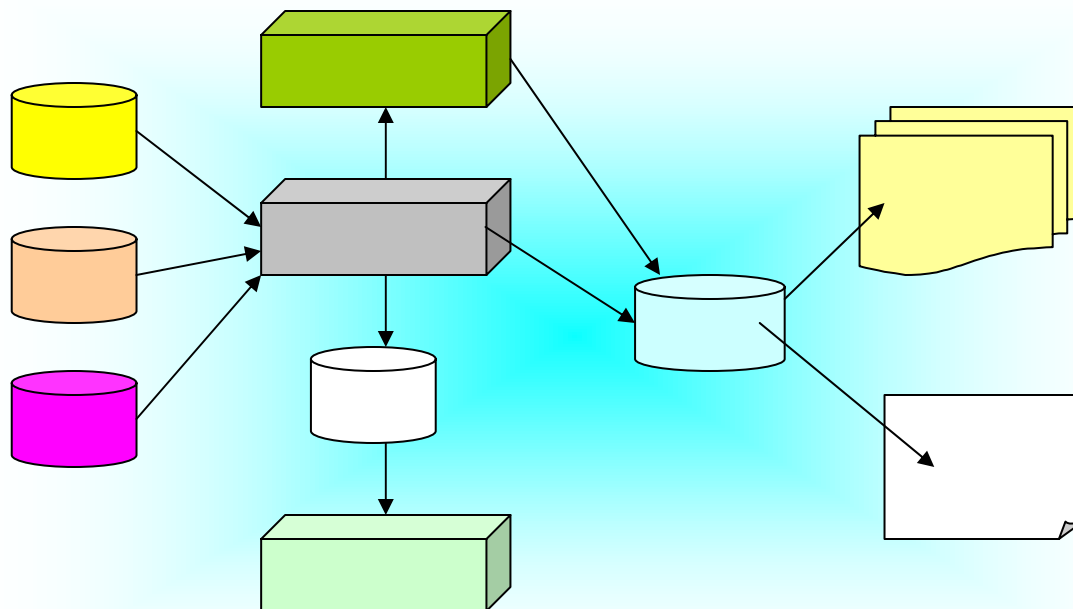




# **Feasibility Study for an Integrated Network of Data Sources**

**Final Report**

**Contract BC-354-61  
UF Project 49104504880**

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

January 2004

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                     |                                                                                                                                                                        |          |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| 1. Report No.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 2. Government Accession No.                         | 3. Recipient's Catalog No.                                                                                                                                             |          |
| 4. Title and Subtitle<br><br>Feasibility Study for an Integrated Network of Data Sources                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                     | 5. Report Date<br>January 2004                                                                                                                                         |          |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                     | 6. Performing Organization Code                                                                                                                                        |          |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                     | 8. Performing Organization Report No.<br>UFTRC 880-1                                                                                                                   |          |
| 7. Author(s)<br>K. Courage, I. Li, J Hammer, F. Ji , Q Yu                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                     | 10. Work Unit No. (TRAIS)                                                                                                                                              |          |
| 9. Performing Organization Name and Address<br>University of Florida<br>Department of Civil and Coastal Engineering<br>365 Weil Hall / P.O. Box 116580<br>Gainesville, FL 32611-6580                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                                                     | 11. Contract or Grant No.<br>BC-354-61                                                                                                                                 |          |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                     | 13. Type of Report and Period Covered<br>Final Report                                                                                                                  |          |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                     | 14. Sponsoring Agency Code                                                                                                                                             |          |
| 12. Sponsoring Agency Name and Address<br>Florida Department of Transportation<br>Research Management Center<br>605 Suwannee Street, MS 30<br>Tallahassee, FL 32301-8064                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                     |                                                                                                                                                                        |          |
| 15. Supplementary Notes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                     |                                                                                                                                                                        |          |
| 16. Abstract<br><br>Florida's Statewide Intelligent Transportation Systems (ITS) Architecture contains a central data warehouse (CDW) component in its Statewide Concept of Operations. The principal function of a CDW is to collect and store statewide information from many data sources and make it available for use by different applications. The CDW will allow data from these separate and disparate data sources to be aggregated.<br><br>Florida Department of Transportation Contract BC-354-61 was initiated to perform a feasibility study for a central data warehouse (CDW) for data generated and used by various ITS applications in Florida and to provide preliminary recommendations on the configuration and functionality of a CDW in preparation for a future project phase.<br><br>This report contains 1) A summary the resources available for the development of a central data warehouse for ITS applications; 2) A discussion of the common concepts and standards that apply to the design of a CDW; 3) Observations and recommendations that can be offered at this time, given the relatively coarse level of the investigations that have been done to this point; 4) A description of prototype for a CDW, designed to illustrate the concepts and recommendations contained in this report. and 5) A prospectus for a proof-of-concept study as a subsequent phase of this project. |                                                     |                                                                                                                                                                        |          |
| 17. Key Words<br><br>ITS, ADUS, Data warehouse                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                                                     | 18. Distribution Statement<br><br>No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA, 22161 |          |
| 19 Security Classif. (of this report)<br>Unclassified                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 20 Security Classif. (of this page)<br>Unclassified | 21.No. of Pages<br>90                                                                                                                                                  | 22 Price |

### **Acknowledgements**

The project described in this report was carried out for the Florida Department of Transportation (FDOT) by the University of Florida Transportation Research Center in cooperation with the Department of Computer and Information Science and Engineering.

The overall effort was coordinated By Prof. Ken Courage with significant technical support from Dr. Joachim Hammer, Dr. Irene Li and Dr. Charles Wallace, Mr. Feng Ji, and Mr. Qingyong Yu.

The FDOT technical coordinator was Liang Hsia, P.E.

### **Disclaimer**

The opinions, findings and conclusions presented in this report are those of the authors and not necessarily those of the Florida Department of Transportation or any other government agency

**Table of Common Acronyms Used in this Report**

| <b>Name</b> | <b>Description</b>                                                 | <b>Comments</b>                                                                         |
|-------------|--------------------------------------------------------------------|-----------------------------------------------------------------------------------------|
| AADT        | Annual Average Daily Traffic                                       |                                                                                         |
| AASHTO      | American Association of State Highway and Transportation Officials |                                                                                         |
| ADMS        | Archived Data Management Subsystem                                 |                                                                                         |
| API         | Application Programming Interface                                  |                                                                                         |
| ATMS        | Advanced Transportation Management System                          |                                                                                         |
| BMI         | Bridge Management Inventory                                        |                                                                                         |
| C2C         | Center to Center                                                   |                                                                                         |
| CART        | Classification and Regression Trees                                |                                                                                         |
| CCTV        | Closed Circuit Television                                          |                                                                                         |
| CDW         | Central Data Warehouse                                             |                                                                                         |
| CORBA       | Common Object Request Broker Architecture                          |                                                                                         |
| DATEX       | Data Exchange                                                      |                                                                                         |
| DATEX-ASN   | Data Exchange in Abstract Syntax Notation One                      |                                                                                         |
| DBMS        | Database Management System                                         |                                                                                         |
| DEP         | Department of Environmental Protection                             |                                                                                         |
| DOT         | Department of Transportation                                       |                                                                                         |
| DSS         | Decision Support System                                            |                                                                                         |
| EIS         | Enterprise Information System                                      |                                                                                         |
| EMC         | Emergency Management Center                                        |                                                                                         |
| EMS         | Emergency Management Service                                       |                                                                                         |
| ETMCC       | External Traffic Management Center Communications                  |                                                                                         |
| FDOT        | Florida Department of Transportation                               |                                                                                         |
| FTP         | File Transfer Protocol                                             |                                                                                         |
| GIS         | Global Information System                                          |                                                                                         |
| GRIP        | Geo-referenced Information Portal                                  |                                                                                         |
| HTTP        | Hypertext Transfer Protocol                                        | A client-server protocol by which two machines can communicate over a tcp/ip connection |
| IDL         | Interface Definition Language                                      |                                                                                         |
| IEEE        | Institute of Electrical and Electronics Engineers                  |                                                                                         |
| IIMS        | Integrated Incident Management System                              |                                                                                         |
| IIOP        | Internet ORB Protocol                                              |                                                                                         |
| ITE         | Institute of Transportation Engineers                              |                                                                                         |
| ITS         | Intelligent Transportation Systems                                 |                                                                                         |

|            |                                                         |                                                                                                   |
|------------|---------------------------------------------------------|---------------------------------------------------------------------------------------------------|
| ITS-DR     | Transportation Systems Data Registry                    | The ITS DR is intended to serve as a common or shared data reference for the national ITS domain. |
| JDBC       | JAVA Data Base Connectivity                             | A Java API for access to a relational database                                                    |
| JSP        | Java Server Pages                                       | A technology for mixing static HTML with dynamically-generated HTML                               |
| LAN        | Local Area Network                                      |                                                                                                   |
| MS         | Message Sets                                            |                                                                                                   |
| NEMA       | National Electrical Manufacturers Association           |                                                                                                   |
| NTCIP      | National Transportation Communications for ITS Protocol | A Family of Communication Standards and Data Definition Standards                                 |
| OIS        | Office of Information Systems                           |                                                                                                   |
| OLAP       | On-line Analysis Processing                             |                                                                                                   |
| OMG        | Object Management Group                                 |                                                                                                   |
| OOCEA      | Orlando-Orange County Expressway Auth.                  |                                                                                                   |
| Oracle 8.X | Oracle DBMS version 8 series                            |                                                                                                   |
| ORB        | Object Request Broker                                   |                                                                                                   |
| OSI        | Open System Interconnection                             |                                                                                                   |
| PCSs       | Permanent Count Stations                                |                                                                                                   |
| RCI        | Roadway Characteristics Inventory                       |                                                                                                   |
| RDBMS      | Relational Database Management System                   |                                                                                                   |
| RHC        | Rail-Highway Crossing                                   |                                                                                                   |
| SNMP       | Simple Network Management Protocol                      |                                                                                                   |
| SQL        | Structured Query Language                               |                                                                                                   |
| STMP       | Simple Transportation Management Protocol               |                                                                                                   |
| TCP/IP     | Transmission Control Protocol / Internet Protocol       | A suite of protocols Rules for sending and receiving data across networks                         |
| TFTP       | Trivial File Transfer Protocol                          |                                                                                                   |
| TMDD       | Traffic Management Data Dictionary                      |                                                                                                   |
| UDP        | User Datagram Protocol                                  | A transport protocol layered on top of the Internet Protocol                                      |
| UI         | User Interface                                          |                                                                                                   |
| USGS       | US Geological Survey                                    |                                                                                                   |
| VMT        | Vehicle Miles Traveled                                  |                                                                                                   |
| VPN        | Virtual Privacy Networks                                |                                                                                                   |
| WSDOT      | Washington State Department of Transportation           |                                                                                                   |

## Table of Contents

|       |                                                                                                              |    |
|-------|--------------------------------------------------------------------------------------------------------------|----|
| 1     | INTRODUCTION .....                                                                                           | 1  |
| 1.1   | Problem Statement .....                                                                                      | 1  |
| 1.2   | The Nature, Benefits and Challenges of a Central Data Warehouse .....                                        | 2  |
| 1.3   | Project Objectives & Tasks .....                                                                             | 5  |
| 2     | SUMMARY OF CLIENT INFORMATION .....                                                                          | 7  |
| 2.1   | ITS-related Applications Developed in the FDOT Districts .....                                               | 7  |
| 2.2   | The FDOT Enterprise Information System .....                                                                 | 10 |
| 2.3   | Established Florida Statewide ITS Architecture .....                                                         | 10 |
| 2.4   | Available Data at the Traffic Management Centers and Other Centers .....                                     | 12 |
| 2.5   | Data Available from GRIP [6] .....                                                                           | 13 |
| 2.5.1 | Data Collection Sites and Devices .....                                                                      | 16 |
| 2.5.2 | Data Collection and Data Products .....                                                                      | 16 |
| 2.5.3 | Data collection quality control .....                                                                        | 17 |
| 2.5.4 | Data storage .....                                                                                           | 17 |
| 2.6   | Emergency Management Center Information .....                                                                | 18 |
| 2.7   | Other Data .....                                                                                             | 18 |
| 2.8   | Future Research Tasks .....                                                                                  | 18 |
| 3     | CDW DESIGN CONCEPTS AND STANDARDS .....                                                                      | 20 |
| 3.1   | Information Flow Concepts .....                                                                              | 20 |
| 3.2   | General Industry Standards .....                                                                             | 23 |
| 3.2.1 | Common Object Request Broker Architecture (CORBA) [1] .....                                                  | 23 |
| 3.2.2 | Data Exchange Using Abstract Syntax Notation (DATEX-ASN) [2] .....                                           | 25 |
| 3.2.3 | Extensible Markup Language (XML) .....                                                                       | 26 |
| 3.2.4 | Other Protocols .....                                                                                        | 26 |
| 3.3   | Standards Developed for ITS Applications .....                                                               | 27 |
| 3.3.1 | Archived Data Management Services (ADMS) [3] .....                                                           | 27 |
| 3.3.2 | Traffic Management Data Dictionary (TMDD)[4] .....                                                           | 28 |
| 3.3.3 | Message Sets for External Traffic Management Center Communications (MS/ETMCC) [5] .....                      | 29 |
| 3.3.4 | IEEE 1512 Standard for Common Incident Management Message Sets for Use by Emergency Management Centers ..... | 30 |
| 3.3.5 | NTCIP .....                                                                                                  | 30 |
| 3.3.6 | Center to Center (C2C) Communications .....                                                                  | 31 |
| 3.3.7 | ITS Registry .....                                                                                           | 32 |
| 3.4   | Future Research Tasks .....                                                                                  | 33 |
| 4     | PRELIMINARY DESIGN RECOMMENDATIONS .....                                                                     | 34 |
| 4.1   | High-level CDW Requirements .....                                                                            | 34 |
| 4.1.1 | Data Presentation Requirements .....                                                                         | 34 |
| 4.1.2 | Decision Support Requirements .....                                                                          | 35 |
| 4.1.3 | Technology Requirements .....                                                                                | 37 |
| 4.2   | CDW Multi-tier Architecture .....                                                                            | 40 |
| 4.3   | Data Stored In the CDW .....                                                                                 | 41 |
| 4.4   | Data Warehouse Components [9] [10] .....                                                                     | 41 |
| 4.4.1 | Data Mapping .....                                                                                           | 41 |
| 4.4.2 | Data Source Extraction .....                                                                                 | 44 |

|        |                                                                 |    |
|--------|-----------------------------------------------------------------|----|
| 4.4.3  | Data Scrubbing.....                                             | 44 |
| 4.4.4  | Data Fusion .....                                               | 45 |
| 4.4.5  | Data Loading.....                                               | 46 |
| 4.4.6  | Table Summarization / Aggregation.....                          | 46 |
| 4.4.7  | Data Replication / Distribution .....                           | 47 |
| 4.5    | Schema Design.....                                              | 47 |
| 4.5.1  | CDW Application Background.....                                 | 47 |
| 4.5.2  | The Detailed Facts Table .....                                  | 48 |
| 4.5.3  | Dimension Table.....                                            | 48 |
| 4.5.4  | Aggregation Tables.....                                         | 49 |
| 4.5.5  | History Tables.....                                             | 50 |
| 4.5.6  | Status Table.....                                               | 50 |
| 4.5.7  | Sample Schema Tables for CDW application .....                  | 50 |
| 4.6    | Implementation Considerations .....                             | 51 |
| 4.7    | Risk Analysis .....                                             | 52 |
| 4.7.1  | Risk Type: Inadequate Project Scope .....                       | 54 |
| 4.7.2  | Risk Type --- Inadequate Control of Project Costs.....          | 54 |
| 4.7.3  | Risk Type --- Lack of Funding .....                             | 55 |
| 4.7.4  | Risk Type --- Inadequate Human Resources .....                  | 55 |
| 4.7.5  | Risk Type --- Resistance to Change.....                         | 56 |
| 4.7.6  | Risk Type – Lack of Training.....                               | 56 |
| 4.7.7  | Risk Type --- Data Conversion Uncertainty .....                 | 57 |
| 4.7.8  | Risk Type --- Hardware and Software Failures .....              | 57 |
| 4.7.9  | Risk Type – Inadequate Data Processing Capacity .....           | 58 |
| 4.7.10 | Risk Type – Bad Data Quality .....                              | 58 |
| 4.8    | Future Research Tasks .....                                     | 59 |
| 5      | PROTOTYPE CDW DEMONSTRATION.....                                | 61 |
| 5.1    | What Is Emergency Management in ITS? .....                      | 61 |
| 5.2    | Participating Agencies .....                                    | 64 |
| 5.3    | System Architecture and Data Flow .....                         | 66 |
| 5.3.1  | District-level EMS .....                                        | 66 |
| 5.3.2  | State-level EMS .....                                           | 71 |
| 5.4    | An Application of a CDW for Emergency Management Services ..... | 73 |
| 5.4.1  | EMS as A Candidate Component of CDW within FDOT EIS.....        | 75 |
| 5.4.2  | Data Organization .....                                         | 77 |
| 5.4.3  | Incident Rerouting Plan .....                                   | 78 |
| 5.4.4  | User Interface for EMS Demo .....                               | 88 |
| 5.5    | Future Research Tasks .....                                     | 89 |

Appendix A: Research Needs Statement for Phase II

Appendix B: CDW Prototype Demonstration

# 1 INTRODUCTION

Florida Department of Transportation Contract BC-354-61 was initiated to perform a feasibility study for a central data warehouse (CDW) for data generated and used by various ITS applications in Florida and to provide preliminary recommendations on the configuration and functionality of a CDW in preparation for a future project phase. This report includes documentation of the progress made to this point on the preliminary design recommendations, an Internet-based demonstration of a prototype CDW function, and a specific prospectus for a proof of concept study in the format of a research needs statement in the prescribed FDOT format.

