

Deliverable 4

Team Name: DAAP-SecDov-518

MILESTONE REPORT

Fixes Implemented:

Securing the Database:

The application database was initially exposed on port 27017:27017, and lacked authentication controls. This allowed for easy inspection of the database during early development, however it exposed the database to anyone with network access to the host environment. Having removed this port mapping, access into the database container is no longer allowed by simply executing mongosh with no parameters.

```
[alex@Mac ~ % mongosh
Current Mongosh Log ID: 691a7480583c48ffb854b298
Connecting to:      mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.5.9
MongoNetworkError: connect ECONNREFUSED 127.0.0.1:27017
alex@Mac ~ % ]
```

The database container can now be entered using docker exec -it daap_secgov_mongo mongosh, however authentication is now required to execute commands.

```
[alex@Mac ~ % docker exec -it daap_secgov_mongo mongosh
Current Mongosh Log ID: 691a7801829bfd77e1b1ddf3
Connecting to:      mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.5.9
Using MongoDB:    7.0.25
Using Mongosh:    2.5.9

For mongosh info see: https://www.mongodb.com/docs/mongodb-shell/
[test> show dbs
MongoServerError[Unauthorized]: Command listDatabases requires authentication
test> ]
```

Access control was implemented by reconfiguring MongoDB to create a root administrative user upon initialization. Database operations now require authentication through prior input of defined username and password credentials. Unauthenticated operations are now rejected as shown in the image below:

```
[test]> use admin
switched to db admin
[admin]> show dbs
MongoServerError[Unauthorized]: Command listDatabases requires authentication
[admin]> db.auth("SecDovDB_Admin", "StrongDBKey_98hvA1bQ")
{ ok: 1 }
[admin]> show dbs
admin      100.00 KiB
config      60.00 KiB
daap_seciov 304.00 KiB
local      72.00 KiB
[admin]> use daap_seciov
switched to db daap_seciov
[daap_seciov]> show collections
friend_requests
group_invitations
group_messages
groups
messages
users
[daap_seciov]> db.users.find()
[
  {
    _id: ObjectId('691a7a9f996b6bc3ebab7d6b'),
    username: 'amillerg',
    email: 'amillerg@buffalo.edu',
    password: '123',
    friends: [ 'test' ],
    login_count: 1,
    login_times: [ '2025-11-17 01:31:02' ]
  },
  {
    _id: ObjectId('691a7aab996b6bc3ebab7d6c'),
    username: 'test',
    email: 'test@gmail.com',
    password: '456',
    friends: [ 'amillerg' ],
    login_count: 1,
    login_times: [ '2025-11-17 01:30:25' ]
  }
]
daap_seciov> █
```

Password Hashing and Brute Force Mitigation:

The application database originally stored the passwords of users in plaintext. This posed a significant security concern as anyone with access to the database could directly view and reuse user credentials. The password-hashing function Bcrypt was used to remedy this vulnerability, as recommended by the attacking team. During registration, users' passwords are now salted and hashed with Bcrypt. This hash is then stored in the application database, replacing the previously stored plaintext version. We also use the same process during password reset. To avoid timing differences between valid and invalid usernames, we hash all attempts, not just those of correct usernames (we use a dummy hash for invalid users). These hashes can be viewed in the database through the authenticated admin account as shown

below:

```
[alex@Mac ~ % docker exec -it daap_secdov_mongo mongosh
Current Mongosh Log ID: 691a9bcabfc27bce2eb1ddf3
Connecting to:      mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.5.9
Using MongoDB:      7.0.25
Using Mongosh:      2.5.9

For mongosh info see: https://www.mongodb.com/docs/mongodb-shell/

To help improve our products, anonymous usage data is collected and sent to MongoDB periodically (https://www.mongodb.com/legal/privacy-policy).
You can opt-out by running the disableTelemetry() command.

[test> use admin
switched to db admin
[admin> db.auth("SecDovDB_Admin", "StrongDBKey_98hvA1bQ")
{ ok: 1 }
[admin> use daap_secdov
switched to db daap_secdov
[daap_secdov> db.users.find()
[
  {
    _id: ObjectId('691a9b82b5910887772d8a74'),
    username: 'amillerg',
    email: 'amillerg@buffalo.edu',
    password: '$2b$12$EqgDzJ4nIN88Cq0xWwxAx.f1KyGH9b9zf0580/WEpvgAKxiV0/6e',
    friends: [ 'test' ],
    login_count: 1,
    login_times: [ '2025-11-17 03:51:17' ]
  },
  {
    _id: ObjectId('691a9b91b5910887772d8a75'),
    username: 'test',
    email: 'test@gmail.com',
    password: '$2b$12$45Pfnd0k/u3YIBv01C60m.qcRkzAA751R1gb37938zI53XvrdezU6',
    friends: [ 'amillerg' ],
    login_count: 1,
    login_times: [ '2025-11-17 03:50:46' ]
  }
]
daap_secdov> ]
```

