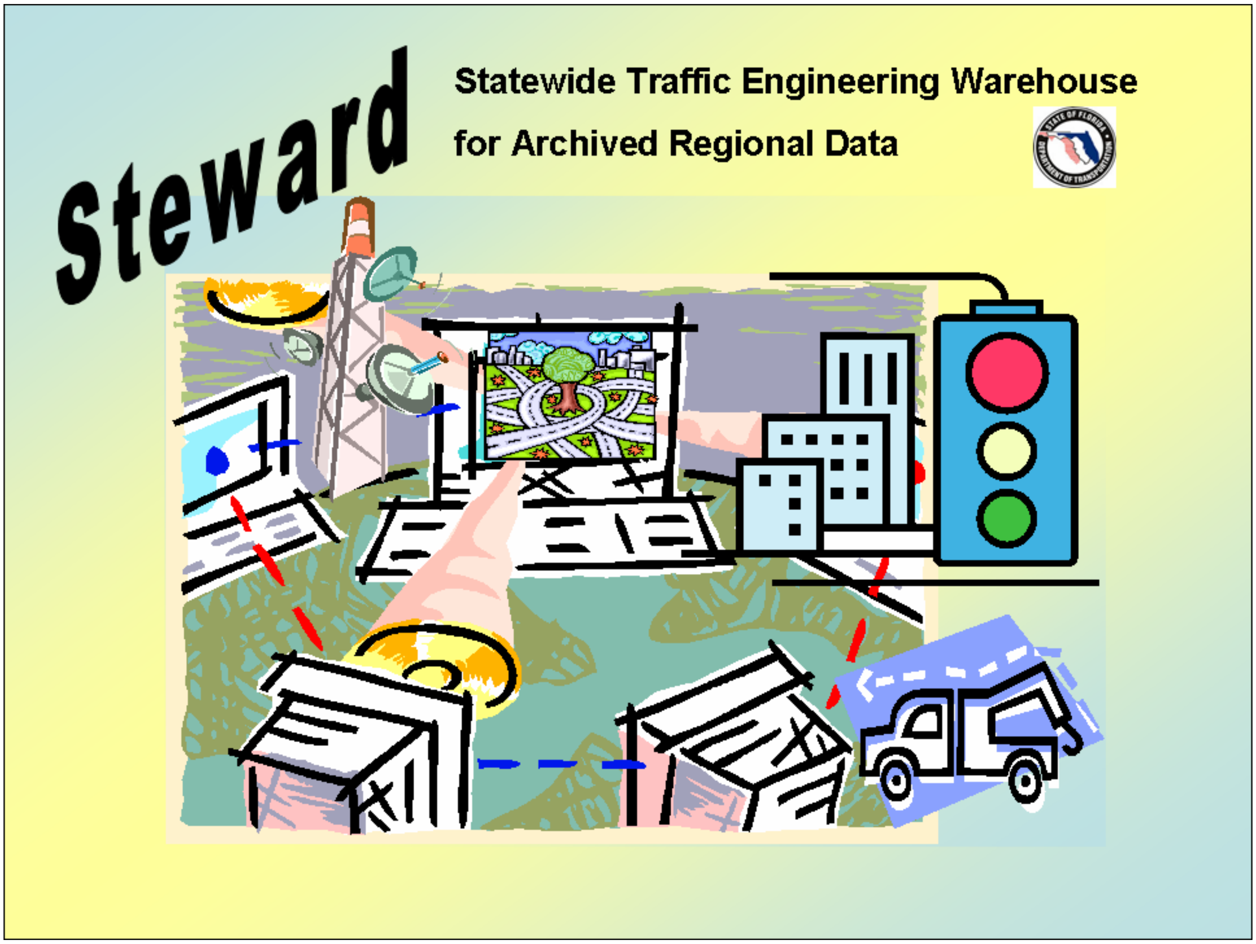


Development of a Central Data Warehouse for Statewide ITS and Transportation Data in Florida

Phase II: Proof of Concept
Final Report

Prepared for the Florida Department of Transportation
By the University of Florida Transportation Research Center

Contract # BD545, RPWO # 37
UF Project 051449

Revision Date: April 15, 2008

Disclaimer

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the State of Florida Department of Transportation.

APPROXIMATE CONVERSIONS FROM SI / METRIC UNITS TO STANDARD / US CUSTOMARY UNITS				
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
LENGTH				
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
AREA				
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
VOLUME				
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
MASS				
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
TEMPERATURE				
^oF	Fahrenheit	(F-32) x 5 / 9 or (F-32) / 1.8	Celsius	^oC
ILLUMINATION				
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
FORCE and PRESSURE or STRESS				
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI / METRIC UNITS TO STANDARD / US CUSTOMARY UNITS (Continued)				
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
LENGTH				
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
AREA				
mm^2	millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
VOLUME				
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
MASS				
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
TEMPERATURE				
^oC	Celsius	1.8C + 32	Fahrenheit	^oF
ILLUMINATION				
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
FORCE and PRESSURE or STRESS				
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Development of a Central Data Warehouse for Statewide ITS and Transportation Data in Florida Phase II: Proof of Concept		5. Report Date April 15, 2008 6. Performing Organization Code
7. Author(s) Kenneth G Courage and Seokjoo Lee		8. Performing Organization Report No. FR-51449
9. Performing Organization Name and Address University of Florida Transportation Research Center PO Box 116580 Gainesville, FL 32611-6580		10. Work Unit No. (TRAIS) 11. Contract or Grant No. BD545, RPWO # 37
12. Sponsoring Agency Name and Address Florida Department of Transportation 605 Suwannee Street, MS 30 Tallahassee, FL 32399		13. Type of Report and Period Covered Final Report 2005-2008 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

This report documents the development and operation of a prototype for the Statewide Transportation Engineering Warehouse for Archived Regional Data (STEWARD) as a proof of concept. It demonstrates that data from traffic management centers around the state can be centrally archived in a practical manner and that a variety of useful reports and other products can be produced. STEWARD will provide TMC managers, District ITS program managers, traffic engineers, and management with several useful functions for periodic reporting and other decision support.

The effort described in this report has concentrated on the SunGuide traffic sensor subsystem (TSS) and the travel time subsystem (TVT). The SunGuide-generated archive data from Interstate 95 in District 2 has been used as a model for this development. Other SunGuide TMCs will be added as their archive data become available. The District 2 system is configured with approximately 450 detectors at 120 RTMS detection stations covering a 25 mile section of the freeway. Other systems in FDOT Districts 4, 5, 6 and 7 are now being configured for STEWARD.

The major topics covered in this report include an overview of STEWARD, a description of the procedure for configuring the SunGuide facility archive, documentation of the traffic sensor subsystem and travel time subsystem databases, reports available from STEWARD, Internet access features and useful applications of the archived data.

17. Key Word	18. Distribution Statement
ITS data archive, Central data warehouse	No Restrictions

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	253	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Acknowledgements

This project was funded by the Florida Department of Transportation Research Office. Technical support was provided by the FDOT Traffic Engineering and Operations Office. The project manager was Liang Y. Hsia, P.E. Administrator, and Deputy State Traffic Engineer.

The SunGuide-generated archive data from Interstate 95 in District 2 has been used as a model for this development. The enthusiastic cooperation of the District 2 staff, especially Peter Vega and Jason Summerfield, contributed greatly to the success of the project.

The ability to generate traffic counts from the archived data is an essential element in the proof of concept for the central data warehouse. This ability was developed with the help and cooperation of Richard Reel and the Transportation Statistics Office.

The authors also wish to acknowledge the contributions to this project from a group of graduate students who were key members of the research team. Special thanks are due to Jongwon Shim, Incheol Shin and Young Hwan Namkoong.

Preface

The potential benefits of maintaining an archive of data produced by traffic management systems are well recognized. This report explores those benefits and demonstrates that they are achievable.

With that in mind, it is important to note that the prototype developed by this project represents a proof of concept and not a fully operational system. The research team hopes that support for statewide implementation of a fully configured central data warehouse will not be too far in the future.

Executive Summary

The potential benefits of maintaining an archive of data produced by traffic management systems are well recognized. The Florida Statewide Intelligent Transportation System Architecture contains an archived data management subsystem that serves as a central data warehouse. The Statewide Transportation Engineering Warehouse for Archived Regional Data (STEWARD) will be the implementation of that feature.

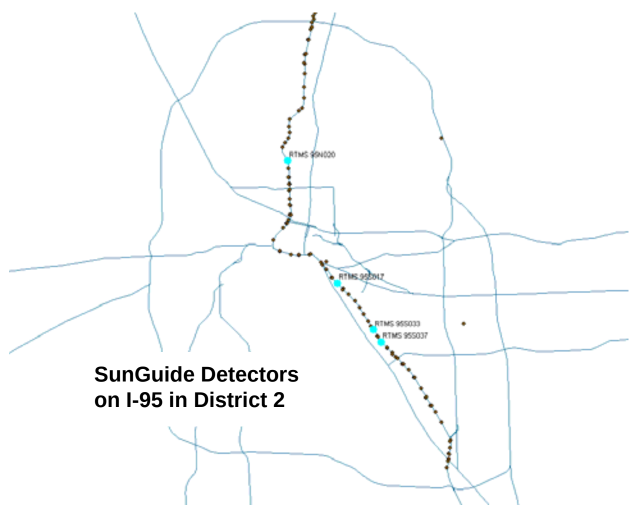
This report documents the development and operation of a prototype for STEWARD as a proof of concept. It demonstrates that data from regional traffic management centers around the state can be centrally archived in a practical manner and that a variety of useful reports and other products can be produced. STEWARD will provide TMC managers, District ITS program managers, traffic engineers, and management with the following useful functions:

- Identify detector malfunctions
- Provide calibration guidance for detectors
- Perform quality assessment data reliability tests on data
- Provide daily performance measures for system, and statewide performance measures
- Facilitate periodic reporting requirements
- Provide data for research and special studies

The effort described in this report has concentrated on the SunGuide traffic sensor subsystem (TSS) and the travel time subsystem (TVT). The SunGuide-generated archive data from Interstate 95 in District 2 has been used as a model for this development. Other SunGuide TMCs will be added as the archive data become available. The District 2 system is configured with approximately 450 detectors at 120 RTMS detection stations covering a 25 mile section of the freeway. Other systems in Districts 4, 5, 6 and 7 are now being configured for STEWARD.

Configuring the STEWARD Facility Data

The archived data are stored in comma-delimited flat files, with each file representing a 24 hour day. Assimilation of the data into the STEWARD database involves the addition of several data items that relate the individual detectors to the system in which they are located. In addition, the STEWARD database includes the results of some analyses that aggregate certain measures over periods ranging from 15 to 60 minutes.


SunGuide Detectors on I-95 in District 2

As a part of the proof of concept, STEWARD can supply data to support FDOT efforts to develop a framework for reporting travel time reliability on arterials. In addition to supporting the specific study, it is anticipated that STEWARD will continue to supply travel time reliability data for long term monitoring and reporting.

The STEWARD TSS Database

The STEWARD database contains the TSS data representing all days at all stations from all TMCs. The SunGuide raw data archive contains one record for each lane for each 20 second polling interval. To streamline the processing and retrieval, the STEWARD data base contains one record per station representing all of the lanes at that station. The values contained in the STEWARD database represent the 20 second archived data values accumulated over periods of 5, 15 and 60 minutes.

TSS Report Generation

Reports are generated from the TSS data at the detector, station and system levels. Station level reports require additional facility information indicating how the lane detectors are grouped into stations. System level reports require information on the location (milepost) of each station. All of this information is contained in the facility configuration databases that must be developed for each TMC. The following reports are produced at the detector level.

1. *Maximum Flow Rates:* A very high flow rate (e.g., greater than 2400 vph in any lane) could be an indication of a detector calibration problem.
2. *Effective Vehicle Lengths:* The effective vehicle length is defined as the combined length of the vehicle plus the length of the detection zone. It may be calculated from the volume, speed and occupancy values for each time interval. The consistency of effective vehicle length provides another quality assessment indicator.
3. *Lane Volume Balance:* The lane volume balance is expressed as the ratio of the highest to lowest lane volume at each station. If all lane volumes at a given station were identical, then the lane balance value would be 1.0. During periods of moderately heavy flow, lane balance values above 1.5 might indicate detection problems unless a reasonable explanation can be found.

The following reports are produced at the system level:

1. *Volume map:* A table is produced with the stations arranged in geographical order showing all freeway and ramp volumes for the period.
2. *Input/Output balance:* Another table is produced showing the total volume entering and leaving each link in the system. Over reasonable time periods, an unbalance between inputs and outputs would suggest volume counting errors unless (as is the case in District 2) there are entrance or exit ramps without detectors.

In addition to the volume analysis tables, several operational performance measures are generated, including:

1. *Vehicle Miles of Travel (VMT):* This is a measure of productivity of the freeway, typically accrued over a peak period or longer.
2. *Vehicle Hours of Travel Time (VHTT):* This is the accumulated travel time of all vehicles in the system over the analysis period.
3. *Average speed:* A figure representing the average speed of all vehicles in the system is computed by dividing the VMT by the VHTT.
4. *Delay:* There are several definitions of delay, each with its own method of computation. For a freeway system, the most appropriate delay measure is obtained by subtracting the VHTT that would have accrued at some desired speed from the measured VHTT. The result is expressed in vehicle hours of delay.

5. *Kinetic Energy:* Kinetic energy is proportional to the product of speed and volume. Higher values of kinetic energy are obtained when heavy volumes are carried at high speeds. For this reason, kinetic energy has been suggested as the "bottom line" performance indicator for a freeway facility. It has also been suggested that high values of kinetic energy could be associated with safety hazards. This measure is produced to support future research.
6. *Level of Service (LOS):* By the Highway Capacity Manual definition, the LOS on a freeway is expressed in terms of the density of traffic. Traffic density may be estimated by dividing volume by speed. So, with the available data, we can estimate the LOS on each freeway segment during each period. LOS reports are generated for each segment indicating the proportion of time that the LOS exceeded each of the letter grades.

The STEWARD TVT Database

District 2 has established approximately 30 travel time links. The one minute travel times for each link are accumulated and grouped into 5, 15 and 60 minute intervals. Other districts are configuring their systems now

In addition to the estimated travel times, three performance measures derived from the travel times are contained in the database:

- *Congestion Delay:* based on a travel time index of 1.5. The travel time index is defined as the ratio of the actual travel time to the travel time at the free flow speed. The unit of measurement is "accumulated minutes of delay."
- *On Time Delay:* referenced to a travel speed of 10 mph below the speed limit. This threshold has been specified for purposes of travel time reliability reporting in Florida. The unit of measurement is also "accumulated minutes of delay."
- *Percent of on-time trips*: defined as the percent of trips that were made at a speed no less than ten mph below the speed limit.

Traffic Count Generation

Given accurate traffic counts from SunGuide, the STEWARD data could provide a very useful supplement to the FDOT traffic counting programs. With this in mind, a feature has been built into STEWARD to provide count files in the same format as the count stations. This capability was tested on selected stations in the District 2 system and eventually will be expanded to other districts. Comparison of the counts provided by the Statistics Office with the SunGuide archived counts from nearby stations has indicated excellent agreement in District 2.

The ability to provide the Central Office with supplemental traffic counts should be a useful STEWARD feature. However, a potentially more useful feature would be created by the ability to provide the district planning offices with supplemental counts for their portable stations. Because the district stations have no permanent counters, obtaining the required periodic counts is much more labor intensive.

The capability to generate traffic counts in the formats now used by the FDOT central and district offices has been implemented in STEWARD. Sample traffic counts have been provided to the both offices and are now under review.

Internet Access for STEWARD

The capability to access the STEWARD TSS and TVT databases is now under development and a prototype web site has been created. Users who access STEWARD via the internet will be able to query the database by location and time and generate a variety of reports.

Future Development

STEWARD will continue to operate with support from the UF Transportation Research Center and FDOT. Additional SunGuide TMCs will be brought on board as their archive data subsystems become active. Data from other sources, such as the FDOT RCI, crash records and Statistics Office traffic counts will be examined to investigate correlations with the SunGuide data.

FDOT has initiated a continuation project with the following objectives:

- Integrate the CDW functions with other FDOT data management programs.
- Transfer the CDW operations to the TERL center in Tallahassee
- Automate and fine tune the transfer of daily archive data from district TMCs and the statewide monthly, quarterly, and annual reporting processes.
- Expand the CDW to include other data sources.

At some point in the future, FDOT hopes to incorporate the STEWARD functionality into future SunGuide versions.

Table of Contents

List of Acronyms	
Name	**Meaning**
ADMS	archived data management subsystem
ASP	active server page
ATRI	American Transportation Research Institute
CARS	Crash Analysis Reporting System
CCTV	Closed-circuit television
CDW	central data warehouse
CMS	Center for Multimodal Solutions for Congestion Mitigation
CSS	cascading style sheets
CSV	comma separated values
ETL	extraction transformation and loading
FDOT	Florida Department of Transportation
FHWA	Federal Highway Administration
FTP	File Transfer Protocol
HOV	high-occupancy vehicle
HTTP	Hypertext Transfer Protocol
ID	identification
IIS	Internet Information Services
ITS	Intelligent transportation system(s)
LOS	level of service
MM	mobility monitoring
NCHRP	National Cooperative Highway Research Program
NITSA	Florida Statewide Intelligent Transportation System Architecture
OWB	Oracle Warehouse Builder
QA	quality assurance
RCI	roadway characteristics inventory
RTMC	regional traffic management center
RTMS	Road Traffic Microwave Sensor
SIS	Strategic Intermodal System
SITSA	Florida Statewide Intelligent Transportation System Architecture
SQL	Structured Query Language
SR	State Road
STEWARD	Statewide Transportation Engineering Warehouse for Archived Regional Data
TCP	Transmission Control Protocol
TERL	Traffic Engineering Research Laboratory
TMC	traffic management center
TSS	traffic sensor subsystem (of SunGuide)
TTI	Texas Transportation Institute
TVT	travel time subsystem (of SunGuide)
UF	University of Florida
VHTT	vehicle hours of travel time
VMT	vehicle miles of travel

List of Figures Page

List of Tables

1 Introduction

The Florida Statewide Intelligent Transportation System Architecture (SITSA), which is based on the National Intelligent Transportation System Architecture (NITSA), contains an archived data management subsystem (ADMS) that serves as a central data warehouse. The Statewide Traffic Engineering Warehouse for Regionally Archived Data (STEWARD) will be the implementation of that ADMS. STEWARD will archive data from a variety of sources in a database that will support the generation of reports and queries. A prototype has been developed as a proof of concept for a fully operational version to be deployed in the future. The work was done by the University of Florida (UF) Transportation Research Center with support from the Florida Department of Transportation (FDOT) Research Office. This document constitutes the final report for the project.

1.1 Current Status of STEWARD

A functioning version of STEWARD now exists on the UF campus in Gainesville. The database schema, the extraction transformation and loading (ETL) functions and the report generation functions are all covered in later sections of this report and in the appendices.

The system was developed with support from District 2 and currently contains several months of volume, speed and travel time data provided by District 2 on a daily basis. Efforts have been initiated to bring the data from Districts 4, 5 and 6 into the system using the procedures developed in District 2 as a model. These efforts are also described in this report.

1.2 Continuing Efforts for STEWARD Operation and Expansion

It is our understanding that FDOT plans to expand the prototype central data warehouse (CDW) into a fully operational system in the future. With that in mind, the operation and expansion of STEWARD will be continued with support from FDOT and UF. The expansion to cover additional FDOT districts that was initiated under this project will be continued under separate projects described later in this report. The functionality of the system will also be expanded and the system will be transferred to the FDOT Traffic Engineering Research Laboratory (TERL)

1.3 Organization of this Report

This report documents the accomplishments in the development and operation of STEWARD. It describes the STEWARD design and operational features and demonstrates that data from traffic management centers (RTMCs) around the state can be centrally archived in a practical manner and that a variety of useful reports and other products can be produced.

The main topics addressed in the body of the report include

- An overview of STEWARD,
- The current status of each district with respect to their SunGuide traffic management center (TMC) data archive configuration,
- Useful applications for the archived data,
- Data quality assurance (QA) considerations and
- Recommendations for continued operation.

A substantial amount of detailed material on the design and operation of STEWARD is presented in this report. To facilitate review and assimilation of this material by those who are interested in the details, most of the material has been incorporated into technical appendices which can function as "stand alone" documents. The following appendices are included as a part of this report:

- Appendix A: SunGuide Configuration Instructions
- Appendix B: STEWARD Installation Instructions
- Appendix C: STEWARD Operating Instructions
- Appendix D: Internet Access Documentation
- Appendix E: STEWARD Design Document
- Appendix F: Description of delivered software and data files
- Appendix G: Supplemental Material Delivered on CD

2 STEWARD Overview

STEWARD will provide TMC managers, District intelligent transportation system (ITS) program managers, traffic engineers, and management with the following useful functions:

- Identify detector malfunctions
- Provide calibration guidance for detectors
- Perform quality assessment data reliability tests on data
- Provide daily performance measures for system, and statewide performance measures
- Facilitate periodic reporting requirements
- Provide data for research and special studies

2.1 Configuring the STEWARD Facility Data

The overall data flow for STEWARD is shown in Figure 1. The archived data are stored in comma-delimited flat file, with each file representing a 24 hour day. Zipped versions of these files are posted periodically by the district staff. Assimilation of the data into the STEWARD database involves the addition of several data items that relate the individual detectors to the system in which they are located. In addition, the STEWARD database includes the results of some analyses that aggregate certain measures over the following periods:

- 5 minutes, for correlation with incidents and other short term phenomena
- 15 minutes, for compatibility with design and traffic engineering requirements
- 60 minutes, for compatibility with the Statistics Office traffic counts

The following information must be collected and maintained in a database for the traffic sensor subsystem (TSS) archive data:

- Station identification (ID) and Lane ID: Unique identifier assigned by STEWARD
- Description: A physical description of the location
- Road: The name given to the facility
- Coordinates: Latitude and longitude
- Milepost: Required for sequential ordering of stations
- Lane Type (freeway, entrance or exit ramp)
- Direction
- Detector type: Loop, Road Traffic Microwave Sensor (RTMS), etc.

- Maximum speed: Normally the speed limit
- Count station: The number assigned by the FDOT Statistics Office or District Planning Office for generating traffic count data files from the SunGuide detectors.

As a part of the proof of concept, STEWARD will supply data to support FDOT efforts to develop a framework for reporting travel time reliability on arterials. In addition to supporting the specific study, it is anticipated that STEWARD will continue to supply travel time reliability data for long term monitoring and reporting.

The current threshold of delay for purposes of travel time reliability analysis is established at 10 mph under the speed limit. Reporting of delay under this definition requires knowledge of the speed limit for each travel time link. This requirement is complicated by the fact that multiple speed limits may apply over a given travel time link.

The following information must be collected and maintained in a database for each travel time link in the travel time (TVT) archive data:

- A description of the origin and destination
- Direction
- Distance weighted average speed limit (required to compute the travel time variability measures)
- Link Length

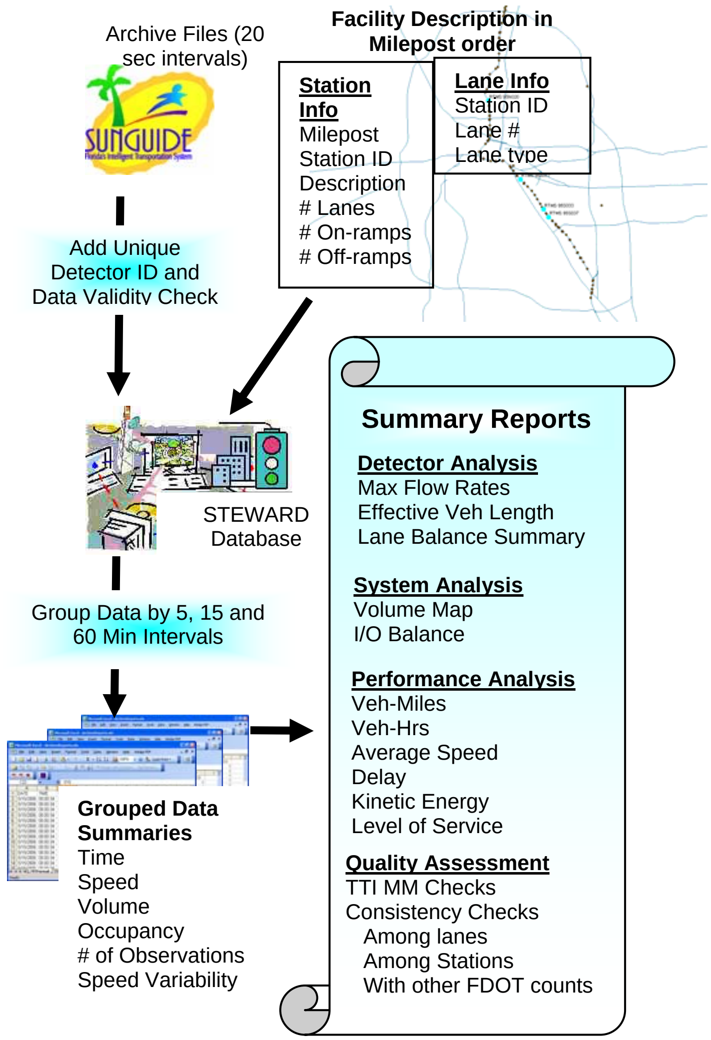


Figure 1: SunGuide Archived TSS Data Flow

2.2 The STEWARD TSS Database

The STEWARD database contains the TSS data from all days at all stations from all TMCs. The SunGuide raw data archive contains one record for each lane for each 20 second polling interval. The STEWARD database content differs from the raw data as follows:

- To streamline the processing and retrieval, the STEWARD data base contains one record per station representing all of the lanes at that station. Individual lane data are archived and lane balance measures based on the ratio of the highest lane value to the lowest lane value are recorded for both volume and speed.
- The values contained in the STEWARD database represent the 20 second archived data values accumulated over periods of 5, 15 and 60 minutes.
- A quality assessment process is applied to identify bad or suspicious data.

2.3 TSS Report Generation

Reports are generated from the TSS database at the detector, station and system levels. Station level reports require additional facility information indicating how the lane detectors are grouped into stations. System level reports require information on the location (milepost) of each station. All of this information is contained in the facility databases that must be developed for each TMC.

The following reports are produced at the lane level.

- *Maximum Flow Rates:* A very high flow rate (e.g., greater than 2400 vph in any lane) could be an indication of a detector calibration problem.
- *Effective Vehicle Lengths:* The effective vehicle length is defined as the combined length of the vehicle plus the length of the detection zone. It may be calculated from the volume, speed and occupancy values for each time interval. The consistency of effective vehicle length provides another quality assessment indicator.
- *Lane Volume Balance:* The lane volume balance is expressed as the ratio of the highest to lowest lane volume at each station. If all lane volumes at a given station were identical, then the lane balance value would be 1.0. During periods of moderately heavy flow, lane balance values above 1.5 might indicate detection problems unless a reasonable explanation can be found.

The following reports are produced at the station level:

- *Volume map:* A table is produced with the stations arranged in geographical order showing all freeway and ramp volumes for the period.
- *Input/Output balance:* Another table is produced showing the total volume entering and leaving each link in the system. Over reasonable time periods, an unbalance between inputs and outputs would suggest volume counting errors unless (as is the case in District 2) there are entrance or exit ramps without detectors.

In addition to the volume analysis tables, several operational performance measures are generated, including:

- *Vehicle Miles of Travel (VMT):* This is a measure of productivity of the freeway, typically accrued over a peak period or longer.
- *Vehicle Hours of Travel Time (VHTT):* This is the accumulated travel time of all vehicles in the system over the analysis period.
- *Average speed:* A figure representing the average speed of all vehicles in the system is computed by dividing the VMT by the VHTT.
- *Delay:* There are several definitions of delay, each with its own method of computation. For a freeway system, the most appropriate delay measure is obtained by subtracting the VHTT that would have accrued at some desired speed from the measured VHTT. The result is expressed in vehicle hours of delay.
- *Kinetic Energy:* Kinetic energy is proportional to the product of speed and volume. Higher values of kinetic energy are obtained when heavy volumes are carried at high speeds. For this reason, kinetic energy has been suggested as the "bottom line" performance indicator for a freeway facility. It has also been suggested that high values of kinetic energy could be associated with safety hazards. This measure is produced to support future research.
- *Level of Service (LOS):* By the Highway Capacity Manual definition, the LOS on a freeway is expressed in terms of the density of traffic. Traffic density may be estimated by dividing volume by speed. So, with the available data, we can estimate the LOS on each freeway segment during each period. LOS reports are generated for each segment indicating the proportion of time that the LOS exceeded each of the letter grades.

2.4 The STEWARD TVT Database

District 2 has established approximately 30 travel time links. The one-minute travel times for each link are accumulated and grouped into 5, 15 and 60 minute intervals. Other districts are configuring their systems now

In addition to the estimated travel times, three performance measures derived from the travel times are contained in the database:

- *Congestion Delay:* based on a travel time index of 1.5. The travel time index is defined as the ratio of the actual travel time to the travel time at the free flow speed. The speed limit is used to represent the free flow speed. The unit of measurement is "accumulated minutes of delay."
- *On Time Delay:* referenced to a travel speed of 10 mph below the speed limit. This threshold has been specified for purposes of travel time reliability reporting in Florida. The unit of measurement is also "accumulated minutes of delay."
- *Percent of on-time trips*: defined as the percent of trips that were made at a speed no less than ten mph below the speed limit.

2.5 TVT Report Generation

The TVT summary report includes average travel time and delay as well as maximum travel times and delays, along with their time of occurrence. A recently added field reports the percentage of on-time trips. This information is presented for each link and time interval. A summary report may be requested for a specific link, for all origins to a specific destination or for all destinations from a specific origin.

2.6 Traffic Count Generation

Given accurate traffic counts from SunGuide, the STEWARD data could provide a very useful supplement to the FDOT permanent traffic counts. With this in mind, a feature has been built into STEWARD to provide count files in the same format as the count stations. This capability has been tested on selected stations in the District 2 system and eventually expanded to other districts. The results of the District 2 comparisons are reported later in this document.

The ability to provide the Central Office with supplemental traffic counts should be a useful STEWARD feature. However, a potentially more useful feature would be created by the ability to provide the district planning offices with supplemental counts for their portable stations. Because the district stations have no permanent counters, obtaining the required periodic counts is much more labor intensive.

The capability to generate traffic counts in the formats now used by the FDOT central and district offices has been implemented in STEWARD. Sample traffic counts have been provided to the Statistics Office and the District 2 Planning Office and are now under review.

2.7 Internet Access for STEWARD

The capability to access the STEWARD TSS and TVT databases is now under development and a prototype web site has been created. Users who access STEWARD via the internet will be able to query the database by location and time and generate a variety of reports. The internet access features are described in Appendix D to this report.

3 Current District TMC Configuration Status

This section summarizes the current status of the TSS and TVT data collection configuration in each of the districts. Appendix A provides detailed step by step instructions for configuring the facility data to allow the daily archive files to be transformed and loaded into the STEWARD database.

Table 1 summarizes the current status of the steps described in Appendix A for all districts. Note that minimal progress has been made in the past few months in this area. A stable system of detectors is required to support the archive function. The detector systems in most districts have been undergoing development and did not reach the level of stability required to support a useful archive within the life of the project. Other district priorities related to system implementation, especially SunGuide enhancement, have precluded the advancement of the data archive features of SunGuide. It appears that the districts are ready for the next step now.

Table 1: STEWARD SunGuide Archive Configuration Status: 9-18-07

	District 2	**District 4**	**District 5**	**District 6**	**District 7**
Contact:	Pete Vega Jason Summerfield	Dong Chen	Nathan Ruckert	Manuel Fontan Julio Orozco	Bill Wilshire James Bitting
Note: Step numbers below refer to the Configuration Instructions					
TSS Data					
Sample data received	Done	Done	Received 7-13	Received 7-19	Contact pending
1. Stations extracted	Done	Done	Done 7-15	Preliminary: Some issues need to be resolved with the district	
2. Subsystems configured	Done	Done	Done 7-20		
3. Station ID numbers assigned	Done	Done	Done 7-20		
4. Station position assigned	Done	Preliminary, needs district review	Done 7-20		
5. Station status assigned	Done	Preliminary, needs district review	Done 7-20		
6. Lane ID numbers assigned	Done	Preliminary, needs district review	Preliminary, needs district review		
7. Operating parameters assigned	Done				
Configuration reviewed by district	Done				
Archive data tests in progress	Done				
Archive process is operational	Done				
TVT Data					
Sample data received	Done				
1. Links extracted	Done				
2. Direction and status added	Done				
3 Origins and destinations created	Done				
4. Location descriptions added	Done				
5. Position information added	Done				
6. Lengths and speed limits computed	Done				
Configuration reviewed by district	Done				
Archive data tests in progress	Done				
Archive process is operational	Done				

3.1 District 2

The procedures for data transfer from the SunGuide TSS and TVT subsystem archives to the STEWARD database were developed using the I-95 SunGuide system in Jacksonville as a model. The District 2 ITS personnel made significant contributions to this effort. The system is fully configured and new data records are being added to the database as they are received. A TSS detector map for District 2 is presented in Figure 2a. A TVT link map is presented in Figure 2b.

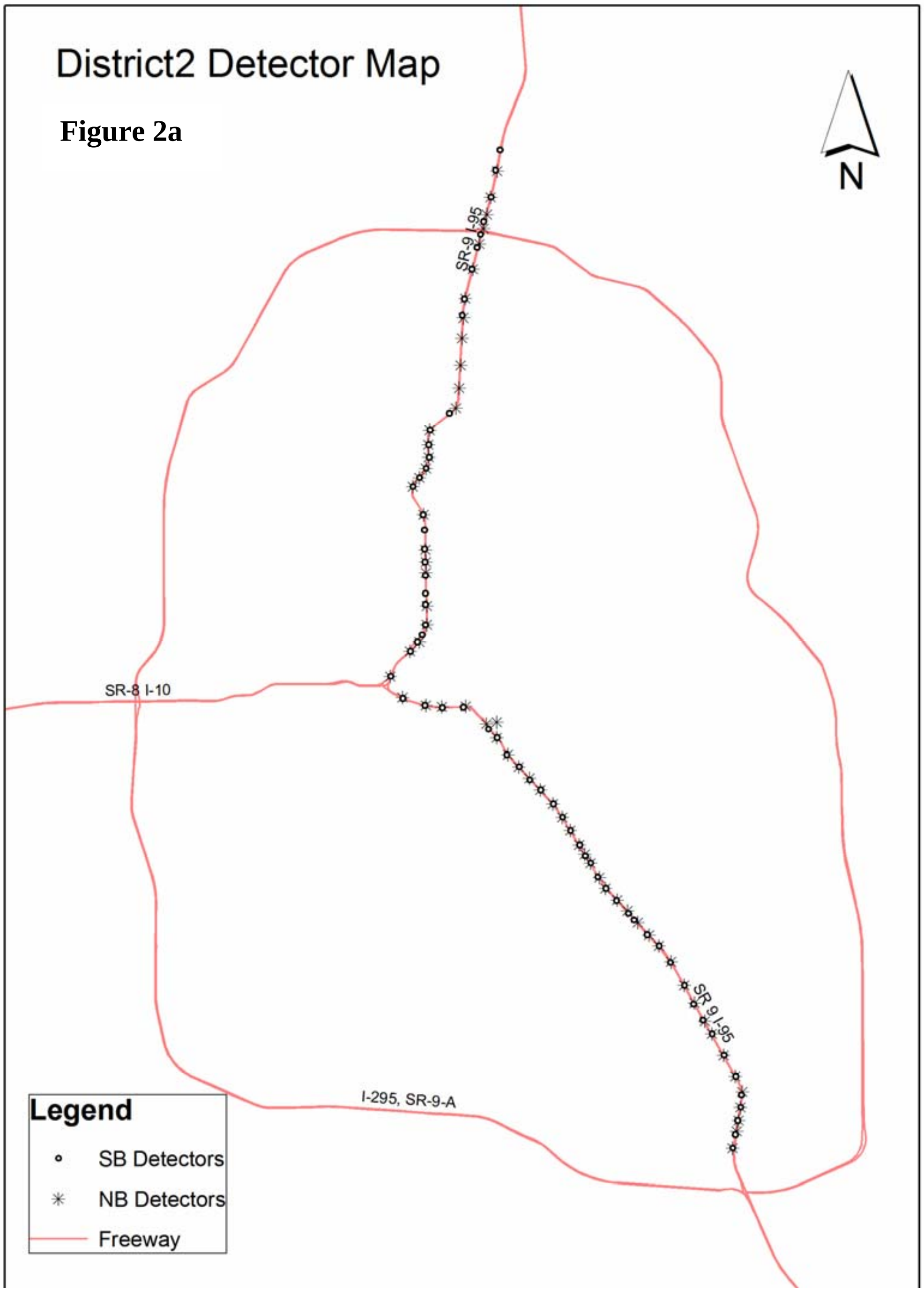
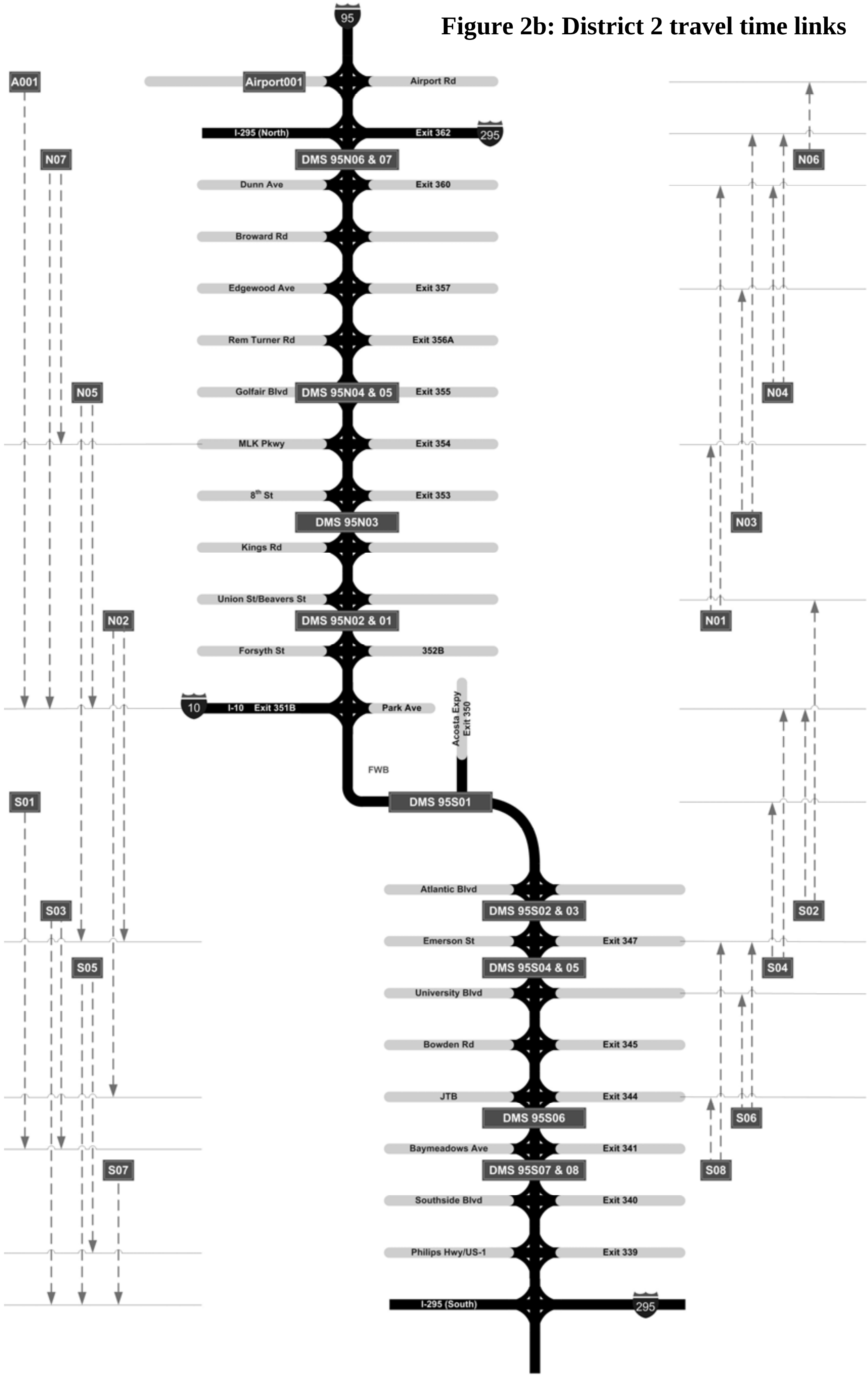


Figure 2b: District 2 travel time links



3.2 District 4

A set of SunGuide TSS archive files was provided by District 4 to begin the task of bringing their SunGuide archives into STEWARD. This section describes the data that we have received and our progress in the conversion to STEWARD.

The archive files contain only a code to identify the station and lane numbers. Additional static information is needed to determine where the station is, what type of lane is represented, etc. This “data about the data”, i.e., metadata, must be compiled from locally available information to support STEWARD’s analysis and reporting capability.

Fortunately, some of the information is embedded in the District 4 naming conventions for each station and lane, including:

- The facility (I-95 or I-595)
- The direction of traffic in the lane
- The milepost for the station

We have made the best possible use of this information in constructing the facility metadata files.

SunGuide archive TSS files were received from FDOT District 4 for twelve consecutive days between 4-1-07 and 4-12-07. The facility metadata files were partially constructed using the information in these files. Additional facility description information will be needed from the district. A detector map showing all detector locations reporting archive data is presented in Figure 3.

Separate facility description databases are required for the station data and the lane data. The current status of the station information is described in Table 2. The current status of the station information is described in Table 3. Each table describes the required data fields, the current development status and any issues that arose in extracting the information from the TSS archive files.

An Excel file containing all of the facility metadata developed from the archive files has been forwarded separately to District 4. Some further discussion with the district will be required to complete the metadata files.

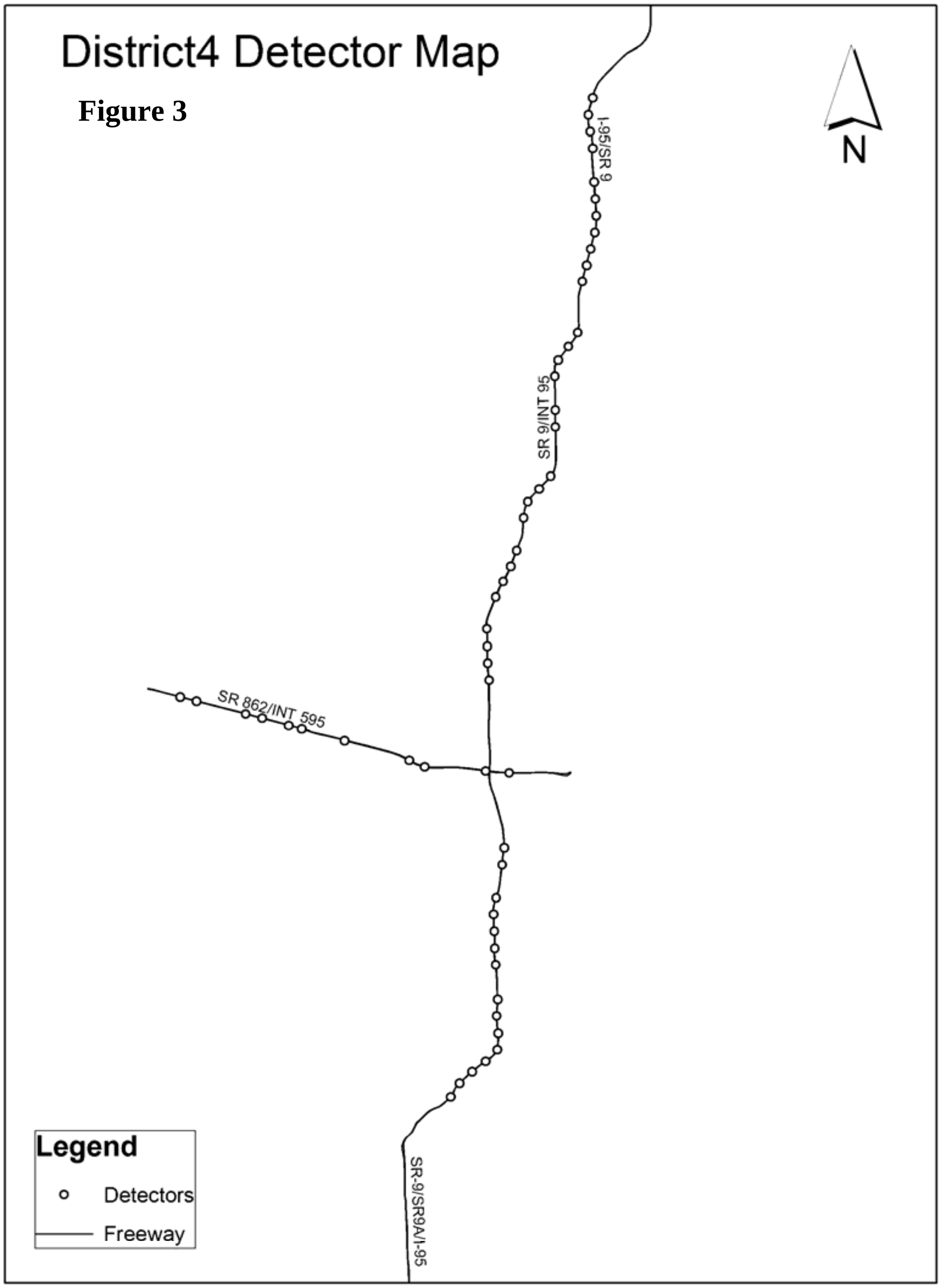

District4 Detector Map
Figure 3
N
I-95/SR 9
SR 9/INT 95
SR 862/INT 595
SR-9/SR9A/I-95
Legend
Detectors
Freeway

Table 2: Status and Issues for the Required Station Description Data for District 4		
Required Station Data	**Description**	**Status and Issues**
STATION_INDEX	An index number assigned sequentially for internal computations	
STATIONCDW_NUM	A statewide-unique station identifier in the format DFnn, where: • D is the district number • F is the facility number within the district • nn is the sequence number of the station within the facility	This was constructed from the "detector_id" field in the TSS archive. The assigned facility numbers are: • 1 = I-595 • 2 = I-95 The sequence numbers were assigned automatically
Description	A narrative description of the station location	Constructed from the information embedded in the TSS archive detector_id and lane_id fields: Example "I-595 WB at Milepost 011.0" Question: Is there additional location information such as the cross street available?
Status	A code to indicate if the station is on line	All stations were assigned as on line (Code 0)
Road	Facility description	Either "I-95" or "I-595" was assigned based on the information embedded in the detector_ID field of the TSS archive data
LATITUDE LONGITUDE	Station coordinates	Are these available? If not, we can probably compute them from the milepost info
STATE_MILEPOST		We assumed that the last number embedded in the detector-id field was the state milepost. Is this correct?
ROADWAY_ID	From roadway characteristics inventory (RCI)	We can find this
ROADWAY_MILEPOST	County milepost	We can find this
MAX_SPEED	Need for Delay at Performance Measure	We can find this
NUM_OF_LANES	Need for V/C ratio at Performance Measure	We can find this
LANE_CAPACITY	Need for V/C ratio at Performance Measure	This needs to be estimated as vphpl
UPNODE	Need for average volume at Performance Measure	We can find this
DETECTOR_TYPE		Is this available?
DETECTOR_UNIT	This was used in D2 to indicate the number of detector units at the station	Not used for any data processing purposes

Table 3: Status and Issues for the Required Lane Description Data for District 4		
Required Lane Data	**Description**	**Status and Issues**
TMC_ID	The station ID as it appears in the “lane_id” field of the TSS archive Example: 95SB044.5-Lanes SB-lane1	All lanes that were contained in the 12 TSS data files are now represented in the facility data.
CDWSTATION	Same as the STATIONCDW_NUM field in the station data description	All lane numbers have been assigned
LANE	A statewide-unique lane identifier in the format SSSSDTn, where: • SSSS is the station number described above • S is direction (1=increasing mileposts, 2=decreasing mileposts) • T is a code for the type of lane • n is the lane number from the “lane_id” field	This information was extracted from the lane_id field of the TSS archive data. Since the lane types are not known, all lanes were assigned type 3, which indicates a general use freeway lane Question: Are there any ramps or HOV lanes? This information was not embedded in the lane_id field.
DET_TYPE	The type of detector in this lane	All detectors were identified as “Wave” for Wavetronics. Are there other detector types?
DIRECTION	1=increasing mileposts 2=decreasing mileposts	Obtained from the lane_ID field in the TSS archive data
STATUS	A code to indicate if the detector is on line	All lanes were assigned as on line (Code 0)
ROADWAY_ID	From the RCI database	Info still needed
ROADWAY_MILEPOST	The county milepost	Info still needed
MAX_SPEED	The speed limit at the station	Info still needed
TSS_SPEED	The upper limit of speed for travel time computation	Info still needed
OCCUPANCY	We are not sure what this is, but it was in the facility data obtained from District 2	
COUNT_STATION	The district or central office traffic count station number to be assigned if traffic count files are desired from this lane	This will require coordination with the district planning office and the central statistics office.

3.3 District 5

A set of SunGuide TSS archive files was provided by District 5 to begin the task of bringing their SunGuide archives into STEWARD. This section describes the data that we have received and our progress in the conversion to STEWARD.

SunGuide archive TSS files were received from FDOT District 5 for several days in June and July. The facility metadata files were partially constructed using the information in these files and in the detector inventory spreadsheet. Some additional facility description information will be needed from the district. A detector map showing all detector locations reporting archive data is presented in Figure 4.

Separate facility description databases are required for the station data and the lane data. The current status of the station information is described in Table 4. The current status of the lane information is described in Table 5. Each table describes the required data fields, the current development status and any issues that arose in extracting the information from the TSS archive files.

An Excel file containing all of the facility metadata developed from the archive files has been forwarded separately to District 5. Some further discussion with the district will be required to complete the metadata files.

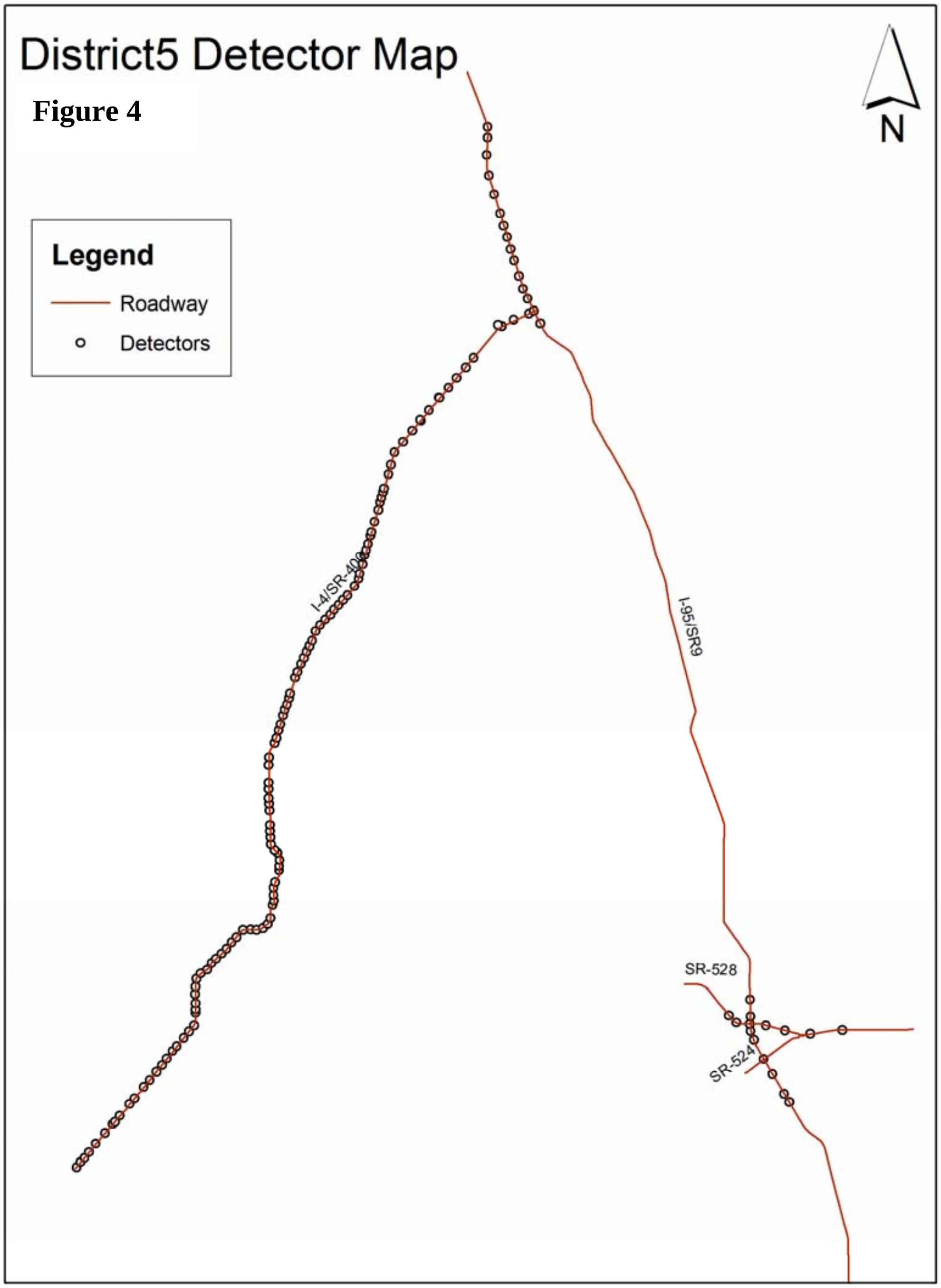

District5 Detector Map
Figure 4
N
Legend
Roadway
Detectors
I-4/SR-400
I-95/SR9
SR-528
SR-524

Table 4: Status and Issues for the Required Station Description Data for District 5		
Required Station Data	**Description**	**Status and Issues**
STATION INDEX	An index number assigned sequentially for internal computations	Done
STATION ID	A statewide-unique station identifier in the format DFnn, where: • D is the district number • F is the facility number within the district • nn is the sequence number of the station within the facility	Facility numbers were assigned from the “Road” field below Station numbers were assigned sequentially within each facility
Description	A narrative description of the station location	Assigned from station ID names in the archive data
Status	A code to indicate if the station is on line	All stations that appeared in the archive files were assigned as on line (Code 0) All stations that appeared in the detector inventory but not in the archive were assigned as off-line (Code 1)
Road	This is the facility number (0-9) Detectors assigned to #9 included a group in the area of Sr 520 and 528 which did not appear to be on a cohesive route. Detectors that appeared in the archive files but were not in the inventory were also assigned to #9, pending resolution of their actual location.	Facility numbers were assigned as follows: • 0 = I-4 W of Turnpike • 1 = I-4 Turnpike to 417 • 2 = I-4 E of 417 to I-95 • 3 = I-95 Brevard • 4 = I-95 Volusia • 9 = Unassigned detectors
LATITUDE LONGITUDE	Station coordinates	Obtained directly from the inventory spreadsheet.
STATE MILEPOST		Obtained directly from the inventory spreadsheet.
ROADWAY_ID	From RCI	We can find this
ROADWAY MILEPOST	County milepost	We can find this
MAX_SPEED	Need for Delay at Performance Measure	We can find this
NUM_OF_LANES	Need for V/C ratio at Performance Measure	We can find this
LANE_CAPACITY	Need for V/C ratio at Performance Measure	This needs to be estimated as vphpl
UPNODE	Need for average volume at Performance Measure	We can find this
DETECTOR_TYPE		Is this available?
DETECTOR_UNIT	This was used in D2 to indicate the number of detector units at the station	Not used for any data processing purposes

Table 5: Status and Issues for the Required Lane Description Data for District 5		
Required Lane Data	**Description**	**Status and Issues**
TMC ID	The station ID as it appears in the "lane_id" field of the TSS archive Example: I-4 @ L Mary Ramp-WB-lane1	All lanes that were contained in the TSS data files are now represented in the facility data.
CDWSTATION	Same as the STATIONCDW_NUM field in the station data description	All lane numbers have been assigned
LANE	A statewide-unique lane identifier in the format SSSSDTn, where: • SSSS is the station number described above • D is direction (1=increasing mileposts, 2=decreasing mileposts) • T is a code for the type of lane • n is the lane number from the "lane_id" field	Directions were determined from the station names as follows: • link1 = NB or EB • link2 = SB or WB When the direction (e,g,. "EB" was embedded in the name, it was used directly. All lanes were assigned as freeway lanes except when "ramp" was embedded in the name. All ramps were assigned as entrance ramps pending resolution of their function
DET_TYPE	The type of detector in this lane	Obtained from the inventory where available
DIRECTION	1=increasing mileposts 2=decreasing mileposts	Determined as shown above for the lane direction
STATUS	A code to indicate if the detector is on line	All lanes were assigned the same status code as the station in which they are located
ROADWAY_ID	From the RCI database	Info still needed. We will try to get it from the RCI
ROADWAY_MILE POST	The county milepost	Info still needed. We will try to get it from the RCI
MAX_SPEED	The speed limit at the station	Info still needed. We will try to get it from the RCI
TSS_SPEED	The upper limit of speed for travel time computation	Info still needed
OCCUPANCY	We are not sure what this is, but it was in the facility data obtained from District 2	
COUNT_STATION	The district or central office traffic count station number to be assigned if traffic count files are desired from this lane	This will require coordination with the district planning office and the central statistics office.

3.4 District 6

A set of SunGuide TSS archive files was provided by District 6 to begin the task of bringing their SunGuide archives into STEWARD. This section describes the data that we have received and our progress in the conversion to STEWARD.

SunGuide archive TSS files were received from FDOT District 6 for several days in July. We have also received a detector inventory spreadsheet file from providing more information about each location, including detector type and coordinates. The facility metadata files were partially constructed using the information in these files. A map showing the known detector locations is presented in Figure 5. Some additional facility description information will be needed from the district.

Separate facility description databases are required for the station data and the lane data. The current status of the station information is described in Table 6. The current status of the lane information is described in Table 7. Each table describes the required data fields, the current development status and any issues that arose in extracting the information from the TSS archive files.

An Excel file containing all of the facility metadata developed from the archive files has been forwarded separately to District 6. Some further discussion with the district will be required to complete the metadata files.

3.5 District 7

District 7 is in the early stages of operation and no data have been received. The configuration of the District 7 data has therefore been deferred to the next phase of the project.

District6 Detector Map

Figure 5

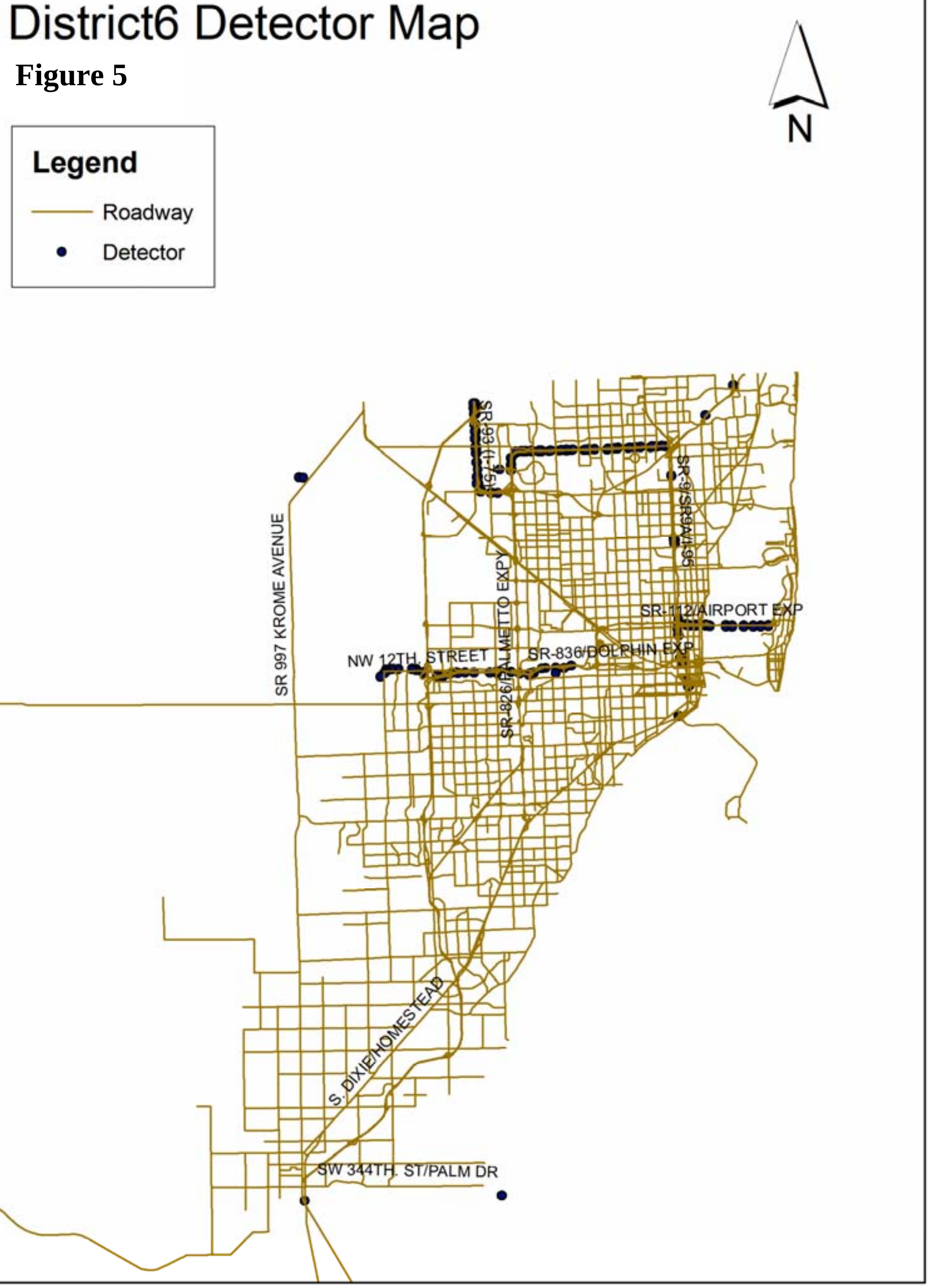


Table 6: Status and Issues for the Required Station Description Data for District 6		
Required Station Data	**Description**	**Status and Issues**
STATION INDEX	An index number assigned sequentially for internal computations	Waiting on updated files
STATION ID	A statewide-unique station identifier in the format DFnn, where: • D is the district number • F is the facility number within the district • nn is the sequence number of the station within the facility • Direction code (1=Increasing MP, 2=Decreasing MP)	Facility numbers will be assigned from the “Road” field below Station numbers will be assigned sequentially within each facility Direction code will be assigned from the “Direction Facing” field in the inventory
Description	A narrative description of the station location	Some stations already have names assigned in the inventory. More names are needed
Status	A code to indicate if the station is on line	Will be assigned from the “Status” field in the inventory
Road	This is the facility number (0-9) Detectors assigned to #9 included a group in the area of Sr 520 and 528 which did not appear to be on a cohesive route. Detectors that appeared in the archive files but were not in the inventory were also assigned to #9, pending resolution of their actual location.	Facility numbers will be assigned as follows: • 0 = I-75 • 1 = I-95 • 2 = I-195 • 3 = SR-826 • 4 = SR-836 • 9 = Unassigned detectors
LATITUDE LONGITUDE	Station coordinates	Obtained directly from the inventory spreadsheet.
STATE MILEPOST		This information is needed
ROADWAY_ID	From RCI	We can find this
ROADWAY MILEPOST	County milepost	We can find this if we know the state milepost
MAX_SPEED	Need for Delay at Performance Measure	We can find this
NUM_OF_LANES	Need for V/C ratio at Performance Measure	We can find this
LANE_CAPACITY	Need for V/C ratio at Performance Measure	This needs to be estimated as vphpl
UPNODE	Need for average volume at Performance Measure	We can find this
DETECTOR_TYPE		Is this available?
DETECTOR_UNIT	This was used in D2 to indicate the number of detector units at the station	Not used for any data processing purposes

Table 7: Status and Issues for the Required Lane Description Data for District 6		
Required Lane Data	**Description**	**Status and Issues**
TMC ID	The station ID as it appears in the "lane_id" field of the TSS archive Example: "578-link1-lane1"	All lanes that were contained in the TSS data files will be represented in the facility data.
CDWSTATION	Same as the STATIONCDW_NUM field in the station data description	All station numbers will be assigned
LANE	A statewide-unique lane identifier in the format SSSSDTn, where: • SSSS is the station number described above • D is direction (1=increasing mileposts, 2=decreasing mileposts) • T is a code for the type of lane • n is the lane number from the "lane_id" field	Direction code will be assigned from the "Direction Facing" field in the inventory More info will be needed on the laned type (Mainline, Ramp, HOV, etc.) Lane number is embedded in the TMC ID
DET_TYPE	The type of detector in this lane	Obtained from the "Type" field in the inventory
DIRECTION	1=increasing mileposts 2=decreasing mileposts	Determined as shown above for the lane direction
STATUS	A code to indicate if the detector is on line	All lanes will be assigned the same status code as the station in which they are located
ROADWAY_ID	From the RCI database	Info still needed. We will try to get it from the RCI
ROADWAY_MILEPOST	The county milepost	Info still needed. We will try to get it from the RCI if we can get the station milepost
MAX_SPEED	The speed limit at the station	Info still needed. We will try to get it from the RCI
TSS_SPEED	The upper limit of speed for travel time computation	Info still needed
OCCUPANCY	We are not sure what this is, but it was in the facility data obtained from District 2	Probably not needed
COUNT_STATION	The district or central office traffic count station number to be assigned if traffic count files are desired from this lane	This will require coordination with the district planning office and the central statistics office.

4 STEWARD Applications

The principal objective of the current project was to provide a "proof of concept" for a central data warehouse. There are two requirements for a proof of concept for a data archive. First, it must be demonstrated that the data can be collected and archived in a practical manner. Second, it must be demonstrated that the data can serve a useful purpose. The development of a functioning system as documented in this report should satisfy the first requirement.

The question of the usefulness of the data is addressed in this section. Several applications are documented that should indicate that STEWARD is capable of providing productive support to a variety of users. The original intent of the project was to have a large volume of data available for analysis at this point. While we anticipate receiving daily data from several district TMCs in the next phase of the project, the quantity of data that actually became available during the project period was disappointingly small. There were several delays in the implementation of satisfactory systems of sensors and communications in most of the districts. We believe that most of the problems have been overcome and we look forward to a significant expansion of the STEWARD database in the coming year.

In spite of the data limitations, STEWARD has been able to demonstrate a reasonable degree of utility. We have been able to interact with several current and potential users with a need for the archived data that STEWARD can supply. The current and future applications for the STEWARD data are discussed as follows:

4.1 Support for Work-Zone Crash Analysis

The primary focus of this research, which is sponsored by the Southeastern Transportation Center, is to determine the impact of reduced capacity on crashes in work zone queues in Jacksonville. The study is being performed by the University of Florida, in cooperation with the Florida Department of Transportation, Jacksonville Traffic Management Center and the Florida Highway Patrol. The study focuses on the I-95 Trout River Bridge reconstruction project and the Interstate-10/Interstate-95 interchange project in Jacksonville.

Crash data were obtained from the Florida Highway Patrol and those crashes occurring in the vicinity of the identified work zones were isolated from the larger crash data set. STEWARD data were used to confirm the traffic impacts that are caused by incidents near the work area. The dates used for this project were from June 2007 to December 2007. The STEWARD data included 15 minute aggregations of traffic volumes and speeds from the stations closest to the work zone.

4.2 Support for Analysis of Breakdown at a Freeway Ramp

The objective of National Cooperative Highway Research Program (NCHRP) Project 3-87, which started in October 2006, is to develop procedures for selecting ramp-management strategies for a freeway section under the threat of flow breakdown. These procedures will be evaluated using simulation in conjunction with field data. One of the current sites in the data collection plan will be within the District 2 SunGuide facility on Interstate 95. The archived volume, speed and occupancy data is well suited to that project's data needs. We expect to work closely with the NCHRP project team to supply the required data. This project provides a good example of a research application that will use short interval aggregations to model the breakdown of traffic flow on a freeway in the vicinity of an entrance ramp.

4.3 Support for Identification of Recurring Congestion

The consulting firm of RS&H is currently conducting a “Bus in Shoulder” study for the Jacksonville Transportation Authority. One of their tasks is to identify recurring congestion on I-95 in Jacksonville “Congestion” has been designated in terms of speeds below 35 mph. Their initial request was for monthly station-level and lane-level, volume and speed data on I-95 in the Jacksonville area. The STEWARD web site and documentation were provided to access and retrieve the traffic data via the STEWARD web pages.

This is an ongoing activity. RS&H has made some constructive suggestions regarding possible improvement of the data report formats to facilitate their use. We will pursue the need for changes with RS&H and other users during the next phase of the project. We have been advised that similar studies will be conducted in other districts.

4.4 Support for Travel Time Reliability Reporting

As part of Strategic Intermodal System (SIS) management, two research projects on travel time reliability models were developed for freeways travel time reliability. The first project used data from Philadelphia, PA and the second project is evaluating the feasibility of using truck travel time data collected by the Federal Highway Administration (FHWA) and the American Transportation Research Institute (ATRI) to estimate travel times and determine the travel time reliability for freeways in Florida.

The UF research team for that project obtained data for the I-95 freeway in Florida from the STEWARD website and are using it for model development. In addition to supporting the specific study, we anticipate a continuing involvement with the supply of travel time reliability data in support of ongoing research projects and FDOT’s periodic data reporting requirements

4.4.1 Additional Facility Data Required to Support TT Reliability Reporting

The current threshold of delay for purposes of travel time reliability analysis is established at 10 mph under the speed limit. Reporting of delay under this definition requires knowledge of the speed limit for each travel time link. This requirement is complicated by the fact that multiple speed limits may apply over a given travel time link.

Prior to the initiation of this support effort, the TSS facility data included the only speed limit and milepost for each SunGuide sensor station. To provide the supporting data for travel time reliability analysis, it was necessary to create a new field in the TVT facility data representing the speed limit for each link. A macro-utility was added to the TVT facility data spreadsheet to process the speed limit data in the following manner:

- Obtain the speed limit for each TSS station included in the travel time link from the TSS facility data spreadsheet
- Assign an area of influence within the link, based on the station mileposts
- Compute a distance-weighted average speed limit over the whole travel time link
- Insert the computed speed limit in the new “speed limit” field for the TVT facility data

This task was completed and the TVT facility database is now able to support the requirements of the travel time analysis study.

4.4.2 Modifications to STEWARD TVT Database

The previous STEWARD TVT database included a "delay" field that is referenced to a constant speed of 40 mph. The ETL Utility was modified to use the new criterion based on the distance-weighted speed limit. Three performance measures are now presented:

- Congestion Delay:
- On Time Delay:
- Percent of on-time trips

These performance measures were defined previously in the "Report Generation" section of this document.

4.5 Traffic Volume Data for the District and Statistics Office Traffic Count Program

The FDOT Statistics Office maintains several continuous telemetered traffic count stations on Florida highways. Three permanent traffic counters are located on I-95 within the District 2 SunGuide system. With the cooperation of the Statistics Office, the research team was able to compare the data from one count station to the archived counts generated by SunGuide and stored in the STEWARD database. The permanent count station was located in the southbound lanes of Interstate 95 between Emerson Street and University Blvd. The two adjacent SunGuide detector stations were located approximately 1000 ft north and 700 ft south of the permanent count station.

Figure 6 shows an example comparison between the hourly counts from the permanent count station and the two SunGuide detectors. Note that a near perfect agreement is apparent here. This will not always be the case and comparison of data from the two sources could potentially improve the accuracy of both sources. This possibility will be explored in the next phase of the project. There is clearly a potential benefit that could be derived from a mutual exchange of traffic count data between the ITS centers and the Statistics Office. The Statistics Office data could provide an important reference for calibrating the ITS detectors, most of which are microwave based. The ITS data could provide a useful supplement to the statewide traffic count coverage now in place.

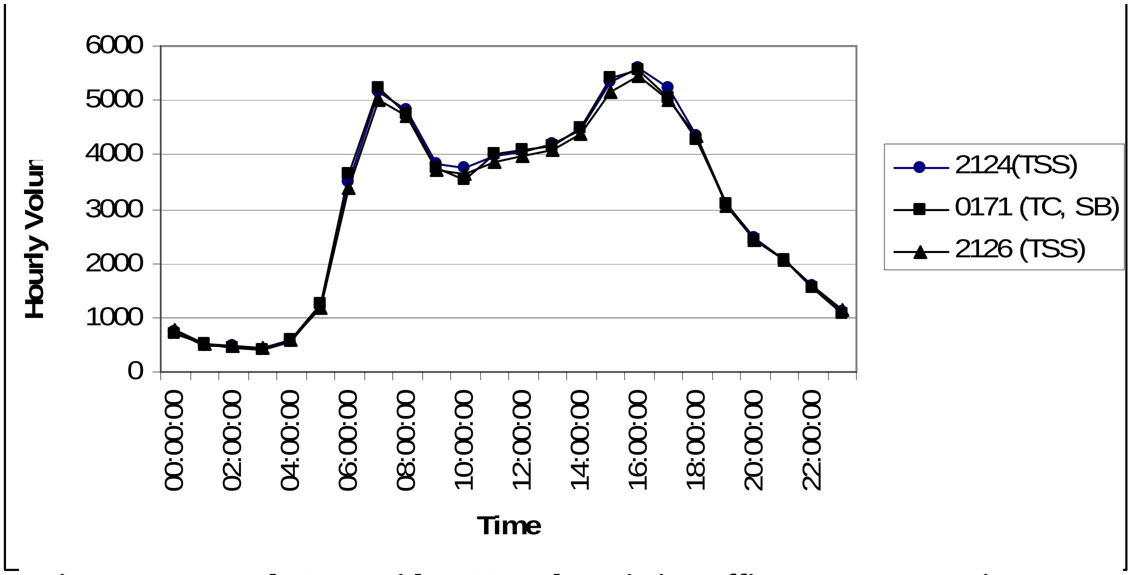


Figure 6: Example SunGuide TSS and Statistics Office count comparison

4.5.1 The Statistics Office Count Format

The traffic counts in the SunGuide archives have essentially the same content as the FDOT Statistics Office and District Planning Office traffic count files. A utility program has been developed to convert the count data in the SunGuide Data Archive to the Statistics Office count format.

The Statistics office has also prescribed a standard for counts derived from ITS detectors. The ITS format uses the “ITS” tag in place of the normal “CNT” tag, which denotes a traffic count file. This format also includes a validity check for each lane (N = normal, B = bad). As a preliminary criterion for generating ITS files, the detector data are assumed to be normal if the detectors reported volumes for at least 90% of the 20 second SunGuide polling intervals. More detailed QA criteria are under development.

4.5.2 District Planning Office Count Format

The district planning offices perform periodic counts at selected locations using various types of portable equipment. The SunGuide archive offers a potentially rich source for this type of data. The district count file format differs from the central office format because of differences in the processing software. The district format requires one record for each lane in format that differs slightly from the Statistics office format. Two different conversion processes are therefore required to accommodate both formats

The TSS lane data described earlier contains a field for the count data station. The station number begins with a “C” or a “D”, followed by a 4 digit number. Stations beginning with “C” are processed in the central office format and those beginning with “D” are processed as district count files. Count files are generated for all lanes with a valid entry in the station number field. Counts are not generated if the field is blank.

4.6 Integration with Statewide Crash Data Records

The Crash Analysis Reporting System (CARS) is implemented in the FDOT mainframe computer. The CARS database includes crash data dating back to 2001. An example of the CARS crash data presentation showing 5 crash records is shown in Figure 7. These data were collected from Duval county (Count ID: 72), section 290 (North of Jacksonville) on SR 9 (I-95). The columns represent:

- crash number,
- milepost,
- crash date,
- 1st harmful event,
- 2nd harmful event,
- lighting conditions,
- road surface conditions,
- site location,
- side road,
- lane number,
- total number of traffic fatalities and
- total number of injuries and influence area.

