

Development Assignment Research

1. The access to a file or directory is managed using what is called an 'Access Control List (ACL)'. ACL is a list that contains the following fields:

a) type – it gives us a set of users that are permitted to access the file.

b) role – it gives us a specific task that a given user is allowed to perform. For eg – reading the file, adding comments, writing the file, etc.

c) email address or domain – it specifies the individual users that have access to that file.

For each role, the other fields are specified which gives us the complete set of access permissions. Following is an example of ACL :

Permitted operation	organizer/owner	writer	commenter	reader
Read the metadata (e.g. name, description) of the file or folder	✓	✓	✓	✓
Read the content of the file	✓	✓	✓	✓
Read the list of items in the folder	✓	✓	✓	✓
Add comments to the file	✓	✓	✓	
Modify the metadata of the file or folder	✓	✓		
Modify the content of the file	✓	✓		
Access historical revisions	✓	✓		
Add items to the folder	✓	✓		
Remove items from the folder	✓	✓		
Delete the file or folder	✓			
Share a Team Drive item	✓	✓		
Add files to Team Drives	✓	✓		
Delete items in Team Drives ¹	✓			
Move items in Team Drives ¹	✓			
Edit Team Drive metadata	✓			
Add Team Drive members	✓			
Delete an empty Team Drive	✓			

A file should contain an entire ACL as its attributes to be able to decide whether a given user is authentic or not.

In my opinion, a good design to handle concurrent editing would be where we have a database storing a given file. Whenever a user requests access to a file, the server verifies the credentials of the user against the ACL of that file and if the user is

authentic, the server reads out the file to the user. Since there might be other users trying to edit the file at the same time, we cannot take the file from the first user once he had completed making all the changes and then update it in our database. Rather, the server will take every edit request made by the user, update the file in the database and then read out the updated file to the user. This way changes made by two user simultaneously will also show up to both the users.

2. A relational database (RDB) is a collection of data organised in the form of tables such that the data can be reassembled in different ways without having to reorganise the tables. For eg. In case of MySQL, the join feature allows joining columns from different tables based on some condition that they satisfy.

A database schema is a blueprint of all the data present in the database. It tells us all the relations between the data and the constraints that are to be applied on the data.

Schema for academic system at IITD:

S.No	Table Name	Columns
1	StudentInfo	Entry Number, Name, Department, Enrolled Courses
2	CourseList	Course Code, Name of Course, Course Coordinator, Instructors, Credits

Apart from these, there will be a separate table for each of the course whose name will be the course code. It will have the following columns – EntryNumber of Student, Group, Grade.

3. Reverse tunnelling refers to creating a connection between a specific device to the ssh host on the same network.

Transmission Control Protocol (TCP) is a standard way of creating and managing a connection through which application programs can exchange data.

User Datagram Protocol (UDP) is another protocol for sending packets over the internet. UDP is faster than TCP as it does not verify whether the packets have been successfully received. So it is used where speed is required. TCP is more reliable and checks for errors and successful delivery of packets.

When we enter a URL, the browser first searches it in the cache. If present it decodes the response that it had got earlier else get the IP address of the server by a DNS call made by the operating system (domain name service converts domain name to IP address). Then the browser opens a TCP connection to the server and sends an HTTP request via it. On receiving the response, the browser checks what type of response it is and acts accordingly. The response is stored in cache if possible. Then the response

is decoded and the browser renders the response in case of HTML, image, etc. or opens a download dialogue for unrecognized types.

To show files over the same wifi without uploading, first create a local server from the source computer by typing (`python -m http.server (port_number)`) in the terminal. Now find the ip address of this device by typing (`ifconfig`) in terminal. From another device over the same wifi, type (`ip_address):(port_number)` in the browser. This will open the directory in which localhost server was set up.

This will not work if the two devices are on two different networks for eg. IITD wifi and mobile network. This is because the Local IP address of the source computer is known only to the wifi it is connected to. Here we are not using the public IP of the computer (which will be known across all networks). Hence only this wifi will be able to route to the source computer.

NginX and Apache, both are web server softwares. A web server is a accepts an HTTP request from a web browser and returns the response. It contains the source code of the websites and gives them to the browser in the form of response.

Think

According to me, the best approach would be to already have the zip file for each course stored on the hard disk of the server and instantly serve the zip of the course when requested. Since zipping is a costly operation, our motive is to reduce the number of times we zip the files.

Zipping the files upon download request would have been better in case the frequency of uploads(changes) to the folder was higher than the frequency of download(ie. Zipping) because then we would save on the computation as we would not zip the files repeatedly on each upload.

But when the number of downloads is much higher than the number of uploads, then zipping will be less frequent if we keep a zipped folder in the hard disk of server and update it whenever there is a new upload.

In case of Citadel, the upload happens only 2-3 times in a semester but every student downloads the zip. Hence upload frequency is much less than download frequency so having a zipped folder in hard disk of server will be better.

Time to Code

Database model – relational model

database used = MySQL

Tables:

1. name = login
columns = username(primary key), password, usertype (prof, student or admin)
2. name = reg
columns = course, 1 column for each student (type = timestamp)
If a student is registered in a course, the corresponding cell will have the timestamp of the time of registration, otherwise it will be null.
3. name = messages
columns = course, message, time (type = timestamp)
time contains the timestamp of the time when message is posted

To provide access control, the username and password are checked in the table login. If such a record is present, the user is redirected to the respective page else the login page reloads.