Bcrypt includes a tunable work factor that determines how computationally expensive the hashing process is. This work factor is encoded into each Bcrypt hash string along with the salt and derived hash output. Our application defines a work factor of 12, meaning Bcrypt performs $2^{12} = 4096$ rounds of its key derivation process for each password. This intentionally slows the password hashing process. Normal users are minimally affected by the slowdown, but large scale brute force cracking attempts now take significantly more time to execute.

HTTPS/TLS



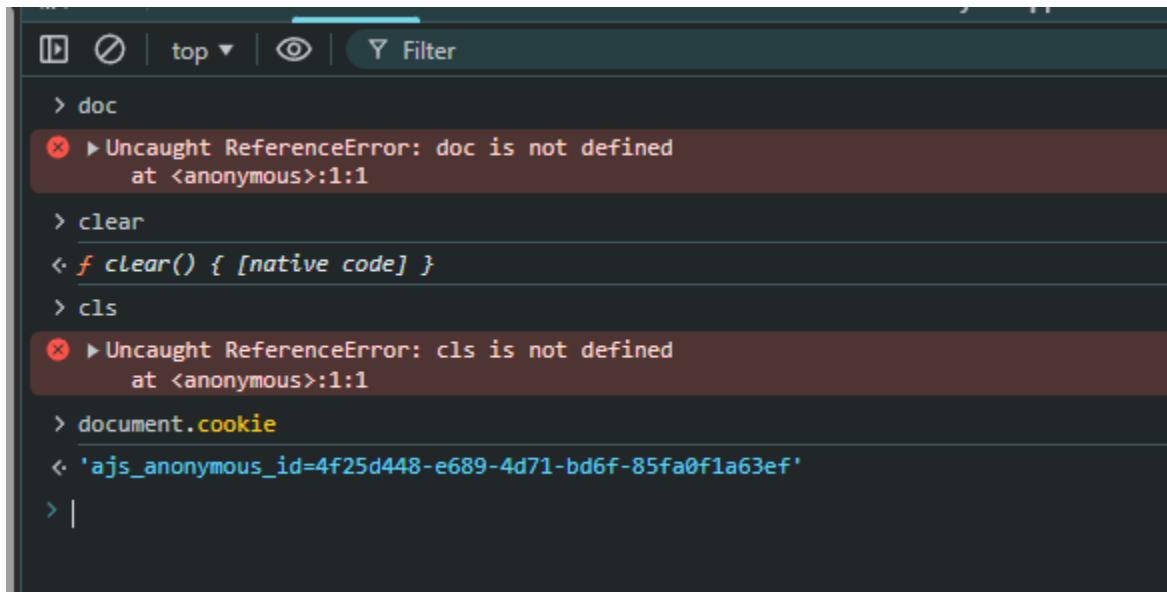
Welcome to Secure Messenger

[Login](#) | [Register](#)

Ready for some secure messaging fun?

To provide secure communication, HTTPS was implemented using Transport Layer Security (TLS). This ensures that all data transmitted between the user's browser and the server is encrypted in transit, preventing attackers from intercepting or reading sensitive information such as login credentials, messages, and session data. A TLS certificate was generated and configured inside the Dockerized Flask application, allowing the web server to serve content over HTTPS rather than plain HTTP. This significantly improves confidentiality and protects against man-in-the-middle (MITM) attacks by guaranteeing encrypted and authenticated connections.