All of the details are described in page 112 of “User Manual for the Crash Analysis Reporting System”.

```
      CARB058             ON-LINE DETAIL DISPLAY          02/12/2007  15:17:53
 FROM 09/10/2006 TO 09/23/2006      A - ASCENDING          PAGE:     1 OF     1
 FROM CO/SEC/SUB: 72 290 000       MP: 000 . 000
 TO   CO/SEC/SUB: 72 290 000       MP: 010 . 513
 SR#: SR      9            CO/SEC/SUB: 72 290 000          TOT CRASHES:       5

                                  HARM EVNT      RDSF SITE SIDE LN  #   #  INFL
    CRASH#    -MP-     CRASH DATE  1ST 2ND  LGHT COND LOC  ROAD #  FTL INJ AREA

   769859390 001.193   09/21/2006   18 00    01   01   01   R   S       1
   769865750 002.180   09/10/2006   08 22    01   01   01   L   S   1   1
   720987470 004.328   09/17/2006   01 00    01   01   01   L   3
   769865780 006.271   09/19/2006   77 00    01   01   01   R   3
   769865200 006.355   09/20/2006   01 00    04   01   01   R   2       2
```

Figure 7: Crash data presentation from CARS

While there appears to be a wealth of information in the CARS database, the potential for integration with the archived data in STEWARD is severely constrained by the omission of the time of the occurrence of the crash. The raw SunGuide archives are generated with a 20 second and 60 second resolution for the TSS and TVT data, respectively. The STEWARD database aggregates these values into groups with 5 minute granularity.

The archived data could offer an excellent potential for integration with the crash records if the time of each crash were captured in the record. On the other hand, without this information, the potential for integration is very limited. Because of this limitation, no further exploration was carried out on the possible links between CARS and STEWARD. One of the recommendations at the end of this document suggests that this possibility be pursued in the next phase of the project.

4.7 Integration with the Roadway Characteristics Inventory

FDOT maintains a comprehensive roadway characteristics inventory (RCI) database containing several descriptive fields for each roadway segment. For example, the RCI data for I-95 in Jacksonville indicates that there are 106 segments, including 3 mainline segments, 9 one way segments and, 94 ramps. Some detailed examples of the RCI record format are presented in the following pages. Figure 8 shows a sample summary of some of the RCI records from this area. Figure 9 shows a sample of the detailed presentation of the characteristics for a single segment.

The integration of the RCI data with STEWARD is limited at this point to the provision of a field in the STEWARD facility data to indicate the RCI roadway segment for each TSS station. This will facilitate access within FDOT to the RCI Data for any SunGuide detector station. There are some possibilities for the creation of more automated links to this information. This question will be explored with STEWARD users when the user base has expanded sufficiently.

4.8 General Support for Periodic Reporting Requirements

As indicated in the “STEWARD Overview” section of this report, several performance measure reports are generated from both the TSS and TVT data. When the STEWARD user base is expanded, these reports will serve to facilitate the periodic reporting requirements for the districts. Changes to the performance report content and format to meet the district expectations will be made as necessary.

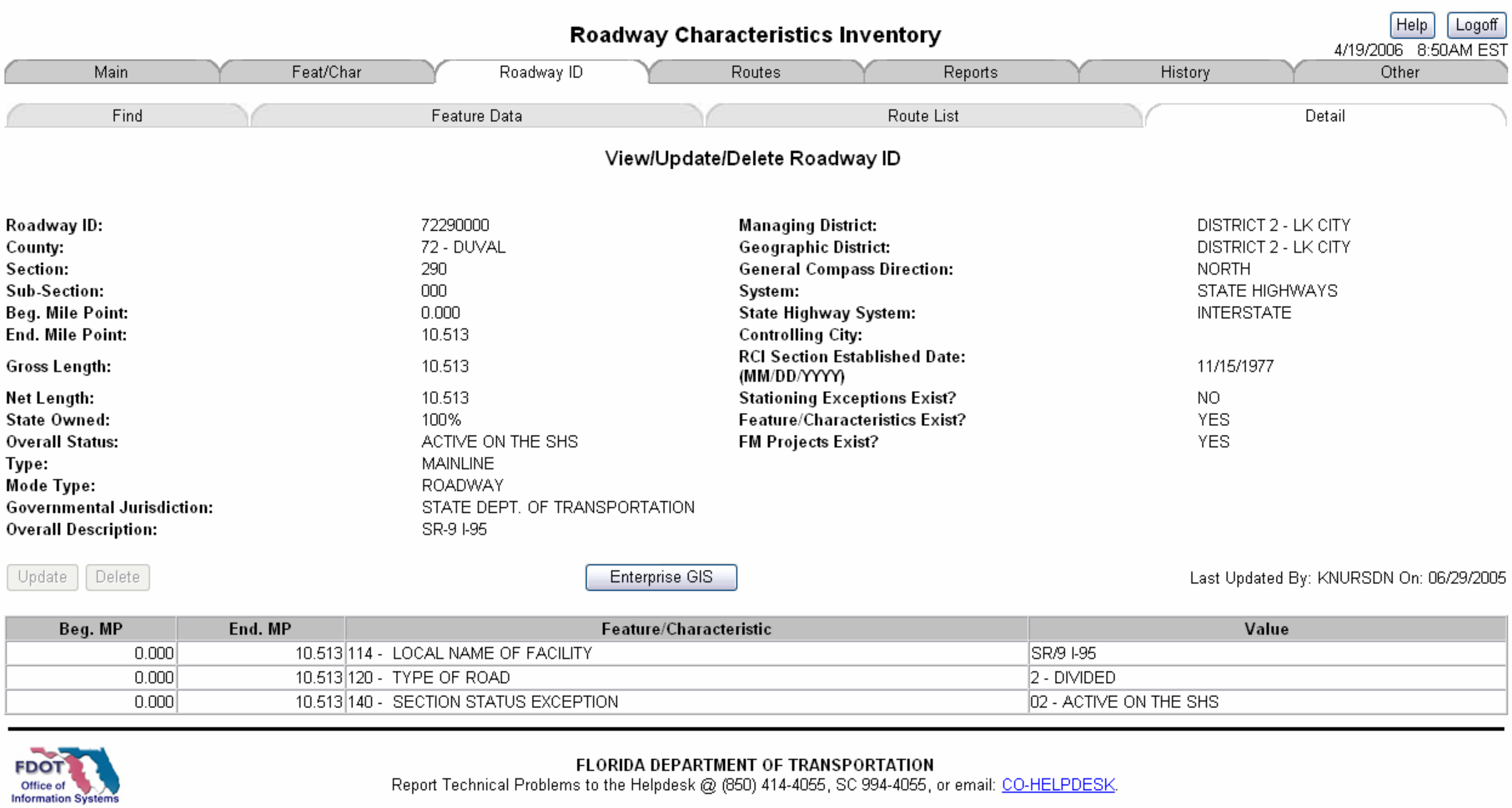
Instead of adhering to a rigid format, all reports are now generated as comma-separated value (CSV) files that may be directly imported into office productivity programs such as Microsoft Excel, etc. This will allow the districts to modify the actual presentation formats to meet their individual preferences.

4.9 Diagnostic Support for TMC Detector Operation and Maintenance

As indicated in the “STEWARD Overview” section of this report, several diagnostic reports are generated from both the TSS and TVT data. In the current phase of development, the data were obtained from District 2 on a more or less monthly basis. When the acquisition of the archived data is streamlined as proposed in the next phase, the diagnostic reports will be able to be generated on a schedule that will give more timely feedback to the personnel at the SunGuide TMCs. This feedback should provide useful support for the maintenance of their detector systems and communications facilities

Order Num.	Beg. Route MP	Roadway ID	Beg. MP	End. MP	Segment Type	Route Side	Distance Calc Exclude Code	Route vs. Section Direction	Connected to Roadway Segment	Connected to MP
16	331.958	72280000	0.000	16.793	MAINLINE		NO	PLUS		
16	338.516	72280215	0.000	0.183	RAMP	L	YES	PLUS	72280000	6.558
16	338.606	72280211	0.000	0.203	RAMP	R	YES	PLUS	72280000	6.648
16	338.680	72280213	0.000	0.115	RAMP	L	YES	PLUS	72280000	6.722
16	338.759	72280214	0.000	0.313	RAMP	R	YES	PLUS	72280000	6.801
16	338.783	72280212	0.000	0.157	RAMP	L	YES	PLUS	72280000	6.825
16	341.094	72280217	0.000	0.165	RAMP	R	YES	PLUS	72280000	9.136
16	341.113	72280216	0.000	0.214	RAMP	L	YES	PLUS	72280000	9.155
16	341.483	72280219	0.000	0.208	RAMP	R	YES	PLUS	72280000	9.525
16	341.491	72280218	0.000	0.174	RAMP	L	YES	PLUS	72280000	9.533
16	343.300	72280223	0.000	0.350	RAMP	R	YES	PLUS	72280000	11.342
16	343.372	72280220	0.000	0.287	RAMP	L	YES	PLUS	72280000	11.414
16	343.554	72280222	0.000	0.212	RAMP	R	YES	PLUS	72280000	11.596
16	343.707	72280238	0.000	0.203	RAMP	L	YES	PLUS	72280000	11.749
16	343.854	72280224	0.000	0.266	RAMP	L	YES	PLUS	72280000	11.896

Figure 8: Sample summary table for RCI Segments in District 2



Roadway Characteristics Inventory

Help | Logoff

4/19/2006 8:50AM EST

Main | Feat/Char | Roadway ID | Routes | Reports | History | Other

Find | Feature Data | Route List | Detail

View/Update/Delete Roadway ID

Roadway ID:	72290000	**Managing District:**	DISTRICT 2 - LK CITY
County:	72 - DUVAL	**Geographic District:**	DISTRICT 2 - LK CITY
Section:	290	**General Compass Direction:**	NORTH
Sub-Section:	000	**System:**	STATE HIGHWAYS
Beg. Mile Point:	0.000	**State Highway System:**	INTERSTATE
End. Mile Point:	10.513	**Controlling City:**	
Gross Length:	10.513	**RCI Section Established Date: (MM/DD/YYYY)**	11/15/1977
Net Length:	10.513	**Stationing Exceptions Exist?**	NO
State Owned:	100%	**Feature/Characteristics Exist?**	YES
Overall Status:	ACTIVE ON THE SHS	**FM Projects Exist?**	YES
Type:	MAINLINE		
Mode Type:	ROADWAY		
Governmental Jurisdiction:	STATE DEPT. OF TRANSPORTATION		
Overall Description:	SR-9 I-95		

Update | Delete | Enterprise GIS

Last Updated By: KNURSDN On: 06/29/2005

Beg. MP	End. MP	Feature/Characteristic	Value
0.000	10.513	114 - LOCAL NAME OF FACILITY	SR/9 I-95
0.000	10.513	120 - TYPE OF ROAD	2 - DIVIDED
0.000	10.513	140 - SECTION STATUS EXCEPTION	02 - ACTIVE ON THE SHS

FDOT Office of Information Systems

FLORIDA DEPARTMENT OF TRANSPORTATION

Report Technical Problems to the Helpdesk @ (850) 414-4055, SC 994-4055, or email: CO-HELPDESK.

Figure 9: Sample of the RCI description for a roadway segment

4.10 Data for Congestion Modeling Research

Research support will become a more important and visible function of STEWARD when the database is expanded to include more districts and more daily archive data. The extent of congestion on a given facility may now be assessed at any point in time. The ability to relate larger quantities of congestion data with external variables will be of great support to the research community. Some examples of such relationships include:

- General quantification of congestion
- Delay caused by incidents
- Development of flow propagation models based on traffic flow theory
- Investigation of the prediction of internal breakdown at high v/c ratios
- Investigation of the prediction of crashes related to turbulence in the traffic stream
- Investigation of the prediction of future conditions at a location based on current upstream conditions.

These and other phenomena will be explored in the next phase of the project

5 Data Quality Assurance Considerations

One of the objectives of the project was to develop QA procedures to apply to the archived data. This objective has been partially met as described below. The lack of adequate data from the districts has precluded complete development and testing of a comprehensive QA procedure. This effort will be continued in subsequent project activities.

5.1 Data Overview

The quality control methods offered in the FHWA report on "Monitoring Urban Freeways in 2003" [2] will be used initially as the basis of the STEWARD data quality assessment. These methods have been widely applied to traffic data in 29 cities in US and the results have been. In this phase of the project, these methods are applied to our project and the results will be validated.

At this point, STEWARD archives traffic data from District 2. A description of the SunGuide TSS detector configuration was presented previously.

5.2 Completeness Test for the District 2 Traffic Data

Completeness is defined as *the degree to which data values are present in the attributes that require them*. This is a percentage value calculated from the available number of data values as a percent of the number of total expected data values.

$$Percent\ Complete\ (\%) = \frac{n_{available values}}{n_{total expected}} \times 100$$

Where

$n_{\text{available values}}$ = the number of records or rows with available values present and
$n_{\text{total expected}}$ = the total number of records or rows expected

In this calculation, completeness is defined to include both valid and invalid data values as long as both types of data values are present in the version of data being evaluated.

Completeness of traffic data from District 2 were examined on five-day data from 9/11/06 (Monday) to 9/15/06 (Friday). The results are as follows:

Completeness = (total number of traffic data during the period)
/ (total number of detector-lane * 5 days * (24 hours / 20 sec))
= (1,656,105 + 1,670,684 + 1,675,965+1,683,734+ 1,698,437)
/ (447 * 5* (24 * 60 * 60) /20)
= 0.86844

But the SunGuide system has an issue on the scan interval. District 2 data are preset to have the scan interval of 20 second but there are several drifts on traffic data. For example, the completeness of traffic data on 9/6/06 would be as follows:

Completeness = (3,226,033) / (447 * 1* (24 * 60 * 60) /20) = 1.67062, which violates the basic definition of completeness. This happens because 49.55% of the scan intervals are less than than 14 seconds. It also has 17,943 observations data that have two or more traffic data reports for the same lane during the same second.

Completeness greater than 1.0 occurs on 12 days from August to September of 2006. This presents a critical issue for quality assurance and is an example of one of the problems, as identified in Section 1.2 of this report, which have delayed our progress in the development of this aspect of STEWARD.

5.3 Data Validity Check for the District 2 Traffic Data

There are four basic rules and thirteen quality check criteria in Reference 2. They are generated from the current practices in other traffic centers or data archiving systems, suggested practices recommended in the literature and the practices found to be necessary from project team analysis of the archived data. Four basic rules are summarized in Table 8. Thirteen quality check criteria are summarized in Table 9.

Table 8: Basic Rules for the Data Validity Check

Quality Control Rules	Sample Code with Threshold Values	STEWARD
Controller error codes	If VOLUME={code} or OCC={code} or SPEED={code} where {code} typically equals “-1” or “255”	N/A
No vehicles present	If SPEED=0 and VOLUME=0 (and OCC=0)	
Consistency of elapsed time between records	Elapsed time between consecutive records exceeds a predefined limit or is not consistent	TBR
Duplicate records	Detector and date/time stamp combination are identical.	TBR

The first rule (Controller error codes) cannot be applied to STEWARD because the SunGuide archive data do not include such error codes. The second rule was applied to STEWARD to verify the detector malfunctions. The third and fourth rules are related to the complete test problems that need to be resolved before they can be applied.

The second quality control rule was applied to District 2 traffic data between 9/11/06 (Monday) and 9/15/06 (Friday). The following results suggest that four stations (#29, #33, #35 and #40) need to be verified.

	9/11/06	9/12/06	9/13/06	9/14/06	9/15/06
Null lanes	14	14	14	14	14
Null stations	4	4	4	4	4

Table 9: Quality Control Criteria for the Data Validity Check

Quality Control Test	Sample Code with Threshold Values	STEWARD
QC1-QC3: Logical consistency tests	If DATE={valid date value} (QC1) If TIME={valid time value} (QC2) If DET_ID={valid detector location value} (QC3)	Verified during ETL process
QC4: Maximum volume	If VOLUME > 17 (20 sec.) If VOLUME > 25 (30 sec.) If VOLUME > 250 (5 min.) If VPHPL > 3000 (any time period length)	Implemented[1]
QC5: Maximum occupancy	If OCC > 95% (20 to 30 sec.) If OCC > 80% (1 to 5 min.)	Implemented[1]
QC6: Minimum speed	If SPEED < 5 mph	Implemented[1]
QC7: Maximum speed	If SPEED > 100 mph (20 to 30 sec.) If SPEED > 80 mph (1 to 5 min.)	Implemented[1]
QC8: Multi-variate consistency	If SPEED = 0 and VOLUME > 0 (and OCC > 0)	Implemented[1]
QC9: Multi-variate consistency	If VOLUME = 0 and SPEED > 0	Implemented[1]
QC10: Multi-variate consistency	If SPEED = 0 and VOLUME = 0 and OCC > 0	Implemented[1]
QC11: Truncated occupancy values of zero	If OCC = 0 and VOLUME > MAXVOL where MAXVOL=(2.932*ELAPTIME*SPEED)/600	Implemented[1]
QC12: Maximum estimated density	IF ((VOLUME*(3600/NOM_POLL))/SPEED) > 220 where NOM_POLL is the nominal polling cycle length in seconds.	Implemented[1]
QC13: Consecutive identical volume-occupancy speed values	No more than 8 consecutive identical volume-occupancy-speed values. That is, the volume AND occupancy AND speed values have more than 8 consecutive identical values, respectively. Zero ("0") values are included in this check.	Implemented[1]
[1]: Implemented but will be accessible to the general user as part of next project phase.		

For STEWARD purposes the thirteen quality check criteria have been applied only to the period from 6:00AM to midnight. In general, the traffic volumes are low in early morning time and their exclusion was necessary to avoid problems associated with detection accuracy issues.

All of the quality rules are applied to District 2 traffic data between 9/11/06 (Monday) and 9/15/06 (Friday) and the results are as follows:

- QC1 ~QC3: 100% of traffic data pass
- QC4: 0.15% of traffic data fail (2235+ 1817 +2133+ 2542 + 2351)/ 7241400 = 0.15%
- QC5: 0.01% of traffic data fail (77+182+136+147+241)/ 7241400 = 0.01%
- QC6: 2.27% of traffic data fail (668+1577+1023+159561+1397)/ 7241400 = 2.27%
- QC7: 100% of traffic data pass
- QC8: 0.01% of traffic data fail (75+233+106+144+247)/ 7241400 = 0.01%
- QC9: 0.00% of traffic data fail (11+136+26+53+117)/ 7241400 = 0.00%
- QC10: 1.43% of traffic data fail (2176+2187+2157+1994+1823)/ 7241400 = 1.43%
- QC11: 100% of traffic data pass
- QC12: 0.06% of traffic data fail (641+1119+572+805+1000)/ 7241400 = 0.06%
- QC13: 0.00% of traffic data fail (11+51+72+9+10)/ 7241400 = 0.00%

Overall, 2.65% of traffic data fail thirteen quality check criteria. This suggests very good results in comparison to those reported in Reference 2, which would be in the range of 55% to 100% in the 2003 freeway data from 29 US cities.

5.4 Preliminary QA Assessment

There are several developmental issues in the STEWARD quality control process.

- The archived data from the District 2 SunGuide system need to be validated to generate traffic data that satisfy the completeness test all the times.
- Each criterion needs be investigated and adjusted for STEWARD. For example, QC7 maximum speed might not be appropriate for SunGuide data archive because the system imposes a maximum speed threshold to constrain the reported travel times. Other numbers also need to be calibrated.
- An integrated and automated QA process needs be developed. Existing methods are executed manually but they should be automated for productivity.
- A strategy for processing the invalid data needs be established. Some imputation rules should be considered.

5.5 Future QA Tests

The existing QA tests focus on each individual sensor as a separate entity. With the addition of the facility configuration data it is possible to expand the QA process to include additional tests that use relationships among the lane detectors. Examples of such relationships include traffic volume distribution by lane at a given station, and consistency of volumes between adjacent stations. The additional tests are being developed now and will be investigated in the coming months.

6 Recommendations for STEWARD Operation and Expansion

This report and its appendices have demonstrated that TMC data can be archived in a practical manner and that the results can be of value to a variety of current and potential users. It is therefore recommended that FDOT proceed towards the implementation and expansion of a STEWARD. Some specific recommendations are presented as follows:

6.1 Future Support

The University of Florida Transportation Research Center is committed to the continued operation and expansion of STEWARD. Our federally funded "Center for Multimodal Solutions for Congestion Mitigation" (CMS) has agreed to provide internal funding to carry out a project with the following tasks:

- Present a series of workshops for potential providers and users of archived data
- Expand the scope of the CDW database to include additional SunGuide detectors
- Operate the system to provide data and reports to agencies and researchers
- Analyze the data to explore congestion modeling relationships

The UF project will allow the CDW to operate, collect data and serve its stakeholders for another year; however additional implementation support from FDOT will be required to bridge the gap between the prototype system and a fully operational CDW. The following objectives that are not addressed by the UF-sponsored project have been identified by FDOT for continued support under a separate project:

- Integrate the CDW functions with other FDOT data management programs.
- Transfer the CDW operations to the TERL center in Tallahassee
- Automate and fine tune the transfer of daily archive data from district TMCs and the statewide monthly, quarterly, and annual reporting processes.
- Expand the CDW to include other data sources.
- Incorporate the CDW functionality into requirements for future SunGuide versions.

To ensure a smooth transition from research to operational status, the FDOT ITS office has indicated that it is willing to commit resources within the TERL center to support the Tallahassee-based activities related to the CDW. These resources will be instrumental to the success of the project in collecting data and preparing reports that are useful to stakeholders.

It is important that the efforts of the two research projects and the TERL support be coordinated in a manner that will avoid duplication and serve the interests of the overall CDW implementation. Table 10 presents an overview of the implementation objectives and summarizes the role of the three activities in support of these objectives. Note that an important element of the TERL support will be liaison and interface between the university-based research and the stakeholders themselves.

6.2 Outreach to SunGuide TMCs

The expansion of the data archive to include all SunGuide TMCs that produce archive data should be a high priority for the near future. To promote continuity with the current project, the summary sheet shown in Figure 10 has recently been distributed to all present and future SunGuide TMCs, along with the archive configuration instructions presented in Appendix A of this report. The continuing project activities should follow up with the districts to ensure that they give adequate attention to the configuration of their archive features.

6.3 Recommended FDOT Actions

During the course of the project a number of possible actions by FDOT were identified that could improve the operation and usefulness of a central data warehouse. The specific actions are discussed separately.

6.3.1 Embedding the Source TMC in the Archive Data File Name

The current SunGuide naming convention for the daily archive files includes only the date in the form “mmddyyyy” to distinguish the file from those collected on other days. While this convention serves the internal needs of each TMC, it creates the potential for confusion in the loading of the data into the central database because multiple files with identical names must be processed at the central location. This is likely to become a significant problem as STEWARD expands to deal with large amounts of data from several TMCs. The solution to the problem would be to embed a district identifier in the archive file names in a future release of SunGuide.

6.3.2 Standardizing the Station and Lane Naming Conventions for the SunGuide Detectors

Each record in the TSS daily archive file contains one field for the station identification and another field for the lane identification. Each district now applies its own conventions in establishing the required identifiers. The only requirement is that the names be unique within each district. The STEWARD database must attach an identifier to each station and lane that is unique within the entire state. The scheme adopted for this project assigns a four character code to each station and a seven character code to each lane. The station number format is “DFnn”, where:

- D is the district number
- F is the facility number assigned by the district
- nn is the station number within the facility

The compositors of the lane number are

- Station number (4 characters)
- Direction (1 character)
- Function code (1 character)
    1. Left entrance ramp
    2. Left exit ramp
    3. Freeway main lane
    4. Right entrance ramp
    5. Right exit ramp
    6. Auxiliary lane
    7. HOV Lane
- Lane Number (1 character)

All new archive data to be added to the STEWARD database must have the identifiers assigned by the district converted to this format to ensure that detected lane will have a unique identifier. The adoption of a standard to be applied to SunGuide systems that have not yet assigned their identification codes would greatly facilitate assimilation of their data by STEWARD. The scheme described here could be adopted easily by each new SunGuide system.

6.3.3 Adding a Date and Time Stamp to Crash Records in CARS

As noted previously, the potential for integration of STEWARD data with crash records is severely limited by the lack of a time stamp in the individual record for each crash. This information is available on the crash report form but crash times cannot be correlated productively with the archived traffic data without automated access to this information. It is therefore recommended that FDOT consider adding a time stamp to the CARS crash record data.

Table 10: Coordination of Proposed Central Data Warehouse Activities for 2008

Objective/Task	UF CMS Project	FDOT Project	Possible TERL Support Activities
Present a series of workshops for potential providers and users of archived data	Requirement of Task 1		1. Participate in workshop organization and presentation. 2. Perhaps expand workshop to cover other TERL activities
Expand the scope of the CDW database to include additional SunGuide detectors	Requirement of Task 2		Serve as a focal point for facility data collection
Operate the system to provide data and reports to agencies and researchers	Requirement of Task 3		1. Serve as a focal point for operation 2. Perform ETL and QA operations after transfer of operation from UF Campus 3. Provide reports to stakeholders 3. Maintain the CDW web site
Analyze the data to explore congestion modeling relationships	Requirement of Task 4:		
Integrate the CDW Functions with Other FDOT Data Management Programs		Requirement of Task 1:	1. Provide liaison with other FDOT offices that produce or use data 2. Organize and schedule stakeholder meetings
Transfer the CDW Operations to the TERL Center in Tallahassee		Requirement of Task 2:	1. Collaborate with the UF Team to ensure a smooth transfer of operations 2. Review the software documentation to ensure maintainability
Automate and fine tune the transfer of daily archive data from district TMCs and the statewide reporting processes.		Requirement of Task 3:	Provide FDOT network support for data transfer
Expand the CDW to Include Other Data Sources		Requirement of Task 4:	Advise the UF research team on potential data sources.
Incorporate the CDW Functionality into Requirements for Future SunGuide Versions		Requirement of Task 5:	1. Identify stakeholder requirements 2. Prepare the framework for requirements documents. 3. Review the draft documents submitted by the UF Team

Figure 10
FDOT District SunGuide TMC Archive Summary

TMC Information: District _______ TMC ____

Contact Name _________________________ Phone ______________ email ______________________________

TSS Subsystem Data

SunGuide generates a TSS archive file at the end of each day containing a comma delimited text representation of the volume, speed and occupancy for each detector in the system. Is your SunGuide system currently producing daily TSS archive files? ____ Yes _____ No _____ Don't know

If yes, the reporting interval is _______ seconds. (The recommended interval is 20 seconds.)

If no, when do you expect to activate the SunGuide TSS archive feature? ___________________________________

Each district may have up to 10 facilities, numbered 0 through 9. Facility #9 is generally reserved for isolated stations. Please list your facilities as you would like them to appear in the STEWARD database. If you have not yet subdivided your system into facilities, please just list the totals below

Status of TSS Detectors for Archive Data						
ID	**Description**	**Current**		**Anticipated 2008**		**Detector Types**
		Lanes	**Stations**	**Lanes**	**Stations**	**RTMS, Loop etc.**
0						
1						
2						
3						
4						
5						
6						
7						
8						
9	(Unassigned)					
Total						

Detector Station and Lane Data Requirements

Some additional lane and station information is required to ensure that each record in the STEWARD database represents a globally unique location, to support the analysis and reporting of system based measures and to relate the measures obtained from a specific location to other forms of data, such as RCI, Statistics Office counts, crash records etc.

Required Data for Each Station:

A station includes one or more detected lanes covering the facility in one direction. Can you provide the following information for each station?

Yes ____ No ____ Description (Example: I-95 NB at Forest St)

Yes ____ No ____ Status (Normal or Offline)

Yes ____ No ____ Facility Number (Assigned above)

Yes ____ No ____ Coordinates (Latitude, Longitude)

Yes ____ No ____ State Milepost

Yes ____ No ____ RCI Roadway Id

Yes ____ No ____ RCI Roadway Milepost

Yes ____ No ____ Maximum Speed

Required Data for Each Detected Lane:

Can you provide the following information for each lane that reports TSS Data?

Yes ____ No ____ Lane number at the station

Yes ____ No ____ Direction

Yes ____ No ____ Lane type or function (Mainline, entrance ramp, exit ramp, auxiliary, HOV)

Yes ____ No ____ Detector type (RTMS Loop, etc.)

Yes ____ No ____ Max Speed (Required for travel time reliability reporting)

TVT Subsystem Data

The TVT subsystem reports daily travel times for each travel time link configured by the TMC. The time resolution is normally one minute. Is your SunGuide system currently producing daily TVT archive files? ____ Yes _____ No _____ Don't know

If no, when do you expect to activate the SunGuide TVT archive feature? ______________________________

Can you provide the following information for each travel time link that you have configured?

Yes ____ No ____ Link ID for the archive file

Yes ____ No ____ Link Description

Yes ____ No ____ Link Direction

Yes ____ No ____ Origin ID (usually a DMS location)

Yes ____ No ____ Destination ID (Usually a TSS Station)

Yes ____ No ____ Lane number at the station

7 Description of Appendix Material

To facilitate review, most of the material in this report is provided as separate appendices. We believe that separate modules will be desirable because of the different focus of various users. The following modules are presented as appendices:

7.1 Instructions and Guidance for Configuring the SunGuide Facility Data for STEWARD

A description of the configuration procedures is presented in Appendix A. This module has been developed to assist TMC personnel in configuring their SunGuide facility data. The procedures followed by the project staff have in the configuration of the District 2 archive been documented to support future configuration efforts in other systems during the continuation phases of the project. The topics include:

- TSS Archived Data Format
- Conversion of TSS Data to the STEWARD Database
- Required TSS Facility Data
- TVT Archived Data Format
- Conversion of TVT Data to the STEWARD Database
- Conversion of SunGuide Archive Data to the FDOT Count Formats
- Transferring Archive Data to STEWARD

7.2 Instructions and Guidance for Installing STEWARD at TERL

A description of the steps required to install the STEWARD software and databases is presented in Appendix B. This module has been developed to assist FDOT and ITS contract personnel in setting up STEWARD in the TERL facility. The topics include:

- Oracle Database Program Installation
    - Installing Oracle 10g Release 2
    - Installing Oracle Workflow
    - Oracle Warehouse Builder (OWB) Setup
- STEWARD Deployment
    - Prerequisites:
    - Importing metadata
    - Registration on the Control Center Manager
    - Deployment and data loading process

7.3 Instructions and Guidance for Operating STEWARD

A description of the steps required to obtain data from the districts and add it to the STEWARD databases has been developed to assist FDOT personnel in operating and maintaining STEWARD in the TERL facility. This information is presented in Appendix C. The topics include:

- Physical architecture of STEWARD
- The ETL process
- Exporting metadata from the database
- Database table configurations

7.4 Instructions and Guidance for Accessing STEWARD Data on the Internet

While some use will be made within FDOT by accessing the STEWARD databases directly, most users in the future will gain access to STEWARD via the internet. A document describing the internet based features of STEWARD in presented in Appendix D. The topics include:

- An overview of the STEWARD web interface.
- An explanation of the STEWARD web pages
- Instructions for accessing the district data and reports
- Examples demonstrating the operations to be performed

7.5 STEWARD Design Document

A detailed description of the design features of STEWARD is presented in Appendix E. This document will be of interest to those who have a need to be familiar with the details of the system design and operation. The topics include:

- STEWARD system overview
- STEWARD database development and management
    - External tables
    - Dimension tables
    - Fact tables
    - Materialized views
- The STEWARD ETL Process
    - Mappings
    - Process flows
    - Schedules
- STEWARD Web Interface
    - Program architecture
    - STEWARD web implementation

This appendix will serve as the definitive STEWARD document for future maintenance and enhancement.

7.6 Workshop Material

Future workshops will be presented to cover the use of the STEWARD data by TMC operators, analysts and managers. The workshop agenda, which covers a half-day period, is shown In Figure 11. Each of the six sessions covers a nominal half-hour period.

The PowerPoint presentation files and other workshop materials are included in the supplemental CD material in Appendix G.

Figure 11: STEWARD workshop agenda

1. Executive Summary
- Benefits of a central data warehouse
- STEWARD features
- District support requirements

2. Facility Data Configuration
- Archive file contents
- Need for metadata
- Station data file content
- Lane detector file content
- Facility data development process

3. The ETL Process
- ETL utility
- Diagnostic reports
- Database development
- Oracle Scheduler operation

4. Available Reports
- Diagnostic
- Quality assessment
- Performance measures

5. Database Operations
- Internet site overview
- Queries
- Selection criteria
- Query results

6. Interaction with Other Data
- Traffic Count Data
- RCI Data
- Crash records Data
- TMC Simulation

7.7 Description of Delivered Software and Data Files

Several software items and data files were produced to support the development and operation of STEWARD. All products developed in connection with this project are delivered as a part of the final report. Their documentation is incorporated into Appendix F and the program and data files are included on the supplemental CD of Appendix G: The items that fall into this category include:

- MPConverter: A set of milepost converter routines required to convert milepost information to coordinates and *vice versa*. This product will be useful for configuring the facility metadata for additional TMCs to be brought into STEWARD.

- TSSBuilder: A utility that reads the raw data from a SunGuide archive over several days and constructs a list of all of the stations and lanes that have reported volume, speed and occupancy data. This program performs the first step in configuring the STEWARD facility data for a SunGuide TMC. The documentation for this program has also been incorporated into the workshop material

- Internal Utility Programs**:.** The Oracle Warehouse Builder has several limitations with respect to the ETL process. To work around these limitations, several utility programs have been developed to give STEWARD the required flexibility and expandability.

- SimTMC: A utility that uses a microscopic traffic simulation model to emulate a SunGuide freeway traffic management center. The program produces input data files for the FRESIM freeway simulation model and creates data archive files in the prescribed SunGuide format. It was originally developed for testing the STEWARD database before the actual data became available. It could be useful in the future for general simulation of a SunGuide TMC.

7.8 Supplemental Material Delivered on CD

The following files are included in the CD delivery:

- PDF version of this report, including appendices
- PowerPoint files for the STEWARD workshop
- MPConverter software: source and installation files
- TSSBuilder software: source and installation files
- SIMTMC software: source and installation files
- Current versions of all facility metadata files for Districts 2, 4, 5 and 6. The files for each district are contained in a separate Excel workbook. Each workbook contains three spreadsheets covering the facility, station and lane description data.
- STEWARD ETL Utility programs: Source and installation files
    - Move2SunETLUtilityFolder
    - Move2StewardFolder
    - PrepareFileLoading
- STEWARD ETL Data Warehouse Builder Metadata File (STEWARD backup file)
- STEWARD database backup file
- Facility configuration data files for all districts

8 Summary and Closure

This report and its appendices have demonstrated that TMC data can be archived in a practical manner and that the results can be of value to a variety of current and potential users. It is suggested that these findings justify further implementation and expansion of STEWARD. The ongoing support for the development and implementation will carry the project forward for another year. An important expansion of STEWARD and a visible demonstration of its effectiveness will result from this effort.

It must be emphasized that the prototype developed by this project represents a proof of concept and not a fully operational system. The research team hopes that support for statewide implementation of a fully configured central data warehouse will not be too far in the future.

9 References

1. University of Florida Transportation Research Center. "Travel Time Reliability and Truck Level Of Service on the Strategic Intermodal System", Final Report, FDOT Contract BD-545-48, April 2007

2. Texas Transportation Institute, "Monitoring Urban Freeways in 2003: Current Conditions and Trends from Archived Operations Data" Report Number: FHWA-HOP-05-018, *December 2004.*

10 Technical Appendices

Appendix A: SunGuide Configuration Instructions

Appendix B: STEWARD Installation Instructions

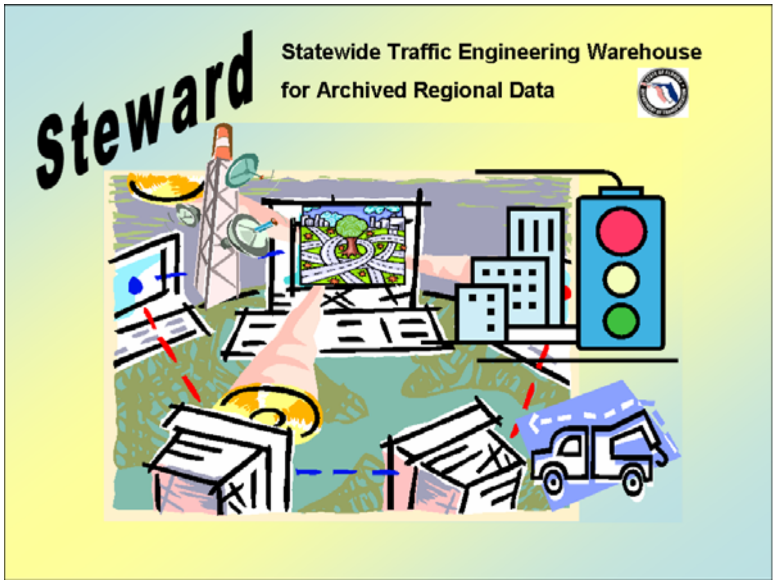
Appendix C: STEWARD Operating Instructions

Appendix D: Internet Access Documentation

Appendix E: STEWARD Design Document

Appendix F: Description of Delivered Software and Data Files

Appendix G: Supplemental Material Delivered on CD



Development of a Central Data Warehouse for Statewide ITS and Transportation Data in Florida

Phase II: Proof of Concept Final Report

Appendix A
Instructions and Guidance for Configuring the SunGuide Facility Data for STEWARD

Prepared for the Florida Department of Transportation
By the University of Florida Transportation Research Center

Contract # BD545, RPWO # 37
UF Project 051449

Revision Date: April 15, 2008

Table of Contents

List of Figures

List of Tables

1 Introduction

The Statewide Traffic Engineering Warehouse for Regionally Archived Data (STEWARD) will archive data from a variety of sources in a database that will support the generation of reports and queries. A prototype is being developed now as a proof of concept for a fully operational version to be deployed in the future. The STEWARD database will eventually reside in Tallahassee at the Florida Department of Transportation (FDOT) Traffic Engineering Research Laboratory (TERL). A prototype system is now being developed in Gainesville by the University of Florida (UF) Transportation Research Center.

This document deals with the requirements and procedures for transferring the daily archive data from the SunGuide traffic sensor subsystem (TSS) and travel time (TVT) subsystem to the STEWARD database. The procedures for data transfer were developed using the I-95 SunGuide system in Jacksonville as a model. Other SunGuide traffic management centers (TMCs) will be added as the archive data become available

The system, as shown in Figure A-1, is configured with approximately 450 detectors at 120 Road Traffic Microwave Sensor (RTMS) detection stations covering a 25 mile section of the freeway. Figure 1 also shows an overview of the archived data flow. The following discussion describes the required facility data for each TMC and the data formats for the raw archived data and the STEWARD database.

2 TSS Archived Data Format

The TSS data are stored in a comma-delimited flat file, with each file representing a 24 hour day. Zipped versions of these files are posted periodically by the District 2 staff. The structure of the raw archive data file is shown Figure A-2. Each record in the file represents the volume, speed and occupancy data from one lane of the freeway over a single 20 second period. So, a file representing 24 hours of data at 450 detectors will contain (450 x 24 x 60 x 3) or approximately 2 million records. The size of each file is approximately 85 megabytes. The zipped version is reduced to approximately 10 megabytes, which is a reasonable size for retrieval from file transfer protocol (FTP) site.

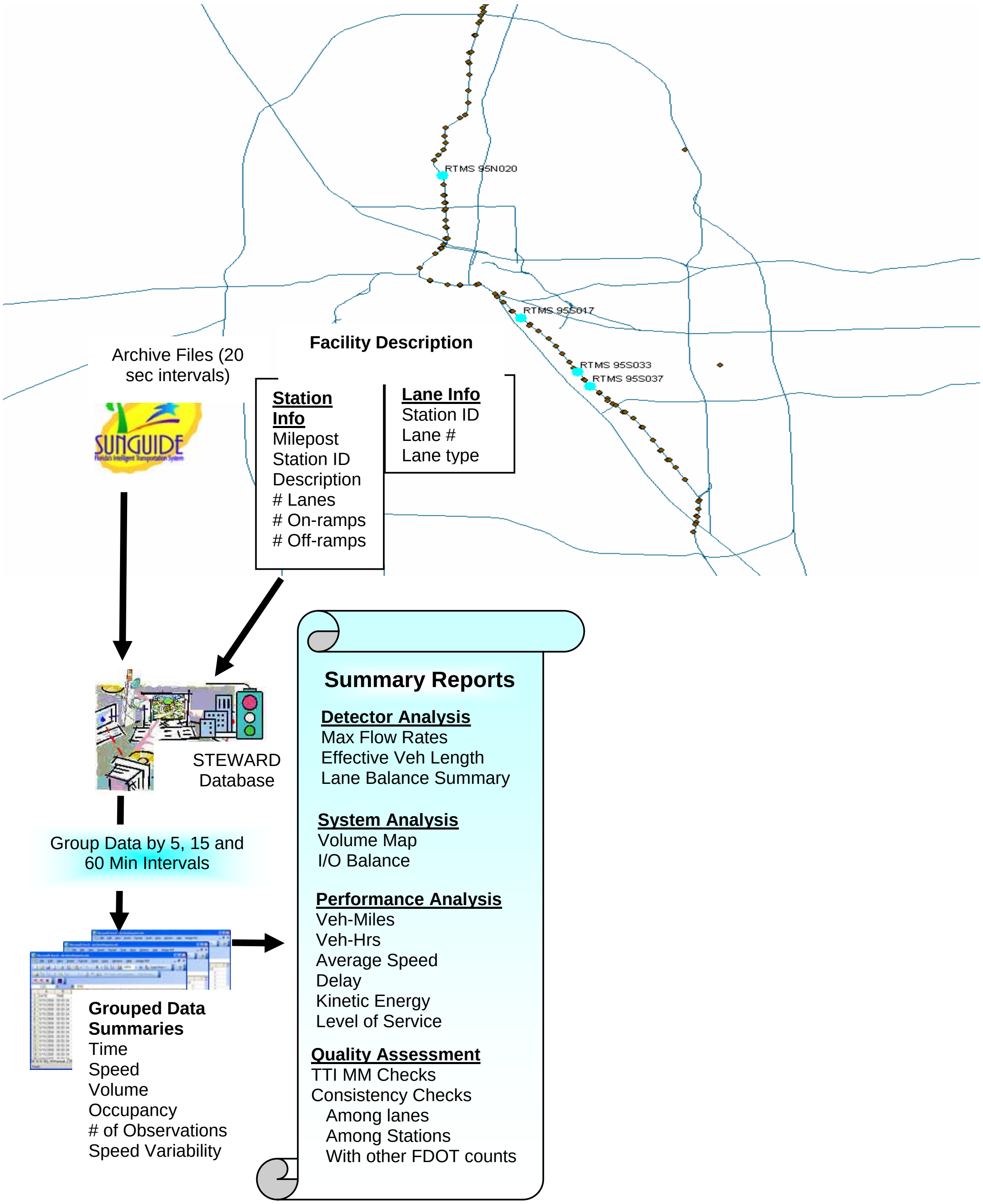


Figure A-1: SunGuide archived TSS data flow

FigureA-2: Example of the raw data from the SunGuide TSS archive

timestamp	detector_id	lane_id	speed	volume	occupancy
00.00.04	RTMS 95N003	R95N003_01Lane_01	55	1	1
00.00.09	RTMS 95S004A	R95S004A_01Lane_03	55	1	1
00.00.09	RTMS 95N006	R95N006_01Ramp_01	0	0	0
00.00.09	RTMS 95N026	R95N026_01LaneN_01	0	0	0
00.00.09	RTMS 95N026	R95N026_04LaneS_01	55	2	6

The data fields in the raw archive file are described as follows:

2.1 Timestamp

The time stamp, based on a 24 hour format "hh.mm.ss", indicates the time at which the record was received from the field hardware. This field is placed in the STEWARD database with global replacement of the period by a colon to give the more conventional time format of "hh:mm:ss". The separator character is user-specified in SunGuide. We will need to know what separator has been specified for each system so it can be converted to a colon.

2.2 Detector_id

The detector ID field identifies the RTMS detector station from which the data was received. The interpretation of this field is somewhat complicated by the fact that more than one RTMS unit may be required to cover all lanes at a given station and, at any station; one RTMS unit may cover lanes in one or both directions.

In District 2, the detector ID format is RTMS 95Fnnn[A]: where:

- "RTMS 95" is the designation used by District 2 for RTMS detector stations on the I-95 facility in Jacksonville. Note that all detector stations on this facility use RTMS units.
- F is a facility descriptor, with "N" representing stations north of I-10 and "S" representing stations south of I-10. Note that this designation does not imply a direction of travel (i.e., Northbound or Southbound).
- nnn is a three digit number identifying the station. The numbers run more or less sequentially in both directions from I-10.
- "A" is an optional identifier used to indicate a second RTMS unit at this station.

2.3 Lane_ID

The lane ID identifies the individual lane at the RTMS Station. The format is R95Fnnn[A]_cctttt[D]_LL, where:

- R95F identifies the detector station from the detector ID field. "R95N" in this field conforms to "RTMS 95N" in the detector ID and "R95S" conforms to "RTMS 95S."
- nnn has the same significance as in the detector ID field
- "A" also has the same significance as in the detector ID field.
- _cc indicates the physical RTMS channel assigned to this detector.
- tttt is either "Lane" or "Ramp" depending on the type of the lane detected. "Lane" refers to a freeway mainline lane. Note that no distinction is made between entrance and exit ramps.
- [D] is an optional indicator of the direction of the detected lane (either N or S in this case)

- _LL is the lane or ramp number, sequentially from left to right.

2.4 Data Fields

Separate data fields are provided for speed, volume and occupancy. These numbers represent the raw data values accumulated over the last 20 second period.

3 Conversion of TSS Data to the STEWARD Database

The STEWARD database will contain the TSS data from all days at all stations in all TMCs. As illustrated in Figure 2, the SunGuide raw data archive contains one record for each lane for each 20 second polling interval. The STEWARD database content differs from the raw data in two ways:

1. To streamline the processing and retrieval, the STEWARD data base contains one record per station representing all of the lanes at that station. Individual lane data are not archived but lane balance measures based on the ratio of the highest lane value to the lowest lane value are recorded for both volume and speed.
2. The values contained in the STEWARD database represent the 20 second archived data values accumulated over periods of 5, 15 and 60 minutes.

The STEWARD database structure is shown in Table A-1.

Table A-1: STEWARD Database Structure

Field	Description
DATE	Date of collection (MM/DD/YYYY)
TIME	Time Stamp (HH:MM:SS)
CDWID	Station ID (4 characters)
Direction	1 = increasing mileposts, 2 = decreasing mileposts
FwySpd	Volume weighted average speed for mainline
FwyVol	Total count for mainline
FwyOcc	Average occupancy for mainline (unweighted)
Spd_CV	Speed standard deviation / mean for all mainline lanes
VolRatio	Highest lane volume / lowest lane volume for mainline lanes
SpdRatio	Highest lane speed / lowest lane speed for mainline lanes
EntryVol	Sum of entrance ramp counts
ExitVol	Sum of exit ramp counts
FwyQA	Placeholder for quality assessment measure. Currently the percent of expected observations that were received during the time period.
EntryQA	
ExitQA	
HOVSpdl	Corresponding measures of freeway speed, volume occupancy and data quality in the HOV lanes, if present
HOVVol	
HOVOcc	
HOVQA	

Some conversion of the raw archive data will be required to create the STEWARD database. The following steps are involved in the conversion:

- A "date" field must be added to each record. The date information in the raw archive data is embedded in the file name.
- The "time" field must be converted to a universal format "hh:mm:ss."
- A unique station identifier must be constructed from the 'detectorID" field that was shown in Figure A-2. The station ID format is "DFnn" where:
    - D indicates the FDOT district
    - F indicates the facility within the district (0=North of I-10, 1=South of I-10)
    - nn is the station number within the facility assigned by the TMC
- A direction code is added (1 = increasing mileposts, 2 = decreasing mileposts)
- The speeds are averaged (volume weighted) for all mainline lanes over the period.
- The counts are summed for all mainline lanes over the period.
- The occupancies are averaged (unweighted) for all mainline lanes over the period.
- A field indicating the variability of speed within the time period is added. The speed variability could indicate turbulence in the traffic stream. The speed variability is represented by the coefficient of variation of speed (i.e., standard deviation/mean speed).
- A field indicating the variability of volume among lanes within the time period is added. The "volume ratio" is defined as the ratio of the highest lane volume / lowest lane volume for all mainline lanes.
- A field indicating the variability of speed among lanes within the time period is added. The "speed ratio" is defined as the ratio of the highest lane speed / lowest lane speed for all mainline lanes.
- Entrance and exit ramp volumes are added in separate fields. The speed and occupancy measures for ramps are not archived in the STEWARD database.
- Quality assessment fields are added for the mainline, entrance ramps and exit ramps in separate fields. The current entry in these fields is the percent of expected observations that were received during the time period. It is likely that the values in these fields will be modified to include the results of a more extensive quality assessment procedure that is now under development.

An example of the STEWARD database content is presented in Figure A-3

Figure A-3: Example of the STEWARD database content

Date	Time	CDWID	Dir	Fwy Spd	Fwy Vol	Fwy Occ	Spd CV	Vol Ratio	Spd Ratio	Entry Vol	Exit Vol	Fwy QA	Entry QA	Exit QA
2/21/2007	08:45:00	2123	1	61.3	679	3.8	7.36	2.3	1.13	0	0	100	0	0
2/21/2007	08:45:00	2124	2	61.5	1146	10	4.67	1.22	1.2	0	0	100	0	0
2/21/2007	08:45:00	2125	1	68.8	1469	8.6	3.82	1.26	1.07	0	0	100	0	0
2/21/2007	08:45:00	2126	2	65.8	1129	9.9	4.86	1.06	1.16	0	0	100	0	0
2/21/2007	08:45:00	2127	1	57.7	1080	9.2	5.64	1.19	1.28	0	0	100	0	0
2/21/2007	08:45:00	2128	2	66.8	985	7.4	8.38	1.29	1.02	0	155	100	0	100
2/21/2007	08:45:00	2129	1	71	897	2.6	5.71	1.44	1.47	166	0	100	100	0
2/21/2007	08:45:00	2130	1	59.2	1276	9.5	3.9	1.82	1.13	99	0	100	100	0
2/21/2007	08:45:00	2131	2	64.8	937	8.3	7.5	1.82	1.15	0	84	100	0	100

4 Required TSS Facility Data

Not all of the information required to convert the raw data to the STEWARD database is contained in the raw data. Some additional information is required for three purposes:

1. To ensure that each record in the STEWARD database represents a globally unique location
2. To support the analysis and reporting of system based measures and quality assessment.
3. To relate the measures obtained from a specific location to other forms of data, such as RCI, Statistics Office counts, crash records etc.

Two facility information databases must be created for each facility to be included in the STEWARD database. This information must be presented in two Excel spreadsheets:

4.1 The Station Data Spreadsheet

The station data spreadsheet must include the following fields for each station on the facility.

4.1.1 Station_Index

This is a number assigned sequentially to all stations in a facility. It is required for internal processing purposes and does not appear in the database or the reports.

4.1.2 Stationcdw_Num

This is the first 4 characters in the lane ID field in the STEWARD database, representing the district, facility and station number (Example: 2001).

4.1.3 Description

A physical description of the station (Example: I-95 NB at Forest St)

4.1.4 Status

This indicates the known status of the station (0=Normal, 1=Offline, 2=Undetected). The offline stations will not be reported as defective. The undetected station locations are required for the input/output analysis to indicate that the inputs and outputs for a specific link should not be expected to balance.

4.1.5 Road

The name given to the facility (Either I-95N or I-95S in District 2)

4.1.6 Latitude

Expressed in degrees and decimal degrees (Example 30.32339)

4.1.7 Longitude

Expressed in degrees and decimal degrees (Example -81.6807)

4.1.8 State_Milepost

Required for sequential ordering of stations (Example 351.451)

4.1.9 Roadway_Id

Required for correlation with the roadway characteristics inventory (RCI) and crash data and for identifying the county number for generating traffic count files (Example 72020000)

4.1.10 Roadway_Milepost

Required for correlation with RCI and crash data and to identify the county number for generating traffic counts (Example 2.7)

4.1.11 Max_Speed

The maximum speed at each station is required for estimating delay as a performance measure because the FDOT definition of delay is referenced to a threshold 10 mph below the speed limit

4.1.12 Num_Of_Lanes

The number of lanes at each station is required for estimating v/c ratio as a performance measure.

4.1.13 Lane_Capacity

The per-lane capacity at each station is required for estimating v/c ratio as a performance measure.

4.1.14 UpNode

The upstream station number for each station is required to determine the average volume in each link bounded by two stations

4.1.15 Detector_Type

The detector type (RTMS, Loop, etc.) is not currently used for processing, but it will be of interest to future studies for comparison of detector performance.

4.1.16 Detector_Unit

The number of detector units was included in the facility data originally supplied by District 2. It is included here as a future provision

4.2 The Lane Data Spreadsheet

The station data spreadsheet must include the following fields for each detected lane on the facility:

4.2.1 CDWstation

This is the same 4 digit station number as in the station data spreadsheet. It is used as a key to relate the station and lane data (Example 2001).

4.2.2 Lane

The lane number reference in the STEWARD database (Example: 2001131)
The compositors of the lane number are

- CDW Station number (4 characters)
- Direction (1 character)
- Function code (1 character)

1. Left entrance ramp
2. Left exit ramp
3. Freeway main lane
4. Right entrance ramp
5. Right exit ramp
6. Auxiliary lane
7. HOV Lane

- Lane Number (from left to right)

4.2.3 Tmc_Id

The lane ID used by the archive file generated by the TMC (Example: R95N001_01Lane_01). Note that this must match the "lane_id" field in the archive database. Archive data records in which the lane_id is not found in the lane data spreadsheet are reported as "Orphan Lanes." Records in the lane data spreadsheet that have no matches in the archive data file are reported as "Null Lanes."

4.2.4 Det_Type

Always "RTMS" for District 2. This is not used in any analysis at present, but is provided for future use.

4.2.5 Direction

The direction of the traffic detected on this lane (1=Increasing mileposts, 2=Decreasing mileposts)

4.2.6 Status

This indicates the known status of the station (0=Normal, 1=Offline, 2=Undetected). The offline stations will not be reported as defective. The undetected station locations are required for the input/output analysis to indicate that the inputs and outputs for a specific link should not be expected to balance.

4.2.7 Roadway_Id

Required for correlation with RCI and crash data (Example 72020000). Also required to obtain the county number for generating traffic count files compatible with the FDOT Statistics Office files.

4.2.8 Roadway_Milepost

Required for correlation with RCI and crash data. Note that these fields are also in the station data file. They are required here because stations that detect traffic in both directions may have different roadways assigned.

4.2.9 Max_Speed

Normally the speed limit

4.2.10 Count_Station

The number assigned by the FDOT Statistics Office or District Planning Office for generating traffic count data files from the SunGuide detectors.

5 TVT Archived Data Format

The TVT archive data are much simpler to acquire, store and analyze than the TSS data. The raw archive data format is shown in Figure A-4.

Figure A-4: Example of the raw SunGuide archive TVT data content

timestamp	travel_link_id	travel_link_timestamp	travel_time
00.00.55	Arpt - MLK	2006.08.31 00.00.53	14
00.00.55	Arpt - Philips	2006.08.31 00.00.53	38
00.00.55	S04 - I-10	2006.08.31 00.00.53	8
00.00.55	S04 - 8th	2006.08.31 00.00.53	12
00.00.55	N07 - MLK	2006.08.31 00.00.53	9
00.00.55	N07 - 10	2006.08.31 00.00.53	13
00.00.55	S08 - Bowden	2006.08.31 00.00.53	6
00.00.55	Arpt - 10	2006.08.31 00.00.53	18

The raw archive data fields are described as follows:

5.1 Timestamp

Same as TSS Data

5.2 Travel_link_id

This is the name given by the TMC to the travel link

5.3 Travel_link_timestamp

We do not use this field. It was noted that some of the records at the beginning of a day were from the previous day. This caused irresolvable confusion.

5.4 Travel_time

This is the travel time in minutes on the link.

6 Conversion of TVT Data to the STEWARD Database

The conversion of the TVT data is a simpler process than the TSS data. The TVT facility data are contained in three tables:

- The link data table, which identifies the required properties for each link. A sample of the link data table is presented in Figure A-5.
- The origin data table which identifies the required properties for each origin point. One origin point may be common to several links. A sample of the origin data table is presented in Figure A-6.
- The destination data table which identifies the required properties for each destination point. One destination point may be common to several links. A sample of the origin data table is presented in Figure A-7.

Archive_Ref	OriginID	DestinationID	Status	Direction	Speed_Limit	Link_Length
95n01- Dunn	201101	201204	0	NB	58.4	7.733
95n01- MLK	201101	201202	0	NB	55.0	1.733
95n02- emerson	202101	212202	0	SB	54.4	5.267
95n02- JTB	202101	212203	0	SB	58.4	8.267
95n03- Edgewood	201102	201203	0	NB	55.0	4.127
95n03- I-295	201102	201205	0	NB	60.5	9.127
95n04- Dunn	201103	201204	0	NB	59.5	5.428
95n04- I-295	201103	201205	0	NB	61.7	7.428
95n05- emerson	202102	212202	0	SB	54.6	7.559
95n05- I-10	202102	212201	0	SB	51.3	3.559

Figure A-5. Example of a TVT link data table

The link data fields are described as follows:

6.1 Archive_Ref

This field is used to search for the travel time link and therefore must match an entry in the archive data travel_link_id field exactly.

6.2 From

A description of the origin for the travel time link

6.3 To

A description of the destination for the travel time link

6.4 Status

0=Normal, 1=Disabled. Other status codes may be desirable.

6.5 Direction

NB or SB. All links are unidirectional.

6.6 Speed Limit

The distance-weighted average speed limit for this link is computed automatically by spreadsheet macro routine that obtains speed limit information from the TSS facility data. The speed limit is required to compute the travel time variability measures.

6.7 Link Length

The length of this link is computed automatically by spreadsheet macro routine that obtains milepost information from the TSS facility data. The speed limit is required to compute the travel time variability measures.

7 The STEWARD TVT Database

An example of the STEWARD travel time database content is shown in Figure A-8. The same database structure is used for the original archive data at 1 minute intervals, plus the 5, 15 and 60 minute aggregations.

Three performance measures derived from the travel times are contained in the database:

- *Congestion Delay:* based on a travel time index of 1.5. The travel time index is defined as the ratio of the actual travel time to the travel time at the free flow speed. The speed limit will be used to represent the free flow speed. The unit of measurement is accumulated minutes of delay.

- *On Time Delay:* referenced to a travel speed of 10 mph below the speed limit. This threshold has been specified for purposes of travel time reliability reporting in Florida. The unit of measurement is accumulated minutes of delay.

- *Percent of on-time trips*: defined as the percent of trips that were made at a speed no less than ten mph below the speed limit.

Most of the other fields are self-explanatory. The fields that require explanation apply to the 5, 15 and 60 minute aggregations. These fields are described as follows:

- *AV_TT*: The average travel time over the period
- *MAX_TT*: The maximum travel time in any 1 minute interval during the period
- *CHECK:* A placeholder for future quality assessment checks

OriginID	Archive_Ref	Sign Name DMS	Location Description	Roadway	Direction	Latitude	Longitude	Milepost	Station_ID
202104	Airport01	Airport01	Airport Road	I-95N	2	30.48213889	-81.6485	363.149	2050
201104	95N06	DMS 95N06	DMS	I-95N	1	30.44197222	-81.65422222	360.365	2037
202103	95N07	DMS 95N07	I-95 North of Dunn Ave SB	I-95N	2	30.44202778	-81.65463889	360.363	2036
201103	95N04	DMS 95N04	DMS	I-95N	1	30.36458333	-81.66861111	354.572	2017
202102	95N05	DMS 95N05	DMS	I-95N	2	30.36438889	-81.66897222	354.559	2018
201102	95N03	DMS 95N03	DMS	I-95N	1	30.34002778	-81.66786111	352.873	2009
201101	95N01	DMS 95N01	DMS	I-95N	1	30.33238889	-81.67225	352.267	2006
202101	95N02	DMS 95N02	DMS	I-95N	2	30.33302778	-81.67291667	352.267	2005
212101	95S01	DMS 95S01	DMS	I-95S	2	30.31388889	-81.66022222	349.812	2105
211102	95S02	DMS 95S02	DMS	I-95S	1	30.29713889	-81.63644444	347.825	2117
212103	95S03	DMS 95S03	DMS	I-95S	2	30.29555556	-81.63527778	347.698	2116

Figure A-6: Example of a TVT origin data table

DestinationID	Archive_Ref	Location Description	Roadway	To_Dir	Latitude	Longitude	Milepost	Station_ID
201206	Airport Rd	Airport Rd - Exit 363	I-95N	1	30.47922094	-81.64403147	363	2048
201205	I-295	I-295 (North) - Exit 362	I-95N	1	30.46505593	-81.6477458	362	2043
201204	Dunn	Dunn Ave - Exit 360	I-95N	1	30.43679506	-81.65553113	360	2037
201203	Edgewood	Edgewood Ave - Exit 357	I-95N	1	30.39719141	-81.66790056	357	2028
201202	MLK	MLK - Exit 354	I-95N	1	30.35632188	-81.6685168	354	2014
202201	MLK	MLK - Exit 354	I-95N	2	30.35632188	-81.6685168	354	2016
201201	Union St	Union St - Exit 353	I-95N	1	30.34187445	-81.66805376	353	2009
211201	I-10	I-10 - Exit 351	I-95S	1	30.31803344	-81.67920523	351	2102
212201	I-10	I-10 - Exit 351	I-95S	2	30.31803344	-81.67920523	351	2101
211202	Downtown	Downtown/Acosta - Exit 350	I-95S	1	30.31405856	-81.66334964	350	2106
211203	Emerson	Emerson St - Exit 347	I-95S	1	30.28829202	-81.62710929	347	2118
212202	Emerson	Emerson St - Exit 347	I-95S	2	30.28829202	-81.62710929	347	2120
211204	University	University Blvd - Exit 346	I-95S	1	30.27663549	-81.61718891	346	2127

Figure A-7: Example of a TVT origin data table

DATE	TIME	FROM	TO	AV_TT	CONGESTION DELAY	MAX_TT	CHECK	ONTIME PERCENT	ONTIME DELAY
2/16/2007	17:00:00	211106	211203	4.3	0	5	15	100	0
2/16/2007	17:00:00	211106	211204	2	0	2	15	100	0
2/16/2007	17:00:00	212107	212206	2	0	2	15	100	0
2/16/2007	17:00:00	211108	211203	6	0	6	15	100	0
2/16/2007	17:00:00	211108	211205	2	0	2	15	100	0
2/16/2007	17:00:00	202104	212201	11.9	0	13	15	100	0
2/16/2007	17:15:00	201101	201204	8.2	0	9	15	100	0
2/16/2007	17:15:00	201101	201202	2	0	2	15	100	0
2/16/2007	17:15:00	202101	212202	7.1	0.3	9	15	73.3	4.6
2/16/2007	17:15:00	202101	212203	12.3	2.8	14	15	0	30.5
2/16/2007	17:15:00	201102	201203	5.7	0	6	15	33.3	5
2/16/2007	17:15:00	201102	201205	9.4	0	10	15	100	0
2/16/2007	17:15:00	201103	201204	3.9	0	4	15	100	0
2/16/2007	17:15:00	201103	201205	6.3	0	7	15	100	0
2/16/2007	17:15:00	202102	212202	10.7	0	12	15	40	10.2
2/16/2007	17:15:00	202102	212201	4.5	0	6	15	93.3	0.8
2/16/2007	17:15:00	201104	201206	2.8	0	3	15	20	4.8

Figure A-8: Sample STEWARD travel time data base content

8 Conversion of SunGuide Archive Data to the FDOT Count Formats

The traffic counts in the SunGuide archives have essentially the same content as the FDOT Statistics Office and District Planning Office traffic count files. There is clearly a potential benefit that could be derived from a mutual exchange of traffic count data between the ITS centers and the Statistics Office. The Statistics Office data could provide an important reference for calibrating the ITS detectors, most of which are microwave based. The ITS data could provide a useful supplement to the statewide traffic count coverage now in place.

Figure A-9 shows an example comparison between the daily counts from a Statistics Office station and the upstream and downstream ITS stations. Note that a near perfect agreement is apparent here. This will not always be the case and comparison of data from the two sources could potentially improve the accuracy of both sources.

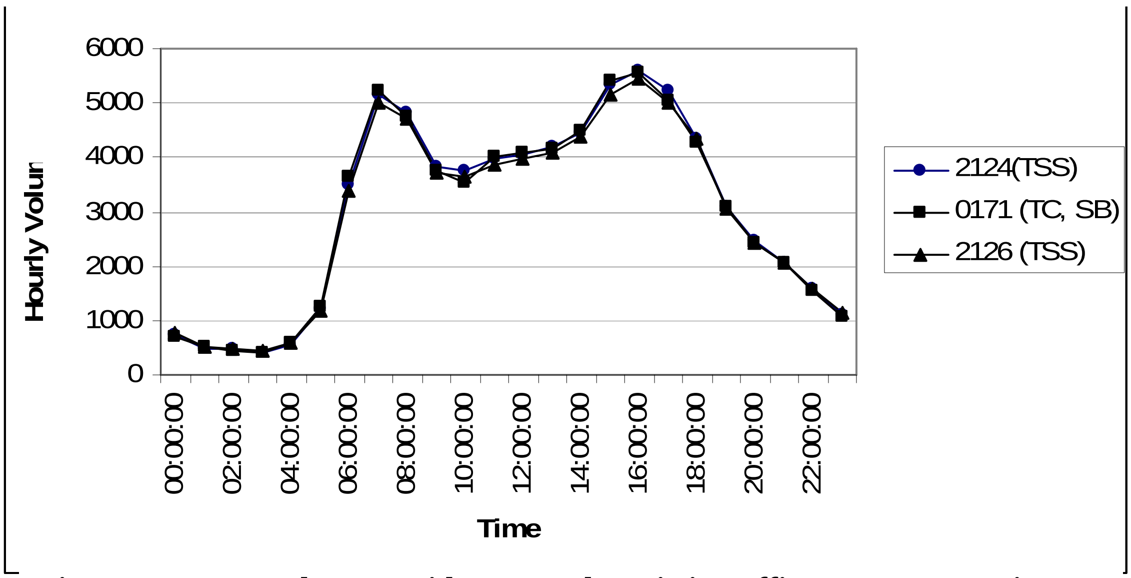


Figure A-9. Example SunGuide TSS and Statistics Office count comparison

8.1 The Statistics Office Count Format

A utility program has been developed to convert the count data in the SunGuide Data Archive to the Statistics Office count format. This format is presented in Table A-2, along with the rules used for conversion from the SunGuide Archive. Note that each filed in this format has a fixed width.

Table A-2: Rules for Converting District 2 ITS Data to FDOT Statistics Office Count Files

TAG	CNT	String constant
CO	County Number	Duval County = 72.
STAT	Station Number	Four characters are available. The ITS station ID is made up as dfnn, where • d = District (1 char) • f = Facility: 0=North of I-10, 1=South of I-10 (1 char) • nn = Assigned station number (2 char) Since there are only 4 characters available, and a given station can have lanes in both directions, we reconfigured the station number as fnnd, where: • f = Facility • nn = Station Number • d = direction (1=increasing MP, 2= Decreasing MP) The redundant district number was eliminated to make room for the direction
	Filler	
YY	Survey Year	
MM	Survey Month	
DD	Survey Day	
HR	Survey Hour	End of reporting interval
MN	Survey Minute	End of reporting interval
	Filler	
INT	Survey Interval	60 minutes
	The following values are repeated for each lane up to a total of 8 lanes	
	Filler	
L	Lane Number	TMG standard: right to left. Note that this is inverted from the normal ITS standard
	Filler	
IVOL	Interval Volume	As computed

The Statistics office has also prescribed a standard for counts derived from ITS detectors. The ITS format uses the "ITS" tag in place of the "CNT" tag. This format also includes a validity check for each lane (N = normal, B = bad). As a preliminary criterion for generating ITS files, the detector data are assumed to be normal if the detectors reported volumes for at least 90% of the 20 second SunGuide polling intervals. More detailed quality assurance (QA) criteria are under development.

8.2 District Planning Office Count Format

The district planning offices perform non-continuous counts periodically at selected locations using various types of portable equipment. The SunGuide archive offers a potentially rich source for this type of data. The district count file format differs from the central office format because of differences in the processing software. The district format requires one record for each lane in the following format:

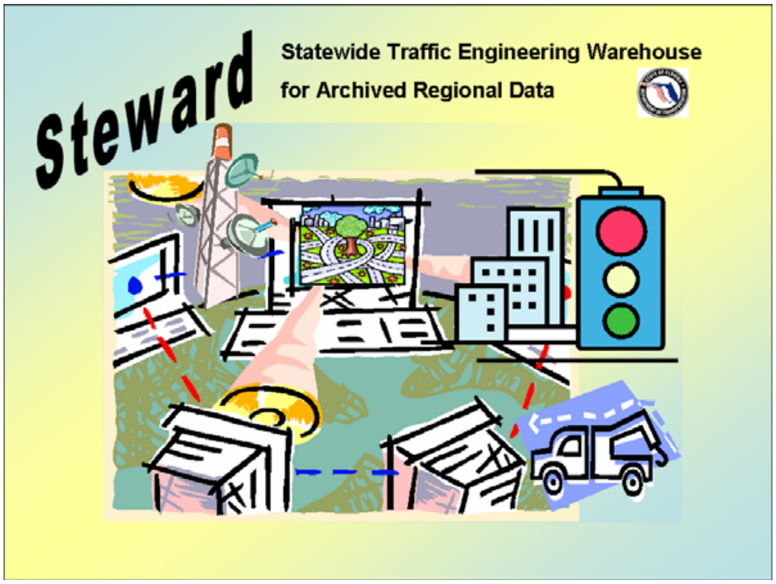
Field	Length	Comment	Example
County	2 Numeric	FDOT County Number	72
Station	4 Numeric	Supplied by District	0004
Date	8 Numeric	YYYYMMDD	20020115
Direction	1 Alpha	Direction of traffic	E
Lane	1 Numeric	1=Outside,2=Middle,etc.	1
Survey Type	1 Numeric	2=Vehicle count	2
Survey Program	1 Numeric	9=ITS program	9
Interval	2 Numeric	Minutes	15
Time	4 Numeric	Military time, leading zeroes	0830
Volume	6 Numeric	leading zeroes required	000357

Note: Blanks between Fields, Carriage Return/Line Field at end of each line of data.

The TSS lane data described earlier contains a field for the count data station. The station number begins with a “C” or a “D”, followed by a 4 digit number. Stations beginning with “C” are processed in the central office format and those beginning with D are processed as district count files. Count files are generated for all lanes with valid data in the station number field. Counts are not generated if the field is blank. The prefix letter is stripped from the station number in the count data file.

9 Transferring Archive Data to STEWARD

At this point we receive periodic data from District 2 through an FTP site. We will have to examine the possible transmission methods with each district to determine what will work best. Probably the first transmission of data would be easiest by “snail mail” CD.



Development of a Central Data Warehouse for Statewide ITS and Transportation Data in Florida

Phase II: Proof of Concept Final Report

Appendix B
Instructions and Guidance for Installing STEWARD at TERL

Prepared for the Florida Department of Transportation
By the University of Florida Transportation Research Center

Contract # BD545, RPWO # 37
UF Project 051449
Revision Date: April 15, 2008

Table of Contents

List of Figures

1 Introduction

A description of the steps required to install the Statewide Traffic Engineering Warehouse for Regionally Archived Data (STEWARD) software and databases is presented in this document as an appendix to the STEWARD Final Report. This material has been developed to assist FDOT and ITS contract personnel in setting up STEWARD in the Traffic Engineering Research Laboratory (TERL) facility. The topics include Oracle database program installation, STEWARD deployment and the STEWARD web site installation. A working knowledge of the Oracle data base manager and internet site management is assumed in the discussion that follows.

2 Oracle Database Program Installation

Oracle 10gR2 and the Oracle Workflow Server 2.6.4 are required for STEWARD database installation. The Oracle Workflow Server is included in the Oracle Database 10g Companion CD. The installation steps are as follows:

2.1 Install Oracle 10*g* Release 2

Oracle 10*g* Release 2 (10.2) for Microsoft Windows Enterprise edition (32-Bit) is installed for STEWARD.

2.1.1 Run the setup.exe file

This is a standard Windows procedure

2.1.2 Select Advanced Installation

As shown in Figure B-1, select *Advanced Installation* for the installation method.

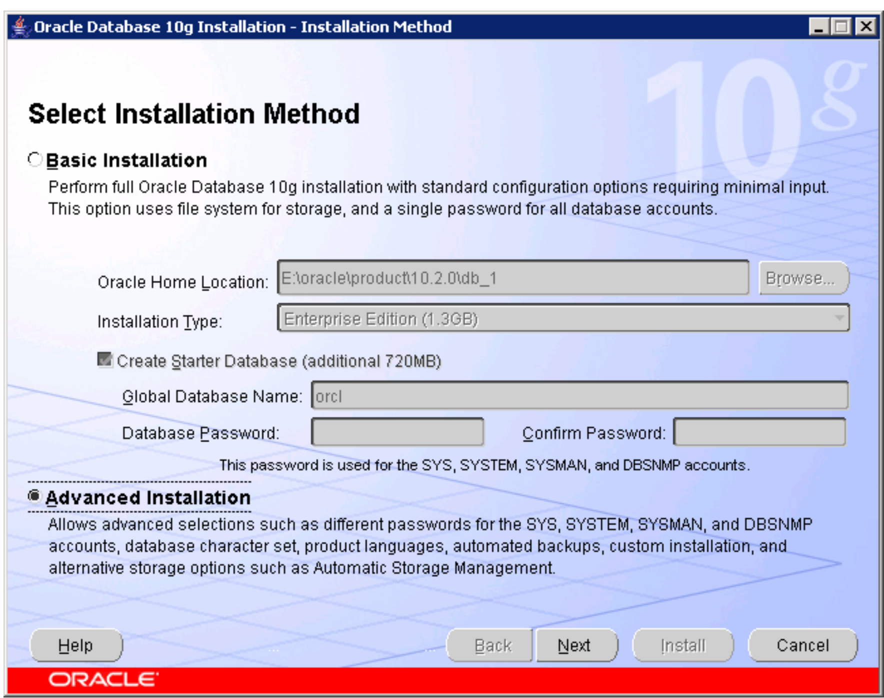


Figure B- 1: Oracle DB installation- Installation method

2.1.3 Select Enterprise Edition

As shown in Figure B-2, select *Enterprise edition* for the installation type.

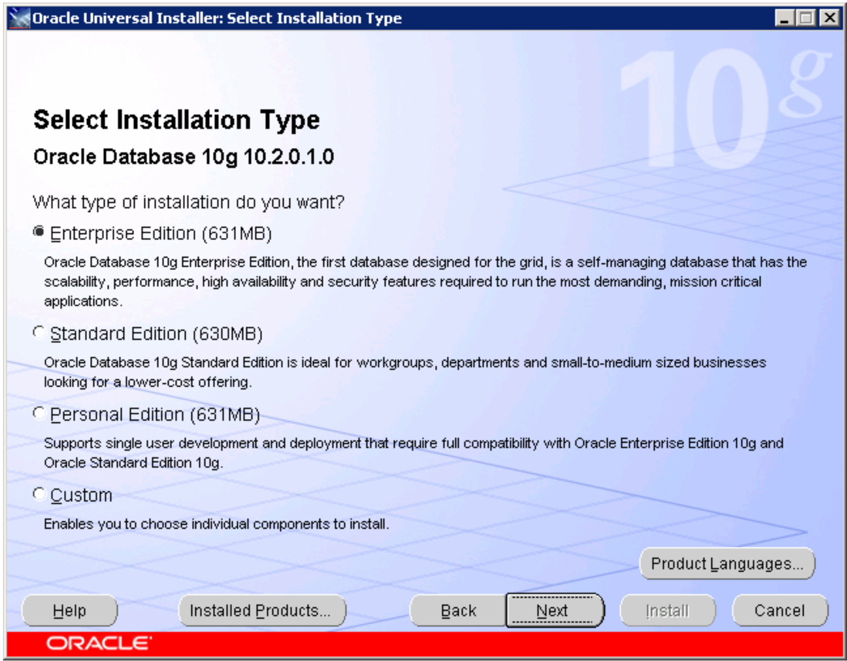


Figure B- 2: Oracle DB installation- Installation type

2.1.4 Select default location

As shown in Figure B-3, select *Home name and path*.

