PROJECT INTRODUCTION -

Implement a distributed relational database management system with limited functionality – Any programming language of your preference. The features of this distributed database management system are -

- Homogeneous – all DBMSs same
- Machines are connected by LAN
- Should work at a fragmentation transparency level
- Should support both read-only and update queries
- Four sites with MySQL on them. Could use Postgress too, if you want
- No data replication
- Support for (One Table at least for each of the following kind)
    - Horizontal fragmentation
    - Derived Horizontal fragmentation
    - Vertical fragmentation
- Should work on
    - Distributed join
    - Aggregate/group-by queries

The project components are as follows -

- System Catalog
- Query Decomposer
- Data Localizer/Fragment Query Generator
- Distributed Query Optimizer
- Distributed Query Executor
- Distributed Update Executor

PROJECT PHASE-I -

- System Catalog-
  - Maintains the information about the fragmentation and allocation schema
  - Meta-data about DDBMS
  - Fully replicated (available on all sites)
  - No need to support updates to the system catalog
  - Processing of every query requires system catalog access
  - Design and implement system catalog
  - Parts:
    - Design system catalog (ER and relational schema)
    - Choose a DBMS application (like sales, library, hospital, etc) and prepare schema.
    - Come up with horizontal, derived horizontal and vertical fragments on the chosen application
    - Create and populate system catalog with fragmentation schema and allocation schema on all sites

- Application DBMS -
  - Choose your favorite DBMS application
    - Design its schema
    - ER-Model -> Relational model -> Normalization
  - Fragment and allocate the relations based on common sense/intuition/applications
    - Should support all three types of fragmentation
  - Think of useful applications on it and write the corresponding SQL queries
  - Choose it carefully
    - Going to be used till the last phase

- Phase I - Deliverables -
  - In hard-copy (to be submitted/uploaded):
    - ER Model and relational model of system catalog
    - ER Model and relational model of application DBMS
    - Fragmentation and allocation of application DBMS relations
    - Printout of system catalog containing the fragmentation and allocation information of the application DBMS relations
  - Demo (schedule will be announced later)

        - Create a system catalog on all three machines and fill its appropriately
        - Create application DBMS relations on appropriate sites based on fragmentation and allocation information
    - Codes to create Application Database and System Catalog on all the sites.
    - A detailed pdf report with proper reasoning for choosing the type of fragmentation and allocation.
    - Upload the zip file with all files.
    - **Submission file name format - Teamname_phase1.zip**
    - Only one person from each team needs to submit on the portal.

Deadline for Phase 1: 11:59 pm, 4th February 2022