JS cookie protection



The screenshot shows a browser developer tools console with the following content:

```
> doc
✖ > Uncaught ReferenceError: doc is not defined
      at <anonymous>:1:1

> clear
< f clear() { [native code] }

> cls
✖ > Uncaught ReferenceError: cls is not defined
      at <anonymous>:1:1

> document.cookie
< 'ajs_anonymous_id=4f25d448-e689-4d71-bd6f-85fa0f1a63ef'

> |
```

To prevent session hijacking through JavaScript-based attacks, security flags were added to session cookies. Specifically, the **HttpOnly** flag was enabled to ensure that cookies cannot be accessed by client-side JavaScript, even if an attacker injects malicious scripts via XSS vulnerabilities. Additionally, the **Secure** flag was enabled to ensure cookies are only transmitted over encrypted HTTPS channels. Together, these settings significantly reduce the risk of cookie theft and unauthorized account access by protecting sensitive information on the client side.

Password Complexity for Users

Register

[Home](#) | [Login](#)

Username:

Email:

Password:



Please match the requested format.

Must be at least 8 characters, include 1 uppercase letter, 1 number, and 1 special character.

The registration system was enhanced by implementing strong password complexity requirements. Users are now required to create passwords that are at least eight characters long and include at least one uppercase letter, one numeric digit, and one special character. This validation is enforced on the backend to prevent bypass attacks and on the client side for better user experience. These controls reduce the risk of successful brute-force and credential stuffing attacks by making passwords significantly harder to guess or crack. Even if there is a database leak and the hashed passwords are found they will be significantly harder to crack since they will not be found in something like a rainbow table.

We also increased the complexity of the password reset codes from being 6 digits to being 8 digits and, per the TA suggestion from our original suggestion, they now last for only 5 minutes. We are also using the secrets module instead of random to make our codes more secure.

Security Audit Logs

```
daap_secdov_web | 2025-11-22 08:39:00,046 - werkzeug - INFO - 172.27.  
224.1 - - [22/Nov/2025 08:39:00] "GET /login HTTP/1.1" 200 -  
daap_secdov_web | 2025-11-22 08:39:09,850 - security_audit - WARNING  
- FAILED_LOGIN: Username 'taco' from IP 172.27.224.1  
daap_secdov_web | 2025-11-22 08:39:09,852 - werkzeug - INFO - 172.27.  
224.1 - - [22/Nov/2025 08:39:09] "POST /login HTTP/1.1" 302 -  
daap_secdov_web | 2025-11-22 08:39:09,870 - werkzeug - INFO - 172.27.  
224.1 - - [22/Nov/2025 08:39:09] "GET /login HTTP/1.1" 200 -  
[]
```

```
daap_secдов_web | 2025-11-22 08:37:59,010 - security_audit - INFO - S  
UCCESSFUL_LOGTN: User 'pb44' from IP 172.27.224.1
```

```
daap_secдов_web | 2025-11-22 08:37:59,012 - werkzeug - INFO - 172.27.  
224.1 - - [22/Nov/2025 08:37:59] "POST /login HTTP/1.1" 302 -
```

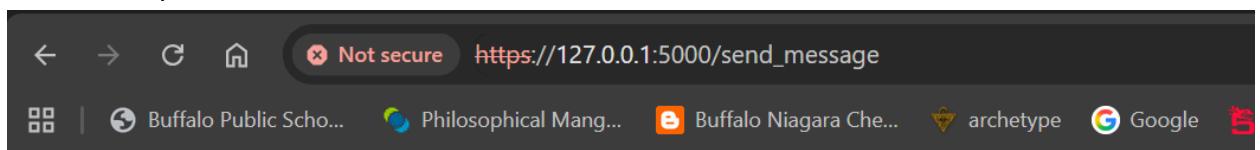
```
daap_secдов_web | 2025-11-22 08:37:59,042 - werkzeug - INFO - 172.27.  
224.1 - - [22/Nov/2025 08:37:59] "GET / HTTP/1.1" 200 -
```

New error auditing logs were added to record audit things like logins and failed logins. Currently these are visible in system logs and on the browser console but in practice in the live system, it would only be visible to admins in the server logs.