Select name as OraDb10g_home1 and path as C:\oracle\product\10.2.0\db_1

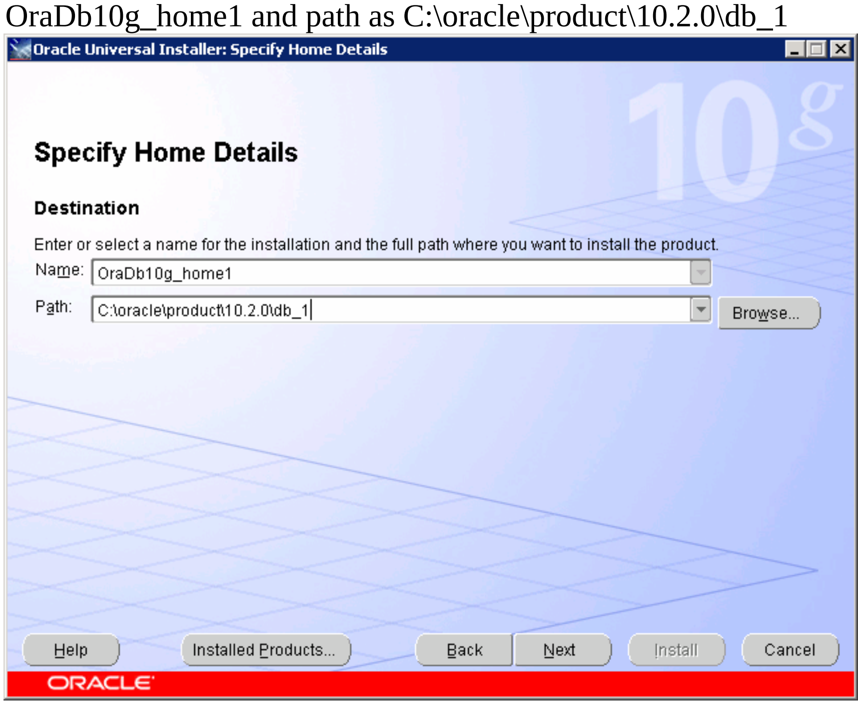


Figure B- 3: Oracle DB installation- Installation home details

2.1.5 Run the prerequisite checks

As shown in Figure B-4, prerequisite will be checked.

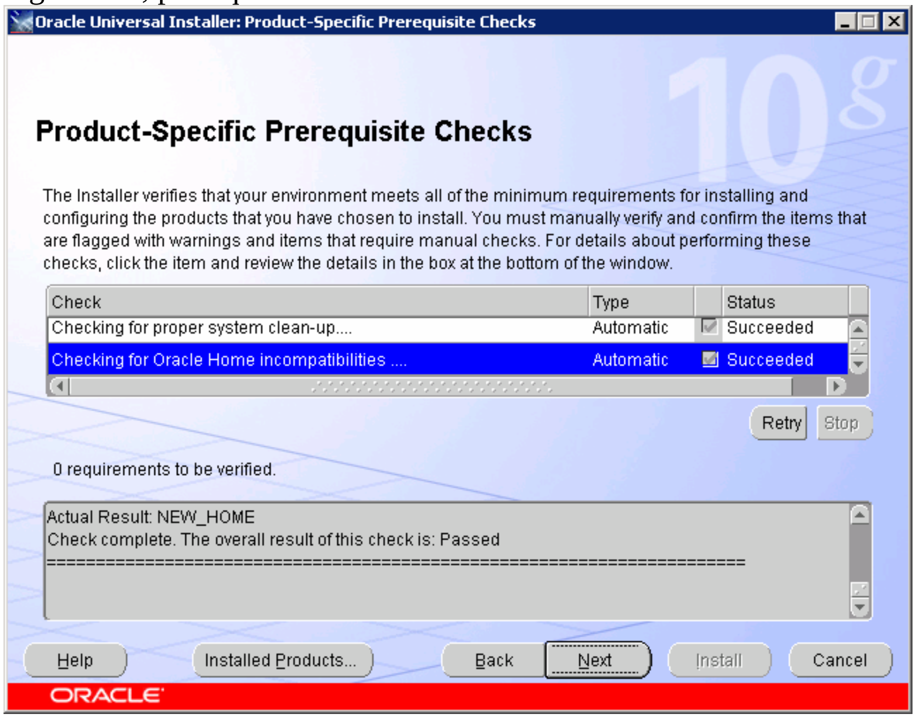


Figure B- 4: Oracle DB installation- Prerequisite checks

2.1.6 Select Database Software only option for configuration option

As shown in Figure B-5, select *Database Software only* option.

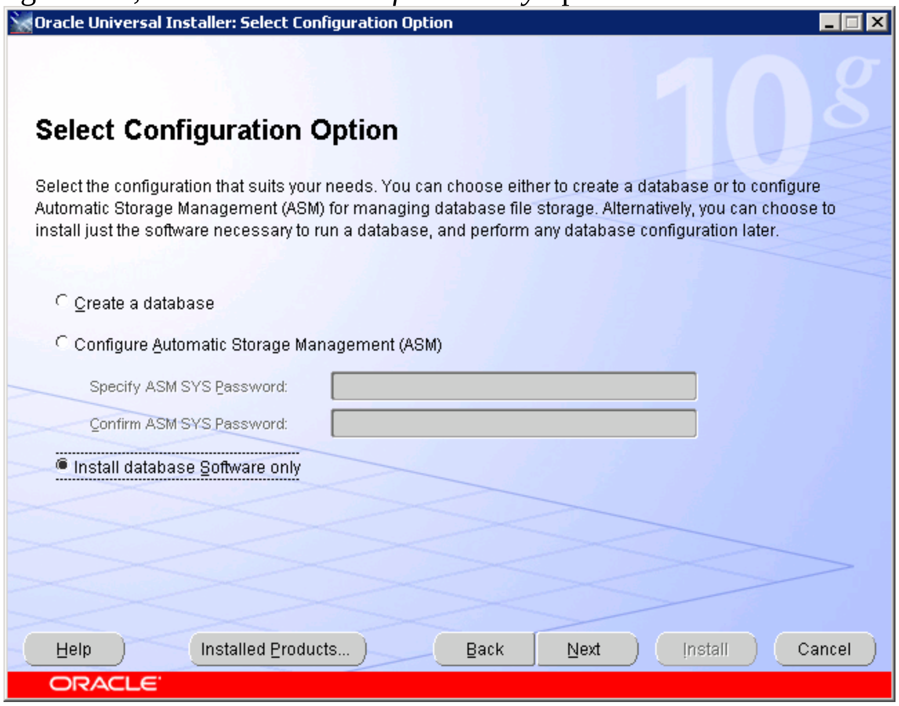


Figure B- 5: Oracle DB installation- Configuration option

2.1.7 Oracle Database installation

As shown in Figure B-6, Oracle Database installation summary is displayed before installation.

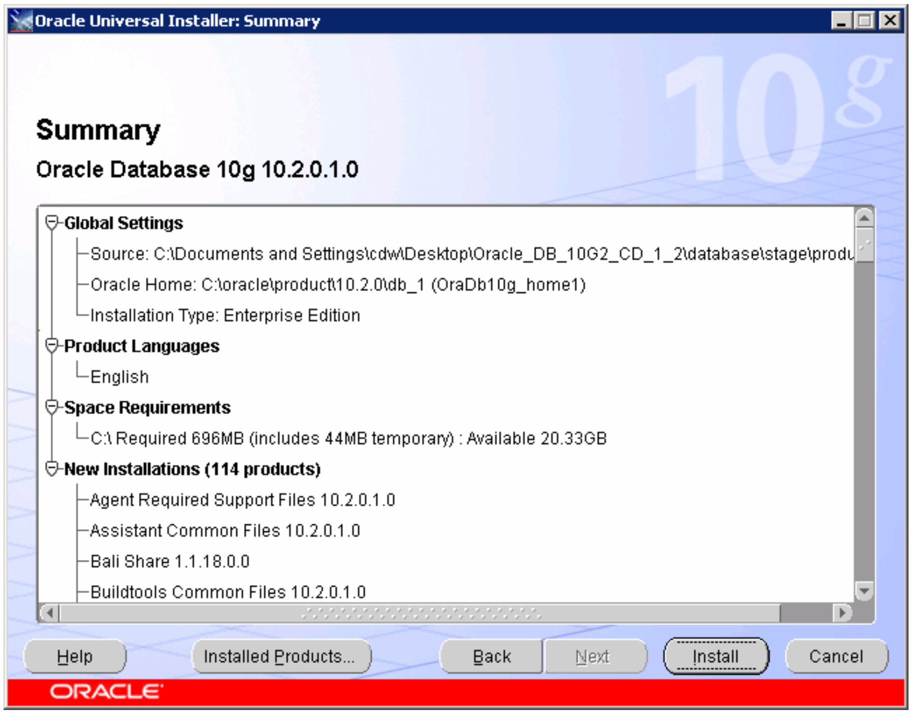


Figure B- 6: Oracle DB installation- Installation Summary

As shown in Figure B-7, Oracle Database is installed.

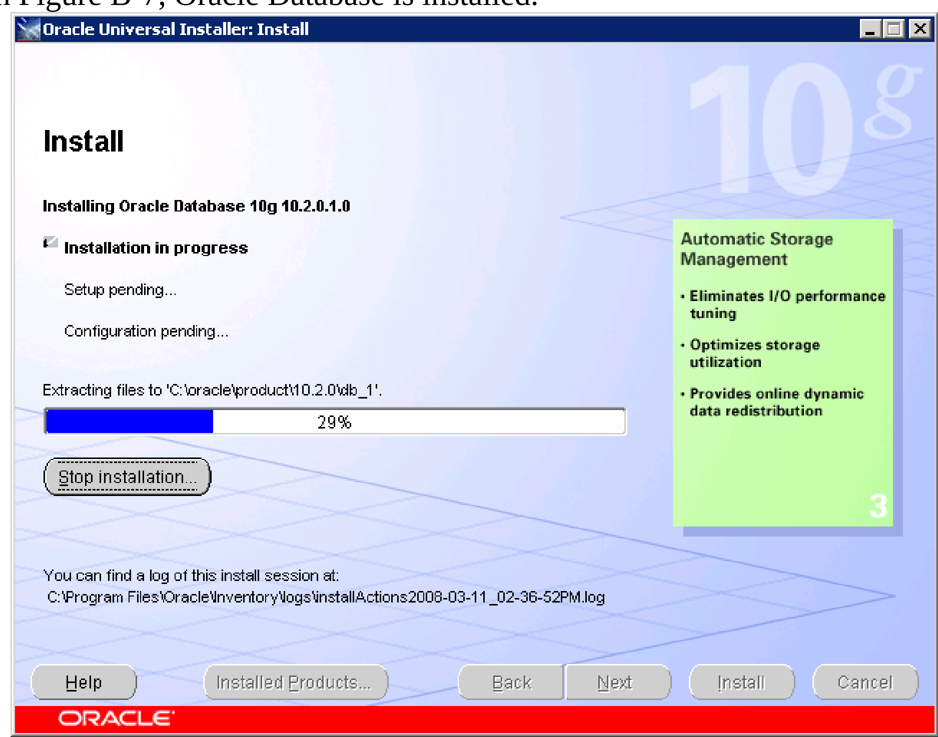


Figure B- 7: Oracle DB installation- Installation

As shown in Figure B-8, Oracle Database installation is done.

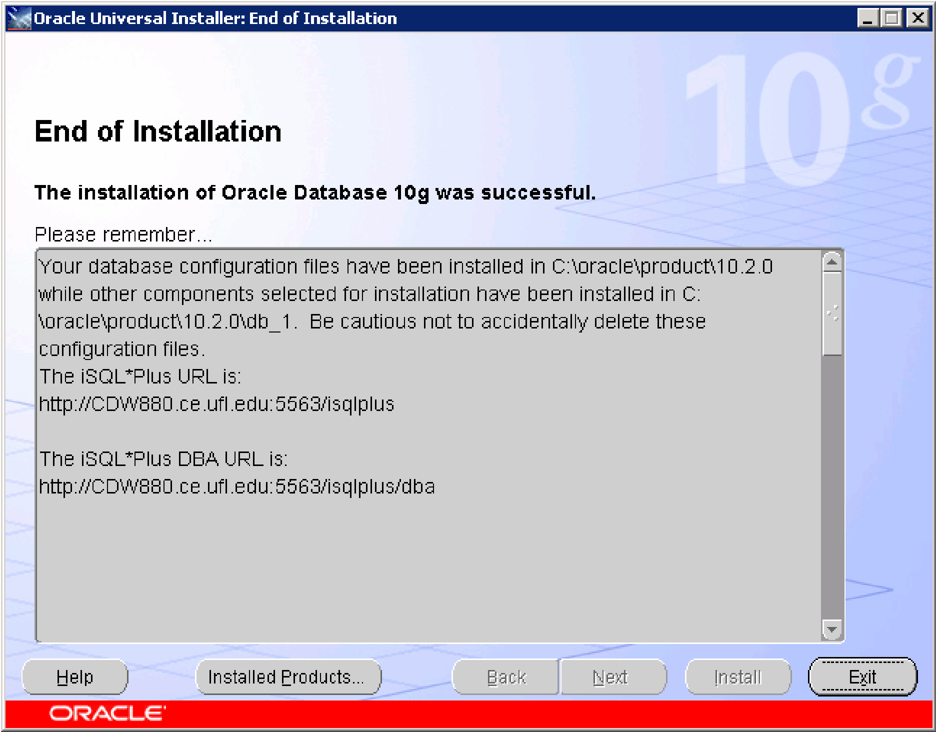


Figure B- 8: Oracle DB installation- End of Installation

2.2 Install Oracle Workflow

Install the Oracle Workflow from the Companion CD for Oracle 10*g* Release 2 (10.2) for Microsoft Windows Enterprise edition (32-Bit)

2.2.1 Run the setup.exe file

As shown in Figure B-9, Oracle Workflow is installed with Oracle universal installer.

Figure B- 9: Oracle Workflow installation

2.2.2 Select the Oracle database 10g Product

As shown in Figure B-10, select Oracle database 10g Product for Oracle Workflow installation.

Figure B- 10: Oracle Workflow installation - Select a product to install

2.2.3 Select Home name and address as follows:

As shown in Figure B-10, select *Home name and path*.
Select name as OraDb10g_home1 and path as C:\oracle\product\10.2.0\db_1

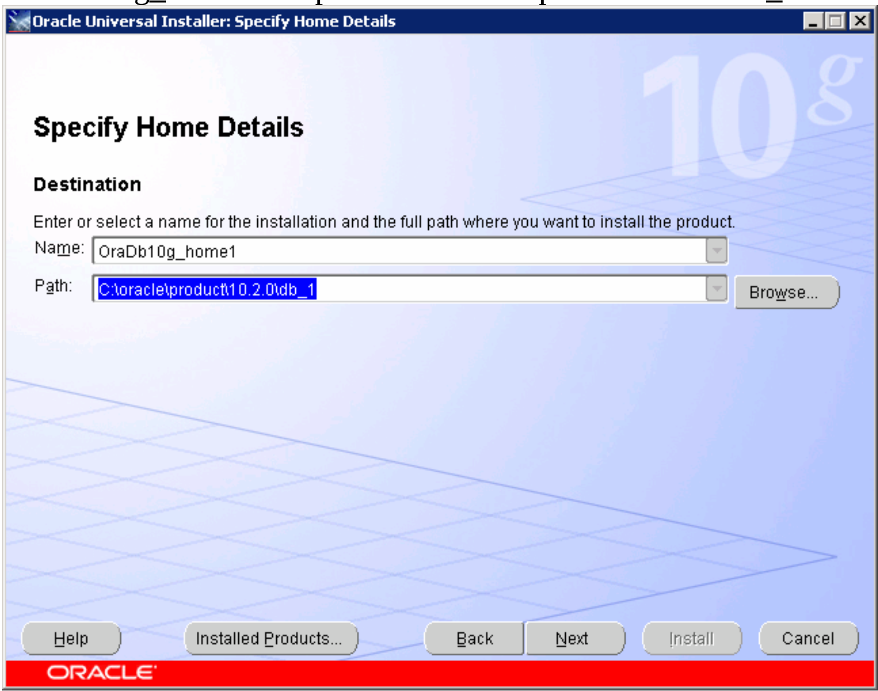


Figure B- 11: Oracle Workflow installation- Installation home details

2.2.4 Run the prerequisite checks

As shown in Figure B-12, prerequisite will be checked.

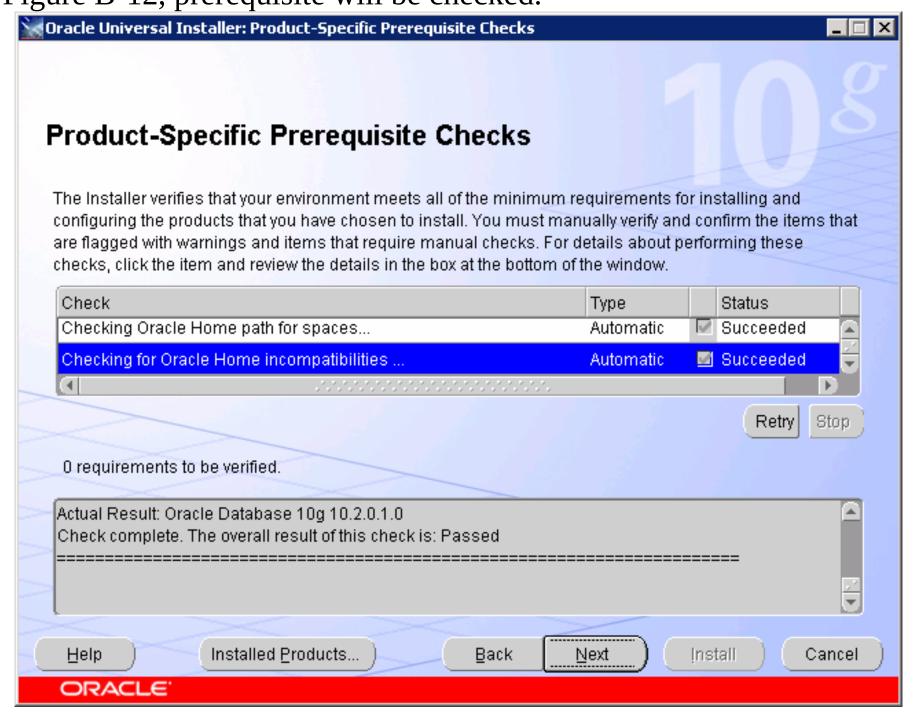


Figure B- 12: Oracle Workflow installation- Prerequisite checks

2.2.5 Oracle Workflow installation

As shown in Figure B-13, Oracle Workflow installation summary is displayed before installation.

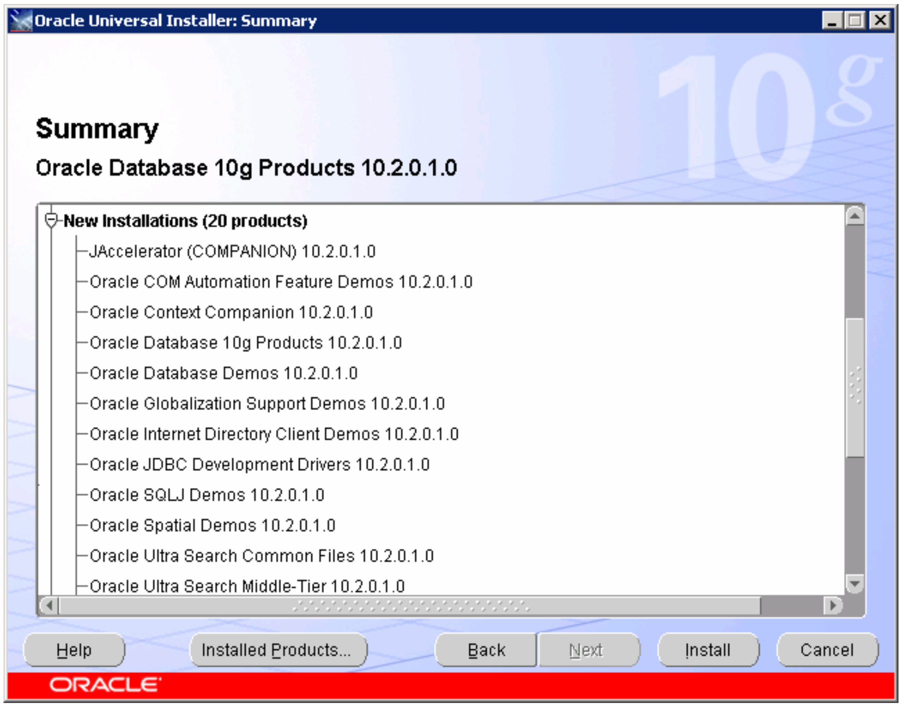


Figure B- 13: Oracle Workflow installation- Installation Summary

2.2.6 Oracle Workflow installation

As shown in Figure B-8, Oracle Database installation is done.

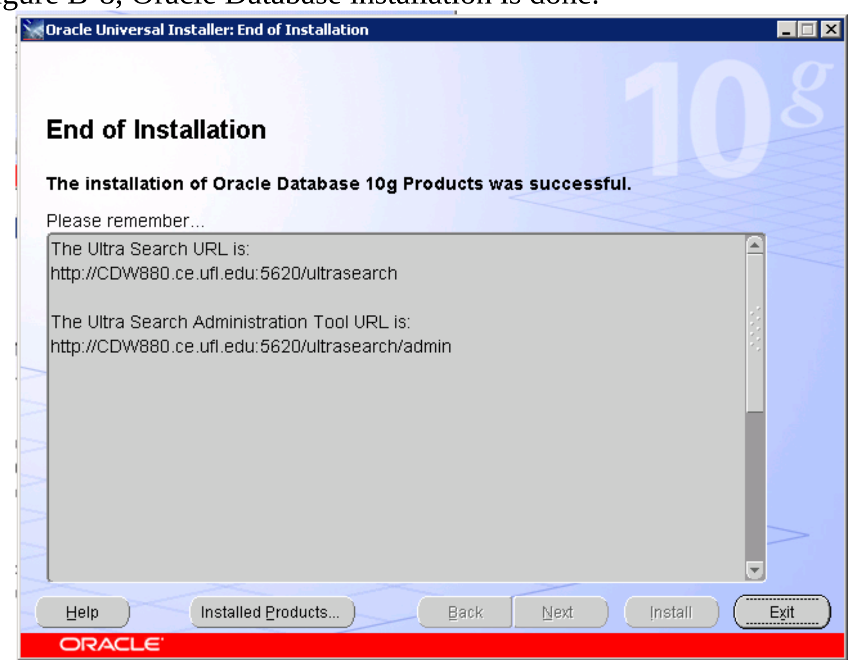


Figure B- 14: Oracle Workflow installation- End of Installation

2.3 Oracle Warehouse Builder Setup

The Design Center in the Oracle Warehouse Builder need be configured during the installation. Design center is executable as shown in Figure B-15.

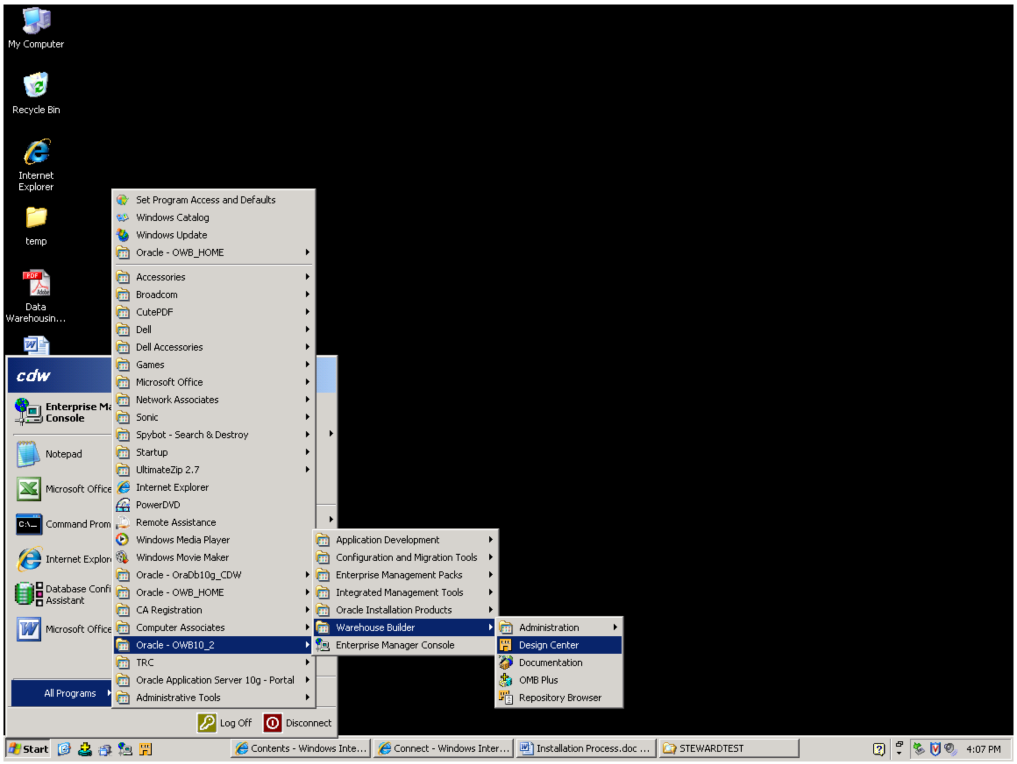


Figure B- 15: Oracle warehouse builder configuration

2.3.1 Create the Design Center User and its Repository

Run the ***Design Center*** and create a "Warehouse Builder User" and "Warehouse Builder Repository". They will be used to archive information for the new data warehouse. As shown in Figure B-16, click ***Get Started.*** If the ***Get Started*** button is not visible, click ***Show Details***.

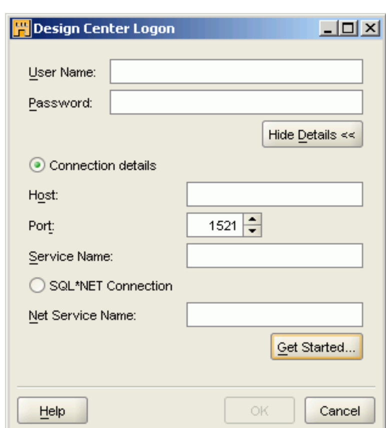


Figure B- 16: Oracle warehouse builder login window

As shown in Figure B-17, choose ***Basic Install*** in Install Type window.

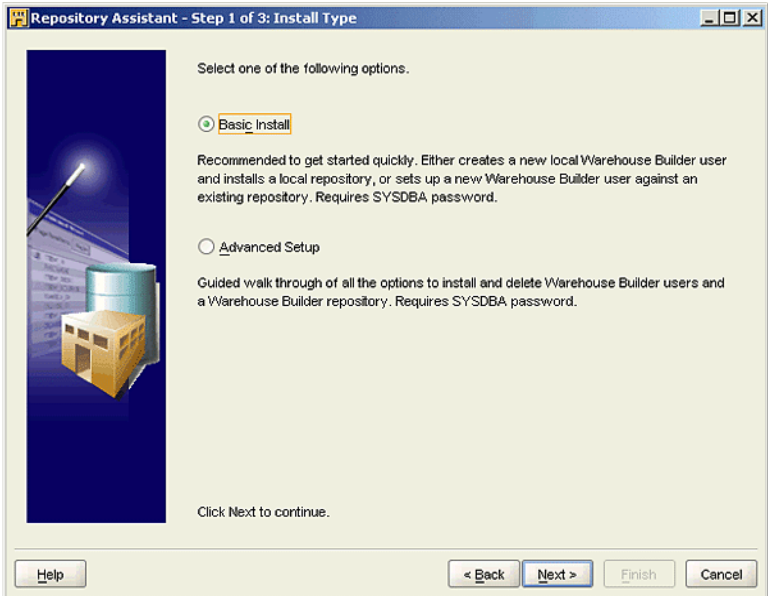


Figure B- 17: Oracle warehouse builder- Installation type

As shown in Figure B-18, create a user named steward_user and other items as follows in the Repository User and Connection Information window:

Repository User Name: **steward_user**
Repository User Password: **trc513**

SYSDBA User Name: **SYS**
SYSDBA Password: **trc513**

Host Name: **LOCALHOST**
Port Number: **1521**
Oracle Service Name: **steward2**

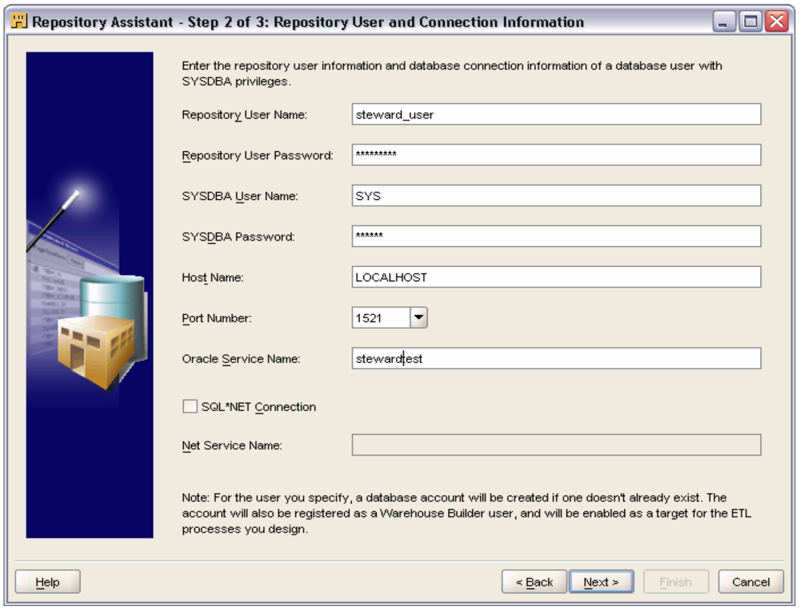


Figure B- 18: Oracle warehouse builder- Repository User and Connection Information

As shown in Figure B-19, re-enter the password for steward_user in the Password Confirmation window.

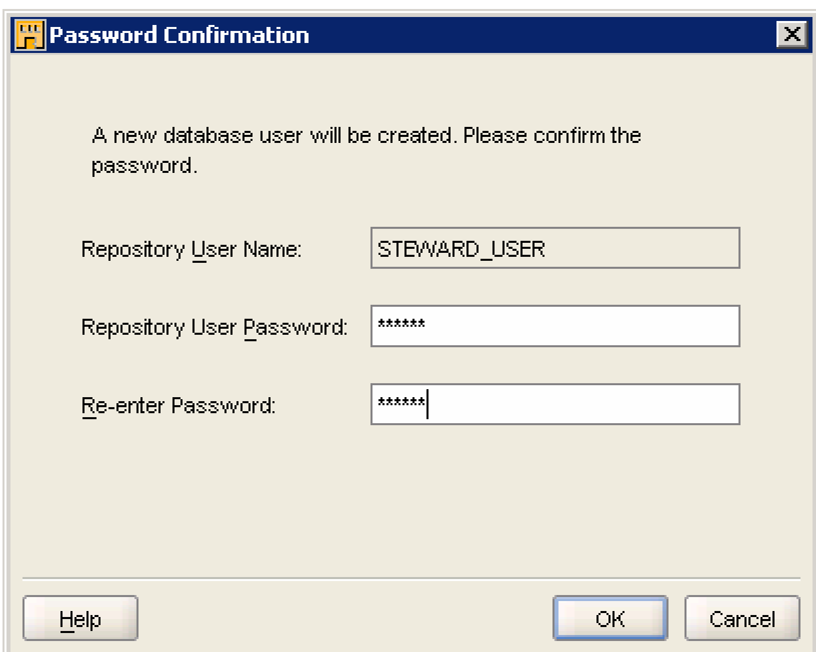


Figure B- 19: Oracle warehouse builder- password confirmation

As shown in Figure B-20, type the owner name Repository Owner Information window. The repository owner is a highly privileged Warehouse Builder user with access to additional security features. Enter “steward_owner” as the username and password (trc513).

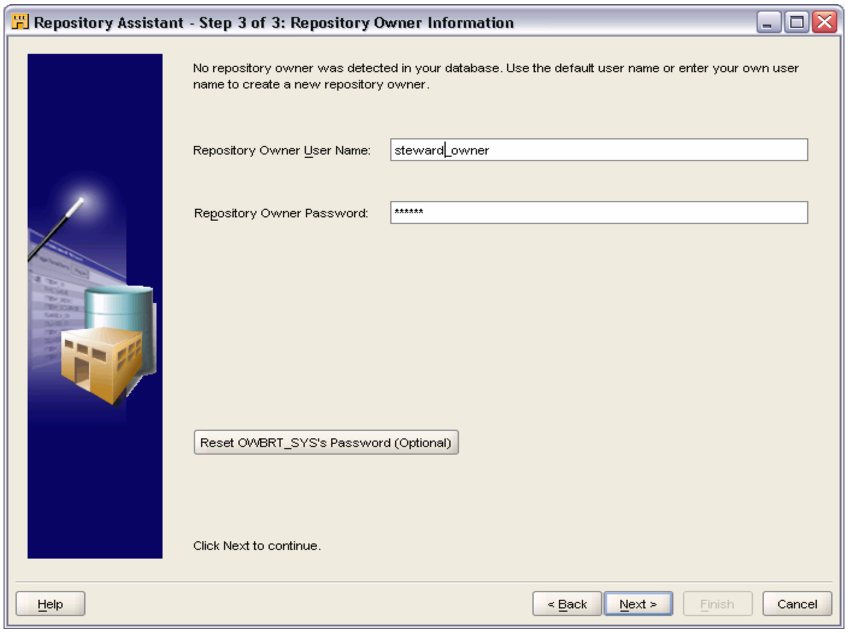


Figure B- 20: Oracle warehouse builder- Repository owner information

The repository installation will take several minutes and an ***Installation Successful*** message will appear as Figure B-21.

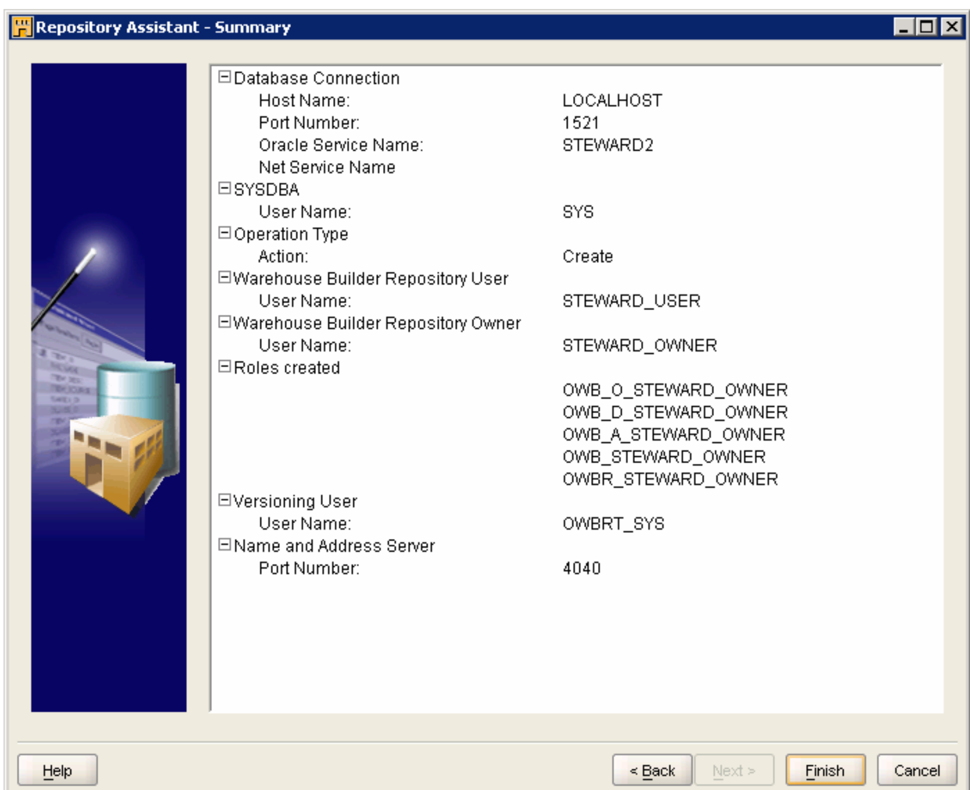


Figure B- 21: Oracle warehouse builder- Installation Summary

2.4 Oracle Workflow Configuration

2.4.1 Create the owf_mgr Workflow schema using the Workflow Configuration Assistant.

Run the Oracle Workflow Configuration Assistant program as shown in Figure B-

START > All Programs > {Oracle_Home} > Configuration and Migration tools > Workflow Configuration Assistant.

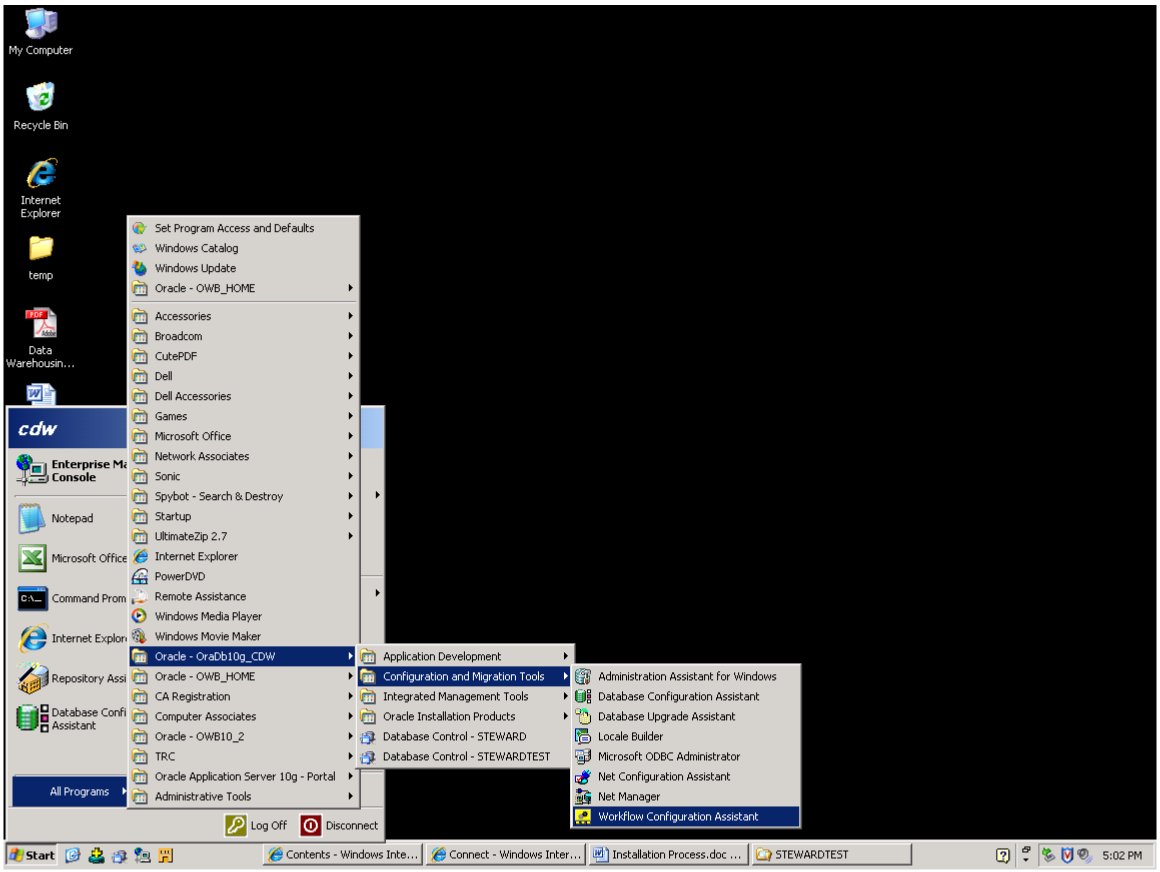


Figure B- 22: Oracle workflow configuration

2.4.2 Set the workflow configuration as follows

Set the workflow account as follows. Workflow configuration assistant window is as Figure B-23.

- Workflow Account: owf_mgr
- Workflow Password: **trc513**
- Sys Password: **trc513**
- TNS Connect descriptor: **localhost:1521:steward2**

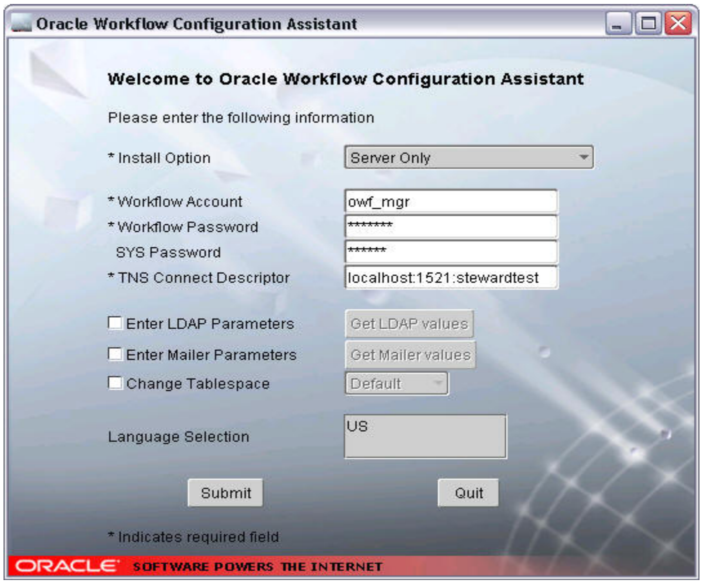


Figure B- 23: Oracle workflow configuration assistant

2.4.3 Unlock the owf_mgr account

After the Workflow Configuration Assistant has completed successfully, unlock the owf_mgr account. Log into SQL Plus as a SYSDBA(sys/trc513). As shown in Figure B-24, enter the following commands:

- alter user owf_mgr account unlock;
- grant execute any procedure to owf_mgr;
- grant owb_o_steward_owner to owf_mgr;

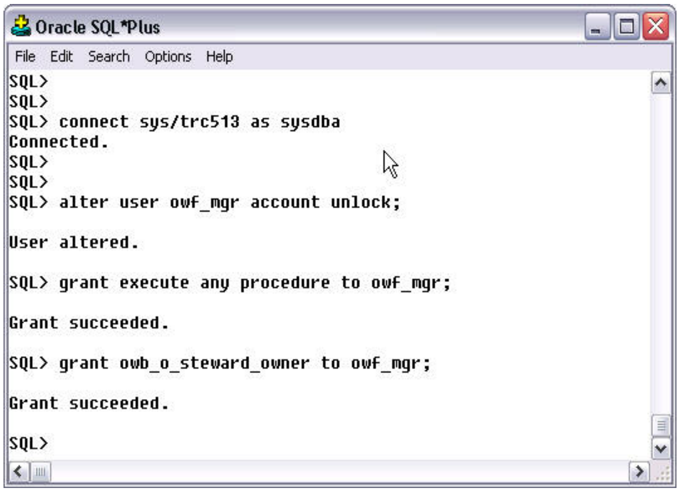


Figure B- 24: Oracle workflow configuration - Unlock the owf_mgr

3 STEWARD Deployment

3.1 Prerequisites: Create a Target User

The Steward2 project has a STEWARD module. Every target module must be mapped to a target user schema. This target schema physically stores target objects on deployment. Also, each target module references a target schema by an assigned location. To create the STEWARD target schema users, perform the following steps:

Expand the ***Security*** node in the Global Explorer panel. Right-click **Users node** and select *New*. Then select ***Create DB User*** to show Figure B-25.

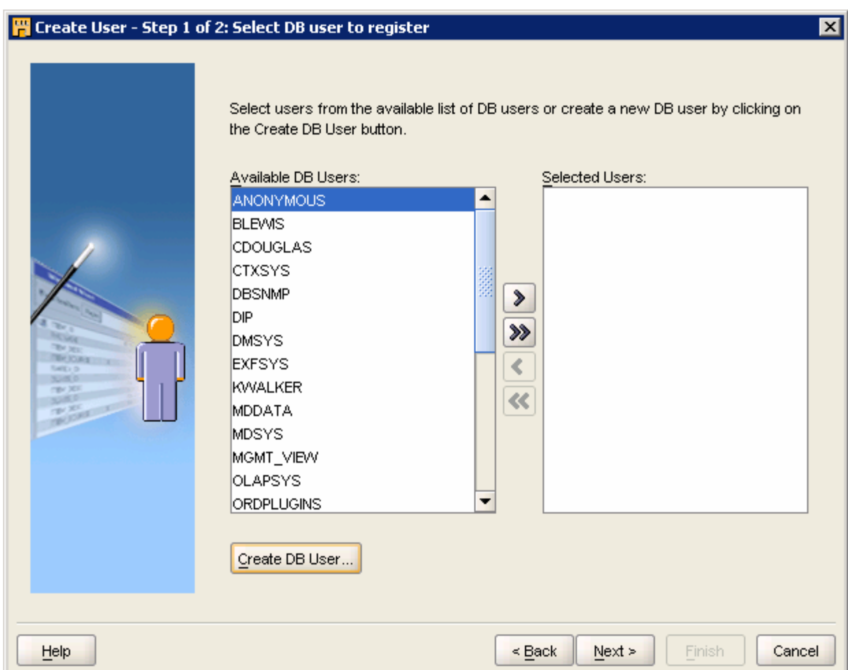


Figure B- 25: Oracle warehouse builder - Create target user

Create the DB user as follows. Figure B-26 shows the screenshot of database user creation window:
Name-GATOR
Password-TRC513

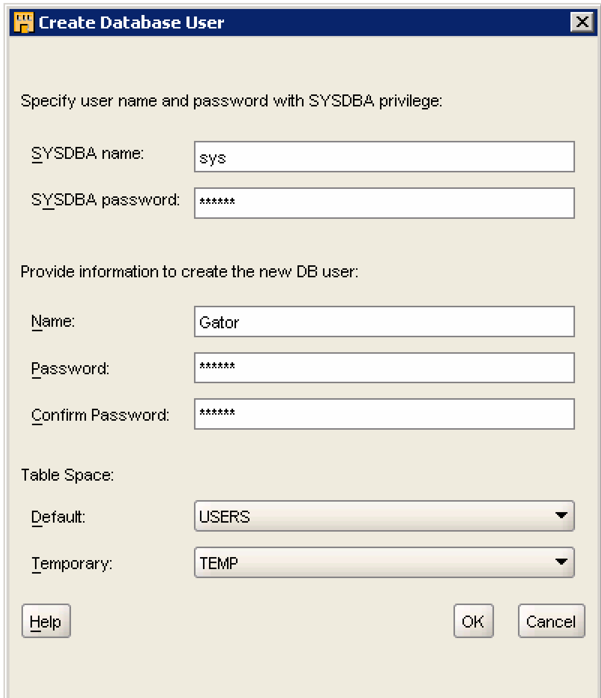


Figure B- 26: Oracle warehouse builder - Create database user

In the ***Connection Explorer*** panel, expand the ***Oracle Locations*** node. Verify that a new location, "Steward_user _LOCATION." has been added. (Connection Explorer → Locations →Databases→Oracle)

Delete the GATOR_LOCATION in the ***Oracle Locations*** to import Metadata.

3.2 Importing Metadata

- From the menu, select [***Design*** > ***Import*** > ***Warehouse Builder Metadata***]
- Select the File Name and Log File that will be imported using ***Browse*** button.
- Click the ***Import*** button as shown in B-27.

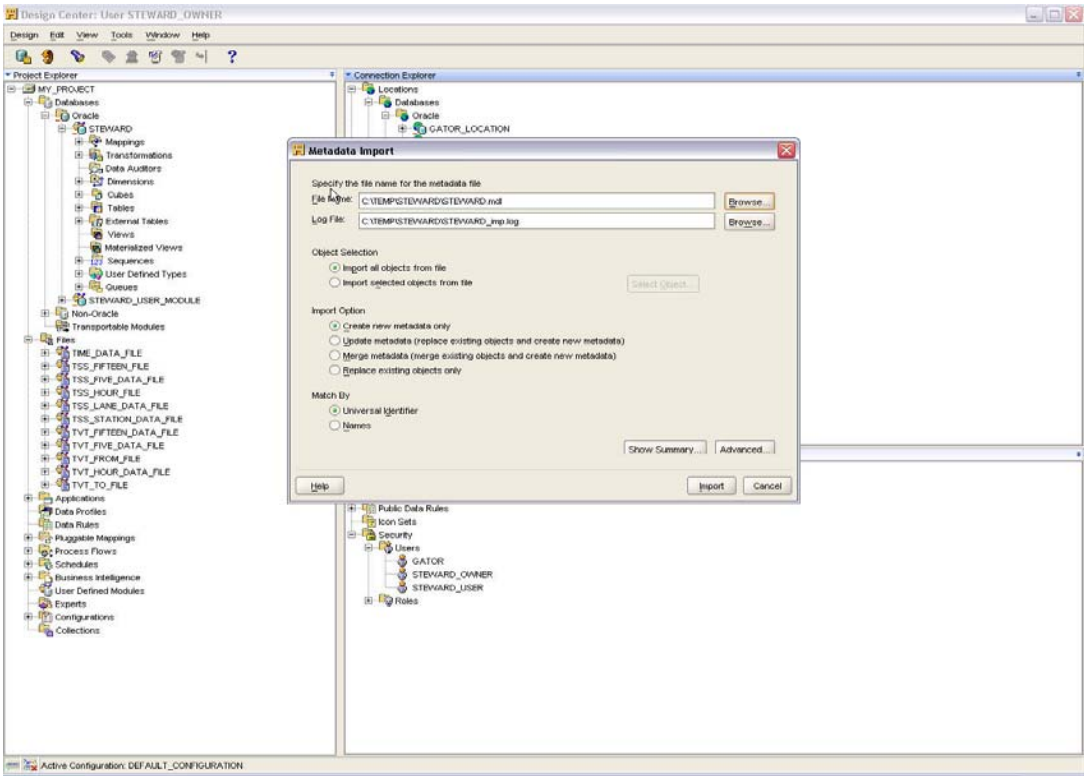


Figure B- 27: Oracle warehouse builder - Metadata import

The metadata need be registered using *Control Cente*r tool. To register metadata, right click on “GATOR_LOCATION,” “LOAD_SCHEDULE_LOCATION,” and “STEWARD_PF_MODLUE_LOCATION.” Open the editor in the *connection Explorer* as shown in Figure B-28. Click *Test* button.

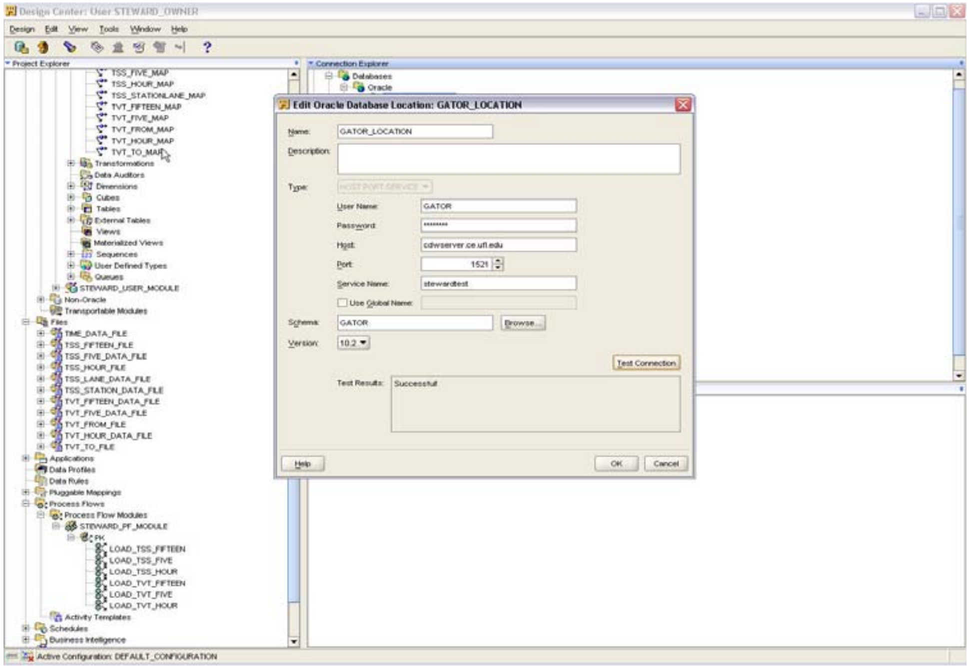


Figure B- 28: Oracle warehouse builder – Edit database location

To open the *Control Center*, Click *Tools* in the menu and click *Control Center Manager*.

In the Control Center, the same locations will be shown on the left side and they are registered when saved as shown in Figure B-29. For the file locations, when the external tables are deployed, they can be registered automatically.

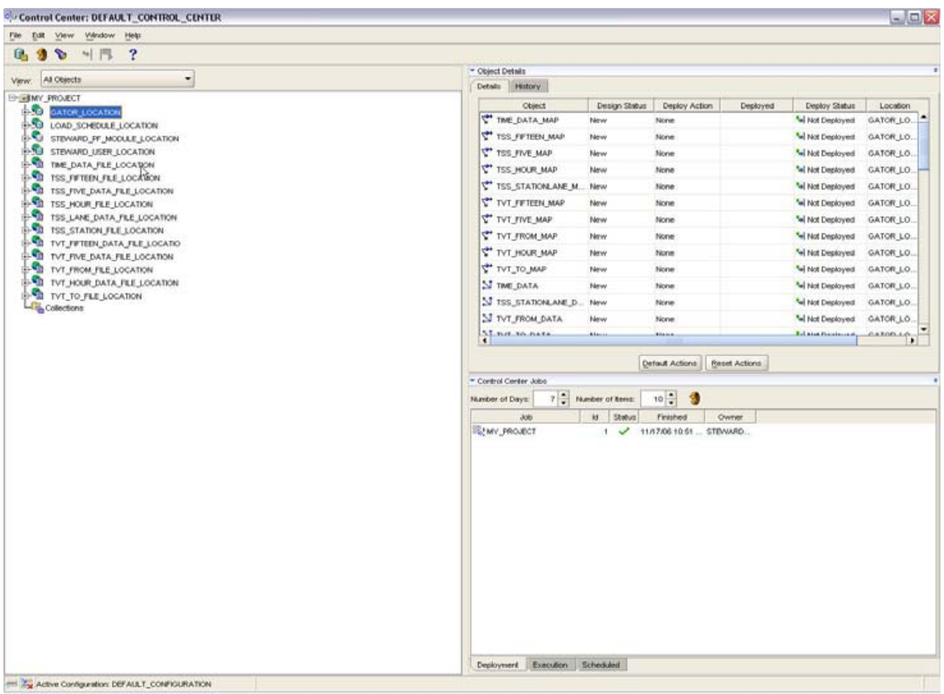


Figure B- 29: Oracle warehouse builder – File location in control center

3.3 Registration of the Control Center Manager

All locations need to be registered at first.

As shown in Figure 30, select the *Oracle* tab and select **.LOCATION* one by one in the *Connection Explorer*. Double-click on each *LOCATION*. Select *Open Editor*. Enter the user name, password, host, and service name. Press *Test Connection*.

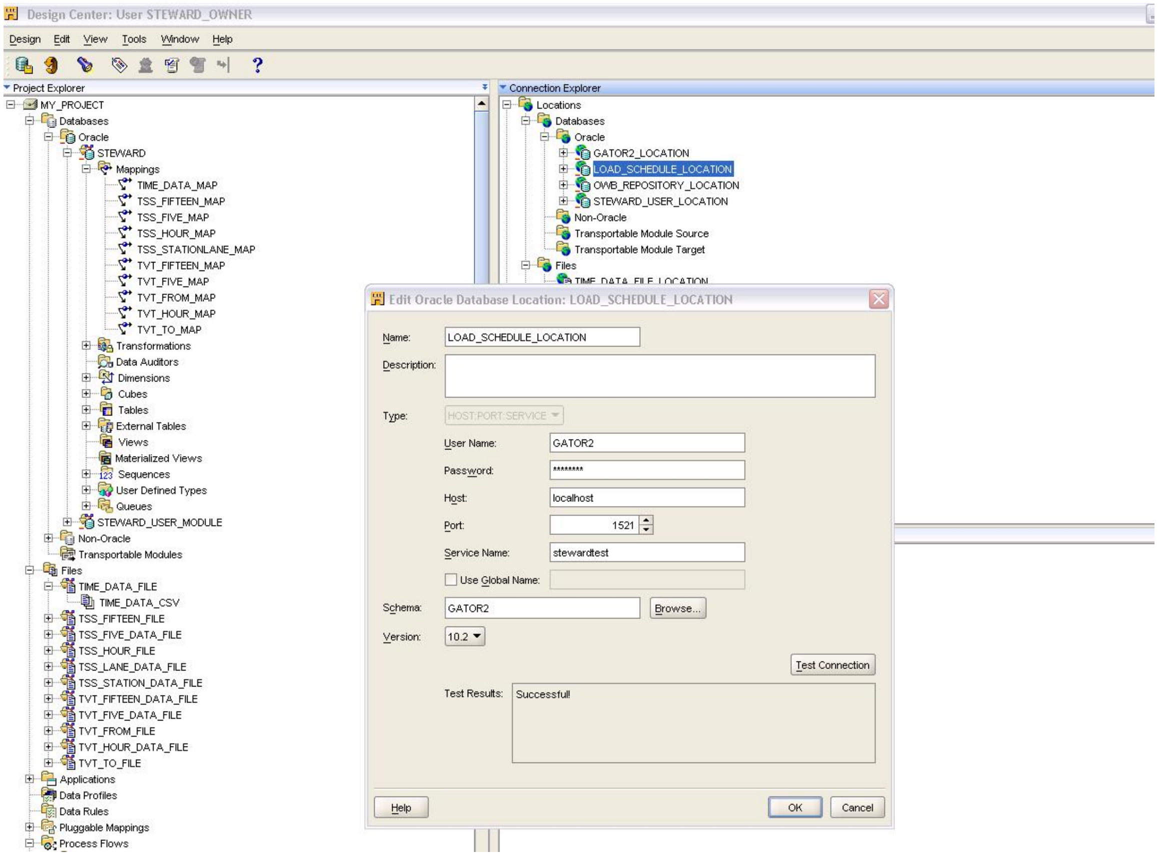


Figure B- 30: Oracle warehouse builder – Edit *Connection Explorer*

As shown in Figure 31, open the *Control Center Manager* and press the *Refresh* Button. You can now see each *LOCATION* item on the left side.

In the *Connection Explorer*, select the *File* tab. Then select **.LOCATION* one by one. Double-click each *LOCATION*. Press the *OK* button.

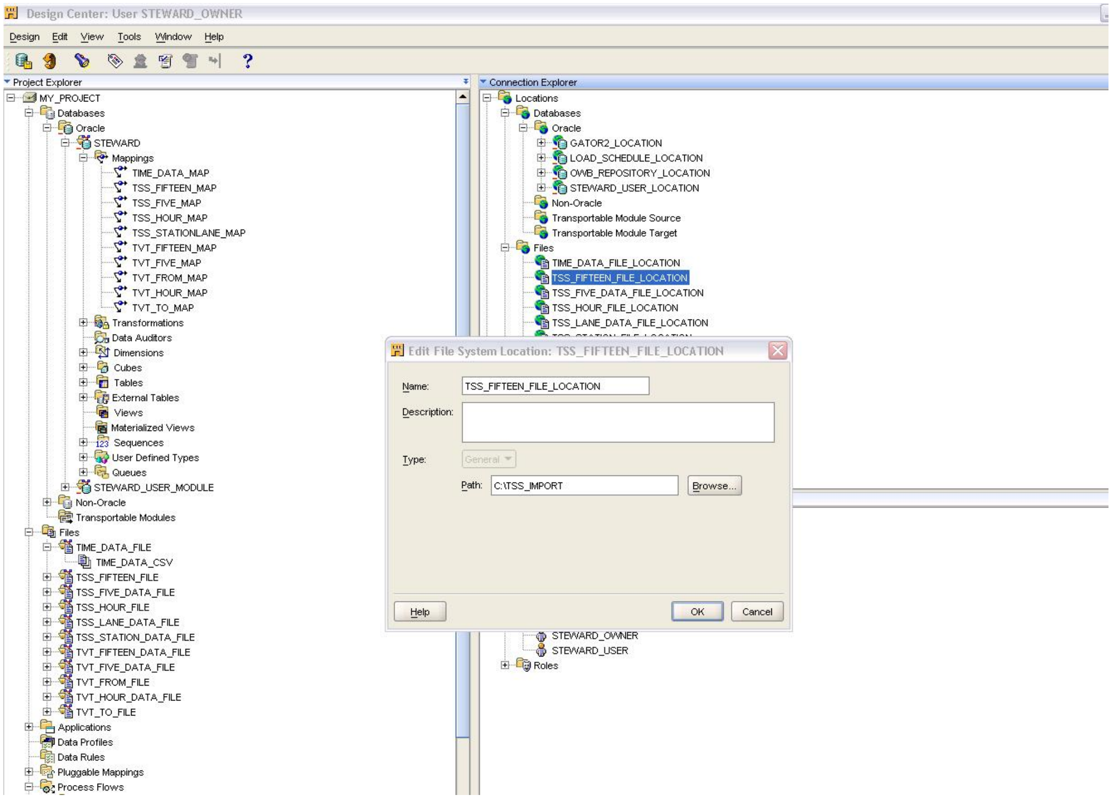


Figure B- 31: Oracle warehouse builder – Edit *File system location*

As shown in Figure 32, open the *Control Center Manager* and press the *Refresh* button. You will now see each *LOCATION* item on the left side.

In the *Connection Explorer*, select *Process Flow* and the *Schedule* tab. Then select *Oracle Workflow*. Now select *.LOCATION* one by one. Double-click each *LOCATION*. Enter suitable values in the blank fields and press *Test Connection*.

Open the *Control Center Manager* and press the *Refresh* button. You will now see each *LOCATION* item on the left side.

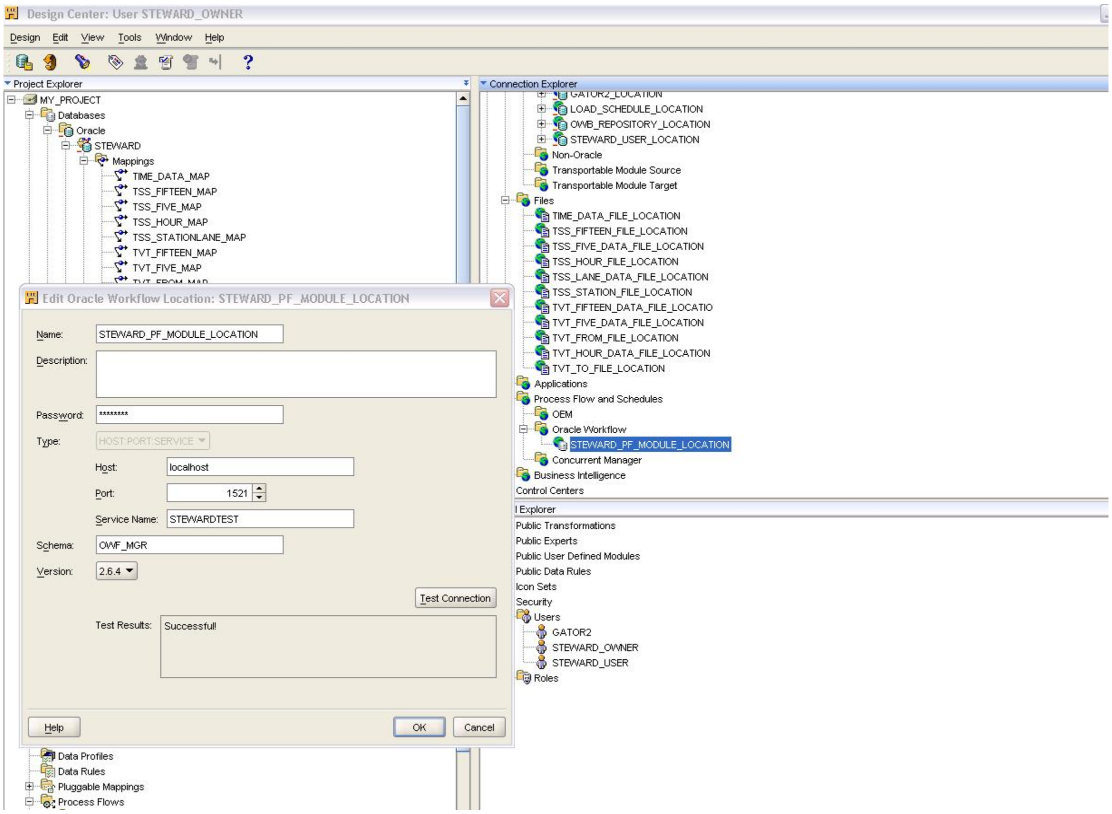


Figure B- 32: Oracle warehouse builder – Edit File system location

The updated *Control Center Manager is shown in Figure B-33.*

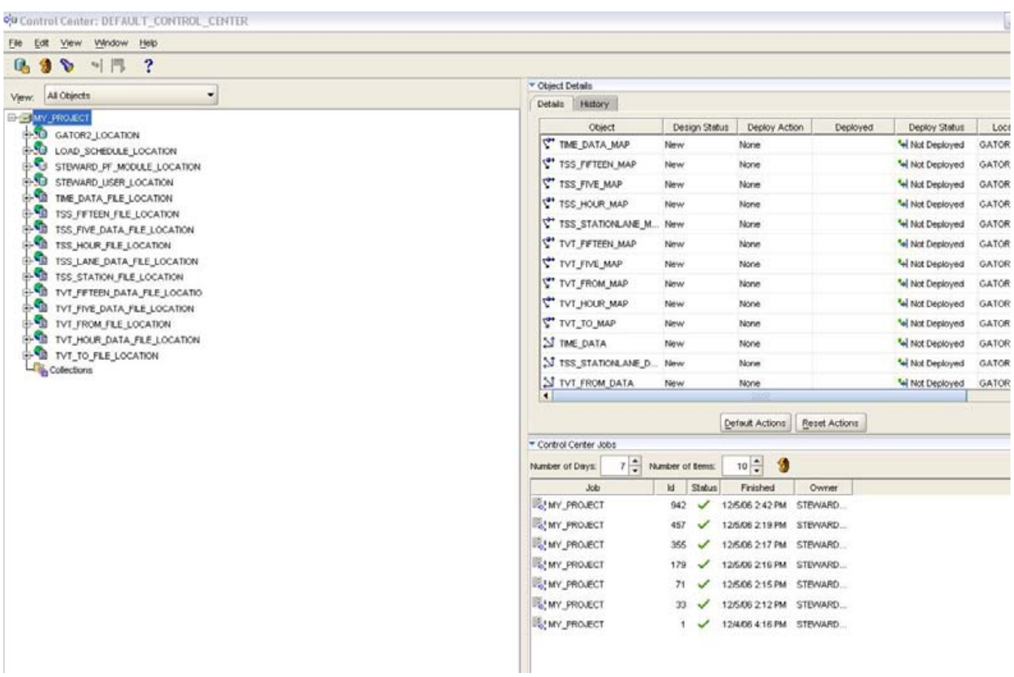


Figure B- 33: Oracle warehouse builder – Updated Control Center

3.4 Deployment and Data loading process

3.4.1 External Table Deployment

As shown in Figure B-34, select *External Tables* and click the *Default Actions* button. Then, click the *Deploy* button to deploy the DB.

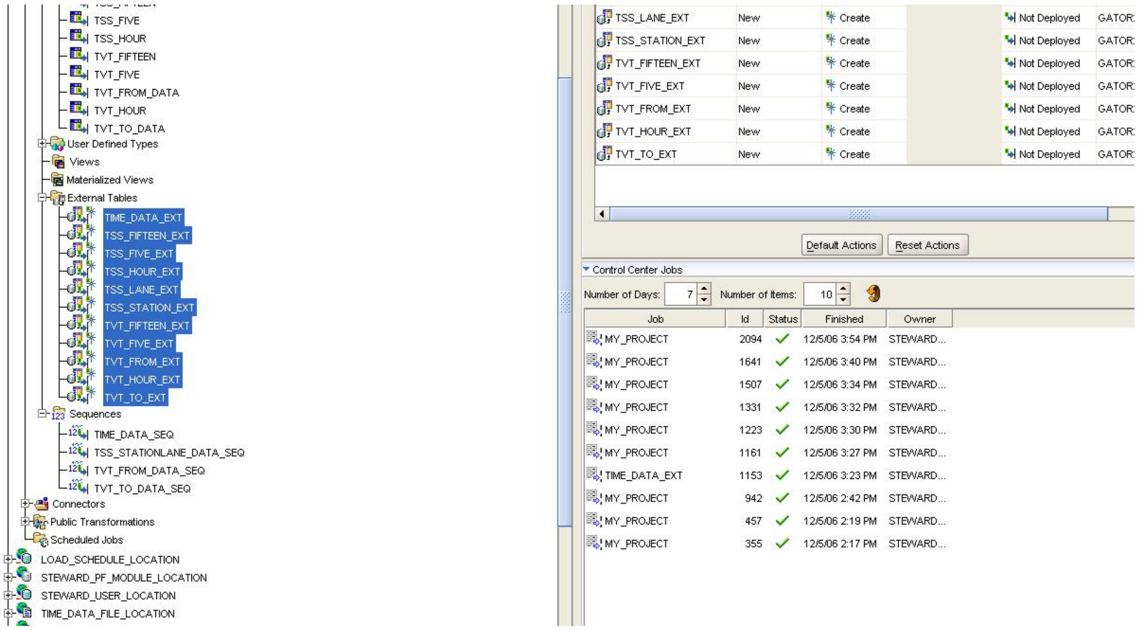


Figure B- 34: Oracle warehouse builder – External Table Deployment

3.4.2 Dimension Table Deployment

As shown in Figure B-35, select all Dimension tables and all sequences and click the *Default Actions* button. Then, click the *Deploy* button to deploy the DB.

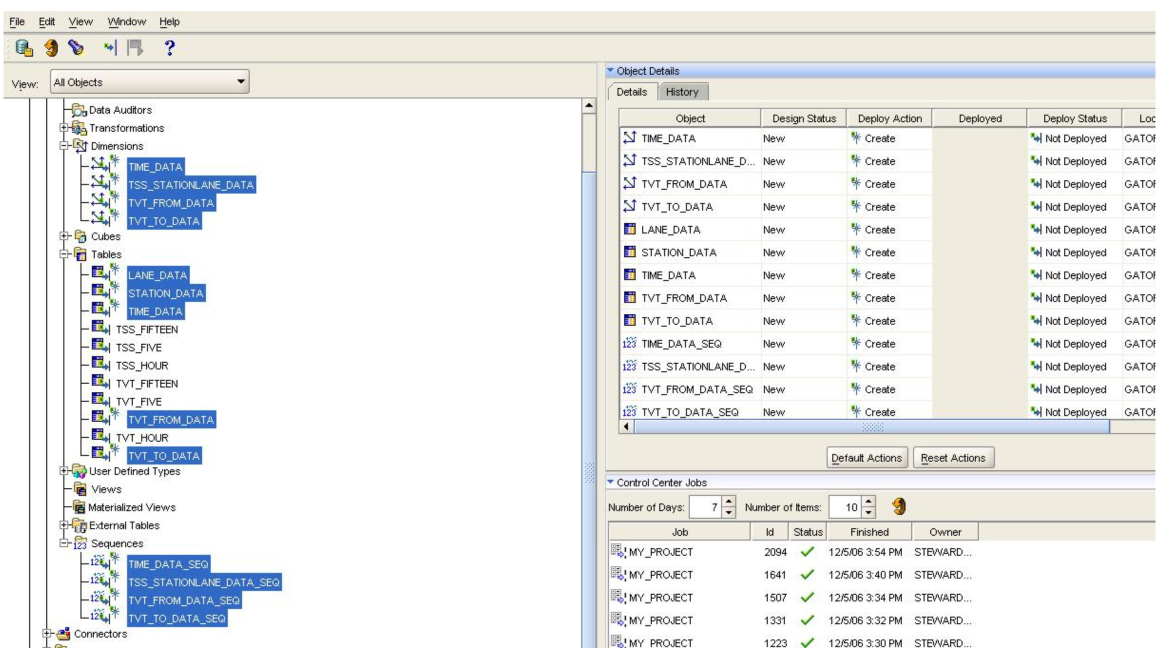


Figure B- 35: Oracle warehouse builder – Dimension Table Deployment

3.4.3 Fact Table Deployment

As shown in Figure B-36, select all cubes and tables for the Fact Table and click the *Default Actions* button. Then, click the *Deploy* button to deploy the DB.

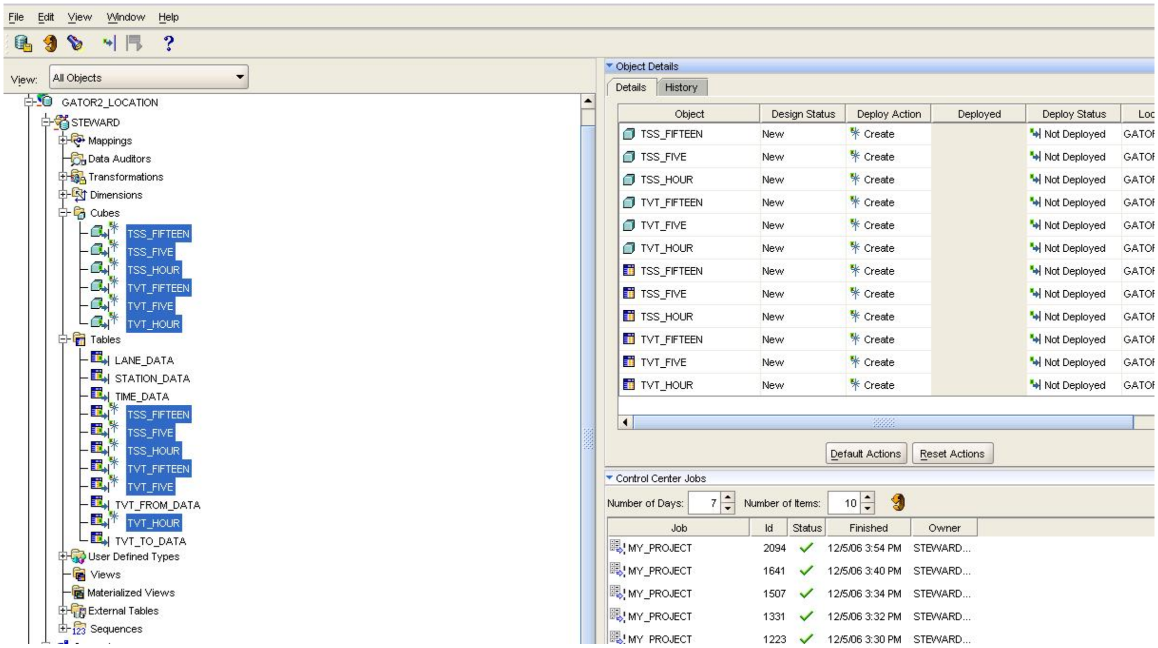


Figure B- 36: Oracle warehouse builder – Fact Table Deployment

3.4.4 Mappings Deployment

As shown in Figure B-37, select all mappings and click then ***Default Actions*** button. Then click the ***Deploy*** button to deploy the DB.

