

Access Control Models: DAC, RBAC and Their Secure Practical Implementation

Majji Hari Krishna

Dept. of Computer Science & Engineering

Indian Institute of Technology Madras

Chennai, India

cs25m028@smail.iitm.ac.in

Abstract—Access control is a fundamental pillar of information security, ensuring that users can only access resources for which they are explicitly authorized. Among the myriad of access control models, Discretionary Access Control (DAC) and Role-Based Access Control (RBAC) are two of the most foundational. This report provides a comprehensive examination of these two protocols, beginning with their theoretical underpinnings and concluding with a detailed report on a practical, secure implementation. For DAC, we explore its owner-centric philosophy and ACL-based implementation. For RBAC, we delve into its organization-centric approach and its alignment with the principle of least privilege. This paper then details the architecture and security mechanisms of a hybrid system that practically implements both protocols. We describe the approach, including a "Zero Trust" data model, a secure API, specific cipher algorithms (AES-GCM) and integrity-protection techniques (HMAC-SHA256) to protect data and metadata, and the core enforcement logic. The result is a prototype that bridges the gap between the abstract theory of access control and its tangible, secure application.

Index Terms—Access Control, Information Security, DAC, RBAC, Access Control List (ACL), Least Privilege, Implementation, Flask, SQLite, Envelope Encryption, HMAC, Zero Trust, AES-GCM

I. INTRODUCTION

In the digital age, the protection of sensitive information is paramount. Organizations, governments, and individuals rely on computer systems to store, process, and transmit vast amounts of data. Unauthorized access to this data can lead to catastrophic consequences. The primary mechanism for preventing such unauthorized access is the implementation of a robust access control policy.

Access control is the selective restriction of access to a resource. It is the process by which a system grants or denies requests based on a predefined security policy. The entities requesting access are subjects (users, processes), and the resources being protected are objects (files, databases).

Over the decades, several models have been developed, including Discretionary Access Control (DAC), Mandatory Access Control (MAC), and Role-Based Access Control (RBAC). DAC is characterized by its flexibility, placing control in the hands of the resource owner. Conversely, RBAC takes a structured, centralized approach, assigning permissions based on an individual's role.

This paper is presented in two phases. Phase 1 (Sections II-VIII) provides a detailed comparative study of the DAC

and RBAC protocols, dissecting their architectures, formal models, and theoretical applications. Phase 2 (Sections IX-XVI) presents a detailed report on the design and construction of a secure software prototype that implements a hybrid of both DAC and RBAC, focusing on the practical challenges and security solutions required to build such a system. Phase 3 (Section XVII) provides a critical security analysis of the implementation and understanding the weakness in the protocols.

II. ABBREVIATIONS AND ACRONYMS

TABLE I
ABBREVIATIONS AND ACRONYMS

Acronym	Description
ACL	Access Control List
AEAD	Authenticated Encryption with Associated Data
AES-GCM	Advanced Encryption Standard - Galois/Counter Mode
AIK	Application Integrity Key
API	Application Programming Interface
DAC	Discretionary Access Control
FEK	File Encryption Key
HMAC	Hash-based Message Authentication Code
KMS	Key Management Service
MAC	Mandatory Access Control
MEK	Master Encryption Key
NIST	National Institute of Standards and Technology
RBAC	Role-Based Access Control
SoD	Separation of Duty
WSGI	Web Server Gateway Interface

III. FORMAL MODELS

To understand the precise behavior of access control systems, it is useful to define them with mathematical formalism.

A. Formalizing Discretionary Access Control (DAC)

The state of a DAC system can be described by an access control matrix (ACM), as proposed by Lampson [1]. This matrix, M , describes the rights of subjects over objects.

Let S be the set of subjects and O be the set of objects. Let R be the set of rights (e.g., read, write, own). An entry $M[s, o]$ where $s \in S$ and $o \in O$, contains a subset of rights $r \subseteq R$.

$$M[s, o] = \{r_1, r_2, \dots, r_k\} \quad \text{where } r_i \in R \quad (1)$$

The "discretionary" aspect comes from rules that allow subjects to modify the matrix. For example, if $\text{'own'} \in M[s, o]$, then subject s can execute commands to add or remove rights for another subject s' over object o .

$$\text{grant}(s, s', o, r) \implies \text{add } r \text{ to } M[s', o] \quad (2)$$

This formalization clearly shows how permissions are tied directly to subject-object pairs.

B. Formalizing Role-Based Access Control (RBAC)

The NIST standard for RBAC provides a widely accepted formalization [2]. The core RBAC model consists of several sets and relations:

- U, R, P, S : sets of users, roles, permissions, and sessions.
- $PA \subseteq P \times R$: a many-to-many permission-to-role assignment relation.
- $UA \subseteq U \times R$: a many-to-many user-to-role assignment relation.

A user can exercise a permission p if and only if that permission is authorized for one of the user's active roles.

$$\exists r \in \text{session_roles}(s) \text{ such that } (p, r) \in PA \quad (3)$$

This model introduces a layer of abstraction—the role—between users and permissions, which is the key to its administrative scalability.

