1. Ongoing Project Background

Many businesses have the goal to establish an Enterprise Data Architecture (EDA) and to promote subsequent activities related to the integration of existing and new projects with the EDA. There are typically three separate efforts that are part of the creation of an EDA:

- Modeling – Creation of a diagram and/or blueprint that support the design of enterprise storage systems
- Operational Data Store (ODS) – Creation of physical database(s) that conform to the model.
- Roadmap – Means to move applications / operations to integrate with the ODS

The motivation that drives businesses towards adopting an EDA approach results from a myriad of data-related problems within the enterprise, including:

- Disjointed / flawed data sources
- Disjointed knowledge about data (experts in a single area / aspect, etc.)
- Too much data (no clear means of moving data as value changes over time, legal aspects, etc.)
- Crisis of data confidence (dirty data, difficult to uniquely ID customers / policies / associates, etc.)

As a result the architectural objectives for EDA include the need to establish the following:

- Data Retention Standards – manages data as its value changes over time (uniform migration / removal data aging data, consensus driven, in process with IT officers, etc.)
- Enterprise Data Model - a diagram (blueprint) that will support the design of storage systems (single integrated source for corporate data, consensus driven, etc.)

The following data storage terminology should be clearly understood when working with EDA:
• TDS (Transactional Data Store)
  o A simple, unambiguous means to track a company’s business
  o Actively being worked on by customers, customer service groups, etc.
  o EDM (long-term future state)
• ODS (Operational Data Store)
  o Daily snapshot of TDS to drive reporting, analytics, etc.
  o EDM (near-term future state)
• DW (Data Warehouse)
  o Tracks data that changes over time
• DM (Data Mart)
  o Stores aggregated and derived ODS / DW data for reporting purposes
• ETL (Extract Transform Load)
  o Provide opportunities for data cleansing

From a cost-benefit analysis standpoint, the following benefits can be attained via EDA:

• Single Source of Data
  o Increased consistency of reports, analytics, etc.
  o Reduction in Nightly Batch
  o Easier to affect changes (living model)
• Single Source of Knowledge about Data
  o Increased accuracy of reports, analytics, etc. (Data Quality Circles)
  o Easier to implement retention policies

2. EDA Modeling Questions

Refer to the sample powerpoint slides provided for this assignment under demo programs on the course Web site.

1. Analyze the sample blueprint and create a corresponding entity-relationship diagram with full documentation. Please use CA Erwin Data Modeler (http://erwin.com/) or an equivalent tool of your choice to draw the entity-relationship diagram.

2. Verify that the various cases described in the powerpoint slides are supported by the model design and if not amend the design accordingly.

3. Deliverables

Please provide an electronic copy of your homework submission as one zip archive by sending it to the course TA by the assignment deadline as noted. The archive should include your E-R model and your homework report (in word or text format). You should name your archive using the following convention for the homework archives: lastname1_lastname2_hw5_fa11.zip. You are also required to provide a hard copy of your homework report at the beginning of the class session on the date the homework is due.

4. Grading

All project assignments are graded on a maximum scale of 10 points. Your grade will be based equally on:
a. The overall quality of your documentation.

b. The understanding and appropriate use of database systems related technologies.

c. Your ability to submit well documented solutions.

d. Extra credit may be granted for solutions that are particularly creative.

5. Additional Information

Please let the TA know as soon as possible about teaming arrangements (only two people per team). You will need to stay with the same team for the duration of the course. You should only submit one report/archive per team for each assignment. To balance things out, the final grading for the course project will take into account the fact that you are working as a team instead of individually, so you should feel free to work individually as well.