

Gunari Bharath Chandrashekar
Graduate Student
Master in Computer Science
Arizona State University. Tempe AZ 85281

Email: bgunari@asu.edu

Phone: (480)-329-5244

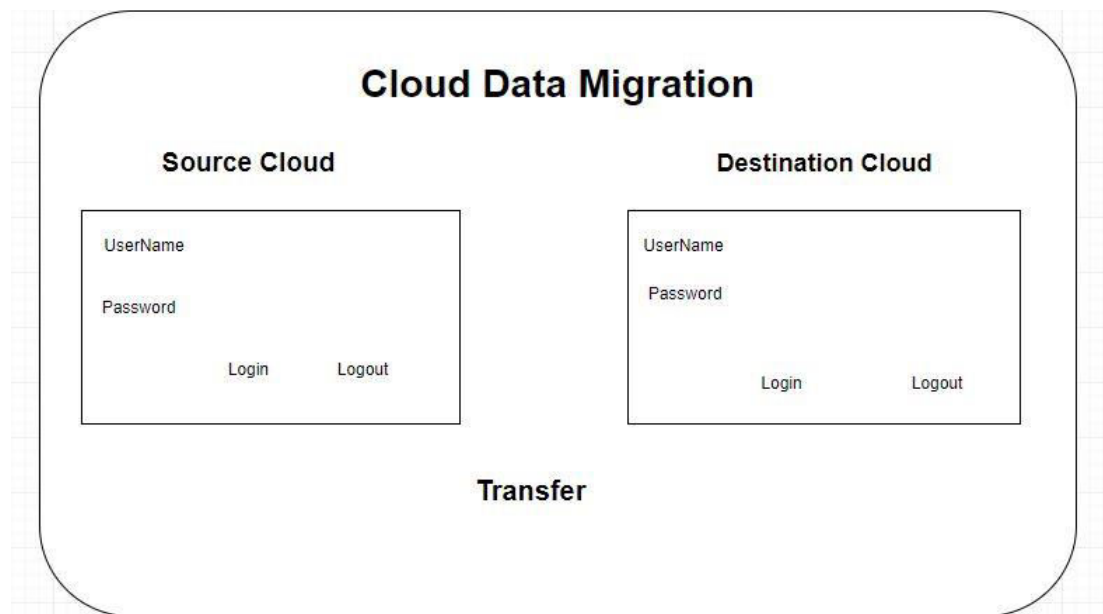
Data Migration Between the Clouds over the network.

Day to Day the amount of data being generated is growing in exponential rate. And people and companies are relying on the cloud storages for their data to be dynamically accessible. Our basic idea is to build a web platform where a user can transfer his data from one cloud storage service provider to another cloud. For example, from google cloud to amazon cloud drive.

Existing Services: Google Cloud Storage Transfer Services, cloudsfer.com, multcloud.com. All the existing cloud transfer service providers will download the data onto their servers/storage units and then upload the data to the destination servers/storage units (as per my knowledge). There is a downside to these. The user data is being passed through third party apart from source and destination cloud which gives rise to many security and privacy issues.

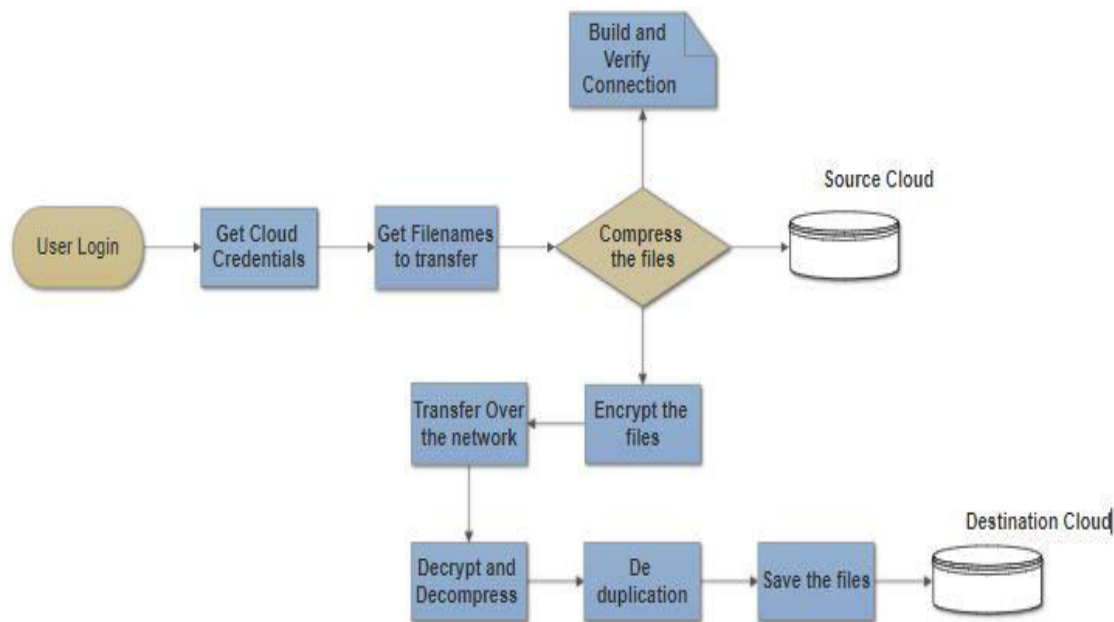
Our Idea in Brief: We transfer the user data from one cloud to another over the network. Our platform will not have user data privacy issue because the transfer happens over the network and files are encrypted while being transferred and report is generated for the user to analyze the transfer. This is a direct transfer from one cloud to another cloud. Our web platform will never access any files. One more best feature of our platform will be providing content based Deduplication feature. Means if a file that is selected to transfer is already existing in the destination cloud. The deduplication is content based. In the sense if two files which have same content but are with different file names it won't be transferred and user is informed about this in the end report generated. Also, the user has wide range of clouds to transfer between.