IV. DISCRETIONARY ACCESS CONTROL (DAC)

Discretionary Access Control is an access control policy determined by the owner of an object. The owner has the discretion to grant or revoke access permissions for that object to other subjects. This model is one of the simplest and most widely used, forming the basis for security in most commercial operating systems.

A. Core Principles and Implementation

The fundamental principle of DAC is **owner-controlled access**. When a subject creates a new object (e.g., a file or directory), they become its owner and are granted full control over it. This control is typically managed through an Access Control List (ACL).

An ACL is a data structure associated with an object that lists all subjects (users or groups) that have been granted access rights to it. Each entry in the ACL, known as an Access Control Entry (ACE), specifies a subject and the set of operations they are permitted to perform.

For instance, in a UNIX-like operating system, file permissions are a classic example of DAC. Each file has an owner and a group, and permissions are defined for three classes of subjects:

- 1) **Owner**: The user who created the file.
- 2) **Group**: A group of users who share permissions.
- 3) **Others**: All other users on the system.

Permissions are specified for reading ('r'), writing ('w'), and executing ('x'). A command like `chmod 754 file.txt` sets the owner's permissions to 'rwx' (read, write, execute), the group's to 'rx' (read, execute), and others' to 'r' (read). The owner can change these permissions at any time.

B. Advantages of DAC

- 1) **Flexibility and Ease of Use**: DAC is highly flexible. Owners can share resources with collaborators quickly and dynamically without needing to go through a central administrator. This fosters a collaborative environment.
- 2) **Simplicity**: The concept of ownership is intuitive and easy for end-users to understand and manage for their own files.
- 3) **Granularity**: Permissions can be assigned on a per-user, per-object basis, allowing for very fine-grained control if needed.

C. Disadvantages of DAC

- 1) **Susceptibility to Trojan Horses**: DAC's major vulnerability lies in its handling of program execution. If a user runs a program (a Trojan horse) that has malicious code, that program inherits all of the user's permissions. It can then perform unauthorized actions on behalf of the user, such as deleting files or copying sensitive data to an unauthorized location, without the user's knowledge [3].
- 2) **Lack of Centralized Control**: Because permissions are managed by individual users, it is difficult to enforce a consistent, organization-wide security policy. This can lead to security gaps and inconsistent permission settings across the system.
- 3) **Administrative Overhead**: In a large system, tracking and managing permissions can become incredibly complex. Revoking access for a user who has left the organization can be a tedious and error-prone process, as an administrator must check the ACL of every object they might have had access to.

V. ROLE-BASED ACCESS CONTROL (RBAC)

Role-Based Access Control is a policy-neutral access control model that has gained widespread acceptance as a best practice for managing permissions in medium-to-large-scale enterprises. In RBAC, access decisions are based on the roles that individual users have as part of an organization.

A. Core Principles and Implementation

The central idea of RBAC is to introduce a level of indirection between users and permissions. Rather than assigning permissions directly to users, permissions are assigned to roles, and users are then assigned to roles. A role is a semantic construct that represents a job function or title within an organization (e.g., 'Accountant', 'HR Manager', 'Database Administrator').

The implementation of RBAC involves three primary mappings:

- 1) **Permission-Role Assignment (PA)**: Specific system permissions (e.g., 'create-invoice', 'view-salary-data', 'reboot-server') are assigned to the relevant roles.
- 2) **User-Role Assignment (UA)**: Users are assigned to one or more roles based on their job responsibilities.

- 3) **Session Management:** When a user logs in, they create a session and can choose to activate a subset of the roles they are assigned. This allows them to operate with only the permissions necessary for the immediate task, adhering to the principle of least privilege.

This structure simplifies security administration. When a new employee joins, the administrator simply assigns them the appropriate role(s). When an employee's job changes, their role assignment is updated, and their permissions change automatically. When an employee leaves, their role assignments are simply revoked, effectively and cleanly severing all access.

B. Advantages of RBAC

- 1) **Simplified Administration:** Managing roles is far simpler than managing individual permissions for thousands of users. This significantly reduces administrative workload and the potential for error [4].
- 2) **Enforcement of Least Privilege:** RBAC makes it easy to implement the principle of least privilege. By defining roles with the minimal set of permissions required for a job function, organizations can reduce the risk of users having excessive, unnecessary access rights.
- 3) **Scalability:** RBAC scales exceptionally well. As an organization grows, new users are simply slotted into existing roles. New job functions can be accommodated by creating new roles.
- 4) **Policy Compliance:** RBAC helps organizations meet regulatory and statutory requirements (e.g., HIPAA, SOX) that mandate privacy and confidentiality, as it allows for systematic enforcement and auditing of access policies.

C. Disadvantages of RBAC

- 1) **Initial Complexity:** The initial implementation of RBAC can be a complex and resource-intensive project. It requires a thorough analysis of the organization's business processes and job functions to define a logical and effective set of roles.
- 2) **Role Explosion:** In large, complex organizations, the number of distinct roles can become unmanageably large. This "role explosion" can reintroduce the complexity that RBAC was designed to solve. Careful role engineering and management are required to prevent this.
- 3) **Static Nature:** Core RBAC models can be static. They do not easily account for contextual attributes, such as time of day, user location, or the state of a resource, which may be required for more dynamic access control decisions.

