Model for estimating large scale data conversion efforts

Sagar Gollapudi

Business Challenges

- M&A , ITIL best practices & SoA has lead to better application portfolio management leading to application rationalization
- Data architecures are gaining more focus industry wide
- Development of Business Intelligence is leading to more need to scalable data architectures

Salient features of projects focusing on data rationalization

- Large collaborative effort in understanding the needs of data conversion
- Tight budget and time criterion
- Large variation in complexity perception
- Lack of good estimates
- Large teams corresponding to skill sets in various diciplines – tools and business objectives

Challenges of data conversion projects

- Constantly changing requirements
- Unsteady data models
- Inadequate requirement emphasis
- Challenges from Data Quality
- Lack of acceptable test data
- Lack of adequate risk identification and mitigation plans

Data conversion challenge

- Lack of investment on enterprise class tools
- Changing budget estimates
- Expanding time horizons on the project
- Lack of definite estimation guidelines for large efforts
- Large variation in estimates between various vendors – credibility of design and approach

Key parameters for estimating the size of a data conversion project - 1

- Number of data stores to be converted
- Size of each data store – design and volume
- Clarity in knowing the final data model
- Complexity of the transformation process
- Design changes with the data models
- Consideration for bi-temporal models
- Platform architecture of source and target

Key parameters for estimating the size of a data conversion project - 2

- Availability of computing resources and inter-platform connectivity
- Availability of the most appropriate tool
- Age of the data – nature of data quality issues
- Availability of knowledgeable resources
- Privacy/ sensitivity content in the data to be converted
- Final use of the converted data
- Rigor needed for the conversion accuracy

Complexity factors

- Number of platform hops needed for transforming the data
- Performance expectation for the conversion process
- Tolerance factors in the data conversion
- BI use versus Application enablement
- Political environment at the business level
- Skill & experience of the resources

Effort estimation guidelines

- Heuristic Graph analysis method
  - Various graphing profiles are created for efforts based off the complexity factors
  - Development effort is estimated based of the conversion size
  - Total project size is calculated based on the profile selected
  - Contineous refinement of the heuristic model is performed by complexity adjustment

Effort Estimation guidelines

- Core factors for profile selection
  - Size of effort
  - Nature of conversion
  - Sensitivity of conversion
- Non Core factors for influencing profile
  - Team knowledge
  - Business Partner knowledge and cooperation
  - Use of ETL tools versus conventional data handling

Criteria for Global development

- Norm rather than exception to leverage low cost location
- Size of the effort
- Sensitivity of the effort
- Availability of best practice library for offshore based delivery

Globally leveraged cost estimates

- Onsite to offshore ratio
- Design to development scaling
- Test development & execution
- Data quality management

Demonstration

- Demonstration of a heuristic estimate model

Case Study of a large conversion effort

- Gee-wiz figures
- Cost model
- Alterations to the cot model to include conversion effectiveness test
- Lessons learned
- Evolution of a data conversion discipline

Summary

- The heuristic graph analysis method has been empirically tested and has proven to provide consistent results