IBC Project Report 1
Ghanendra Singh MT19213

Progress for March 2021

So far as per the project plan, objectives for the month of March, following things were completed and the link for the source codes is mentioned at the end of report.
First process for sharing human aging genomics data was conceptualized and how it can be implemented using Linux Foundation based open source distributed HyperLedger Fabric blockchain framework. And then the human aging data was gathered from the available website.
Initially the plan is to implement very basic genomic data private sharing between two organizations consisting of two peers each and develop on top of the existing fabric examples and later extend to the complete broader picture of the project at a later stage.

Next, startups were identified and read their case studies to understand the blockchain ecosystem in genomics and how the technology is being used and its future use cases.

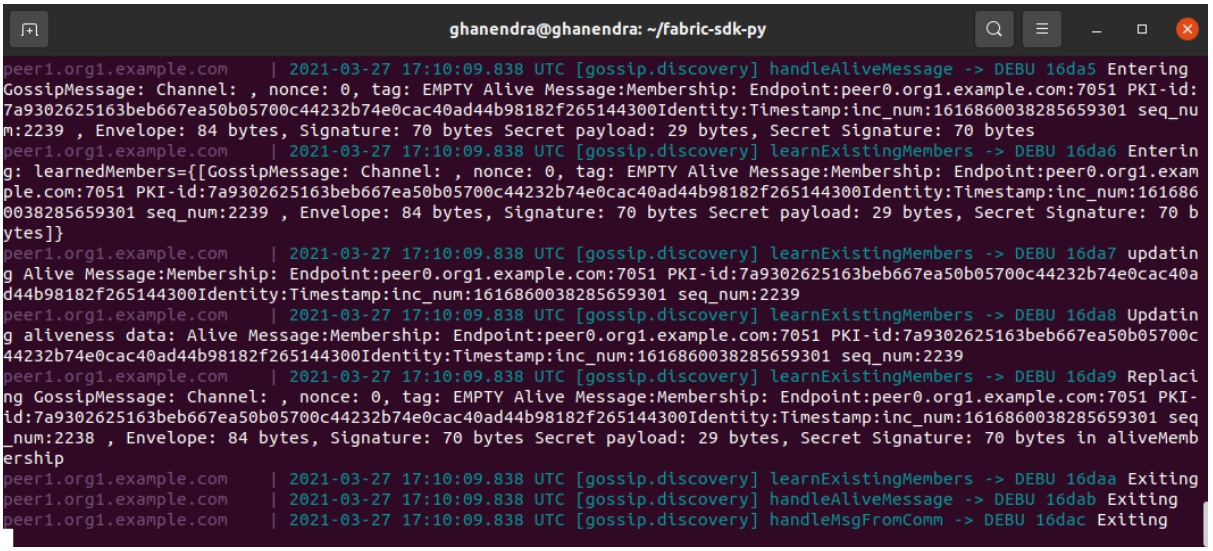
Below is a list of few of the startups utilizing blockchain in genomics.

- **Zenome** is a Russian biotech startup which is using blockchain and genomic data science to empower the genetic testing consumer to control the use and sale of their own data.
    - White paper - The Zenome Project: Blockchain-based genomic ecosystem
- **Nebula Genomics** is a startup out of Harvard Medical School which provides a genomic data sharing and analysis platform based on blockchain technology.
    - Technology - Technology
    - Data Privacy in the Age of Personal Genomics
    - 'Fit-for-purpose?' – challenges and opportunities for applications of blockchain technology in the future of healthcare
- **DNAtix** is Israili biotech startup one of those using blockchain to allow DNA data exchange between individuals.
- **EncrypGen** is a cryptocurrency based genomics data sharing and storing company.
- **Shivom** is a cloud computing based multi-omics informatics platform.
    - Case Studies