VI. COMPARATIVE ANALYSIS: DAC VS. RBAC

While both models aim to secure resources, their philosophical and practical differences are stark. The choice between them depends heavily on the context of the system being secured.

TABLE II
COMPARISON OF DAC AND RBAC CHARACTERISTICS

Feature	Discretionary Access Control (DAC)	Role-Based Access Control (RBAC)
Control Loc	Decentralized; at the discretion of the resource owner.	Centralized; managed by system administrators based on organizational policy.
Access Basis	Based on the identity of the subject (user or group).	Based on the assigned role(s) of the subject.
Primary Mechanism	Access Control Lists (ACLs).	Role-Permission and User-Role assignments.
Flexibility	Very high; permissions can be changed dynamically by users.	Moderate; requires administrative action to change roles or role permissions.
Scalability	Poor; becomes unmanageable in large organizations.	Excellent; scales efficiently with organizational growth.
Security Principle	Focuses on ease of sharing and collaboration.	Enforces the principle of least privilege and separation of duty.
Policy Enforcement	Inconsistent; depends on the diligence of individual users.	Consistent; enforces a single, organization-wide security policy.
Typical Environment	Personal computing, small workgroups, academic projects.	Enterprise systems (ERP, CRM), government, healthcare, financial institutions.

The most fundamental difference lies in how permissions are managed. In DAC, permissions are attached to an object, and control flows from the object's owner. In RBAC, permissions are attached to a role, and control flows from a central administrative policy that maps users to these roles.

This leads to a trade-off: DAC offers unparalleled flexibility for the end-user, making it ideal for environments where information sharing is fluid and dynamic. However, this comes at the cost of centralized oversight and policy enforcement. RBAC enforces a rigid, centrally managed policy, which is less flexible for the end-user but provides superior security, auditability, and administrative efficiency for an organization.

VII. USES AND REAL-LIFE APPLICATIONS

The theoretical differences between DAC and RBAC are best understood through their practical implementations in real-world systems.

A. Applications of DAC

- 1) **Desktop Operating Systems:** The most common examples of DAC are found in file systems of OSs like Microsoft Windows (NTFS permissions), macOS, and Linux. When you create a document on your computer, you are its owner. You can then right-click the file, go to 'Properties' or 'Get Info', and set specific permissions for other users or groups on that computer or network.
- 2) **Social Media Platforms:** On platforms like Facebook or Google Drive, when you upload a photo or create a document, you can often choose who gets to see it—"Only Me," "Friends," or "Public." This is a form of DAC, where you, the owner of the data object, are making discretionary decisions about its access.

- 3) **Small-Scale Collaborative Tools:** Tools like Trello or shared folders in Dropbox often use a DAC-like model. The creator of a board or folder can invite specific collaborators and assign them different levels of access (e.g., view-only, editor).

B. Applications of RBAC

- 1) **Enterprise Resource Planning (ERP) Systems:** In a large company using a system like SAP or Oracle, thousands of employees need access. An employee in the finance department is assigned the 'Accountant' role. This role automatically grants them access to view financial ledgers, create invoices, and process payments, but not to view HR records or modify production schedules.
- 2) **Cloud Computing Platforms:** Services like Amazon Web Services (AWS) and Microsoft Azure rely heavily on RBAC through their Identity and Access Management (IAM) systems. An administrator can create roles like 'Developer', 'DatabaseAdmin', or 'Auditor'. A user assigned the 'Developer' role can spin up and manage virtual machines but cannot change billing information or access sensitive security logs, which are permissions reserved for other roles.
- 3) **Healthcare Information Systems:** To comply with regulations like HIPAA, hospitals use RBAC to control access to Electronic Health Records (EHR). A doctor is assigned a 'Physician' role, allowing them to view and modify the records of patients under their care. A 'Billing Clerk' role allows access only to the insurance and payment information associated with a patient, not their clinical diagnoses or treatment history. This ensures that employees only access the minimum necessary information required to perform their jobs.

VIII. CHALLENGES AND FUTURE TRENDS

Neither DAC nor RBAC is a perfect solution. DAC's weakness is its lack of central control. RBAC's main challenge is managing complexity and potential "role explosion." The successor to RBAC is widely considered to be Attribute-Based Access Control (ABAC), which makes dynamic access decisions based on attributes of the subject, object, and environment [5].

IX. PHASE 2: SECURE IMPLEMENTATION APPROACH

The preceding sections (Phase 1) established the formal theory of DAC and RBAC. This second phase of the project focuses on the practical challenge: implementing a secure, hybrid system that combines both protocols.

The theoretical models are clean and abstract, but a real-world implementation must contend with numerous security threats. An attacker will not just try to bypass the logic; they will try to attack the *implementation itself*. Our approach was therefore based on a "Zero Trust" posture [9], which assumes that no component of the system, including our own database, is inherently trustworthy.