## 1.1 Problem Statement

The Statewide Intelligent Transportation Systems (ITS) Architecture [1] contains in its Statewide Concept of Operations a “central data warehouse” component as follows:

*The ultimate purpose of the central data warehousing would be to provide both traveler information, but also traffic management services. It could likely expand to more widespread service, such as emergency management, disaster mitigation and even disaster management.*

While the name “central data warehouse” (CDW) implies a centralized database system that collects data from many sources, it will more practically operate as a networked system of databases. These databases may range from those in traffic control centers and transportation management centers to emergency management centers, weather centers, disaster management centers, ports (air and sea), motor carrier control centers and a variety of others. Data needs would range from traffic, transit, equipment inventory (both deployed as well as stockpiled) and

traveler information, to weather, event and emergency management data, all superimposed and presented in a geographic information system (GIS).

All, or certainly most, of these are (at least presently) unique, autonomous, independent and stand-alone database systems. They are generally not integrated or even linked to provide any meaningful data warehousing services. The proposed system must support the integration and merging of schema and data from the underlying repositories.

There are several candidate alternatives that arise at all levels in the design of a CDW especially with respect to the level of centralization or decentralization. The advantages and disadvantages of each alternative must be weighed carefully against the requirements of the project. The problem to be addressed in this study is the development of a high level view of the scope of the proposed CDW, in terms of the information needs, the data availability, the hardware and software platforms, the communications requirements and the data processing requirements.

## **1.2 The Nature, Benefits and Challenges of a Central Data Warehouse**

A CDW is one of the major components in a typical ITS application. Its principal function is to collect and store statewide information from many data sources and make it available for use by different applications or by the operator.

The main features of the CDW include:

- A statewide data warehouse to allow data from what would normally be separate and disparate applications to be aggregated, enabling extra functions at reduced cost.
- Flexible data protection and security to ensure that the data in the CDW is secure with access to it being set by the user [2].
- Common data services that not only store data but also manage their delivery to other applications

The main reasons for the development of a CDW are:

- To provide more and better information in managing and operating the transportation system: ITS-generated data can be stored in a CDW at different levels of detail to meet the needs of different applications. The data can be used to predict when and where problems may occur again, as well as to evaluate alternative strategies for preventing or mitigating problem(s). In addition, as the focus of transportation policy shifts away from large-scale, long-range capital improvements and toward better management of existing facilities, ITS-generated data stored in a CDW can support the creation and use of new system performance measures.
- To promote the use of archived data for multiple purposes: The continuous nature of most data stored in a CDW removes sampling bias from estimates and supports the study of variability required for transportation planning, performance management, congestion management and research studies. Multi-purpose use of archived data offers a way to stimulate stakeholder support for ITS initiatives.
- To maximize cost-effectiveness of the data collection infrastructure: Data archiving permits transportation agencies to optimize their investments in data collection infrastructure by re-using the same data for numerous transportation planning, design, operations, and research needs.
- To reduce data acquisition costs: Data archiving is significantly less expensive than having a planning or design group re-collect even a small percentage of the data using manual methods or special studies.
- To support automated government reporting systems: A major use of ITS-generated data is to support periodic updates to government reporting systems. Once common data have been defined, the CDW can provide an automated process for updating government reporting systems.
- To conform to established business practice in other industries: The retention and analysis of operational data is an established practice in most competitive industries that use data to manage their business activities. For example, the retail sales industry uses data warehouses full of customer transactions and inventories to better

understand the basics of supply and demand in numerous markets around the world [3].

- To facilitate the introduction of new applications: With a data warehouse the introduction of a new application is simple and generally requires only an interface to the central data warehouse.

A CDW requires a centralized database system that collects, archives, and analyzes data from many sources to generate powerful and useful applications. However, the concept of a networked system of databases in the transportation industry is fairly new. The following challenges are known to face the design and implementation of a CDW [4]:

- Available transportation data are collected by many stakeholders with no common standards. The means by which data are stored range from hard copy files to those managed by commercial databases. It is very difficult to fuse all kinds of data with different formats and storage means.
- In the national ITS architecture, Archived Data User Service (ADUS) standardization for data quality and accessibility is just getting underway. The newly developed ASTM standard for ADUS offers a potential solution to this problem.
- Current data available to support needed short-term application are insufficient. Little experience exists for multiple uses of data generated by ITS projects.
- Administrative and cross jurisdiction problems arise when trying to collect data maintained by different agencies. Some stakeholders are reluctant to share data among jurisdictions due to cost, liability, and lack of personnel resources. In addition, some agencies are reluctant to adopt new technology, partly because of the lack of staff qualified in advanced technology.
- A centralized data warehousing system has a very high requirement for data storage and communication infrastructure. The costs can run as high as \$1.8 million. Some of the noticeable costs associated with data warehousing are time spent in careful analysis of measurable needs, design and implementation effort, hardware cost, software costs, on-going support and maintenance, and resulting re-engineering effort.

### **1.3 Project Objectives & Tasks**

The primary objectives of this project were to develop a conceptual design for a statewide CDW covering the full range of ITS functions and services and to produce a roadmap for further development. The following chapters and appendices summarize the tasks performed in support of these objectives:

- Chapter 2 summarizes the resources available for the development of a central data warehouse for ITS applications. It examines the current FDOT activities and data infrastructure and identifies the key data items that are available from a variety of sources. The function of a future CDW will be to place the available data within an expanded version of the current infrastructure to make the information more usable to the ITS-community.
- Chapter 3 presents a discussion of the common concepts and standards that apply to the design of a CDW. The concepts are taken from general industry literature. Several standards are identified that must also be considered in the development of any ITS component. Some are general industry-wide standards, while others have been developed specifically with ITS functionality in mind.
- Chapter 4 presents the observations and recommendations that can be offered at this time, given the relatively coarse level of the investigations that have been done to this point. These recommendations define the nature of the system that should be pursued in the next phase of the study, during which substantially more detail will be required.
- Chapter 5 describes a physical prototype for a CDW, designed to illustrate the concepts and recommendations contained in this report. While the prototype is functional and accessible from the Internet, it is not capable of doing useful work. Its sole purpose is that of demonstration.
- Appendix A contains a prospectus for a proof-of-concept study in the FDOT research needs statement format.
- Appendix B includes the operating instructions and CD-ROM for the prototype demonstration based on emergency management described in Chapter 5. The full demonstration exists physically on an Internet-based server in the form of an ORACLE application. The main content of Appendix B is delivered on a CD-ROM that contains the HTML files included in the web based demonstration. Without the ORACLE support, the CD-ROM version is only able to indicate the general nature of the demonstration. It is not able to perform any of the actual queries implemented in the Internet version.

## **Chapter 1 References**

[1]: Needs Statement for Phase II Study for an Integrated Network of Data Sources

[2]: Shaw M. Turner, Guidelines for developing ITS Data Archiving Systems, Texas Transportation Institute, September 2001

[3]: Mott MacDonald, “Achieving UTMC Integration using a Common Database – User Handbook”, <http://www.utmc.gov.uk/utmc10/pdf/u10uh.pdf>. Access in June 2003

[4]: Archived Data User Service in the National ITS Architecture, <http://www.iteris.com/itsarch>. Access in July 2003

## **2 SUMMARY OF CLIENT INFORMATION**

This chapter summarizes the resources available for the development of a central data warehouse for ITS applications. It examines the current FDOT activities and data infrastructure and identifies the key data items that are available from a variety of sources. The function of the CDW will be to place the available data within an expanded version of the current infrastructure to make the information more usable to the ITS-community.

### **2.1 ITS-related Applications Developed in the FDOT Districts**

Intelligent transportation systems involve the integrated application of advanced information, electronics, communications, and other technologies to address surface transportation problems. In the last decade, the transportation industry has made great progress in the utilization of ITS tools to enhance the nation's transportation systems.

Within the Florida regional ITS architectures, every district has developed many ITS applications to enhance the management and control of traffic. The recently completed or on-going ITS-related projects [1] in the each district are summarized in Table 2-1.

Table 2-1: Completed or On-going FDOT ITS-related Projects

| District Name  | Project Title                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| District One   | <p>Incident Management Plan for the I-75 Corridor in Lee and Collier Counties</p> <p>Approach for ITS Regional Architecture, Sarasota and Manatee County</p> <p>Incident Management Plan, I-75 Corridor, Sarasota, Manatee and Charlotte Counties</p> <p>Incident Management Project Plan, I-75 Corridor, Sarasota, Manatee and Charlotte Counties</p> <p>Road Ranger Service Contract, Lee County</p> <p>Road Ranger Service Contract, Charlotte County/Sarasota County</p> <p>Road Ranger Service Contract, Polk County</p> <p>I-4 Portable ITS (PITS) Project:</p> <p>Sarasota Retiming Project</p> <p>Manatee Retiming Project</p> <p>Traffic Signal System Analysis</p> <p>Collier County/City of Naples Computerized Traffic Signal System, Phase I Construction</p> <p>Punta Gorda/Charlotte County Computerized Traffic Signal System, Construction of Design Group I</p> |
| District Two   | <p>Jacksonville Interstate Surveillance and Control System, Phase 2</p> <p>Road Ranger Service Patrol</p> <p>I-95 Surveillance</p> <p>Closed Loop System</p> <p>I-10 Reverse Lane Implementation Plan</p> <p>ITS Maintenance</p> <p>SunGuide ITS System Database</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| District Three | <p>I-10/I-110 ITS Architecture:</p> <p>ITS Plan for the Interstate System</p> <p>Okaloosa County ATMS System</p> <p>Bay County ITS Earmark</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| District Four  | <p>Advanced Incident Information System (AIIS)</p> <p>Broward County ATMS</p> <p>Interim Traffic Management System</p> <p>Freeway Video Monitoring System for Broward and Palm Beach Counties</p> <p>Broward County Regional ATIS</p> <p>I-95/595 Dynamic Message Sign System (DMSS)</p> <p>Broward County ICWW CMSS</p> <p>I-595/I-95 CMSS</p> <p>Broward County ATIS Operations</p> <p>Broward County Service Patrol</p> <p>Palm Beach County ATMS Groups 1 &amp; 2</p> <p>Boca Raton ATMS Master Plan</p> <p>Palm Beach County Regional ATIS</p> <p>Palm Beach County ATIS Operations</p> <p>Palm Beach County I-95 Dynamic Message Sign System</p> <p>Palm Beach County Service Patrol</p> <p>Broward County ITS Operations Facility</p> <p>Palm Beach County Traffic Operations Facility Master Plan</p> <p>Broward and Palm Beach Counties APTS Master Plans</p>            |

|                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| District Five          | I-4 Surveillance and Motorist Information System (SMIS) Phase 3<br>I-4 Surveillance and Motorist Information System (SMIS) Phase 4:<br>Volusia ITS Integration Project<br>I-4 SMIS Phase 5<br>Upgrade of Existing I-4 SMIS Infrastructure<br>Expansion of I-95 SMIS<br>Phase 3 - I-95<br>Phase 4 - I-95<br>Expansion of I-75<br>Regional Traffic Management Center<br>City of Orlando Regional Computerized Signal System<br>Seminole County ATMS Project<br>Orange County ATMS Project<br>Fiber Optic Interconnection to District Maintenance Offices                                                                                       |
| District Six           | I-95 SunGuide/Intelligent Corridor System<br>Advanced Traveler Information System (ATIS) for Miami-Dade,<br>Broward and Palm Beach Counties<br>ATIS Public Involvement<br>SunGuide/District VI ITS Work Program<br>ITS Deployment in the Florida Keys<br>I-95 Golden Glades Interchange Post Construction Evaluation<br>SunGuide Total Integration Phase<br>SunGuide Software System Technical Support, Enhancement, Upgrades and Training<br>ITS Management, Technical Operations, Maintenance & Public Involvement<br>Service Patrol (Road Ranger)<br>SmarTraveler - ATIS                                                                  |
| District Seven         | Intelligent Transportation Systems Study and Implementation Plan for District Seven<br>Interstates<br>Update of the Tampa Bay Regional Architecture<br>St. Petersburg I-275 Dynamic Message Sign System<br>US 19 Pasco County ATMS<br>Clearwater/Countywide ATMS Signal System Project<br>Hernando County Computerized Traffic Signal System<br>District Seven Regional ITS Architecture<br>Skyway Video Monitoring System Modifications<br>Urban Area Traveler Information and Traffic Management System for I-275 in<br>Tampa<br>ITS Infrastructure for Freeway Management System in the Downtown<br>Interchange of I-4 and I-275 in Tampa |
| Turnpike<br>Enterprise | Electronic Toll Collection System (Sunpass) Interoperability<br>Turnpike ATIS (Prime Co)<br>SunGuide ATIS<br>Microwave Tower Grounding System Upgrade<br>Turkey Lake Communication System                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| ITS Office             | Statewide TMC Software Library System<br>Center to Center Communication<br>Statewide ITS Architecture and Standard                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |

Source: ITS Progress Report (December 2001), SUNGUIDE Florida Intelligent Transportation System website.

## **2.2 The FDOT Enterprise Information System**

FDOT has been developing a spatially enabled web-based Enterprise Information System (EIS) [2] to provide accurate, up-to-date information to technical staff and decision-makers for the entire department. The EIS uses a data-centric approach that takes advantage of the major developments in database technology including advances that make it possible to store, index and query spatial data using object-relational data base management system (DBMS) technology. It is an Intranet/Internet information system with GIS components. The Oracle 8.X DBMS provides the data storage infrastructure.