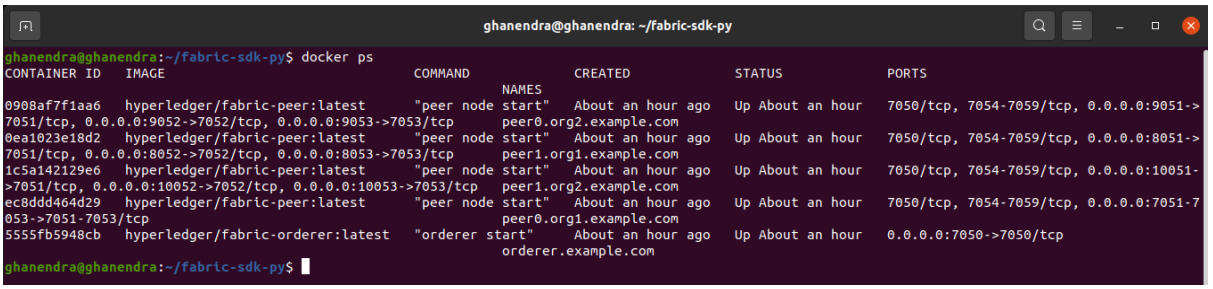
Thereafter, hyperledger fabric fundamentals were learnt to get a rough idea on how to install and use the framework. Identified Fabric related different tools and components which plays an important role and how they interact with each other. Further executed tutorial to learn.

Installation: https://fabric-sdk-py.readthedocs.io/en/latest/index.html
Important note: Installation of docker and docker compose is very essential.
Docker: https://docs.docker.com/engine/install/ubuntu/
Docker Compose: https://docs.docker.com/compose/install/
Tutorials: https://fabric-sdk-py.readthedocs.io/en/latest/tutorial.html

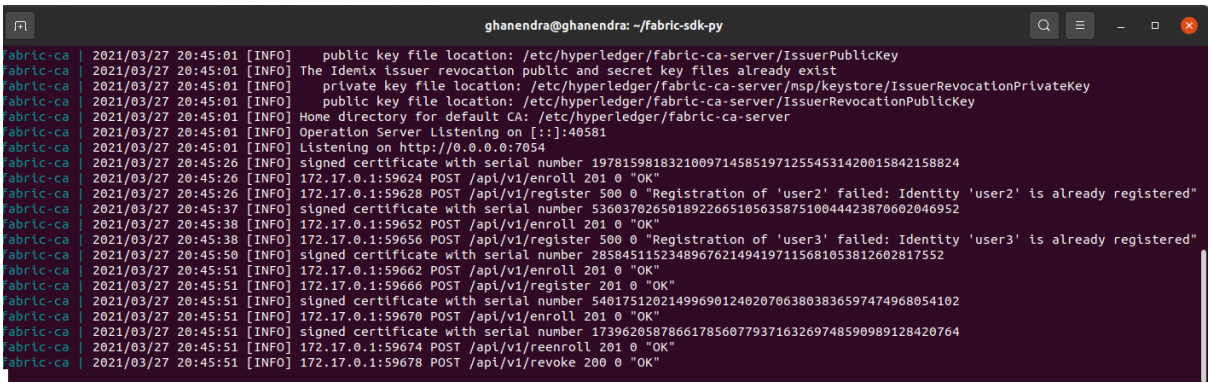
Fabric Network up and Running.



Docker Container up and running



Fabric Certificate Authority (CA)