This led to two primary security goals for the implementation:

- 1) **Confidentiality:** All sensitive data (user files) and sensitive metadata (access control permissions) must be unreadable to an attacker, even one with full access to the database.
- 2) **Integrity:** The system must be able to detect and prevent any unauthorized modification of access control metadata by an attacker with database access.

To achieve these goals, we designed a system composed of a secure API, an encrypted database, and a robust client-side simulator. The real-world scenario modeled is a "Secure Hybrid Document Management System" (SH-DMS) where an organization sets baseline permissions via RBAC (e.g., 'Engineers' can write to the /Engineering folder), but individual users can then use DAC to share their specific files with others (e.g., an Engineer shares a file with a user from 'Marketing').

X. DEVELOPMENT ENVIRONMENT & TOOLING

The selection of a proper development environment is the first step in building a secure system. The primary goal was isolation and the use of industry-standard tools.

A. Python Virtual Environment (venv)

All development was conducted within a Python `venv`. This is a critical best practice for secure development, as it isolates project dependencies (e.g., Flask, cryptography) from the host operating system's Python installation. This prevents package version conflicts and ensures that the cryptographic libraries are explicitly version-managed for the project. For this project, `python -m venv venv` was used, and all packages were installed via `pip` from a `requirements.txt` file. This guarantees a reproducible and auditable build environment.

B. Core Technologies

- **Python 3.10:** The core programming language, chosen for its clarity and extensive support for security libraries.
- **Flask [6]:** A lightweight Web Server Gateway Interface (WSGI) web framework. It was used to build the secure REST API (`app.py`). Its "micro-framework" nature is ideal for this project, as it carries no unnecessary overhead and forces all security logic to be implemented explicitly.
- **SQLite 3 [8]:** A single-file, serverless database engine (`secure-data.db`). This was chosen for its simplicity of setup (no separate database server required) and ease of use in a prototype setting. Despite its simplicity, it is a robust, ACID-compliant database.
- **Cryptography Library [7]:** The `cryptography` library for Python. This is the *de facto* standard, providing modern, peer-reviewed, and secure implementations of cryptographic primitives. Using this library prevents common, self-implemented crypto flaws.
- **Visual Studio Code:** The IDE used for all development, providing integrated terminal and debugging capabilities.

XI. SYSTEM ARCHITECTURE

The system is designed as a classic three-tier architecture, but with security considerations at each layer. The components are:

- 1) **Presentation/Client Tier:** A command-line script (`client.py`) that simulates user actions.
- 2) **Application/Logic Tier:** The Flask server (`app.py`) that contains all business logic.
- 3) **Data Tier:** The SQLite database (`secure-data.db`) and the file system (`uploads/` folder).

A. Server Model (Flask/Werkzeug)

The server in this project is the Flask application, which is run by executing `python app.py`. By default, Flask uses the **Werkzeug WSGI development server**. It is important to understand its operation:

- **How it Works:** When started, it binds to a local port (e.g., 127.0.0.1:5000) and listens for incoming HTTP requests. As a development server, it is single-threaded and processes one request at a time.
- **Debug Mode:** We ran the server with `app.run(debug=True)`. This enables two key features: 1. *The Reloader*: Automatically restarts the server when a code file is changed. 2. *The Debugger*: Provides a detailed, interactive traceback in the browser if an error occurs. This was essential for the iterative debugging process.
- **WSGI:** Flask (and Werkzeug) implement the WSGI standard, which provides a simple interface between the web server and the Python application.

In a production environment, this server would be replaced by a production-grade WSGI server like Gunicorn or uWSGI, which can handle multiple concurrent requests.

B. API-Client Architecture

The system is built on a strict API-centric design.

- **The Server (`app.py`):** This is the "brain" and the "fortress" of the system. It exposes a set of REST API endpoints (e.g., `/api/files/upload`, `/api/files/share`). Crucially, it contains 100% of the security and access control logic. No client is ever trusted to make a security decision.
- **The Client (`client.py`):** This script is "dumb." It holds no logic, no keys, and no permissions. It is merely a command-line tool to send HTTP requests to the server. For this project, authentication was simulated by sending a trusted `X-User-ID` header. In a real system, this would be replaced by an unforgeable JWT (JSON Web Token).

This separation ensures that even if the client application is compromised, the attacker cannot bypass the server's enforcement engine.

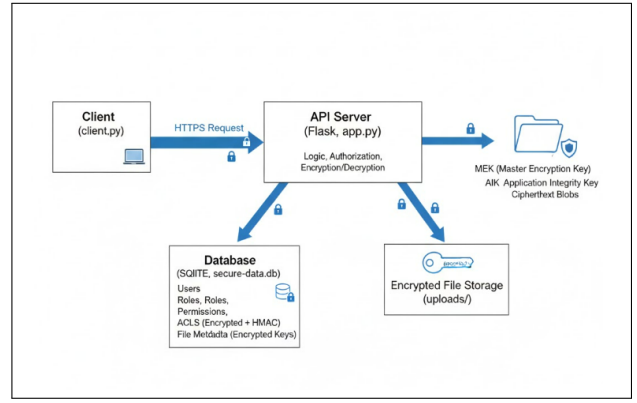


Fig. 1. System Architecture Diagram.