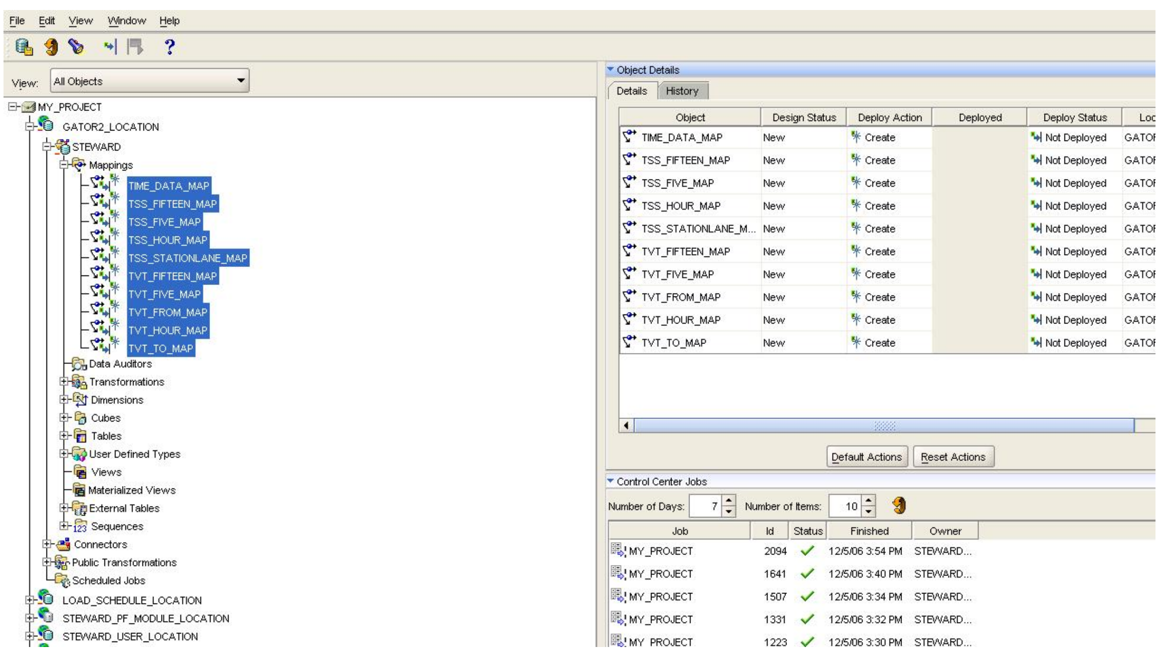


Figure B- 37: Oracle warehouse builder – Mappings Deployment

3.4.5 Dimension Table Data loading

The dimension table doesn't change at all. As shown in Figure B-38, select mappings for each dimension table one by one. Then, click the *Start* button to begin loading data.

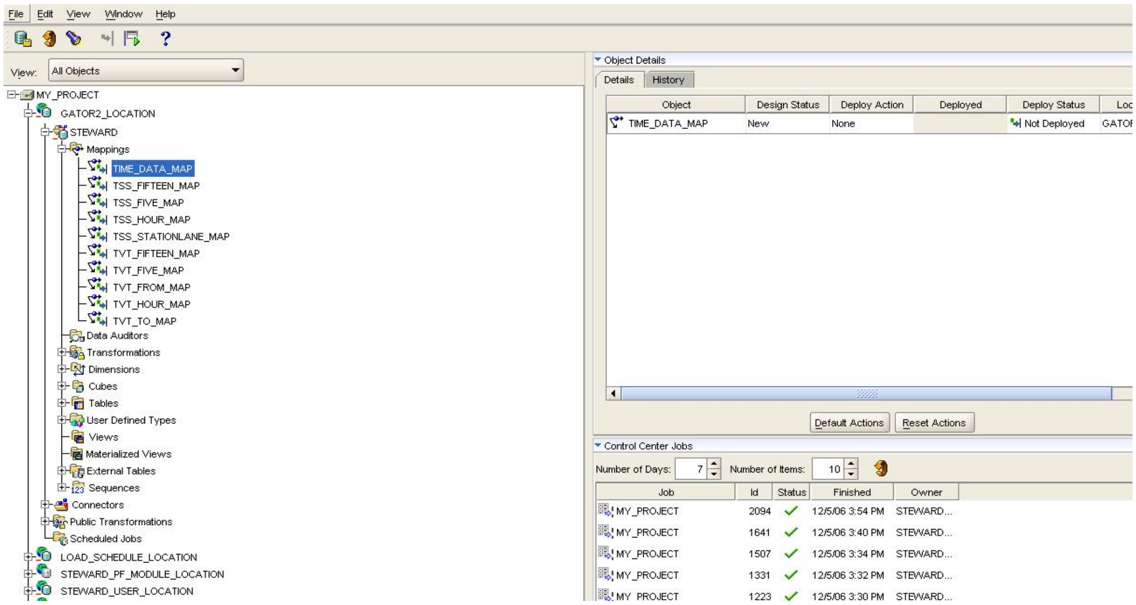


Figure B- 38: Oracle warehouse builder – Loading dimension table data

3.4.6 Loading Fact Table Data

As shown in Figure B-39, select *STEWWARD_PF_MODULE_LOCATION* and deploy it first. Then select each item one by one and click the *Start* button. This will load one day's data. Remember to click *Start* button only one time.

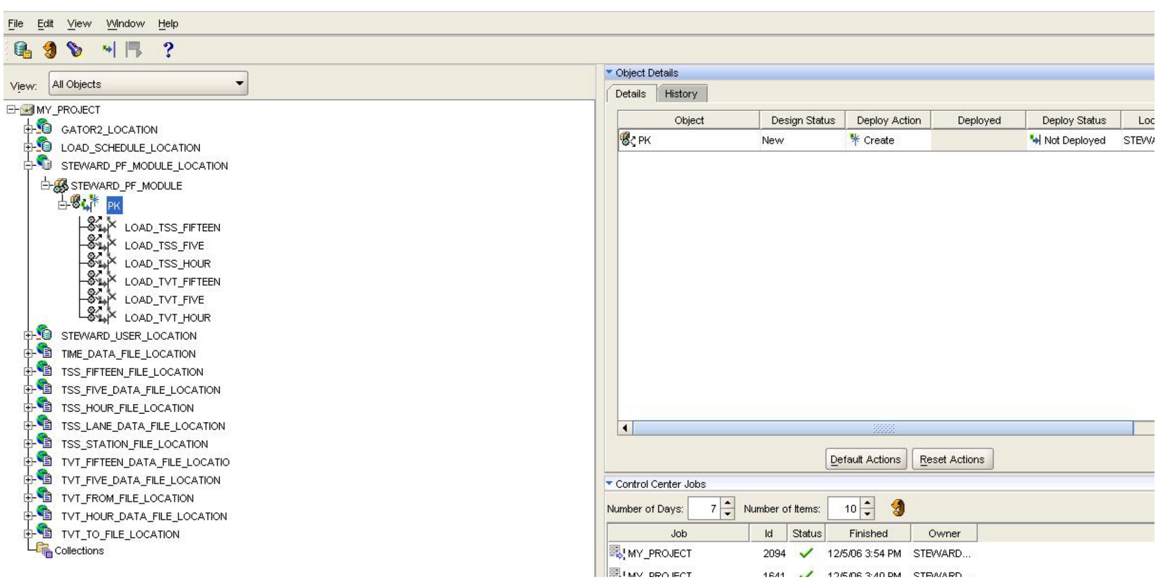


Figure B- 39: Oracle warehouse builder – Loading fact table data

To load the next day's data, change the ***.csv** file in the source destination. Then, click the ***Start*** button. You can check the table contents using ***SQL*PLUS***.

4 STEWARD Web Installation

To communicate with the Oracle database, the ***Net configuration*** for Oracle data needs to be set up. This is accomplished through the Net Configuration Assistant for Oracle. Access to the Net Configuration Assistant is illustrated in the Figure B-40 .

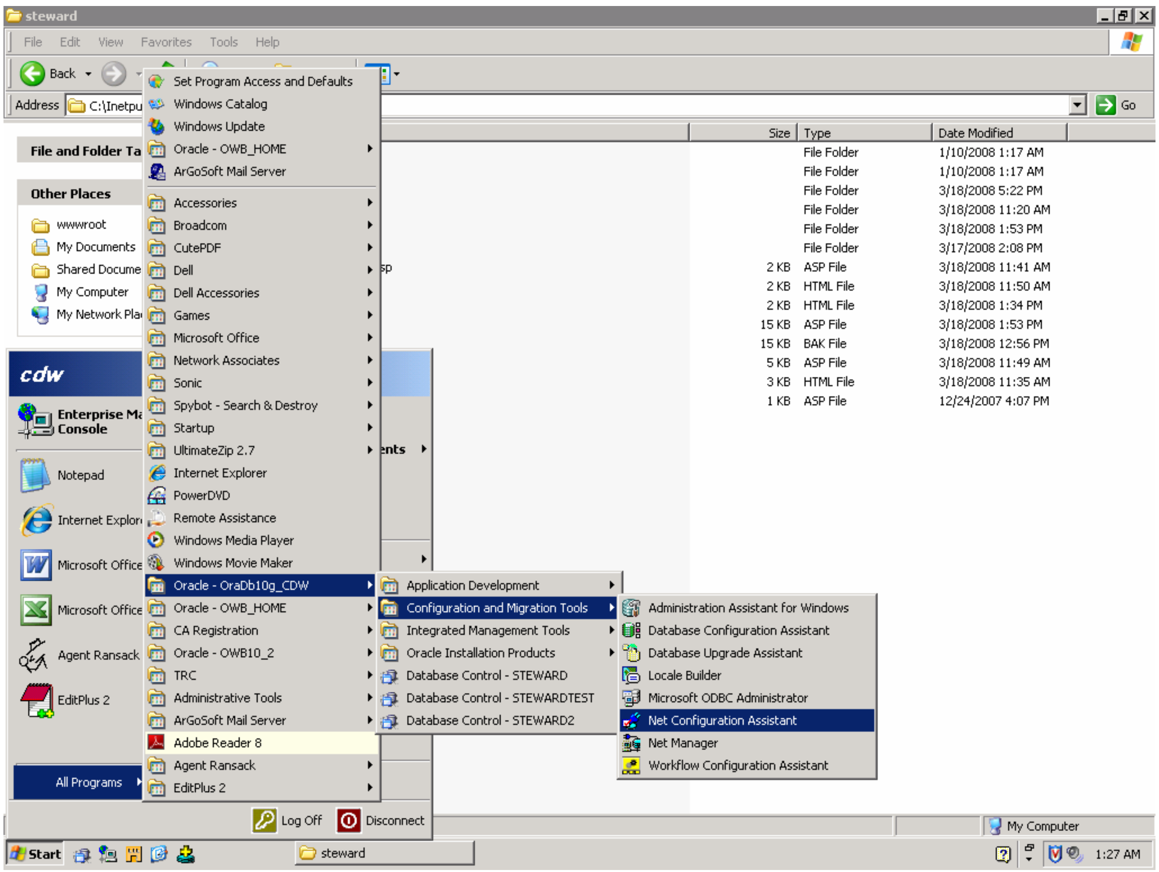


Figure B- 40: Steward web installation – run web configuration assistant

4.1 Net Configuration Assistant Steps

The following steps must be performed using the Oracle Net Configuration Assistant:

Step1: As shown in Figure B-41, select local net service name configuration.

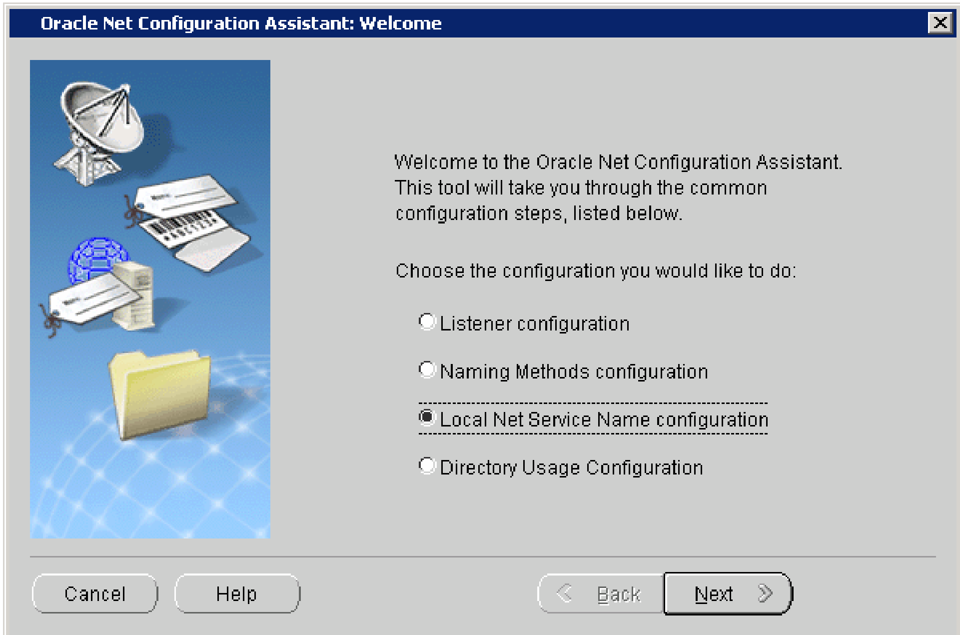


Figure B- 41: Steward web installation – Oracle net configuration1

Step2: As shown in Figure B-42, select service name as Steward2.

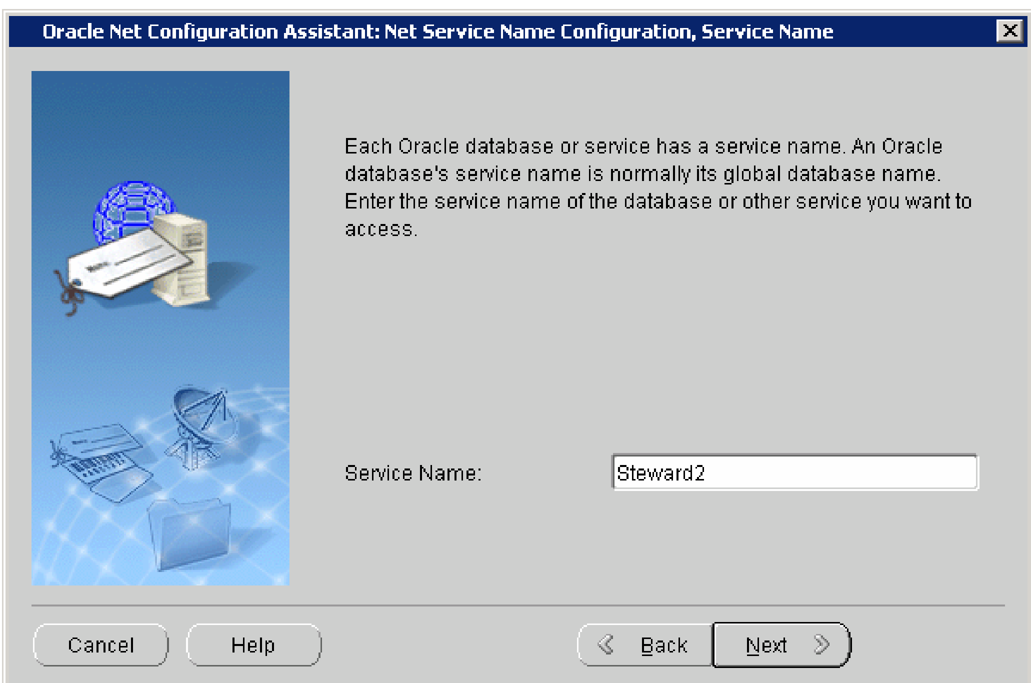


Figure B- 42: Steward web installation – Oracle net configuration2

Step3: As shown in Figure B-43, select the communication protocol as TCP.

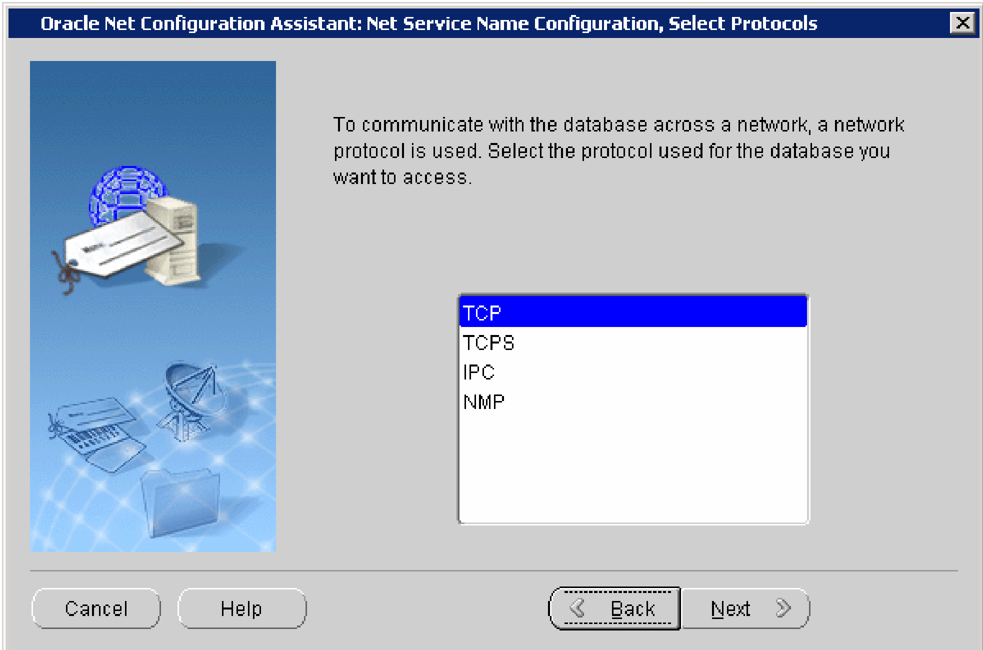


Figure B- 43: Steward web installation – Oracle net configuration3

Step4: As shown in Figure B-44, edit the Host name as cdwserver.ce.ufl.edu.

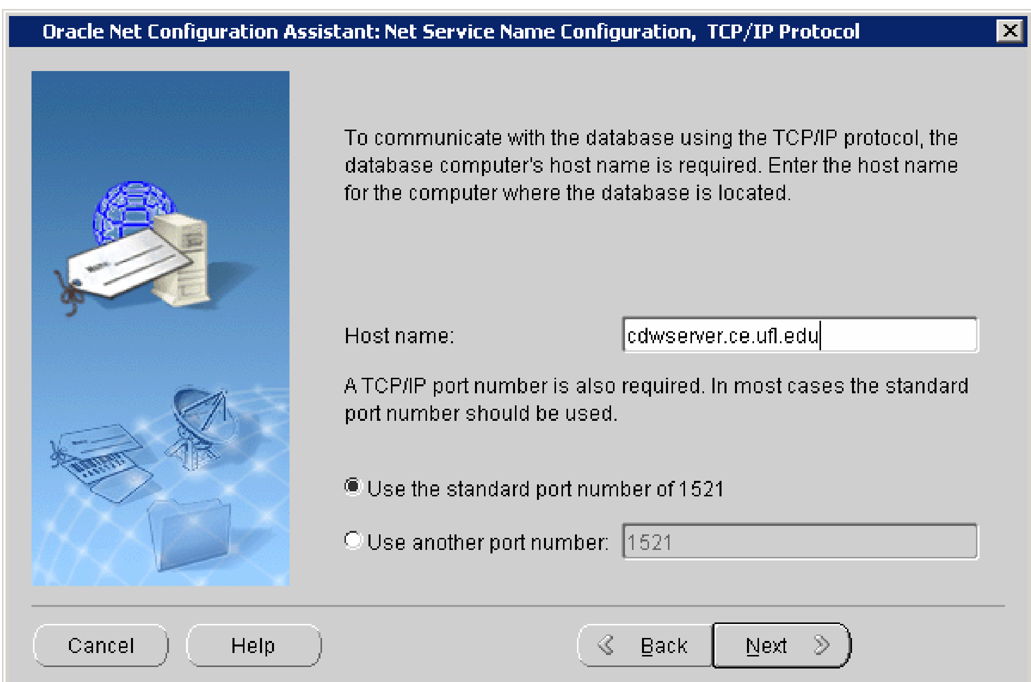


Figure B- 44: Steward web installation – Oracle net configuration4

Step5: As shown in Figure B-45, perform the test for the new Oracle net configuration

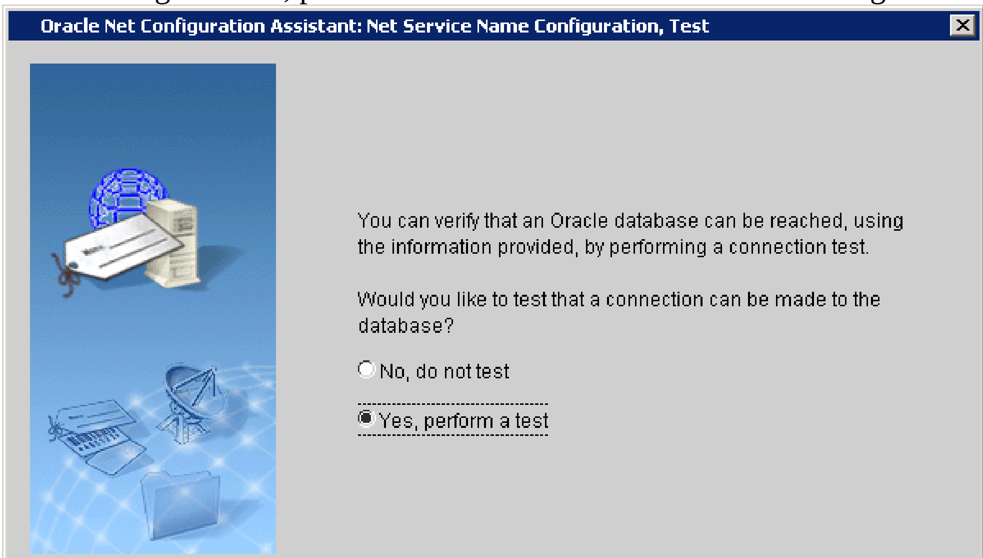


Figure B- 45: Steward web installation – Oracle net configuration5

Step 6: If the test fails as Figure B-46 for invalid login ID or password, change the login ID and password by clicking change login.

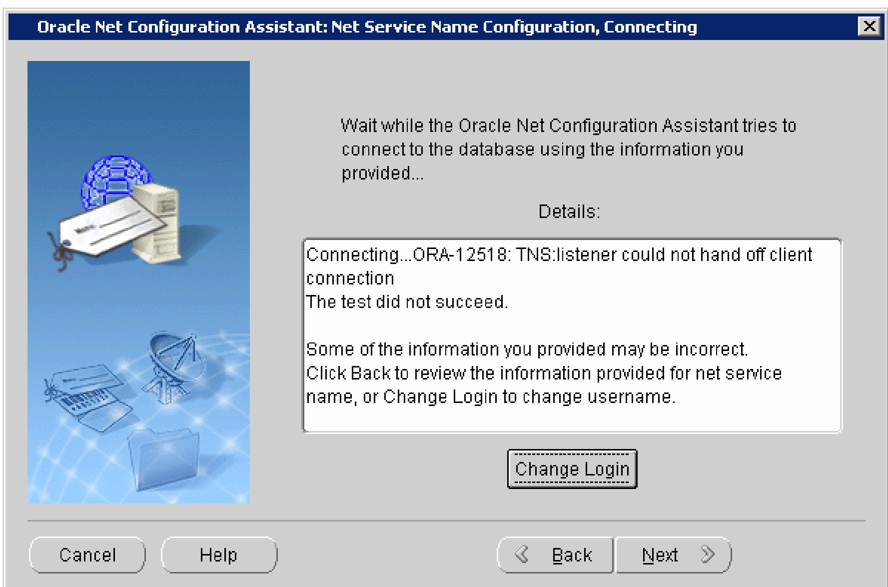


Figure B- 46: Steward web installation – Oracle net configuration6

Step 7: As shown in Figure B-47, set the user name/password as “gator/trc513”

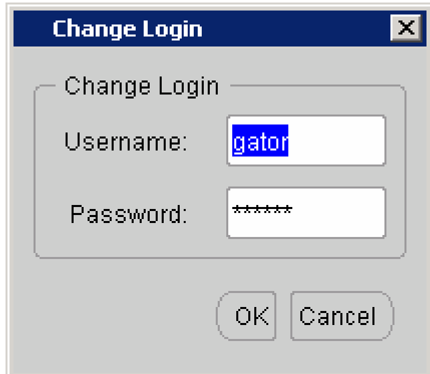


Figure B- 47: Steward web installation – Oracle net configuration7

Step 8: As shown in Figure B-47, the remote login should be successful.

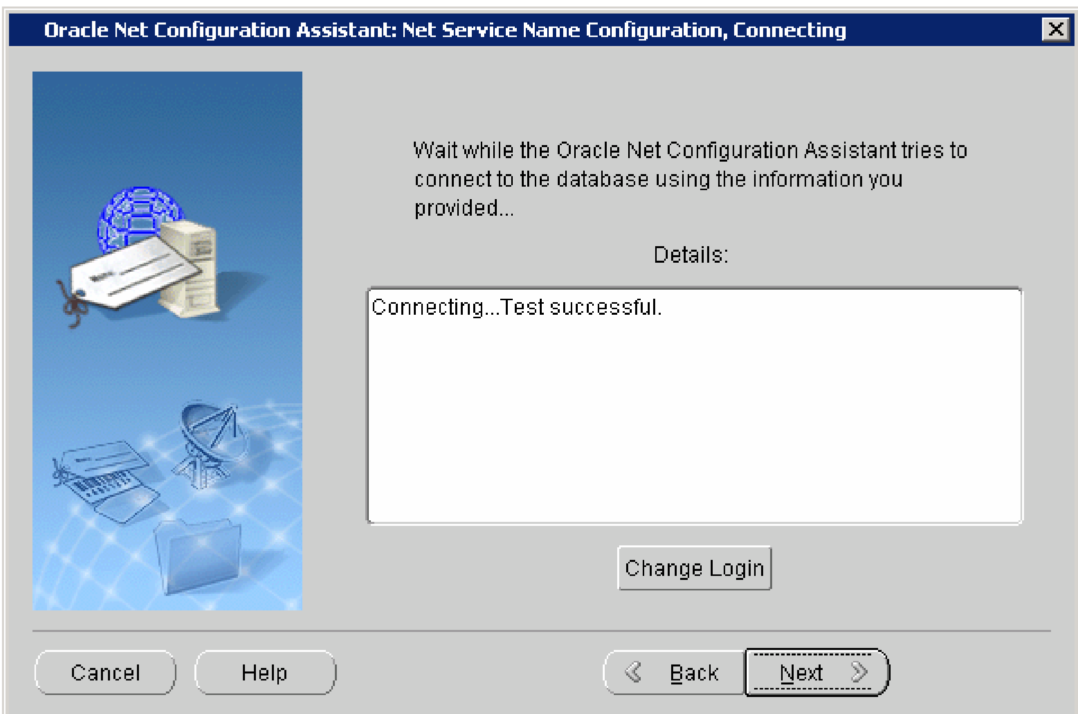


Figure B- 48: Steward web installation – Oracle net configuration8

Asp codes and Java Scripts don’t need any specific installation procedures. The program works as the source codes are copied to the directory (steward) as the web root directory. In case of Microsoft Internet Information Services version 5.1, the default root directory is located at C:\Inetpub\wwwroot and the web program need be copied to STEWARD directory.

4.2 STEWARD Web Program Installation

There are no special installers or routines for the STEWARD web program. The program itself needs be extracted to the specific directory (C:\Inetpub\wwwroot: Root directory for Microsoft Internet Information Services). If the database connection is successful, the STEWARD Web program will respond to the remote Microsoft Internet Explorer.

4.3 System configuration

After the STEWARD web program is installed on web server, the system configuration needs be updated. There are two main configuration operations. One is to enable the web server on the system firewall. The other is to add the sharing permission for the public user to access the traffic data directory

4.3.1 Firewall setting

The STEWARD web server is installed on Windows XP service pack 2 with the default firewall. In this case, the local area connection needs be updated to enable the web service. The procedures are as follows:

1. Open the Windows firewall icon from windows control panel as shown in Figure B-49.

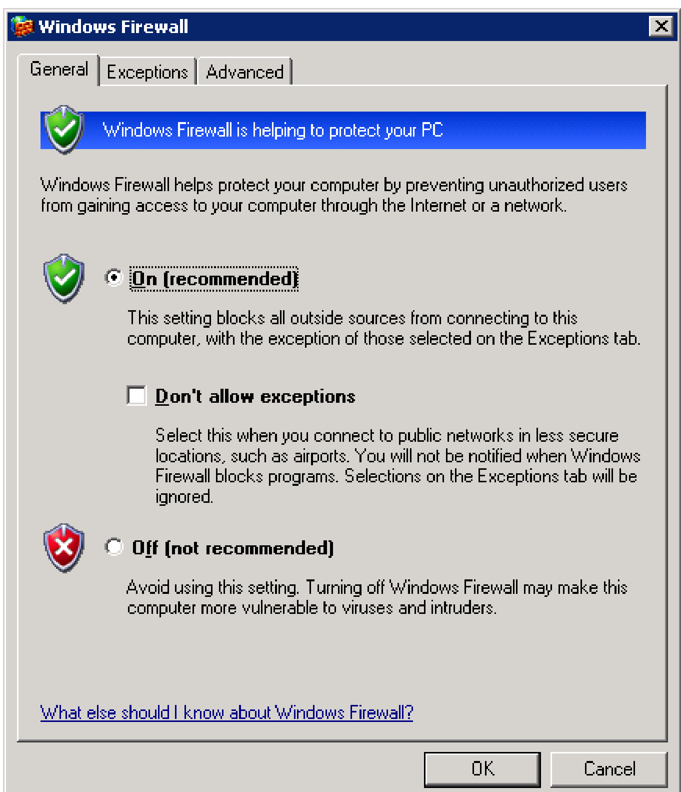


Figure B- 49: Steward web installation – System configuration1

2. Select the ***Advanced*** tab as shown in Figure B-50.

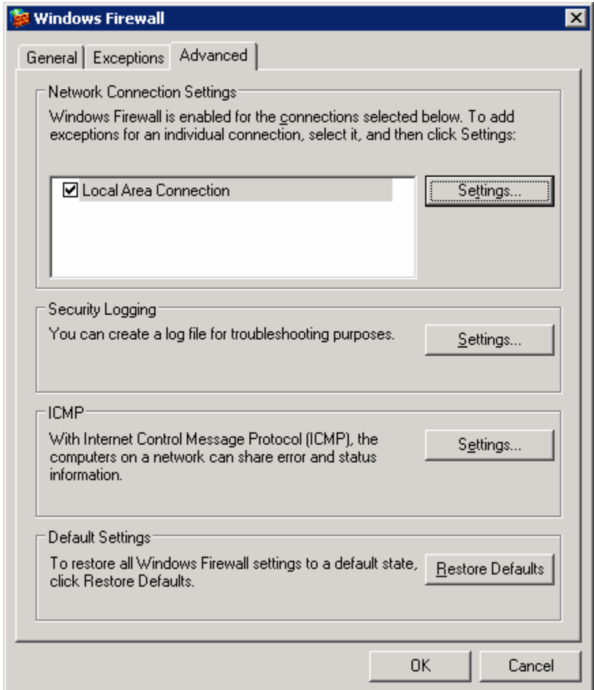


Figure B- 50: Steward web installation – System configuration2

Settings for the Local Area Connection

Check ***Web Server*** (HTTP) as shown in Figure B-51.

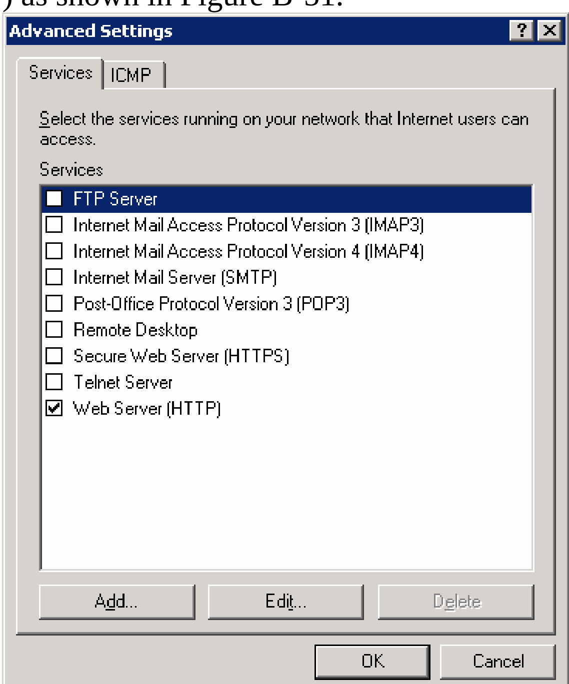


Figure B- 51: Steward web installation – System configuration3

4.3.2 Permission for the file sharing

The files folder is located in the STEWARD root directory (wwwroot\Steward\files). This directory needs to update the web sharing options for the web user to open or download data file. The procedure is as follows:

Open the ***Properties*** screen for the ***Files*** directory as shown in Figure B-52.

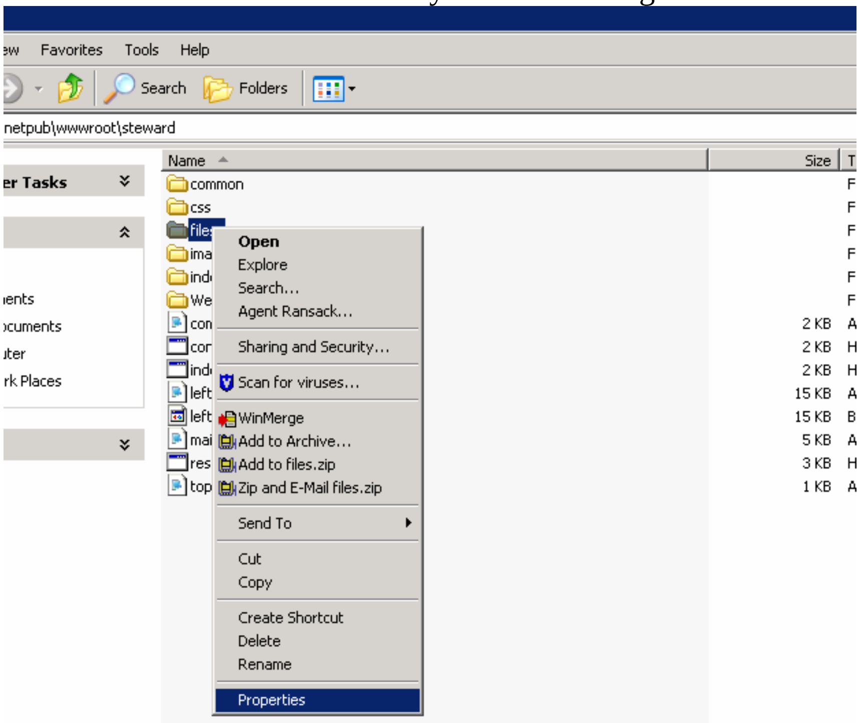


Figure B- 52: Steward web installation – System configuration4

Select the web sharing tab and change the sharing option to allow sharing of this folder as shown in Figure B-53.

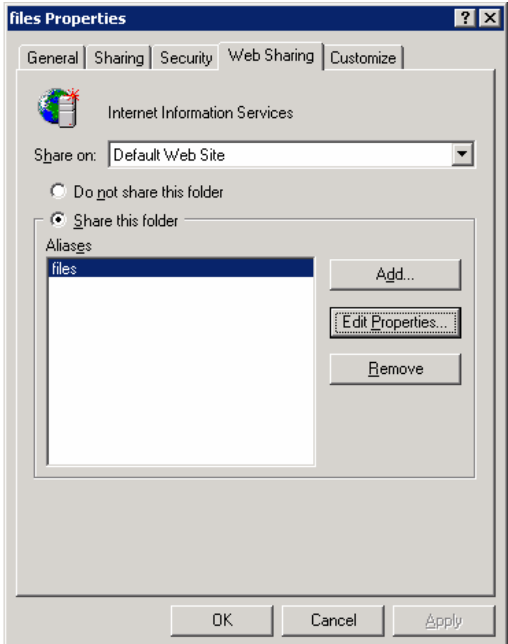


Figure B- 53: Steward web installation – System configuration4

Edit properties as shown in Figure B-52:

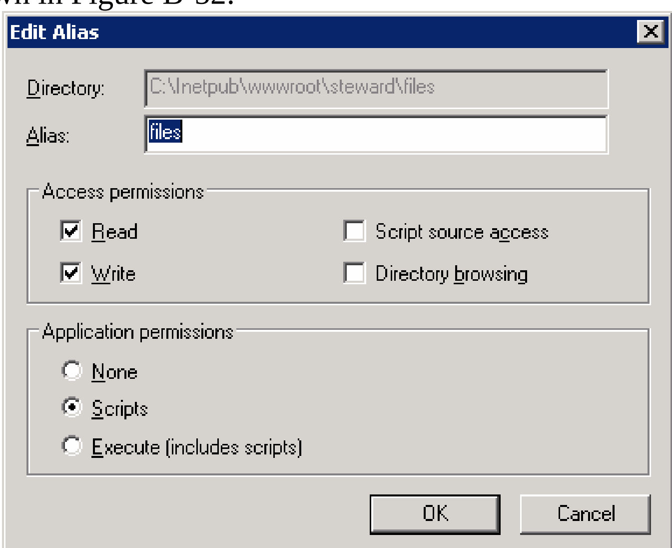
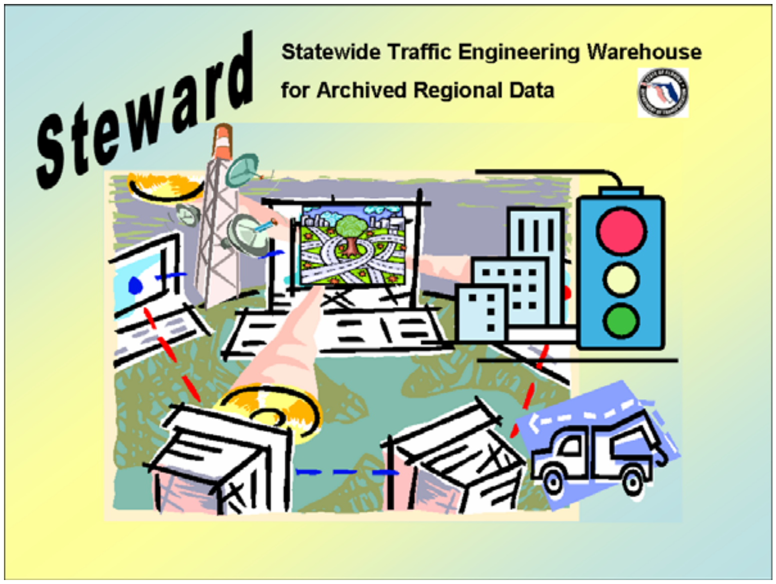


Figure B- 54: Steward web installation – System configuration5

This will complete the set up of the STEWARD Internet access



Development of a Central Data Warehouse for Statewide ITS and Transportation Data in Florida

Phase II: Proof of Concept Final Report

Appendix C
Instructions and Guidance for Operating STEWARD

Prepared for the Florida Department of Transportation
By the University of Florida Transportation Research Center

Contract # BD545, RPWO # 37
UF Project 051449

Revision Date: April 15, 2008

Table of Contents

List of Figures

List of Tables

1 Introduction

A description of the steps required to obtain archived data from the districts and add it to the Statewide Traffic Engineering Warehouse for Regionally Archived Data (STEWARD) databases is presented in this document as an appendix to the STEWARD Final Report. This material has been developed to assist FDOT and the ITS contract personnel in operating STEWARD in the FDOT Traffic Engineering Research Laboratory (TERL) facility. The topics include Oracle database program installation, STEWARD deployment and the STEWARD web site installation. A working knowledge of the Oracle data base manager and the STEWARD features and configuration is assumed in the discussion that follows.

The topics include:

- Program configuration
- The extraction, transformation and loading (ETL) process
- Exporting metadata from the STEWARD database
- STEWARD database tables reference

2 STEWARD Program Configuration

STEWARD consists of one database (DB) server and one file server. The DB server includes the Oracle DB, web server and application programs. The file server includes Oracle database files and is designed to be mapped as the network drive.

2.1 Programs Used in the Project

2.1.1 Programs Developed for the STEWARD Project

- SunETLUtility
- Internal Utility programs

2.1.2 Commercial Programs on the Database Server

- Oracle Database 10g 10.2.0.1.0 Enterprise edition
- Oracle Warehouse Builder 10g 10.2.0.1.31
- Oracle Workflow 2.6.4.0.0
- Microsoft Internet Information Services 5.1

2.2 File Locations

The assumptions on file locations include:

- The commercial programs will be installed in their default locations.
- The STEWARD programs will be installed in the locations specified in the Table C-1.
- The STEWARD TSS data file will be installed in the locations specified in the Table C-2.
- The STEWARD TVT data file will be installed in the locations specified in the Table C-3.
- The STEWARD database file will be installed in the locations specified in the Table C-4.

Table C- 1: STEWARD Program File Locations

Location	Descriptions	Files
C:\Program Files\TRC\SunETLUtility	Default directory for SunETLUtility	
C:\STEWARD\Scripts	Utility programs for the ETL process	Move2SunETLUtilityFolder.bat PrepareFileLoading.exe Move2STEWARDFolder.exe

Table C- 2: STEWARD TSS Data File Locations

Location	Descriptions	Files
C:\STEWARD\Tss_Raw_data	Raw data files (SunGuide archive files from each District)	
C:\STEWARD\Tss_FacilityData	TSS Facility data files in CSV format	TSSMaster-D2.xls TSSMaster-D4.xls TSSMaster-D5.xls TSSStationData.csv
C:\STEWARD\Tss_station_data\Tss_5min_data	SunETLUtility Output files (Input files for 5min station data)	
C:\STEWARD\Tss_station_data\Tss_15min_data	SunETLUtility Output files (Input files for 15min station data)	
C:\STEWARD\Tss_station_data\Tss_1hr_data	SunETLUtility Output files (Input files for 1hr station data)	
C:\STEWARD\Tss_lane_data\Tss_5min_data	SunETLUtility Output files (STEWARD input files for 5min lane data)	
C:\STEWARD\Tss_lane_data\Tss_15min_data	SunETLUtility Output files (Input files for 15min lane data)	
C:\STEWARD\Tss_lane_data\Tss_1hr_data	SunETLUtility Output files (Input files for 1hr lane data)	

Table C- 3: STEWARD TVT Data File Locations

Location	Descriptions	Files
C:\STEWARD\Tvt_Raw_data	Raw data files for STEWARD (SunGuide archive files from each District)	
C:\STEWARD\Tvt_FacilityData	TVT Facility data files in CSV format	TVTMaster-D2.xls TVTMaster-D4.xls TVTMaster-D5.xls
C:\STEWARD\Tvt_data\Tss_5min_data	SunETLUtility Output files (Input files for 5min station data)	
C:\STEWARD\Tvt_data\Tss_15min_data	SunETLUtility Output files (Input files for 15min station data)	
C:\STEWARD\Tvt_data\Tss_1hr_data	SunETLUtility Output files (Input files for 1hr station data)	

Table C- 4: Oracle Database File Locations

Location	Descriptions	Files
K:\ORADATA\STEWARD	Oracle database files for STEWARD	STEWARD2SYSAUX01.DBF STEWARD2SYSTEM01.DBF STEWARD2TEMP01.DBF STEWARD2TEMP02.DBF STEWARD2UNDOTBS01.DBF STEWARD2USERS01.DBF

2.3 Account Information

The following account information applies to the current STEWARD installation. Some of this information may be changed when the operation is migrated to TERL:

2.3.1 User ID and password for Oracle Enterprise Manager

Access Address: http://cdwserver.ce.ufl.edu:5501/em
User ID/Password: sys/trc513, connected as sysdba
User ID/Password: system/trc513, connected as normal

2.3.2 User ID and password for SQL*Plus

User ID/Password, connection string: sys/trc513, steward2
User ID/Password, connection string: system/trc513, steward2
User ID/Password, connection string: steward_owner/trc513, steward2

2.3.3 User ID and password for Oracle Warehouse builder

User ID/Password, connection string: Gator/trc513, steward2

The process flow account ***owf_mgr***, needs be granted execution privileges during the installation. The following command gives ***owf_mgr*** the privilege to upgrade dimension table. This can be executed from SQL*Plus.

>@C:\oracle\product\10.2.0\OWB10_2\owb\rtp\sql\grant_upgrade_privileges.sql owf_mgr

3 The STEWARD ETL process

STEWARD integrates all ETL procedures as processes of the Oracle Warehouse Builder (OWB) Design Center. Users can develop the building blocks (mapping the external files into the STEWARD database tables) using OWB mapping tools. These building blocks can be used to design the ETL process flow. The process flow includes the mappings as a functional block to design and implement the ETL process. These blocks and process flows can be executed using the Oracle Control Center Manager or Scheduler.

3.1 Mapping

In STEWARD, 15 basic mappings are defined. All of them load an external table (virtual database table mapped from the external files) into a dimension table or fact table. These mappings can be executed directly or used as building blocks for the process flow.

3.2 Process Flows

At this time, six main process flows and nine sub-procedures are defined in STEWARD. The ETL process assumes that the traffic data are archived in the predefined location *(C:\STEWARD\TSS_RAW_DATA AND C:\STEWARD\TVT_RAW_DATA).*

3.2.1 Stage1: Pre-transformation

As shown in Figure C-1, the raw data files in the pre-selected folder will be moved to the input data folder for the SunETLUtility. All of the files will be processed internally for the next step. The processing time depends on the number of files and their size. Embedded scripts check the files and move them to the specific locations (C:\Program Files\TRC\SunEtlUtility)

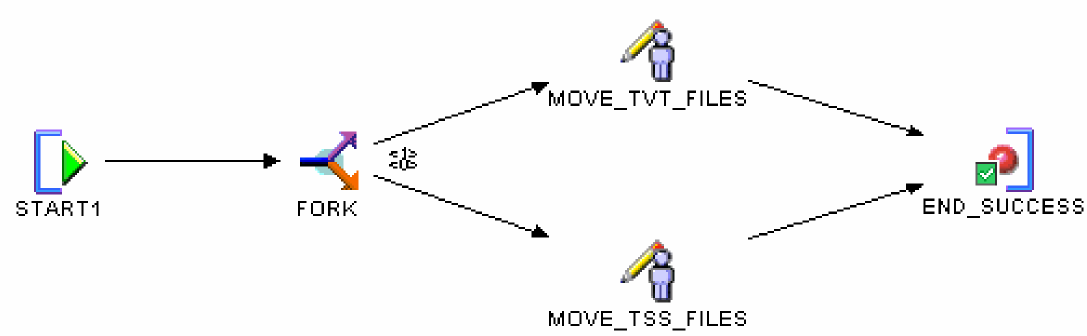


Figure C- 1: Stage1-Pre-transformation

The default input data folder for SunETLUtility are:
>> *C:\PROGRAM FILES\TRC\SunETLUility\ToConvert*

3.2.2 Stage2: Transformation

As shown in Figure C-2, the input TSS and TVT files will be converted to the data format defined for STEWARD. The existence of the input data files will be verified. If the files exist, then the SunEtlUtility will be executed to generate data and reports. If necessary, it will report the errors and quit the process.

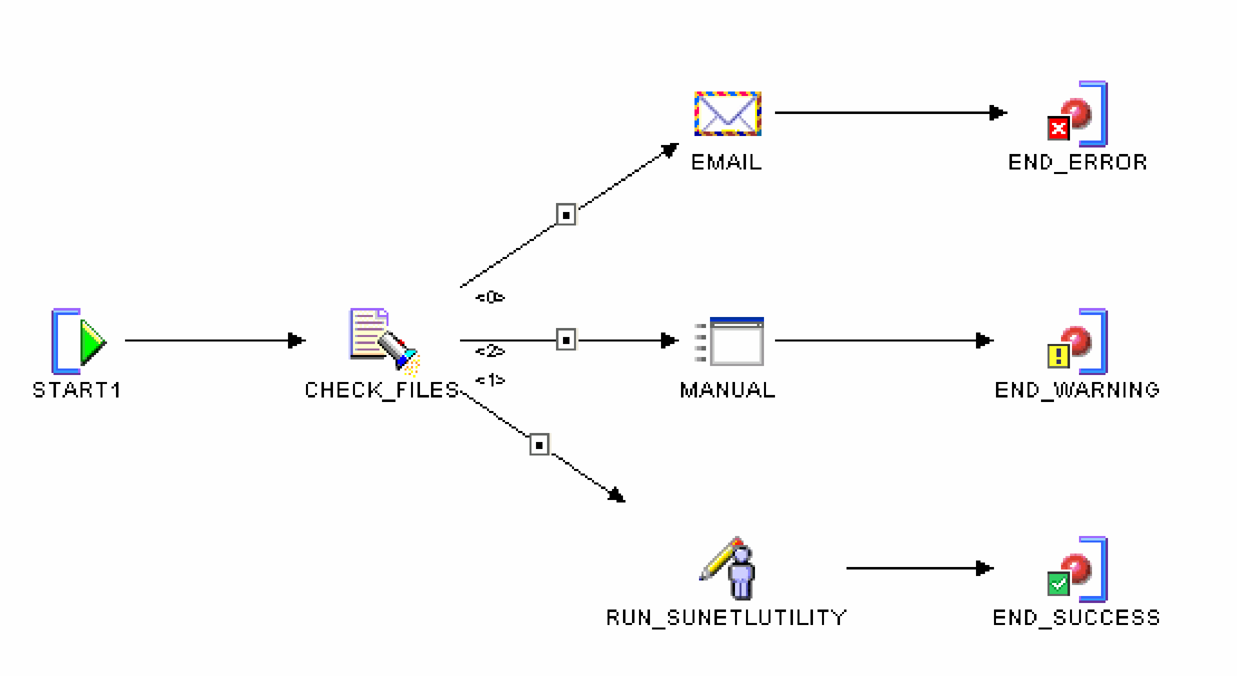


Figure C- 2: Stage2- Transformation

3.2.3 Stage3: Pre-loading

As shown in Figure C-3, all the group files are moved to the dedicated folders for data-loading to the STEWARD database.

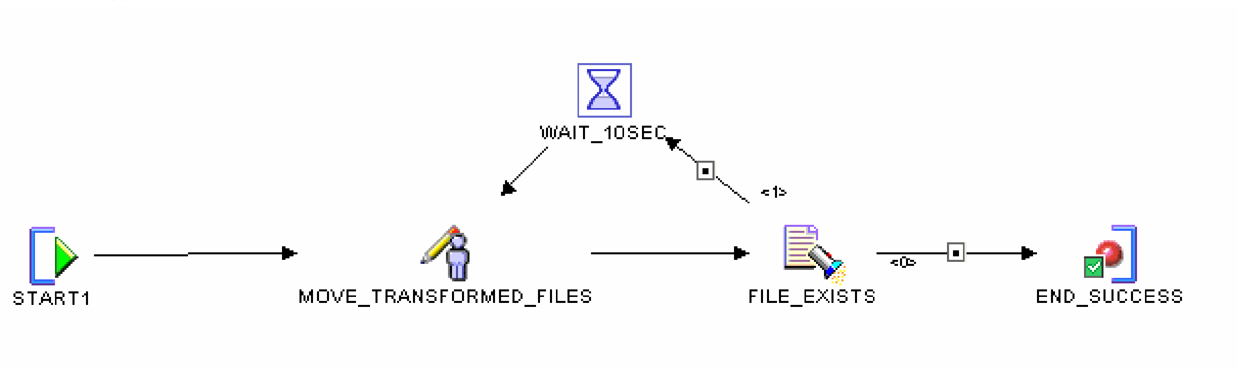


Figure C- 3: Stage3- Pre-loading

3.2.4 Stage4: Loading TSS/TVT data files to STEWARD

As shown in Figure C-4, nine sub-processes are defined (3 file types by 3 grouping intervals). Each sub-process loads one TSS or TVT data file into the STEWARD database. The loading processes are divided into sub-processes to allow individually data file loading.

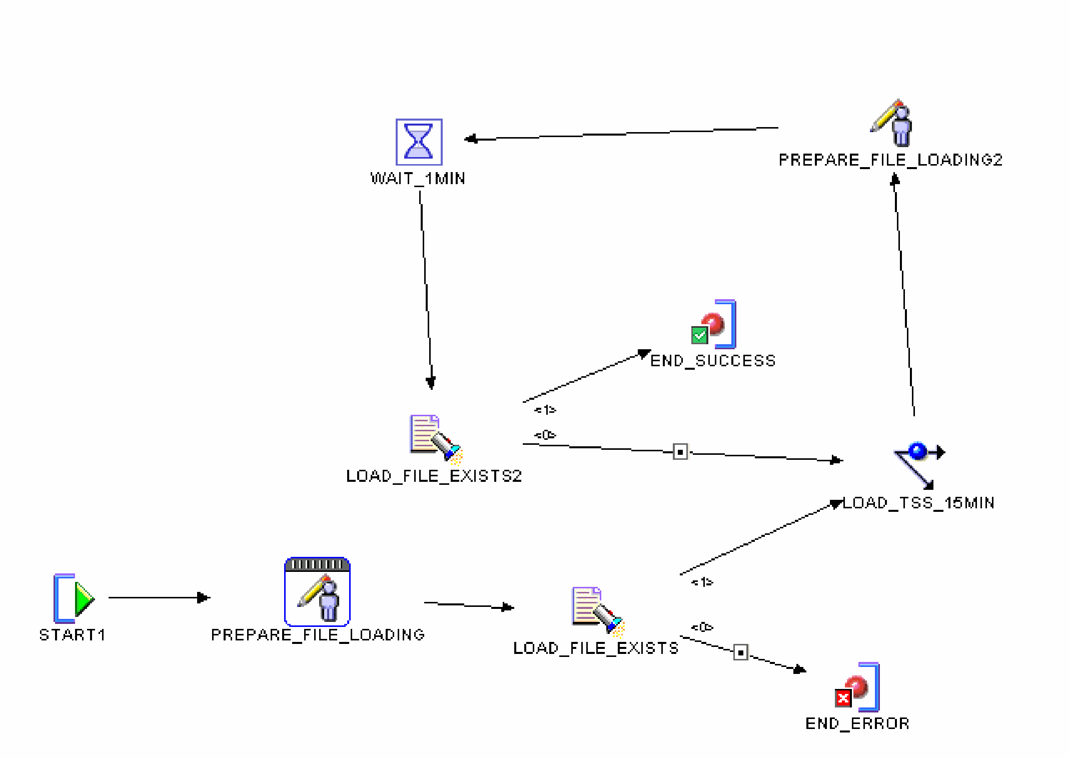


Figure C- 4: Stage4 (sub-process) - Loading the TSS 15Min files into the STEWARD database

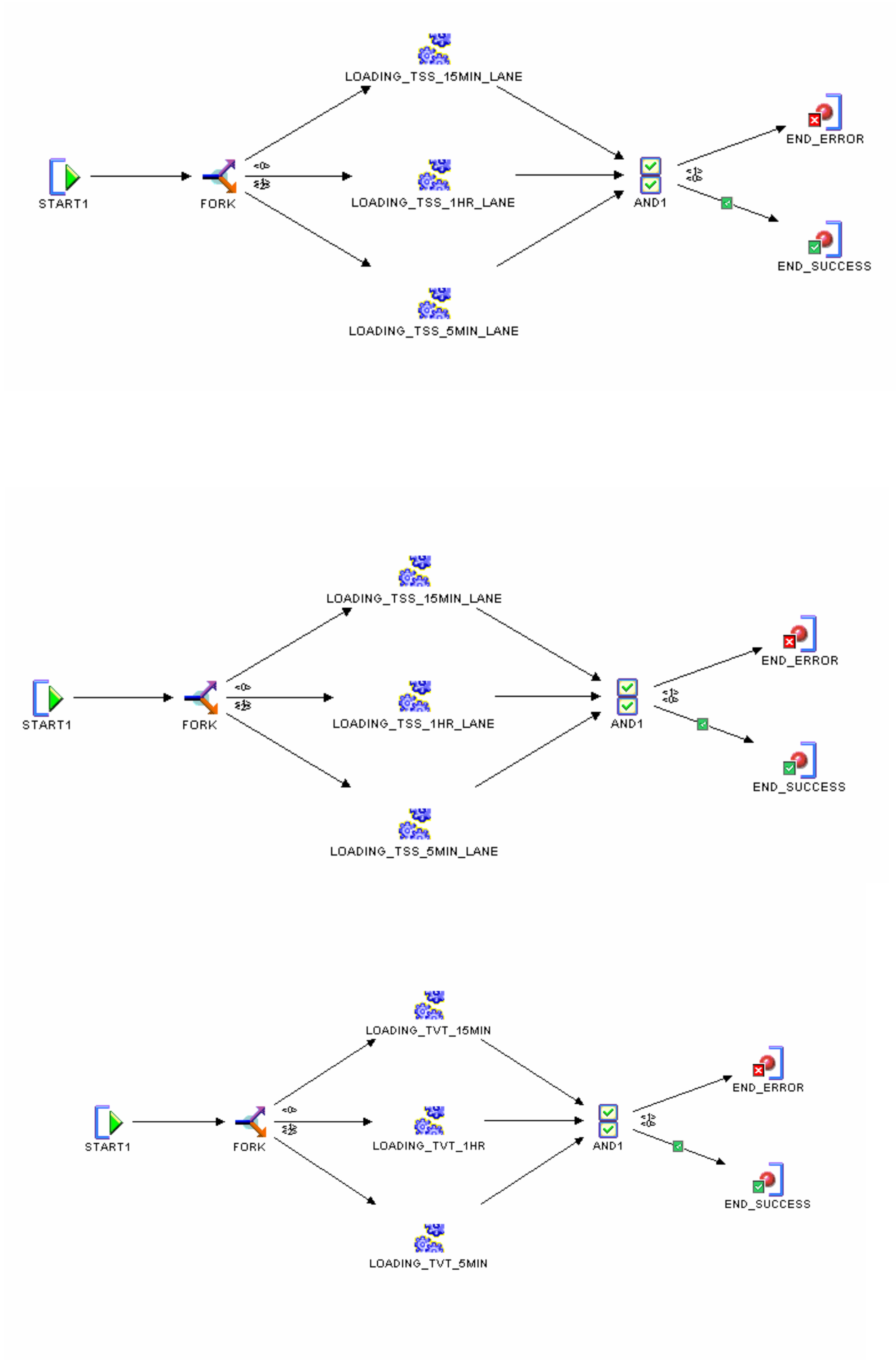


Figure C- 5: Stage4 – Loading the TSS and TVT files into the STEWARD database

As shown in Figure C-4 and Figure C-5, each process can be executed individually or sequentially. In any case, the process can be executed with the scheduler. Using the timer, the operator can coordinate the sequence of the process execution. Most of the processes have the embedded script to work with.

In each step, the process checks the results and decides to proceed or stop if an error happens. The error can be handled by stopping the entire process or sending an automatic notification (email) to the STEWARD operator.

3.3 Control Center Manager

The OWB control center manager is used to execute the mappings or process flows manually. Control Center Manager can be executed from the OWB design center as Figure C-6 shows.

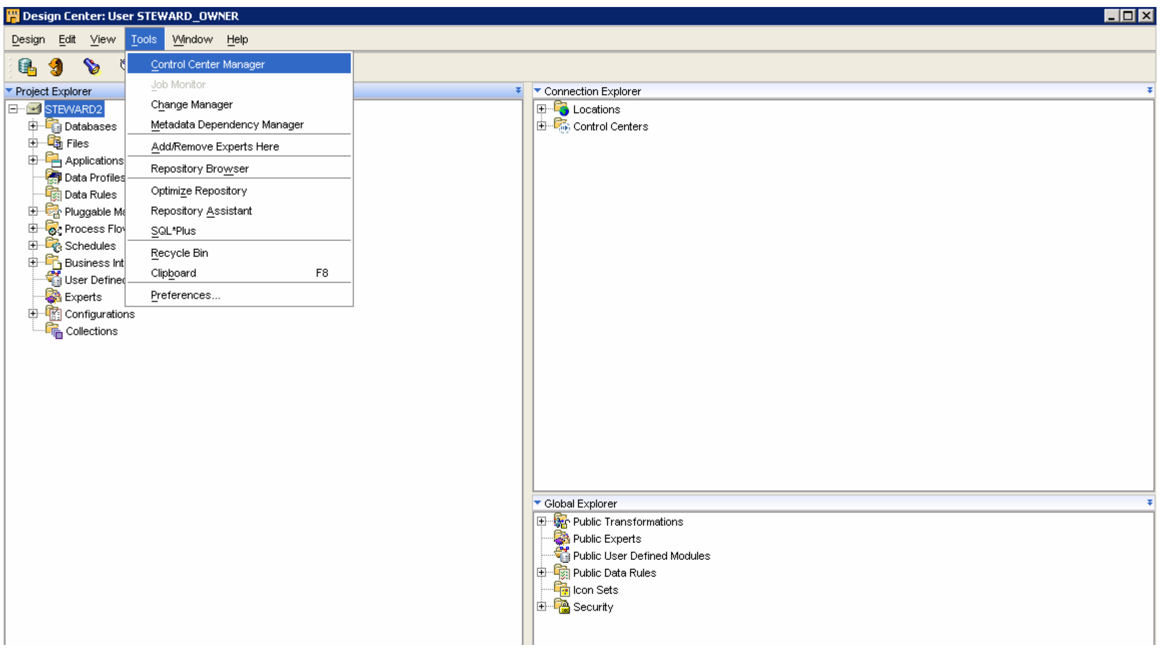


Figure C- 6: Control Center manager from Oracle Warehouse Builder console

To execute the mappings or process flows, select the process in the object window and run the start menu. Figure C-7 below shows that the STAGE4_LOADING_TSS_LANE is selected and start button is ready to run the process.

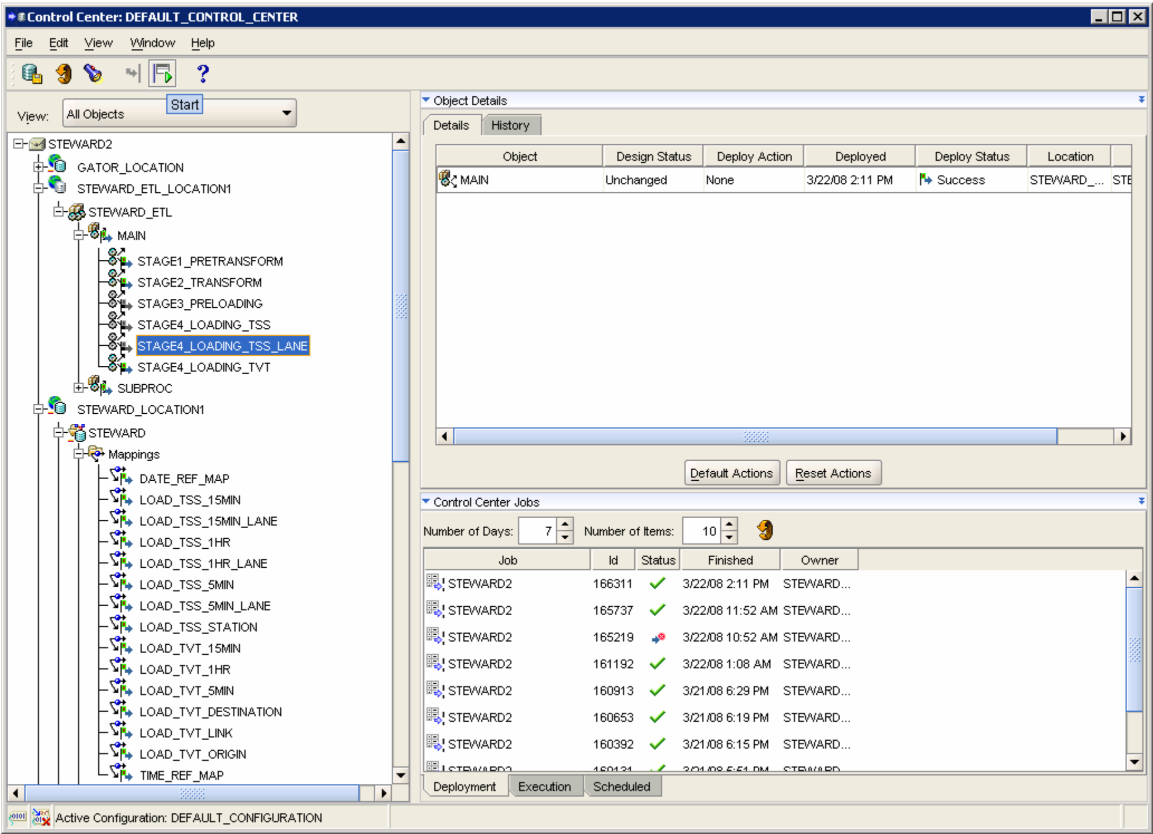


Figure C- 7: File-loading sub-processes

Several issues are involved in the process execution. The Oracle OWB has the known bugs of process flow such that the multiple file loading is not always executed thoroughly. If the number of files to be loaded is not large, loading the files one by one would be sometimes more efficient than loading the entire set of files at once.

Another issue involves the process problems that are running but cannot be controlled (stopped) by the control center manager. In this case, Oracle Workflow module on the Oracle companion CD needs to be installed. This program supports process monitor and provides control to stop the process. This program uses the Apache web server and has problems with MS IIS server.

3.4 Job Scheduling

The OWB design center supports scheduler functions and can be used with the process flow. STEWARD users can run the predefined process flow with this scheduler with the preselected sequences.

3.4.1 Schedule setup

The Operator needs to create the *schedule* module first and associate it with the STEWARD processes to be executed on schedule. The schedule module sets up the start time and the frequency. The start time indicates the time and date when the schedule module starts. The

frequency is used to set the repetition interval from one second to one year. It also includes one time execution.
For example, as illustrated on Figure C-8, schedule module, STAGE1 is defined with the start time of 9/13/2007 at 04:23:15 and the frequency is set as “one time”.

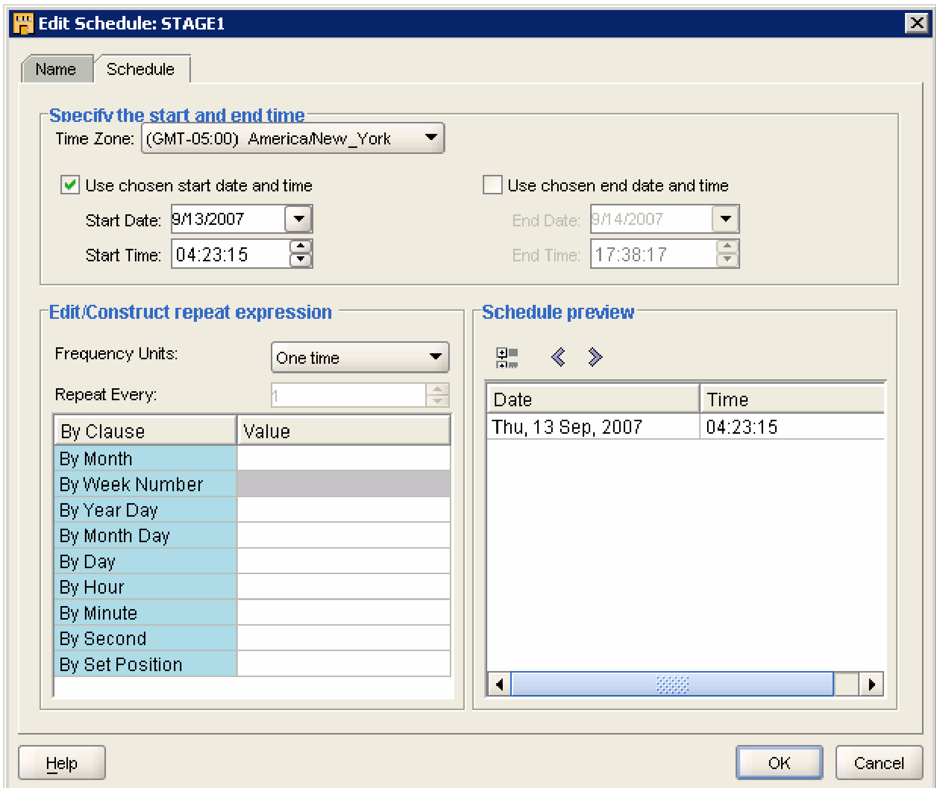


Figure C- 8: Oracle database scheduler

Once this schedule module is designed, it can be attached to the process to invoke the process as a predefined schedule. For example, as illustrated In Figure C-9, process STAGE1_PRETRANSFORM has schedule module STAGE1 as the referred calendar to be executed on its schedule.

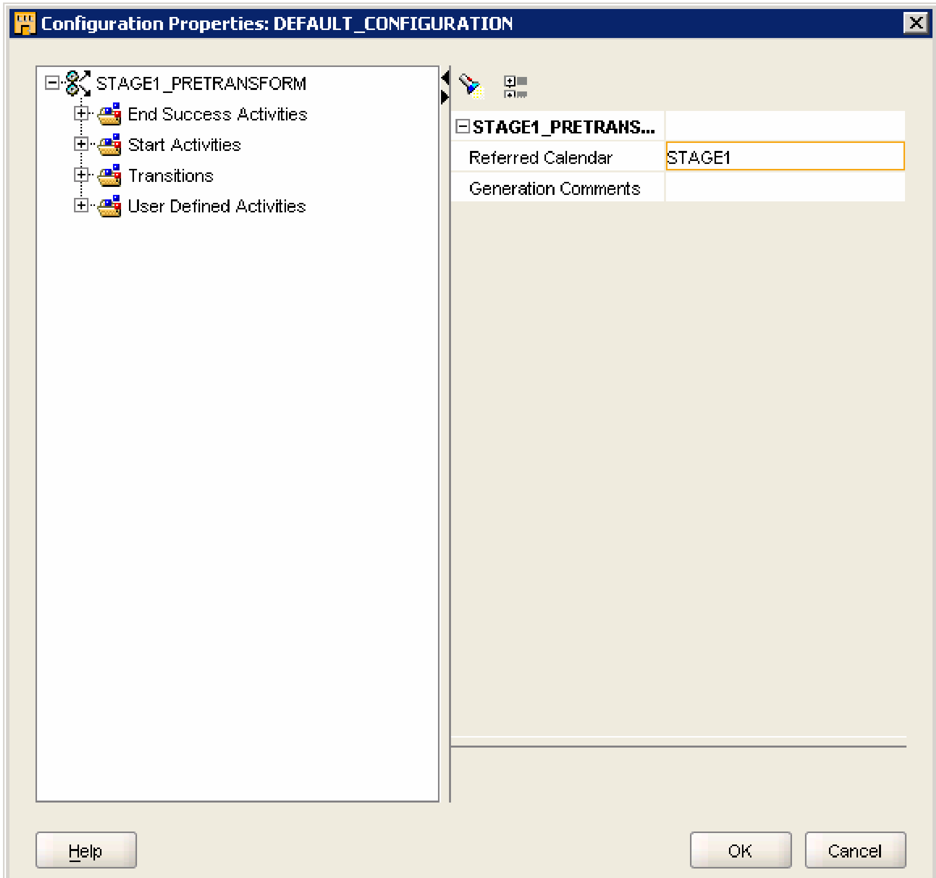


Figure C- 9: Oracle process flow with scheduler setting

One of issues in the scheduler for STEWARD is to estimate the execution time of each process. The transform process could take from 5 minutes to several days depending on the number of data files and their sizes. If the operator sets up the scenario to run the ETL processes in sequences a time-consuming process should be anticipated. Transformation and loading should be done in advance to prevent the next process start before the current process ends.

3.5 Updating the Materialized Views

The materialized views have two options to be updated. First, they can be updated when the dependent tables are updated. Second, they can be updated manually. All of the materialized views in STEWARD are set to be updated manually. There are two reasons for this.

1. It minimizes unexpectedly time-consuming loading process. User loads the data files to STEWARD successfully and updates the materialized views at once.
2. There are dependencies in the materialized views and specific views need to be updated in order. For example, EFF_DEF_LEN_DAY uses materialized view EFF_DEF_LEN_1HR and expects all of the dependent tables to be updated beforehand. At this time, there are four tables that have this kind of dependency.

Materialized views with dependency	Files that need to be pre-existing
EFF_DET_LEN_DAY	EFF_DET_LEN_1HR
LANE_DISTRIBUTION_DAY	LANE_DISTRIBUTION_1HR
PM_LOS_DAY	PM_LOS_5MIN
SYSTEM_LANE_DAY	SYSTEM_LANE_1HR

Materialized views can be updated as follows:

1. In the Oracle warehouse builder, select Tasks → Control Center Manager

2. As shown in Figure C-10, select materialized views in the Control Center Manager's the left pane:
   STEWARD2 → STEWARD_LOCATION1→ STEWARD→Materialized Views

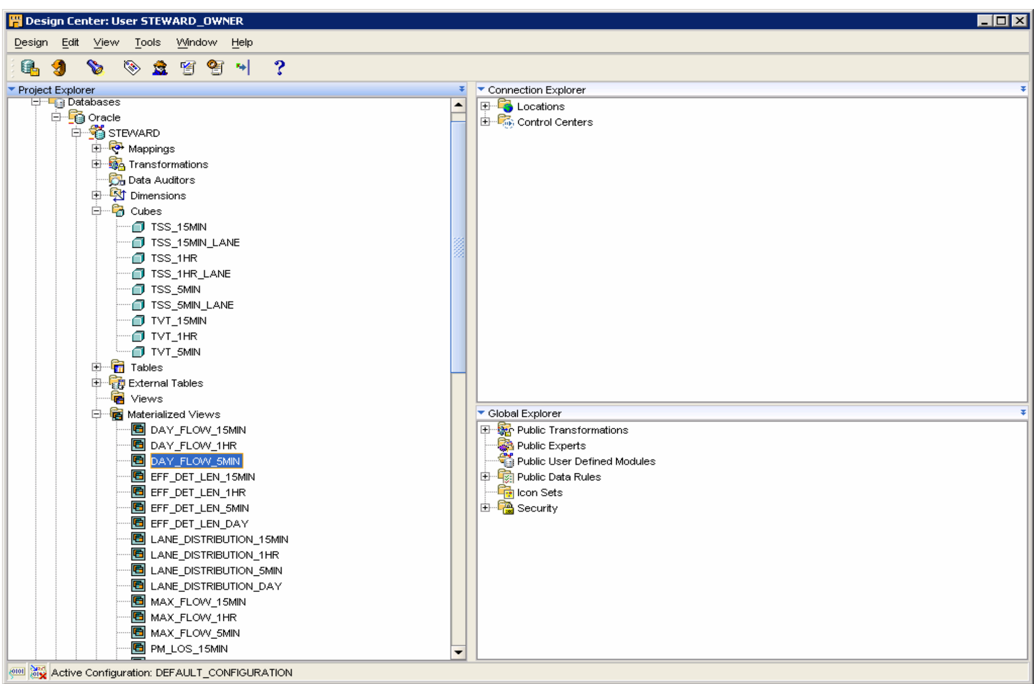


Figure C- 10: Materialized views in the Oracle Control Center Manager

3. As shown in Figure C-11, select the materialized views in the left pane that need to be updated. In the ***Object Details*** window, select the deployment action for each of the materialized views.

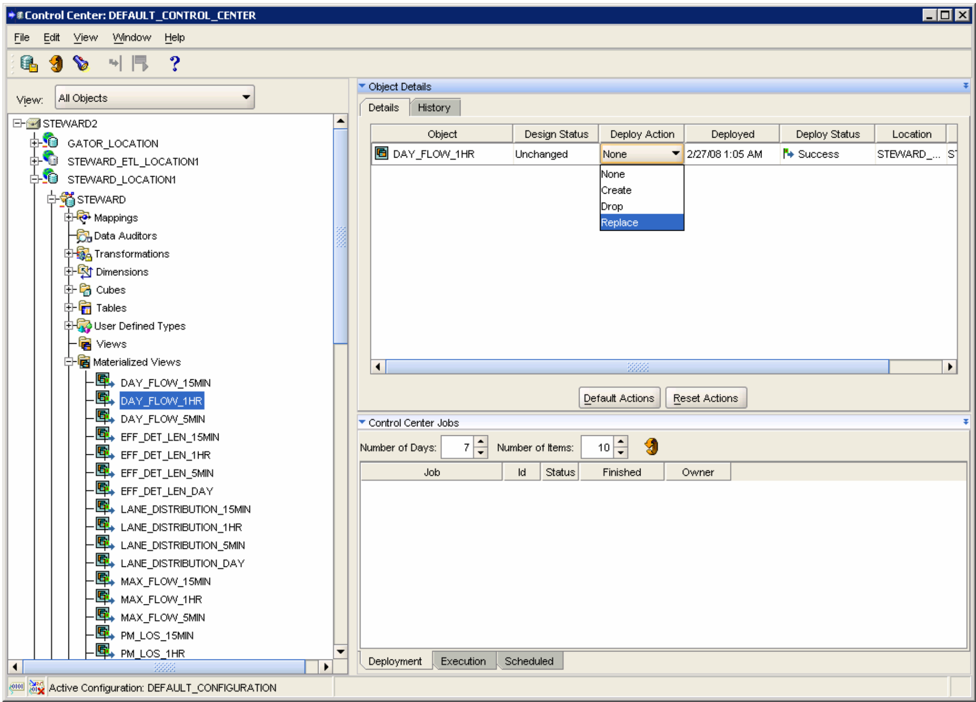


Figure C- 11: Replace materialized views in the Oracle Control Center Manager

As shown in Figure C-12, click the ***Deploy*** button to execute the options.