The EIS will integrate the FDOT's business data with several other identified data themes to satisfy the existing and future information needs of the entire Department.

It is intended that the data content will be created from the "best" data available. These data are often collected by local governments, utilities, regional and field offices of State and Federal agencies [2]. The FDOT EIS is not intended to eliminate or replace any existing applications and no new data are created.

## **2.3 Established Florida Statewide ITS Architecture**

The U.S. Department of Transportation (USDOT) has invested in substantial research and testing of new ITS technologies. The National ITS architecture developed by the USDOT provides a common framework to guide practitioners in developing regional ITS architectures. Florida ITS architecture, as a subset of the National ITS Architecture, has been developed to provide the framework to implement Florida statewide ITS applications.

The Florida Intelligent Transportation System architecture was developed a team led by Kimley-Horn and Associates for the FDOT . The scope of Florida ITS architecture provides the

regional architectures to support the ITS infrastructure for all FDOT purposes [3]. The statewide ITS architecture focuses on inter-urban and rural applications, but also adds value to urban areas. The overall scheme is based on FDOT's seven districts and the Turnpike Enterprise as shown in Figure 2-1.

In 2000, an intensive three-day workshop was conducted in each district with regional stakeholders representing all aspects of ITS in the district. At the end of the workshops, the district and regional ITS architectures were drafted, reviewed and agreed to by the stakeholders. After completing this process for each of the eight workshops covering all Districts, the results were used to define the Florida Statewide ITS Architecture. The following items are included in each FDOT District's ITS Architecture:

- Architecture interconnect ("Sausage") diagram
- Customized market packages
- Inventory by architecture entity or by stakeholder
- Standards Support for Architecture Flow
- Documentation of the structure and data flow in Turbo Architecture format

Four primary goals for the Florida transportation system are specified in the 2020 Florida Transportation Plan [4].

- Safe transportation
- Protection of transportation investment
- Statewide interconnection of the transportation system
- Provision of travel choice

These goals are used to identify the most appropriate type of ITS applications in Florida Statewide ITS Architecture.

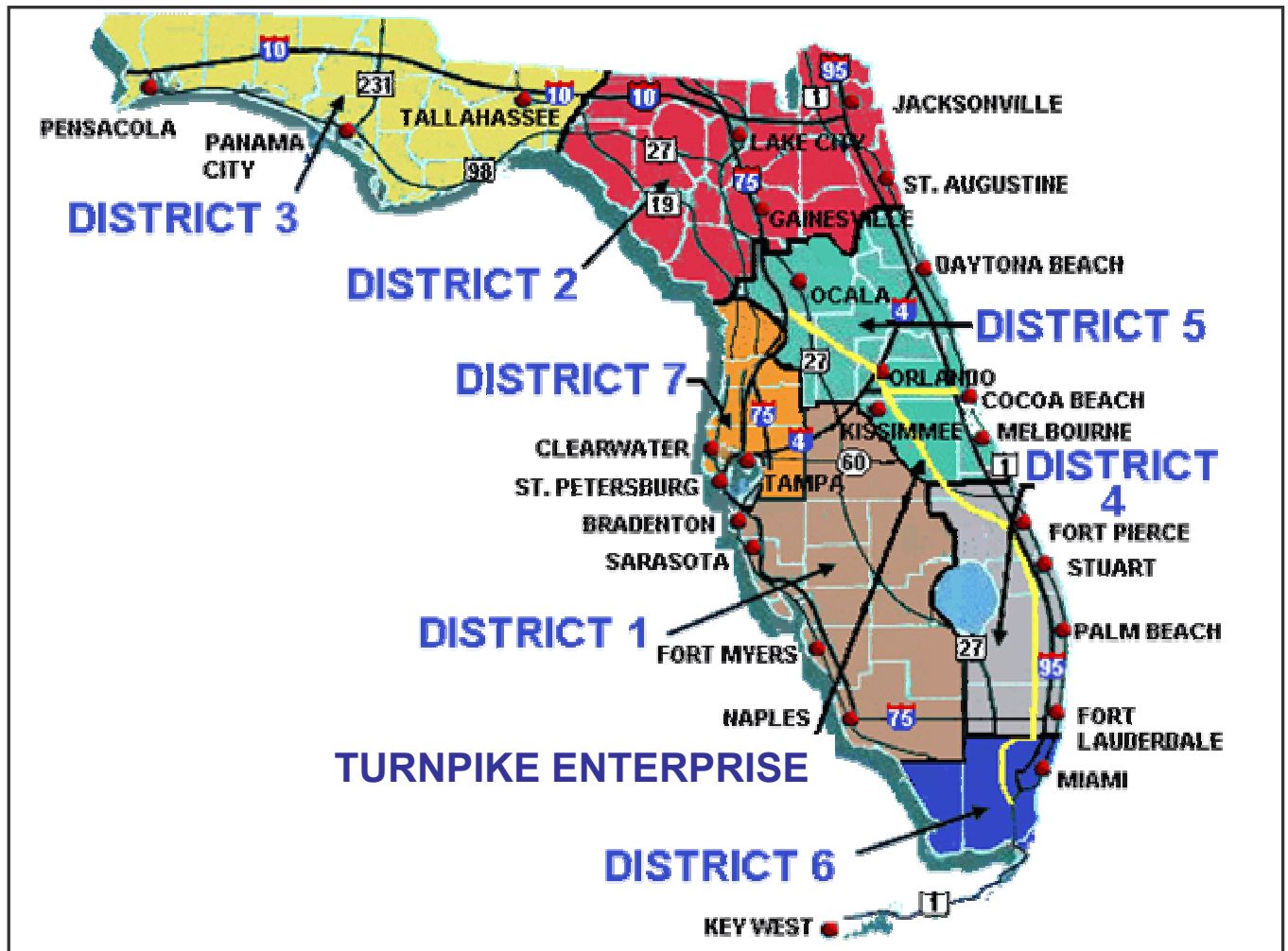


Figure 2-1: Distribution of FDOT Districts

## 2.4 Available Data at the Traffic Management Centers and Other Centers

Based on the results of a survey performed by the University of Central Florida, [5] the current ITS data collected and archived by each of Florida's seven districts and Turnpike Enterprise are summarized in Table 2-2. The various transportation data are collected by traffic centers and there are no standards describing the definitions or storage formats. The collected data are stored in flat file format or in a commercial DBMS. Even the data collected by the same traffic center is not integrated in many cases.

The information presented in Table2-2 also shows the different technologies utilized to collect traffic information, and indicates that there is a significant amount of statewide ITS data currently collected on the major highways.

In most districts, these records of collected data are stored in a commercial DBMS. The primary use of the DBMS is to retain records of collected data rather than to integrate different data sources and provide data warehousing capabilities. Different archiving strategies and data definitions are applied among the applications [5].

## **2.5 Data Available from GRIP [6]**

The Geo-referenced Information Portal (GRIP) system is the result of a major project undertaken by FDOT for the EIS infrastructure. The GRIP system focuses on the integration, dissemination, and leveraging of existing technologies and infrastructures. The stated goals of the GRIP system are to manage the digital assets using a data-centric approach that will promote a single enterprise data resource to encourage data sharing across functional and organizational boundaries, to increase confidence in the Department's digital geospatial data by maintaining data consistency and integrity, to handle temporal issues by providing all users with immediate and easy access to up-to-data information, and to eliminate redundant collection and storage of information.

Table 2-2: Application/Data-Collection Survey

| User Service Category  | Application                         | Detector Type                           | Collection Data                                                         | Collection Interval | Application Place                      |
|------------------------|-------------------------------------|-----------------------------------------|-------------------------------------------------------------------------|---------------------|----------------------------------------|
| Traffic Management     | Real-time travel information system | Inductive loop detector                 | Volume, occupancy, speed                                                | 20~30 seconds       | I-4 , part of I-95                     |
|                        |                                     | Remote Traffic Microwave Systems        | Volume, occupancy, speed                                                | At least 10 seconds | I-95                                   |
|                        |                                     | Video image detectors                   | Volume occupancy, speed                                                 | Real-time           | I-95 and I-10                          |
| Traffic Management     | Incident management                 | Inductive loop detector                 | Volume, occupancy, speed                                                | 20~30 seconds       | I-4 and part of I-95                   |
|                        |                                     | CCTV Camera                             | Image                                                                   | Real-time           | Some districts                         |
| Traffic Management     | Vehicle classification              | Inductive loop detector                 | Volume, occupancy, speed                                                | 20~30 seconds       | I-4 and part of I-95                   |
| Traffic Management     | Travel time study                   | Automatic vehicle identification reader | Travel time information, speed data                                     | Real-time           | On the toll roads of OCEA and Turnpike |
| Traveler information   | Travel information service          | Manual collection                       | Road information and lane closure data                                  |                     | All districts                          |
|                        |                                     | Closed Circuit Television Camera        | Image                                                                   | Real-time           | Some districts                         |
| Traveler information   | Dynamic Message Display             |                                         | Traffic monitoring data                                                 | Not automatically   | All districts                          |
| Information management | Archived data management system     |                                         | Dynamic message sign, incident data, road maintenance data, video image |                     | Most districts                         |

From architecture point of view, GRIP consists of an integrated multi-layered GIS spatial database as well as attribute and image servers residing in each district. Access to each district repository is provided by agreements within and across district boundaries. The user connects to a server via the LAN with a web-browser based user interface. Multiple layers of functionality are implemented to distinguish the service levels for different users, such as privileged users and non-privileged users.

The current transportation data layer in the GRIP system includes the following sub-layers:

**BMI (Bridge Management Inventory (also named Pontis))**

This system maintains structure and condition information pertaining to Florida's bridge infrastructure including location, number, and attributes. This information is reported annually to the Federal Highway Administration and supports the Department's bridge inspection and maintenance programs.

**RCI (Roadway Characteristics Inventory)**

This system maintains an inventory of the characteristics for roadways in Florida. Information tracked by this system includes:

- highway routes,
- ramps,
- intersection and interchange locations,
- inventory items on and adjacent to roadways and
- maintenance responsibility.

The RCI provides a foundation for integrating all roadways related data based upon the county/section/relative point on the roadway identifier and it brings the roadway inventory up-to-date.

**Roadway Inventory File**

This system is presently the responsibility of the Roadway Inventory Unit in the Bureau of Planning. The Districts perform the data gathering efforts and send the information to the central office for input and processing. This system provides the information at the state and county levels. The state level keeps track of the information related to the entire state, such as the highway routes, transportation subsystems, and cities. The county level keeps track of information about each county, such as roadway sections. Its design allows for changes in the section numbers and stationing without any re-entry of existing data.

### **CAR (Crash Analysis Reporting)**

This system retrieves the safety data from the Florida Department of Highway Safety Database. The collected information is submitted to the district office and central office of the Florida Department of Transportation.

### **RHC (Railroad-Highway Crossings)**

This system contains detailed information concerning all types of railroad crossings in the state, set up in a client-server architecture. It tracks all Florida railroad crossings. The data collected are used to update the corresponding data in the USDOT crossing inventory system. The RHC is used to calculate a safety index, which ranks the crossings to prioritize their need for improvement.

In addition, the Systems Planning Office collects the following data to meet the data needs for planning and modeling.

#### **2.5.1 Data Collection Sites and Devices**

There are now approximately 300 permanent count stations (PCSs) for planning. Data are polled daily from these stations, generally around 2-5 AM (but can be polled anytime, especially in emergencies). Other than the PCSs, there are also portable sites, that bring the total number of count stations to more than 8000. Approximately 2000 of the stations have vehicle classification capability.

#### **2.5.2 Data Collection and Data Products**

FDOT collects hourly traffic volume per direction, vehicle classification information and average speed. The Planning Office publishes a data CD every year that contains the final data products, which can produce the following eight reports for any traffic count site:

- Annual Average Daily Traffic Report
- Historical AADT Report
- 200 Highest Hour Report

- Hourly Continuous Count Report
- AADT Forecast Report
- Annual Vehicle Classification Report
- Weekly Axle Factor Category Report
- Peak Season Factor Category Report
- Volume Factor Category Summary Report

### **2.5.3 Data collection quality control**

To control the quality of the data, the newly collected data are compared against historical data for the same day of week, month, and year, and the analyst decides if the data are of acceptable quality. FDOT follows the AASHTO guidelines for count data quality control [7]:

- At least 1 week of data every month;
- The days don't need to be consecutive but must cover each day of the week (FDOT desires contiguous counts), and
- Each day must have a complete, uninterrupted 24 hours of data.

### **2.5.4 Data storage**

The data are stored using a *defacto* standard developed by FDOT. The data CD contains a GIS interface to access traffic data collected from traffic count sites on the State Highway System. To view traffic characteristics, the user can visually locate a traffic site using the interactive map of Florida. For each traffic count site selected, information about the particular site can be instantly displayed and eight reports (as mentioned in Section 2.6.2) can either be printed or exported to a spreadsheet format.

## **2.6 Emergency Management Center Information**

The Emergency Management Center (EMC) also collects and stores data for their own applications. The EMC claims to have the most complete state-wide data set, which includes data from DOT, DEP, USGS, etc. The amount of data currently stored exceeds 320GB and all data reside within the EMC. However, the data are all stored in a file system (not in database), so no direct query is available. The EMC supports outside requests for data by writing specific code to extract desired data.

## **2.7 Other Data**

Additional data will likely be needed to be archived in the CDW data repository to support specific applications. Such data could include weather, regional population and regional business activity. The data could be collected/provided either by the FDOT district agencies or the other public and private agencies.

## **2.8 Future Research Tasks**

This chapter has summarized the resources available for the development of a central data warehouse for ITS applications in very broad terms. The next step in the CDW development will be to examine the data items in more detail with respect to their definitions and formats and to construct the necessary infrastructure to move these data from place to place with the creation of the required communications, storage, retrieval, and presentation functions.

## **Chapter 2 References**

- [1]: ITS Progression Report,  
[http://www.floridait.com/Newsletters/12-01\\_newsletter/newsletter\\_p4.htm#Progress](http://www.floridait.com/Newsletters/12-01_newsletter/newsletter_p4.htm#Progress), Access in March 2003.
- [2]: FDOT Enterprise Information System  
<http://www.bts.gov/gis/reference/gist/2001/61/613/613.pdf> Access in March 2003
- [3]: “Florida Statewide Intelligent Transportation System (ITS) Architecture & Standards”,  
<http://www.jeng.com/html/florida/default.htm>, Access on March 10, 2003
- [4]: “Florida Intelligent Transportation System Strategic Plan Final Report”, August 23, 1999.  
<http://www11.myflorida.com/planning/systems/sm/its/PDFs/ITS%20Strategic%20Plan%20FINAL%2010Mar00.PDF>, Access in March 2003.
- [5]: Haitham Al-Deek, Amr Abd-Elrahman, “An Evaluation Plan For The Conceptual Design Of The Florida Transportation Data Warehouse Final Report”, University of Central Florida, January 2001.
- [6]: John Van Vliet, Jared Lish, GRIP System FDOT, Xmarc Inc April 11 2001,  
<http://www.bts.gov/gis/reference/gist/2001/61/613/613.pdf>, Access in March 2003.
- [7]: AASHTO Guidelines for Traffic Data Programs, American Association of State Highway and Transportation Officials, Washington, D.C. 1992, ISBN 1-56051-054-4.

## **3 CDW DESIGN CONCEPTS AND STANDARDS**

This chapter presents a discussion of the common concepts and standards that apply to the design of a CDW. The concepts are taken from general industry literature. Several standards must also be considered in the development of any ITS component. Some are general industry-wide standards, while others have been developed specifically with ITS functionality in mind.