XII. SECURE DATABASE SCHEMA

The foundation of the access control logic is the database schema, defined in `database.py`. The tables were designed not just for function, but for security.

- `users`: Stores user information. (e.g., 'alice').
- `roles`: Defines the RBAC roles. (e.g., 'Engineer').
- `permissions`: Defines the available actions. (e.g., 'read', 'write').
- `resources`: The "object" table, representing both files and folders. This table's design is critical:
 - `owner_user_id`: This column is the foundation of the entire DAC protocol.
 - `parent_id`: This column enables the folder hierarchy, which is essential for the RBAC-inheritance check.
- `user_roles`: A simple mapping table linking `user_id` to `role_id`. This is the implementation of the *UA* relation from the formal model.
- `role_permissions`: A mapping table linking `role_id` and `permission_id` to a `resource_id` (a folder). This is the *PA* relation.
- `file_metadata`: This table holds the *cryptographic* metadata for files.
 - `encrypted_file_key`: This column stores the encrypted FEK, forming the basis of our envelope encryption strategy.
- `acls`: This is the implementation of the Access Control List for explicit DAC. This table is the most heavily secured:
 - `encrypted_permissions`: Stores the granted permissions (e.g., 'read') as an encrypted blob, ensuring confidentiality.
 - `hmac`: Stores an HMAC tag, ensuring the integrity of the entire row against tampering.

XIII. CRYPTOGRAPHIC DESIGN & IMPLEMENTATION

This project's core security objective was to protect data and metadata from all attackers, including a malicious insider with database access. This was achieved through a multi-layered cryptographic strategy defined in `security.py`.

A. Key Management (Simulated KMS)

In a production system, all cryptographic keys would be stored in a hardware security module (HSM) or a Key Management Service (KMS) like AWS KMS or HashiCorp Vault. For this prototype, we simulated this by storing two master keys in a `secrets.ini` file, which is explicitly forbidden from source control by `.gitignore`.

- 1) **Master Encryption Key (MEK):** An AES-256 key (32 bytes) used *only* to encrypt other, data-specific keys.
- 2) **Application Integrity Key (AIK):** A separate, secret key (32 bytes) used *only* for generating HMACs.

This separation of keys is a critical security principle. The key used for encryption should never be used for signing, as this can open theoretical attack vectors. These keys are loaded into application memory once on startup and represent the system's root of trust.

B. File Encryption (Confidentiality)

To protect the files themselves, we used an **Envelope Encryption strategy**. This avoids reusing the MEK for every file, which would be a cryptographic weakness (reusing a key in GCM mode with different nonces can be catastrophic).

The process, handled by `encrypt_file` and `decrypt_file_stream`, is as follows:

- 1) **On Upload:**
 - a) A new, unique, cryptographically random 32-byte **File Encryption Key (FEK)** is generated.
 - b) A random 12-byte **nonce** (IV) is generated.
 - c) The file's contents are encrypted using this FEK and nonce with our chosen cipher algorithm.
 - d) The FEK itself is then encrypted using the **MEK** (and a different nonce).
 - e) The *encrypted* FEK, the nonce, and the ciphertext are stored.
- 2) **On Download:**
 - a) The server retrieves the *encrypted* FEK and the file's nonce.
 - b) It uses the **MEK** to decrypt the FEK, bringing the plaintext FEK into memory.
 - c) It then uses this plaintext FEK and the file's nonce to decrypt the file's ciphertext, streaming it to the user.
 - d) The plaintext FEK is immediately discarded from memory.

Security Benefit: An attacker who steals the database only gets encrypted files and encrypted keys. Both are useless without the MEK, which only exists in the application server's memory.

C. Cipher Algorithm: AES-256-GCM

The specific cipher algorithm chosen was **AES-256-GCM**. This was a deliberate choice over simpler ciphers like AES-CBC.

- **AES (Advanced Encryption Standard):** The global standard for symmetric encryption, approved by NIST. We use a 256-bit key length for a high security margin.
- **GCM (Galois/Counter Mode):** This is the critical part. GCM is an **AEAD** (Authenticated Encryption with Associated Data) mode. This means it provides both: 1. **Confidentiality:** It encrypts the data (as an output-feedback cipher). 2. **Integrity/Authentication:** It produces an "authentication tag" (a GCM tag) that verifies the data has not been tampered with.

By using AES-GCM, we ensure that an attacker cannot perform a "bit-flipping" attack on the encrypted files. If an attacker modifies even a single bit of the ciphertext, the GCM tag will not match during decryption, and the `cryptography` library will raise an `'InvalidTag'` exception, automatically thwarting the attack.

D. Metadata Integrity (The "Crown Jewels")

The most significant security contribution of this project is the protection of the access control metadata itself. An attacker who can write to the database could simply grant themselves access by bypassing our API.