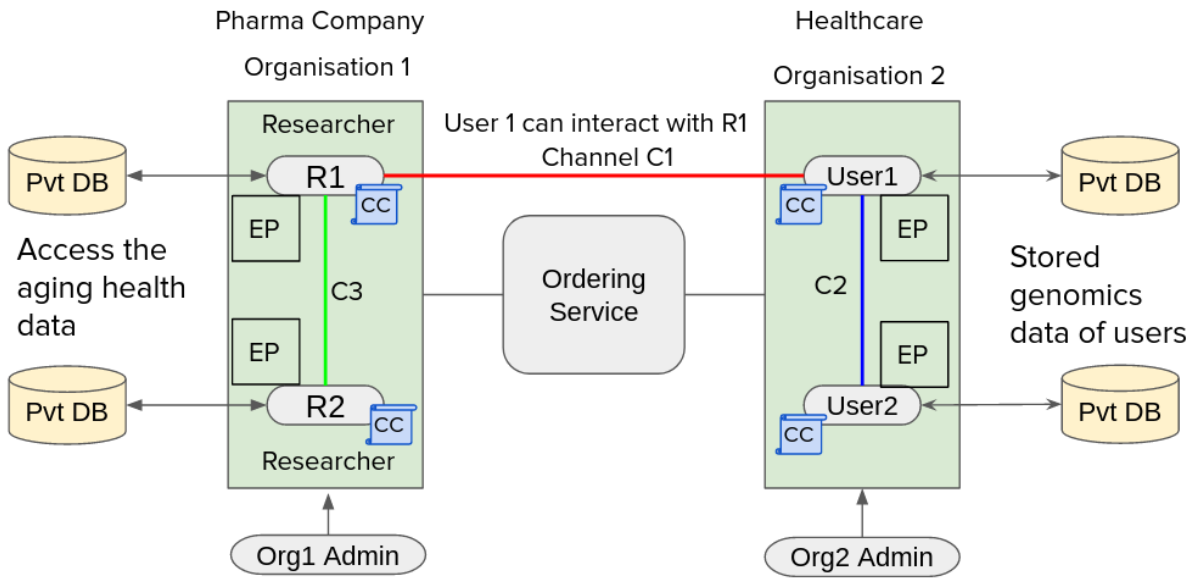
HAGR Data Link: HAGR_data.ipynb
Running Fabric Python SDK Link: Fabric_Network.ipynb

Source code has got many bugs, need to be fixed properly to create chain codes and channels. Next, after getting familiar with fabric networks, aging genomics data must be used as private data in fabric, which requires defining collections which provides an easy way to selectively share data among entities in a channel.

Data Privacy Requirements for Fabric

- Humans aging genomics data should not go through ordering service as part of the transaction.
- All peers (members of the channel) have access to the general genomics information
    - Name of genomics data
    - Type of data
    - Owner of data
- Only a subset of the peers have access to gene therapy pricing information, genomic sharing and read write data accessibility.

Fabric Chaincode Lifecycle



Transaction

It consists of public channel data which goes to orderer and all peers. Hash of the private genomic data is stored in the transaction such that later it can be verified easily with private data by peers as per the defined policy.

Assets

Majorly to share the human aging genomic data across different organizations, different types of aging data are considered as assets which gets transferred between different entities.

Aging Genes of different types

- Longevity genes
- Model genes
- Anti aging Drug
- Dietary Restriction Genes
- Genage models
- Human aging genes

Collection

It consists of private data in which the data is shared between all channel peers but not with the orderer. Currently the collection design is inspired from Marbles asset scenario. At a later stage it will be modified properly

Currently two collections have been defined as a json file.

- collectionAgingGenes
- colllectionAgingGenensPrivateDetails
- **Link:** https://github.com/Ghanendra19213/IBC/blob/main/collections_config.json

```
[
    {
        "name": "collectionAgingGenes",
        "policy": "OR('Org1MSP.member', 'Org2MSP.member')",
        "requiredPeerCount": 0,
        "maxPeerCount": 3,
        "blockToLive":1000000,
        "memberOnlyRead": true
    },

    {
        "name": "collectionAgingGenesPrivateDetails",
        "policy": "OR('Org1MSP.member')",
        "requiredPeerCount": 0,
        "maxPeerCount": 3,
        "blockToLive":3,
        "memberOnlyRead": true
    }
]
```

Summary

Github Project Link: https://github.com/Ghanendra19213/IBC

1. Project conceptualized and process defined to implement it.
2. Identified existing genomics startups and their case studies and some literature review.
3. Building fundamentals of Hyperledger Fabric sdk and its different components.
4. Installed Hyperledger Fabric Python SDK and executed Fabric tutorials.
5. Downloaded HAGR data and converted them into json format in Jupyter Notebook.
6. Defined Fabric chaincode lifecycle for ageing data private sharing. Fabric v2.0
7. **Using Private data in Fabric. Youtube, still its pending not defined properly using the transied field and transient store during chaincode instantiation.**
8. **Creation of database using CouchDB not able to complete as per march plan.**

Thanks