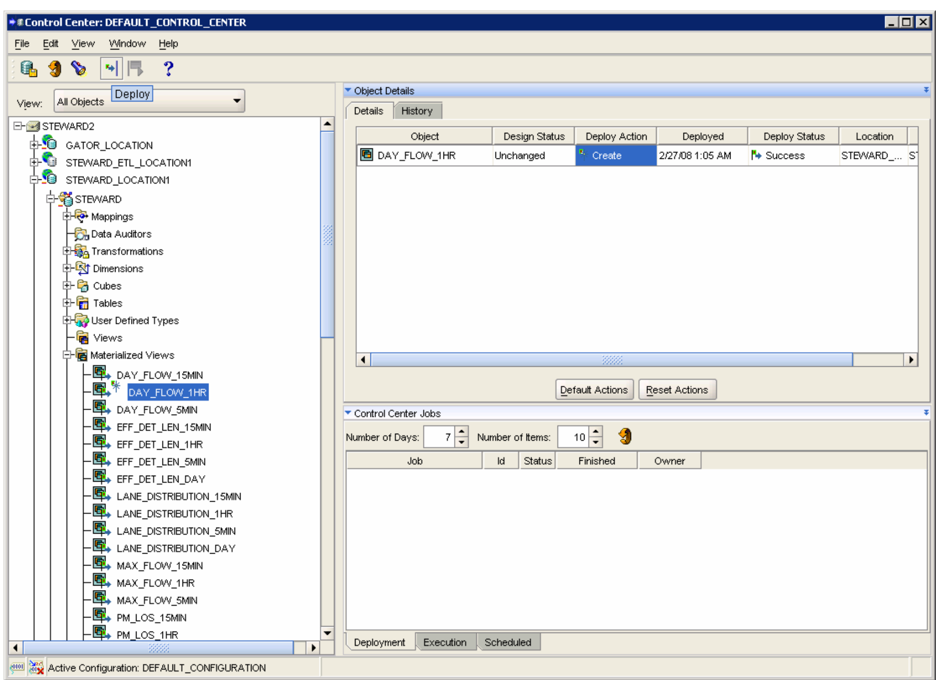


Figure C- 12: Deploy materialized views in Oracle Control Center Manager

4 Exporting Metadata from the STEWARD Database

The Oracle warehouse builder supports the “Export warehouse builder metadata” function. As shown in Figure C-13, select Design → Export → Warehouse builder metadata from the menu.

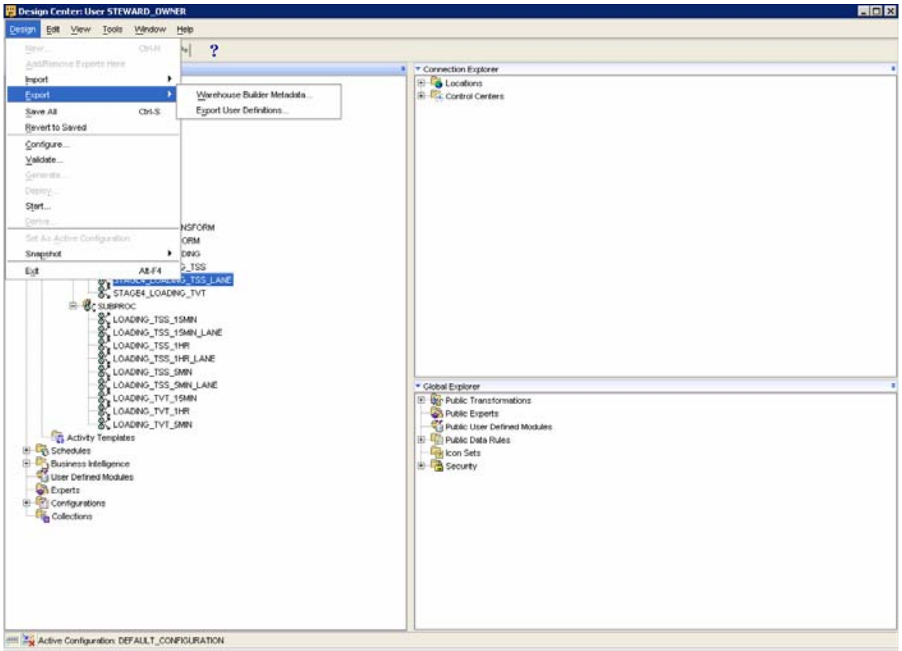


Figure C- 13: Export Warehouse Builder metadata

Check the combo box for ***Export all object dependencies*** and click ***Export***.
As shown in Figure C-14, select the designated location for metadata file and its log file.

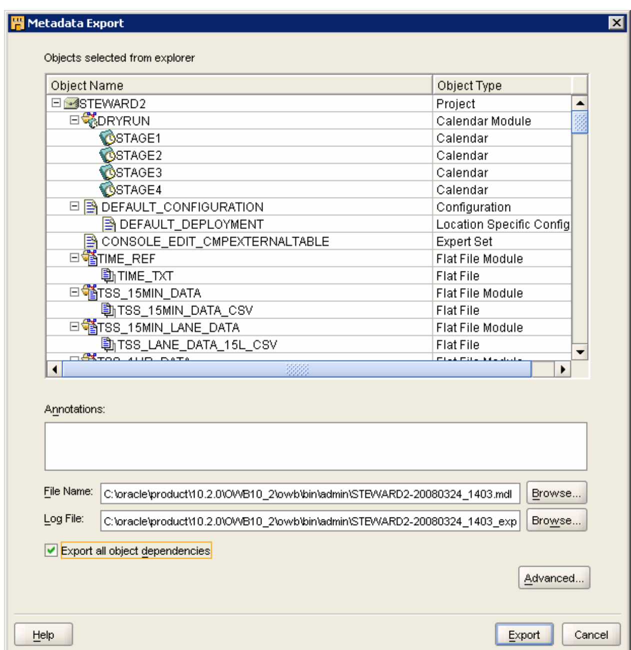


Figure C- 14: Warehouse Builder metadata window

5 STEWARD Database Tables

A description of the STEWARD database tables is presented here to provide a concise reference.

5.1 TSS ETL REPORT

TSS_ETL_REPORTS_EX table (Table C-5) is created to keep a log on ETL reports.

Table C- 5: STEWARD TSS_ETL_REPORTS_EX

Column Name	Data Type
NAME	VARCHAR2(255)
SOURCE	VARCHAR2(255)
PATH	VARCHAR2(255)
CREATION_DATE	DATE
EXPIRATION_DATE	DATE

Examples are as follows. Three files are logged in this example:

NAME	SOURCE	PATH	CREATION_	EXPIRATIO
D2_TSS-09142007-DailyReport.Log	TSS	\steward\files\	01-NOV-07	31-DEC-99
D2_TSS-09152007-DailyReport.Log	TSS	\steward\files\	01-NOV-07	31-DEC-99
D2_TSS-09162007-DailyReport.Log	TSS	\steward\files\	01-NOV-07	31-DEC-99

5.2 TVT ETL REPORT

TVT_ETL_REPORTS_EX table (Table C-6) is created to keep a log on ETL reports.

Table C- 6: STEWARD TVT_ETL_REPORTS_EX

Column Name	Data Type
NAME	VARCHAR2(255)
SOURCE	VARCHAR2(255)
PATH	VARCHAR2(255)
CREATION_DATE	DATE
EXPIRATION_DATE	DATE

Examples are as follows. Three files are logged:

NAME	SOURCE	PATH	CREATION_	EXPIRATIO
D2_TVT-09142007-DailyReport.Log	TSS	\steward\files\	01-NOV-07	31-DEC-99
D2_TVT-09152007-DailyReport.Log	TSS	\steward\files\	01-NOV-07	31-DEC-99
D2_TVT-09162007-DailyReport.Log	TSS	\steward\files\	01-NOV-07	31-DEC-99

5.3 Volume Map and I/O balance

VOLUME_MAP_5MIN, VOLUME_MAP_15MIN, VOLUME_MAP_1HR and VOLUME_MAP_DAY table (Table C-7 and C-8) are created to keep the volume map and I/O balance.

Table names: VOLUME_MAP_5MIN, VOLUME_MAP_15MIN, VOLUME_MAP_1HR

Table C- 7: STEWARD Volume Map and I/O Balance

Column Name	Data Type
DATE1	DATE
TIME	VARCHAR2(25)
FACILITY	NUMBER
STATION_ID	NUMBER
STATION_MP	FLOAT(5)
UPNODE_ID	NUMBER
ENTRY_VOLUME	NUMBER
FWY_VOLUME	NUMBER
EXIT_VOLUME	NUMBER
LINK_INPUT	NUMBER
LINK_OUTPUT	NUMBER
DIFFERENCE	NUMBER
PCNT_DIFF	NUMBER

Table names: VOLUME_MAP_DAY

Table C- 8: STEWARD Volume Map and I/O Balance for Daily Table

Column Name	Data Type
DATE1	DATE
FACILITY	NUMBER
STATION_ID	NUMBER
STATION_MP	FLOAT(5)
UPNODE_ID	NUMBER
ENTRY_VOLUME	NUMBER
FWY_VOLUME	NUMBER
EXIT_VOLUME	NUMBER
LINK_INPUT	NUMBER
LINK_OUTPUT	NUMBER
DIFFERENCE	NUMBER
PCNT_DIFF	NUMBER

5.4 Performance Measures and LOS

PM_LOS_5MIN, PM_LOS_15MIN, PM_LOS_1HR and PM_LOS_DAY table (Table C-9 and C-10) are created to keep the Performance Measures and LOS.

Table names: PM_LOS_5MIN, PM_LOS_15MIN, PM_LOS_1HR

Table C- 9: STEWARD Performance Measures and LOS

Column Name	Data Type
DATE1	DATE
TIME	VARCHAR2(25)
FACILITY	NUMBER
STATION_ID	NUMBER
STATION_DESC	VARCHAR2(200)
UPNODE_ID	NUMBER
UPNODE_DESC	VARCHAR2(200)
STATION_MP	FLOAT(5)
LENGTH	NUMBER
LINK_VOL	NUMBER
NUM_OF_LANES	NUMBER
VOL_PER_LANE	NUMBER
VEH_MI	NUMBER
VEH_HR	NUMBER
SPEED	FLOAT(5)
DELAY	NUMBER
KENERGY	NUMBER
DENSITY	NUMBER
VOL_TO_CAP	NUMBER
LOS	CHAR(1)

Table names: PM_LOS_DAY

Table C- 10: STEWARD Performance Measures and LOS with Daily Table

Column Name	Data Type
DATE1	DATE
FACILITY	NUMBER
STATION_ID	NUMBER
STATION_DESC	VARCHAR2(200)
UPNODE_ID	NUMBER
UPNODE_DESC	VARCHAR2(200)
STATION_MP	FLOAT(5)
LENGTH	NUMBER
LINK_VOL	NUMBER
NUM_OF_LANES	NUMBER
VOL_PER_LANE	NUMBER
VEH_MI	NUMBER
VEH_HR	NUMBER
SPEED	FLOAT(5)
DELAY	NUMBER
KENERGY	NUMBER
F_PLUS	NUMBER
E_PLUS	NUMBER
D_PLUS	NUMBER
C_PLUS	NUMBER

5.5 System Lane Volume

SYSTEM_LANE_5MIN, SYSTEM_LANE_15MIN, SYSTEM_LANE_1HR and SYSTEM_LANE_DAY (Table C-11) are created to keep the system lane volume.

Table names: SYSTEM_LANE_5MIN, SYSTEM_LANE_15MIN, SYSTEM_LANE_1HR, SYSTEM_LANE_DAY

Table C- 11: STEWARD System Lane Volume

Column Name	Data Type
DATE1	DATE
TIME	VARCHAR2(25)
FACILITY	NUMBER
STATION_ID	NUMBER
STATION_DESC	VARCHAR2(200)
STATION_MP	FLOAT(5)
LANE1_VOL	NUMBER
LANE1_SPD	FLOAT(5)
LANE1_OCC	FLOAT(5)
LANE2_VOL	NUMBER
LANE2_SPD	FLOAT(5)
LANE2_OCC	FLOAT(5)
LANE3_VOL	NUMBER
LANE3_SPD	FLOAT(5)
LANE3_OCC	FLOAT(5)
LANE4_VOL	NUMBER
LANE4_SPD	FLOAT(5)
LANE4_OCC	FLOAT(5)
LANE5_VOL	NUMBER
LANE5_SPD	FLOAT(5)
LANE5_OCC	FLOAT(5)
LANE6_VOL	NUMBER
LANE6_SPD	FLOAT(5)
LANE6_OCC	FLOAT(5)
ONRAMP1_VOL	NUMBER
ONRAMP1_SPD	FLOAT(5)
ONRAMP1_OCC	FLOAT(5)
ONRAMP2_VOL	NUMBER
ONRAMP2_SPD	FLOAT(5)
ONRAMP2_OCC	FLOAT(5)
ONRAMP3_VOL	NUMBER
ONRAMP3_SPD	FLOAT(5)
ONRAMP3_OCC	FLOAT(5)
OFFRAMP1_VOL	NUMBER
OFFRAMP1_SPD	FLOAT(5)
OFFRAMP1_OCC	FLOAT(5)
OFFRAMP2_VOL	NUMBER
OFFRAMP2_SPD	FLOAT(5)
OFFRAMP2_OCC	FLOAT(5)
OFFRAMP3_VOL	NUMBER
OFFRAMP3_SPD	FLOAT(5)
OFFRAMP3_OCC	FLOAT(5)

5.6 Station Lane Volume and Distribution

LANE_DISTRIBUTION_5MIN, LANE_DISTRIBUTION_15MIN, LANE_DISTRIBUTION_1HR and LANE_DISTRIBUTION_DAY (Table C-12) are created to keep the station lane volume and distribution.

Table names: LANE_DISTRIBUTION_5MIN, LANE_DISTRIBUTION_15MIN, LANE_DISTRIBUTION_1HR, LANE_DISTRIBUTION_DAY

Table C- 12: STEWARD Station Lane Volume and Distribution

Column Name	Data Type
DATE1	DATE
TIME	VARCHAR2(25)
FACILITY	NUMBER
STATION_ID	NUMBER
STATION_DESC	VARCHAR2(200)
STATION_MP	FLOAT(5)
LANE1_VOL	NUMBER
LANE2_VOL	NUMBER
LANE3_VOL	NUMBER
LANE4_VOL	NUMBER
LANE5_VOL	NUMBER
LANE6_VOL	NUMBER
TOTAL	NUMBER
BALANCE	NUMBER

5.7 Daily Flow

DAY_FLOW_5MIN, DAY_FLOW_15MIN and DAY_FLOW_1HR (Table C-13) are created to keep the daily flow.

Table names: DAY_FLOW_5MIN, DAY_FLOW_15MIN, DAY_FLOW_1HR

Table C- 13: STEWARD Daily Flow

Column Name	Data Type
DATE1	DATE
TIME	VARCHAR2(25)
FACILITY	NUMBER
STATION_ID	NUMBER
STATION_DESC	VARCHAR2(200)
STATION_MP	FLOAT(5)
LANE1_FLOW_RATE	NUMBER
LANE2_FLOW_RATE	NUMBER
LANE3_FLOW_RATE	NUMBER
LANE4_FLOW_RATE	NUMBER
LANE5_FLOW_RATE	NUMBER
LANE6_FLOW_RATE	NUMBER

5.8 Maximum Flow

MAX_FLOW_5MIN, MAX_FLOW_15MIN and MAX_FLOW_1HR (Table C-14) are created to keep the max flow.

Table names: MAX_FLOW_5MIN, MAX_FLOW_15MIN, MAX_FLOW_1HR

Table C- 14: STEWARD Maximum Flow

Column Name	Data Type
DATE1	DATE
FACILITY	NUMBER
STATION_ID	NUMBER
STATION_DESC	VARCHAR2(200)
STATION_MP	FLOAT(5)
LANE_NUM	NUMBER
MAX_FLOW	NUMBER
MAX_TIME	VARCHAR2(25)
MAX_GAP	NUMBER

5.9 Effective Detector Length

EFF_DET_LEN_5MIN, EFF_DET_LEN_15MIN, EFF_DET_LEN_1HR and EFF_DET_LEN_DAY (Table C-15) are created to keep the effective detector length.

Table names: EFF_DET_LEN_5MIN, EFF_DET_LEN_15MIN, EFF_DET_LEN_1HR, EFF_DET_LEN_DAY

Table C- 15: STEWARD Effective Detector Length

Column Name	Data Type
DATE1	DATE
TIME	VARCHAR2(25)
FACILITY	NUMBER
STATION_ID	NUMBER
STATION_DESC	VARCHAR2(200)
STATION_MP	FLOAT(5)
LANE1_VOL	NUMBER
LANE1_SPD	FLOAT(5)
LANE1_OCC	FLOAT(5)
LANE1_EFF_DET_LENGTH	NUMBER
LANE2_VOL	NUMBER
LANE2_SPD	FLOAT(5)
LANE2_OCC	FLOAT(5)
LANE2_EFF_DET_LENGTH	NUMBER
LANE3_VOL	NUMBER
LANE3_SPD	FLOAT(5)
LANE3_OCC	FLOAT(5)
LANE3_EFF_DET_LENGTH	NUMBER
LANE4_VOL	NUMBER
LANE4_SPD	FLOAT(5)
LANE4_OCC	FLOAT(5)
LANE4_EFF_DET_LENGTH	NUMBER
LANE5_VOL	NUMBER
LANE5_SPD	FLOAT(5)
LANE5_OCC	FLOAT(5)
LANE5_EFF_DET_LENGTH	NUMBER
LANE6_VOL	NUMBER
LANE6_SPD	FLOAT(5)
LANE6_OCC	FLOAT(5)
LANE6_EFF_DET_LENGTH	NUMBER

5.10 TVT MAX Delay

TVT_MAX_DELAY_5MIN, TVT_ MAX_DELAY_15MIN and TVT_ MAX_DELAY_1HR (Table C-16) are created to keep the effective detector length.

Table names: TVT_MAX_DELAY_5MIN, TVT_ MAX_DELAY_15MIN, TVT_ MAX_DELAY_1HR

Table C- 16: STEWARD TTS Maximum Delay

Column Name	Data Type
DATE1	DATE
ORIGIN	VARCHAR2(40)
DESTINATION	VARCHAR2(40)
FF_TT	FLOAT(5)
AVG_TT	FLOAT(5)
AVG_DELAY	FLOAT(5)
MAX_DELAY	FLOAT(5)
MAX_DELAY_TIME	VARCHAR2(25)

5.11 TVT MAX TT

TVT_DELAY_5MIN, TVT_DELAY_15MIN and TVT_DELAY_1HR (Table C-17) are created to keep the effective detector length.

Table names: TVT_DELAY_5MIN, TVT_DELAY_15MIN, TVT_DELAY_1HR

Table C- 17: STEWARD TTS Maximum Travel Time

Column Name	Data Type
DATE1	DATE
ORIGIN	VARCHAR2(40)
DESTINATION	VARCHAR2(40)
FF_TT	FLOAT(5)
AVG_TT	FLOAT(5)
AVG_DELAY	FLOAT(5)
MAX_DELAY	FLOAT(5)
MAX_DELAY_TIME	VARCHAR2(25)

6 STEWARD Materialized Views and their Dependent Tables

Table C-18 shows the materialized views for TSS data used in Steward and its dependent tables.

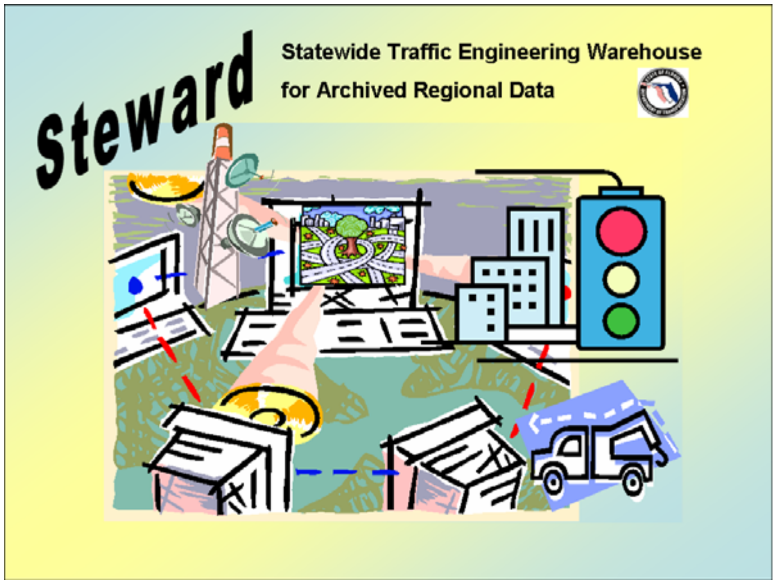
Table C- 18: STEWARD Materialized Views for TSS Data

Materialized views	Dependent tables	Dependent Materialized views
DAY_FLOW_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN_LANE	None
DAY_FLOW_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR_LANE	None
DAY_FLOW_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN_LANE	None
EFF_DET_LEN_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN_LANE	None
EFF_DET_LEN_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR_LANE	None
EFF_DET_LEN_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN_LANE	None
EFF_DET_LEN_DAY	None	EFF_DET_LEN_1HR
LANE_DISTRIBUTION_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN_LANE	None
LANE_DISTRIBUTION_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR_LANE	None
LANE_DISTRIBUTION_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN_LANE	None
LANE_DISTRIBUTION_DAY	None	LANE_DISTRIBUTION_1HR
MAX_FLOW_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN_LANE	None
MAX_FLOW_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR_LANE	None
MAX_FLOW_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN_LANE	None
PM_LOS_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN	None
PM_LOS_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR	None
PM_LOS_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN	None
PM_LOS_DAY	None	PM_LOS_5MIN
SYSTEM_LANE_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN_LANE	None
SYSTEM_LANE_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR_LANE	None
SYSTEM_LANE_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN_LANE	None
SYSTEM_LANE_DAY	None	SYSTEM_LANE_1HR
VOLUME_MAP_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN	None
VOLUME_MAP_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR	None
VOLUME_MAP_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN	None
VOLUME_MAP_DAY	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR	None

Table C-19 shows the materialized views for TVT data used in Steward and its dependent tables.

Table C- 19: STEWARD Materialized Views for TVT Data

Materialized views	Dependent tables	Dependent Materialized views
TVT_MAX_DELAY_15MIN	TVT_ORIGIN, TVT_DESTINATION, TVT_LINK, DATE_REF, TVT_LINK, TIME_REF, TVT_15MIN	None
TVT_MAX_DELAY_1HR	TVT_ORIGIN, TVT_DESTINATION, TVT_LINK, DATE_REF, TVT_LINK, TIME_REF, TVT_1HR	None
TVT_MAX_DELAY_5MIN	TVT_ORIGIN, TVT_DESTINATION, TVT_LINK, DATE_REF, TVT_LINK, TIME_REF, TVT_5MIN	None
TVT_MAX_TT_15MIN	TVT_ORIGIN, TVT_DESTINATION, TVT_LINK, DATE_REF, TVT_LINK, TIME_REF, TVT_15MIN	None
TVT_MAX_TT_1HR	TVT_ORIGIN, TVT_DESTINATION, TVT_LINK, DATE_REF, TVT_LINK, TIME_REF, TVT_1HR	None
TVT_MAX_TT_5MIN	TVT_ORIGIN, TVT_DESTINATION, TVT_LINK, DATE_REF, TVT_LINK, TIME_REF, TVT_5MIN	None



Development of a Central Data Warehouse for Statewide ITS and Transportation Data in Florida

Phase II: Proof of Concept Final Report

Appendix D
Instructions and Guidance for Internet Access to STEWARD Archived Data

Prepared for the Florida Department of Transportation
By the University of Florida Transportation Research Center

Contract # BD545, RPWO # 37
UF Project 051449

Revision Date: April 15, 2008

Table of Contents

List of Figures

List of Tables

1 Introduction

While some use of Statewide Traffic Engineering Warehouse for Regionally Archived Data (STEWARD) will be made within FDOT by accessing the databases directly, most users in the future will gain access to the archived data via the internet. This document describes the internet based features of STEWARD from the perspective of a user who seeks to query the database and produce reports via the Internet.

2 Overview of the STEWARD Web Interface

The STEWARD web site has been developed for general users to access and retrieve the STEWARD data. The web interface allows users to access the database remotely, to retrieve the specific data easily and to download the data to the local computer for further analysis.

At this time, the STEWARD web site can be accessed from the following address: http://cdwserver.ce.ufl.edu/steward/index.html

3 STEWARD Web Architecture

The STEWARD web site consists of three main categories: Overview, Resources and District Data and Reports. The overall architecture is shown in Figure D-1.

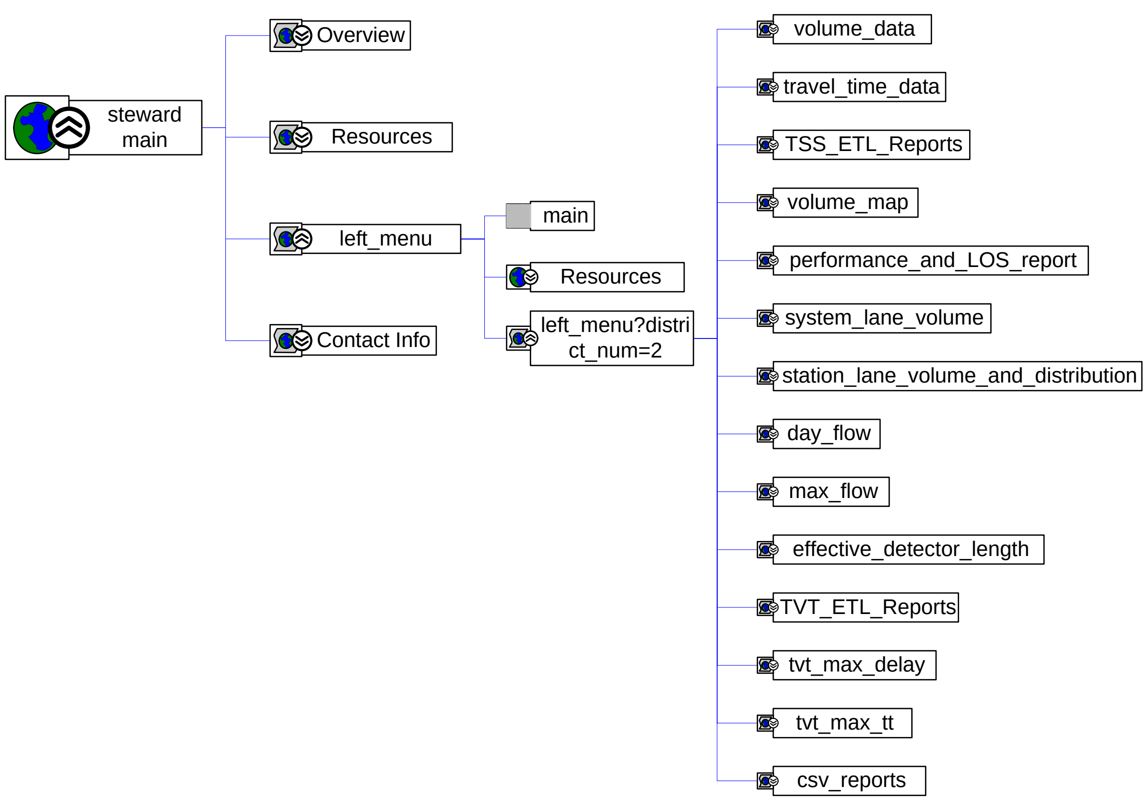


Figure D- 1: STEWARD web architecture

The Overview page provides a general description of the STEWARD project. The Resource page includes STEWARD project reports and utility programs. The SunETLUtility transforms data files from the SunGuide data archive to the STEWARD database format. It executes data verification, data aggregation in time (5, 15 and 60 min intervals) and in space (lane-level and station-level data) and also generates reports. The MPConverter utility converts longitude/latitude data into state/county mile post or vice- versa. The results were validated with GPS data on I-95. It supports I-95 data conversion in Florida only.

From the traffic data and report section, the user can access traffic data (volume/speed/occupancy), travel time data as well as several reports and maps. STEWARD supports system level reports, such as TSS extraction transformation and loading (ETL) Reports, Volume Map/IO Balance, Performance Measure/LOS, System Lane Volume, Station Lane Volume/Distribution, Day Flow, Max Flow, Effective Detector Length, TVT ETL Reports, TVT Max Delay, TVT Max Travel Time, CSV Reports. TSS detector map and TVT detector map are supports for the user convenience.

4 STEWARD Web Pages

This section will review each STEWARD web page and describe its specific functions.

4.1 STEWARD Main page –STEWARD Overview

The STEWARD main page includes two panes as Figure D-2. The main pane displays a brief description of the STEWARD project, objectives and tasks. The left pane is used to navigate to the STEWARD Overview, Resources and District Data/Reports sections.

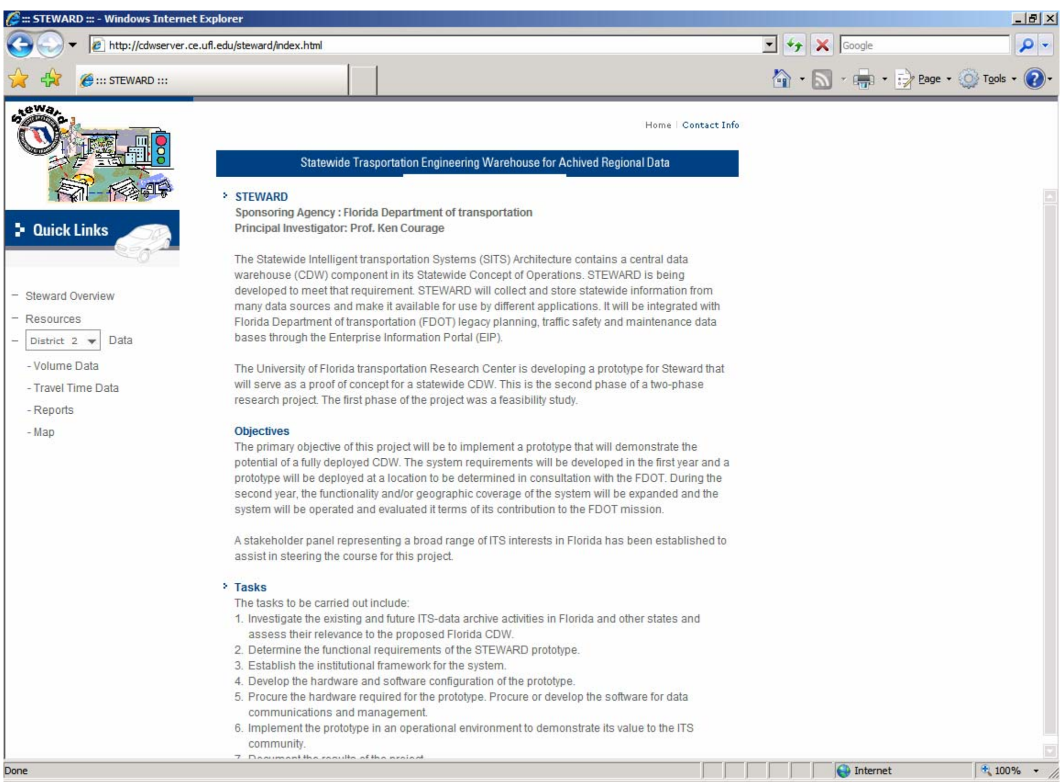


Figure D- 2: STEWARD overview web page

4.2 STEWARD Resources

The STEWARD resource page includes a report section and utilities section as Figure D-3. The report section includes four project progress reports. The utility programs include SunETLUtility and MPConverter. Both include installation programs and descriptive document as a user's guide. Even though these programs are developed as part of STEWARD system these could be executable as stand-alone programs in the Windows environment. The full documentation for these programs is presented in Appendix G of the STEWARD Final Report.

The SunETLUtility performs the transformation process in STEWARD system. The program reads the data files from SunGuide archive files, verifies and aggregates the data and generates reports. The output of the program is used as input to the STEWARD database

The MPConverter converts longitude/latitude data into state/county milepost or vice-versa. This program was validated using GPS data on I-95 in Volusia County. The data conversion is limited to I-95 in Florida only.

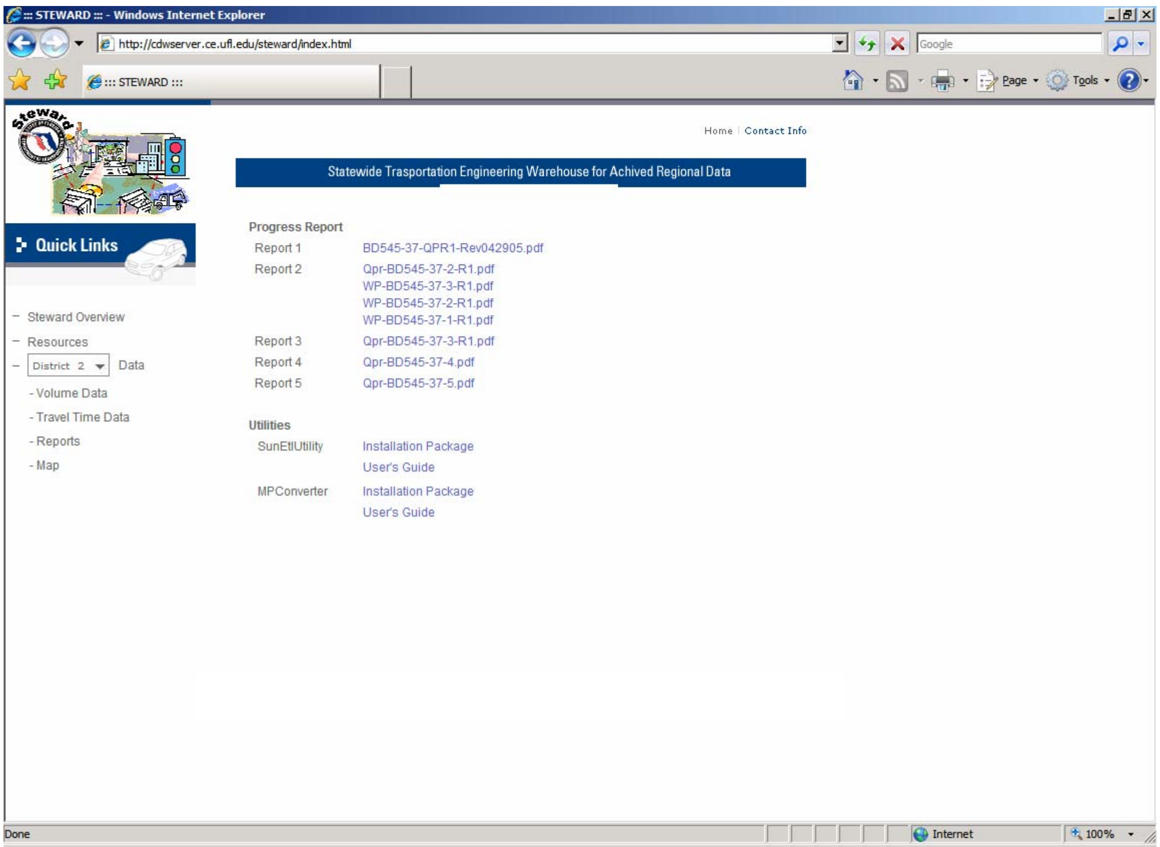


Figure D- 3: STEWARD resources web page

4.3 District Data and Reports

When the District is selected in the left pane, the user can access one of four items (Volume data, Travel time data, Reports, and Map) for the selected District. At this moment, only District 2 data are available. All other District choices are inaccessible.

The structures of each web page are almost identical. In this document, the first page, *volume data* will be described in detail and the other pages will be treated in less detail when their content overlaps the *volume data* pages.

4.3.1 Volume data

As shown in Figure D-4, the user can access traffic data (Volume, Speed and Occupancy) data from this page. The content comes from TSS detector data from the SunGuide system. These data are aggregated into 5min, 15min and 1hr data in the STEWARD database.

The user can query the traffic data with the specific conditions as follows:

- The date range can be selected by the start date and end date. The date in calendar has a bold font if the data on that day are available. For example, STEWARD database archived traffic volume data from Jun. 28th, 2007 to Feb. 6th, 2008 and Figure D-6 shows bold fonts for those dates. The start date can be selected by clicking on that date on the calendar. Then the color of the selected date will be changed to red.
- The daily time range can be specified by the starting and ending hour.
- The aggregation level can be specified in terms of the day of the week (weekday, weekend, all days, Mondays, Tuesdays, Wednesdays, Thursdays, Fridays, Saturdays and Sundays) and aggregation of the data in 5min, 15min and 1hr. The defaults are all days and 15min.
- Facility and Stations can also be selected. In case of District 2 volume data, I-95 is assumed as one facility and user can select stations. The station selection allows multiple selections by using the shift key or mouse.

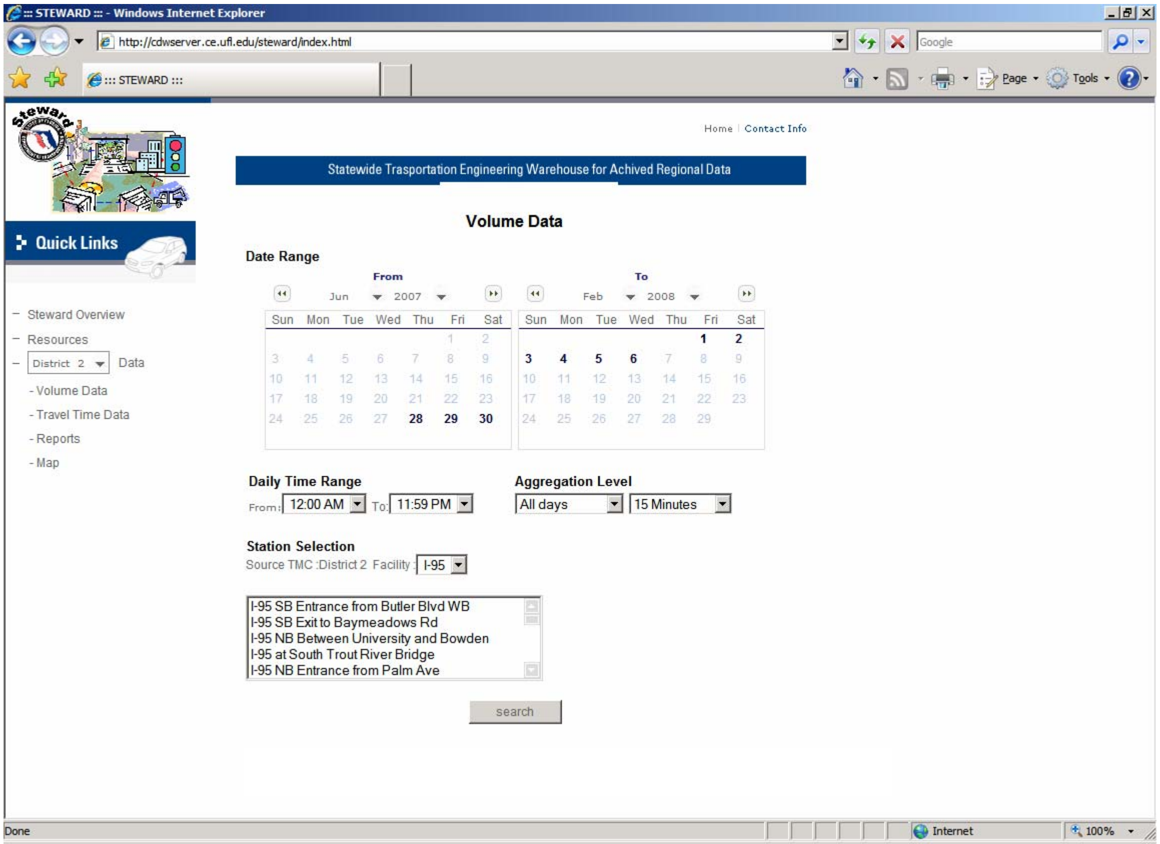


Figure D- 4: STEWARD volume map web page

Figure D-5 shows the resulting screen when the user selects the following criteria:

Start date: Nov. 23rd, 2007
End date: Nov. 28
Time range: 8:00AM ~9:59AM
Day of the week: Monday
Aggregation level: 1hour
Stations: I-95 SB Between Univ and Bowden,
I-95 SB Entrance from Bowden Rd
I-95 SB South of Univ Blvd

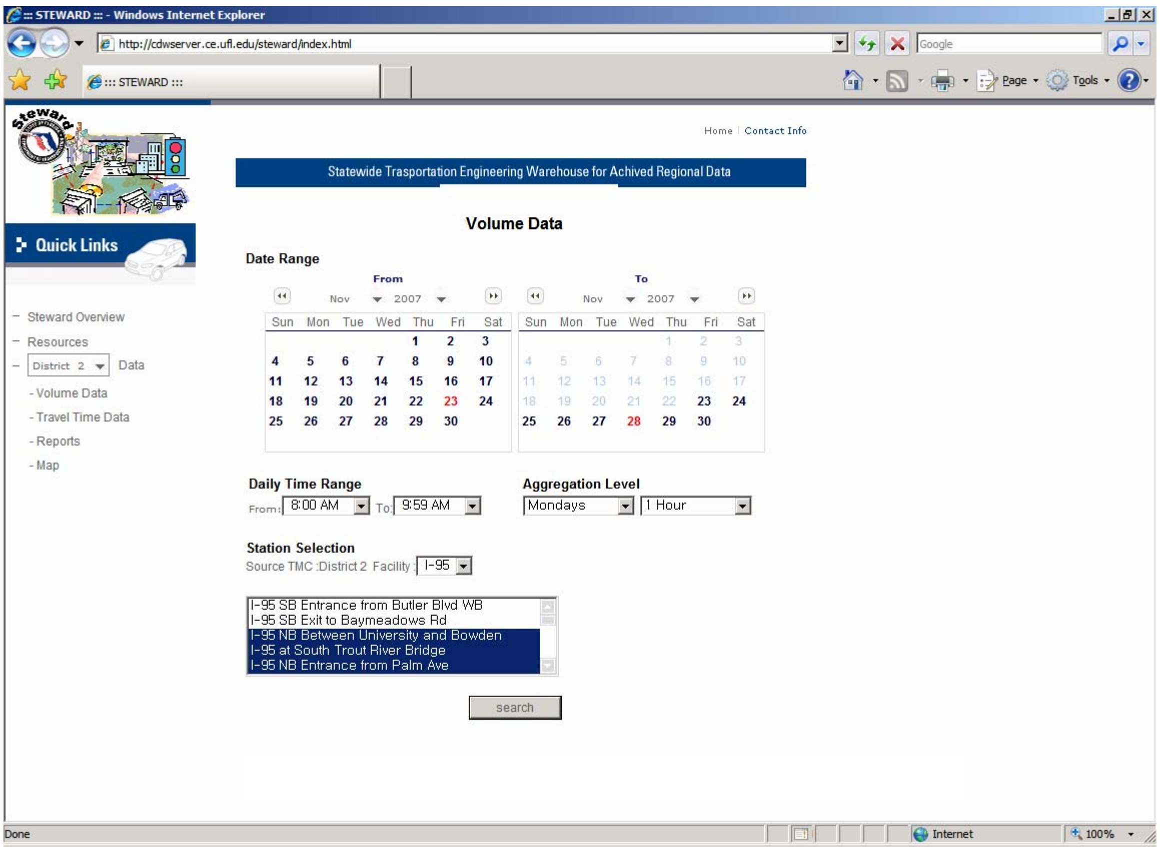


Figure D- 5: STEWARD volume map web page (interactive)

When the ***Search*** button is clicked, a new popup window will appear as shown in Figure D-6 with retrieved volume data. The user can save the results to a CSV file in the STEWARD server and download it to a local computer. Figure D-7 shows the CSV archives for the volume data. Newly generated files are archived on the top of the CSV file list. The file name is generated using the time stamp for convenience.

When a new CSV file is generated, it is also registered to the CSV Reports web page. Therefore, the user can access this CSV file later without a repeated data request. For example, when a new

CSV file (20080319174649.csv in Figure D-7) is generated, the same CSV file is archived in CSV Report page (20080319174649.csv in Figure D-8) .

The ***Prev Station*** and ***Next Station*** buttons may be used to navigate through the stations.

The detailed contents of the Volume Data tables are described in Table D-1.as follows:

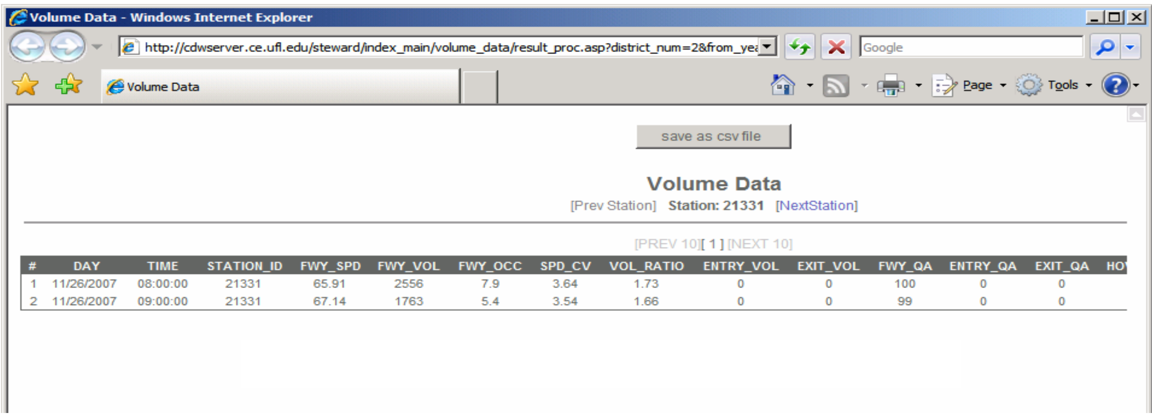


Figure D- 6: STEWARD volume map web page (data window)

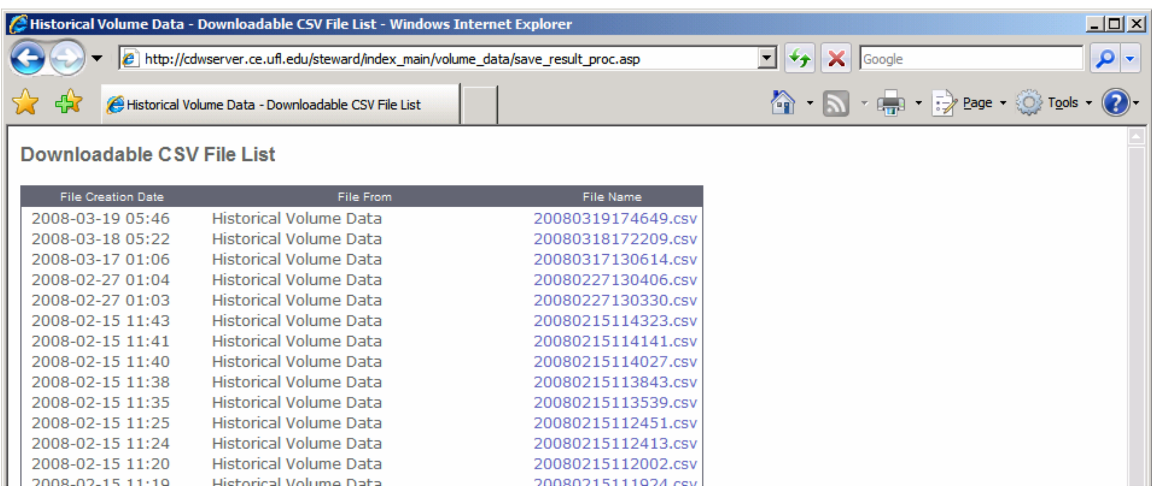


Figure D- 7: STEWARD volume map web page (CSV file archives)

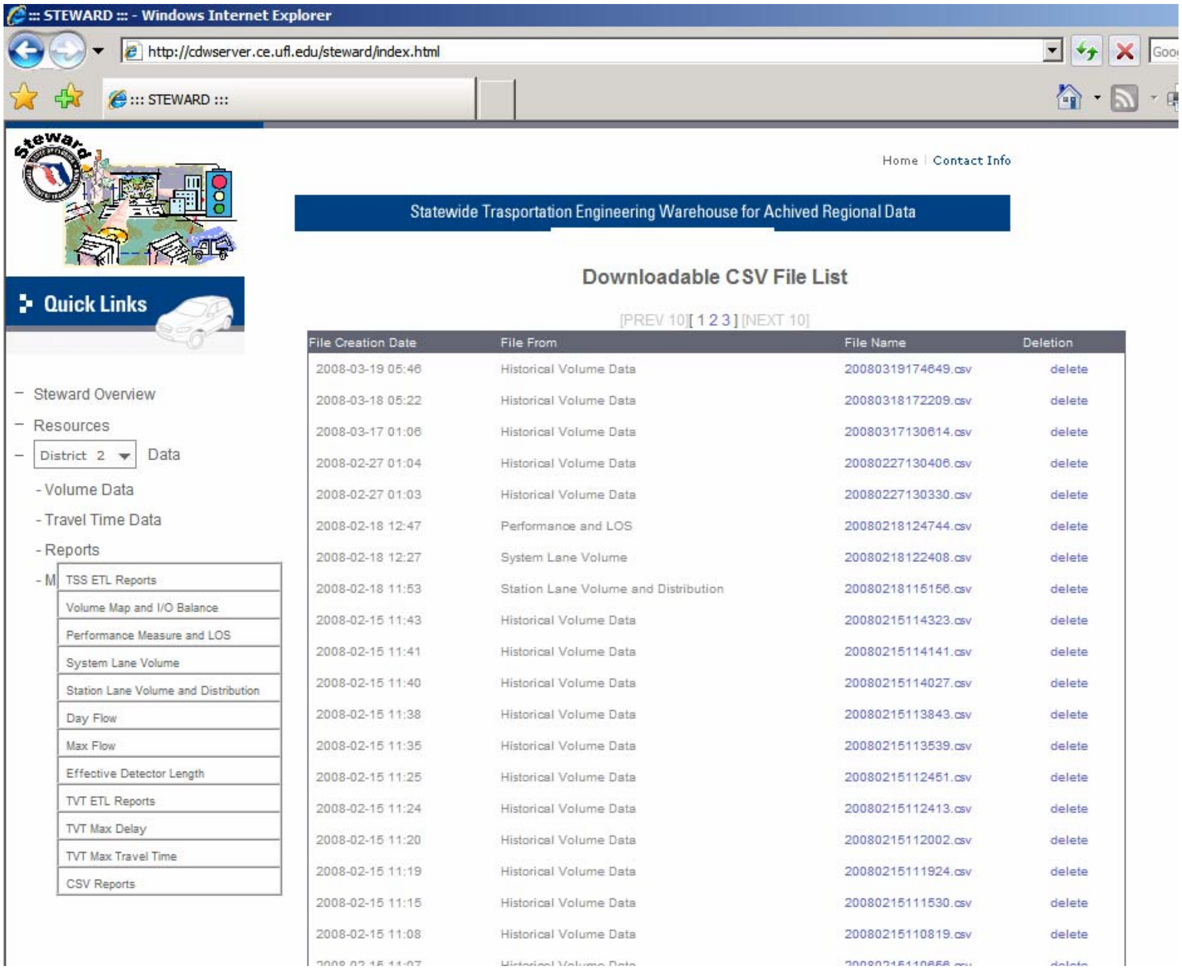


Figure D- 8: STEWARD CSV file list report web page

Table D-1 describes each column of the Volume data table in Figure D-6.

Table D- 1: Volume Data Table

Column	Data Type
#	Column number
DAY	Date of collection (MM/DD/YYYY)
TIME	Time Stamp (HH24:MM:SS)
STATION_ID	Station ID (5 characters)
FWY_SPD	Volume weighted average speed for mainline
FWY_VOL	Total count for mainline
FWY_OCC	Average occupancy for mainline (unweighted)
SPD_CV	Speed standard deviation / mean for all mainline lanes
VOL_RATIO	Highest lane volume / lowest lane volume for mainline lanes
SPD_RATIO	Highest lane speed / lowest lane speed for mainline lanes
ENTRY_VOL	Sum of entrance ramp counts
EXIT_VOL	Sum of exit ramp counts
FWY_QA	Currently the percent of expected observations that were received during the time period.
ENTRY_QA	
EXIT_QA	
HOV_SPD	Corresponding measures of freeway speed, volume occupancy and data quality in the HOV lanes, if present
HOV_VOL	
HOV_OCC	
HOV_QA	

4.3.2 Travel time data

As shown in Figure D-9, Travel time data (travel time in links) may be accessed from this page. The data content comes from travel time data from the SunGuide system. These data are aggregated into 5min, 15min and 1hr data in STEWARD database.

The travel time data page has a similar user interface to the volume data page. The user can query the traffic data with the specific conditions. The main differences are in the station selections. In the case of volume data, the user can select both the facility and stations. In the case of travel time data, I-95 in District 2 is assumed as one facility and only the origin and destination stations may be selected. The destination station selection allows multiple selections by using shift key or mouse.

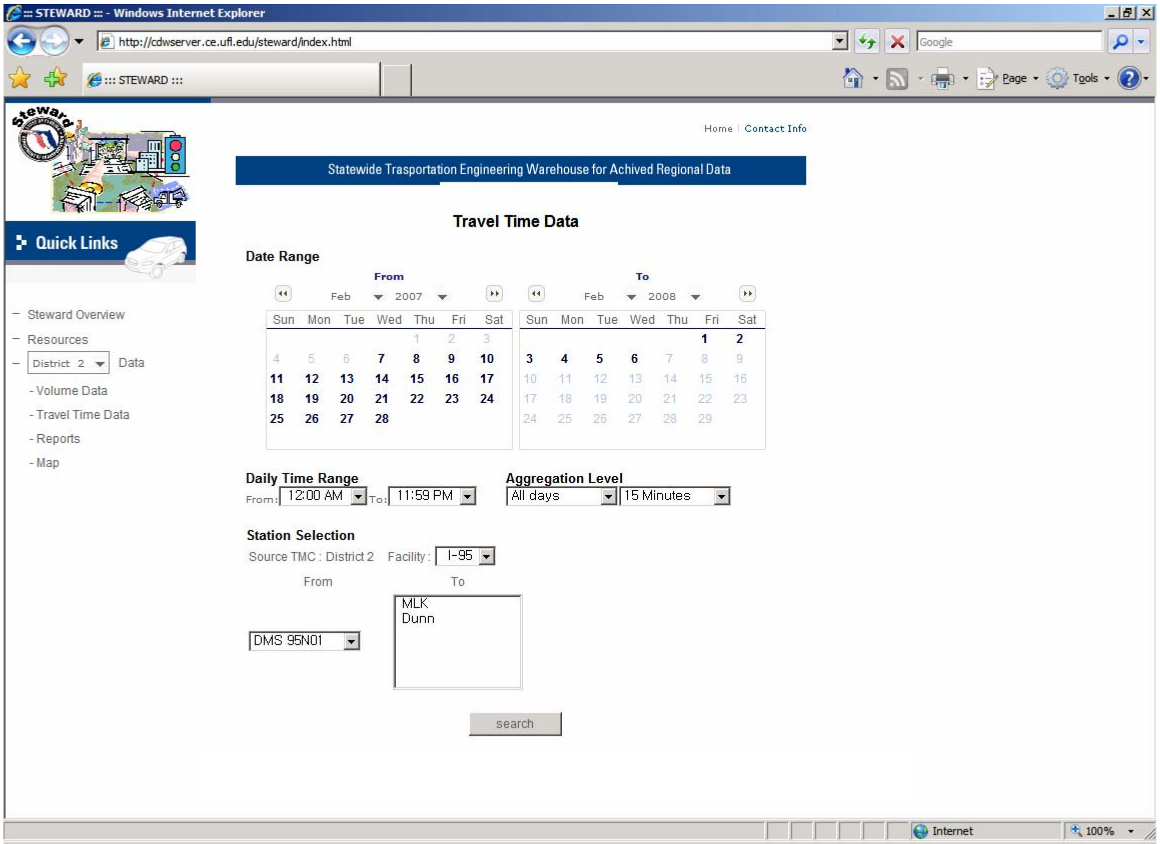


Figure D- 9: STEWARD travel time data web page

Table D-2 describes each column of the travel time data table.

Table D- 2: Travel Time Data Table

Column	Data Type
#	Column number
DAY	Date of collection (MM/DD/YYYY)
TIME	Time Stamp (HH24:MM:SS)
AVG_TT	(Sum of LinkTravelTime) /(Number of Observations)
DELAY	Sum of (Travel Time – Free Flow Travel Time *1.5)
MAX_TT	Max Travel Time
OCCURRENCE	Number of Observations
PCNT_ONTIME	OnTimePercentage / Observations
ONTIME_DELAY	Sum of (Travel Time – TT Reliability time*) *TT Reliability time = 60*(Linklength/(Speedlimit-10)
TVT_ORIGIN	Origin ID
TVT_DESTINATION	Destination ID

4.3.3 TSS ETL Report

As shown in FigureD-10, The user can access TSS ETL reports generated by the SunETLUtility. This program generates a log file for every input data file and appends a summary record to the conversion log file. The log file has the name of Dx_TSS-mmddyyyy-DailyReport.Log, where x stands for district number and mmddyyyy stands for the date data are collected. For example, D2_TSS-10042007-DailyReport.Log is the log file for District 2 TSS data collected on Oct. 4th, 2007. The conversion log file has the name of TSSConversionHistory-Dx.csv, where x stands for district number. These log files includes the status of each station and can be used as diagnostic tools.

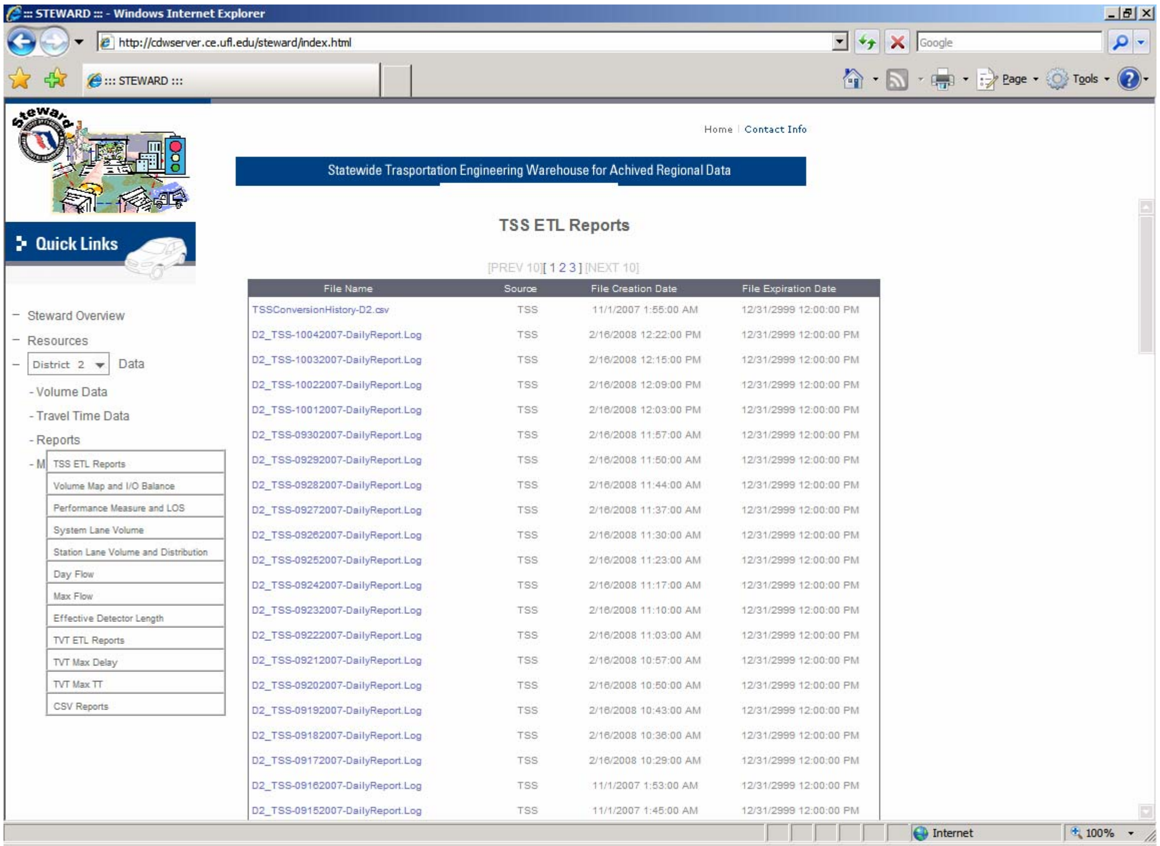


Figure D- 10: STEWARD TSS ETL report web page

Table D-3 describes each column of the TSS ETL Report table.

Table D- 3: TSS ETL Report Table

Column	Data Type
File name	Log file name
Source	Data source
File creation date	Date the log file was generated
File expiration date	Date the log file could be removed

4.3.4 Volume Map and I/O Balance Report

As shown in Figure D-11, the user can access a system-level volume map and I/O balance report from this page. The data content is generated from the STEWARD database and this page retrieves the data from the database. These data have the aggregation level of 5min, 15min, 1hr, and all day data.

The volume map and I/O balance report have a similar user interface to the volume data page. The main differences are aggregation level and station selections.

- Aggregation levels of 5min, 15min, 1hr and all day may be selected. In the “all day” case a daily summary may be retrieved.

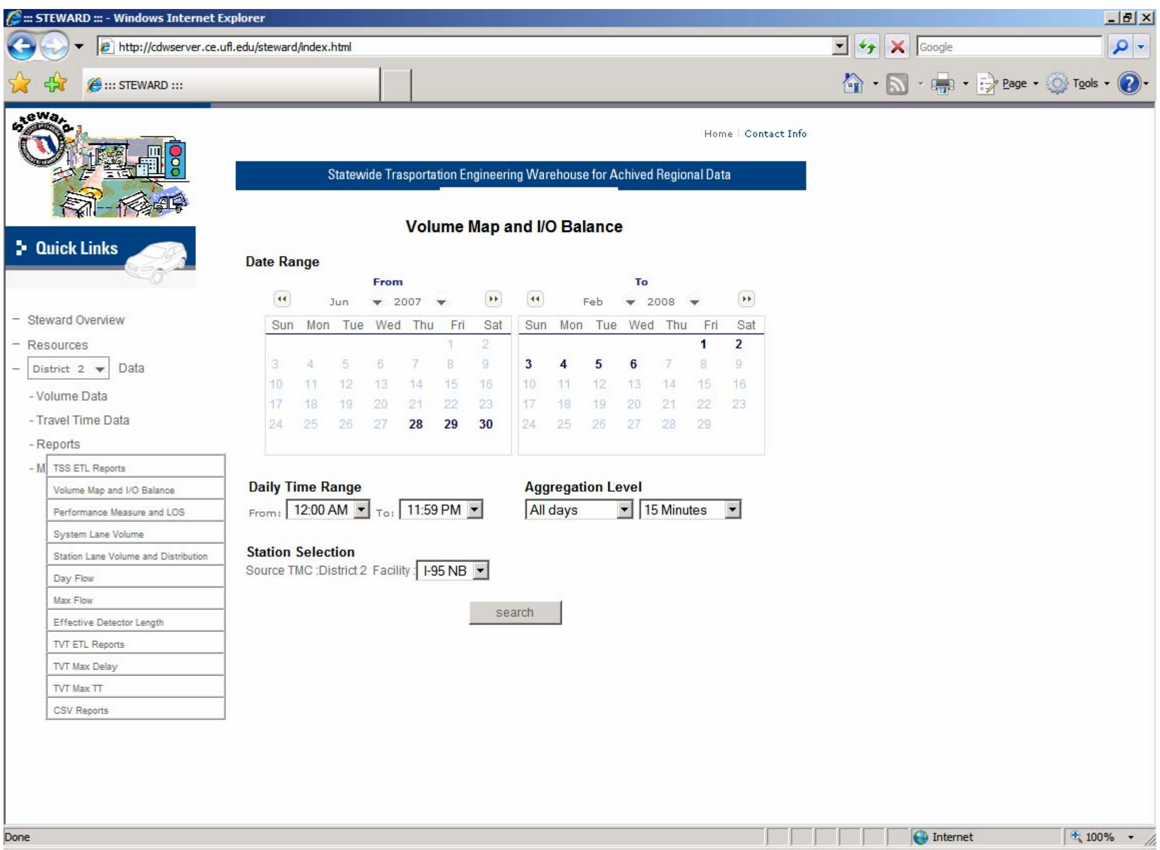


Figure D- 11: STEWARD volume map and I/O balance report web page

- There is no specific station selection. The nature of the volume map and I/O balance is such that it displays the results for the entire facility. In the case of District 2, the available facilities are I-95 NB and I-95 SB.

Table D-4 describes each column of the Volume Map and I/O balance table

Table D- 4: Volume Map and I/O Balance Report Table

Column	Data Type
#	Column number
DATE	Date of collection (MM/DD/YYYY)
TIME	Time Stamp (HH24:MM:SS)
Staiton_ID	Station ID
Station_MP	Station Milepost
Upnode_ID	Upstream node station ID
Entry_Volume	Entrance ramp volume from the current node
FWY_Volume	Freeway volume from the current node
Exit_Volume	Exit ramp volume from the current node
Link_input	Sum of upnode freeway volume and upnode entrance ramp volume
Link_output	Sum of freeway volume and exit ramp volume at current node
Difference	Link_input – Link_output
Pcnt_Diff	100 * 2*(Link_input – Link_output) /(Link_input + Link_output)

4.3.5 Performance Measure and LOS Report

As shown in Figure D-12, the user can access performance measure and LOS data from this page. The data content is generated from the STEWARD database and this page retrieves the data from the database. These data have aggregation levels of 5min, 15min, 1hr and all day data.

The performance measure and LOS data page has the similar user interface to volume data page. The main difference is aggregation level which is treated the same as the aggregation level for the volume map and IO Balance report.

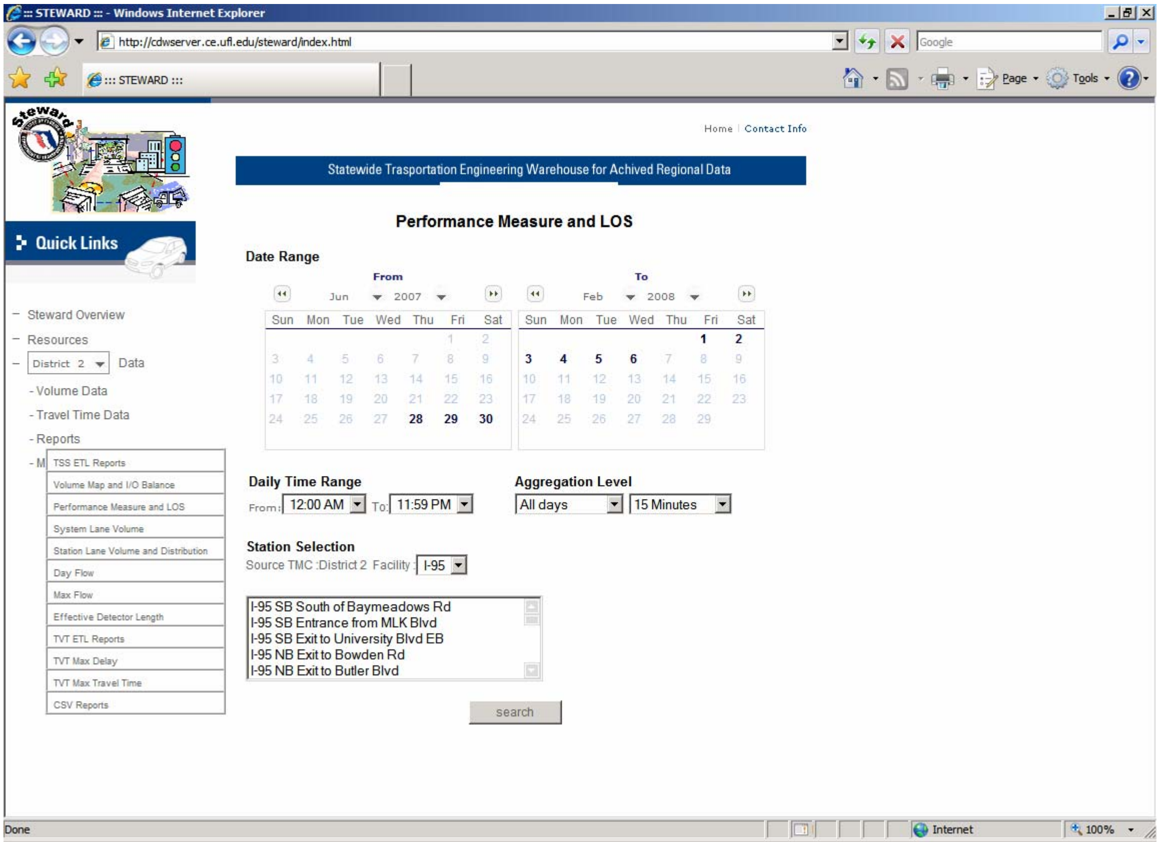


Figure D- 12: STEWARD performance measure and LOS report web page

Table D-5 describes each column of the performance measure and LOS table

Table D- 5: Performance Measure and LOS Report Table

Column	Data Type
#	Column number
DATE	Date of collection (MM/DD/YYYY)
TIME	Time Stamp (HH24:MM:SS)
Staiton_ID	Station ID
Station Description	Station Description
Upnode ID	Upstream node station ID
Upnode description	Upstream node station description
Station_MP	Station Milepost
Length	Distance between upnode and current node
Link volume	Average of Link_In* and Link_out volume* *Link_In volume = Sum of upnode freeway volume and upnode entrance ramp volume *Link_out volume = Sum of freeway volume and exit ramp volume at current node
Num of Lanes	Number of Lanes
Volume/Lane	Link_volume / Number of Lanes
Veh-Mi	Link_volume * Length
Veh-Hr	Link_volume * Length/(freeway speed at current node
Speed	Average speed
Delay	If freeway speed is 10mph below the speed limit, delay = (freeway volume) * length *(1/freeway_speed – 1/(freeway_speed-10))
Kinetic Energy	Link_volume * freeway_speed
Density	Hourly Link_volume * freeway_speed This column exists for 5min, 15min and 1hr aggregation level only.
V/C	100 * Density /(Lane_Capacity * Number of Lanes) This column exists for 5min, 15min and 1hr aggregation level only.
LOS	F when Density > 45 E when Density > 35 D when Density > 26 C when Density > 18 B when Density > 11 A otherwise This column exists for 5min, 15min and 1hr aggregation level only.
F_Plus	Percentage of LOS counts that is above ‘F’ in that day This column exists for all day aggregation level only.
E_Plus	Percentage of LOS counts that is above ‘E’ in that day This column exists for all day aggregation level only.
D_Plus	Percentage of LOS counts that is above ‘D’ in that day This column exists for all day aggregation level only.
C_Plus	Percentage of LOS counts that is above ‘C’ in that day This column exists for all day aggregation level only.

4.3.6 System Lane Volume Report

As shown in Figure D-13, the user can access system lane volume data from this page. The data content is generated from the STEWARD system and this page retrieves the data from the database. These data has the aggregation level of 5min, 15min, 1hr, and all day data.

The system lane volume data page has the same user interface as the volume map and I/O balance report page

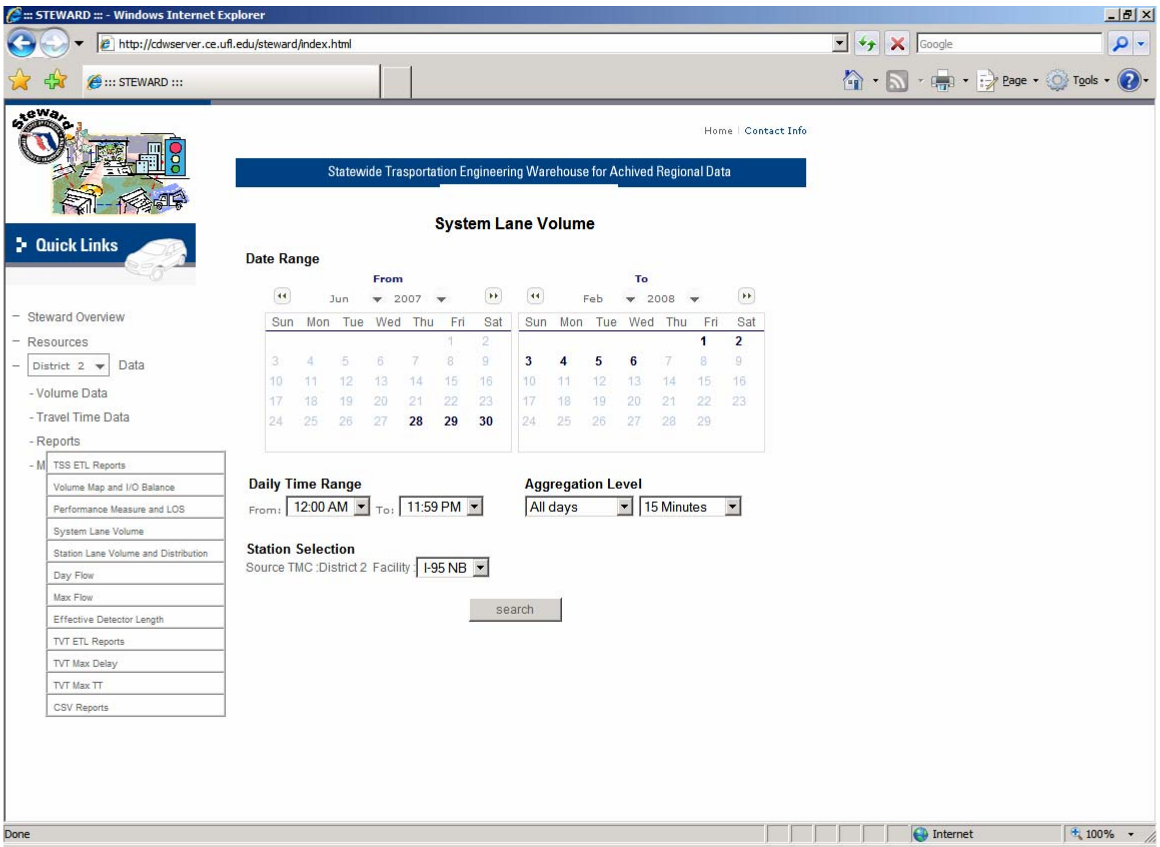


Figure D- 13: STEWARD system lane volume report web page

Table D-6 describes each column of the system lane volume table

Table D- 6: System Lane Volume Report Table

Column	Data Type
#	Column number
DATE	Date of collection (MM/DD/YYYY)
TIME	Time Stamp (HH24:MM:SS)
Facility	Facility ID
Staiton_ID	Station ID
Station Description	Station Description
Station_MP	Station Milepost
Lane1 Vol	Lane1 Volume
Lane1 Occ	Lane1 Occupancy
Lane1 Spd	Lane1 Speed
Lane2 Vol	Lane2 Volume
Lane2 Occ	Lane2 Occupancy
Lane2 Spd	Lane2 Speed
Lane3 Vol	Lane3 Volume
Lane3 Occ	Lane3 Occupancy
Lane3 Spd	Lane3 Speed
Lane4 Vol	Lane4 Volume
Lane4 Occ	Lane4 Occupancy
Lane4 Spd	Lane4 Speed
Lane5 Vol	Lane5 Volume
Lane5 Occ	Lane5 Occupancy
Lane5 Spd	Lane5 Speed
Lane6 Vol	Lane6 Volume
Lane6 Occ	Lane6 Occupancy
Lane6 Spd	Lane6 Speed
Onramp1 Vol	Onramp1 Volume
Onramp1 Occ	Onramp1 Occupancy
Onramp1 Spd	Onramp1 Speed
Onramp2 Vol	Onramp2 Volume
Onramp2 Occ	Onramp2 Occupancy
Onramp2 Spd	Onramp2 Speed
Onramp3 Vol	Onramp3 Volume
Onramp3 Occ	Onramp3 Occupancy
Onramp3 Spd	Onramp3 Speed
Offramp1 Vol	Offramp1 Volume
Offramp1 Occ	Offramp1 Occupancy
Offramp1 Spd	Offramp1 Speed
Offramp2 Vol	Offramp2 Volume
Offramp2 Occ	Offramp2 Occupancy
Offramp2 Spd	Offramp2 Speed
Offramp3 Vol	Offramp3 Volume
Offramp3 Occ	Offramp3 Occupancy
Offramp3 Spd	Offramp3 Speed

4.3.7 Station Lane Volume and Distribution Report

As shown in Figure D-14, users can access station lane volume and distribution data from this page. The data content is generated from the STEWARD database and this page retrieves the data from the database. These data have the aggregation levels of 5min, 15min, 1hr and all day data.