### **3.1 Information Flow Concepts**

In terms of its architecture, the functionality of a CDW is similar to that of Archived Data Management. The CDW helps to achieve the ITS information goals of unambiguous interchange and reuse of data and information throughout all functional areas. It facilitates the efficient use of ITS data for real time traffic management, planning, developing new traffic management strategies as well as for myriads of transportation analysis. It collects, archives, manages and distributes data from ITS and other sources. It connects other transportation layer subsystems including travelers, vehicles, transportation management centers and field devices mainly through wire line communication or wide area wireless communication. Figure 3-1 illustrates a typical top-level architecture interconnect sausage diagram.

Figure 3-2 displays the interconnect diagram between a typical CDW and other transportation operation centers. The CDW supports all of the major themes of the ITS system. Corridor-wide implementation is desired with real-time and archived data available to support various user services including operations, management, planning and other stakeholder needs. Each operating agency does own data processing and shares data with the CDW.

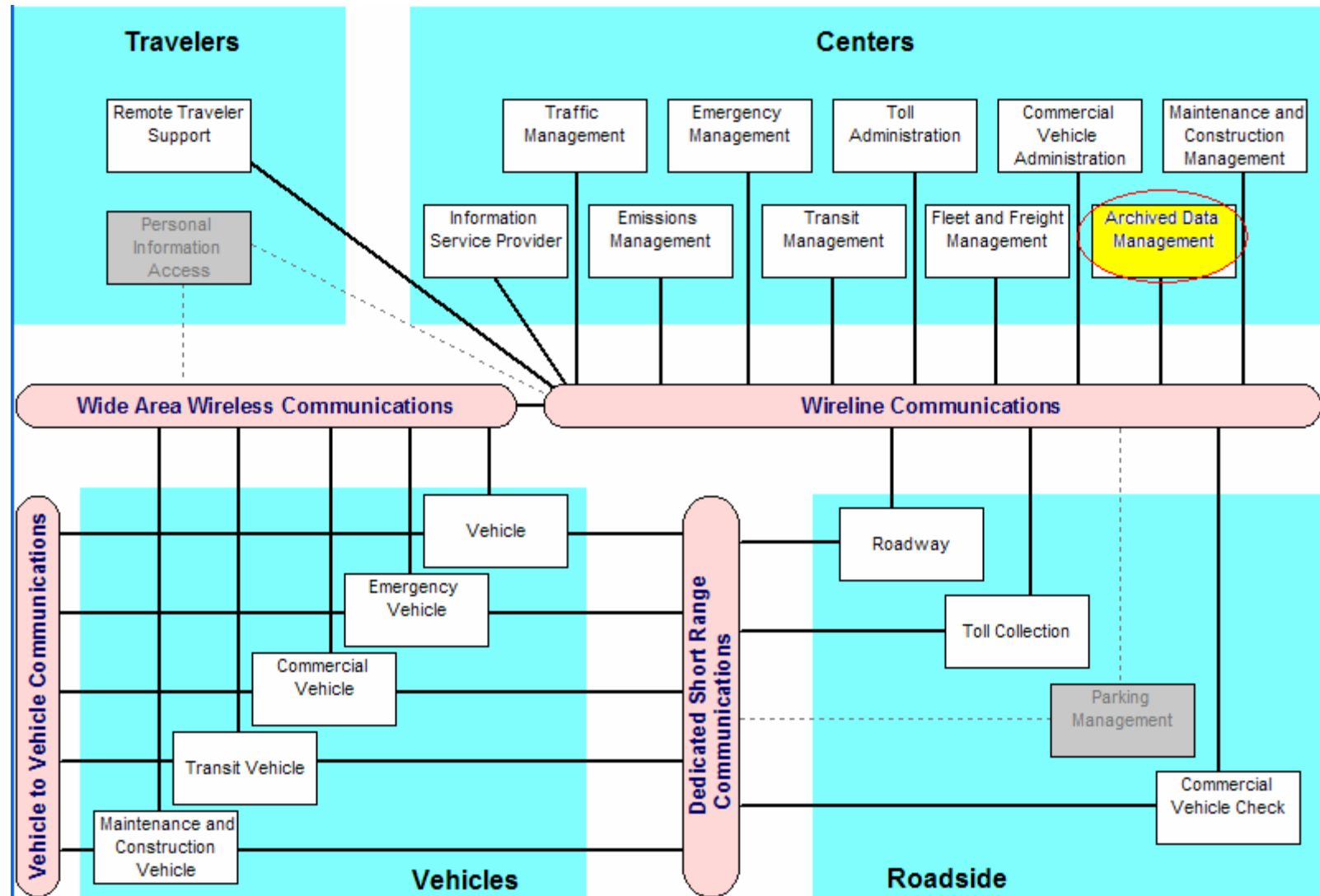
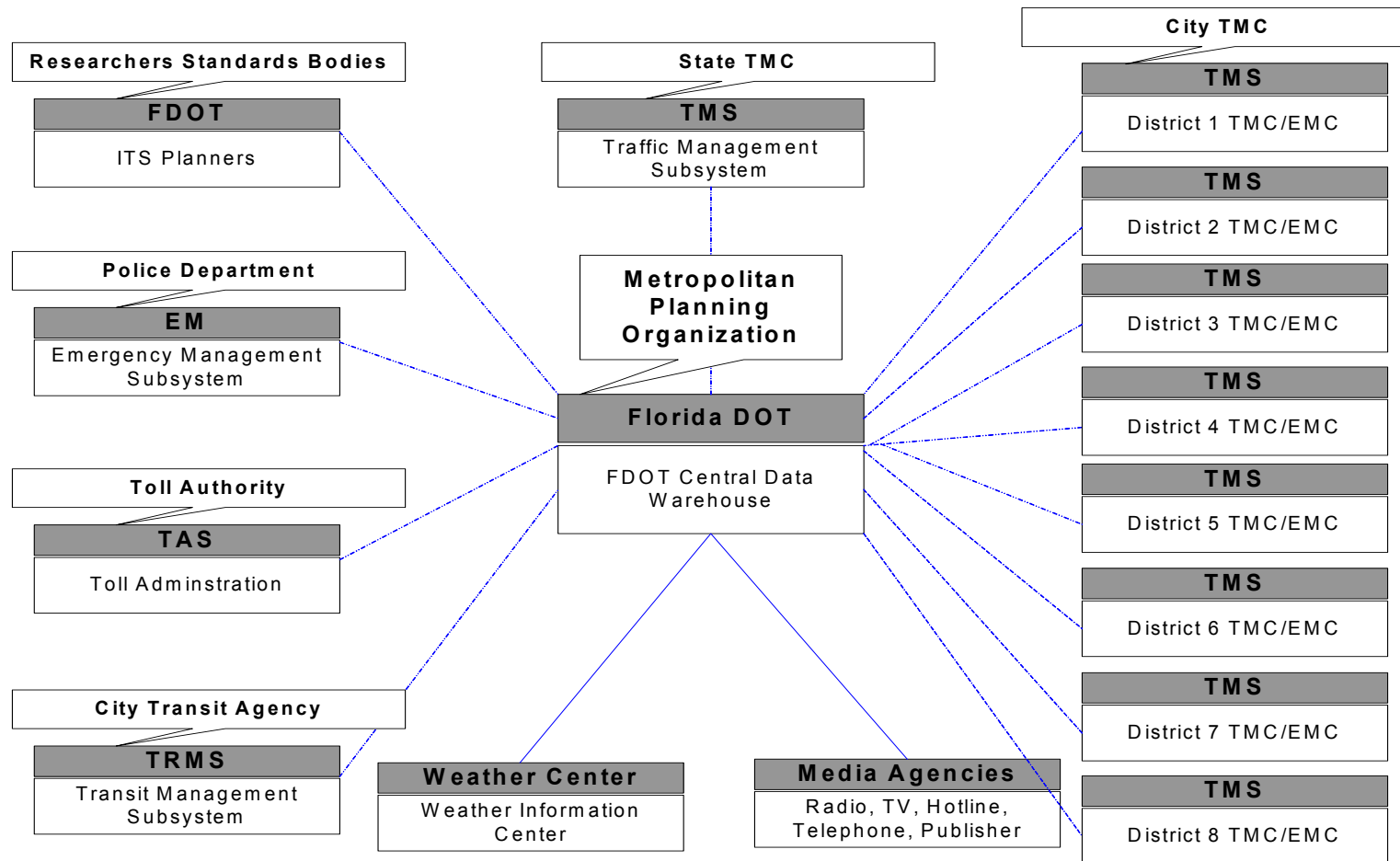


Figure 3-1: Architecture Interconnect Sausage Diagram.



**Figure 3-2: CDW Interconnect Diagram**

New sets of useful information may be drawn into the CDW for distribution to other agencies. Some information centers outside of the ITS core, such as weather centers, and the media can share information from the CDW and provide useful information for their respective purposes.

A CDW system must support the integration and merging of schema and data from the underlying repositories. This can be achieved in one of two ways:

1. By building an information mediation system that does not contain any data itself but can understand and process queries against the global schema. In response to an incoming query, the mediator generates one or more new queries against the standalone databases that together fetch the information necessary to satisfy the original query. The mediator integrates and merges the individual answers and returns the final result to the application that submitted the original request. The approach is also known as the “just-in-time” or “virtual warehouse” approach since the mediator gives its users the illusion of managing an integrated data warehouse.
2. By building a real data warehouse that contains a copy of the data in the distributed databases. As in the case of the mediator, the schema of the warehouse contains all of the important and useful concepts of the underlying sources in an integrated fashion. However, unlike the mediator, the warehouse contains the integrated and cleansed data that are ready for querying.

### **3.2 General Industry Standards**

The following industry-wide standards apply to all systems, including the CDW to be developed in Florida:

#### **3.2.1 Common Object Request Broker Architecture (CORBA) [1]**

As more and more distributed systems are implemented, CORBA becomes the suitable communication infrastructure because it is able to provide a truly open, vendor-independent architecture for distributed object computing. CORBA has been standardized by the Object Management Group (OMG). Using the standards of Internet ORB Protocol (IIOP), a CORBA based program can operate with any other CORBA-based program residing on a computer in the network, independent of computer, platform, programming language and network configuration.

It provides interoperability between objects in a heterogeneous distributed environment and in a way that is transparent to the programmer.

CORBA is based on the OMG Object Model, which defines common object semantics for specifying the externally visible characteristics of objects in a standard and implementation-independent manner. In this model, clients request services from servers through a strong-type-defined interface, which is specified using the Interface Definition Language (IDL). When a client intends to access an object, it issues a request to the object. The request is composed of an object reference of the service provider and the required parameters associated with the services. The central focus of CORBA is the Object Request Broker (ORB), also known as Object Reference Model. It encompasses the entire communication infrastructure required for identifying and locating objects, managing connection, and delivering data.

CORBA provides a full-scale architecture for an object-oriented distributed system. Object-oriented software implemented in traffic management centers can take full advantage of CORBA technology that traditional software can't offer. However, due to its complexity, CORBA requires significant resources for implementations in the areas of programmer knowledge, hardware requirements and funds. In short, CORBA is a good candidate for high-end system with plenty of resources.

CORBA not only enables systems to retrieve data from other systems, but also to remotely activate methods embedded in objects. The network connection is peer-to-peer.

CORBA enables the central server calling object implementations to provide the execution of requested services to retrieve data from other systems. At the same time, CORBA has the ability to remotely activate methods embedded in client objects. Within the CORBA

communication infrastructure, the network connection between server and client object(s) is peer-to-peer.

### **3.2.2 Data Exchange Using Abstract Syntax Notation (DATEX-ASN) [2]**

DATEX offers a logical data model with an extensive data dictionary. DATEX combines the “data object” and the “phrase” approach. The data object carries the major attributes of the particular object while a distinguished mandatory attribute called "phrase" specifies the state of the data object. The DATEX data dictionary contains a substantial amount of work on collecting pieces of information from the transportation telematics field.

DATEX-ASN is based on the client-server architecture. Data elements are exchanged between a client and a server. Any traffic center requesting information from other centers is defined as the client. Any traffic center that receives requests to provide information to other centers is defined as the server. In a CDW application, the CDW will become the client when it requests and retrieves traffic data from other traffic centers.

For datagram publication, publication data is embedded into a publication data packet and transmitted to the client. If the resulting datagram is over the pre-defined maximum datagram size, the publication data will be stored in a file and the publication data packet will indicate the path and filename of the file that stores the publication data. When the client receives the publication data packet, it can activate a file transfer mechanism, such as FTP (File Transfer Protocol), to receive the publication data from the server. This can be viewed as a file-transferring mechanism for a CDW source data extraction operation.

DATEX-ASN provides a simple, general-purpose center to center data exchange stack for basic data exchange needs. It is a cost-effective solution for systems that request fast data transfer and high data volumes, but with limited bandwidth. DATEX-ASN is not designed for

and cannot take advantages of high object-oriented systems. The network connection can be peer-to-peer or dial-up.

### **3.2.3 Extensible Markup Language (XML)**

XML has proven to be a universal structured data transfer methodology that facilitates e-Business and e-Government. Originally designed to meet the challenges of large-scale electronic publishing, XML is also playing an increasingly important role in the exchange of a wide variety of data on the web and elsewhere. XML data structures, known as schemas, provide a mechanism to develop and adopt common formats for data exchange.

A substantial interest in XML as a data transfer mechanism has emerged within the transportation profession. Schemas have already been developed for representing data structures for traffic modeling, crash records, highway infrastructures and geospatial features. A new project has been initiated by the Transportation Research Board (NCHRP 20-64) “to develop broadly accepted public domain XML schemas for exchange of transportation data and to develop a framework for development, validation, dissemination, and extension of current and future schemas.”

### **3.2.4 Other Protocols**

There are other communication protocols suitable for a CDW application, such as, Simple Network Management Protocol (SNMP), Simple Transportation Management Protocol (STMP}, File Transfer Protocol (FTP), Trivial File Transfer Protocol (TFTP). With the exception of STMP, all application protocols are widely accepted industrial standard application layer protocols.

### **3.3 Standards Developed for ITS Applications**

In addition to the general industry standards, several ITS-specific standards have been developed to guide the design and deployment of ITS applications. The principal standards development organizations (SDOs) that are active in the development of ITS-specific standards include:

- The Institute of Transportation Engineers (ITE),
- The American Association of State Highway and Transportation Officials (AASHTO),
- The National Electrical Manufacturers Association (NEMA),
- The American Society for Testing Machinery ASTM,
- The American Public Transit Association (APTA),
- The Transit Standards Consortium (TSC) and
- The Society of Automotive Engineers (SAE)
- The Institute of Electrical and Electronic Engineers (IEEE)

#### **3.3.1 Archived Data Management Services (ADMS) [3]**

A guide developed by ASTM, covers desired approaches to be considered and followed in planning, developing, and operating specific ADMS for the archiving and retrieval of ITS-generated data. The scope of the guide anticipates incremental or modular implementation of an ADMS, which over time and with a series of investment of resources will approach or exceed desired practice. However, it is recognized that programmatic constraints of time and budget resources do not always allow practitioners to follow a more desirable course of action and that during interim periods the ability to implement a particular fully functioning system may be less than desired.

The ADMS subsystem gathers information about surface transportation for system analysis. There are many entities in the National ITS Architecture that can provide data to the ADMS. Each interface to a segment of transportation data operates in a similar fashion. The ADMS can request a catalog of the transportation data and/or the data itself from each data source. Data definition in any data warehouse implementation should follow some standards.