User Constraints On Functionality For Security and Sanitization

We restricted a lot of different features in our system to help protect users while still maximizing regular usability as much as possible. This ranged from limiting where/how javascript code could be executed on the site to constraints on how many messages a user could send or how big the files they could send are. We also sanitized user inputs to prevent possible avenues for user code. Here is a list of changes to carry this out:

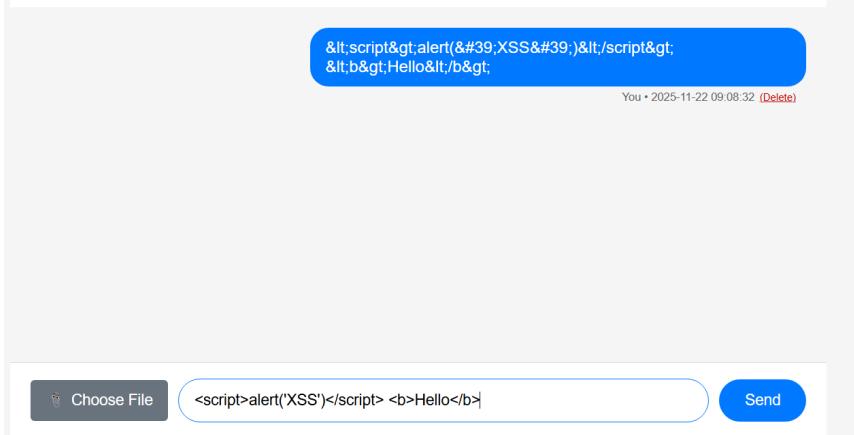
- Limited file uploads to be 16mb in size to help prevent the system from being overloaded from large files (potentially from a DoS attack)
- Set specific limits over a certain period of time for each of the POST routes. For example, messages are limited to 30/minute as that is pre-defined by the route. See below example:



Too Many Requests

30 per 1 minute

- We added security headers to restrict the browser to only use the file type specified, we restricted the resources that a browser was able to pre-load to be a small set (allowing things like the necessary components for Google's reCAPTCHA, prevented most referral information from being sent to other sites (if navigating to our site) outside our own site, and specifically prevented API calls to the camera, microphone, and geolocation.
- Escaped both upload file names and all messages inputted by the user to prevent malicious access to internal files and prevent users from coming in contact with malicious code.



Above: Example of escaped Text as part of sanitization

Security Enhancement with Testing summary

Based on the security testing undertaken in Deliverable 3, we discovered a number of major security issues with our messaging application and developed a comprehensive plan to handle them.

We conducted thorough testing to validate that all security improvements work correctly against the original attack methods.