The station lane volume and distribution data page has a similar user interface to the volume data page. Users can query the traffic data with the specific conditions. The main difference is aggregation level. Aggregation levels of 5min, 15min, 1hr and all day may be selected. In the "all day" case a daily summary may be retrieved.

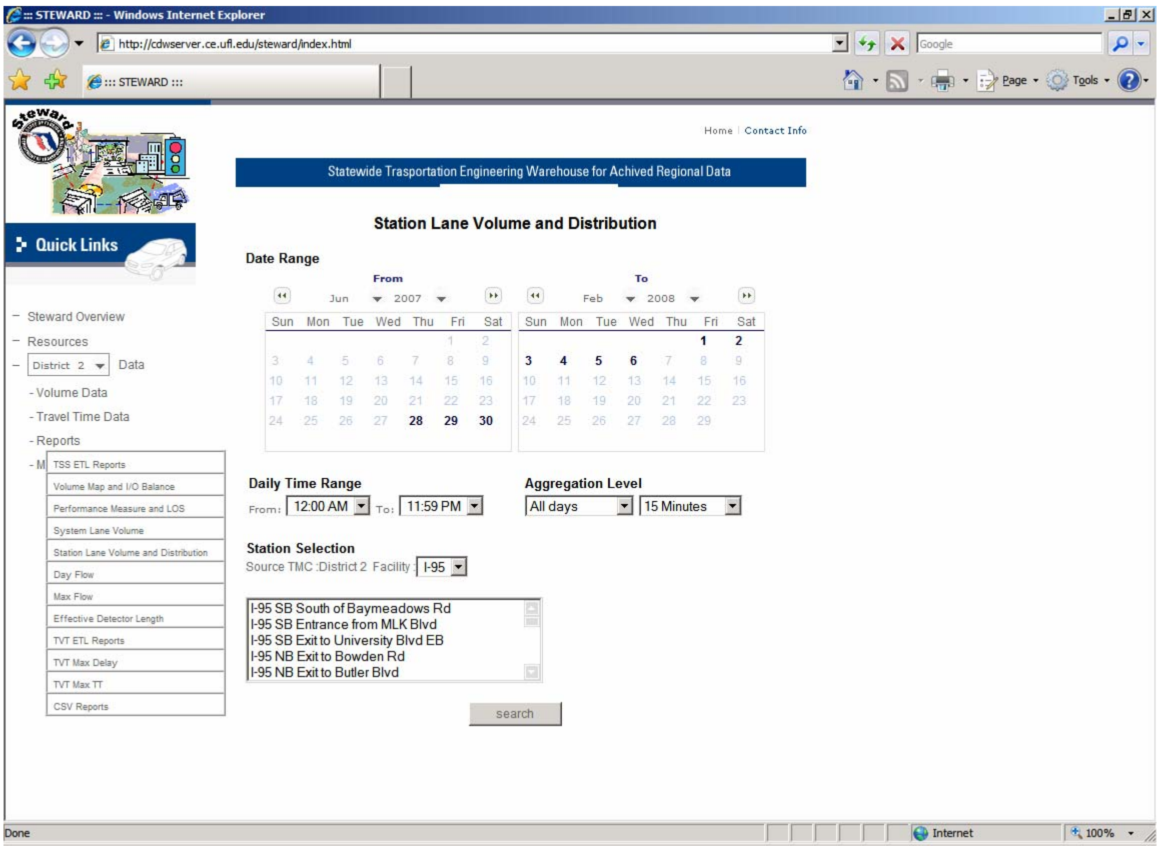


Figure D- 14: STEWARD station lane volume and distribution report web page

Table D-7 describes each column of the station lane volume and distribution table

Table D- 7: Station Lane Volume and Distribution Report Table

Column	Data Type
#	Column number
DATE	Date of collection (MM/DD/YYYY)
TIME	Time Stamp (HH24:MM:SS)
Facility	Facility ID
Staiton_ID	Station ID
Station Description	Station Description
Station_MP	Station Milepost
Lane1 Vol	Lane1 Volume
Lane2 Vol	Lane2 Volume
Lane3 Vol	Lane3 Volume
Lane4 Vol	Lane4 Volume
Lane5 Vol	Lane5 Volume
Lane6 Vol	Lane6 Volume
Total	Sum of lane volumes
Balance	Max lane volume/Min lane volume

4.3.8 Daily Flow Report

As shown in Figure D-15, users can access daily flow data from this page. The data content is generated from the STEWARD database and this page retrieves the data from the database. These data have aggregation levels of 5min, 15min, and 1hr.

The daily flow data page has a similar user interface to volume data page. Users can query the traffic data with the specific conditions. The main difference is daily time range. This report returns the daily flow for the day. Therefore there is no need to specify the time range.

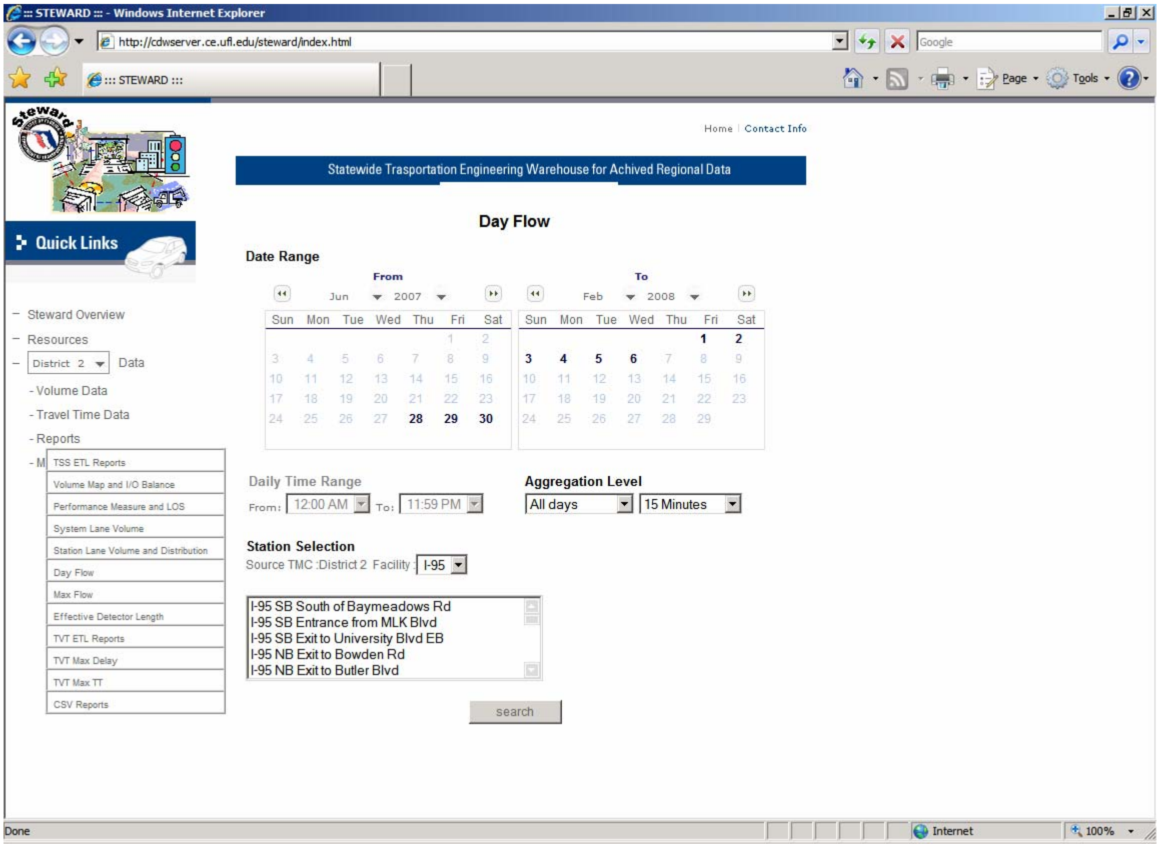


Figure D- 15: STEWARD daily flow report web page

Table D-8 describes each column of the daily flow report table

Table D- 8: Daily Flow Report Table

Column	Data Type
#	Column number
DATE	Date of collection (MM/DD/YYYY)
Facility	Facility ID
Staiton_ID	Station ID
Station Description	Station Description
Station_MP	Station Milepost
Lane1 flow rate	Lane1 flow rate
Lane2 flow rate	Lane2 flow rate
Lane3 flow rate	Lane3 flow rate
Lane4 flow rate	Lane4 flow rate
Lane5 flow rate	Lane5 flow rate
Lane6 flow rate	Lane6 flow rate

4.3.9 Maximum Flow Report

As shown in Figure D-16, users can access the maximum flow data from this page. The data content is generated from the STEWARD database and this page retrieves the data from the database. These data have aggregation levels of 5min, 15min, and 1hr.

The maximum flow data page has the similar user interface to volume data page. Users can query the traffic data based on specified conditions. The main difference is daily time range. This report returns the maximum flow for the day. Therefore there is no need to specify the time range.

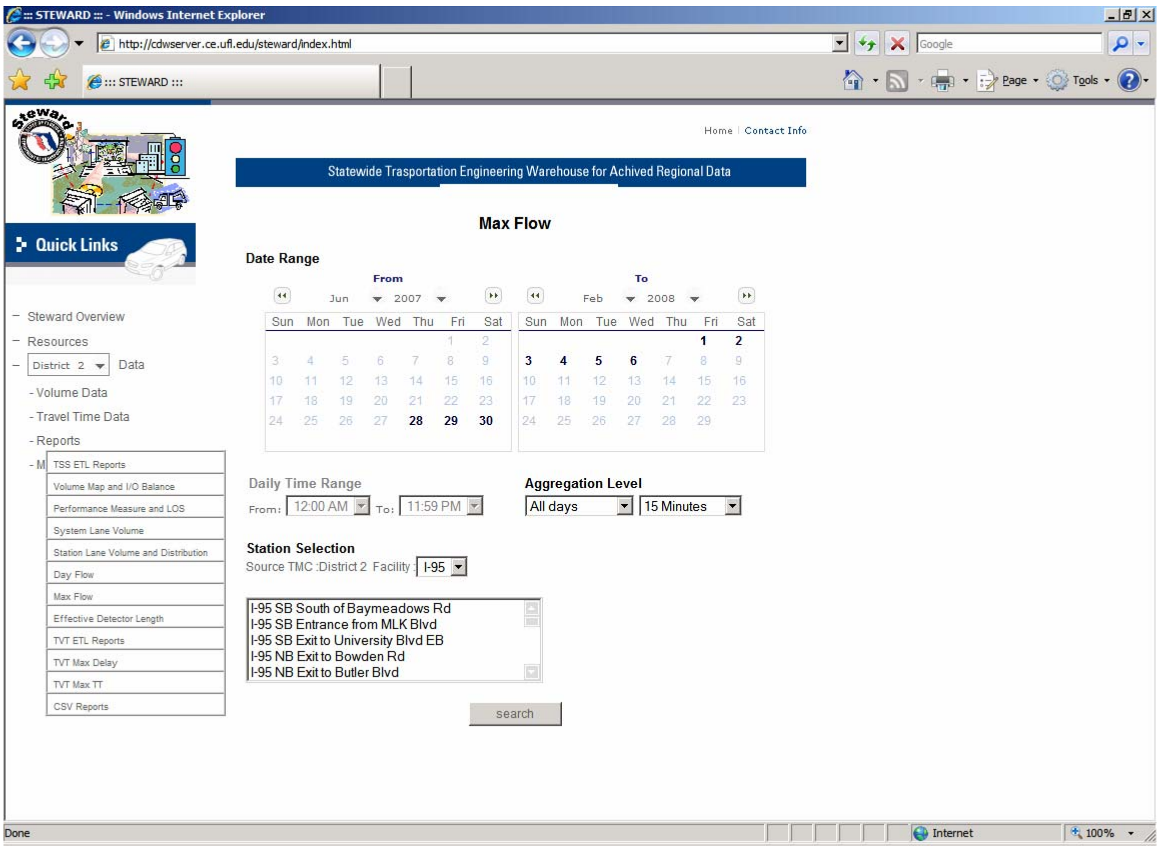


Figure D- 16: STEWARD maximum flow report web page

Table D-9 describes each column of the maximum flow table

Table D- 9: Maximum Flow Report Table

Column	Data Type
#	Column number
DATE	Date of collection (MM/DD/YYYY)
Facility	Facility ID
Staiton_ID	Station ID
Station Description	Station Description
Station_MP	Station Milepost
Lane num	Lane number with max flow
Max flow	Max flow
Max time	Time when max flow occurs
Max gap	N/A

4.3.10 Effective Detector Length Report

As shown in Figure D-17, users can access effective detector length data from this page. The data content is generated from the STEWARD database and this page retrieves the data from the database. These data have aggregation levels of 5min, 15min, and 1hr and all day data.

The effective detector length data page has a similar user interface to volume data page. Users can query the traffic data with the specified conditions. The main difference is the aggregation level. Users can select aggregation levels of 5min, 15min, 1hr and all day.

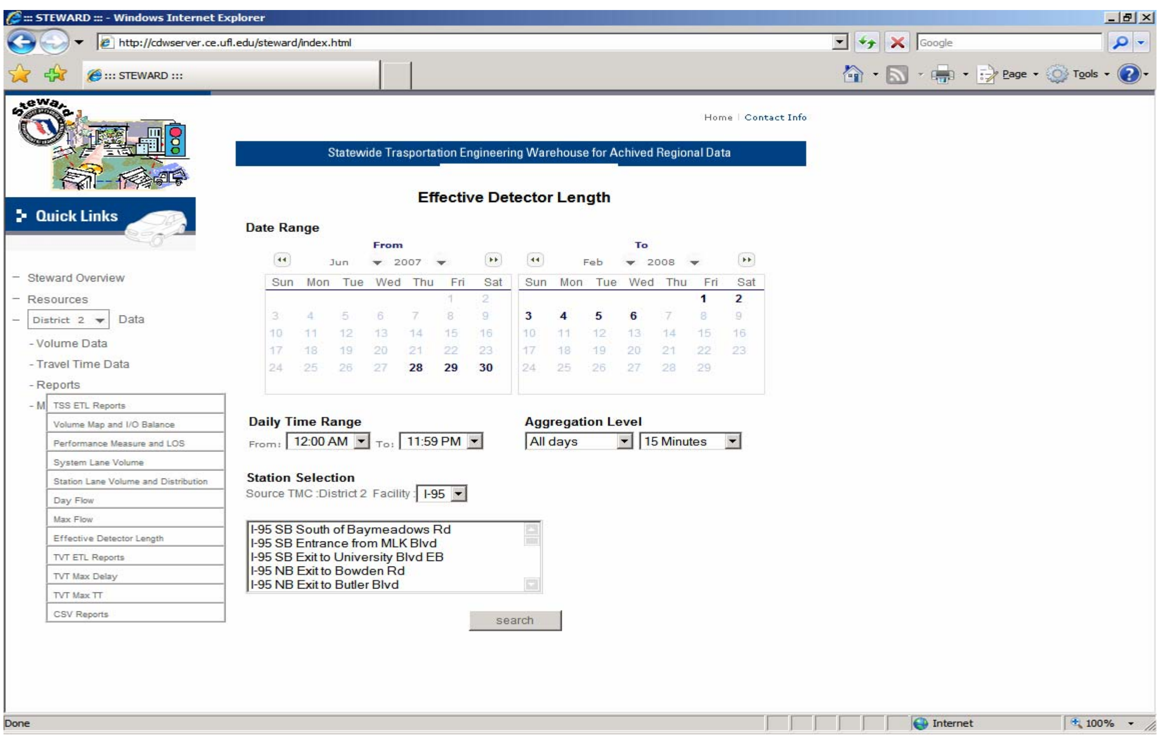


Figure D- 17: STEWARD effective detector length report web page

Table D-10 describes each column of the effective detector length table

Table D- 10: Effective Detector Length Report Table

Column	Data Type
#	Column number
DATE	Date of collection (MM/DD/YYYY)
TIME	Time Stamp (HH24:MM:SS)
Facility	Facility ID
Staiton_ID	Station ID
Station Description	Station Description
Station_MP	Station Milepost
Lane1 Vol	Lane1 Volume
Lane1 Occ	Lane1 Occupancy
Lane1 Spd	Lane1 Speed
Lane1 DL	Lane1 effective detector length (SPD * 5280) * (OCC /100) / VOL
Lane2 Vol	Lane2 Volume
Lane2 Occ	Lane2 Occupancy
Lane2 Spd	Lane2 Speed
Lane2 DL	Lane2 effective detector length (SPD * 5280) * (OCC /100) / VOL
Lane3 Vol	Lane3 Volume
Lane3 Occ	Lane3 Occupancy
Lane3 Spd	Lane3 Speed
Lane3 DL	Lane3 effective detector length (SPD * 5280) * (OCC /100) / VOL
Lane4 Vol	Lane4 Volume
Lane4 Occ	Lane4 Occupancy
Lane4 Spd	Lane4 Speed
Lane4 DL	Lane4 effective detector length (SPD * 5280) * (OCC /100) / VOL
Lane5 Vol	Lane5 Volume
Lane5 Occ	Lane5 Occupancy
Lane5 Spd	Lane5 Speed
Lane5 DL	Lane5 effective detector length (SPD * 5280) * (OCC /100) / VOL
Lane6 Vol	Lane6 Volume
Lane6 Occ	Lane6 Occupancy
Lane6 Spd	Lane6 Speed
Lane6 DL	Lane6 effective detector length (SPD * 5280) * (OCC /100) / VOL

4.3.11TVT ETL report

As shown in Figure D-18, users can access TVT ETL reports generated by SunETLUtility. This program generates a log file for every input data file and appends a summary record to the conversion log file. The log file has the name of Dx_TVT-mmddyyyy-DailyReport.Log, where x stands for district number and mmddyyyy stands for the date data are collected. For example, D2_TVT-05142007-DailyReport.Log is the log file for District 2 TVT data collected on May 14th, 2007. The conversion log file has the name of TVTConversionHistory-Dx.csv, where x stands for district number. These log files includes the status of each station and can be used as the tools for status check.

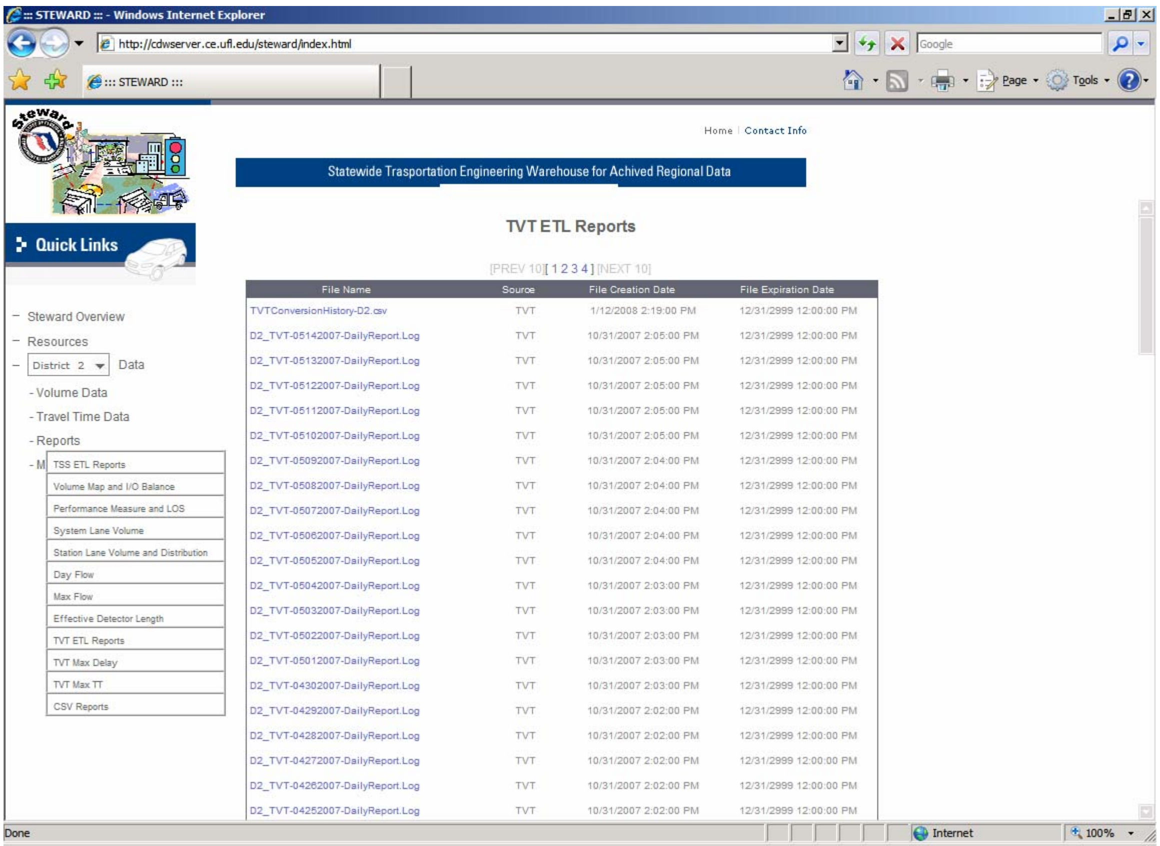


Figure D- 18: STEWARD TVT ETL report web page

Table D-11 describes each column of the TVT ETL Report table.

Table D- 11: TSS ETL Report Table

Column	Data Type
File name	Log file name
Source	Data source
File creation date	Date the log file was generated
File expiration date	Date the log file could be removed

4.3.12 TVT Max Delay Report

As shown in Figure D-19, users can access the maximum delay data from this page. The data content is generated from the STEWARD database and this page retrieves the data from the database. These data have aggregation levels of 5min, 15min, and 1hr.

The maximum delay data page has a similar user interface to volume data page. Users can query the traffic data with the specified conditions. The main difference is daily time range. This report returns the maximum delay for the day. Therefore there is no need to specify the time range.

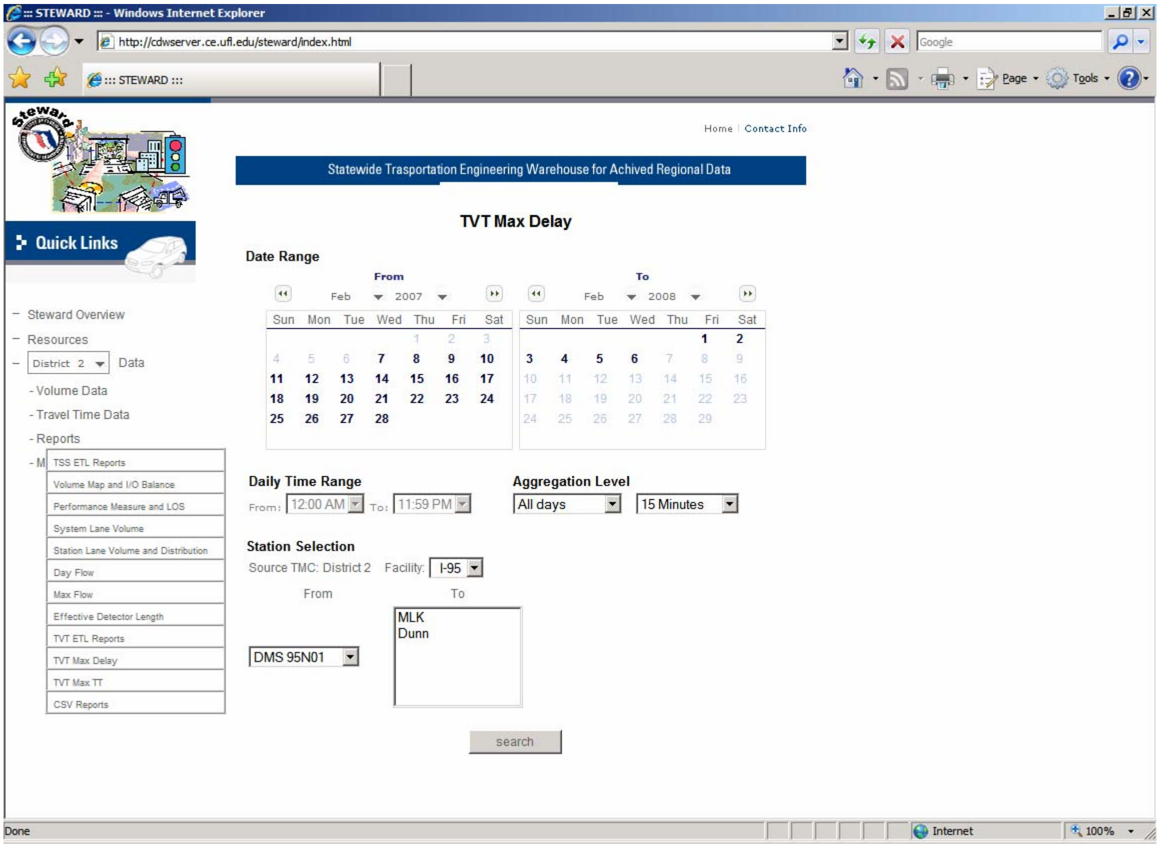


Figure D- 19: STEWARD TVT maximum delay report web page

Table D-12 describes each column of the TVT maximum delay table

Table D- 12: TVT Maximum Delay Report Table

Column	Data Type
#	Column number
DATE	Date of collection (MM/DD/YYYY)
Origin	Description of origin
Destination	Description of destination
FF TT	Free flow travel time
Avg TT	Average travel time
Max_delay	Max delay
Max delay time	Timestamp when max delay occurs

4.3.13 TVT Max Travel Time Report

As shown in Figure D-20, users can access maximum travel time data from this page. The data content is generated from the STEWARD database and this page retrieves the data from the database. These data have aggregation levels of 5min, 15min, and 1hr.

The maximum travel time data page has a similar user interface to volume data page. Users can query the traffic data with the specified conditions. The main difference is daily time range. This report returns the maximum travel time for the day. Therefore there is no need to specify the time range.

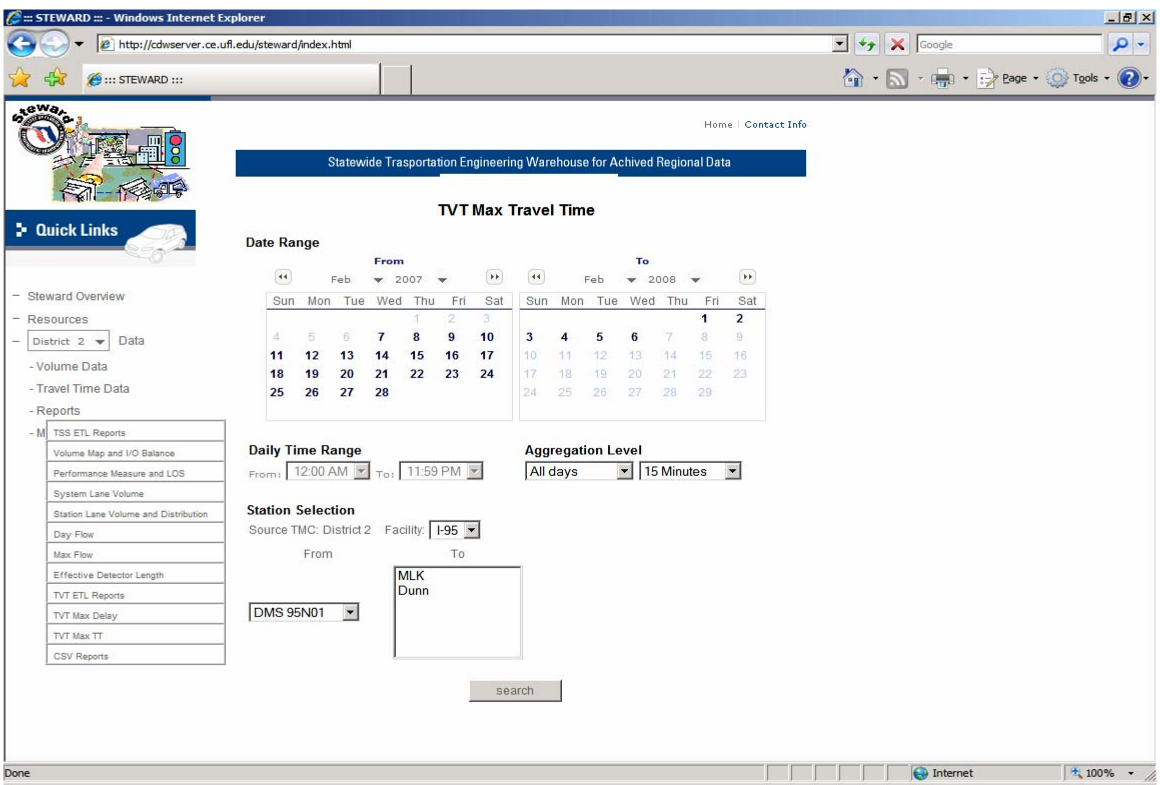


Figure D- 20: STEWARD TVT maximum travel time report web page

D-13 describes each column of the TVT maximum travel time table

Table D- 13: TVT Maximum Travel Time Report Table

Column	Data Type
#	Column number
DATE	Date of collection (MM/DD/YYYY)
Origin	Description of origin
Destination	Description of destination
FF TT	Free flow travel time
Avg TT	Average travel time
Max_TT	Max travel time
Max TT time	Timestamp when max travel time occurs
Avg delay	Average delay

4.3.14CSV File List

As described in the *volume data* page, the CSV files for each report are generated in the STEWARD server. As shown in Figure D-21, these files are accessible from the CSV file list page. From this page, users can delete the CSV files from the STEWARD server or download the CSV files again without queries.

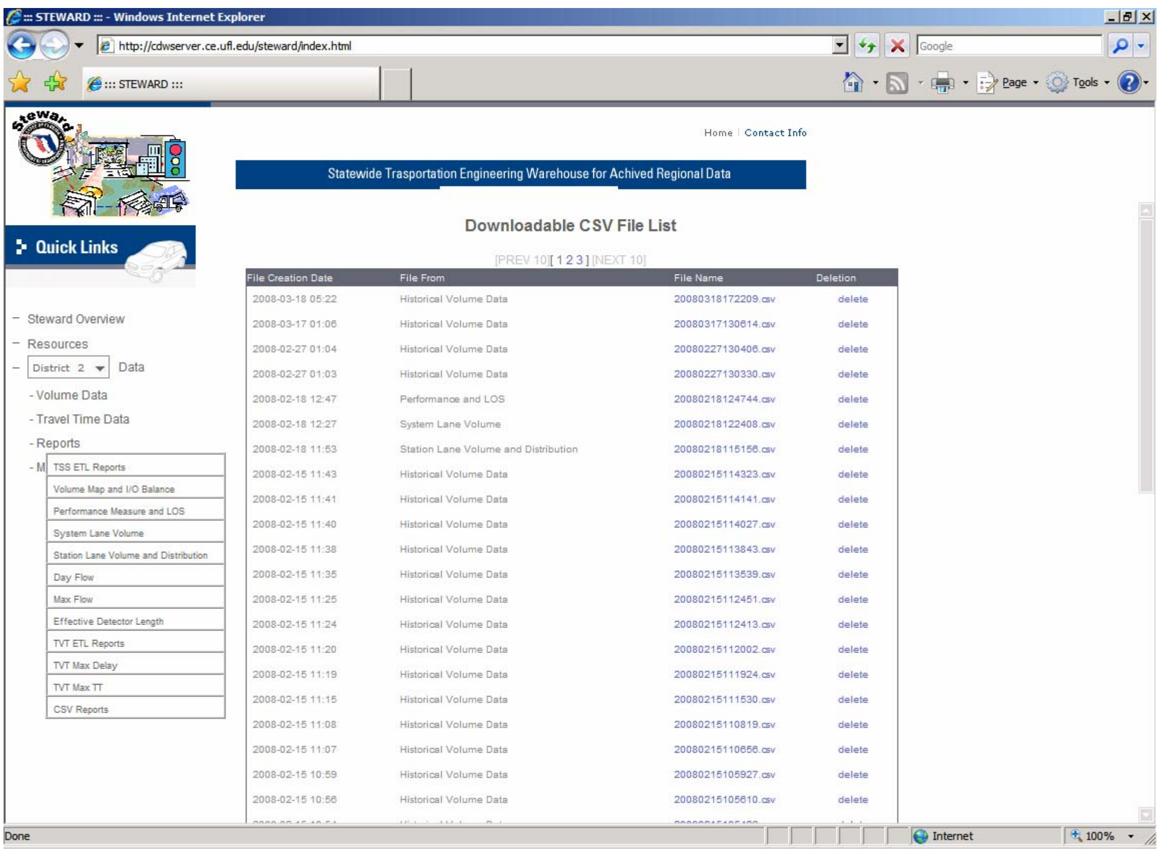


Figure D- 21: STEWARD CSV file list web page

4.3.15 District Detector Maps

From this page, users can download TSS/TVT detector maps. Examples of these maps are presented in Figures D-22 and D-23.

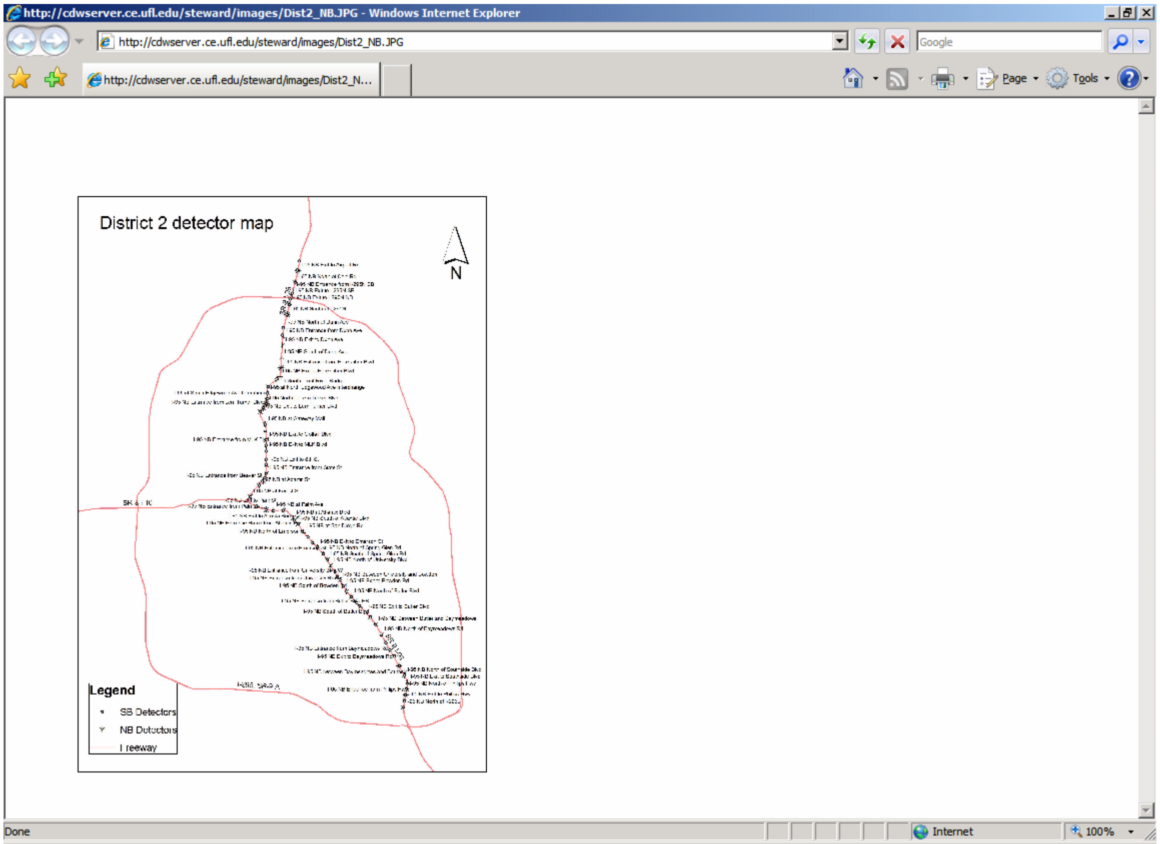


Figure D- 22: STEWARD Sample TSS detector map web page

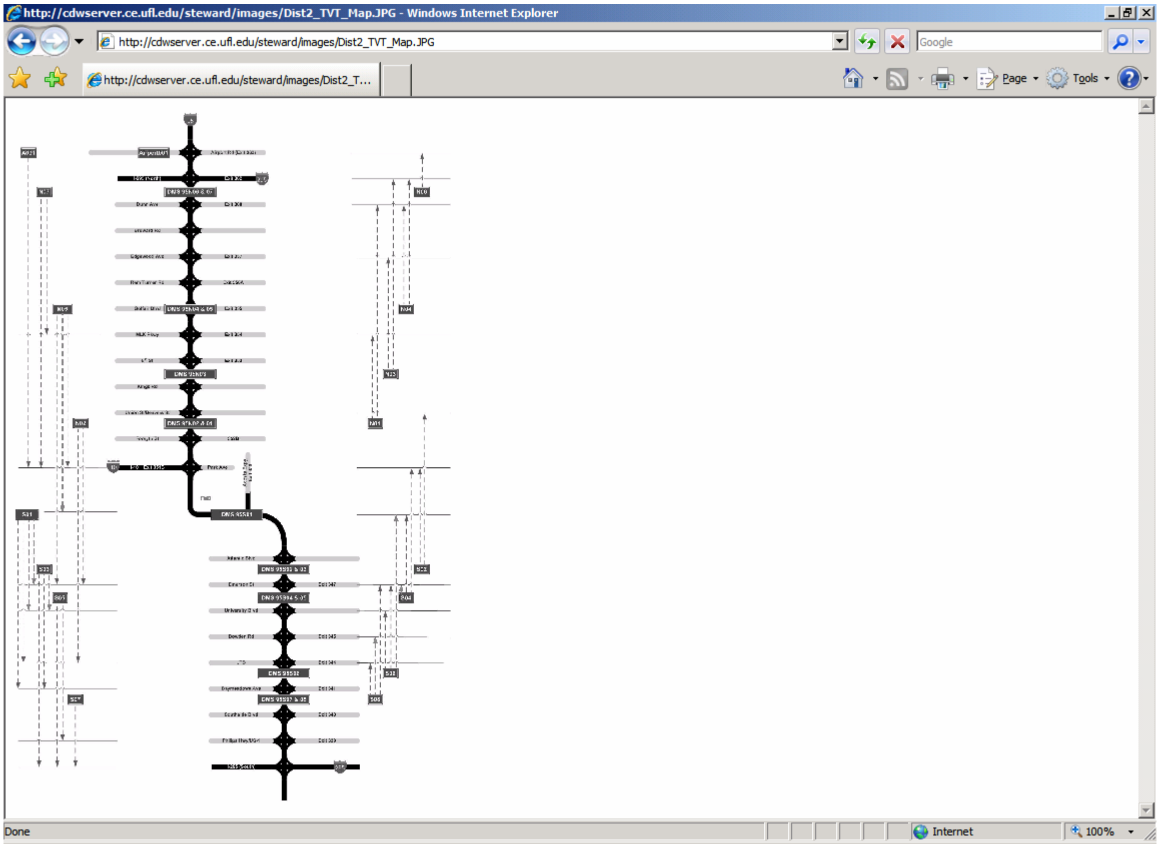


Figure D- 23: STEWARD sample TVT detector map web page

4.3.16 Contact Information

As shown in Figure D-24, the project contact information is presented on this page.

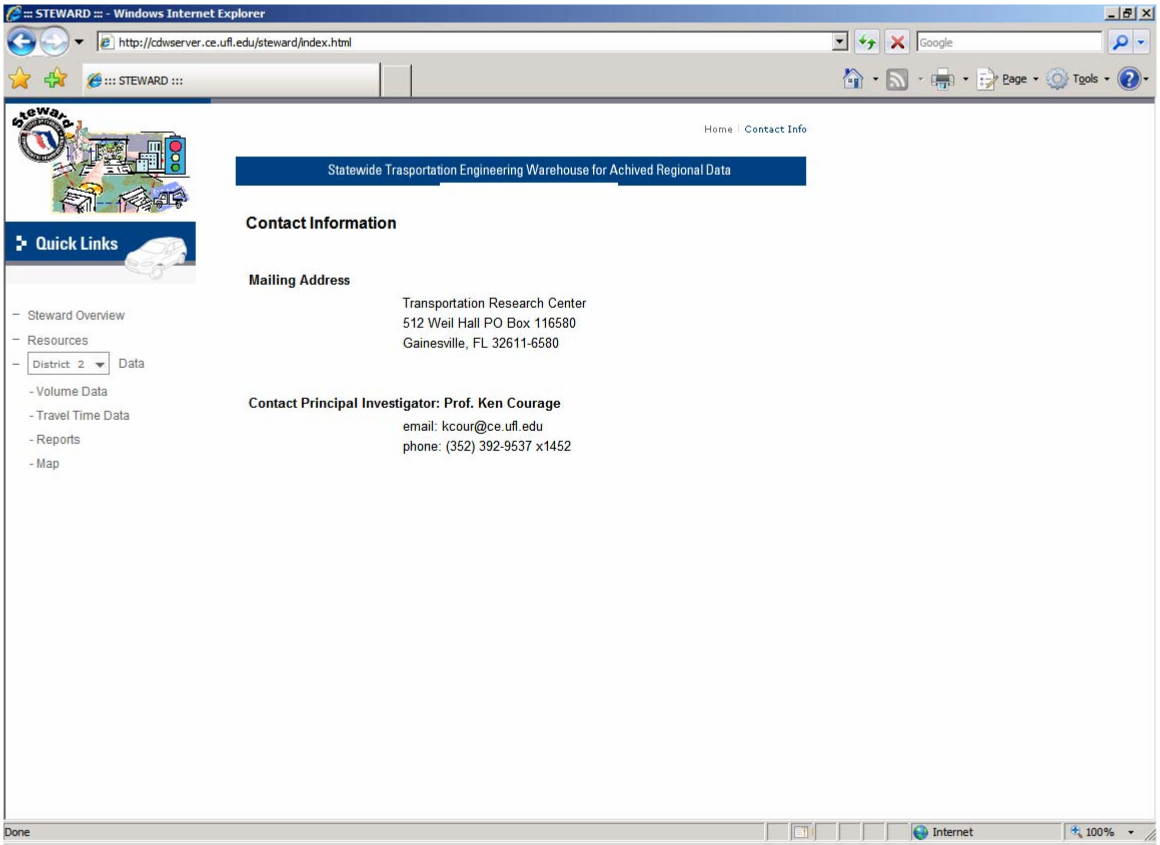


Figure D- 24: STEWARD contact information web page

5 Examples

Two examples are provided here to illustrate the use of the web site to obtain data from the STEWARD database

5.1 Example 1 Retrieve station-level hourly volume data from District 2

5.1.1 Open the following web site:

http://cdwserver.ce.ufl.edu/steward/index.html

5.1.2 Select the followings from left pane to get the volume data:

District 2 data → Volume data

The result will be as Figure D-25.

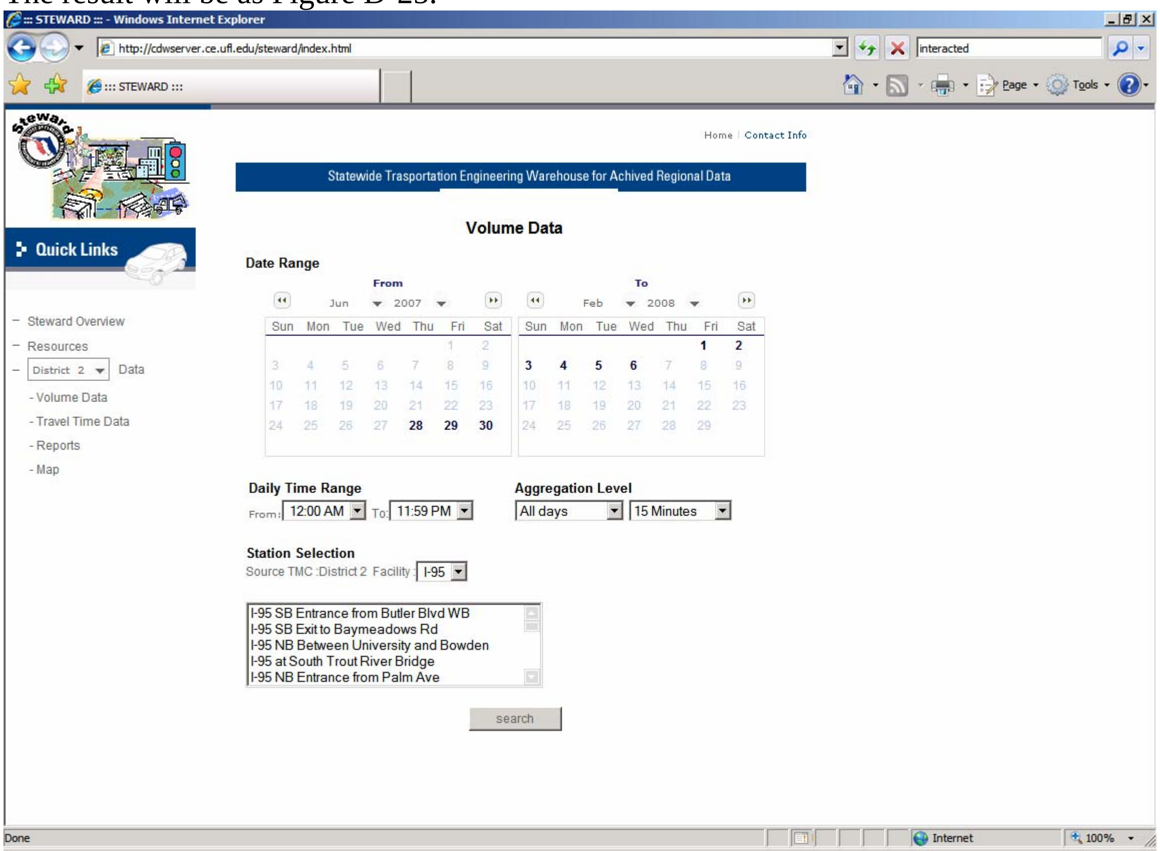


Figure D- 25: Example 1 - Select volume data page

5.1.3 Select the starting and ending days from Date Range

As shown in Figure D-26, the “From” calendar shows the month that includes the first day of the data. The “To” calendar shows the month that includes the last day of the data. The days with bold font indicate that the data are available on these days.

Click the starting and ending days on the calendar as July, 1st 2007 and July, 31st 2007. You can navigate the month and year using the pull-down menu or arrow keys.

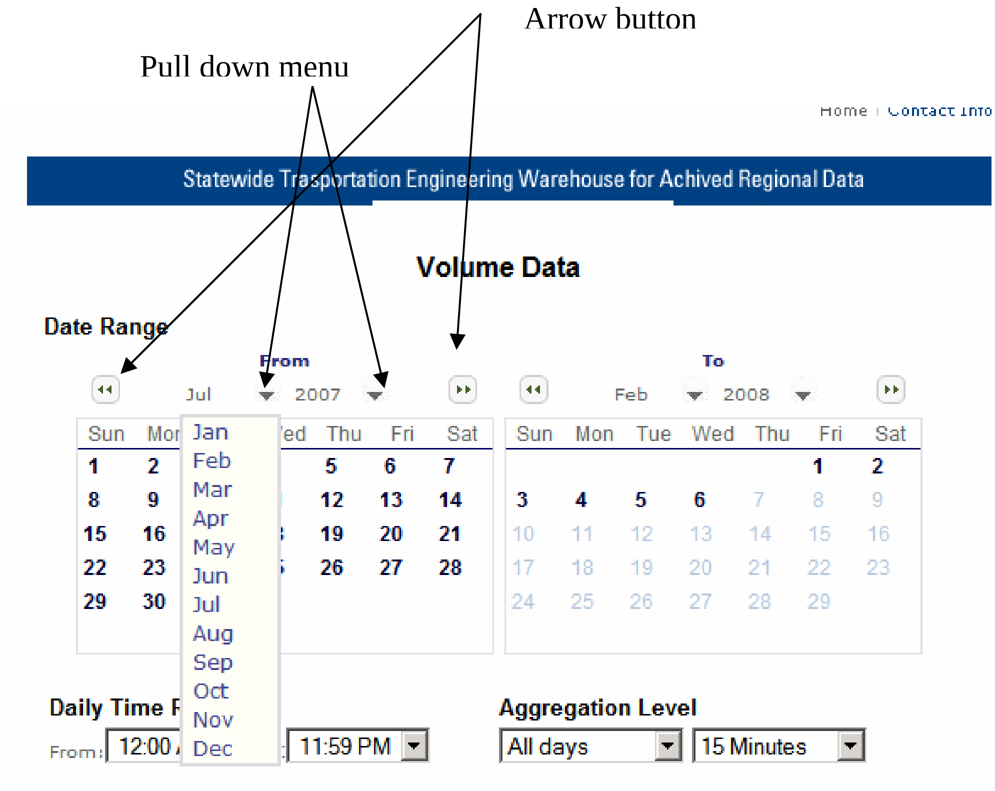


Figure D- 26: Example 1 - Select date range

5.1.4 Select the Daily Time Range and Aggregation Level

As shown in Figure D-27, keep the default daily time range as 12:00AM and 11:59PM to retrieve the entire day's data.
Select ***Weekday*** and ***1 Hour*** for the aggregation level.

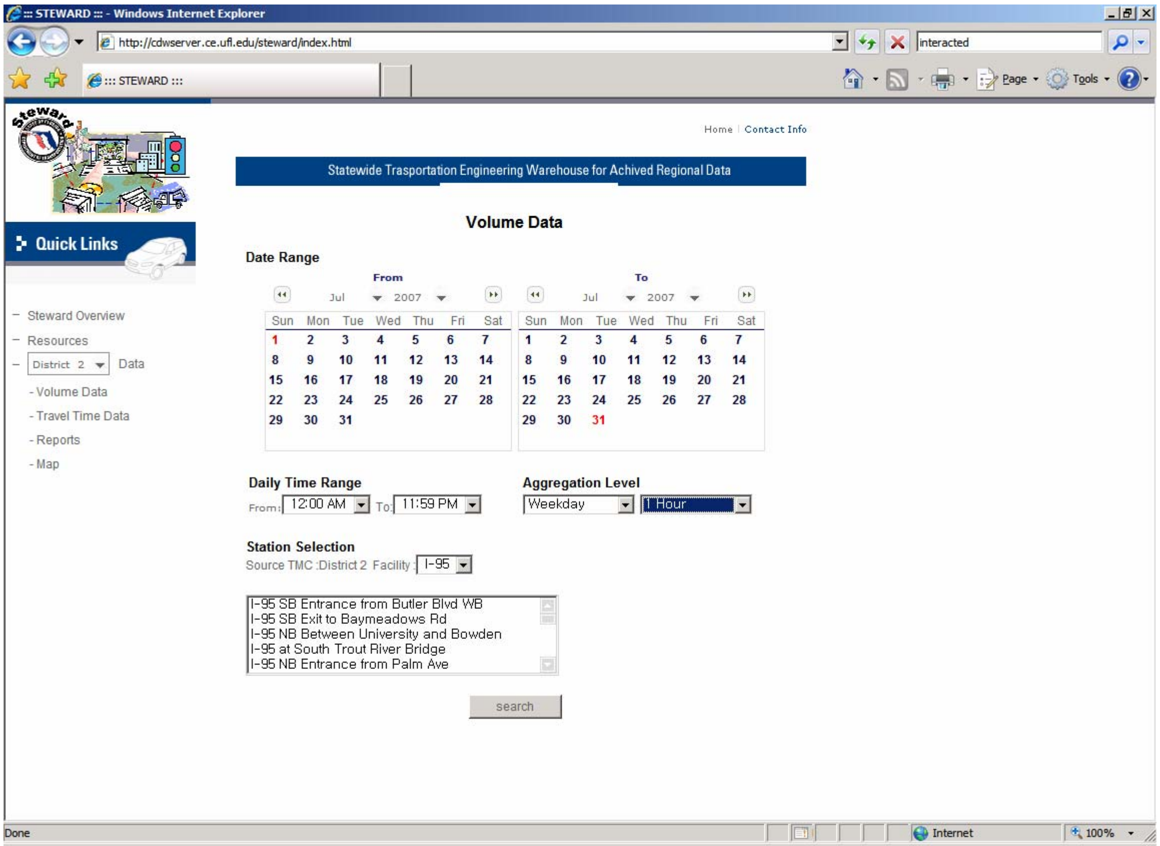


Figure D- 27: Example 1 - Select time and aggregation level

5.1.5 Set the Station Selection

As shown in Figure D-28, select all the stations from the station selection box.

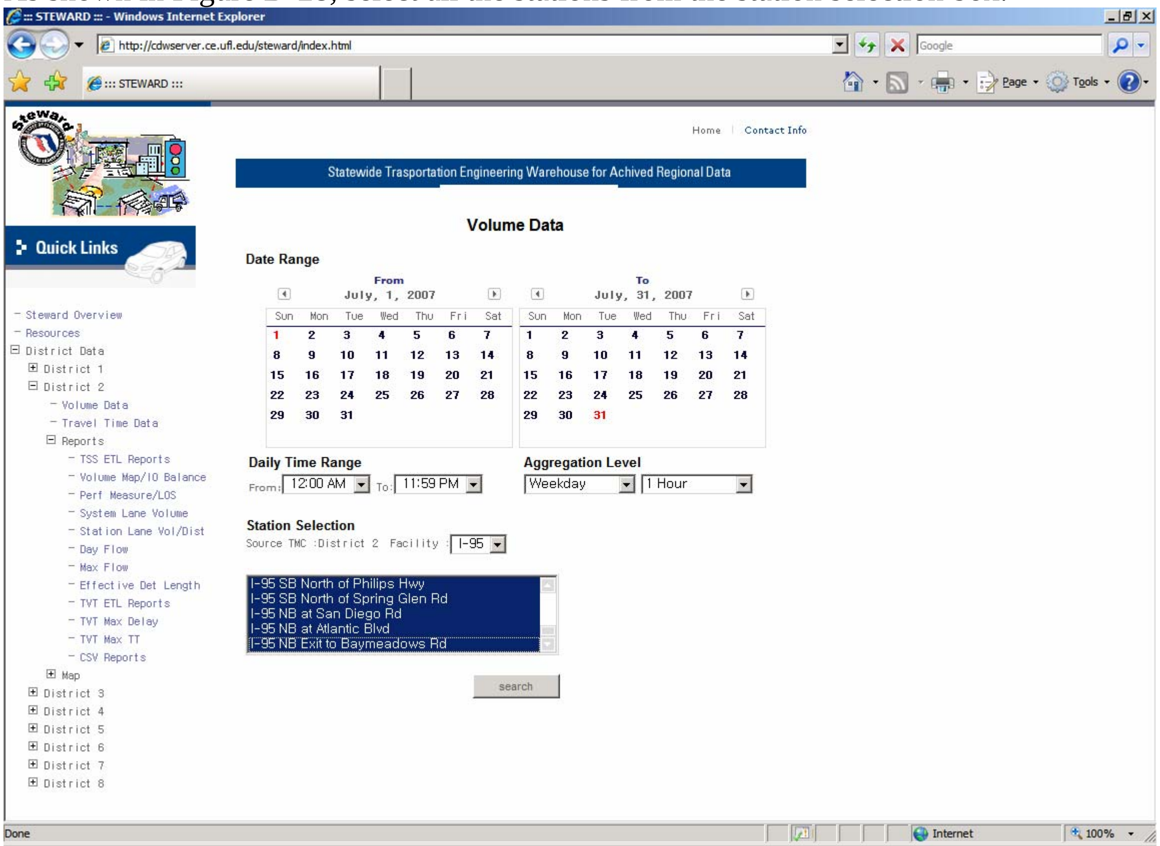
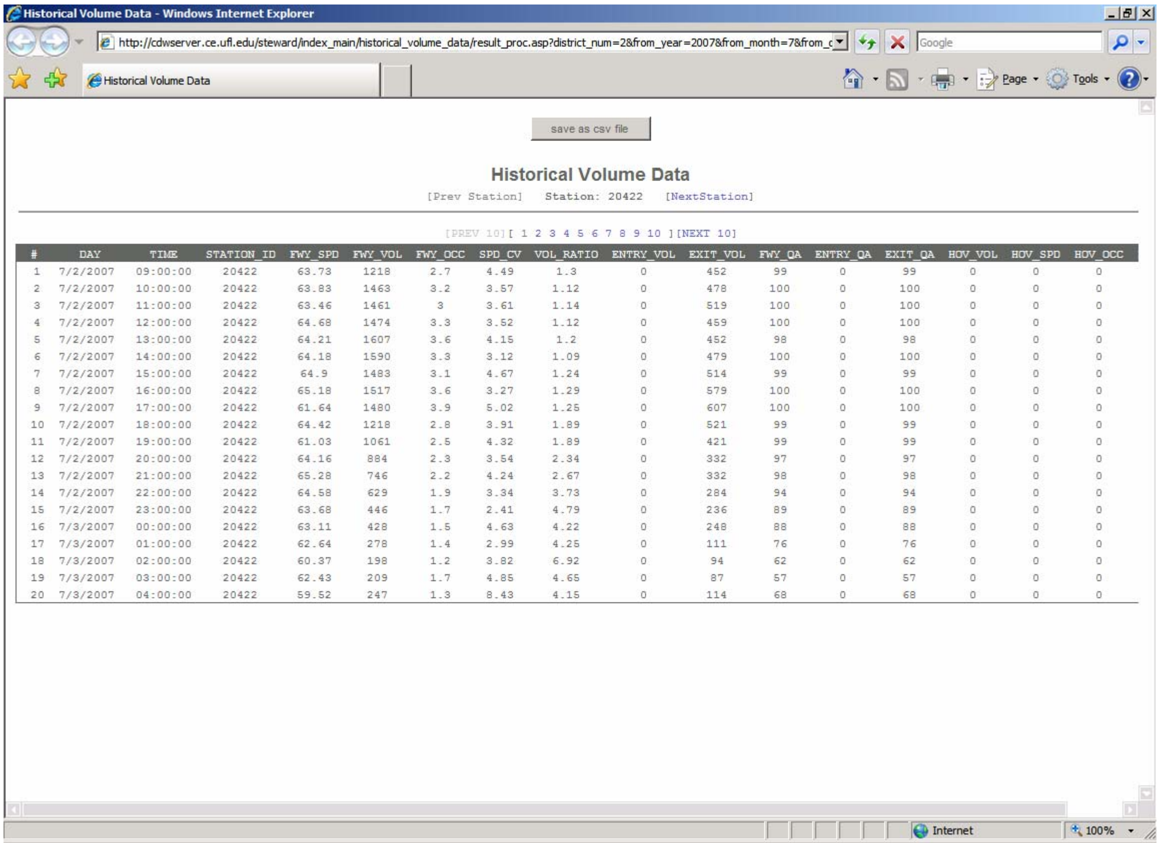


Figure D- 28: Example 1 - Select the stations

5.1.6 Retrieve the data from STEWARD

Click the ***Search*** button. The result will be shown as Figure D-29. The column explanations are given in Table D-1 (Volume Data Table) of this document



#	DAY	TIME	STATION_ID	FWY_SPD	FWY_VOL	FWY_OCC	SPD_CV	VOL_RATIO	ENTRY_VOL	EXIT_VOL	FWY_QA	ENTRY_QA	EXIT_QA	HOV_VOL	HOV_SPD	HOV_OCC
1	7/2/2007	09:00:00	20422	63.73	1218	2.7	4.49	1.3	0	452	99	0	99	0	0	0
2	7/2/2007	10:00:00	20422	63.83	1463	3.2	3.57	1.12	0	478	100	0	100	0	0	0
3	7/2/2007	11:00:00	20422	63.46	1461	3	3.61	1.14	0	519	100	0	100	0	0	0
4	7/2/2007	12:00:00	20422	64.68	1474	3.3	3.52	1.12	0	459	100	0	100	0	0	0
5	7/2/2007	13:00:00	20422	64.21	1607	3.6	4.15	1.2	0	452	98	0	98	0	0	0
6	7/2/2007	14:00:00	20422	64.18	1590	3.3	3.12	1.09	0	479	100	0	100	0	0	0
7	7/2/2007	15:00:00	20422	64.9	1483	3.1	4.67	1.24	0	514	99	0	99	0	0	0
8	7/2/2007	16:00:00	20422	65.18	1517	3.6	3.27	1.29	0	579	100	0	100	0	0	0
9	7/2/2007	17:00:00	20422	61.64	1480	3.9	5.02	1.25	0	607	100	0	100	0	0	0
10	7/2/2007	18:00:00	20422	64.42	1218	2.8	3.91	1.89	0	521	99	0	99	0	0	0
11	7/2/2007	19:00:00	20422	61.03	1061	2.5	4.32	1.89	0	421	99	0	99	0	0	0
12	7/2/2007	20:00:00	20422	64.16	884	2.3	3.54	2.34	0	332	97	0	97	0	0	0
13	7/2/2007	21:00:00	20422	65.28	746	2.2	4.24	2.67	0	332	98	0	98	0	0	0
14	7/2/2007	22:00:00	20422	64.58	629	1.9	3.34	3.73	0	284	94	0	94	0	0	0
15	7/2/2007	23:00:00	20422	63.68	446	1.7	2.41	4.79	0	236	89	0	89	0	0	0
16	7/3/2007	00:00:00	20422	63.11	428	1.5	4.63	4.22	0	248	88	0	88	0	0	0
17	7/3/2007	01:00:00	20422	62.64	278	1.4	2.99	4.25	0	111	76	0	76	0	0	0
18	7/3/2007	02:00:00	20422	60.37	198	1.2	3.82	6.92	0	94	62	0	62	0	0	0
19	7/3/2007	03:00:00	20422	62.43	209	1.7	4.85	4.65	0	87	57	0	57	0	0	0
20	7/3/2007	04:00:00	20422	59.52	247	1.3	8.43	4.15	0	114	68	0	68	0	0	0

Figure D- 29: Example 1 - Volume data

5.1.7 Click "Save as the CSV file" button

As shown in Figure D-30, the CSV file is generated and ready to download.
The new file is located on the top of the table.
Right click on the file name to save it as a CSV file.

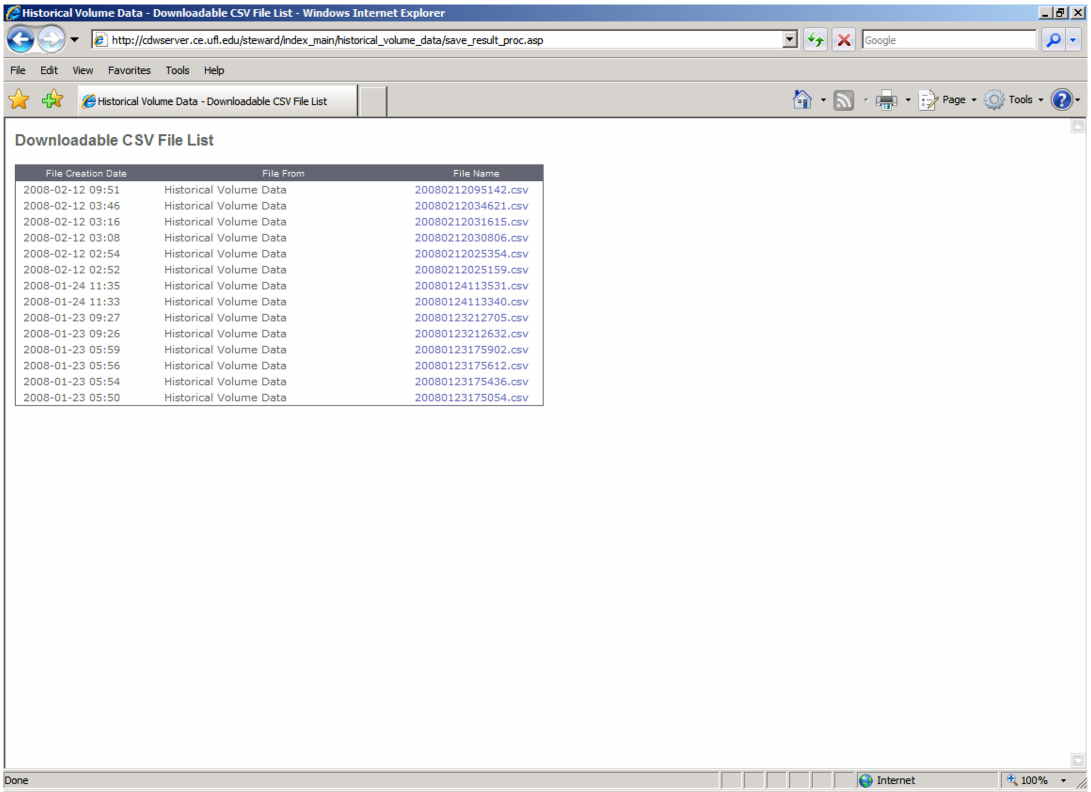


Figure D- 30: Example 1 - Save the volume data to a CSV file

5.2 Example 2 Retrieve Lane-by-lane volume data for District 2

5.2.1 Open the following web site:

http://cdwserver.ce.ufl.edu/steward/index.html

5.2.2 Select the following sequence from the left pane to get the volume data:

District 2 data → Reports → Station Lane Volume and Distribution
The result will be as Figure D-31.

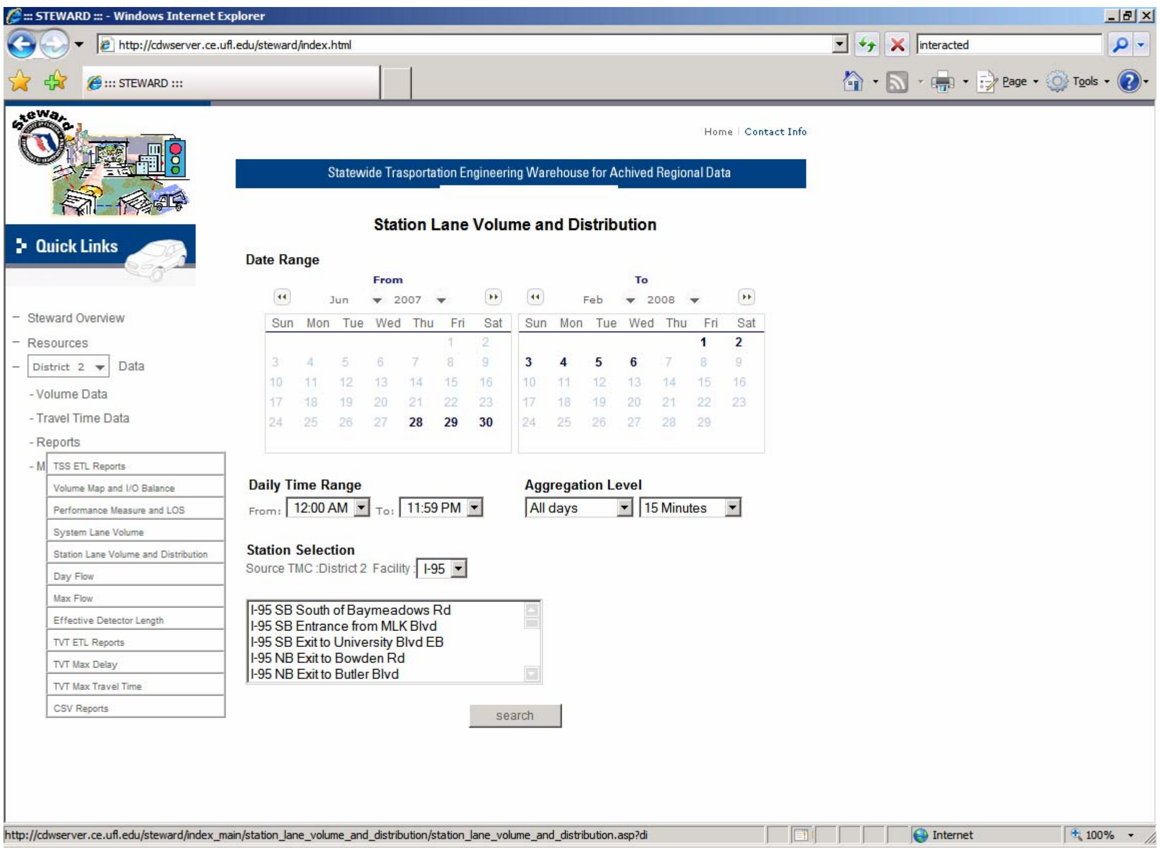


Figure D- 31: Example 2 - Select station lane volume and distribution

5.2.3 Set up the user requirements

As shown in Figure D-32, select the starting and ending dates from ***Date Range***
The “From” calendar shows the month that includes the first day of the data.
The “To” calendar shows the month that includes the last day of the data.
The days with bold font in the calendar indicate that data are available on those days.

Click the starting and ending days on the calendar as July, 1st 2007 and July, 31st 2007.
You can navigate the month and year using the pull-down menu or arrow keys.

Select the daily time range and aggregation level
Keep the default daily time range as 12:00AM and 11:59PM to retrieve the entire day’s data.
Select ***Weekday*** and ***1Hour*** for the aggregation levels.

Set the Station Selection
Select all the stations from the station selection box.

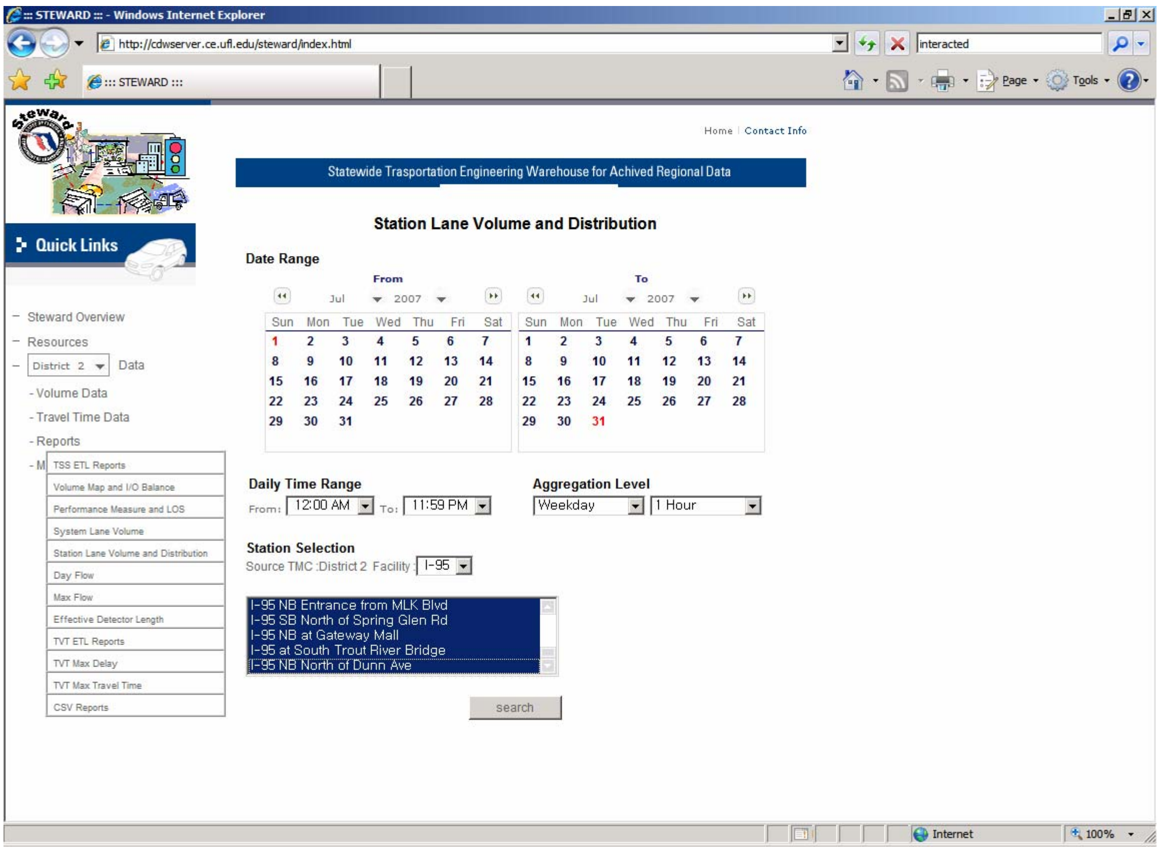
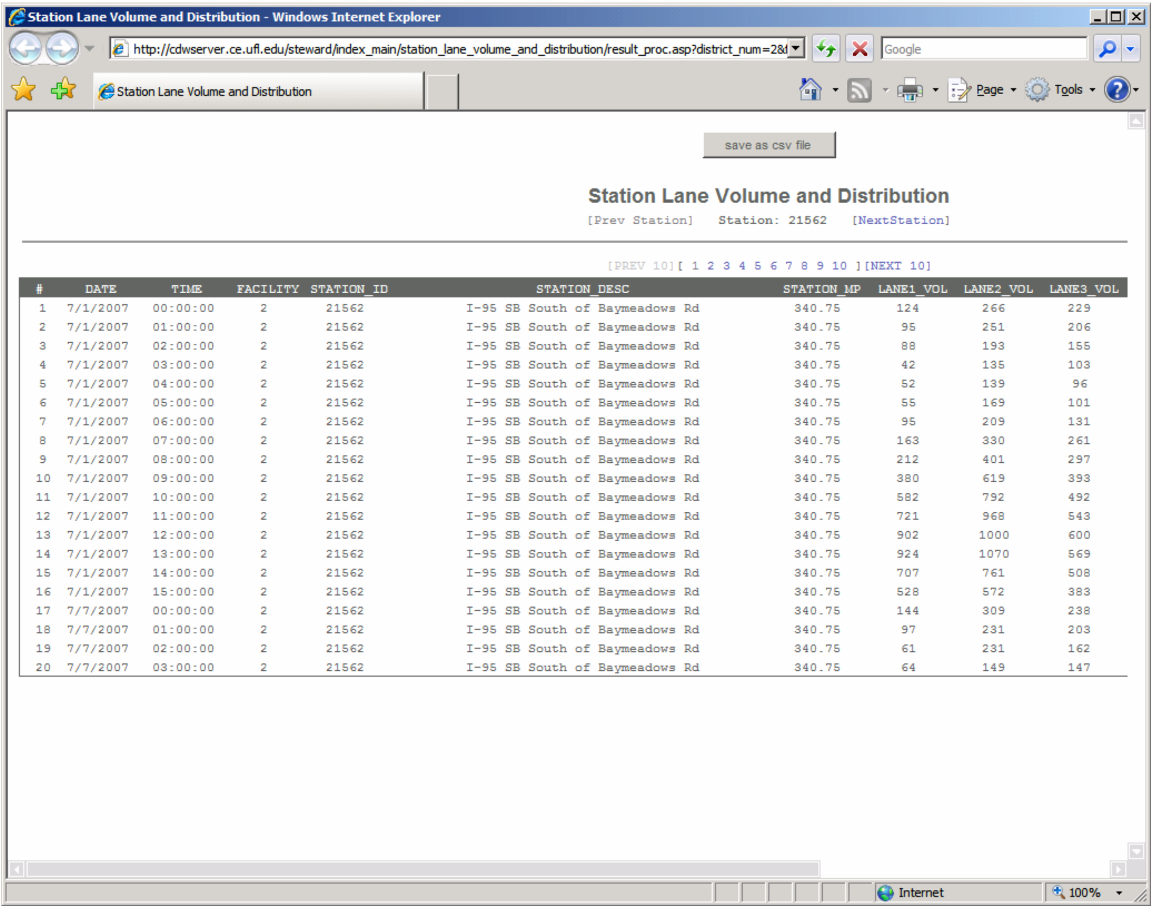


Figure D- 32: Example 2 - Select the user requirements

5.2.4 Click the *Search* button

Click the ***Search*** button. The results will be as Figure D-33. The column explanations are given in Table D-7 of this document.