The ADMS Standard Guide sets forth eleven guiding principles should that be adhered to in the development of a CDW:

1. Reliance on user needs and requirements processes
2. Get archived data from other centers
3. Integrate selected other transportation data including roadside data collection
4. Manage the archive to account for data quality
5. Provide security for the ADMS
6. Specify and maintain metadata to support the ADMS
7. Manage the Interfaces of the Archive Data Administrator
8. Interact with other archives and monitor other standards
9. Process user requests for data
10. Support analysis of the archived data
11. Prepare data for government reporting systems

Each of these principles is discussed in more detail in Reference 3.

### **3.3.2 Traffic Management Data Dictionary (TMDD)[4]**

This standard was developed by ITE and AASHTO for ITS projects that manage traffic. It is being developed under the oversight of a national steering committee with representation from both organizations. The primary message set for the TMDD, is the companion standard described in the next subsection.

This standard provides a functional level data dictionary consisting of and defining a set of data elements necessary to support data flows within and among traffic management systems. Specifically, as a data dictionary standard, it provides meta attributes for each data element, including definitions (semantics) and specific format (syntax) for individual data elements. The TMDD, as a national functional level data dictionary, provides a standardized national set of data elements that are intended to be the basis of individual application-level data dictionaries implemented at specific sites.

At present four sections have been developed into the following partitions:

- Traffic Links and Nodes
- Events, Incidents and Notification Alarms

- Traffic Network, Traffic Signal Control, Traffic Detectors, Vehicle Probes, Ramp Metering, and Traffic Modeling
- Closed Circuit Television, Dynamic Message Signs, Environmental Sensor Stations, Highway Advisory Radio, and Parking

### **3.3.3 Message Sets for External Traffic Management Center Communications (MS/ETMCC) [5]**

This standard, which was developed by ITE and AASHTO, includes message sets that were developed specifically for ITS traffic management systems. It consists of nineteen message sets organized into six message groups. It was developed under the oversight of a national steering committee composed of representatives of both ITE and AASHTO and is being published as a joint standard.

This standard provides formal message sets for six message groups necessary to convey key data within and between traffic management centers and other ITS centers. As a message set standard, it provides a list of specific data elements for each message plus other format information such as message name, message number, and certain other mandatory and optional message attributes. The MS/ETMCC is designed to be independent of any specific communications protocol. The MS/ETMCC, as a national ITS standard, provides a set of messages intended to be the core of individual messages implemented at specific sites.

The MS/ETMCC standard contains the following message groups:

- Roadway-Network,
- Network-State,
- Network-Events,
- Traffic-Requests,
- Traffic-Device-Status,
- Traffic-Control.

### **3.3.4 IEEE 1512 Standard for Common Incident Management Message Sets for Use by Emergency Management Centers**

This standard addresses the exchange of vital data about transportation-related incidents among emergency management centers through common incident management message sets. Message sets specified are consistent with the National Intelligent Transportation Systems Architecture and are described using Abstract Syntax Notation One syntax. This standard provides the base standard of a family of incident management standards; specific incident management message sets for traffic, public safety, and HAZMAT centers may be found in forthcoming companion volumes which build upon and augment this base standard [6].

The 1512 standards family supports the operational functions of basic incident information exchange, situational awareness, asset management, plan dissemination and warning. These functions can be combined in many ways, depending on the operational needs of an evolving incident. The basic standard handles a wide variety of incident related data, including basic incident details (type, location, time, impact, estimated duration, etc.), incident detail updates, merging of incidents, splitting of incidents, deletion of incidents, etc. Sufficient flexibility was built into this standard to handle virtually any multi-agency incident management protocol. In some cases the use of these message sets causes agencies to critically evaluate how they accomplish multi-agency incident management [7].

### **3.3.5 NTCIP**

NTCIP stands for “National Transportation Communications for ITS Protocol.” It is a family of communication standards/data definition standards based on Open Systems Interconnection (OSI) seven-layer structure. NTCIP was developed by three organizations, NEMA (National Electrical Manufacturers Association), ITE (Institute of Transportation

Engineers) and AASHTO (American Association of State Highway and Transportation Officials). Among them, NEMA is the lead agency and NTCIP standards are disseminated through them. [8]

NTCIP supports data communication for both center-to-field and center-to-center applications. There are six application layer protocols proposed and standardized by NTCIP:

- CORBA---Common Object Request Broker Architecture [1]
- DATEX-ASN---Data Exchange in Abstract Syntax Notation One [2]
- SNMP---Simple Network Management Protocol
- STMP---Simple Transportation Management Protocol
- FTP---File Transfer Protocol
- TFTP---Trivial File Transfer Protocol

Among them, DATEX-ASN and CORBA are the protocols applicable in center-to-center communications for CDW application. SNMP and STMP are protocols appropriate for center-to-field communications. FTP and TFTP could be used to retrieve files stored on a server when the client receives a DATEX-ASN file publication message or any file transfer requests. NTCIP adopts a number of existing computer industry standards for the sub-network layers, such as TCP/IP or UDP.

### **3.3.6 Center to Center (C2C) Communications**

Under the NTCIP standard, the center-to-center communications are mainly for four service requirements, including traffic coordination, event notification, data sharing and regional command discretion. For a CDW application, the transportation data will be transferred from local traffic sources to the CDW remotely. The supported data sharing functionality for center-to-center communications is applicable for a CDW requirement.

The shared data may include historical data, real-time data and static data. The historical data are the legacy data stored in the information system of the district traffic centers. The real-time data can be detector data or CCTV video images. The static data can be event and detector historical information.

The communication message is constructed based on data definition and message sets defined in ITS Data Dictionary and ITS Message Sets (TMDD, MS/ETMCC). These messages are wrapped using the DATEX standard. The network layer can be a TCP/IP or UDP and PPP standard. Normally, DATEX is applicable for low-end systems. For the high-end systems, CORBA is generally used. FTP, a computer industrial communication standard, could be also used for file transferring between the traffic center and ITS CDW.

### **3.3.7 ITS Registry**

The ITS Data Registry is mainly operated by the Institute of Electrical and Electronics Engineers (IEEE). It is a centralized data dictionary or repository for all ITS data elements and other data concepts that have been formally specified and established. The registry is intended to serve as a common or shared data reference for the national ITS domain. The primary objective of the registry is to support the unambiguous interchange and reuse of data and data concepts among functional-areas of ITS (i.e., among associated ITS subsystems and their application systems) by recording unambiguous definitions of data concepts. As new data dictionaries are developed, they are entered in the registry. If similar data elements or concepts from a submitting data dictionary already exist, any differences need to be reconciled via a consensus process. By this process, the ITS Data Registry ensures a reliable and consistent set of ITS data elements and message sets.

### **3.4 Future Research Tasks**

This chapter has presented a discussion of the common concepts and standards that apply to the design of a CDW. The main effort in the next phase of the CDW development will focus on the design aspects. Future research on concepts and standards should be limited to an update of the material presented in this chapter to accommodate changes to the standards described in this chapter that could affect ITS deployment. The results of ongoing FDOT efforts, such as C2C communications, should be monitored and incorporated into the CDW design.

#### **Chapter 3 References:**

- [1]: CORBA Definition. <http://www.puder.org/corba/matrix/>, Access in March 2003
- [2]: DATEX Definition <http://www.vdot-itsarch.com/nova/docs/comm.pdf> Access in March 2003
- [3]: Document E2259-03: “Standard Guide for Archiving and Retrieving ITS-Generated Data” ASTM, 2003
- [4]: “ITE-AASHTO TM1.03—Standards for Functional Level Traffic Management Data Dictionary (TMDD)”, Intelligent Transportation systems  
[http://www.its-standards.net/Documents/FS\\_TM1.01-1.pdf](http://www.its-standards.net/Documents/FS_TM1.01-1.pdf), Access in August 2003.
- [5]: “ITE-AASHTO TM2.01—Message Sets for External Traffic Management Center Communication”, Intelligent Transportation systems  
[http://www.its-standards.net/Documents/FS\\_TM2.pdf](http://www.its-standards.net/Documents/FS_TM2.pdf), Access in August 2003.
- [6]: IEEE Std 1512-2000 IEEE Standard for Common Incident Management Message Sets for Use by Emergency Management Centers  
[http://standards.ieee.org/reading/ieee/std\\_public/description/intellitrans/1512-2000\\_desc.html](http://standards.ieee.org/reading/ieee/std_public/description/intellitrans/1512-2000_desc.html), access in December 2003
- [7]: IEEE 1512 Incident Management Standards Frequently Asked Questions (FAQ),  
<http://www.nawgits.com/ieee/1512faq.html>, access in December 2003.
- [8]: NTCIP Standard. [www.nema.org](http://www.nema.org). Access in March 2003

## **4 PRELIMINARY DESIGN RECOMMENDATIONS**

This chapter presents the observations and recommendations that can be offered at this time, given the relatively coarse level of the investigations that have been done to this point. These recommendations define the nature of the system that should be pursued in the next phase of the study, during which substantially more detail will be required.

### **4.1 High-level CDW Requirements**

The following observations are offered with respect to the high-level system requirements:

#### **4.1.1 Data Presentation Requirements**

It is important that the traffic data acquired by the CDW can be linked with geographic location information. Therefore, the GIS features should be fully integrated to the CDW structure to provide the location dimension information. GIS becomes a critical tool to be used to assemble, store, manipulate, analyze, and display information about relevant traffic data with spatial relations in the state of Florida.

Integrated GIS will provide a significant part of the CDW functionality in addition to its traditional role in providing the graphic interface. The GIS aids the decision making process of locating potential connection points for sharing of ITS data. Due to the very dynamic nature of ITS data, GIS can be used over time to help grow and make even more useful the initial inventory and analysis effort. FDOT has adopted ESRI's suite of core GIS products (ArcInfo and ArcView) as the basis for its GIS development efforts involving the statewide repository of data. [1]

The GIS role in the CDW should not only be viewed as a traditional display tool, but also to provide significant part of the data warehouse infrastructure and functionality. The GIS data

can be interpreted as location dimension in the meta-data of data warehousing data. This location dimension provides a common linkage to almost all the data subjects in the data warehouse. GIS by itself has the innate capacity to describe spatial relationships among the data elements. This capacity is hard to achieve via the traditional relational DBMS technology. Integrating the GIS with the CDW enables complex spatial queries. Furthermore, the integrated GIS feature will impact on emergency management situations, such as fire and flood evacuation planning. [2]

While GIS features are extremely important to the CDW operation, the development of an extensive dedicated GIS capability neither necessary nor desirable as a project task. Instead, the proposed GIS features of the FDOT EIS system should be relied upon for this purpose. The CDW should operate within the EIS, using the EIS capabilities wherever possible.

#### **4.1.2 Decision Support Requirements**

A decision support system (DSS) is an information system that analyzes the data created and updated by other systems. The outputs are presented to be used by human decision makers. The DSS component should be built to utilize the information contained in the CDW information repository. This function will be very important to the Enterprise Information System [3]. The decision support system details will vary according to the specific ITS application [4].

If the existing FDOT infrastructure has adequate capacity, this system can be implemented inside the existing computer system. Otherwise, the approach of a dedicated client/server model could be adopted such that the decision support system will be implemented as a new hardware box. The Internet could be used to provide decision support data to the users, thereby offering the advantages of universal accessibility and familiarity. [3]

The data required by the DSS must be consolidated into the CDW database. All of the data will generally be retrieved from other real-time data sources, i.e., the DSS will not access the data sources directly.

There are two basic ways to analyze the content of a data warehouse, including on-line analytical processing and data mining [5]. On-line analytical processing, or OLAP, requires the user to ask a series of questions to be answered using data from the warehouse database. Several OLAP products are available. The OLAP operates in a client/server mode and can access data from the warehouse. It will let users select subsets of the data for analysis and then provide a variety of tabular, and graphical representations of the selected data [5].

The second approach, data mining, will allow users to specify target objectives (i.e., what information is sought under which conditions. The system will then perform the analysis automatically. Data mining approaches include classification and regression trees (CART), correlation and other types of statistical analysis of the data, neural networks, and nearest neighbor approaches [5].

Since a DSS is an information system, the general principles of the system architecture will still be effective, including system interoperability, system compatibility and system expandability. The DSS architecture must include the following elements [5]:

- Database sub-system (if applicable)
- Data models
- DSS users
- Software tools
- Hardware and operating system
- Networking and communications

The major categories of specialized software used to assist DSS development are the following:

- Database management packages
- Information retrieval (query and reporting) packages
- Specialized modeling package (including spreadsheet) and languages
- Statistical data analysis packages
- Forecasting packages
- Graphing packages

Since DSS functions are intended to work closely with human decision makers in carrying out their tasks, a DSS can only be effective as its interface with those the human element. Several user interface styles should be considered, such as, Command-line Interface, Graphical User Interface, Question-and-answer Dialogues, Animation etc. [5]. The use of a commercial expert system shell, of which there are several choices, should be considered seriously in the next phase of the CDW development.

#### **4.1.3 Technology Requirements**

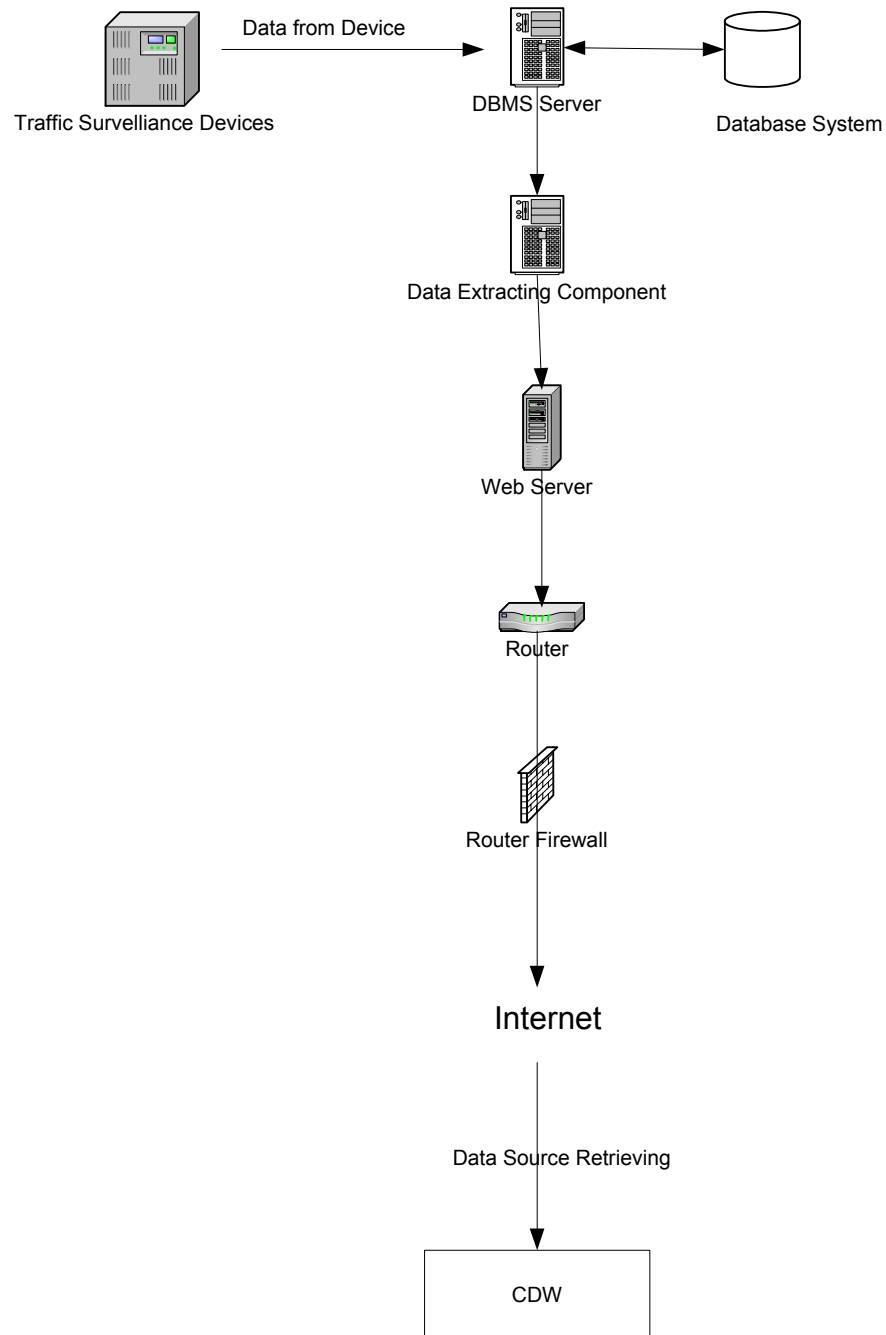
The technology for the hardware and software should be chosen such that it can be applied in the data warehouse environment with extensibility, flexibility and ease of management. It was pointed out in Chapter 3 that there are two alternative high level configurations for a data warehouse. These alternatives are generally identified by the terms “real data warehouse” and “virtual data warehouse.” While the virtual configuration appears to have a strong conceptual appeal in the industry, it is a generally viewed as a somewhat abstract concept at this point. Implementation problems, including the difficulty of defining the mediation schema, the incompatibility of the individual data sources and the availability of a general purpose commercial software framework, have limited the practicality of virtual data

warehouses. Therefore, it is recommended that the real data warehouse concept be pursued in the design phase of the Florida CDW.