Idea in detail: (on front-end)



1. The user registers in our web platform by signing up and we verify the user through verification link or OTP. And the user details will be stored in backend database
2. Next page, User selects the source and destination clouds from the dropdown list provided.
3. We take the user credentials of both clouds and access the clouds and send acknowledgment for successful access in both clouds.
4. The files names from the source clouds will be retrieved and displayed to user in form of checking list. User will check all the files which he desires to be transferred. And user will also be given the option for whether the files in source cloud to be deleted or kept even after transfer.
5. The user clicks "Transfer" and all the files corresponding to selected files names will be transferred to destination cloud and it will be deleted if user has given instruction to delete the files in source cloud.
6. A Report will be generated to indicating which all files are successfully transferred and which files were unsuccessful transfer. (this might happen because no all file formats are supported in all clouds).
7. The Report will also show if any files were renamed or changed in any sorts. And it will also give the directory structure of transferred files so that user don't have to waste time searching for files in cloud after transfer.
8. The user will also have a "Logout" button under the respective user credentials block. The user will logout from both clouds by clicking logout button. If a user tries to logout of our platform without logging out from both clouds, he must be alerted advising to logout from respective.
9. The credentials given by the user for cloud access will not be saved by our platform.

Idea in detail: (back end)



1. Our platform transfers the data over network instead of being downloaded onto the third-party storage.
2. The user credentials of our platform are stored or authenticated as user tries to login or signup.
3. When user enters source cloud credentials, the cloud is accessed and the files are retrieved for the user to be displayed.
4. And parallelly the destination cloud is accessed. And a connection between the source cloud and destination cloud is established. Here either we can test the connection between 2 clouds by transferring a file back and forth and set it ready for the transfer.
5. When user clicks "Transfer". The transfer happens over the network in following steps:
 - The files are compressed using lossless compression algorithms. Lossless because the integrity of the user data being transferred is of key importance here. We might have to employ different compression algorithms for different types of files.
 - The compressed files are encrypted using encryption algorithms like AES 256 etc. Files are compressed because to increase the rate of successful transfer of files which are considerably large in size.
 - The compressed and encrypted files are transferred over verified built connection. While files are transferring the user is displayed with progress bar.
 - At the destination cloud the files are decrypted and decompressed and saved. Before saving a file, it is checked for content based deduplication in order save the cloud space.
 - Renaming and other tasks are taken care of and informed to user regarding all the changes in the end report.

Scalability:

1. Our services can be extended to transfer between all the clouds that are existing.
2. We can extend the range of file types supported for the transfer.
3. We can improve our encryption, compression and deduplication features.
4. And also, we can improve the transfer quality and transfer success rate.
5. We can improve handling network failure and other conditions and make our system more robust.
6. We can allow data transfer between 2 users with user 1's cloud as source cloud and user 2's cloud as destination cloud by mutual authorization.

Different Application of same idea:

The data transfer between clouds give rise many other similar applications. We can do data transfer between different application of similar platform. Before starting this project, this was big picture we had in mind. For example, consider apps like WhatsApp and Facebook messenger. Both are messenger apps. If there is a chat between person A and person B in messenger. We build a platform through which we can able to transfer that chat from messenger to WhatsApp between same 2 people. We can try out this data transfer idea across different applications.

Questionnaire:

1. What was your inspiration to develop this project?

Answer: We were searching for a project for our Final year project in our Bachelor's. Even though we were not sure what project, but we were knocking a specific door i.e cloud computing. After constant searching we came across this topic. And we started working on this project.

2. What makes it different from what currently exists?

Answer: This has been explained initially in this document.

3. What was your approach to develop this project using GCP?

Answer: Our initial stage was developing a web platform and do the transfer between 2 specific clouds. Between google drive and amazon cloud drive to be specific. Before we try out transfer between other clouds, we were trying to build a basic and core functioning part which handles the transfer. We got encryption, compression and deduplication only for ".txt" files. We were trying to integrate the functioning to web platform.

4. Have you thought about other practices of developing this project?

Answer: As soon as I finished my undergrad, I came to ASU for my master's. I really didn't get a chance to continue working on this project. And also, all the work we did belong to the undergrad university so that our juniors can continue the work. We don't have the access to the project. I have to start from scratch if I plan to work on this project.

5. While automating the entire deployment process in GCP, what were the challenges you had to face?

Answer: This was a bit complicated project for undergraduate level. So we weren't able to build complete end to end system. So we treated 2 laptops as 2 clouds and demonstrated the file transfer between 2 laptops with compression, encryption and content based deduplication features.

Disclaimer: Me, Gunari Bharath is not responsible if this document raises any patent issues or any kind of issue. This is an honest project idea which I worked on under mentorship of a professor in my undergraduate. If someone/some organization is working or has any rights reserved causing inconvenience , I am not responsible.

Thank you