#	DATE	TIME	FACILITY	STATION_ID	STATION_DESC	STATION_MP	LANE1_VOL	LANE2_VOL	LANE3_VOL
1	7/1/2007	00:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	124	266	229
2	7/1/2007	01:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	95	251	206
3	7/1/2007	02:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	88	193	155
4	7/1/2007	03:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	42	135	103
5	7/1/2007	04:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	52	139	96
6	7/1/2007	05:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	55	169	101
7	7/1/2007	06:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	95	209	131
8	7/1/2007	07:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	163	330	261
9	7/1/2007	08:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	212	401	297
10	7/1/2007	09:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	380	619	393
11	7/1/2007	10:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	582	792	492
12	7/1/2007	11:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	721	968	543
13	7/1/2007	12:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	902	1000	600
14	7/1/2007	13:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	924	1070	569
15	7/1/2007	14:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	707	761	508
16	7/1/2007	15:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	528	572	383
17	7/7/2007	00:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	144	309	238
18	7/7/2007	01:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	97	231	203
19	7/7/2007	02:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	61	231	162
20	7/7/2007	03:00:00	2	21562	I-95 SB South of Baymeadows Rd	340.75	64	149	147

Figure D- 33: Example 2 - Station lane column and distribution data

5.2.5 Click "Save as the CSV file" button

As shown in Figure D-34, the comma separated values (CSV) file is generated and ready to download. The new file is located on the top of the table. Right click on the file name to save it as a CSV file.

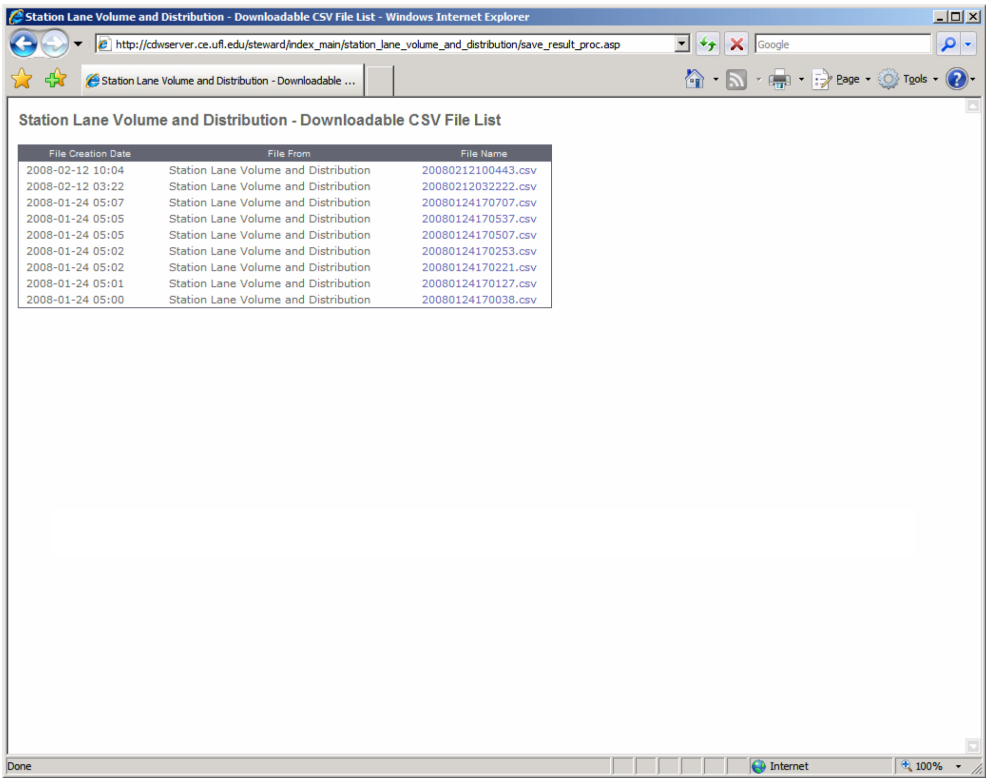
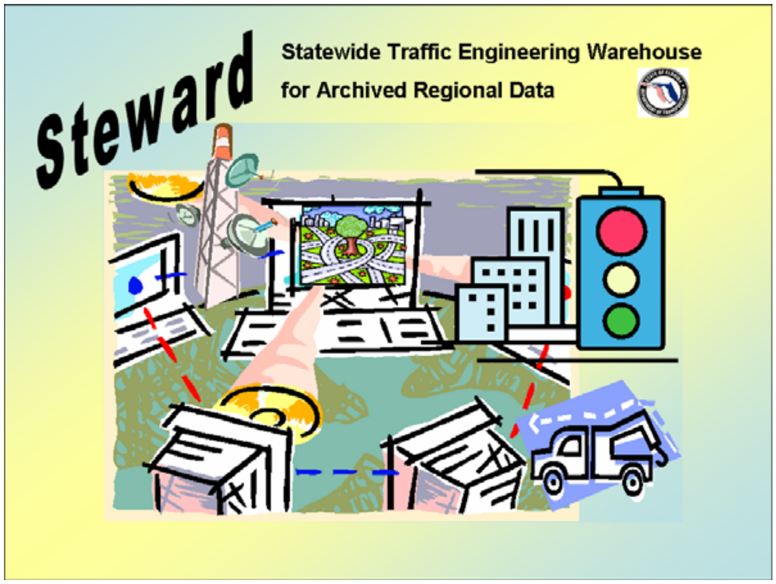


Figure D- 34: Example 2 - Save the lane column and distribution data to CSV file



Development of a Central Data Warehouse for Statewide ITS and Transportation Data in Florida

Phase II: Proof of Concept Final Report

Appendix E
STEWARD Design Document

Prepared for the Florida Department of Transportation
By the University of Florida Transportation Research Center

Contract # BD545, RPWO # 37
UF Project 051449

Revision Date: April 15, 2008

Table of Contents

List of Figures

List of Tables

1 Introduction

This document describes the overall design process of Statewide Transportation Engineering Warehouse for Archived Regional Data (STEWARD). The overall system design was developed with three categories: extraction, transformation and loading (ETL) process, Database design/implementation/management and web user interface. This document will describe the design process, system architecture and implementation.

One of the main design tasks was to select the appropriate tools to develop, deploy and maintain the STEWARD system efficiently. The Oracle database was selected as the base requirement at the beginning of the project. At the initial STEWARD project meeting with FDOT, The Windows 2003 Server and Microsoft Internet Information Services were selected for the operation system and web server. Based on these decisions, The Oracle Warehouse Builder 10g2, Oracle Enterprise manager and ASP/JavaScript were selected for the integrated ETL processes, the database management and the web development.

Even though the traffic data from each District are not yet fully available, the STEWARD system was designed and developed to accommodate the statewide traffic data. Therefore, the naming conventions, database structure and ETL process have the requirements for the statewide data warehouse.

The overall operation of STEWARD is divided into three categories: ETL process, database management system (DBMS) and web user interface. The ETL process acquires traffic data from local databases – District databases or FDOT databases, to transform the data into predefined format and to load into the STEWARD database.

The design and implementation of the database are closely related to that of ETL process. Therefore, these tasks were developed together using the Oracle Warehouse Builder 10g2 platform. For example, the loading process transfers the external data files into the target database and the Oracle Warehouse Builder supports the design of the external data file, target data table and transfer process between these two items.

The web interface was developed for user convenience and speed. Therefore, most of the data processing is executed and archived beforehand. When the user requests specific data, the web interface delivers the data from the pre-calculated data tables.

The detail design of these three categories will be reviewed in the following sections. By the nature of database design, the STEWARD database design and management will be explained first, STEWARD ETL and web interface will be explained afterward.

2 STEWARD Database Development and Management

All of the database design and implementation was developed using the Oracle Warehouse Builder, which is an integrated tool to support warehouse development using a graphic user interface.

This tool includes predefined rules that are generally required in the warehouse design. For example, the keys for the dimensional tables have a predefined name (same as the table name) and data type (integer). The contents of these keys are sequential numbers assigned by the Oracle Warehouse Builder. For these limitations, Station_ID for the fact table, TSS_5MIN cannot be the primary key and user needs to refer the dimension table TSS_STATION to retrieve the Station_ID of the traffic data.

Table types used in STEWARD include:

- External tables,
- Dimension tables,
- Fact tables and
- materialized views

2.1 External Tables

These tables are used to import text files into the Oracle database. The file name and all the columns attributes (name, data type and length) need be defined to be loaded into the database. The details of each table used in STEWARD are as follows:

2.1.1 TIME_TABLE

This table is used to import the time text file into the Oracle database.

In a data warehouse, date (day and time) is generally used as one of the keys in the fact table. In STEWARD, date (Date and Time) is also used as one of the keys to identify the traffic data. But they are divided into Date and Time fields in STEWARD to minimize the reference table size. For example, five-years of traffic data with 5 minute resolution are to be archived, it would need 525,600 rows (5years*365days * 24hrs * 12) in one date table. But if the time and date are divided into two tables, it requires 1825 rows of date table (5years*365days) and 288 rows of time table. This would decrease lookup time.

As shown in Table E-1, the time field includes the character stings of timestamps from "00:00:00" to "23:59:00" with 1 minute resolution.

Table E- 1: STEWARD external table – Contents of TIME_TABLE

Name	Data Type (size)
TIME	Varchar2(20)

2.1.2 TSS_STATION_EX

As shown in Table E-2, this table is used to import the station information from the district facility file into the Oracle database as a dimension table.

Table E- 2: STEWARD external table – Contents of TSS_STATION_EX

Name	Data Type (Length/Precision)
STATION_INDEX	NUMBER
STATIONCDW_NUM	NUMBER
DESCRIPTION	VARCHAR2(200)
STATUS	NUMBER
ROAD	VARCHAR2(40)
DETECTOR_UNIT	NUMBER
LATITUDE	FLOAT(5)
LONGITUDE	FLOAT(5)
STATE_MILEPOST	FLOAT(5)
ROADWAY_ID	VARCHAR2(40)
ROADWAY_MILEPOST	FLOAT(5)
MAX_SPEED	FLOAT(5)
NUM_OF_LANES	NUMBER
UPNODE	NUMBER
LANE_CAPACITY	NUMBER
DETECTOR_TYPE	VARCHAR2(40)

2.1.3 TSS_5MIN_EX, TSS_15MIN_EX and TSS_1HR_EX,

As shown in Table E-3, these tables are used to import the station-level 5min, 15min and 1hr traffic data files into the Oracle database as fact tables.

Table E- 3: STEWARD external table – Contents of TSS_5MIN_EX, TSS_15MIN_EX, and TSS_1HR_EX

Name	Data Type (Length/Precision)
DATE1	VARCHAR2(255)
TIME	VARCHAR2(255)
CDWID	NUMBER
DIRECTION	NUMBER
FWYSPD	FLOAT(5)
FWYVOL	NUMBER
FWYOCC	FLOAT(5)
SPDCV	FLOAT(5)
VOLRATIO	FLOAT(5)
SPDRATIO	FLOAT(5)
ENTRYVOL	NUMBER
EXITVOL	NUMBER
FWYQA	FLOAT(5)
ENTRYQA	FLOAT(5)
EXITQA	FLOAT(5)
HOVVOL	NUMBER
HOVSPD	FLOAT(5)
HOVOCC	FLOAT(5)
HOVQA	FLOAT(5)

2.1.4 TSS_5MIN_LANE_EX, TSS_15MIN_LANE_EX and TSS_1HR_LANE_EX

As shown in Table E-4, these tables are used to import the lane-level 5min, 15min and 1hr traffic data files into the Oracle database as fact tables.

Table E- 4: STEWARD external table – Contents of TSS_5MIN_LANE_EX, TSS_15MIN_LANE_EX, and TSS_1HR_LANE_EX

Name	Data Type (Length/Precision)
DATE1	VARCHAR2(255)
TIME	VARCHAR2(255)
LANE_ID	NUMBER
SPEED	FLOAT(5)
VOLUME	NUMBER
OCCUPANCY	FLOAT(5)
OBSERVATIONS	FLOAT(5)
SPD_CV	FLOAT(5)

2.1.5 TSS_ETL_REPORTS_EX

As shown in Table E-5, this table is used to provide the TSS ETL report information to web user. All the report information that web users can access is provided by the STEWARD database for easy maintenance.

Table E- 5: STEWARD external table – Contents of TSS_ETL_REPORTS_EX

Name	Data Type (Length/Precision)
NAME	VARCHAR2(255)
SOURCE	VARCHAR2(255)
PATH	VARCHAR2(255)
CREATION_DATE	DATE
EXPIRATION_DATE	DATE

2.1.6 TVT_LINK_EX, TVT_ORIGIN_EX, and TVT_DESINATION_EX

As shown in Table E-6, 7 and 8, these tables are used to import the link, origin and destination information from the district facility file into the Oracle database as a dimension table.

Table E- 6: STEWARD external table – Contents of TVT_LINK_EX

Name	Data Type (Length/Precision)
ARCHIVE_REF	VARCHAR2(255)
ORIGINID	NUMBER(38)
DESTINATIONID	NUMBER(38)
STATUS	NUMBER(38)
DIRECTION	VARCHAR2(40)
FF_TIME	FLOAT(5)
SPEED_LIMIT	FLOAT(5)
LINK_LENGTH	FLOAT(5)

Table E- 7: STEWARD external table – Contents of TVT_ORIGIN_EX

Name	Data Type (Length/Precision)
ORIGIN_ID	NUMBER(38)
ARCHIVE_REF	VARCHAR2(40)
SIGN_NAME	VARCHAR2(40)
LOCATION	VARCHAR2(200)
ROADWAY	VARCHAR2(40)
DIRECTION	NUMBER(38)
LATITUDE	FLOAT(5)
LONGITUDE	FLOAT(5)
MILEPOST	FLOAT(5)
STATION_ID	NUMBER(38)

Table E- 8: STEWARD external table – Contents of TVT_DESTINATION_EX

Table E- 8	
Name	Data Type (Length/Precision)
DESTINATION_ID	NUMBER(38)
ARCHIVE_REF	VARCHAR2(40)
LOCATION	VARCHAR2(200)
ROADWAY	VARCHAR2(40)
TO_DIR	NUMBER(38)
LATITUDE	FLOAT(5)
LONGITUDE	FLOAT(5)
MILEPOST	FLOAT(5)
STATION_ID	NUMBER(38)

2.1.7 TVT_5MIN_EX, TVT_15MIN_EX and TVT_1HR_EX,

As shown in Table E-9, these tables are used to import the 5 min, 15min and 1hr travel-time data files into the Oracle database as fact tables.

Table E- 9: STEWARD external table – Contents of TVT_5MIN_EX, TVT_15MIN_EX, and TVT_1HR_EX

Name	Data Type (Length/Precision)
DATE1	VARCHAR2(20)
TIME	VARCHAR2(20)
ORIGIN	NUMBER(38)
DESTINATION	NUMBER(38)
AV_TT	FLOAT(5)
DELAY	FLOAT(5)
MAX_TT	FLOAT(5)
OCCURRENCE	NUMBER(38)
PCNT_ONTIME	FLOAT(5)
ONTIME_DEL	FLOAT(5)

2.1.8 TVT_ETL_REPORTS_EX

As shown in Table E-10, this table is used to provide the TVT ETL report information to web users. All the report information that web users can access is provided by the STEWARD database for easy maintenance.

Table E- 10: STEWARD external table – Contents of TVT_ETL_REPORTS_EX

Name	Data Type (Length/Precision)
NAME	VARCHAR2(255)
SOURCE	VARCHAR2(255)
PATH	VARCHAR2(255)
CREATION_DATE	DATE
EXPIRATION_DATE	DATE

2.2 Dimension Tables

These tables are used to archive the attributes used to constrain and group data for the fact table. In the Oracle Warehouse Builder, these are called as dimensions.

In STEWARD, facility data and date/time data are used as the dimension tables. All the tables except DATE_REF have no hierarchies and therefore, there is only one level for all dimension tables. DATE_REF is created using time wizard for dimension tables and four levels of hierarchies. The details of each table used in STEWARD systems are as follows:

2.2.1 TIME_REF

As shown in Table E-11, this table is used for the timestamp text from “00:00:00” to “23:59:00” with 1 minute resolution.

Table E- 11: STEWARD dimension table – Definitions of TIME_REF

Name	Data Type (Length/Precision)	Key
DIMENSION_KEY	NUMBER	Primary
ID	NUMBER	Surrogate
TIME	VARCHAR2(25)	
DESCRIPTION	VARCHAR2(40)	

2.2.2 DATE_REF

As shown in Table E-12, this table is used for the date data from 2006 for 10 years. This table is designed and deployed in the default format in Warehouse Builder program.

Table E- 12: STEWARD dimension table – Definitions of DATE_REF

Name	Data Type (Length/Precision)	Key
DIMENSION_KEY	NUMBER	Primary
ID	NUMBER	Surrogate
DAY	DATE	
CODE	NUMBER	
START_DATE	DATE	
END_DATE	DATE	
TIME_SPAN	NUMBER	
JULIAN_DATE	NUMBER	
DESCRIPTION	VARCHAR2(2000)	
NAME	VARCHAR2(25)	
DAY_OF_CAL_WEEK	NUMBER	
DAY_OF_CAL_MONTH	NUMBER	
DAY_OF_CAL_QUARTER	NUMBER	
DAY_OF_CAL_YEAR	NUMBER	
CAL_MONTH_NUMBER	NUMBER	
MONTH_OF_QUARTER	NUMBER	
MONTH_OF_YEAR	NUMBER	
CAL_QUARTER_NUMBER	NUMBER	
QUARTER_OF_YEAR	NUMBER	
CAL_YEAR_NUMBER	NUMBER	

2.2.3 TSS_STATION

As shown in Table E-13, this table is used for the TSS station data from TSS facility data file.

Table E- 13: STEWARD dimension table – Definitions of TSS_STATION

Name	Data Type (Length/Precision)	Key
DIMENSION_KEY	NUMBER	Primary
ID	NUMBER	Surrogate
STATION_ID	NUMBER	
DESCRIPTION	VARCHAR2(200)	
STATUS	NUMBER	
ROAD	VARCHAR2(40)	
DETECTOR_TYPE	VARCHAR2(40)	
DETECTOR_UNIT	NUMBER	
LATITUDE	FLOAT(5)	
LONGITUDE	FLOAT(5)	
STATE_MILEPOST	FLOAT(5)	
ROADWAY_ID	VARCHAR2(40)	
ROADWAY_MILEPOST	FLOAT(5)	
SPEED_LIMIT	FLOAT(5)	
NUM_OF_LANES	NUMBER	
UPNODE	NUMBER	
LANE_CAPACITY	NUMBER	

2.2.4 TVT_LINK, TVT_ORIGIN, and TVT_DESTINATION

As shown in Table E-14, 15 and 16, these tables are used for the TVT link, origin and destination from TVT facility data file.

Table E- 14: STEWARD dimension table – Definitions of TVT_LINK

Name	Data Type (Length/Precision)	Key
DIMENSION_KEY	NUMBER	Primary
LINK_ID	NUMBER	Surrogate
LINK_REF	VARCHAR2(255)	
ORIGIN	NUMBER(38)	
DESTINATION	NUMBER(38)	
STATUS	NUMBER(38)	
DIRECTION	VARCHAR2(40)	
FF_TIME	FLOAT(5)	
SPEED_LIMIT	FLOAT(5)	
LINK_LENGTH	FLOAT(5)	

Table E- 15: STEWARD dimension table – Definitions of TVT_ORIGIN

Name	Data Type (Length/Precision)	Key
DIMENSION_KEY	NUMBER	Primary
ID	NUMBER	Surrogate
ORIGIN_ID	NUMBER(38)	
ARCHIVE_REF	VARCHAR2(40)	
SIGN_NAME	VARCHAR2(40)	
LOCATION	VARCHAR2(200)	
ROADWAY	VARCHAR2(80)	
DIRECTION	NUMBER(38)	
LATITUDE	FLOAT(5)	
LONGITUDE	FLOAT(5)	
MILEPOAT	FLOAT(5)	
STATION_ID	NUMBER(38)	

Table E- 16: STEWARD dimension table – Definitions of TVT_DESTINATION

Name	Data Type (Length/Precision)	Key
DIMENSION_KEY	NUMBER	Primary
ID	NUMBER	Surrogate
DESTINATION_ID	NUMBER(38)	
ARCHIVE_REF	VARCHAR2(40)	
LOCATION	VARCHAR2(200)	
ROADWAY	VARCHAR2(40)	
TO_DIR	NUMBER(38)	
LATITUDE	FLOAT(5)	
LONGITUDE	FLOAT(5)	
MILEPOST	FLOAT(5)	
STATION_ID	NUMBER(38)	

2.3 Fact tables

Fact tables archive business facts or measures and foreign keys which refer to candidate keys (normally primary keys) in the dimension tables. In the Oracle Warehouse Builder, these are called as cubes. In STEWARD system, TSS fact tables includes detector data in 5 min, 15 min and 1 hr resolutions and TVT data includes travel time data in 5 min, 15 min and 1 hr resolutions

2.3.1 TSS_5MIN, TSS_15MIN and TSS_1HR

As shown in Table E-17, these tables archive the station-level TSS data with day and timestamp. They have the same foreign keys and attributes. TIME_REF, DATE_REF and TSS_STATION are used as foreign keys.

Table E- 17: STEWARD fact tables – Definitions of TSS_5MIN, TSS_15MIN and TSS_1HR

Name	Data Type (Length/Precision)	Key
DATE_REF	NUMBER	Foreign
TIME_REF	NUMBER	Foreign
TSS_STATION	NUMBER	Foreign
FWY_SPD	FLOAT(5)	
FWY_VOL	NUMBER	
FWY_OCC	FLOAT(5)	
SPD_CV	FLOAT(5)	
VOL_RATIO	FLOAT(5)	
SPD_RATIO	FLOAT(5)	
ENTRY_VOL	NUMBER	
EXIT_VOL	NUMBER	
FWY_QA	FLOAT(5)	
ENTRY_QA	FLOAT(5)	
EXIT_QA	FLOAT(5)	
HOV_SPD	FLOAT(5)	
HOV_VOL	NUMBER	
HOV_OCC	FLOAT(5)	
HOV_QA	FLOAT(5)	

2.3.2 TSS_5MIN_LANE, TSS_15MIN_LANE and TSS_1HR_LANE

As shown in Table E-18, these tables archive the lane-level TSS data with day and timestamp. They have the same foreign keys and attributes. TIME_REF, DATE_REF and TSS_STATION are used as foreign keys.

Table E- 18: STEWARD fact tables – Definitions of TSS_5MIN_LANE, TSS_15MIN_LANE and TSS_1HR_LANE

Name	Data Type (Length/Precision)	Key
DATE_REF	NUMBER	Foreign
TIME_REF	NUMBER	Foreign
TSS_STATION	NUMBER	Foreign
LANE1_SPD	FLOAT(5)	
LANE1_VOL	NUMBER	
LANE1_OCC	FLOAT(5)	
LANE1_OBS	FLOAT(5)	
LANE2_SPD	FLOAT(5)	
LANE2_VOL	NUMBER	
LANE2_OCC	FLOAT(5)	
LANE2_OBS	FLOAT(5)	
LANE3_SPD	FLOAT(5)	
LANE3_VOL	NUMBER	
LANE3_OCC	FLOAT(5)	
LANE3_OBS	FLOAT(5)	
LANE4_SPD	FLOAT(5)	
LANE4_VOL	NUMBER	
LANE4_OCC	FLOAT(5)	
LANE4_OBS	FLOAT(5)	
LANE5_SPD	FLOAT(5)	
LANE5_VOL	NUMBER	
LANE5_OCC	FLOAT(5)	
LANE5_OBS	FLOAT(5)	
LANE6_SPD	FLOAT(5)	
LANE6_VOL	NUMBER	
LANE6_OCC	FLOAT(5)	
LANE6_OBS	FLOAT(5)	
ONRAMP1_SPD	FLOAT(5)	
ONRAMP1_VOL	NUMBER	
ONRAMP1_OCC	FLOAT(5)	
ONRAMP1_OBS	FLOAT(5)	
ONRAMP2_SPD	FLOAT(5)	
ONRAMP2_VOL	NUMBER	
ONRAMP2_OCC	FLOAT(5)	
ONRAMP2_OBS	FLOAT(5)	
ONRAMP3_SPD	FLOAT(5)	
ONRAMP3_VOL	NUMBER	
ONRAMP3_OCC	FLOAT(5)	
ONRAMP3_OBS	FLOAT(5)	
OFFRAMP1_SPD	FLOAT(5)	
OFFRAMP1_VOL	NUMBER	
OFFRAMP1_OCC	FLOAT(5)	
OFFRAMP1_OBS	FLOAT(5)	
OFFRAMP2_SPD	FLOAT(5)	
OFFRAMP2_VOL	NUMBER	
OFFRAMP2_OCC	FLOAT(5)	
OFFRAMP2_OBS	FLOAT(5)	
OFFRAMP3_SPD	FLOAT(5)	
OFFRAMP3_VOL	NUMBER	
OFFRAMP3_OCC	FLOAT(5)	
OFFRAMP3_OBS	FLOAT(5)	

2.3.3 TVT_5MIN, TVT_15MIN and TVT_1HR

As shown in Table E-19, these tables archive the TVT data with day and timestamp. They have the same foreign keys and attributes. TIME_REF, DATE_REF, TVT_ORIGIN and TVT_DESTINATIN are used as foreign keys.

Table E- 19: STEWARD fact tables – Definitions of TVT_5MIN, TVT_15MIN and T_1HR

Name	Data Type (Length/Precision)	Key
DATE_REF	NUMBER	Foreign
TIME_REF	NUMBER	Foreign
TVT_ORIGIN	NUMBER	Foreign
TVT_DESTINATION	NUMBER	Foreign
AVG_TT	FLOAT(5)	
DELAY	FLOAT(5)	
MAX_TT	FLOAT(5)	
OCCURRENCE	NUMBER(38)	
PCNT_ONTIME	FLOAT(5)	
ONTIME_DELAY	FLOAT(5)	

2.4 Materialized views

While a view in database system is a virtual table representing the result of a database query. A materialized view takes a different approach in which the query result is cached as a concrete table that may be updated from the original base tables from time to time. This enables much more efficient access, at the cost of some data being potentially out-of-date. It is most useful in data warehousing scenarios, where frequent queries of the actual base tables are extremely expensive.

In STEWARD, reports in the web user interface are developed with this materialized view. There are two options to update these tables- *Refresh on demand* and *Refresh on commit.* All these tables are defined as *Refresh on demand.* There are two reasons to select *Refresh on demand.* First, if the materialized views are updated when their dependant tables, such as fact tables are updated, it's quite difficult to estimate the execution time of the data loading process. STEWARD system expects the system operators to upload the low level tables, such as fact tables or dimensional tables first and updates related materialized view later. Second, some materialized views have dependencies with other materialized views. Most of the materialized views depend on the fact tables and dimension tables but some of them (PM_LOS_DAY, SYSTEM_LANE_DAY, LANE_DISTRIBUTION_DAY and EFF_DET_LEN_DAY) depend on its materialized views. It's not easy to set the sequences on *Refresh on commit* to update the materialized views in the order of table dependencies.

Table E-20 shows STEWARD materialized views for TSS reports and its dependent tables

Table E- 20: STEWARD materialized views and its dependency for TSS reports

Materialized views	Dependent tables	Dependent Materialized views
VOLUME_MAP_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN	None
VOLUME_MAP_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN	None
VOLUME_MAP_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR	None
VOLUME_MAP_DAY	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR	None
PM_LOS_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN	None
PM_LOS_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR	None
PM_LOS_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN	None
PM_LOS_DAY	None	PM_LOS_5MIN
SYSTEM_LANE_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN_LANE	None
SYSTEM_LANE_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN_LANE	None
SYSTEM_LANE_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR_LANE	None
SYSTEM_LANE_DAY	None	SYSTEM_LANE_1HR
LANE_DISTRIBUTION_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN_LANE	None
LANE_DISTRIBUTION_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN_LANE	None
LANE_DISTRIBUTION_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR_LANE	None
LANE_DISTRIBUTION_DAY	None	LANE_DISTRIBUTION_1HR
DAY_FLOW_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN_LANE	None
DAY_FLOW_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN_LANE	None
DAY_FLOW_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR_LANE	None
MAX_FLOW_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN_LANE	None
MAX_FLOW_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN_LANE	None
MAX_FLOW_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR_LANE	None
EFF_DET_LEN_5MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_5MIN_LANE	None
EFF_DET_LEN_15MIN	TSS_STATION, DATE_REF, TIME_REF, TSS_15MIN_LANE	None
EFF_DET_LEN_1HR	TSS_STATION, DATE_REF, TIME_REF, TSS_1HR_LANE	None
EFF_DET_LEN_DAY	None	EFF_DET_LEN_1HR

Table E-21 shows STEWARD materialized views for TVT reports and its dependent tables

Table E- 21: STEWARD materialized views and its dependency for TVT reports

Materialized views	Dependent tables	Dependent Materialized views
TVT_MAX_DELAY_5MIN	TVT_ORIGIN, TVT_DESTINATION, TVT_LINK, DATE_REF, TVT_LINK, TIME_REF, TVT_5MIN	None
TVT_MAX_DELAY_15MIN	TVT_ORIGIN, TVT_DESTINATION, TVT_LINK, DATE_REF, TVT_LINK, TIME_REF, TVT_15MIN	None
TVT_MAX_DELAY_1HR	TVT_ORIGIN, TVT_DESTINATION, TVT_LINK, DATE_REF, TVT_LINK, TIME_REF, TVT_1HR	None
TVT_MAX_TT_5MIN	TVT_ORIGIN, TVT_DESTINATION, TVT_LINK, DATE_REF, TVT_LINK, TIME_REF, TVT_5MIN	None
TVT_MAX_TT_15MIN	TVT_ORIGIN, TVT_DESTINATION, TVT_LINK, DATE_REF, TVT_LINK, TIME_REF, TVT_15MIN	None
TVT_MAX_TT_1HR	TVT_ORIGIN, TVT_DESTINATION, TVT_LINK, DATE_REF, TVT_LINK, TIME_REF, TVT_1HR	None

2.4.1 VOLUME_MAP_5MIN, VOLUME_MAP_15MIN, VOLUME_MAP_1HR, and VOLUME_MAP_DAY

These tables provide volume map for web user queries. Table E-22 shows STEWARD materialized views definition for VOLUME_MAP_5MIN, VOLUME_MAP_15MIN, VOLUME_MAP_1HR, and VOLUME_MAP_DAY.

Table E- 22: STEWARD materialized views – Definitions of VOLUME_MAP_5MIN, VOLUME_MAP_15MIN, VOLUME_MAP_1HR, and VOLUME_MAP_DAY

Name	Data Type (Length/Precision)	
DATE1	DATE	
TIME	VARCHAR2(25)	N/A for VOLUME_MAP_DAY
FACILITY	NUMBER	
STATION_ID	NUMBER	
STATION_MP	FLOAT(5)	
UPNODE_ID	NUMBER	
ENTRY_VOLUME	NUMBER	
FWY_VOLUME	NUMBER	
EXIT_VOLUME	NUMBER	
LINK_INPUT	NUMBER	
LINK_OUTPUT	NUMBER	
DIFFERENCE	NUMBER	
PCNT_DIFF	NUMBER	

Table E-23 shows the query to generate the materialized view- VOLUME_MAP_5MIN.

Table E- 23: Query for STEWARD materialized views-VOLUME_MAP_5MIN

```
SELECT
      DATE_TBL.DAY,
      TIME_TBL.TIME,
      MOD(STATION.STATION_ID,10),
      STATION.STATION_ID,
      STATION.STATE_MILEPOST,
      UPNODE.STATION_ID,
      TSS_5MIN.ENTRY_VOL,
      TSS_5MIN.FWY_VOL,
      TSS_5MIN.EXIT_VOL,
      UPNODE_TSS_5MIN.ENTRY_VOL + UPNODE_TSS_5MIN.FWY_VOL,
      TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL,
      UPNODE_TSS_5MIN.FWY_VOL + UPNODE_TSS_5MIN.ENTRY_VOL - TSS_5MIN.EXIT_VOL -
TSS_5MIN.FWY_VOL,
      CASE
            WHEN (UPNODE_TSS_5MIN.FWY_VOL + UPNODE_TSS_5MIN.ENTRY_VOL + TSS_5MIN.EXIT_VOL
+ TSS_5MIN.FWY_VOL) > 0
           THEN (100 * 2*(UPNODE_TSS_5MIN.FWY_VOL + UPNODE_TSS_5MIN.ENTRY_VOL -
TSS_5MIN.EXIT_VOL - TSS_5MIN.FWY_VOL) / (UPNODE_TSS_5MIN.FWY_VOL +
UPNODE_TSS_5MIN.ENTRY_VOL + TSS_5MIN.EXIT_VOL + TSS_5MIN.FWY_VOL))
           ELSE NULL
      END

FROM GATOR.TSS_STATION STATION, GATOR.TSS_STATION UPNODE, GATOR.TSS_5MIN TSS_5MIN,
GATOR.TSS_5MIN UPNODE_TSS_5MIN, GATOR.DATE_REF DATE_TBL, GATOR.TIME_REF TIME_TBL

WHERE TSS_5MIN.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND STATION.UPNODE = UPNODE.STATION_ID
            AND UPNODE_TSS_5MIN.TSS_STATION = UPNODE.DIMENSION_KEY
            AND UPNODE_TSS_5MIN.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND UPNODE_TSS_5MIN.TIME_REF = TIME_TBL.DIMENSION_KEY
```

2.4.2 PM_LOS_5MIN, PM_LOS_15MIN, PM_LOS_1HR, and PM_LOS_DAY

These tables provide performance measure and LOS for web user queries. Table E-24 shows STEWARD materialized views definition for PM_LOS_5MIN, PM_LOS_15MIN, and PM_LOS_1HR. Table E-25 shows STEWARD materialized views definition for PM_LOS_DAY

Table E- 24: STEWARD materialized views – Definitions of PM_LOS_5MIN, PM_LOS_15MIN, and PM_LOS_1HR

Name	Data Type (Length/Precision)
DATE1	DATE
TIME	VARCHAR2(25)
FACILITY	NUMBER
STATION_ID	NUMBER
STATION_DESC	VARCHAR2(200)
UPNODE_ID	NUMBER
UPNODE_DESC	VARCHAR2(200)
STATION_MP	FLOAT(5)
LENGTH	NUMBER
LINK_VOL	NUMBER
NUM_OF_LANES	NUMBER
VOL_PER_LANE	NUMBER
VEH_MI	NUMBER
VEH_HR	NUMBER
SPEED	FLOAT(5)
DELAY	NUMBER
KENERGY	NUMBER
DENSITY	NUMBER
VOL_TO_CAP	NUMBER
LOS	CHAR(1)

Table E- 25: STEWARD materialized views – Definitions of PM_LOS_DAY

Name	Data Type (Length/Precision)
DATE1	DATE
FACILITY	NUMBER
STATION_ID	NUMBER
STATION_DESC	VARCHAR2(200)
UPNODE_ID	NUMBER
UPNODE_DESC	VARCHAR2(200)
STATION_MP	FLOAT(5)
LENGTH	NUMBER
LINK_VOL	NUMBER
NUM_OF_LANES	NUMBER
VOL_PER_LANE	NUMBER
VEH_MI	NUMBER
VEH_HR	NUMBER
SPEED	FLOAT(5)
DELAY	NUMBER
KENERGY	NUMBER
F_PLUS	NUMBER
E_PLUS	NUMBER
D_PLUS	NUMBER
C_PLUS	NUMBER

Table E-25 shows the query to generate the materialized view- PM_LOS_5MIN.

Table E- 26: Query for STEWARD materialized views-PM_LOS_5MIN

```
SELECT
      DATE_TBL.DAY,             /* Date */
      TIME_TBL.TIME,            /* Time */
      MOD(STATION.STATION_ID,10),       /* FACILITY */
      STATION.STATION_ID,       /* Destination */
      STATION.DESCRIPTION,      /* Destination description */
      UPNODE.STATION_ID,        /* Origin */
      UPNODE.DESCRIPTION,       /* Origin description */
      STATION.STATE_MILEPOST,   /* Destination milepost */
      ABS(STATION.STATE_MILEPOST- UPNODE.STATE_MILEPOST),   /* Length */

      ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE_TSS_5MIN.FWY_VOL +
UPNODE_TSS_5MIN.ENTRY_VOL))/2,  /*MainVol*/
      STATION.NUM_OF_LANES,   /*Main Lane*/
      ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE_TSS_5MIN.FWY_VOL +
UPNODE_TSS_5MIN.ENTRY_VOL))/2 /STATION.NUM_OF_LANES,  /*Vol/Lane */
      ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE_TSS_5MIN.FWY_VOL +
UPNODE_TSS_5MIN.ENTRY_VOL))/2* ABS(STATION.STATE_MILEPOST- UPNODE.STATE_MILEPOST),
/*Veh-Mi*/
      ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE_TSS_5MIN.FWY_VOL +
UPNODE_TSS_5MIN.ENTRY_VOL))/2* ABS(STATION.STATE_MILEPOST- UPNODE.STATE_MILEPOST)/
TSS_5MIN.FWY_SPD,       /*Veh-Hr*/
      TSS_5MIN.FWY_SPD,   /*AvSpeed */
      CASE    /*Delay*/
         WHEN TSS_5MIN.FWY_SPD >= (STATION.SPEED_LIMIT-10) then 0
         ELSE TSS_5MIN.FWY_VOL * ABS(STATION.STATE_MILEPOST- UPNODE.STATE_MILEPOST)*(1/
TSS_5MIN.FWY_SPD -1/(STATION.SPEED_LIMIT-10))
      END,
      ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE_TSS_5MIN.FWY_VOL +
UPNODE_TSS_5MIN.ENTRY_VOL))/2* TSS_5MIN.FWY_SPD,    /*KEnergy*/
      4 * ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE_TSS_5MIN.FWY_VOL +
UPNODE_TSS_5MIN.ENTRY_VOL))/2 / TSS_5MIN.FWY_SPD, /*Density*/
      100* 4* ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE_TSS_5MIN.FWY_VOL +
UPNODE_TSS_5MIN.ENTRY_VOL))/2 /(STATION.LANE_CAPACITY* STATION.NUM_OF_LANES),  /*V/C*/

      CASE    /*LOS*/
         WHEN 4 * ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE_TSS_5MIN.FWY_VOL +
UPNODE_TSS_5MIN.ENTRY_VOL))/2 / TSS_5MIN.FWY_SPD > 45 then 'F'
         WHEN 4 * ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE_TSS_5MIN.FWY_VOL +
UPNODE_TSS_5MIN.ENTRY_VOL))/2 / TSS_5MIN.FWY_SPD > 35 then 'E'
         WHEN 4 * ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE_TSS_5MIN.FWY_VOL +
UPNODE_TSS_5MIN.ENTRY_VOL))/2 / TSS_5MIN.FWY_SPD > 26 then 'D'
         WHEN 4 * ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE_TSS_5MIN.FWY_VOL +
UPNODE_TSS_5MIN.ENTRY_VOL))/2 / TSS_5MIN.FWY_SPD > 18 then 'C'
         WHEN 4 * ((TSS_5MIN.FWY_VOL + TSS_5MIN.EXIT_VOL)+ (UPNODE_TSS_5MIN.FWY_VOL +
UPNODE_TSS_5MIN.ENTRY_VOL))/2 / TSS_5MIN.FWY_SPD > 11 then 'B'
         ELSE 'A'
      END
FROM GATOR.TSS_STATION STATION, GATOR.TSS_STATION UPNODE, GATOR.TSS_5MIN TSS_5MIN,
GATOR.TSS_5MIN UPNODE_TSS_5MIN, GATOR.DATE_REF DATE_TBL, GATOR.TIME_REF TIME_TBL
WHERE TSS_5MIN.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND STATION.UPNODE = UPNODE.STATION_ID
            AND UPNODE_TSS_5MIN.TSS_STATION = UPNODE.DIMENSION_KEY
            AND UPNODE_TSS_5MIN.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND UPNODE_TSS_5MIN.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN.FWY_SPD >0
            AND STATION.SPEED_LIMIT-10 >0
            AND STATION.NUM_OF_LANES >0
            AND STATION.LANE_CAPACITY >0
```

2.4.3 SYSTEM_LANE_5MIN, SYSTEM_LANE_15MIN, SYSTEM_LANE_1HR, and SYSTEM_LANE_DAY

These tables provide system lane volume for web user queries. Table E-27 shows STEWARD materialized views definition for SYSTEM_LANE_5MIN, SYSTEM_LANE_15MIN, SYSTEM_LANE_1HR, and SYSTEM_LANE_DAY.

Table E- 27: STEWARD materialized views – Definitions of SYSTEM_LANE_5MIN, SYSTEM_LANE_15MIN, SYSTEM_LANE_1HR, and SYSTEM_LANE_DAY

Name	Data Type (Length/Precision)	
DATE1	DATE	
TIME	VARCHAR2(25)	N/A for VOLUME_MAP_DAY
FACILITY	NUMBER	
STATION_ID	NUMBER	
STATION_DESC	VARCHAR2(200)	
STATION_MP	FLOAT(5)	
LANE1_VOL	NUMBER	
LANE1_SPD	FLOAT(5)	
LANE1_OCC	FLOAT(5)	
LANE2_VOL	NUMBER	
LANE2_SPD	FLOAT(5)	
LANE2_OCC	FLOAT(5)	
LANE3_VOL	NUMBER	
LANE3_SPD	FLOAT(5)	
LANE3_OCC	FLOAT(5)	
LANE4_VOL	NUMBER	
LANE4_SPD	FLOAT(5)	
LANE4_OCC	FLOAT(5)	
LANE5_VOL	NUMBER	
LANE5_SPD	FLOAT(5)	
LANE5_OCC	FLOAT(5)	
LANE6_VOL	NUMBER	
LANE6_SPD	FLOAT(5)	
LANE6_OCC	FLOAT(5)	
ONRAMP1_VOL	NUMBER	
ONRAMP1_SPD	FLOAT(5)	
ONRAMP1_OCC	FLOAT(5)	
ONRAMP2_VOL	NUMBER	
ONRAMP2_SPD	FLOAT(5)	
ONRAMP2_OCC	FLOAT(5)	
ONRAMP3_VOL	NUMBER	
ONRAMP3_SPD	FLOAT(5)	
ONRAMP3_OCC	FLOAT(5)	
OFFRAMP1_VOL	NUMBER	
OFFRAMP1_SPD	FLOAT(5)	
OFFRAMP1_OCC	FLOAT(5)	
OFFRAMP2_VOL	NUMBER	
OFFRAMP2_SPD	FLOAT(5)	
OFFRAMP2_OCC	FLOAT(5)	
OFFRAMP3_VOL	NUMBER	
OFFRAMP3_SPD	FLOAT(5)	
OFFRAMP3_OCC	FLOAT(5)	

Table E-28 shows the query to generate the materialized view- SYSTEM_LANE_5MIN.

Table E- 28: Query for STEWARD materialized views-SYSTEM_LANE_5MIN

```
SELECT
      DATE_TBL.DAY,
      TIME_TBL.TIME,
      MOD(STATION.STATION_ID,10),
      STATION.STATION_ID,
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      TSS_5MIN_LANE.LANE1_VOL,
      TSS_5MIN_LANE.LANE1_SPD,
      TSS_5MIN_LANE.LANE1_OCC,

      TSS_5MIN_LANE.LANE2_VOL,
      TSS_5MIN_LANE.LANE2_SPD,
      TSS_5MIN_LANE.LANE2_OCC,

      TSS_5MIN_LANE.LANE3_VOL,
      TSS_5MIN_LANE.LANE3_SPD,
      TSS_5MIN_LANE.LANE3_OCC,

      TSS_5MIN_LANE.LANE4_VOL,
      TSS_5MIN_LANE.LANE4_SPD,
      TSS_5MIN_LANE.LANE4_OCC,

      TSS_5MIN_LANE.LANE5_VOL,
      TSS_5MIN_LANE.LANE5_SPD,
      TSS_5MIN_LANE.LANE5_OCC,

      TSS_5MIN_LANE.LANE6_VOL,
      TSS_5MIN_LANE.LANE6_SPD,
      TSS_5MIN_LANE.LANE6_OCC,

      TSS_5MIN_LANE.ONRAMP1_VOL,
      TSS_5MIN_LANE.ONRAMP1_SPD,
      TSS_5MIN_LANE.ONRAMP1_OCC,

      TSS_5MIN_LANE.ONRAMP2_VOL,
      TSS_5MIN_LANE.ONRAMP2_SPD,
      TSS_5MIN_LANE.ONRAMP2_OCC,

      TSS_5MIN_LANE.ONRAMP3_VOL,
      TSS_5MIN_LANE.ONRAMP3_SPD,
      TSS_5MIN_LANE.ONRAMP3_OCC,

      TSS_5MIN_LANE.OFFRAMP1_VOL,
      TSS_5MIN_LANE.OFFRAMP1_SPD,
      TSS_5MIN_LANE.OFFRAMP1_OCC,

      TSS_5MIN_LANE.OFFRAMP2_VOL,
      TSS_5MIN_LANE.OFFRAMP2_SPD,
      TSS_5MIN_LANE.OFFRAMP2_OCC,

      TSS_5MIN_LANE.OFFRAMP3_VOL,
      TSS_5MIN_LANE.OFFRAMP3_SPD,
      TSS_5MIN_LANE.OFFRAMP3_OCC

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE, GATOR.DATE_REF
DATE_TBL, GATOR.TIME_REF TIME_TBL

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
```

2.4.4 LANE_DISTRIBUTION_5MIN, LANE_DISTRIBUTION_15MIN, LANE_DISTRIBUTION_1HR, and LANE_DISTRIBUTION_DAY

These tables provide station lane distribution for web user queries. Table E-29 shows STEWARD materialized views definition for LANE_DISTRIBUTION_5MIN, LANE_DISTRIBUTION_15MIN, LANE_DISTRIBUTION_1HR, and LANE_DISTRIBUTION_DAY

Table E- 29: STEWARD materialized views – Definitions of LANE_DISTRIBUTION_5MIN, LANE_DISTRIBUTION_15MIN, LANE_DISTRIBUTION_1HR, and LANE_DISTRIBUTION_DAY

Name	Data Type (Length/Precision)	
DATE1	DATE	
TIME	VARCHAR2(25)	N/A for VOLUME_MAP_DAY
FACILITY	NUMBER	
STATION_ID	NUMBER	
STATION_DESC	VARCHAR2(200)	
STATION_MP	FLOAT(5)	
LANE1_VOL	NUMBER	
LANE2_VOL	NUMBER	
LANE3_VOL	NUMBER	
LANE4_VOL	NUMBER	
LANE5_VOL	NUMBER	
LANE6_VOL	NUMBER	
TOTAL	NUMBER	
BALANCE	NUMBER	

Table E-30 shows the query to generate the materialized view- LANE_DISTRIBUTION_5MIN.

Table E- 30: Query for STEWARD materialized views-LANE_DISTRIBUTION_5MIN

```
SELECT
      DATE_TBL.DAY,
      TIME_TBL.TIME,
      MOD(STATION.STATION_ID,10),
      STATION.STATION_ID,
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      TSS_5MIN_LANE.LANE1_VOL,
      TSS_5MIN_LANE.LANE2_VOL,
      TSS_5MIN_LANE.LANE3_VOL,
      TSS_5MIN_LANE.LANE4_VOL,
      TSS_5MIN_LANE.LANE5_VOL,
      TSS_5MIN_LANE.LANE6_VOL,

      (COALESCE(TSS_5MIN_LANE.LANE1_VOL,0) + COALESCE(TSS_5MIN_LANE.LANE2_VOL,0) +
COALESCE(TSS_5MIN_LANE.LANE3_VOL,0) + COALESCE(TSS_5MIN_LANE.LANE4_VOL,0) +
COALESCE(TSS_5MIN_LANE.LANE5_VOL,0) + COALESCE(TSS_5MIN_LANE.LANE6_VOL,0)) AS TOTAL,

      CASE
          WHEN TSS_5MIN_LANE.LANE1_VOL IS NULL THEN NULL

          WHEN (TSS_5MIN_LANE.LANE2_VOL IS NULL) AND (TSS_5MIN_LANE.LANE1_VOL > 0) THEN 1

          WHEN (TSS_5MIN_LANE.LANE3_VOL IS NULL) AND (LEAST(TSS_5MIN_LANE.LANE1_VOL,
TSS_5MIN_LANE.LANE2_VOL) >0 ) THEN GREATEST (TSS_5MIN_LANE.LANE1_VOL,
TSS_5MIN_LANE.LANE2_VOL)/ LEAST(TSS_5MIN_LANE.LANE1_VOL, TSS_5MIN_LANE.LANE2_VOL)

          WHEN (TSS_5MIN_LANE.LANE4_VOL IS NULL) AND (LEAST(TSS_5MIN_LANE.LANE1_VOL,
TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL) >0 ) THEN GREATEST
(TSS_5MIN_LANE.LANE1_VOL, TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL)/
LEAST(TSS_5MIN_LANE.LANE1_VOL, TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL)

          WHEN (TSS_5MIN_LANE.LANE5_VOL IS NULL) AND (LEAST(TSS_5MIN_LANE.LANE1_VOL,
TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL) >0 ) THEN
GREATEST (TSS_5MIN_LANE.LANE1_VOL, TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL,
TSS_5MIN_LANE.LANE4_VOL)/ LEAST(TSS_5MIN_LANE.LANE1_VOL, TSS_5MIN_LANE.LANE2_VOL,
TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL)

          WHEN (TSS_5MIN_LANE.LANE6_VOL IS NULL) AND (LEAST(TSS_5MIN_LANE.LANE1_VOL,
TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL,
TSS_5MIN_LANE.LANE5_VOL) >0 ) THEN GREATEST (TSS_5MIN_LANE.LANE1_VOL,
TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL,
TSS_5MIN_LANE.LANE5_VOL)/ LEAST(TSS_5MIN_LANE.LANE1_VOL, TSS_5MIN_LANE.LANE2_VOL,
TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL, TSS_5MIN_LANE.LANE5_VOL)

          WHEN LEAST(TSS_5MIN_LANE.LANE1_VOL, TSS_5MIN_LANE.LANE2_VOL,
TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL, TSS_5MIN_LANE.LANE5_VOL,
TSS_5MIN_LANE.LANE6_VOL) >0 THEN GREATEST (TSS_5MIN_LANE.LANE1_VOL,
TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL,
TSS_5MIN_LANE.LANE5_VOL, TSS_5MIN_LANE.LANE6_VOL)/ LEAST(TSS_5MIN_LANE.LANE1_VOL,
TSS_5MIN_LANE.LANE2_VOL, TSS_5MIN_LANE.LANE3_VOL, TSS_5MIN_LANE.LANE4_VOL,
TSS_5MIN_LANE.LANE5_VOL, TSS_5MIN_LANE.LANE6_VOL)

          ELSE NULL
      END

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE, GATOR.DATE_REF
DATE_TBL, GATOR.TIME_REF TIME_TBL

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
```

2.4.5 DAY_FLOW_5MIN, DAY_FLOW_15MIN, and DAY_FLOW_1HR

These tables provide daily flow for web user queries. Table E-31 shows STEWARD materialized views definition for DAY_FLOW_5MIN, DAY_FLOW_15MIN, and DAY_FLOW_1HR.

Table E- 31: STEWARD materialized views – Definitions of DAY_FLOW_5MIN, DAY_FLOW_15MIN, and DAY_FLOW_1HR

Name	Data Type (Length/Precision)	
DATE1	DATE	
TIME	VARCHAR2(25)	N/A for VOLUME_MAP_DAY
FACILITY	NUMBER	
STATION_ID	NUMBER	
STATION_DESC	VARCHAR2(200)	
STATION_MP	FLOAT(5)	
LANE1_FLOW_RATE	NUMBER	
LANE2_FLOW_RATE	NUMBER	
LANE3_FLOW_RATE	NUMBER	
LANE4_FLOW_RATE	NUMBER	
LANE5_FLOW_RATE	NUMBER	
LANE6_FLOW_RATE	NUMBER	

Table E-32 shows the query to generate the materialized view- DAY_FLOW_5MIN.

Table E- 32: Query for STEWARD materialized views-DAY_FLOW_5MIN

```
SELECT
      DATE_TBL.DAY,
      TIME_TBL.TIME,            /* Time */
      MOD(STATION.STATION_ID,10),
      STATION.STATION_ID,
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      12*(TSS_5MIN_LANE.LANE1_VOL),
      12*(TSS_5MIN_LANE.LANE2_VOL),
      12*(TSS_5MIN_LANE.LANE3_VOL),
      12*(TSS_5MIN_LANE.LANE4_VOL),
      12*(TSS_5MIN_LANE.LANE5_VOL),
      12*(TSS_5MIN_LANE.LANE6_VOL)

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE, GATOR.DATE_REF
DATE_TBL, GATOR.TIME_REF TIME_TBL

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
```

2.4.6 MAX_FLOW_5MIN, MAX_FLOW_15MIN, and MAX_FLOW_1HR

These tables provide max flow for web user queries. Table E-33 shows STEWARD materialized views definition for MAX_FLOW_5MIN, MAX_FLOW_15MIN, and MAX_FLOW_1HR.

Table E- 33: STEWARD materialized views – Definitions of MAX_FLOW_5MIN, MAX_FLOW_15MIN, and MAX_FLOW_1HR

Name	Data Type (Length/Precision)
DATE1	DATE
FACILITY	NUMBER
STATION_ID	NUMBER
STATION_DESC	VARCHAR2(200)
STATION_MP	FLOAT(5)
LANE_NUM	NUMBER
MAX_FLOW	NUMBER
MAX_TIME	VARCHAR2(25)
MAX_GAP	NUMBER

Table E-34 shows the query to generate the materialized view- MAX_FLOW_5MIN.

Table E- 34: Query for STEWARD materialized views-MAX_FLOW_5MIN

```
SELECT     /*Lane1*/
      DATE_TBL.DAY,
      MOD(STATION.STATION_ID,10),
      STATION.STATION_ID,
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      1,
      12*TSS_5MIN_LANE.LANE1_VOL,
      TIME_TBL.TIME,
      0

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE, GATOR.DATE_REF
DATE_TBL, GATOR.TIME_REF TIME_TBL,
(SELECT MAX(TSS_5MIN_LANE.LANE1_VOL) as MAX_VOL, TSS_STATION, DATE_REF
FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE1_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
UNION

SELECT     /*Lane2*/
      DATE_TBL.DAY,
      MOD(STATION.STATION_ID,10),
      STATION.STATION_ID,
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      2,
      12*TSS_5MIN_LANE.LANE2_VOL,
      TIME_TBL.TIME,
      0

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE, GATOR.DATE_REF
DATE_TBL, GATOR.TIME_REF TIME_TBL,
(SELECT MAX(TSS_5MIN_LANE.LANE2_VOL) as MAX_VOL, TSS_STATION, DATE_REF
FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW
```

```
WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE2_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
UNION

SELECT     /*Lane3*/
      DATE_TBL.DAY,
      MOD(STATION.STATION_ID,10),
      STATION.STATION_ID,
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      3,
      12*TSS_5MIN_LANE.LANE3_VOL,
      TIME_TBL.TIME,
      0

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE, GATOR.DATE_REF
DATE_TBL, GATOR.TIME_REF TIME_TBL,
(SELECT MAX(TSS_5MIN_LANE.LANE3_VOL) as MAX_VOL, TSS_STATION, DATE_REF
FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE3_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
UNION

SELECT     /*LANE4*/
      DATE_TBL.DAY,
      MOD(STATION.STATION_ID,10),
      STATION.STATION_ID,
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      4,
      12*TSS_5MIN_LANE.LANE4_VOL,
      TIME_TBL.TIME,
      0

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE, GATOR.DATE_REF
DATE_TBL, GATOR.TIME_REF TIME_TBL,
(SELECT MAX(TSS_5MIN_LANE.LANE4_VOL) as MAX_VOL, TSS_STATION, DATE_REF
FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE4_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
UNION

SELECT     /*Lane5*/
      DATE_TBL.DAY,
      MOD(STATION.STATION_ID,10),
      STATION.STATION_ID,
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,
```

```
      5,
      12*TSS_5MIN_LANE.LANE5_VOL,
      TIME_TBL.TIME,
      0

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE, GATOR.DATE_REF
DATE_TBL, GATOR.TIME_REF TIME_TBL,
(SELECT MAX(TSS_5MIN_LANE.LANE5_VOL) as MAX_VOL, TSS_STATION, DATE_REF
FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE5_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
UNION

SELECT     /*Lane6*/
      DATE_TBL.DAY,
      MOD(STATION.STATION_ID,10),
      STATION.STATION_ID,
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      6,
      12*TSS_5MIN_LANE.LANE6_VOL,
      TIME_TBL.TIME,
      0

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE, GATOR.DATE_REF
DATE_TBL, GATOR.TIME_REF TIME_TBL,
(SELECT MAX(TSS_5MIN_LANE.LANE6_VOL) as MAX_VOL, TSS_STATION, DATE_REF
FROM GATOR.TSS_5MIN_LANE GROUP BY TSS_STATION, DATE_REF) MAX_FLOW

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.LANE6_VOL = MAX_FLOW.MAX_VOL
            AND TSS_5MIN_LANE.TSS_STATION = MAX_FLOW.TSS_STATION
            AND TSS_5MIN_LANE.DATE_REF = MAX_FLOW.DATE_REF
            AND MAX_FLOW.MAX_VOL IS NOT NULL
            AND MAX_FLOW.MAX_VOL >0
```

2.4.7 EFF_DET_LEN_5MIN, EFF_DET_LEN_15MIN, EFF_DET_LEN_1HR and EFF_DET_LEN_DAY

These tables provide effective detector length for web user queries. Table E-35 shows STEWARD materialized views definition for EFF_DET_LEN_5MIN, EFF_DET_LEN_15MIN, EFF_DET_LEN_1HR and EFF_DET_LEN_DAY.

Table E- 35: STEWARD materialized views – Definitions of EFF_DET_LEN_5MIN, EFF_DET_LEN_15MIN, EFF_DET_LEN_1HR and EFF_DET_LEN_DAY

Name	Data Type (Length/Precision)	
DATE1	DATE	
TIME	VARCHAR2(25)	N/A for VOLUME_MAP_DAY
FACILITY	NUMBER	
STATION_ID	NUMBER	
STATION_DESC	VARCHAR2(200)	
STATION_MP	FLOAT(5)	
LANE1_VOL	NUMBER	
LANE1_SPD	FLOAT(5)	
LANE1_OCC	FLOAT(5)	
LANE1_EFF_DET_LENGTH	NUMBER	
LANE2_VOL	NUMBER	
LANE2_SPD	FLOAT(5)	
LANE2_OCC	FLOAT(5)	
LANE2_EFF_DET_LENGTH	NUMBER	
LANE3_VOL	NUMBER	
LANE3_SPD	FLOAT(5)	
LANE3_OCC	FLOAT(5)	
LANE3_EFF_DET_LENGTH	NUMBER	
LANE4_VOL	NUMBER	
LANE4_SPD	FLOAT(5)	
LANE4_OCC	FLOAT(5)	
LANE4_EFF_DET_LENGTH	NUMBER	
LANE5_VOL	NUMBER	
LANE5_SPD	FLOAT(5)	
LANE5_OCC	FLOAT(5)	
LANE5_EFF_DET_LENGTH	NUMBER	
LANE6_VOL	NUMBER	
LANE6_SPD	FLOAT(5)	
LANE6_OCC	FLOAT(5)	
LANE6_EFF_DET_LENGTH	NUMBER	

Table E-36 shows the query to generate the materialized view- EFF_DET_LEN_5MIN.

Table E- 36: Query for STEWARD materialized views-EFF_DET_LEN_5MIN

```
SELECT
      DATE_TBL.DAY,
      TIME_TBL.TIME,
      MOD(STATION.STATION_ID,10),
      STATION.STATION_ID,
      STATION.DESCRIPTION,
      STATION.STATE_MILEPOST,

      TSS_5MIN_LANE.LANE1_VOL,
      TSS_5MIN_LANE.LANE1_SPD,
      TSS_5MIN_LANE.LANE1_OCC,
      /* vehicle length = velocity*5280*occupancy/100/flow */
      CASE
            WHEN TSS_5MIN_LANE.LANE1_VOL > 0 THEN (TSS_5MIN_LANE.LANE1_SPD * 5280) *
(TSS_5MIN_LANE.LANE1_OCC /100) / (12*TSS_5MIN_LANE.LANE1_VOL)
            WHEN TSS_5MIN_LANE.LANE1_VOL = 0 THEN 0
            ELSE NULL
      END,

      TSS_5MIN_LANE.LANE2_VOL,
      TSS_5MIN_LANE.LANE2_SPD,
```

```
      TSS_5MIN_LANE.LANE2_OCC,
      /* vehicle length = velocity*5280*occupancy/100/flow */
      CASE
            WHEN TSS_5MIN_LANE.LANE2_VOL > 0 THEN (TSS_5MIN_LANE.LANE2_SPD * 5280) *
(TSS_5MIN_LANE.LANE2_OCC /100) / (12*TSS_5MIN_LANE.LANE2_VOL)
            WHEN TSS_5MIN_LANE.LANE1_VOL = 0 THEN 0
            ELSE NULL
      END,

      TSS_5MIN_LANE.LANE3_VOL,
      TSS_5MIN_LANE.LANE3_SPD,
      TSS_5MIN_LANE.LANE3_OCC,
      /* vehicle length = velocity*5280*occupancy/100/flow */
      CASE
            WHEN TSS_5MIN_LANE.LANE3_VOL > 0 THEN (TSS_5MIN_LANE.LANE3_SPD * 5280) *
(TSS_5MIN_LANE.LANE3_OCC /100) / (12*TSS_5MIN_LANE.LANE3_VOL)
            WHEN TSS_5MIN_LANE.LANE1_VOL = 0 THEN 0
            ELSE NULL
      END,

      TSS_5MIN_LANE.LANE4_VOL,
      TSS_5MIN_LANE.LANE4_SPD,
      TSS_5MIN_LANE.LANE4_OCC,
      /* vehicle length = velocity*5280*occupancy/100/flow */
      CASE
            WHEN TSS_5MIN_LANE.LANE4_VOL > 0 THEN (TSS_5MIN_LANE.LANE4_SPD * 5280) *
(TSS_5MIN_LANE.LANE4_OCC /100) / (12*TSS_5MIN_LANE.LANE4_VOL)
            WHEN TSS_5MIN_LANE.LANE1_VOL = 0 THEN 0
            ELSE NULL
      END,

      TSS_5MIN_LANE.LANE5_VOL,
      TSS_5MIN_LANE.LANE5_SPD,
      TSS_5MIN_LANE.LANE5_OCC,
      /* vehicle length = velocity*5280*occupancy/100/flow */
      CASE
            WHEN TSS_5MIN_LANE.LANE5_VOL > 0 THEN (TSS_5MIN_LANE.LANE5_SPD * 5280) *
(TSS_5MIN_LANE.LANE5_OCC /100) / (12*TSS_5MIN_LANE.LANE5_VOL)
            WHEN TSS_5MIN_LANE.LANE1_VOL = 0 THEN 0
            ELSE NULL
      END,

      TSS_5MIN_LANE.LANE6_VOL,
      TSS_5MIN_LANE.LANE6_SPD,
      TSS_5MIN_LANE.LANE6_OCC,
      /* vehicle length = velocity*5280*occupancy/100/flow */
      CASE
            WHEN TSS_5MIN_LANE.LANE6_VOL > 0 THEN (TSS_5MIN_LANE.LANE6_SPD * 5280) *
(TSS_5MIN_LANE.LANE6_OCC /100) / (12*TSS_5MIN_LANE.LANE6_VOL)
            WHEN TSS_5MIN_LANE.LANE1_VOL = 0 THEN 0
            ELSE NULL
      END

FROM GATOR.TSS_STATION STATION, GATOR.TSS_5MIN_LANE TSS_5MIN_LANE, GATOR.DATE_REF
DATE_TBL, GATOR.TIME_REF TIME_TBL

WHERE TSS_5MIN_LANE.TSS_STATION = STATION.DIMENSION_KEY
            AND TSS_5MIN_LANE.DATE_REF = DATE_TBL.DIMENSION_KEY
            AND TSS_5MIN_LANE.TIME_REF = TIME_TBL.DIMENSION_KEY
```

2.4.8 TVT_MAX_DELAY_5MIN, TVT_MAX_ DELAY_15MIN, and TVT_MAX_ DELAY_1HR

These tables provide TVT max delay for web user queries. Table E-37 shows STEWARD materialized views definition for TVT_MAX_DELAY_5MIN, TVT_MAX_ DELAY_15MIN, and TVT_MAX_ DELAY_1HR.

Table E- 37STEWARD materialized views – Definitions of TVT_MAX_DELAY_5MIN, TVT_MAX_ DELAY_15MIN, and TVT_MAX_ DELAY_1HR

Name	Data Type (Length/Precision)
DATE1	DATE
ORIGIN	VARCHAR2(40)
DESTINATION	VARCHAR2(40)
FF_TT	FLOAT(5)
AVG_TT	NUMBER
AVG_DELAY	NUMBER
MAX_DELAY	FLOAT(5)
MAX_DELAY_TIME	VARCHAR2(25)

Table E-38 shows the query to generate the materialized view- TVT_MAX_DELAY_5MIN.

Table E- 38: Query for STEWARD materialized views-TVT_MAX_DELAY_5MIN

```
SELECT
      DATE_REF.DAY,
      TVT_ORIGIN.ARCHIVE_REF AS ORIGIN,
      TVT_DESTINATION.ARCHIVE_REF AS DESTINATION,

      TVT_AVG_FF_TT.FF_TIME,
      TVT_AVG_TT.AVG_TT,

      TVT_AVG_TT.AVG_DELAY,

      TVT_MAX_DELAY.DELAY AS MAX_DELAY,
      TVT_MAX_DELAY.TIME AS MAX_DELAY_TIME

FROM TVT_ORIGIN, TVT_DESTINATION, TVT_LINK, DATE_REF, TVT_5MIN,

     (SELECT TVT_LINK.FF_TIME AS FF_TIME, TVT_ORIGIN.DIMENSION_KEY AS ORIGIN,
             TVT_DESTINATION.DIMENSION_KEY AS DESTINATION
      FROM TVT_LINK, TVT_ORIGIN, TVT_DESTINATION
      WHERE TVT_LINK.ORIGIN = TVT_ORIGIN.ORIGIN_ID
        AND TVT_LINK.DESTINATION = TVT_DESTINATION.DESTINATION_ID
      GROUP BY TVT_LINK.FF_TIME, TVT_ORIGIN.DIMENSION_KEY, TVT_DESTINATION.DIMENSION_KEY)
TVT_AVG_FF_TT,

     (SELECT AVG(TVT_5MIN.AVG_TT) AS AVG_TT,
             AVG(TVT_5MIN.DELAY) AS AVG_DELAY,
             TVT_5MIN.TVT_ORIGIN AS ORIGIN,
             TVT_5MIN.TVT_DESTINATION AS DESTINATION,
             TVT_5MIN.DATE_REF AS DATE_REF
      FROM TVT_ORIGIN, TVT_DESTINATION, DATE_REF, TVT_5MIN
      WHERE TVT_5MIN.TVT_ORIGIN = TVT_ORIGIN.DIMENSION_KEY
        AND TVT_5MIN.TVT_DESTINATION = TVT_DESTINATION.DIMENSION_KEY
        AND TVT_5MIN.DATE_REF = DATE_REF.DIMENSION_KEY
      GROUP BY TVT_5MIN.TVT_ORIGIN, TVT_5MIN.TVT_DESTINATION, TVT_5MIN.DATE_REF)
TVT_AVG_TT,

     (SELECT TVT_5MIN.DELAY AS DELAY, TIME_REF.TIME AS TIME, TVT_5MIN.TVT_ORIGIN,
             TVT_5MIN.TVT_DESTINATION, TVT_5MIN.DATE_REF
      FROM TVT_5MIN, TIME_REF,
          (SELECT MAX(TVT_5MIN.DELAY) AS MAX_DELAY,
```

```
                   TVT_ORIGIN, TVT_DESTINATION, DATE_REF
           FROM TVT_5MIN
           GROUP BY TVT_ORIGIN, TVT_DESTINATION, DATE_REF) TVT_MAX
      WHERE TVT_5MIN.DELAY=TVT_MAX.MAX_DELAY
        AND TVT_5MIN.TVT_ORIGIN=TVT_MAX.TVT_ORIGIN
        AND TVT_5MIN.TVT_DESTINATION =TVT_MAX.TVT_DESTINATION
        AND TVT_5MIN.TIME_REF=TIME_REF.DIMENSION_KEY
        AND TVT_5MIN.DATE_REF=TVT_MAX.DATE_REF
        AND TVT_5MIN.DELAY>0
      GROUP BY TVT_5MIN.DELAY, TIME_REF.TIME, TVT_5MIN.TVT_ORIGIN,
TVT_5MIN.TVT_DESTINATION, TVT_5MIN.DATE_REF) TVT_MAX_DELAY

WHERE TVT_ORIGIN.DIMENSION_KEY = TVT_5MIN.TVT_ORIGIN
      AND TVT_DESTINATION.DIMENSION_KEY = TVT_5MIN.TVT_DESTINATION
      AND TVT_5MIN.DATE_REF = DATE_REF.DIMENSION_KEY
      AND TVT_5MIN.TVT_ORIGIN = TVT_AVG_FF_TT.ORIGIN
      AND TVT_5MIN.TVT_DESTINATION = TVT_AVG_FF_TT.DESTINATION
      AND TVT_5MIN.TVT_ORIGIN = TVT_AVG_TT.ORIGIN
      AND TVT_5MIN.TVT_DESTINATION = TVT_AVG_TT.DESTINATION
      AND TVT_5MIN.DATE_REF = TVT_AVG_TT.DATE_REF
      AND TVT_5MIN.TVT_ORIGIN = TVT_MAX_DELAY.TVT_ORIGIN
      AND TVT_5MIN.TVT_DESTINATION = TVT_MAX_DELAY.TVT_DESTINATION
      AND TVT_5MIN.DATE_REF = TVT_MAX_DELAY.DATE_REF

GROUP BY DATE_REF.DAY, TVT_ORIGIN.ARCHIVE_REF, TVT_DESTINATION.ARCHIVE_REF,
TVT_AVG_FF_TT.FF_TIME,
   TVT_AVG_TT.AVG_TT, TVT_AVG_TT.AVG_DELAY, TVT_MAX_DELAY.DELAY, TVT_MAX_DELAY.TIME
```

2.4.9 TVT_MAX_TT_5MIN, TVT_MAX_TT_15MIN, and TVT_MAX_TT_1HR

These tables provide TVT max travel time for web user queries. Table E-39 shows STEWARD materialized views definition for TVT_MAX_TT_5MIN, TVT_MAX_TT_15MIN, and TVT_MAX_TT_1HR.

Table E- 39: STEWARD materialized views – Definitions of TVT_MAX_TT_5MIN, TVT_MAX_TT_15MIN, and TVT_MAX_TT_1HR

Name	Data Type (Length/Precision)
DATE1	DATE
ORIGIN	VARCHAR2(40)
DESTINATION	VARCHAR2(40)
FF_TT	FLOAT(5)
AVG_TT	NUMBER
MAX_TT	FLOAT(5)
MAX_TT_TIME	VARCHAR2(25)
AVG_DELAY	NUMBER

Table E-40 shows the query to generate the materialized view- TVT_MAX_TT_5MIN.

Table E- 40: Query for STEWARD materialized views-TVT_MAX_TT _5MIN

```
SELECT
      DATE_REF.DAY,
      TVT_ORIGIN.ARCHIVE_REF AS ORIGIN,
      TVT_DESTINATION.ARCHIVE_REF AS DESTINATION,

      TVT_AVG_FF_TT.FF_TIME,
      TVT_AVG_TT.AVG_TT,

      TVT_MAX_TT.TT AS MAX_TT,
      TVT_MAX_TT.TIME AS MAX_TT_TIME,

      TVT_AVG_TT.AVG_DELAY

FROM TVT_ORIGIN, TVT_DESTINATION, TVT_LINK, DATE_REF, TVT_5MIN,

     (SELECT TVT_LINK.FF_TIME AS FF_TIME, TVT_ORIGIN.DIMENSION_KEY AS ORIGIN,
             TVT_DESTINATION.DIMENSION_KEY AS DESTINATION
      FROM TVT_LINK, TVT_ORIGIN, TVT_DESTINATION
      WHERE TVT_LINK.ORIGIN = TVT_ORIGIN.ORIGIN_ID
        AND TVT_LINK.DESTINATION = TVT_DESTINATION.DESTINATION_ID
      GROUP BY TVT_LINK.FF_TIME, TVT_ORIGIN.DIMENSION_KEY, TVT_DESTINATION.DIMENSION_KEY)
TVT_AVG_FF_TT,

     (SELECT AVG(TVT_5MIN.AVG_TT) AS AVG_TT,
             AVG(TVT_5MIN.DELAY) AS AVG_DELAY,
             TVT_5MIN.TVT_ORIGIN AS ORIGIN,
             TVT_5MIN.TVT_DESTINATION AS DESTINATION,
             TVT_5MIN.DATE_REF AS DATE_REF
      FROM TVT_ORIGIN, TVT_DESTINATION, DATE_REF, TVT_5MIN
      WHERE TVT_5MIN.TVT_ORIGIN = TVT_ORIGIN.DIMENSION_KEY
        AND TVT_5MIN.TVT_DESTINATION = TVT_DESTINATION.DIMENSION_KEY
        AND TVT_5MIN.DATE_REF = DATE_REF.DIMENSION_KEY
      GROUP BY TVT_5MIN.TVT_ORIGIN, TVT_5MIN.TVT_DESTINATION, TVT_5MIN.DATE_REF)
TVT_AVG_TT,

     (SELECT TVT_5MIN.MAX_TT AS TT, TIME_REF.TIME AS TIME, TVT_5MIN.TVT_ORIGIN,
             TVT_5MIN.TVT_DESTINATION, TVT_5MIN.DATE_REF
      FROM TVT_5MIN, TIME_REF,
          (SELECT MAX(TVT_5MIN.MAX_TT) AS MAX_TT,
                  TVT_ORIGIN, TVT_DESTINATION, DATE_REF
           FROM TVT_5MIN
           GROUP BY TVT_ORIGIN, TVT_DESTINATION, DATE_REF) TVT_MAX
      WHERE TVT_5MIN.MAX_TT=TVT_MAX.MAX_TT
        AND TVT_5MIN.TVT_ORIGIN=TVT_MAX.TVT_ORIGIN
        AND TVT_5MIN.TVT_DESTINATION =TVT_MAX.TVT_DESTINATION
        AND TVT_5MIN.TIME_REF=TIME_REF.DIMENSION_KEY
        AND TVT_5MIN.DATE_REF=TVT_MAX.DATE_REF
      GROUP BY TVT_5MIN.MAX_TT, TIME_REF.TIME, TVT_5MIN.TVT_ORIGIN,
TVT_5MIN.TVT_DESTINATION, TVT_5MIN.DATE_REF) TVT_MAX_TT

WHERE TVT_ORIGIN.DIMENSION_KEY = TVT_5MIN.TVT_ORIGIN
      AND TVT_DESTINATION.DIMENSION_KEY = TVT_5MIN.TVT_DESTINATION
      AND TVT_5MIN.DATE_REF = DATE_REF.DIMENSION_KEY
      AND TVT_5MIN.TVT_ORIGIN = TVT_AVG_FF_TT.ORIGIN
      AND TVT_5MIN.TVT_DESTINATION = TVT_AVG_FF_TT.DESTINATION
      AND TVT_5MIN.TVT_ORIGIN = TVT_AVG_TT.ORIGIN
      AND TVT_5MIN.TVT_DESTINATION = TVT_AVG_TT.DESTINATION
      AND TVT_5MIN.DATE_REF = TVT_AVG_TT.DATE_REF
      AND TVT_5MIN.TVT_ORIGIN = TVT_MAX_TT.TVT_ORIGIN
      AND TVT_5MIN.TVT_DESTINATION = TVT_MAX_TT.TVT_DESTINATION
      AND TVT_5MIN.DATE_REF = TVT_MAX_TT.DATE_REF

GROUP BY DATE_REF.DAY, TVT_ORIGIN.ARCHIVE_REF, TVT_DESTINATION.ARCHIVE_REF,
TVT_AVG_FF_TT.FF_TIME,
   TVT_AVG_TT.AVG_TT, TVT_MAX_TT.TT, TVT_MAX_TT.TIME, TVT_AVG_TT.AVG_DELAY
```

3 The STEWARD ETL Process

There are three modules in the Oracle Warehouse Builder related to the STEWARD ETL process: mappings, process flows and schedules.