Security measures should be considered by assigning the appropriate permission privileges to the users. Users with different security levels should access system by different methods. The users within FDOT can access system via the Tallahassee intranet without any limitations. The users outside FDOT office who can access the system through a virtual private network (VPN) service will have the same access level as those within Tallahassee office. The public user and outside customer may only gain access through a particular webpage. That means the services provided is limited to those defined in web server, which is behind and protected by firewall. The security methods can be applied in different layers, such as the network layer, the application layer and/or the database layer [6].

To transfer the stored data from traffic data resource centers to the CDW, a data-extraction component should be embedded into the traffic data resource centers. This component will behave as a data bridge between the data resource and the CDW system. Since commercial DBMS systems are used frequently to store data in traffic centers, a strategy of the DBMS data replication between the data resource and the CDW is applicable to extract the real time and historical data into CDW database. Figure 4-1 represents the data-extraction component architecture in a traffic data resource center.

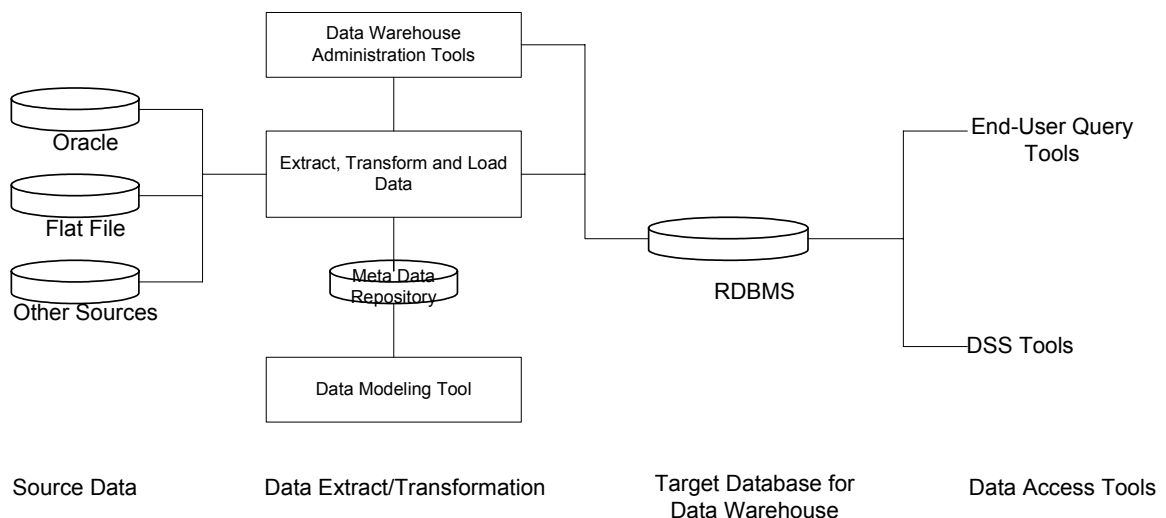
The DBMS in the traffic data resource center can also generate a flat file containing the refresh data. Then the flat file can be transferred to the CDW to replicate the database.



**Figure 4 - 1 Traffic Center Data Extracting Component Architecture**

## 4.2 CDW Multi-tier Architecture

It is expected that a multi-tier architecture will be developed for the CDW to take advantage of system portability and reusability [7]. Figure 4-2 presents a diagram for the multi-tier concept for CDW architecture. The source data will be contained in different information systems with various storage formats. The raw data retrieved from data resources will be stored in the central database system. A relational DBMS approach should be adopted. [8] The data organization within the CDW will be based on the service requirements of different users. The CDW will provide the required data service to the customer via the separate DSS.



**Figure 4 - 2 CDW Architecture Component**

The data will be retrieved from traffic data resource centers within each district. The raw data will be processed, cleaned, reorganized and stored in the CDW. Also, the source data in the GRIP system will be retrieved to the CDW if required. The decision maker/user will access the data via the DSS through a mechanism such as OLAP server. The application processing server with a light client model will be adopted. The users will be distinguished by different privilege

layers and each group will have different access modes. Figure 4-3 presents the data flow for this aspect of the CDW.

### **4.3 Data Stored In the CDW**

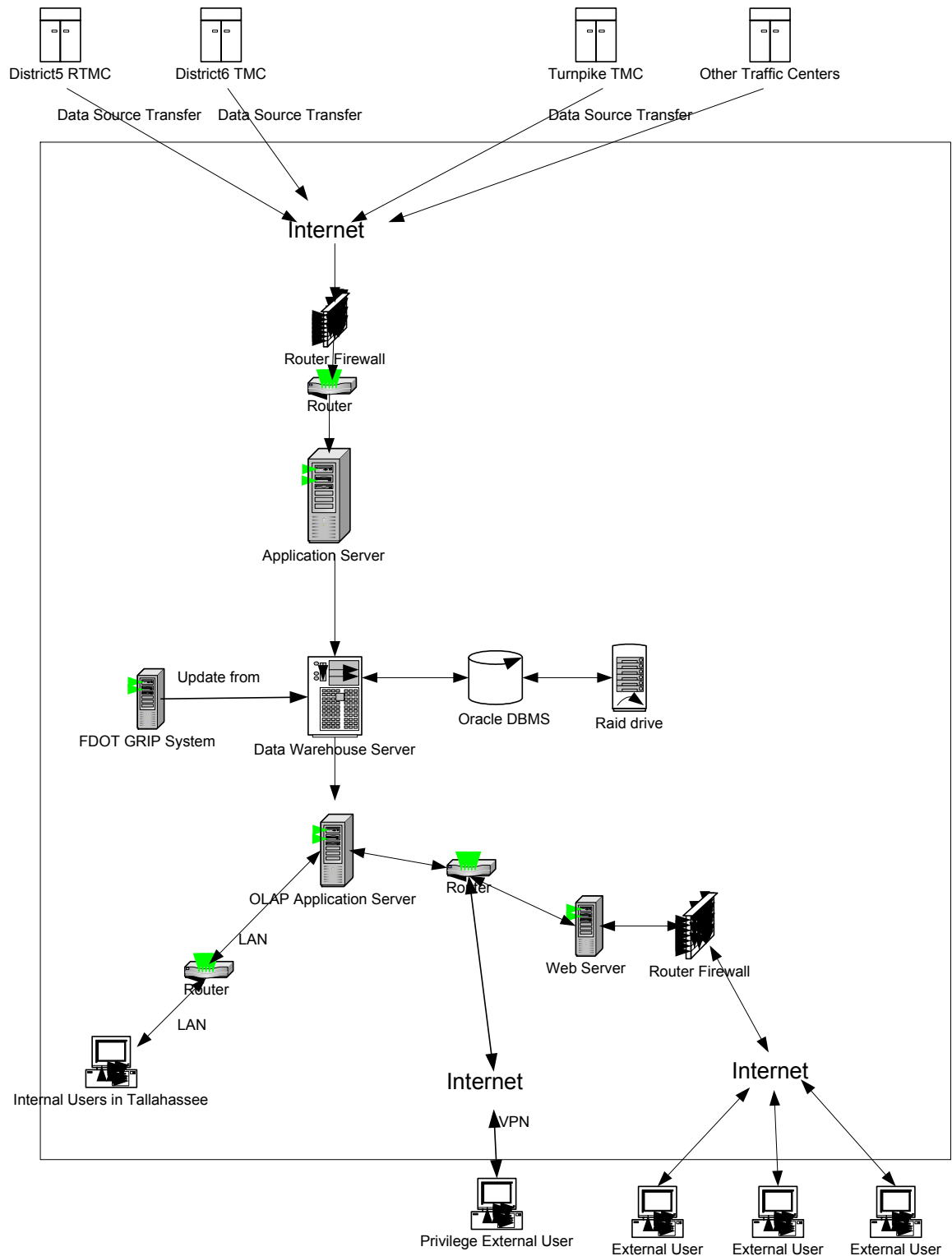
The data of CDW will include all the required traffic data collected by traffic surveillance devices, the required data in legacy information systems and the required data in the existing GRIP database. Source data could be retrieved to the CDW in real time provided that necessary data extraction component is properly implemented in the remote data resource.

### **4.4 Data Warehouse Components [9] [10]**

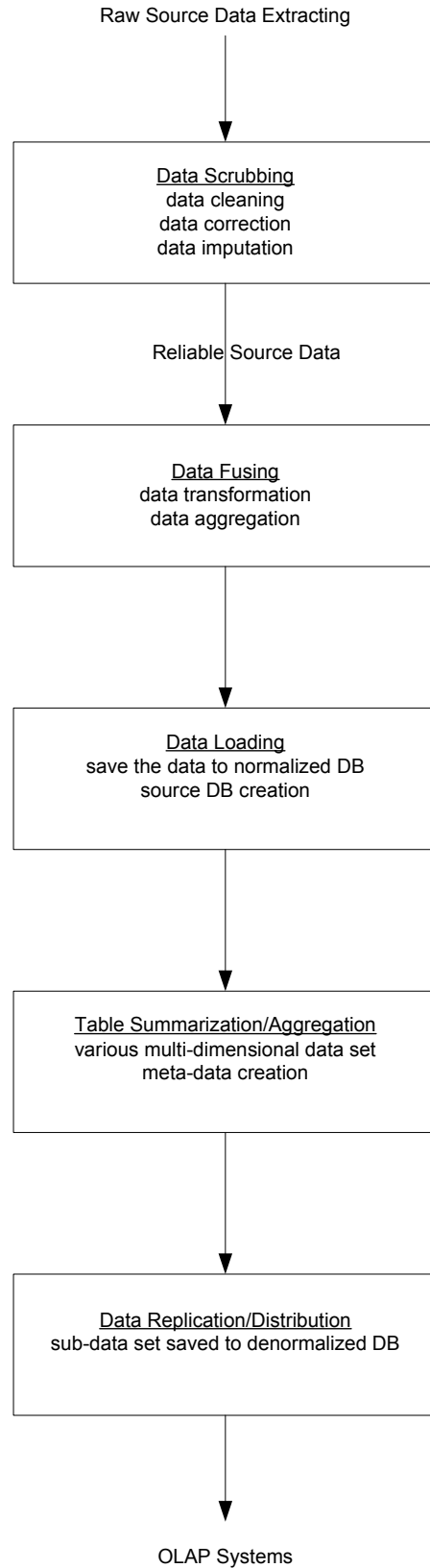
In this section, the conceptual function components that make up the DBMS server will be discussed. Figure 4-4 represents a diagrammatic representation of the functional components of CDW. The functionality of the each component will be discussed as follows.

#### **4.4.1 Data Mapping**

Data mapping defines a rule to specify the relationships between data elements from the data source and data elements in the data warehouse. The rule may define the strategy to choose the data sources and method to deal with overlaps among the data sources. The source data will be retrieved from multiple sites and will be mapped to the integrated data in the CDW.



**Figure 4 - 3 ITS CDW Data Flow**



**Figure 4 - 4 CDW Server Functional Components**

#### **4.4.2 Data Source Extraction**

First, the raw traffic data will be collected by the traffic data source centers including traffic management centers, emergency management centers, weather centers, disaster management centers, ports and motor carrier control centers, etc. Normally, a standalone relational database management system will be used in the individual traffic center to store the raw traffic data. A data-extracting software component needs to be implemented in the existing information system of the traffic center. This component will behave as a data bridge between the traffic center and the CDW. Most of the state-of-the-art commercial database systems provide some level of database replication. In all cases, their solutions are highly tuned to specific environment settings and require a lot of effort in their setup and maintenance. So a system-specific capability should be built for the data-extraction component for each traffic center on a case-by case basis. One plausible solution is to export the source data to flat files. The flat files can be transferred to the CDW using an established communication standard such as CORBA, DATEX, XML or FTP. The data-extraction functionality needs to be done at the lowest data aggregation level so that the greatest data flexibility can be achieved on the CDW side.

#### **4.4.3 Data Scrubbing**

After the raw source data are extracted and before the raw source data are loaded into the data warehouse, the data are scrubbed to correct any errors, omissions, or inaccuracies. According to the different conditions in traffic centers, the source data might already be pre-processed and cleaned by the traffic centers before being extracted to CDW. If this is the case, the scrubbing step could be skipped for the source data from those particular traffic centers. The main purpose of the data scrubbing step is to recognize and remove all of the faulty data. After

the initial CDW is established, the data scrubbing procedure will probably require some refinement. The historical data gradually accumulated in the CDW will provide the information base for the data scrubbing algorithms.

#### **4.4.4 Data Fusion**

A transformation operation will be processed in this step. The transformation operation restructures, reformats, or modifies the scrubbed data to make it ready for use in data warehouse. The retrieved raw data from traffic centers will be processed.

The integrated data warehousing metadata will be created at this stage. There are three types of metadata groups to be developed:

- Informational metadata,
- Physical metadata and
- Transformational metadata.

The informational meta data indicates the information about the content of the data in the data warehouse from business/service point of view. The physical metadata describes the physical format of the data being stored in the CDW database. The transformational metadata describes the procedure by which the data from the source system is processed before being loaded into the data warehouse. The transformational metadata also defines the derivation rules of the data from source system to the data warehouse system as well as the procedure to transform the scrubbed data before storage in the target data warehouse. These rules will apply for the data of the initial data load and for the data of on-going periodic data update. The data with same type but from different sources could be linked together upon the requirement. Data of different types could also be linked together to enhance the cross-subject query within the data warehouse.

Data are often aggregated prior to loading into the data warehouse. Multiple transactions may be combined into a single field, or a range of transactions can be summed or averaged. The data from different sources should be aggregated into the same aggregation level. The transformation/derivation procedure should be processed on source data before it meets the target requirements.

#### **4.4.5 Data Loading**

The data now must be loaded into the data warehousing database. Traditionally, flat files of the integrated, scrubbed, transformed, and aggregated data are prepared and then loaded into the database using a high-performance parallel loader.

There are two design issues. One is for the initial data load and the other is for the ongoing data update for the data warehouse. The main difference between these two issues is the data format. The initial data load must load all data from the source, including the legacy data. The legacy data often has a different format from the current data. This is a data integrity problem in the legacy system. The enforcement rules between the initial data load and the ongoing data update must be established as well. These rules must be applied uniformly to avoid discrepancies between the initial data and the ongoing data updates.