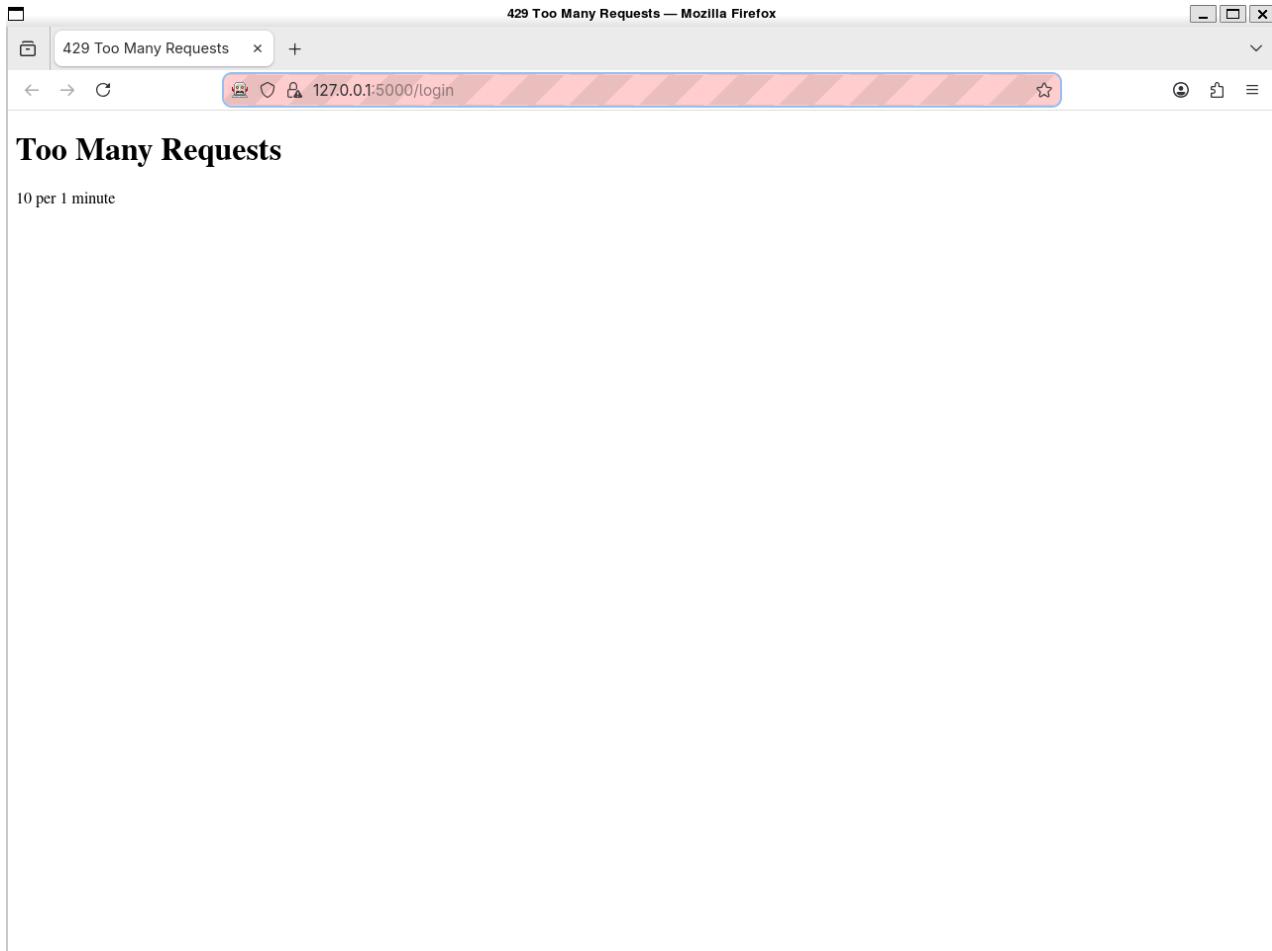
We tried accessing the database without credentials and found that it indeed requires authentication now. Testing password security revealed passwords are hashed appropriately in the database and cannot be accessed in plain text. Brute force attacks that were possible earlier are prevented by our rate limiting implementation. We confirmed that weak passwords, such as "123", are denied at both registration and password change time.

Session security testing confirmed cookies are now flagged with `HttpOnly`, which prohibits JavaScript from accessing them. All previous attacks that succeeded in Deliverable 3 no longer work against our secured implementation.

These security metrics have significantly improved in all aspects. Where previously confidentiality was completely compromised and all messages were exposed, our current

implementation maintains proper message privacy. Integrity protection is restored, and unauthorized modification of messages can no longer take place. System availability has increased owing to the rate limiting protections provided, which maintain service during attempts to attack.

To confirm these results, we reran the exploits and analysis conducted by the partner team. The code we ran and the analysis we performed can be found in the `Attacks_Faced_And_Analysis` folder in our repo. To start with, we ran the same `BruteForce` script that the partner group ran (just updating the url to include the https).

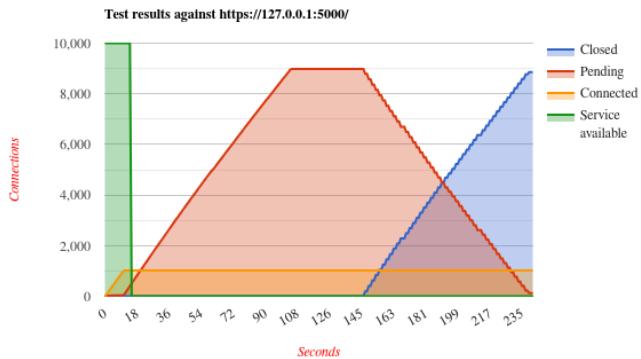


As you can see above, due to our limits on routes as explained previously, this code quickly ran into this “too many requests” block out. This prevented the brute force attack from being able to guess the password due to the block out.