Threat: `UPDATE acls SET target_user_id = [ATTACKER_ID] WHERE resource_id = [TARGET_FILE_ID]`

Solution: We use a HMAC (Hash-based Message Authentication Code) to protect every row in the `acls` table.

- **Chosen Algorithm: HMAC-SHA256.** We use SHA-256 as the underlying hash function, keyed with our secret **AIK**.
- A simple hash (like SHA-256) is **not** sufficient, as the attacker could just re-compute the hash for their tampered data. HMAC requires the secret **AIK**, which the attacker (who only has DB access) does not possess.

The process, handled by `generate_hmac` and `verify_hmac`, is as follows:

- 1) **On Write (Share):** When an ACL is created for a user and resource:
 - The permissions (e.g., `['read']`) are first encrypted to ciphertext.
 - A "message" is constructed that binds the permission to the user and resource:

```
message = f"{encrypted_perms} : {resource_id} : {target_user_id}"
```
 - The system computes:

`hmac_tag = HMAC(AIK, message)`

- This `hmac_tag` is stored in the `acls.hmac` column.
- 2) **On Read (Access Check):** When the system fetches an ACL row:
 - It reconstructs the expected message from the retrieved data.
 - It computes the expected HMAC tag using the **AIK**.

- It securely compares the expected HMAC with the stored HMAC from the database.

Security Benefit: If they do not match, the system knows the row has been tampered with. The request is immediately aborted, a critical security event is logged, and the attack is stopped. This makes the access control rules cryptographically read-only to anyone without the AIK.

XIV. THE HYBRID ENFORCEMENT ENGINE

All security theory and cryptographic mechanisms are centralized into a single, hardened function in `app.py: check_permission()`. This function is the "Policy Decision Point" (PDP) for the entire application. All secure API endpoints *must* call this function before performing any action.

The function executes a hybrid check in a specific, logical order:

- 1) **Check Ownership (Implicit DAC):** The system first checks if the requesting `user_id` matches the `owner_user_id` in the `resources` table. If true, access is granted. This is the fastest and most common check.
- 2) **Check ACL (Explicit DAC):** If the user is not the owner, the system queries the `acls` table for a specific share. If an entry is found:
 - It *first* verifies the HMAC. If invalid, the request is immediately denied.
 - It *then* decrypts the `encrypted_permissions`.
 - If the requested action (e.g., 'read') is in the decrypted list, access is granted.
- 3) **Check Roles (RBAC):** If both DAC checks fail, the system performs an RBAC check. This is the most complex logic, as it must account for permission inheritance.
 - The system identifies the resource's parent folder.
 - It performs a recursive query (using the `get_parent_folders` helper) to find **all** parent folders up to the root.
 - It then executes a complex SQL query that joins `user_roles`, `role_permissions`, and `permissions` to see if the user has **any** role that is granted the requested permission on the resource's folder **or any of its parents**.
 - If a valid RBAC permission is found, access is granted.
- 4) **Default Deny:** If all three checks fail, the request is denied.

This prioritized, multi-step process successfully implements the hybrid logic, allowing organization-level policy (RBAC) to set the baseline, while empowering users with fine-grained control (DAC) over their own resources.

XV. THREAT MODELING & SECURITY ANALYSIS

By adopting this "Zero Trust" architecture, the system is secured against multiple attack vectors.

A. Threat Model

We considered two primary attackers:

- 1) **External Attacker:** A remote user on the internet with no valid credentials.
- 2) **Malicious Insider:** A disgruntled employee (e.g., a database administrator) with direct read/write access to the `secure-data.db` file and `uploads/` folder, but *not* to the application server's memory or secret keys.

B. Protection Against Attack Vectors

• Against External Attacker:

- All access is denied by default. The attacker must guess a valid `X-User-ID`, which is a brute-force problem.
- Even with a valid user ID, the server-side `check_permission` engine prevents them from accessing any resource for which they are not explicitly authorized.
- All communication would be over HTTPS (TLS), preventing man-in-the-middle (MITM) attacks.

• Against Malicious Insider (Read Access):

- **File Confidentiality:** The admin can access the `uploads/` folder but will only find AES-GCM encrypted blobs. They can access the `file_metadata` table but will only find *encrypted* file keys (FEKs). The files are unreadable.
- **Permission Confidentiality:** The admin can see the `acls` table but cannot determine *what* permissions are granted (e.g., 'read' vs 'write'), as they are encrypted.

• Against Malicious Insider (Write Access):

- **File Tampering:** Attempting to modify an encrypted file in the `uploads/` folder will cause the AES-GCM authentication tag to mismatch, resulting in an `InvalidTag` error on download. The attack fails.
- **Permission Tampering:** This is the key defense. If the admin tries to `UPDATE acls SET target_user_id=1` (granting themselves access), the HMAC check will fail on the next read because the stored `hmac` field will not match the new, tampered data. The system will detect this, deny access, and log a critical security event. The admin is powerless to bypass the access control logic.

XVI. IMPLEMENTATION LEARNINGS & PROJECT EXPERIENCE

This project bridged the critical gap between abstract theory and practical application. The process was iterative, and the debugging phase provided the most significant learnings.