The initial data loading may not have a strict time limitation. However, the ongoing data update has a limited time window that is restricted by the data warehouse application service. This performance concern is critical since the data volume is likely to be substantial.

#### **4.4.6 Table Summarization / Aggregation**

To meet the needs of various user groups, the original data source is aggregated into various summary data sets. This process creates the meta-data. The summary data sets can be

simple one or two-dimensional summaries or very large and complex multidimensional data sets. A large selection of summary data sets targeted toward various users will be required in a multidimensional hierarchy. Table aggregations can be accomplished by simple scripts driven by SQL.

#### **4.4.7 Data Replication / Distribution**

The resulting data sets must be replicated and distributed to the target sub data mart. The sub data mart is organized as a denormalized database system which will service user generated queries via an OLAP system.

### **4.5 Schema Design**

The following observations apply to the design of the schema by which the data will be stored and retrieved:

#### **4.5.1 CDW Application Background**

Data warehouse success is predicated on metadata success. Metadata success requires careful resource planning, ongoing team, management expectation management, specific team metrics, detailed project management, sound architectural design and as much enabling technology as can be acquired. The star schema is an ideal way to represent the multiple dimensions and transactional metrics of transportation data in a relational database as an organization of a CDW. All the major commercial relational DBMS vendors are supporting extensions or intrinsic capabilities to support star schema [11], such as Oracle 8i used by FDOT. The basic star schema is composed of fact and dimensional tables. The fact table contains detailed information, and the dimension table contains information regarding the common

dimensions of the transportation data. The full-production implementation of a star schema normally also includes aggregation tables, history tables and status tables.

The key elements of a star schema include detailed fact tables, dimension tables, aggregations, history tables and status tables. [12] [13]

#### **4.5.2 The Detailed Facts Table**

The detailed facts table lies at the center of the star schema and contains the transactional detail from the data sources used for trend analysis. The fact table is highly normalized in large volume implementation. The fact table will consist of keys and metrics of the traffic data model.

#### **4.5.3 Dimension Table**

Dimensions are categories used to classify the data in the fact table. Dimension can be a simple list or a complex organization that is also called spider or snowflake schema. Star schema dimensions contain descriptive information about the dimensions of the traffic data. The common dimensions are time, location, road ID, etc. Dimensions are denormalized flat data structures that contain many attributes regarding diverse subjects. Dimension tables represent the primary viewpoint of the users and are used extensively to analyze current and historical traffic data with single and/or multiple dimensions. Dimension tables contain three types of attributes:

- *Descriptive*: Descriptive attributes contain full-text descriptions of the keys and codes that make up most data sources.
- *Hierarchical*: Hierarchical attributes provide the classifications and hierarchies that are used to categorize and segment the data. Hierarchies form the basis for most OLAP tools' multidimensional classification capabilities. They provide the framework to allow "drill up", "drill down", and "drill across" in OLAP analysis tools. Hierarchies provide most of the replicated data in the denormalized dimensional tables.
- *Metrics*: Metrics attributes are used to store the activity status of the dimension key. A good design will have a standard set of metrics that are shared across all applicable

dimensions. A well-designed set of metrics opens the door for very powerful, very easy utilization of the iterative implementation approach data warehouse system. Dimension tables are typically the criteria used for filtering facts during data analysis and trend analysis.

#### **4.5.4 Aggregation Tables**

Aggregation in many ways forms the front lines of the data warehouse to the OLAP system. The pre-calculated aggregation tables support the majority of the CDW customer utilization. Users prefer the quick response of a single table answer from the aggregation sets over the much slower response of a detailed fact and multiple dimension table queries. The overall design goal of the data warehouse is that the users should be able to get their answers from one or two tables over 80% of the time. The aggregation tables are the key to accomplishing this goal. Aggregation tables are simply the combination of the detailed facts table and one or more dimension tables. It is critical that any attribute that is not contained in the fact or dimension table should never be introduced into an aggregation table. That means all the user service must be defined absolutely well in the data warehouse designing stage. The anticipated user requirements based on the user interviews will be the initial picture of the aggregate tables.

The combination of multidimensional aggregations and a highly aggregated OLAP tool is a very powerful combination for the user community. To be effective, aggregation tables must be highly denormalized and contain a large percentage of replicated data. Aggregations form the primary data interface to the user community. It is critical that they be designed with convenience, ease of use, and user orientation as the top priorities. They must be denormalized and contain a very large amount of replicated data to allow for easy browsing and single table answers.

#### **4.5.5 History Tables**

History tables are used to capture and store the history of a relationship between two or more attributes of a dimension table.

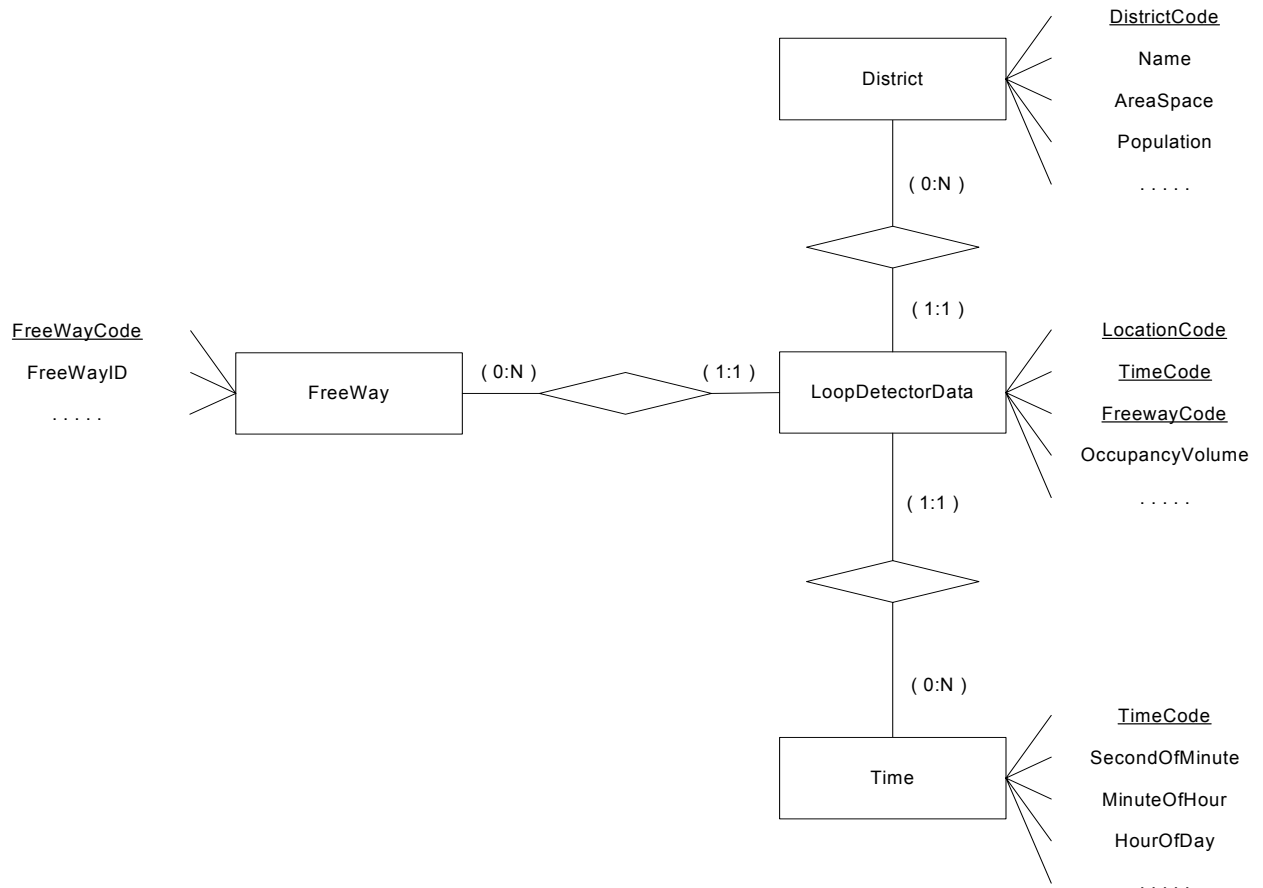
#### **4.5.6 Status Table**

Status tables are used to provide a consistent user interface for all information access needs. Status tables are snapshots of data resources, usually taken nightly. These status tables allow users to check on the status of an operational system for the last refresh cycle.

#### **4.5.7 Sample Schema Tables for CDW application**

For the CDW application, the detailed information of traffic data, such as detector speed, vehicle occupancy and volume is stored in fact tables. The properties of traffic data, such as time, location, etc. are stored in dimension tables. The primary key of the fact table is the composite of the primary keys of the dimension tables which the fact table is linked.

Figure 4-5 represents a sample star schema for the typical loop detector data. The *Loop Detector Data* table has the primary key composed of the primary keys from dimension tables representing District, Time and Freeway. The *Loop Detector Data* table contains the collected data, such as lane occupancy, from loop detectors. The dimension tables contain the general traffic data, such as district location and time. By using this schema, the various and complex statistical data analysis for the loop detector data can be processed efficiently because of the ease with which tables can be combined.



**Figure 4 - 5 Traffic Data Star Schema Sample**

#### **4.6 Implementation Considerations**

The successful and efficient implementation of a CDW is a major challenge. The followings discussion identifies several implementation issues

There is no statewide standard for the source data format. The data stored in the individual traffic are with its own format. Even the data from the same device may be with the different format. The source data are extracted from heterogeneous sources. The information systems used in data resource centers are not standard. The customer requirement is not defined very clearly currently.

An incremental implementation plan with reliability and flexibility will be required. The incremental approach interleaves the design phase and implementation phases of the data warehouse. This approach deals with one data source at a time and yields a much more flexible, manageable project plan. After the initial data load is completed, a usable warehouse will be achieved immediately even though it may not contain all the required data. This first stage data warehouse can produce a fully functional decision support system. After additional features are added to the decision support system or subsequent new data sources are identified, new components may be incorporated into the data warehouse expand its functionality. Summarization rules need not be defined or developed until the new base data is loaded. The benefit of parallel development and implementation activities will be achieved naturally.

The incremental approach will govern the usage pattern of the data warehouse in stages. Figure 4-6 illustrates the high-level task plan for the incremental approach [13].

#### **4.7 Risk Analysis**

A statewide Central Data Warehouse will be extremely complex to build and maintain, and has a high failure potential absent the elements of need, institutional support, and efficient design and implementation. Also, the impact it can have on traditional views of data ownership and organizational structures can be quite large. In this section, the CDW project risks associated with feasibility, sponsorship, motivation, culture, and technology are identified.

The risk analysis process is a formal one that attempts to identify the risk factors that could prevent the project from being implemented within budget and from producing expected benefits, and to manage the identified risks [14]. Based on the features of data warehouse project, the project risks are identified and related mitigation strategies are recommended.

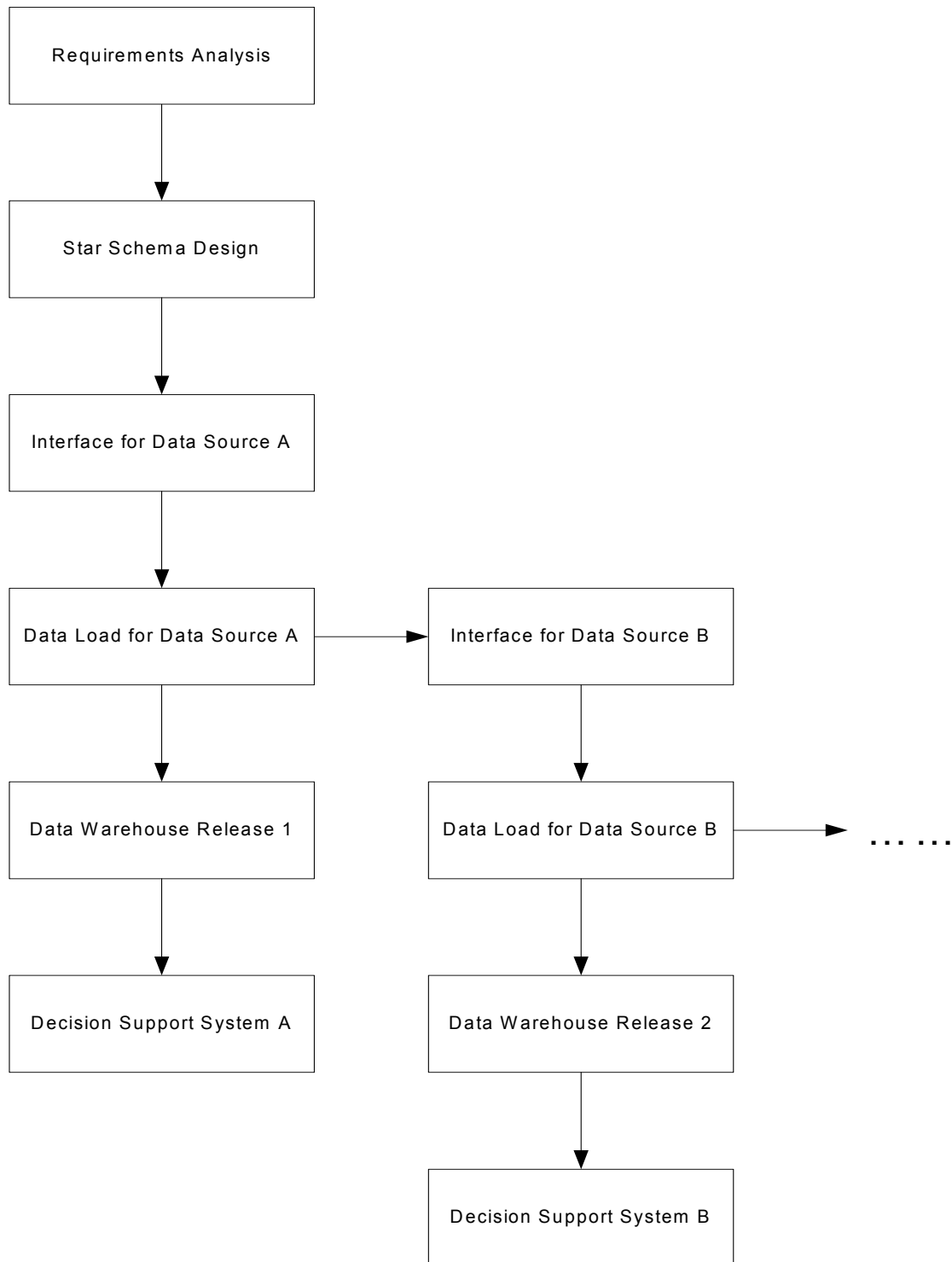


Figure 4 - 6 Iterative Implementation Approach

The impact level is categorized as follows:

- *High* - project objectives and critical path are impacted
- *Medium* - project deliverables and milestones are impacted
- *Low* - a project issue or action item is impacted

#### **4.7.1 Risk Type: Inadequate Project Scope**

The project scope should be identified by the functional requirements and technical requirements in detail. Uncontrolled increases in the number of project requirements can occur, leading to missed deadlines, delays in delivery of products, and potential project failure. Because of this history in the systems development industry, potential “scope creep” is a major risk to the successful completion of the data warehousing project.

##### **Mitigating Action**

- Implement a requirements-management process that identifies and tracks the requirements being implemented within specific project phases to document and control the implementation schedule for newly identified requirements.
- Use appropriate techniques to control project scope while remaining on schedule
- Implement a strong release management program that controls the rollout of certified software products

**Impact Level:** High

#### **4.7.2 Risk Type --- Inadequate Control of Project Costs**

Failure to control all project costs within the structure of a well-defined project budget will lead to project cost overruns and potential project shutdown.