We then also ran the same `slowhttptest` code that the partner group ran (updating again for the URL). This is a simulation of a DoS attack, simulating 10000 connections to our application over

a 240 second period. This is a very extreme attack. The results are shown below:

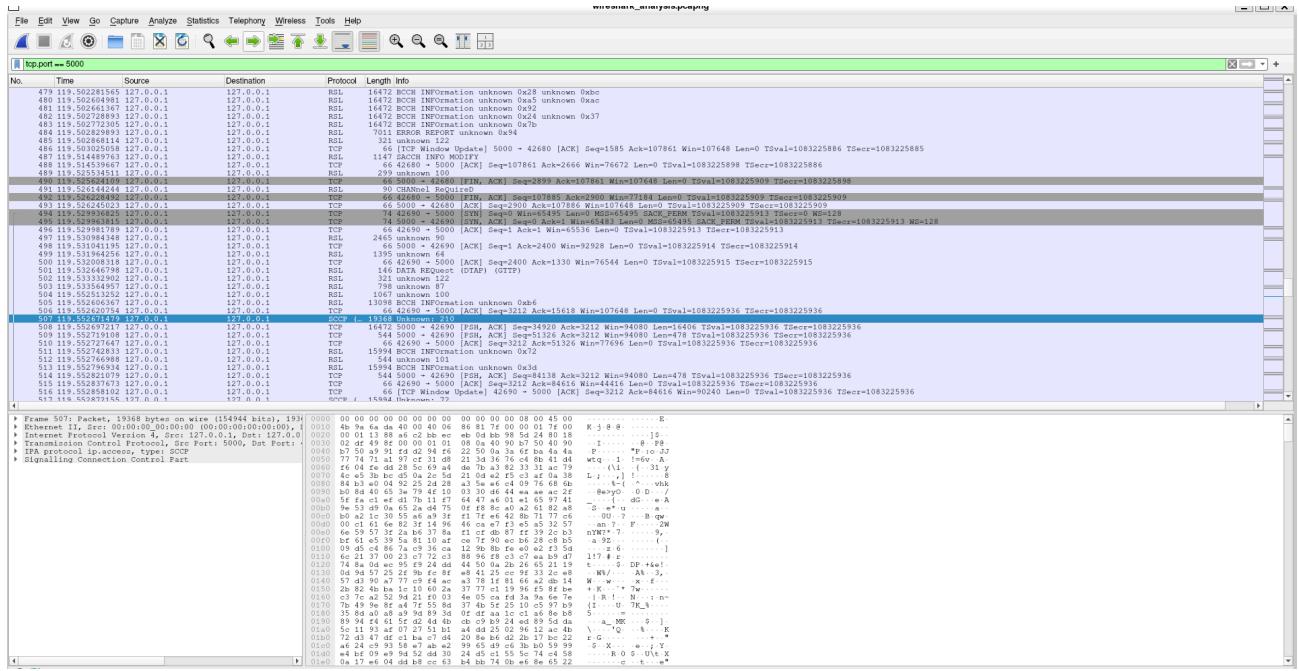
Test parameters	
Test type	SLOW HEADERS
Number of connections	10000
Verb	GET
Content-Length header value	4096
Cookie	
Extra data max length	52
Interval between follow up data	10 seconds
Connections per seconds	200
Timeout for probe connection	3
Target test duration	240 seconds
Using proxy	no proxy



Per the partner team, our previous system did not seem to properly close connections and never fully recovered even well after the test. As you can see, while the service was still taken down during the attack, it did show a marked improvement compared to the partner team's testing. Our system successfully recovered and closed out connection successfully as it slowly caught up after the attack completed. Roughly 20-30 seconds after the end of this test, the application was fully functional again (tested through curl).

For completeness, we still did a wireshark analysis on the site as we used the site to login, send messages, send attachments, and logout. A sample of what information was visible from this

analysis can be seen below:



The partner team was able to see what routes were being used as well as the plain text username and password through this analysis. However, currently our system is sending information through https and important information is being encrypted. Even the names were not discoverable through this analysis. The full analysis file can be found in our [Attacks_Faced_And_Analysis](#) folder.

