|  |  |
| --- | --- |
| **Document:** | ETL Framework Project Guide |
| **Date:** | 01-March-2012 |
| **Author:** | Pavan Keerthi |
| **Contact Details :** | E-mail- [Pavan.Keerthi@3dmxconsulting.co.uk](mailto:Pavan.Keerthi@3dmxconsulting.co.uk) |

**Version History:**

|  |  |
| --- | --- |
| v0.1 | Initial Draft |
|  |  |

Index

[Introduction 3](#_Toc319940321)

[What does it do? 3](#_Toc319940322)

[Design and Architecture 3](#_Toc319940323)

[SQL App 3](#_Toc319940324)

[Enum Definitions 4](#_Toc319940325)

[How to Use? 5](#_Toc319940326)

[Insert Data into Tables 5](#_Toc319940327)

[Understanding Log Tables 5](#_Toc319940328)

[Developer Checklist 5](#_Toc319940329)

[Troubleshooting Tips 5](#_Toc319940330)

# Introduction

Pretty much every SQL Server BI Project comprises ETL(Extract-Transform-Load) as a major and often complex code modules. There are many common development aspects carried out by every developer in every new project. This “Boiler Plate” code is often written across each project even within same Organisation adding up to a lot of Development time.

The goal of this Project is to standardise the common aspects of the ETL development in SSIS Technology and provide it as a framework. We think this saves lot of important development time and improves quality in the Project.

# What does it do?

The Framework standardises the Concept of building ETL workflow in modular design. Child Packages are like worker processes which does the Heavy lifting work. They can be created from various templates provided which follow Kimball Datawarehouse design practices.

The one control package will manage the entire flow of process and handle error handling, flow order and logging

You are expected to pick up child package templates and modify it according to your requirement.

# Design and Architecture

The Project contains basically

* SQL Database with some tables and Procedures
* Visual Studio Solution with SSIS Project (Contains Templates for Control Master and Child Tasks)
* Powershell scripts to deploy the Project

Also ,you could refer to following documents

* ETL Framework Topology.pdf : Summarises high level Components in Entire Project
* Master Control Package Flow.pdf : Process flow for Control Package
* SQLApp Schema Model.png : Database Schema

# SQL App

The SQL Database (called SQLApp) is the heart of the System. The tables contain information about a particular ETL Workflow (E.g.: Line of Business Sales Data Feed), associated Tasks and associated packages.

Procedures provide a very rudimentary API access to manipulate data in these tables. Below diagram explains how the data is setup in the tables.

**.  
.  
.**

Package 1 (Versioned)

Package 2 (Versioned)

Package n (Versioned)

**.  
.  
.**

# Enum Definitions

Workflow\_Status:  
S-Started  
R-Running  
N-Not Running  
  
Workflow\_FinishStatus(Log):  
I-Initialized   
E-Executing   
F-Failed   
S-Successful  
A-Aborted

Workflow\_RecoveryMode:  
R-Recover Last Run   
I-Ignore Failure in Last Run

Task\_Status:  
S-Started  
R-Running  
N-Not Running

Task\_FinishStatus(Log):  
I-Initialised  
R-Running  
F-Failed   
S-Successful  
P-Precedent Failed  
A-Aborted

Task\_FailureAction  
A-Abort Execution  
C-Continue Execution  
R-Retry Execution  
I-Ignore

Task\_RecoveryMode:  
R-Recover   
I-Ignore

ExtractLimit\_Type  
Date Range –Range of Period between Start and End Boundaries  
Key Range –Range of Key(Id) between Start and End Boundaries

# How to Use?

## Insert Data into Tables

* When Updating Task Records for a Workflow, make sure you set IsActive=0 for Tasks you no longer want to include in processing.
* Insert Package Version as MajorVersion.MinorVersion.BuildNumber
* When Inserting New Package Record, Make sure you set IsActive=0 for Old Versions

## Understanding Log Tables

* The datetime columns are logged in Regional time zone using GetDate() function.

# Developer Checklist

* Generate New Package-ID after Template is Copied to a New Solution
* Generate New Package ID when Packages are Refactored

# Troubleshooting Tips