A. The Theory-to-Practice Gap

The formal models for DAC and RBAC are clean and mathematical. The implementation, however, is a series of complex, practical decisions.

- "ACL" in theory is a row in a matrix. In this practice, it is an encrypted, integrity-protected, and complex-to-query row in a SQL table.
- "RBAC" in theory is a simple set relation. In practice, it is a recursive, hierarchical SQL query that must be optimized.
- The theory papers do not often discuss the most critical vulnerability: the protection of the policy data itself. Our Zero Trust model addressed this gap directly.

B. The Iterative Debugging Process

The development was not linear; it was a cycle of coding, testing, and fixing. This process was fundamental to understanding the system's inner workings. Key errors encountered included:

- **Environment Errors:** The first hurdle was a `PSSecurityException` on Windows, where PowerShell's `ExecutionPolicy` blocked the `venv` activation script. This was a lesson in how the host OS itself is a security layer.
- **Configuration Errors:** A `KeyError` from `secrets.ini` highlighted the dependency on a correct configuration; a missing key is a fatal, non-recoverable error.
- **Database Errors:** A `sqlite3.OperationalError: near "ALNULL" was a simple typo (NOT NULL)` that brought the entire database creation script to a halt, reinforcing the need for precision.
- **Application Logic Errors:** The most complex bugs were in `app.py`.
 - An `'AssertionError'` and later a `'TypeError'` related to Flask decorators. This was a deep lesson in Python's metaprogramming, requiring the use of `functools.wraps` to preserve function metadata.
 - A `'TypeError: int() ... not 'NoneType''` occurred because the server did not correctly validate form data before processing, a common but critical vulnerability.
- **Cryptographic Errors:** A `'str' object has no attribute 'hex'` error was a simple data type mismatch. In cryptography, however, such a simple bug can corrupt the entire security chain. It highlighted the need for exact precision when handling keys, nonces, and ciphertext.

C. Key Takeaways

- 1) **Security is Holistic:** A secure system is not just one strong algorithm. It is the *entire ecosystem*: the `venv` for isolation, `.gitignore` to protect secrets, server-side-only logic, data validation, and finally, the cryptographic primitives.
- 2) **Centralization is Strength:** The decision to centralize all security logic in the `check_permission` function and all crypto logic in `security.py` was invaluable. It makes the system auditable and ensures that no matter which API endpoint is called, the same, correct security check is performed.

- 3) **Distrust Your Database:** The most valuable security lesson was learning to treat the database as a potential threat. Encrypting and signing the ACLs is the single most powerful security feature in this prototype, moving it from a standard application to a truly "Zero Trust" system.

XVII. PHASE 3: CRITICAL SECURITY ANALYSIS

A robust security analysis must go beyond describing features and instead critically evaluate the system's limitations and cryptographic integrity. This section provides a detailed examination of the current implementation, highlighting its novel aspects while candidly addressing its inherent weaknesses.

A. Cryptographic Analysis & Novelty

The primary novelty of this implementation lies not in the invention of new cryptographic primitives, but in the *application* of established primitives to secure the access control metadata itself, a commonly overlooked attack vector.

- 1) **Cryptographic Binding of ACLs:** Standard database implementations rely on row-level security or simple trust in the database administrator (DBA). Our approach cryptographically binds the permission to its context. By including `resource_id` and `target_user_id` in the HMAC message:

$$HMAC(AIK, enc_perms || res_id || user_id) \quad (4)$$

We prevent "cut-and-paste" attacks. An attacker cannot take a valid, signed "full control" permission blob from a low-value resource and paste it into the row for a high-value resource. The HMAC verification would fail because the `resource_id` component of the message would not match the moved blob. This effectively makes each ACL row non-fungible.

- 2) **Separation of Cryptographic Duties:** The explicit use of two distinct master keys (MEK for confidentiality, AIK for integrity) adheres to the principle of key separation. If the MEK were compromised, an attacker could decrypt files, but they still could not forge new access permissions without the AIK. This compartmentalization limits the blast radius of a partial key compromise.
- 3) **AEAD as a Standard:** The exclusive use of AES-GCM for all file encryption provides built-in integrity for the data at rest. Many legacy systems use AES-CBC, which is vulnerable to padding oracle attacks if not implemented perfectly. GCM eliminates this class of vulnerability entirely by authenticating the ciphertext before any decryption is attempted.

B. Weaknesses and Limitations

Despite these strengths, the current prototype has several identifiable weaknesses that would need to be addressed in a production environment.

- 1) **Key Storage (The "Bootstrap Problem"):** The most glaring weakness is the storage of the master keys

(MEK, AIK) in a flat `secrets.ini` file. While excluded from source control, this file sits unencrypted on the server's filesystem. Any attacker who gains shell access to the server (e.g., via a remote code execution vulnerability in Flask) can immediately read these keys, rendering all other cryptographic protections moot. In production, these must be moved to a hardware-backed KMS (like AWS KMS), where the keys never leave the secure hardware module.