3.1 Mappings

Mappings are defined in the Oracle Warehouse Builder as a series of operations that extract data from sources, transform it, and load it into targets. In STEWARD, these mappings are confined to the data loading from the external files to STEWARD database. Fifteen mappings are defined in the STEWARD system. All the mappings except the DATE_REF_MAP have the same architecture: Mapping from the external files to dimension and fact tables. In case of DATE_REF_MAP, the mapping was generated automatically using the time wizard for dimensional tables.

Mappings for the dimension tables:
DATE_REF_MAP
TIME_REF_MAP
LOAD_TSS_STATION
LOAD_TVT_LINK
LOAD_TVT_ORIGIN
LOAD_TVT_DESTINATION

Mappings for the fact tables:
LOADING_TSS_5MIN
LOADING_TSS_15MIN
LOADING_TSS_1HR
LOADING_TSS_5MIN_LANE
LOADING_TSS_15MIN_LANE
LOADING_TSS_1HR_LANE
LOADING_TVT_5MIN
LOADING_TVT_15MIN
LOADING_TVT_1HR

Figure E-1 shows a simple example of STEWARD mapping. The external file (tss-15min_data.csv) is shown as the Oracle external table and treated as the table (TSS_15MIN_EX). The first column in the external table (DATE1) is transformed using expression utility. In this case, the column type is changed from VARCHAR2 to DATE with the format changes ('MM/DD/YYYY' → 'YYYYMMDD'). All the columns in the external file are mapped to fact table with appropriate data conversion.

The following PL/SQL code resides within the EXPRESSION_0 utility.

```
>> TO_NUMBER(TO_CHAR(TO_DATE(INGRP1.DATE1,'MM/DD/YYYY'),'YYYYMMDD'),'99999999')
```

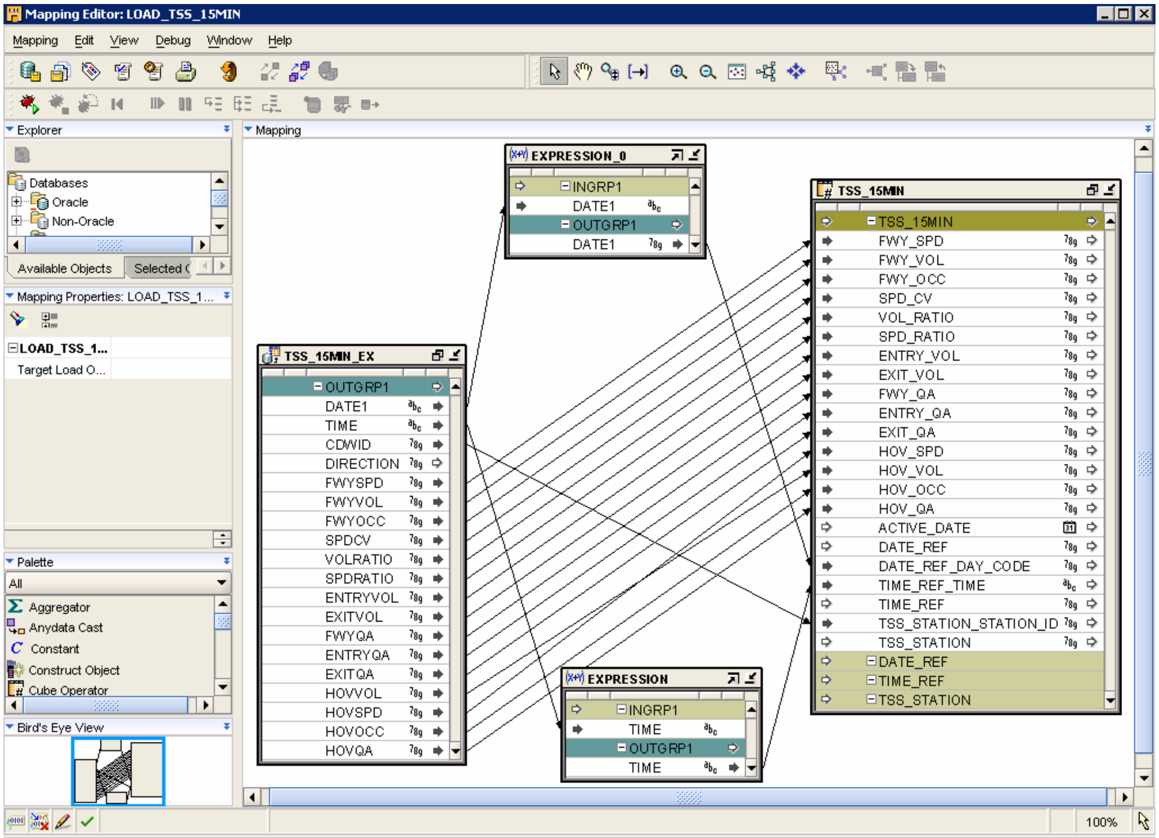


Figure E- 1: STEWARD mapping example – LOAD_TSS_15MIN

3.2 Process flows

Process flows are used to execute the STEWARD ETL process. Process flows can execute the external program and can be executable manually or in a batch job by scheduler. In STEWARD, the transform process is developed using an external program (SunETLUtility), which is executed from this module.

There are six main process flows and nine sub-procedures defined in STEWARD. The six main process flows can be categorized into Stage1 through Stage 4. The details of each stage are as follows:

3.2.1 Stage 1: Pre-transformation

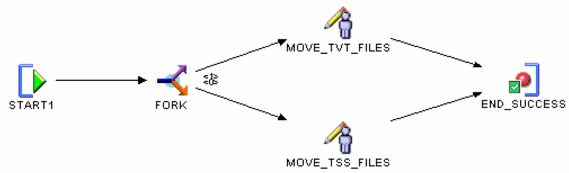


Figure E- 2: STEWARD process flow – Stage1 Pre-transformation

As shown in Figure E-2, the input data files from SunETLUtility are moved from the predefined location (K:\District_Data\District2\TSS_DATA or K:\District_Data\District2\TVT_DATA) to the SunETLUtility input data folder (C:\Program Files\TRC\SunETLUtility\ToConvert). One batch program (Move2SunETLUtilityFolder.bat) is executed for this stage.

3.2.2 Stage 2: Transformation

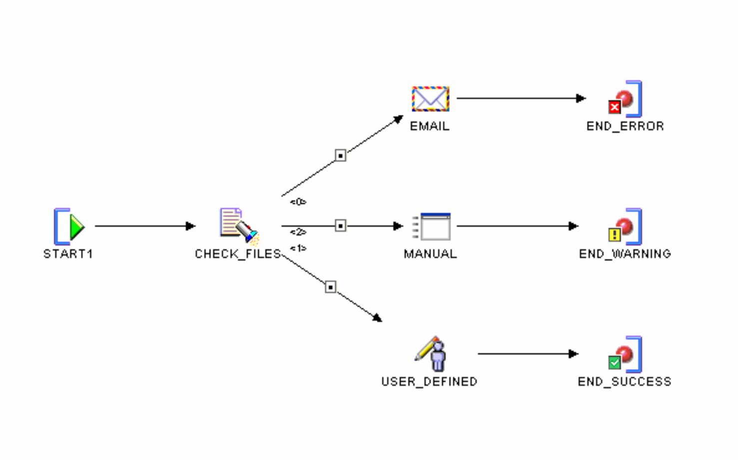


Figure E- 3: STEWARD process flow – Stage2 Transformation

As shown in Figure E-3, the SunETLUtility is executed to transform the input data files to the STEWARD input data format. The execution time of this process depends on number of input files, its contents and host computer. The execution time estimation might be needed for the scheduler operations. One program (SunETLUtility) is executed for this stage.

3.2.3 Stage 3: Pre-loading

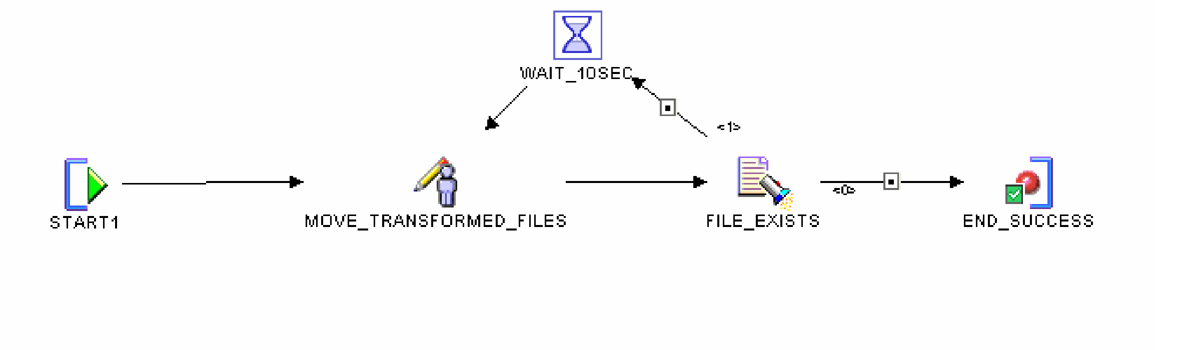


Figure E- 4: STEWARD process flow – Stage3 Pre-loading

As shown in Figure E-4, the outputs of SunETLUtility are determined from its file name and moved to the STEWARD loading folders. The default output folder is c:\progra~1\TRC\SunETLUtility and the target data folders for data loading are as follows:

C:\Steward\TSS_STATION_DATA\TSS_5MIN_Data
C:\Steward\TSS_STATION_DATA\TSS_15MIN_Data
C:\Steward\TSS_STATION_DATA\TSS_1Hr_Data
C:\Steward\TSS_LANE_DATA\TSS_5MIN_Data
C:\Steward\TSS_LANE_DATA\TSS_15MIN_Data
C:\Steward\TSS_LANE_DATA\TSS_1Hr_Data
C:\Steward\TVT_DATA\TVT_5MIN_Data
C:\Steward\TVT_DATA\TVT_15MIN_Data
C:\Steward\TVT_DATA\TVT_1Hr_Data

One program (Move2StewardFolder.exe) is executed for this stage

3.2.4 Stage 4: Loading

As shown in Figure E-5, there are three main process flows at this stage. These are for TSS station data, TSS lane data and TVT data. Each main process includes three sub-process flows for 5min, 15min and 1hr data loading. The name of each process is as follows:

- STAGE4_ LOADG_TSS
- STAGE4_ LOADG_TSS_LANE
- STAGE4_ LOADG_TVT

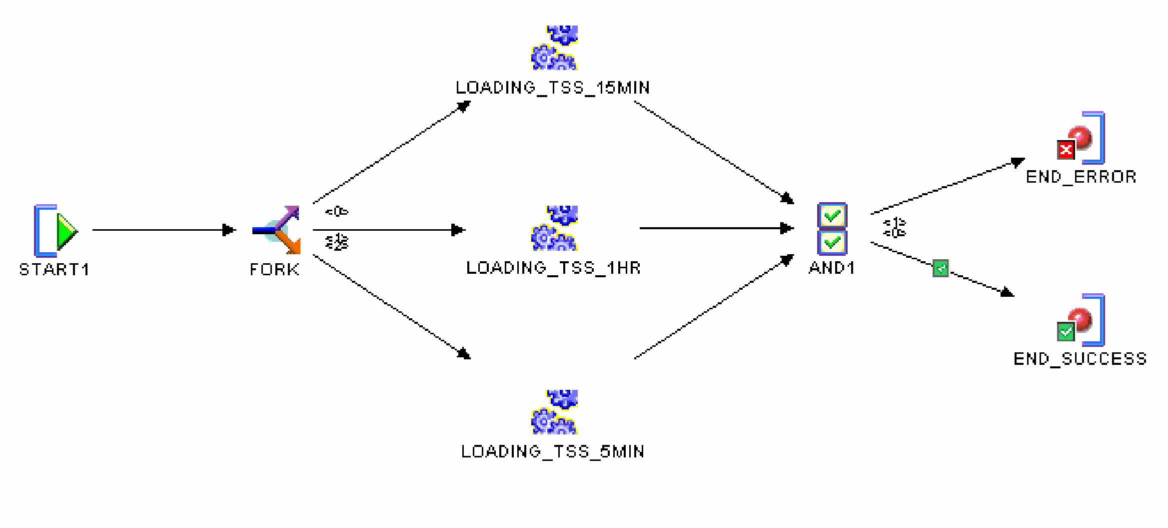


Figure E- 5: STEWARD process flow – Stage4 Loading

Nine sub process flows are defined as follows:

-LOADING_TSS_5MIN
-LOADING_TSS_15MIN
-LOADING_TSS_1HR
-LOADING_TSS_5MIN_LANE
-LOADING_TSS_15MIN_LANE
-LOADING_TSS_1HR_LANE
-LOADING_TVT_5MIN
-LOADING_TVT_15MIN
-LOADING_TVT_1HR

As shown in Figure E-6, each process load one type of files to Steward system.

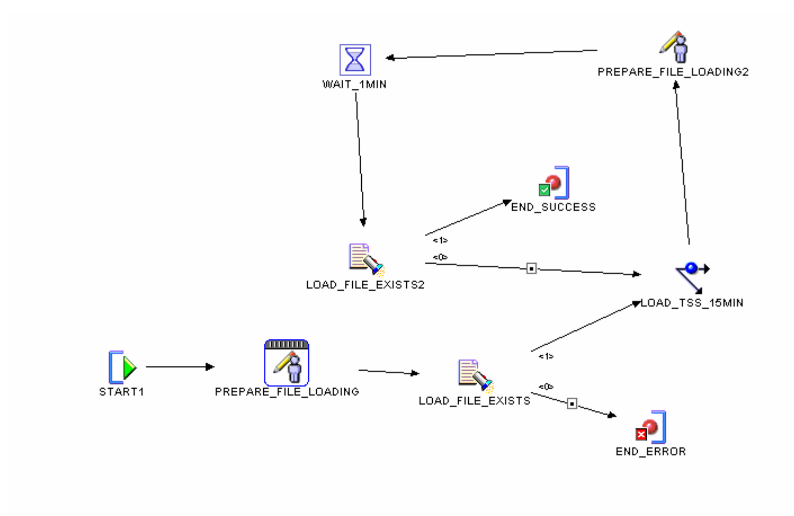


Figure E- 6: STEWARD process flow – Substage Loading_TSS_15MIN

One program (PrepareFileLoading.exe) is executed for this stage

3.3 Schedules

The STEWARD operator can use the scheduler to run the STEWARD process flows at a predetermined time. The details of how to use the scheduler are described in Appendix C to the STEWARD Final Report.

4 STEWARD Web Interface

The STEWARD web program was developed to support the general user interface. It communicates with the STEWARD database to retrieve the traffic data and generates the reports. It is developed on Microsoft Internet Information Services (IIS) 5.1 using Active Server Pages (ASP) and JavaScript.

4.1 Program Architecture

Web user can access the STEWARD web program as the tree architecture. The site map is shown in Figure E-7.

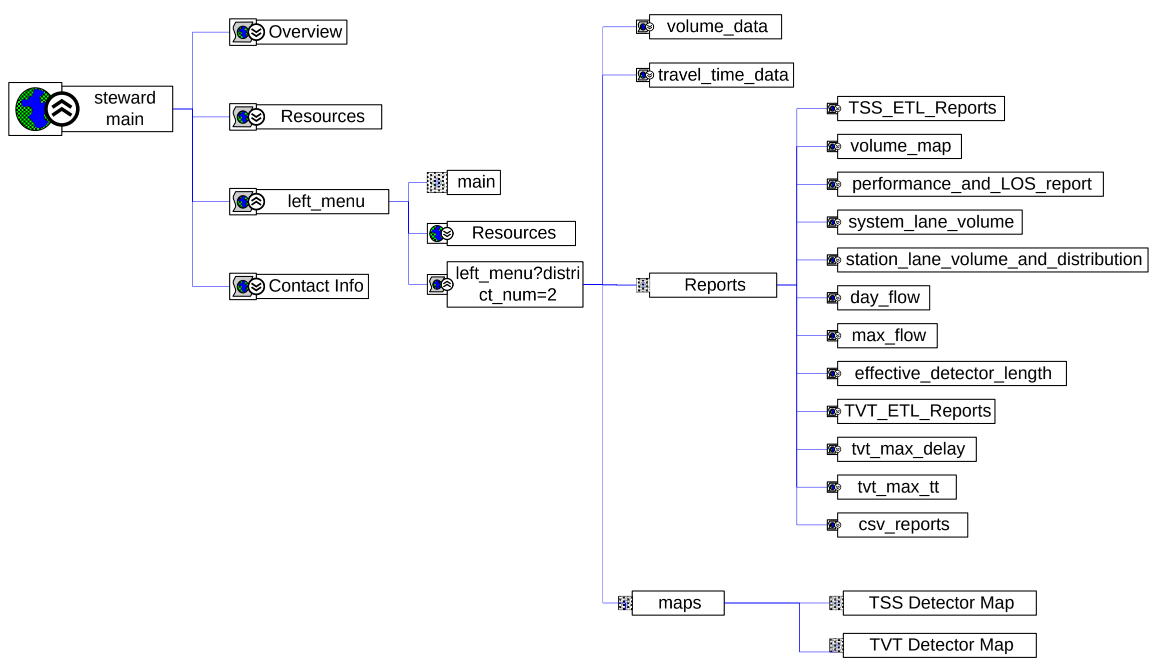


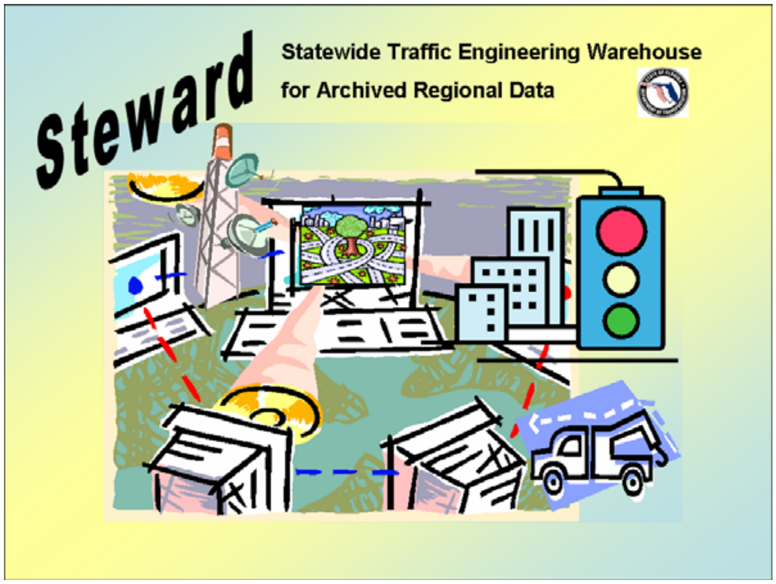
Figure E- 7: STEWARD web architecture

4.2 STEWARD Web Implementation

The STEWARD web program has one main folder and six sub folders. The main folder includes the program main page and resource pages. The Index_main folder includes subfolders for District data and reports. The WebFiles and Files folders include SunETLUtility output and STEWARD web report outputs. Other folders (images, common and css) folders include utility files used in the STEWARD web program. The details are described in the following Table E-41.

Table E- 41: Directory architecture of STEWARD web program

Folder name	Descriptions
wwwroot\steward	Includes STEWARD main page and three sub pages (STEWARD Overview, Resources and Contact Info)
~\WebFiles	Includes the linked contents for the Resources page
~\files	Includes CSV files generated from the district data and reports
~\index_main	Includes subfolders for each reports
~\files\volume_data	Includes asp files for volume data
~\files\travel_time_data	Includes asp files for travel time data
~\files\TSS_ETL_Reports	Includes asp files for TSS ETL Reports
~\files\Volume_Map	Includes asp files for Volume Map and I/O Balance
~\files\performance_and_LOS_report	Includes asp files for Performance Measure and LOS
~\files\system_lane_volume	Includes asp files for System Lane Volume
~\files\station_lane_volume_and_distribution	Includes asp files for Station Lane Volume Map and Distribution
~\files\day_flow	Includes asp files for Day Flow
~\files\max_flow	Includes asp files for Max Flow
~\files\effective_detector_length	Includes asp files for Effective Detector Length
~\files\TVT_ETL_Reports	Includes asp files for TVT ETL Reports
~\files\tvt_max_delay	Includes asp files for TVT Max Delay
~\files\tvt_max_tt	Includes asp files for TVT Max Travel Time
~\files\csv_reports	Includes asp files for CVS Reports
~\files\TSS_REPORTS ~\files\TVT_REPORTS	Includes log files from the SunETLUtility for the TSS/TVT ETL Reports.
~\css ~\common	Utility files for STEWARD web programs



Development of a Central Data Warehouse for Statewide ITS and Transportation Data in Florida

Phase II: Proof of Concept Final Report

Appendix F
Software Utilities Documentation

Prepared for the Florida Department of Transportation
By the University of Florida Transportation Research Center

Contract # BD545, RPWO # 37
UF Project 051449

Revision Date: April 15, 2008

Table of Contents

List of Figures

List of Tables

1 Introduction

Several software products were produced to support the development and operation of STEWARD. All products developed in connection with this project are delivered as a part of the final report. Their documentation is incorporated into this appendix. The program and data files are included on the supplemental CD of Appendix G: The items that fall into this category include:

- ETL Utility Program: The ETL utility performs the first steps of the extraction, transformation and loading (ETL) process by which daily archive data from SunGuide traffic management centers (TMCs) is acquired and transformed into the STEWARD database format. It was developed for use at the STEWARD location by personnel responsible for loading archive data into the STEWARD database. This documentation is provided in Section 2 of this Appendix.

- TSSBuilder: A utility that reads the raw data from a SunGuide archive over several days and constructs a list of all of the stations and lanes that have reported volume, speed and occupancy data. This program, which is described in Section 3 of this Appendix, performs the first step in configuring the STEWARD facility data for a SunGuide TMC. The documentation for this program has been incorporated into the workshop material.

- Internal Utility Programs: The Oracle Warehouse Builder has several limitations with respect to the ETL process. To work around these limitations, several utility programs have been developed to give STEWARD the required flexibility and expandability. These utilities are described in Section 4 of this Appendix.

- SimTMC: A utility that uses a microscopic traffic simulation model to emulate a freeway traffic management center. This program, which is described in Section 5 of this Appendix, creates input data files for the FRESIM freeway simulation model and creates data archive files in the prescribed SunGuide format. It was originally developed for testing the STEWARD database before the actual data became available. It could be useful in the future for general simulation of a SunGuide TMC.

- MPConverter: A set of milepost converter routines required to convert milepost information to coordinates and *vice versa*. This product, which is described in Section 6 of this Appendix, will be useful for configuring the facility metadata for additional TMCs to be brought into STEWARD.

These programs were all developed for project purposes. Their documentation is adequate for these purposes; however they are not supportable at a level that would make them suitable for widespread distribution.

2 ETL Utility Program

2.1 Purpose

The ETL utility performs the first steps of the ETL process by which daily archive data from SunGuide TMCs is acquired and transformed into the STEWARD database format. It was developed for use at the STEWARD location by personnel responsible for loading archive data into the STEWARD database. This documentation is provided primarily for those personnel.

2.2 Concept of Operations

An overview of the ETL Utility data flow is presented in Figure F-1. There are two inputs and four outputs associated with this program. The inputs are described as follows:

2.2.1 Raw SunGuide Archive Data

The raw SunGuide archive data obtained daily from the SunGuide TMC. The method of acquisition may vary among districts. The important point is that the raw archive data must exist in a specified folder (see the section on file locations). The SunGuide file naming convention for archive data files is:

- TSS-mmddyyyy- - 1.dat for TSS archive files
- TVT-mmddyyyy-1.dat for TVT archive files

Each TSS archive file contains multiple records with the following fields:

- timestamp
- detector_id
- lane_id
- speed
- volume
- occupancy

Each TSS record covers 20 seconds of operation for one lane

Each TSS archive file contains multiple records with the following fields:

- timestamp
- travel_link_id
- travel_time

Each TVT record covers 60 seconds of operation for one travel time link

2.2.2 Facility Description Spreadsheets

The configuration of the facility data is described in Appendix A to the STEWARD Final Report. The purpose of the facility description data is to relate the identifiers in the raw archive data to the geometrics and geography of the facility monitored by the TMC The complete facility description is described in an Excel workbook containing three spreadsheets:

- General facility data
- Detector station data
- Detector lane data

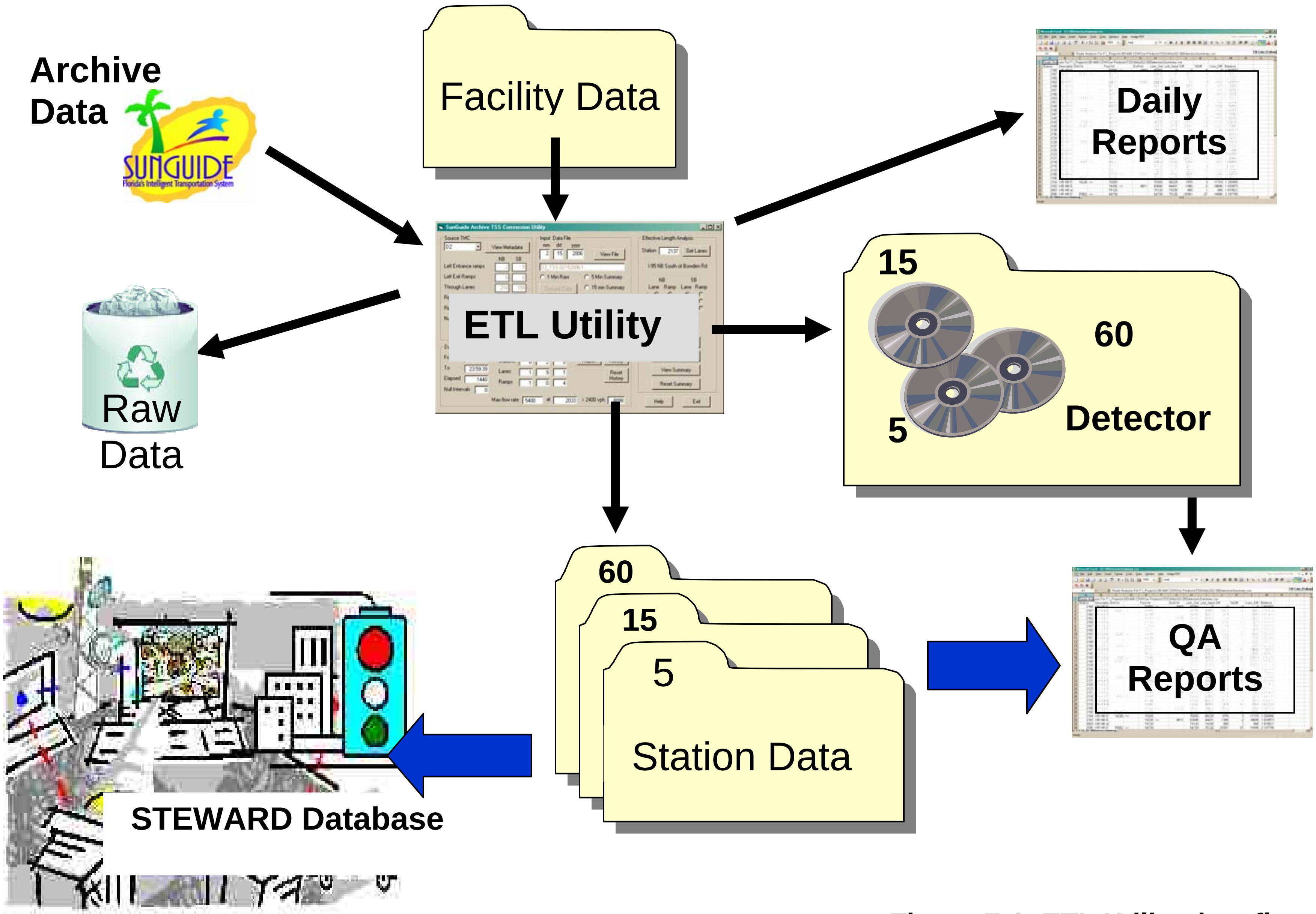


Figure F-1: ETL Utility data flow

The outputs of the ETL Utility include:

2.2.3 Daily Reports

Diagnostic reports are generated to indicate the health of the detector and communications system. The diagnostic information produced by the conversion process is presented in some detail in the daily report. This report identifies the actual stations, ramps and lanes represented in the diagnostic summary. The daily report will be useful in assessing the day- to day detector maintenance and service requirements. A sample of the Daily Report is shown in Figure F-2.

```
00:00:39   RTMS 95N006: Ramp R95N006_01Ramp_01 is not in the TSS Map.
00:00:39   RTMS 95N024: Lane R95N024_05LaneS-02 is not in the TSS Map.
Offline Station #  1 : RTMS 95N001, I-95 NB at Forest St
Offline Lane: R95N001_01Lane_01 at RTMS 95N001, I-95 NB at Forest St
Offline Lane: R95N001_02Lane_02 at RTMS 95N001, I-95 NB at Forest St
Offline Lane: R95N001_03Lane_03 at RTMS 95N001, I-95 NB at Forest St
Offline Lane: R95N001_04Lane_04 at RTMS 95N001, I-95 NB at Forest St
Offline Station #  2 : RTMS 95N002, I-95 SB at Forest St
Offline Lane: R95N002_01Lane_01 at RTMS 95N002, I-95 SB at Forest St
Offline Lane: R95N002_02Lane_02 at RTMS 95N002, I-95 SB at Forest St
Offline Lane: R95N002_03Lane_03 at RTMS 95N002, I-95 SB at Forest St
Offline Lane: R95N002_04Lane_04 at RTMS 95N002, I-95 SB at Forest St
Offline Lane: R95N002_05Lane_05 at RTMS 95N002, I-95 SB at Forest St
Null Ramp: R95N014A_01Ramp_01 at #  14 , RTMS 95N014, I-95 NB Exit to MLK Blvd
Null Ramp: R95N022A_01Ramp_01 at #  22 , RTMS 95N022, I-95 NB Exit to Lem Turner Blvd
Null Lane: R95N024_05LaneS_02 at #  24 , RTMS 95N024, I-95 NB Entrance from Lem Turner Blvd
Null Ramp: R95S054A_01Ramp_01 at #  104 , RTMS 95S054, I-95 SB Entrance from Baymeadows Rd
Null Ramp: R95S054A_02Ramp_02 at #  104 , RTMS 95S054, I-95 SB Entrance from Baymeadows Rd

Conversion Summary for D2_TSS-02152006-1
From 00:00:39
To 23:59:39
Elapsed Time 1440
Null Intervals  0
Orphan  Stations  0
       Lanes     1
       Ramps     1
  Null  Stations  0
       Lanes     1
       Ramps     4
Offline Stations  2
       Lanes     9
       Ramps     0
 Max flow rate    5400
          at    5400
 No. > 2400vph    8098
```

Figure F-2. Sample report for daily TSS file conversion

2.2.4 History Log (CSV Format)

The diagnostic summary produced by the conversion process is appended to the conversion log each time the conversion is performed. By this process, the conversion log presents a history of all of the daily conversions and their diagnostic summary information.

2.2.5 Grouped Data Summaries by Lane (CSV Format)

The lane-specific records in the 1 minute converted data files are grouped into 5, 15 and 60 minute summaries with the same format. These files were produced to support detailed analysis of the data from each lane. Their main purpose is to provide an intermediate step to the aggregation of data by station. As an end product, their eventual role in STEWARD is likely to be limited to supporting research where lane-by lane data must be available. These summaries are not loaded into the STEWARD database because of their limited value to the end user. They are instead retained in CD or DVD storage to support subsequent requests.

2.2.6 Grouped Data Summaries by Station (CSV Format)

The station-based summaries are produced from the lane-based summaries for loading into the STEWARD database. These summaries are much more compact and useful because the performance measures to be reported apply to the detection stations as a whole and not to the individual lanes. The individual lane data does not appear in the STEWARD database. The only lane-related item in the data base is a measure called "lane balance," defined as the ratio of the highest to the lowest lane volume at any station.

2.3 Loading the STEWARD Database

The data from the station-based group files produced by the ETL Utility must be loaded into the STEWARD database. The method by which this step is accomplished is documented in Appendix C of the STEWARD Final Report.

2.4 File Locations

The ETL Utility program expects and/or produces files in the following folders. All of the folders listed except for the program folder must be created as subfolders of the program folder.

- Program Folder
    - Program
    - Metadata File (TSSMetadata-D2.csv)
    - Lastaction.txt: Used to save and restore the screen parameters from the last action to minimize reentry and editing.

- "FacilityData" Folder
  This folder must contain all of the facility data files as Excel workbooks. The facility data requirements are explained in detail in Appendix A of the STEWARD Final Report.

- "CountData" Folder
  This folder contains traffic count files in the FDOT Central and District office formats created by the ETL Utility. Requirements for setting up the configuration for traffic count files are given in Appendix A to the STEWARD Final Report.

- "ToConvert" Folder
  The Raw SunGuide TSS and TVT data files from the district TMCs must be placed in this folder

- "ConvertedData" Folder
  ETL Utility writes three types of files to this folder during the conversion process:
    - Converted data files
    - Daily conversion reports (TXT format)
    - Conversion history, 1 record per day (CSV)

- "GroupData" Folder
  ETLUtility writes the following files into this folder
  - Lane based grouped data files for 5, 15 and 60 minute intervals (same format as converted files except for time stamp)
  - Station-based grouped data files for 5, 15 and 60 minute intervals, in the format required by the STEWARD database

2.5 Operation

When the program is run, the main screen appears as shown in Figure F-3.

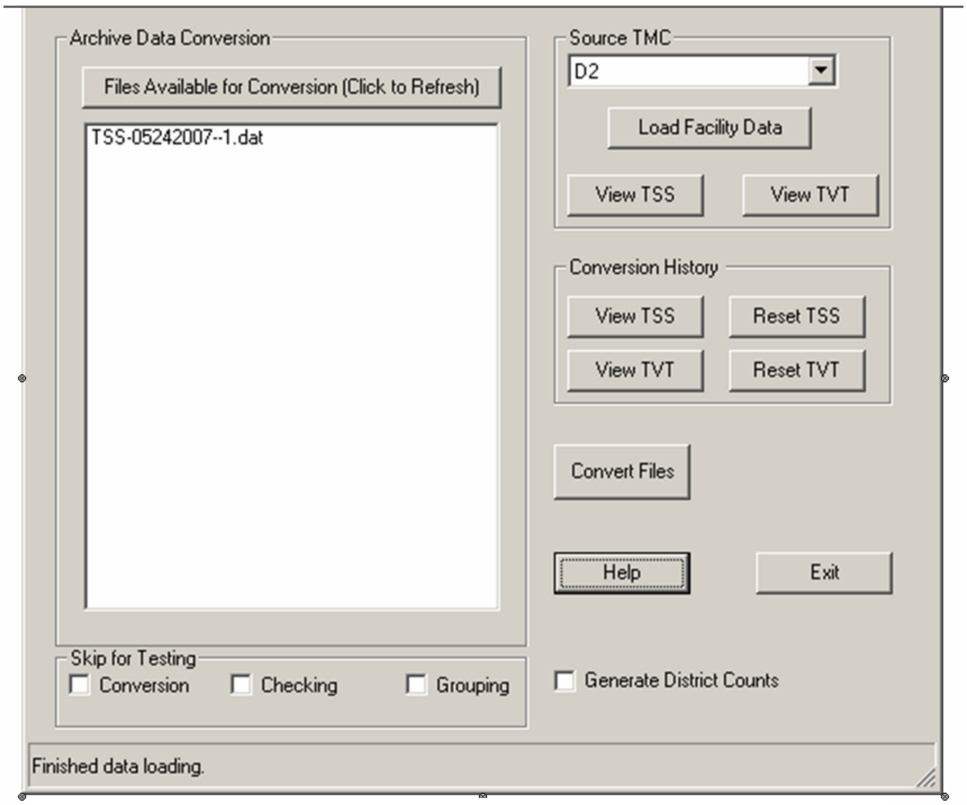


Figure F-3: ETL Utility main screen

2.5.1 "Archive Data Conversion" Frame

All of the files in the "To Convert" folder are shown in the "Archive Data Conversion" list box. Remember, it is up to the user to ensure that the files to be converted are placed in the "To Convert" folder before conversion and removed to some other location after the conversion has been completed. You may click the box above this list to refresh the list after you have made changes.

2.5.2 Source TMC Frame

The drop down list in the "Source TMC" frame will display all of the TMCs for which facility data files exist in the "FacilityData" folder. The facility data for the chosen TMC will be applied to the archive data files to carry out the transformation to the STEWARD database format. Command buttons in this frame may be used to load and view the facility configuration data.

2.5.3 Conversion History Frame

This frame contains command buttons that let you view and reset the conversion history files.

2.5.4 Converting files

Clicking the “Convert Files” button will initiate the conversion of all of the archive data in the “ToConvert” folder into the STEWARD database format. Conversion history and daily report files will be created as a part of this process.

2.6 Generating District Counts

If you select the “Generate District Counts” check box, each day’s data will generate traffic count files in the district planning office format for the count stations specified in the facility data files.

2.7 Command Line Execution for the ETLUtility

The ETL Utility supports command line execution to support the loading process in the Oracle Warehouse Builder. Command line parameters specify the district number, input directory and output directory. The input directory indicates the directory that includes the ***ToConvert*** and ***FacilityData*** folders for input data files and facility data files. The output directory indicates the directory that includes the ***ConvertedData*** and ***GroupData*** for converted and aggregated output data files.

The calling syntax is:
>> SunETLUtility /d”arg1” /i”arg2” /o”arg3”
The arguments are:

- arg1: District name
- arg2: Input directory path and name
- arg3: Output directory path and name

Example:
```
>> SunETLUtility /d”D2” /i”c:\progra~1\trc\SunETLUtility” /o”c:\progra~1\trc\SunETLUtility”
```

2.8 Additional Guidance

A couple of points worth remembering:

1. The archive file conversion handles many megabytes of data. This process is time consuming and is best carried out in an unattended mode (e.g., overnight)
2. The SunGuide system does not distinguish between districts in the naming of archive files. On a given date, each district will produce an archive file that cannot be distinguished by its name from other districts. If you are processing files from multiple districts, careful attention must be given to separating the individual files.

3 TSS Builder Utility

The TSS Builder utility creates the list of TSS stations and lanes from sample TSS archive files. It generates the station data and lane data spreadsheets required for processing the daily archive data and transforming the records to the STEWARD database format.

3.1 Data Flow Overview

An overview of the TSS Builder data flow is presented in Figure F-4. The program reads a series of SunGuide archive data files from several days and creates a list of all stations and lanes that have reported archive data. When the station and lane lists have been created, they are transformed into the corresponding station and lane data spreadsheets with blank fields for the data that must be added later. The complete process for configuring the TSS facility data for STEWARD is presented in Appendix A of the STEWARD Final Report.

The remaining steps in the process described above must be completed by adding the required data from the district inventory records. Each district now maintains its own inventory in a different format. In some cases, it will be possible to merge the data from the inventory to the station and lane data spreadsheets; however some manual data entry will likely be required.

3.2 File Locations

All data are read from and written to the “DATA” folder, which must be created as a subfolder in the program folder.

3.3 Program Operation

When the program is run, the main screen appears as shown in Figure F-5. There are three frames on this screen:

1. The District Information Frame lets you select the district and load any station and lane lists that have been created previously. Each time you process more archive files, the list is expanded cumulatively to represent all of the stations and lanes that have reported data at any time. You may also reset the detector list to begin a new list. Another command button will display the district inventory spreadsheet if one exists. If a district spreadsheet is available it must be placed in the “DATA” folder as a comma-delimited file with the name “TSSInventory-Dn.csv,” where “n” is the district number.

2. The “Build Detector Lists” frame displays the number of stations and lanes in the current lists. It lets you specify the name of an archive file to read or you may choose to have multiple archive files read from a text file list. The file names are specified simply by date (mmddyyyy). If a file list is supplied, it must be placed in the “DATA” folder with the name “FILELIST.TXT.” A command button in this frame lets you edit the file list.

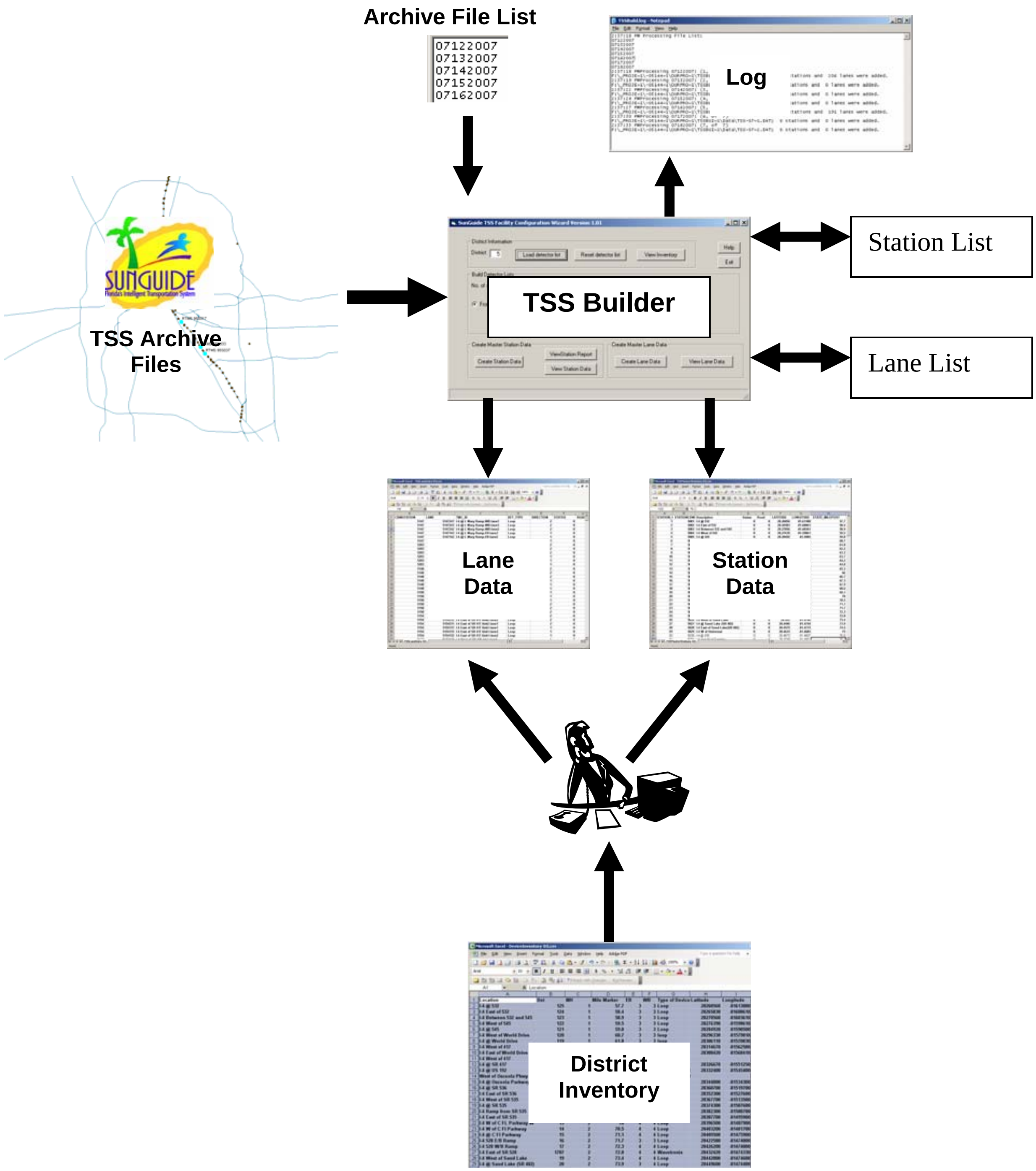


Figure F-4: TSS Builder data flow overview

The “Add TSS Detectors” command button will initiate the process of reading the archive files and appending any new stations or lanes to their respective lists. It will also generate a log showing what new information was added. Command buttons are also provided for viewing and resetting the log.

3. The “Create Master Station Data” and “Create Master Lane Data” frames have command buttons that initiate the conversion of the station and lane lists developed previously to the format of the corresponding station and lane data spreadsheets with blank fields for the data that must be added later. Command buttons are also provided to view and reset these data files, which are written in comma delimited (CSV) format.

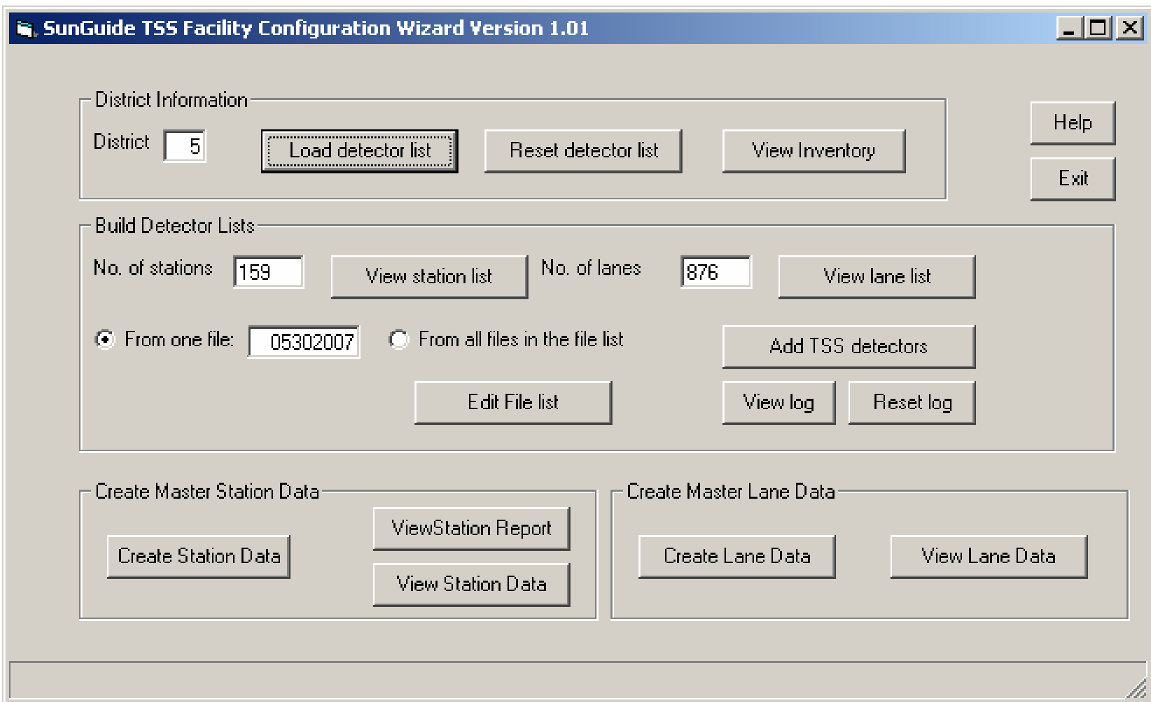


Figure F-5: TSS Builder main screen

4 Internal Utility Programs

The Oracle Warehouse Builder has several limitations for the support of the ETL process. For example, the Oracle Warehouse Builder loads CSV files with predefined names only. But in general, the input data files each have unique names from their data origin. To resolve these kinds of problems, several utility programs have been developed to give STEWARD the required flexibility and expandability.

4.1 Utility Program: Move2SunETLUtilityFolder.bat

This program is used for the Stage1_protransforma process to move raw data files from each district to the SunETLUtility input folder. The calling syntax is:

```
>> Move2SunETLUtilityFolder arg1
```

Where arg1: raw data file name with full directory path.

Example:

```
>> c:\steward\scripts\ Move2SunETLUtilityFolder.bat K:\District_Data\District-2\TSS_DATA\*.dat
```

4.2 Utility Program: Move2StewardFolder.exe

This program moves files in the ***GroupData*** folder in the ETL Utility to the STEWARD database input folders. This program parses the file name in the ***SunETLUtility\GroupData*** folder and distributes files to the ***TSS_STATION_DATA, TSS_LANE_DATA***, and ***TVT_DATA*** folders.

The calling syntax is:

```
>> Move2StewardFolder.exe arg1 arg2
```

The Command line arguments are:

- arg1: data folders where SunETLUtility program is installed.
- arg2: data folder which includes TSS_STATION_DATA, TSS_LANE_DATA, and TVT_DATA folders

Example:

```
>> C:\Steward\Scripts\Move2StewardFolder.exe C:\Progra~1\TRC\SunETLUutility C:\Steward
```

4.3 Utility Program: PrepareFileLoading.exe

The Oracle Data Warehouse requires a fixed file name for the external data file. Therefore, the data files from SunGuide need to be modified to the predefined file name. This program is designed to support this input procedure. Every time the program is executed, it deletes the default input data file and renames one of the input data files to the default input data file.

The calling syntax is:

```
>> PrepareFileLoading arg1 arg2
```

The arguments are:

- arg1: data folders where Oracle warehouse builder input data file is located
- arg2: predefined file name which Oracle warehouse builder use for data loading

Example:

```
>> c:\steward\scripts\PrepareFileLoading.exe c:\steward\TVT_DATA\tvt_15min_data tvt-
15min_data.csv
```

5 SunGuide TMC Data Archiving Subsystem Simulator (SimTMC)

SimTMC uses a microscopic traffic simulation model to emulate a SunGuide freeway traffic management center. The program consists of a preprocessor that creates input data files for the FRESIM freeway simulation model and a postprocessor that reads the output files, extracts the required data and creates data archive files in the prescribed SunGuide format. It serves two purposes in the CDW development process:

1. It makes data available immediately for CDW development. The data archiving system and communications infrastructure for data transmission have not yet been established and it is not know at this time when or how the CDW will start receiving data from RTMCs.
2. It provides a level of controllability in the data that will be useful for the development of a data quality assurance process in the CDW.

5.1 SunGuide Data Archive Files

The following requirements apply to each of the SunGuide data archive files:

1. Format shall be in comma separated values (CSV) that can be converted to fixed field length data by the FDOT.
2. Archive files shall contain 24 hours of data (midnight to midnight).
3. The location of the archive files shall be configurable.
4. The archive file name shall be in the following format: <subsystem name>-<date>-<log interval> Where:
    - <subsystem> acronym for subsystem (e.g. TSS, or RWIS)
    - <date> is the date in MMDDYYYY format
    - <log interval> is in an integer in the range of 1 min to 60 min.
5. The first line (header line) of each data file shall contain comma separated descriptive names for the detailed data to be logged.
6. Each detail line of the archive file shall contain comma separated fields.

Eight types of archive data are described in the SRS Document. Three of the data types may be produced by a microscopic traffic simulation model that simulates the operation of a freeway under TMC control. The file formats for these archives are as follows:

5.1.1 Incident Archives

For each SunGuide incident records the following data (1 incident per record):

- Timestamp (HH:MM:SS 24 hour format)
- Incident ID
- User
- Event details
- History of event

5.1.2 Detector Data archives

Each TSS detector records the following data (1 time slice per record):

- Timestamp (HH:MM:SS 24 hour format) (Presumed to indicate the beginning of the period)
- Detector identifier
- Speed (MPH, range 0 to 65535,in 1 MPH increments) [if not available set to 65535]
- Occupancy (0 to 65535 % in 0.1% increments) [if not available set to 65535]
- Volume (raw counts, 0 to 65535) [if not available set to 65535]

5.1.3 Travel Time Archives

Each travel time link defined records the following data (1 time slice per record):

- Timestamp (HH:MM:SS 24 hour format)
- Travel time link identifier
- Travel time (in minutes)
- Link status (ASCII: "in service" or "failed")

5.2 SimTMC Operation

Each TSS detector is represented by an entry in the detector list. The detector is assumed to be located at the downstream end of the link. Some control over detector placement may be accomplished by inserting nodes in the middle of physical links.

The program modules, data files and data flow are illustrated in Figure F-6.

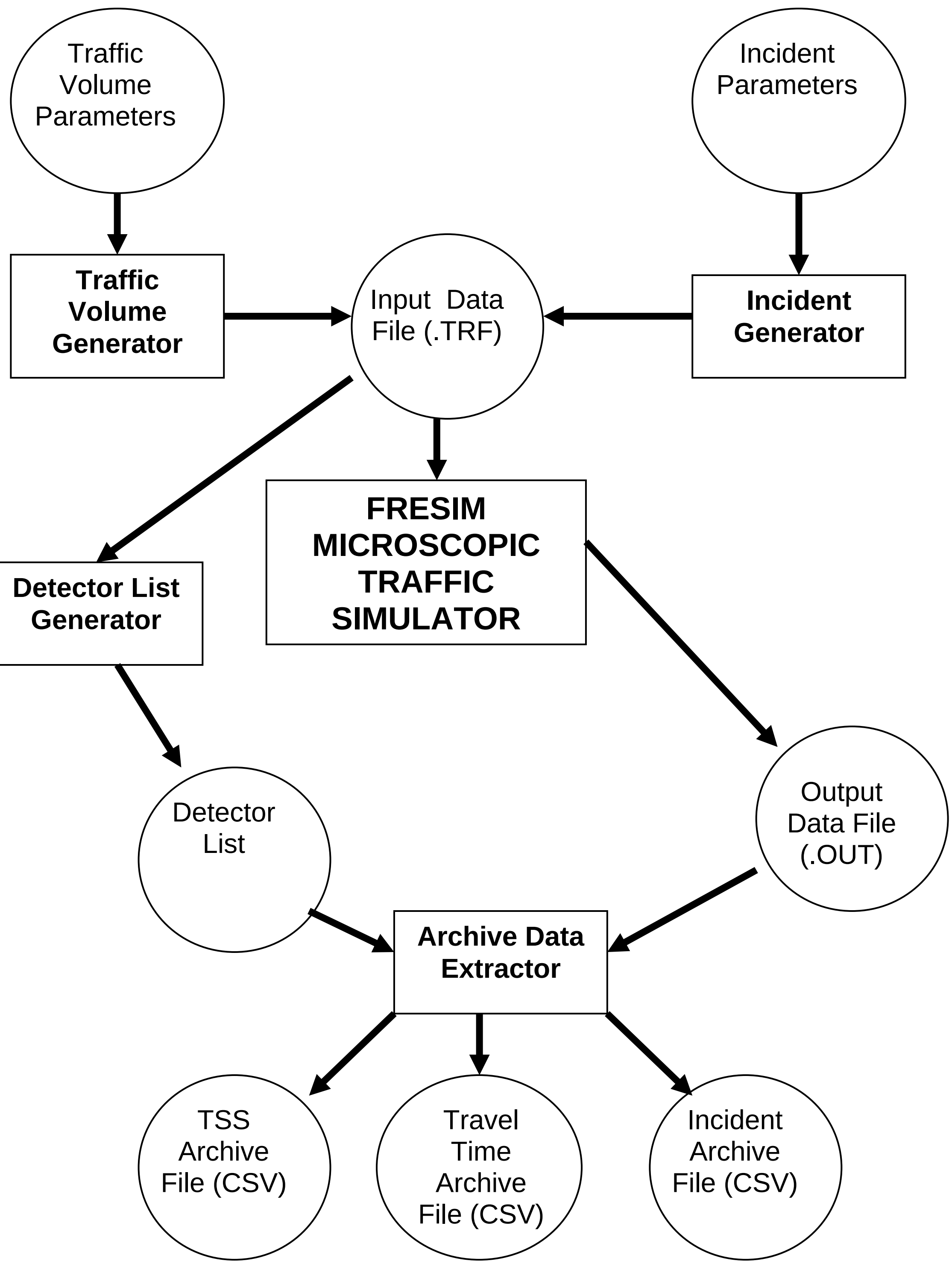


Figure F-6: SimTMC modules, data files and data flow.

5.3 SimTMC Components

The following SimTMC program modules are represented in Figure F-6.

5.3.1 Detector List generator

The detector list generator is a part of the postprocessor. It reads the input data file (TRF) and generates a list of links from the Record Type 19 entries. The parameters in the detector list include

- The upstream node of the link containing the detector
- The downstream node of the link containing the detector
- Length of the link
- The type of link (Mainline or Ramp)
- The number of through lanes on the link

The detector list is stored in a TXT file that may be edited using the Windows Notepad. The main reasons for editing are to eliminate links for which detectors are not desired and to insert detector ID numbers. If no detector ID numbers are inserted, the program will number the detectors consecutively as the links are read.

5.3.2 Data Archive Extractor

The data archive extractor is also a part of the postprocessor. The module reads the detector list and the FRESIM output file and creates the required CSV files in the SunGuide format. There are three archive files:

1. <u>The TSS Archive</u>

This file contains volume, speed and occupancy for each of the detectors for each time interval. Volume, speed and occupancy are computed as follows (i = interval number):

- Volume(i) = (Cumulative vehicle output(i) - Cumulative vehicle output (i-1)
- VehicleMiles(i) = Cumulative VehicleMiles(i) - Cumulative VehicleMiles(i-1)
- VehicleMin(i) = Cumulative VehicleMin(i) - Cumulative VehicleMin(i-1)
- Speed(i) = 60 * VehicleMiles(i) / VehicleMin(i)
- Density(i) = Volume(i) / Speed(i)
- The percent occupancy is given by the probability that a vehicle will be on top of the detector at any point in time. This is computed as (effective detector length, L) / (Vehicle Spacing, S) L is assumed as 30 ft. (20 ft for the vehicle and 10 ft for the loop). S is computed as (5280 / Density) * Number of lanes covered by the detector.

This Scheme seems to give reasonable numbers. Very short links with very low volumes can cause problems because of the roundoff of cumulative veh-miles and veh-mins to 1 decimal place in the FRESIM output file.

2. The Travel Time Archive
   This file contains travel times for each of the detectors for each time interval. The travel times are extracted directly from the FRESIM output file. One data item in this file is the *Link status (ASCII: "in service" or "failed").* We are not certain what this means and more discussion will be required to determine how to treat it.

3. The Incident Archive
   Incidents are described in FRESIM in terms of eight parameters discussed later in connection with the Incident Generator Module. The SunGuide archive format contains a file referred to simply as “event details.” Some further discussion will be required on how the RTMC operators plan to treat this item.

5.3.3 Traffic Volume Generator

This module creates an input data (TRF) file for a 24 hour day given a valid TRF file and a set of specifications indicating how the volumes should vary. Volumes will be randomized around the base volumes plus a bias factor that will apply to the whole day. This module is now under development.

5.3.4 Incident Generator

This module adds incidents to the TRF file. FRESIM incidents are specified on Record Type 29 of the TRF file in terms of:

1. The incident code and its effect for each lane
   - 0 = Normal speed
   - 1 = Traffic capacity reduced by the rubberneck factor at the point of the incident
   - 2 = Blockage at point of incident

2. The longitudinal location of the upstream end of the incident from the upstream node. It must be less than the link length.
3. The length of the roadway affected by the incident. This value must be greater than zero. The affected length can exceed the length of the link.
4. The time of onset of the incident (in seconds). Time is measured from the start of the simulation. For a blockage incident, a value of zero will place the blockage at the beginning of the initialization period. This value must be less than the length of the simulation.
5. The duration of the incident (in seconds). It must be greater than zero.
6. The rubberneck factor (as a percentage). This value must be greater than zero if the incident code is one. The rubberneck factor (in a percentage) represents the reduction in capacity at the point of the incident for vehicles that are in lanes that have an incident code of one.
7. The location of the upstream warning sign for blockage incidents.
8. The location (upstream of the incident) at which vehicles will respond to the blockage by attempting to lane-change away from the lane(s) affected by the blockage. This capability was designed to reflect the fact that signs are usually placed on the roadway to warn

motorists that a work zone is ahead and to indicate which lanes are affected. The warning sign might be placed even further upstream from the blockage if the work zone is a long-term situation and if it is believed that motorists respond to it even before they reach the warning sign. This field should be set to a small value (a few feet) for non-recurring incidents because motorists usually cannot respond to them until they see the blockages. A blank or zero results in the default value being used.

Incidents will be added to the TRF files with some randomization of these parameters. This module is under development.

5.4 Sample TSS Archive File

Figure F-7 presents a sample TSS archive file conforming to the SunGuide archive file specification. The file is named "TSS-08292005-30.csv", as prescribed in this specification. The log interval is 30 seconds. Two consecutive intervals are represented for a system of eleven detectors.

```
Time, Detector, Speed, Volume, Occupancy
07:00:30,  1 ,  63 ,  44 , 15.9
07:00:30,  2 ,  63 ,  42 , 15.1
07:00:30,  7 ,  60 ,  3 , 3.4
07:00:30,  3 ,  59 ,  43 , 16.6
07:00:30,  4 ,  62 ,  36 , 13.2
07:00:30,  9 ,  60 ,  9 , 10.2
07:00:30,  5 ,  62 ,  44 , 16.
07:00:30,  6 ,  62 ,  26 , 9.5
07:00:30,  11 ,  57 ,  18 , 10.7
07:00:30,  10 ,  67 ,  7 , 3.6
07:00:30,  8 ,  0 ,  1 , 0.0
07:01:00,  1 ,  63 ,  45 , 16.3
07:01:00,  2 ,  60 ,  41 , 15.5
07:01:00,  7 ,  60 ,  6 , 6.8
07:01:00,  3 ,  62 ,  43 , 15.8
07:01:00,  4 ,  62 ,  32 , 11.7
07:01:00,  9 ,  64 ,  13 , 13.8
07:01:00,  5 ,  61 ,  40 , 14.8
07:01:00,  6 ,  64 ,  19 , 6.7
07:01:00,  11 ,  54 ,  20 , 12.7
07:01:00,  10 ,  60 ,  8 , 4.5
07:01:00,  8 ,  30 ,  3 , 6.8
```

Figure F-7: Portion of a sample TSS archive file

5.5 SimTMC Data Preparation Wizard

The main screen display for the SimTMC data preparation wizard is shown in Figure F-8. Four separate steps are included in the process:

1. *Create link list files from TRF files:* A valid TRF file name (Base TRF File) must be given to initiate this step. The wizard reads this file and generates a list of detector links and input links for use in the next step.

2. *Make SimTMC Input file:* This step reads the base TRF file and both of the link list files to create a multi-period TRF file that simulates the freeway over the specified time periods. The default time periods are:
   - 0700-0930
   - 0930-1200
   - 1200-1430
   - 1430-1630
   - 16-30-1900
   - 1900-2130
   - 2130-2400

   These time periods are established on the FRRSIM input record . Incidents are randomized and 1 to 3 incidents are added on the FRESIM RT29. One of the limitations of FRESIM is that the field representing the time of the incident is limited to four columns. This imposes a limit of 9999 seconds (i.e., less than three hours) from the beginning of the simulation. This is an unfortunate limitation for long term simulations. It is also an unnecessary limitation considering all of the free space on the RT29 format.

3. *Run FRESIM:* FRESIM must be run to produce an output file (.OUT) from the SimTMC input file. In the current version this is an external step.

4. *Create Archive files:* The three archive files described above are created by reading the specified FRESIM output file, extracting the data and generating the comma-delimited text files in the proper format.

Each of the steps may be executed by clicking the corresponding box. The comma delimited archive files may be displayed by clicking their associated buttons.

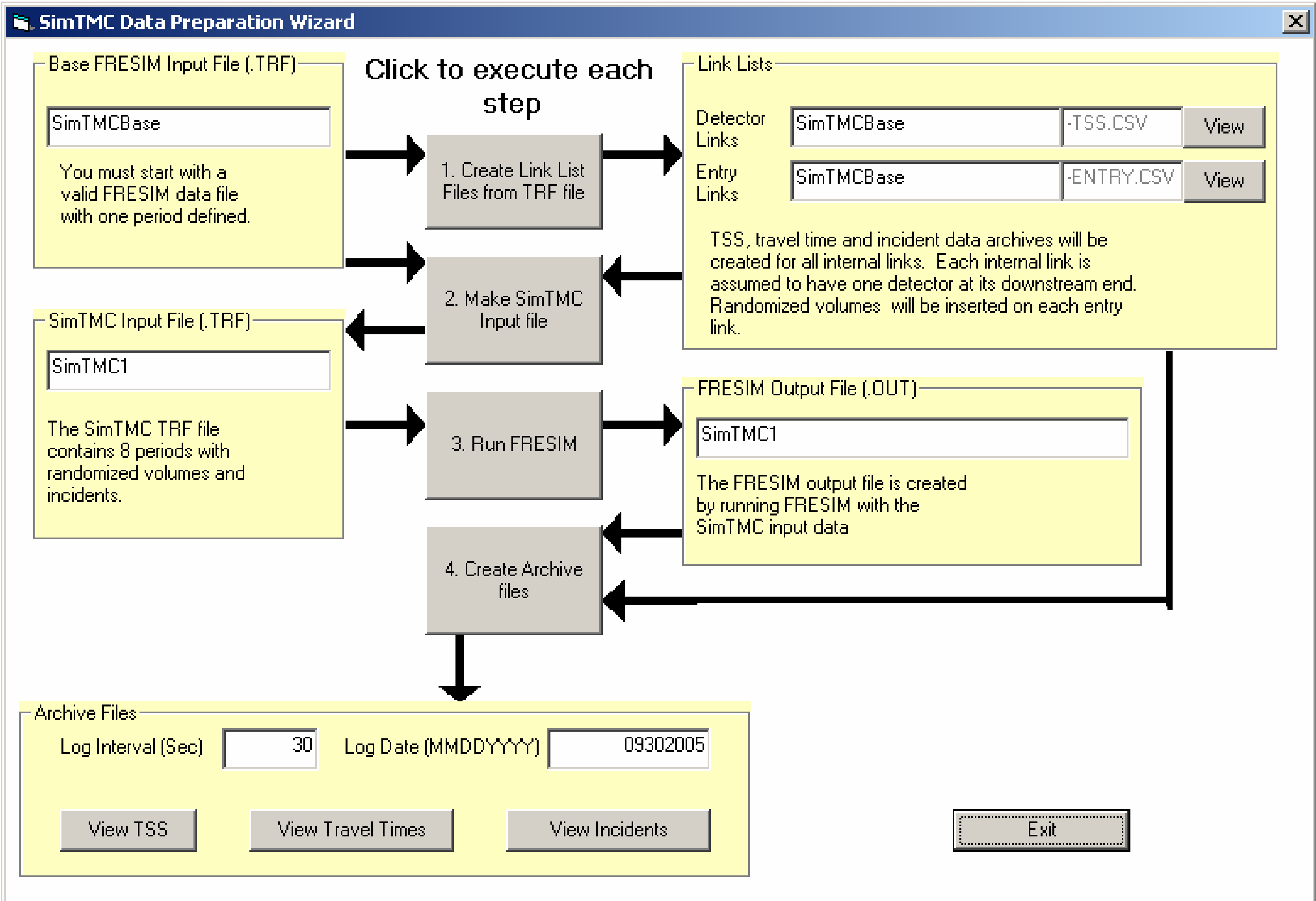


Figure F-8: SimTMC Data Preparation Wizard screen display

6 Milepost Converter Program

This program converts between State mileposts, RCI Roadway mileposts on I-95 and latitude/longitude coordinates. Conversion in both directions (i.e, mileposts to coordinates and coordinates to mileposts) is supported.

6.1 User Interface

The user interface screen is presented in Figure F-9. The user selects one of three input formats by clicking the corresponding radio button for State milepost, RCI Roadway mile post or latitude/longitude coordinates.

- State mileposts on I-95 start from Miami to the Florida/Georgia state line (0 mile ~382.083mile).
- The RCI Roadway mileposts represent the relative milepost on each roadway. The roadways are defined in Table F-1.
- Latitude/longitude data in decimal format are converted to milepost data.
- Calibration factor: For the latitude/longitude to milepost conversion, a “miles per degree” factor is required. In Jacksonville, 1 degree in latitude is converted to 69.24 miles and 1 degree in longitude is converted to 59.82 miles. These values are used as default for the program. In Miami, latitude factor remains the same but 1 degree in longitude is converted to 62.47 miles.
- If the distance from the latitude/longitude location to I-95 is larger than the specified maximum, the distance field in Output Data frame will be displayed in a red font.

The conversion is initiated by clicking the ***Convert*** button and the results appear in the ***Output Data*** frame.

6.2 Program Logic

The program uses reference tables for data conversion. These tables are generated from the Florida roads map in the FDOT web site using the ArcGIS software. Each table represents the location data for each roadway and is archived in the program directory with the roadway name and “CSV” extension. These tables contain the state/roadway mileposts in 0.001mile (5.28ft) resolution and their corresponding longitude/latitude data.

- State/roadway milepost to longitude/latitude conversion: The program looks up the longitude/latitude data from the milepost table and returns its values.
- Latitude/longitude data to mileposts: The program calculates the distance from the input latitude/longitude data to the latitude/longitude data in the reference tables. The milepost with the shortest distance point in the reference tables will be presented as the output. If the distance is larger than the maximum distance from the user input, the distance field in the ***Output Data*** frame will be displayed in a red font.
- All of the milepost and coordinate data are calculated from the north bound direction of I-95. Therefore, the location data from the south bound direction of I-95 could include minor calculation errors during the conversion by its nature. These errors appear to be relatively small.

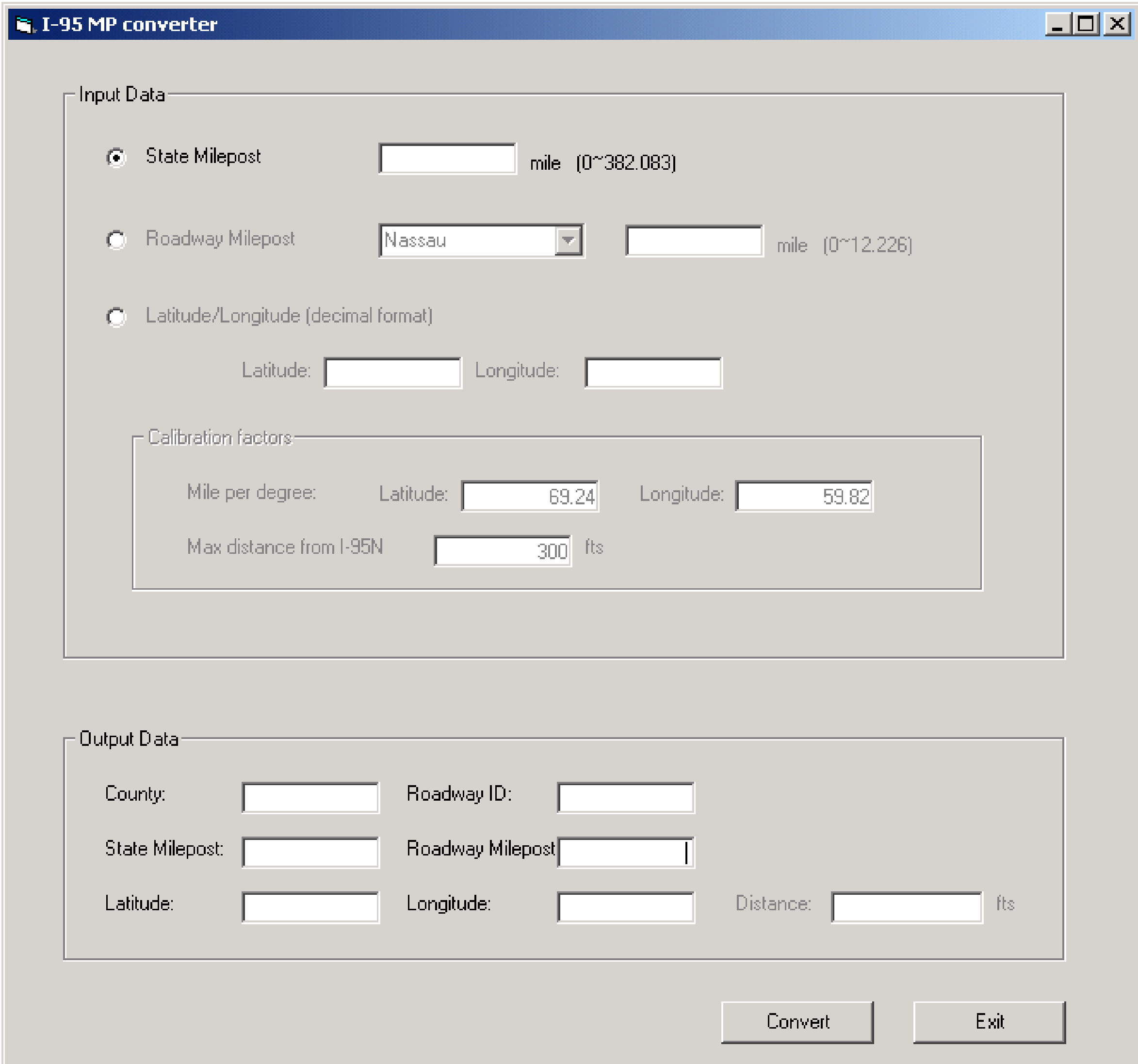


Figure F-9: Milepost Converter user interface screen

Table F-1: RCI Roadways on I-95 in Florida

Roadway ID	Name	County	Length (mi)	State Milepost
74160000	SR 9 (I-95)	Nassau	12.226	369.857
72290000	SR-9 I-95	Duval(upper)	10.513	359.344
72020000	SR-9 I-95	Duval(mid)	10.593	348.751
72280000	SR 9 I-95	Duval(lower)	16.793	331.958
78080000	SR9;FLAG-DUVAL C.L.	St.Johns	34.855	297.103
73001000	I-95(SR-9)VOL-STJOHN	Flagler	18.729	278.374
79002000	SR9;BREVARD-FLAGLER	Volusia	45.804	232.57
70225000	SR9;SR520 TO VOL.CO.	Brevard (north)	31.19	201.38
70220000	SR9;IND.RIV.-SR520	Brevard (south)	41.503	159.877
88081000	SR 9 (I-95)	Indian River	19.198	140.679
94001000	SR-9 I-95	St.Lucie	27.259	113.42
89095000	I-95/SR 9	Martin	24.835	88.585
93220000	I-95/SR 9	Palm Beach	46.018	42.567
86070000	SR 9/INT 95	Broward	25.307	17.26
87270000	SR-9/SR9A/I-95	Miami-Dade	17.26	0

6.3 Program Verification

This program was verified from GPS field measurement data. Two points in Volusia County were used to verify the program results in quantitative detail.

Location 1: SB Entrance from I-4 to I-95

- GPS measurement:
  Latitude: 29.1557, Longitude: -81.07608333
- ArcGIS extimation:
  Latitude: 29.15578, Longitude: -81.075998
- Distance between two data points: 40.51ft

Location 2: Crossing point of SR44 and I-95

- GPS measurement:
  Latitude: 29.01288333, Longitude: -80.98836667
- ArcGIS extimation:
  Latitude: 29.01280312, Longitude: -80.98830365
- Distance between two data points: 35.44ft

Qualitative verification was also performed by inspection of the GPS and ArcGIS coordinates for each milepost covering a 20 mile section of I-95 in this area.