We also did a bandit analysis. Our partner team previously found several issues with our code using this analysis including hard coded passwords, using the random module instead of the secrets module for more security, and having the debug flag set to true which could allow for the execution of arbitrary code. In our current analysis, these issues all resolved (gives a false positive error for hard coded passwords as it misinterprets regex). The only other issue the analysis came up with was “possible binding to all surfaces” which seems to be a result of running the application in the docker.

These validation results confirm that our security enhancements successfully addressed all the critical vulnerabilities identified during the previous testing phase.

Updated Security Goals

Confidentiality

Updated-information is now encrypted when passing through the application

- Users should only be able to view the messages sent by them to others or messages sent to them by others.
 - Message history will show older messages sent to and from the user.

- Users should be able to upload and share small files with whom they want to, and the other users shouldn't be able to tell or see what the contents of the file are.
 - The file upload and sharing will be done safely with encryption so the contents are not seen by a third party.
- Users should only be able to see groups they are invited to and accepted into.
 - Groups will have a list of members that are in them that can't be seen by users outside of the group.
- Only group owners should be able to delete the groups, message owners should be able to delete their messages, and account owners should be able to delete their accounts.
 - Only authorized users can see and delete groups, messages and accounts that will be the owners of all.
- **Updated** Only users should be able to have access to their own passwords. **Passwords are stored hashed inside the database. Brute force attempts get circumvented with route limiting.**

Integrity

- File upload will be checksummed and checked on arrival
 - Checksum will be calculated and sent with the file and will be checked on arrival to make sure that the file wasn't tampered with.
- Users won't be able to edit or delete messages sent by others.
 - Messages should not be editable or deletable by people who didn't send them.
- Users outside of the group should not be able to change the users in a group
 - Only users inside the group can leave or join a group.
 - Only the group owner can invite new people to the group.
 - Users on the outside of the group can't join or tamper with the users inside the group.

Availability

- Users should be able to send messages at all times to whoever they want
 - Messaging is available to users whenever they require it.
- Users should be able to upload and send files at all times to whoever they want.
 - File uploading and sharing are available to users whenever they require.
- Users outside of a group can't change group settings and members
 - Users outside of the group can't see the member inside, and the group is not available to them unless the group owner invites them
- **Updated** If the application does go down, it should be able to recover and return to full availability as the attack passes. We currently have route limiting implemented to prevent attackers from sending unlimited requests.

Security Metrics

Confidentiality -

- Percentage of messages that are viewed by a user other than the intended recipient or recipients. Note that this also includes messages in groups that are either viewable by users not in that group OR messages in groups that are viewable by users in that group, but who became active users (accepted invites) of that group after those messages had been sent (as users should not have access to messages from before they joined the group). Users should also not have access to group messages once they leave a group.
 - Target will be 0% and will be greater than 0% if any vulnerabilities are discovered
- **Updated** Percentage of passwords or login codes that attackers can access due to brute force or database leaks. Any passwords stored as plain text is unacceptable.
 - Target will be 0% and will be greater than 0% if any passwords are exposed (to attackers or even to admins)

Integrity -

- Percentage of messages that have their content modified in any way (either by accident, through malicious means, or through failures/bugs) after being sent by a sender
 - Target will be 0% and will be greater than 0% if any vulnerabilities are discovered
- Percentage of messages that have their sender field modified in any way (either by accident, through malicious means, or through failures/bugs) after being sent by a sender to appear as if they are coming from another user
 - Target will be 0%, and will be greater than 0% if any vulnerabilities are discovered

Availability

- Server downtime percentage
 - **Updated** Target will be **5%**. It must be able to recover on its own after a DoS attack and return to full functionality
- Percentage of server database updates that fail to complete
 - Target will be 0% and will be greater than 0% if the database does not properly update after applicable user actions (sending messages, uploading files, deleting messages, audit logging, etc.). for any reason
- Percentage of database deletions done by attackers
 - Target will be 0% and will be greater than 0% if an unauthorized user can exploit a vulnerability to delete elements from the database