- 2) **Authentication Bypass:** The current system uses a trust-based `X-User-ID` header for authentication. This is a deliberate simplification for the prototype but is completely insecure for real-world use. Any user can impersonate any other user, including the admin, by simply changing this header in their HTTP request. A secure system must implement a robust authentication protocol (like OAuth2 or OIDC) and use cryptographically signed tokens (JWTs) to prove identity.
- 3) **Side-Channel Leakage in RBAC:** The RBAC permission check uses complex, recursive SQL queries. In a very large database, an attacker might be able to perform timing attacks. By measuring how long the `check_permission` function takes to return a "403 Forbidden," they might be able to infer the depth of the folder structure or the complexity of the permission rules, leaking metadata about the organization's structure even without direct access.
- 4) **Lack of Key Rotation:** The current implementation uses static master keys. If a key is compromised, all data must be re-encrypted. A production system needs automated key rotation, where new data is encrypted with a new key, and old data is lazily re-encrypted over time. The envelope encryption scheme used here supports this (only the FEKs would need to be re-encrypted with the new MEK), but the logic is not yet implemented.
- 5) **Database Scalability vs. Security:** The recursive SQL queries used for RBAC inheritance checks can become a performance bottleneck in a deep folder hierarchy. While not strictly a security vulnerability, a Denial of Service (DoS) attack could be launched by an authenticated user repeatedly requesting access to a deeply nested resource, exhausting database CPU resources.

XVIII. CONCLUSION AND FUTURE WORK

This project successfully completed both its theoretical (Phase 1) and practical (Phase 2) objectives. Phase 1 established the formal definitions, advantages, and disadvantages of the Discretionary and Role-Based Access Control models.

Phase 2 built upon this foundation to create a tangible, secure software prototype. By designing a "Zero Trust" architecture, we demonstrated that it is possible to build a system that is secure not only against external attackers but also against malicious insiders with direct database access. The use of modern, authenticated encryption (AES-GCM) and integrity checks (HMAC-SHA256) to protect not just the data, but

the *access control rules themselves*, represents a robust and modern approach to security.

Phase 3 critically analyzed this implementation, confirming its resistance to insider threats while securely identifying areas for future hardening.

The resulting system demonstrates that it is possible to build an access control engine that is secure not only against external attackers but also against malicious insiders with direct database access. The novel application of HMAC to bind ACLs to their context represents a robust, modern approach to database security.

A. Future Work

This prototype is a strong foundation, but for a production-grade system, several enhancements would be necessary:

- 1) **Proper Authentication:** The `X-User-ID` header would be replaced with a secure authentication mechanism, such as OAuth 2.0 and JWTs (JSON Web Tokens).
- 2) **True Key Management:** The `secrets.ini` file would be replaced with a real, network-attached Key Management Service (KMS) or a Hardware Security Module (HSM).
- 3) **Production-Grade Server:** The Flask development server would be replaced with a robust WSGI server like Gunicorn, running behind an Nginx reverse proxy.
- 4) **Evolve to ABAC:** The next logical step would be to enhance the `check_permission` engine to include attributes (e.g., time of day, user location), evolving the system toward an Attribute-Based Access Control (ABAC) model.

ACKNOWLEDGMENT

I would like to express profound gratitude to **Prof. Manikantan Srinivasan** from the Department of Computer Science & Engineering at the Indian Institute of Technology Madras for his exceptional guidance and mentorship throughout this project. His invaluable feedback on the theoretical report and his continuous support during the practical implementation and debugging were essential to the successful completion of this work.

REFERENCES

- [1] M. A. Harrison, W. L. Ruzzo, and J. D. Ullman, "Protection in operating systems," *Communications of the ACM*, vol. 19, no. 8, pp. 461-471, Aug. 1976.
- [2] D. F. Ferraiolo, D. R. Kuhn, and R. Chandramouli, *Role-Based Access Control*, 2nd ed. Norwood, MA: Artech House, 2007.
- [3] W. E. Boebert and R. Y. Kain, "A Practical Alternative to Hierarchical Integrity Policies," in *Proceedings of the 8th National Computer Security Conference*, 1985, pp. 18-27.
- [4] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman, "Role-Based Access Control Models," *IEEE Computer*, vol. 29, no. 2, pp. 38-47, Feb. 1996.
- [5] V. C. Hu, D. F. Ferraiolo, R. Kuhn, et al., "Guide to Attribute Based Access Control (ABAC) Definition and Considerations," *NIST Special Publication 800-162*, Jan. 2014.
- [6] A. Ronacher, "Flask - A Python Framework for Web Applications," *Flask Website*. [Online]. Available: <https://flask.palletsprojects.com/>
- [7] The Cryptography Authors, "Cryptography," *Python Cryptographic Authority*. [Online]. Available: <https://cryptography.io/>

- [8] D. R. Hipp, "SQLite," *SQLite Website*. [Online]. Available: <https://www.sqlite.org/> 2010. [Online].
- [9] J. Kindervag, "Defining The Zero Trust Model," *Forrester Research*,