possible in maintenance.

- **Plugin Software**
  o Intuitive user interface that can be integrated in multiple operating systems.

- **Novel technique**
  o Possible implementation of homomorphic encryption for an efficient cloud system.

Current solutions

Although there are some cloud services targeted for nonprofits, some of them free and some of them not, many of them are proprietary, which discourages nonprofits from placing their data
with them because private organizations may not always have their best interests in mind. Up to May 2018, there existed an open-source solution similar to the one presented here developed by a nonprofit called Benetech. The system consisted of multiple servers that contained encrypted copies of the data uploaded by the nonprofits. Each server was located at a nonprofit. In May 2018, Benetech stopped maintaining the system, called Martus, affirming that the system, which was 15 years old, had long surpassed longevity expectations. Something needs to take its place.

**Literature Review**

Currently, there is no research being done for the development of cloud technologies for nonprofits, but present research on privacy-preserving cloud systems looks into the challenges that nonprofits commonly face.

- **Obladi: Oblivious Serializable Transactions in the Cloud**
  - Authors: Natacha Crooks, Matthew Burke, Ethan Cecchetti, Sitar Harel, Lorenzo Alvisi and Rachit Agarwal
  - First cloud-based key value store that supports transactions while hiding access patterns from cloud providers.

- **A hybrid two layer attribute based encryption for privacy preserving in public cloud**
  - Authors: Shiva Verma, Sachin Ahuja
  - New framework that assures the confidentiality of data and preserves the privacy of users while delegating most of the access control enforcement to the cloud.
  - [https://ieeexplore.ieee.org/document/7824822/authors#authors](https://ieeexplore.ieee.org/document/7824822/authors#authors)

- **The Security Analysis of Data Re-Encryption Model in Cloud Services**
  - Authors: Xin Gu, Zhengquan Xu, Lizhi Xiong, Chunhui Feng
  - This paper studies the collusion attack issue by analyzing re-encryption model and classic encryption methods, and indicate how to choose encryption algorithm to against collusion attack in re-encryption model.

- **Encryption-Based Solution for Data Sovereignty in Federated Clouds**
  - Authors: Christian Esposito, Aniello Castiglione, Kim-Kwang Raymond Choo
  - This column describes a solution that applies encryption to protect data sovereignty in federated clouds rather than restricting the elasticity and migration of data across federated clouds. This paper will provide more insights into the legal aspects of storing data across multiple regions.

- **Homomorphic cloud computing scheme based on hybrid homomorphic encryption**
  - Authors: Xidan Song, Yulin Wang
  - This paper presents a hybrid cloud computing scheme based on the Paillier algorithm which is additively homomorphic, and RSA encryption algorithm which is multiplicative homomorphic. The system, according to the authors, is practical and efficient.

---

4 Martus website: [https://www.martus.org/](https://www.martus.org/)
• Comparative study of homomorphic encryption methods for secured data operations in cloud computing
  o Authors: Kanagavalli Rangasami, S. Vagdevi
  o This paper provides an analysis of multiple homomorphic encryption methods in terms of security provided to data and their usefulness in cloud computing.

• Fully homomorphic encryption application in cloud computing
  o Authors: Baohua Chen, Na Zhao
  o This paper introduces the principle of homomorphic encryption, then analyses some homomorphic encryption scheme and its improved algorithm used in cloud computing.

• Hybrid Encryption Techniques for Secure Sharing of a Sensitive Data for Banking Systems Over Cloud
  o Authors: Prachi More, Shubham Chandugade, Shaikh Mohammad Shafi Rafiq, Priya Pise
  o This work shows a security framework that can give protection and uprightness to trading delicate data through the cloud or the correspondence systems, in view of the utilization of the combination of Attribute-Based Encryption and Byte Rotation Encryption Algorithm. The essence of the work is to build up a straightforward platform that can get protection, integrity, and performance for the data exchange from peer to peer.

• Ensuring consistency file authentication over encrypted files in the cloud
  o Authors: V. Venkatesa Kumar, A. Murugavel
  o The paper presents a fork consistency approach to be integrated with the Deterministic Finite Automata (DFA) authentication method in order to achieve a consistency property.

• Socio Path: Protecting Privacy by Self-Sufficient Data Distribution in User-Centric Networks
  o Authors: Fabian Hartmann, Ingmar Baumgart
  o This paper presents a decentralized protocol for self-sufficient user-to-user communication. It handles device heterogeneity by decoupling data objects from notifications and keeping recurrent state exchanges small. Additionally, it offers a user-centric application interface which abstracts from devices towards the user.

• Secure algorithm for cloud computing and its applications
  o Authors: Akshita Bhandari, Ashutosh Gupta, Debasis Das
  o This paper proposes a schema that uses HE-RSA or hybrid encryption RSA along with Advanced Encryption Standard or AES to ensure efficiency, consistency and trustworthiness in cloud servers.