##### **Mitigating Action**

- Develop a project budget plan that identifies and projects the cost of all work orders within the budget ceiling allowed for the project.

- Support districts in formulating their budget requests for internal hardware and network services.

**Impact Level:** High

#### **4.7.3 Risk Type --- Lack of Funding**

Failure of FDOT to obtain funding to support the CDW rollout as planned will lead to delay or cancellation of the CDW implementation. In addition, failure of districts to adequately fund infrastructure establishment to prepare for the CDW rollout will lead to delays in rollout to specific districts or cancellation of CDW participation by specific districts.

##### **Mitigating Action**

- Provide adequate planning to fully inform districts on timing of the rollout in different districts
- Strive to obtain adequate funding from the Federal Highway Administration
- Communicate the importance of CDW products to FDOT traffic management operation

**Impact Level:** High

#### **4.7.4 Risk Type --- Inadequate Human Resources**

Staffing levels consistent with the project plan are essential to successful completion of the CDW project. Availability of the appropriate technical skill cannot always be guaranteed in today's highly competitive labor market.

##### **Mitigating Action**

- Develop a staffing plan for Central and District office personnel
- Develop a training and mentoring plan for Central and District office personnel
- Develop a strategy for retraining Central and District office personnel after skills have been upgraded for new technologies and management processes
- Develop a support plan for staff at data centers
- Assure that vendors deliver proposed staff through contract management process

**Impact Level:** High

#### **4.7.5 Risk Type --- Resistance to Change**

The introduction of new technology will require a large change in the way business is conducted at the FDOT central office and the districts TMCs. To some degree the resistance by district and FDOT personnel to changes in business operations could pose a risk to the successful implementation of a CDW.

##### **Mitigating Action**

- Identify the existing process for effecting changes to statutes, regulations, and operational procedures that must be changed to allow implementation of a CDW
- If necessary, develop new strategies for changing statutes, regulations and operational procedures to overcome inadequacies of the existing process for effecting changes.
- Develop a strategy for assimilation of the planned changes for personnel affected by the new system

**Impact Level:** Medium

#### **4.7.6 Risk Type – Lack of Training**

There will be a continuing need to train personal not familiar with the fundamental concepts, technologies and methods used in a CDW. Personnel lacking the proper background in data warehousing may not take full advantage of the many features of a CDW and may also jeopardize the system integrity.

##### **Mitigating Action**

- Develop a clearly defined training plan for all personnel
- Encourage information exchange among staff members.
- Develop strict operational regulations

**Impact Level:** Medium

#### **4.7.7 Risk Type --- Data Conversion Uncertainty**

Conversion of data from systems in the field to the centralized analysis system is a significant effort that is easy to overlook or underestimate when planning for implementation of a new system. The risk is that the data will not be transformed correctly because data elements are identified or defined incorrectly or rejected by the new system due to incorrect mapping of old fields to new fields.

##### **Mitigating Action**

- Develop a detailed conversion plan.
- Reconcile legacy data elements with definition of data elements identified in the recognized standards
- Develop a CDW data model that is consistent with national ITS standards and architecture
- Develop detailed mapping of legacy data elements to the CDW data model and physical model

**Impact Level:** Medium

#### **4.7.8 Risk Type --- Hardware and Software Failures**

A CDW places a high requirement on the hardware such as field sensors, communication network facilities and the database server. The failure of communication network facilities or the shutdown of database server can deprive the end-users of timely access to the data.

##### **Mitigating Action**

- Develop contingency plan for work stoppage of one, three, and six week to allow time for stabilizing any disruptions caused by software and hardware
- Develop contingency plans for repair and replacement of CDW components that may fail.
- Develop database backup plans for daily, weekly and monthly periods.

**Impact Level:** High

#### **4.7.9 Risk Type – Inadequate Data Processing Capacity**

A CDW must process large quantities of data in a timely manner. Data processing capacity is therefore a very important consideration. Inadequate capacity can result in unsatisfactory response to user queries. It can also prevent data from being loaded into the data warehouse within the allotted timeframe. Low system performance poses a great risk to the success of a CDW project.

##### **Mitigating Action**

- Ensure that the data model and system architecture are optimized for data processing and adequate to handle the anticipated data loading.
- Develop a test plan to monitor system performance and identify the real bottlenecks
- Periodically clean the redundant and outdated data from the online storage areas.

**Impact Level:** High

#### **4.7.10 Risk Type – Bad Data Quality**

Data collected by field surveillance equipment are prone to errors due to equipment malfunctions, and communication disruptions. Some errors may also occur in the data aggregation process. Inaccurate data can lead to the wrong conclusions from an analysis.

##### **Mitigating Action**

- Perform quality control before data are sampled, summarized and stored for later use.
- Develop mechanisms for correcting problems with data collection equipment
- Set appropriate default values for the missing data and questionable data.
- Improve the management of metadata

**Impact Level:** High

## **4.8 Future Research Tasks**

The observations and recommendations presented in this chapter must be refined substantially in the next phase of the CDW design. Specific hardware and software configurations must be chosen within the framework of the recommendations provided here. More details will be required on the nature and format of the data available from the various sources, and of the current means (which appear to be very limited at present) used to archive those data. Overall system costs and time lines should be projected from the Phase 2 results, and the risk analysis presented in this section should be expanded.

A specific task should focus on gaining a detailed understanding of the recent ASTM publication covering the subject of ITS data archiving, so that the work done in Florida will be compatible with similar efforts in other states. Other tasks should address the mechanisms by which the Safety and ITS office activities and data can be integrated directly into sub layers of GRIP and the EIS and mechanisms by which input from stakeholders at all levels can be solicited formally and used effectively to guide the CDW development process.

## **Chapter 4 References**

[1]: FDOT, GIS Info, <http://www11.myflorida.com/planning/statistics/GIS/default.htm>. Access in March 2003

[2]: Haitham Al-Deek, Amr Abd-Elrahman, "AN EVALUATION PLAN FOR THE CONCEPTUAL DESIGN OF THE FLORIDA TRANSPORTATION DATA WAREHOUSE FINAL REPORT", University of Central Florida, January 2001.

[3]: FDOT Enterprise Information System  
<http://www.bts.gov/gis/reference/gist/2001/61/613/613.pdf>. Access in March 2003

[4]: "Florida Statewide Intelligent Transportation System (ITS) Architecture & Standards", <http://www.jeng.com/html/florida/default.htm>, Access on March 10, 2003

- [5]: *Efrem G. Mallach*, ( 2000 ), “Decision Support and Data Warehouse Systems”, McGraw-Hill Ltd., ISBN 0-07-289981-6
- [6]: *Erik Thomsen*, ( 2002 ), “OLAP Solutions---Building Multidimensional Information Systems”, John Wiley & Sons, Inc., ISBN 0-471-40030-0
- [7]: Data Warehousing, <http://www.billinmon.com/>. Access in March 2003
- [8]: *Paolo Atzeni, Stefano Ceri, Stefano Paraboschi, Riccardo Torlone*, ( 2000 ), “Database Systems concepts, languages & architecture”, McGraw-Hill, Inc, ISBN 007-2353872
- [9]: *Hackney, Douglas*, ( 1997 ), “Understanding and implementing successful data marts”, Addison-Wesley
- [10]: *Ramon C. Barquin, Herbert A. Edelstein*, ( 1997 ), “Planning And Designing The Data Warehouse”, Prentice Hall PTR, ISBN 0-13-255746-0
- [11]: *Ken S.Brathwaite*, ( 1988 ), “Analysis, Design, & Implementation of Data Dictionaries”, McGraw-Hill, Inc, ISBN 0-07-007248-5
- [12]: *Paolo Atzeni, Stefano Ceri, Stefano Paraboschi, Riccardo Torlone*, ( 2000 ), “Database Systems concepts, languages & architecture”, McGraw-Hill, Inc, ISBN 007-2353872
- [13]: *Ulka Rodgers*, ( 1999 ), “Oracle: A Database Developer’s Guide”, Prentice-Hall, Inc., ISBN 0-13-841420-3
- [14]: *North Carolina Windows of Information for Student Education Project Status Report*, Department of Public Instruction, February 2001.

## **5 PROTOTYPE CDW DEMONSTRATION**

This chapter describes a physical prototype for a CDW, designed to illustrate the concepts and recommendations contained in this report. While the prototype is functional and accessible from the Internet, it is not capable of doing useful work. Its sole purpose is that of demonstration.

The Emergency Management ITS application was chosen for the prototype because of FDOT's expressed interest in the EMS area. The project team has limited knowledge of the EMS field itself. The prototype must therefore be viewed as a demonstration of CDW concepts and not of emergency management concepts.

According to Florida's ITS Strategic Plan, emergency management is one of the essential elements of management and operations in Florida ITS. Since a CDW provides a coordinated service for aggressive information sharing, operational cooperation and joint service programs among agencies (across sectors), its implementation would greatly improve the efficiency of the emergency management system (EMS). In this chapter, a conceptual framework for establishing a statewide, integrated EMS based on a CDW will be explored. The topics include an introduction to emergency management from an ITS perspective, participating agencies of the system, and system architecture and data flow. A demonstration of an incident inquiry and rerouting system is also presented.

### **5.1 What Is Emergency Management in ITS?**

Emergency management in ITS is concerned primarily with incidents on the highway system. An incident is an event that causes significant delay and/or damage. Incidents generally include:

- Motor vehicle crashes involving serious personal injury or fatality
- Crashes where a load of cargo is spilled
- Crashes involving hazardous material cargo
- Overturned cars or trucks
- Downed power lines across roadways
- Structural failures of bridges or roads.
- Fire, flood, hurricane and other disasters

Such incidents create hazards as well as delay, frustration, inconvenience, wasted fuel and higher costs for motorists as well as for the system. Stopped traffic can create secondary motor vehicle crashes. Local streets can become gridlocked with motorists trying to avoid the incident scene.

Incident and emergency management is one of the areas that has received numerous benefits from ITS applications. In ITS applications, emergency management mainly focuses on the detection, response, and management of incidents or other disruptions to minimize delay and improve safety. The typical problems that the incident and emergency management system faces are [1]:

- Excessive amount of time to detect/verify, respond to, and clear/remove the incident;
- Difficulties in communicating and sharing data among incident management agencies;
- Congestion, accidents and other traffic management problems occurring on the surface street system due to the incident diversion;
- Lack of comprehensive and timely traveler information for freeway incidents;
- Inconsistent strategies for effective, coordinated traffic incident management planning and evaluation.

To some extent, the above problems are due to failure in communicating and sharing real-time incident-related data and the lack of historical data for designing effective system strategies. Integrated EMS is a new management strategy that is implemented by enhancing communication of incident data among incident management and emergency response personnel, both on-scene and at multi-modal communications and operations centers to achieve the ITS goals of improved safety, reduced congestion, improved mobility, and increased efficiency and productivity.

The Integrated Incident Management System (IIMS) [2] operated by New York State Department of Transportation is an example of an EMS. It is a multi-agency project of New York State Department of Transportation, in partnership with New York City Department of Transportation and the New York City Police Department. The IIMS is unique in its ability to transmit incident scene data real time to incident operations centers. These data include information required to effectively select the proper responders and equipment for clearance.

In the following sections, a statewide, integrated EMS framework will be explored, based on the Florida ITS Architecture. The hypothetical statewide system will link the CDW of FDOT, Transportation Management Centers (TMC) and Emergency Management Center (EMC) of the districts and incorporate other operational components to implement an integrated emergency management plan in order to maximize motorists' safety and traffic efficiency through incident prevention, timely response, site clearance and motorist information.

## **5.2 Participating Agencies**

A critical step in developing a successful EMS is to identify the relevant stakeholders and to build from this group's critical coalitions and institutional framework that will support, protect, and fund the program. The list of stakeholders includes:

- Agencies that are responsible for emergency management functions
  - State Department of Transportation
  - State Police Agencies
  - Local Traffic Management Center
  - Local Emergency Management Center
  - Highway Patrol
  - 911 Dispatch
  - Fire and Rescue agencies
  - Weather Center
  - Department of Natural Resources
  - Disaster Management Centers
- Other departments or individuals within the typical response agencies
  - Motor carrier control centers
  - Transit system
  - Seaports or airports

- Other stakeholders who are affected by, or contribute to, traffic operations
  - Media
  - Towing and recovery
  - Emergency Medical Service (EMS)
  - Hazardous materials treatment
  - Policy Makers

The impacts of incidents are not constrained by the jurisdictional boundaries that delineate agency service areas or operational responsibilities [3]. The impact of major traffic incidents transcends political, jurisdictional, and geographical boundaries, and may affect entire regional road networks and hundreds of thousands of travelers. Typically, emergency management activity is underway in large- or medium-sized urban areas with significant traffic congestion problems. The success of the emergency management program is a direct function of the extent of cooperation and coordination that is achieved. Within the statewide, integrated emergency management system, a key issue is how to cooperate and coordinate so many agencies that should be an integral part of planning for and responding to traffic incidents. Accurate and real-time incident scene data are the basic information, based on which each agency can make appropriate response to the incident. That is, the key to the success of the statewide, integrated emergency management system is the ability to communicate and share incident-related information among agencies in this system. Thus, the statewide, integrated EMS framework will focus on data sharing and communication.

### **5.3 System Architecture and Data Flow**

According to the scope of data sharing and communication, the statewide, integrated EMS can be divided into two levels: district-level EMS and state-level EMS.

The district-level EMS implements the incident data sharing among field units and multiple centers/agencies in real time. The shared data include pictures, location, and dispatch information, which provide instruction for dispatching proper equipment and resources. The state-level EMS involves CDW as open the database infrastructure and it not only shares real-time incident related data with operation centers in the entire state, but also stores incident data for long-term planning analysis.

#### **5.3.1 District-level EMS**

Local EMS mainly includes the local EMC, Highway Patrol, local Police Department, towing and recovery agencies, emergency medical service and hazardous materials treatment agencies. All agencies in the integrated system are connected using communication links. The local TMC is the center of the district-level EMS and it performs the following functions:

- Collection of real-time transportation information;
- Dispatch of traffic management teams for controlling and directing traffic during incidents;
- Coordination of EMC incident management actions with other emergency response agencies and
- Keeping the public informed of traffic and roadway conditions.

The first step in the emergency management process is incident detection and verification. In Florida, there are CCTV cameras, electronic loop detectors, call boxes, telephone numbers for use by drivers with cellular phones, and visual observations through peak-period

motorcycle patrol and dedicated freeway service patrols. All these means can be used for incident detection and verification. Initial planning for integrated EMS should focus on full utilization of existing resources and provision of low-cost enhancements.

Incidents can be reported to the Highway Patrol, by the following means:

- A patrolling officer
- FSP truck
- 911 calls or call boxes by a passing motorist
- 911 calls or call boxes by someone involved in an incident
- Others

In the local TMC, incident information can be detected by the following means:

- Closed circuit television
- Electronic loop detector
- Video-image processing
- Microwave radar detection
- Automatic vehicle location
- Others

An incident is first reported to local Highway Patrol or detected by local TMC. These two agencies are in charge of separating real incidents from false alarms. In the Highway Patrol, incidents are generally verified by the Highway Patrol officers on the site. In the TMC, incidents are generally verified through CCTV or other sensors. After the incident is verified, the incident information collected by the Highway Patrol is transmitted to the local TMC. The TMC then proceeds with data processing and dispatch plan design and implementation, which is a four-step process as described in the following.

### **Step 1. Incident Classification**

Based on the scope and severity of incident, incident can be classified into six classes [4]:

Class 1: Up to 2 vehicles on shoulder or center divider

Class 2: 3 or more vehicles on shoulder or center divider

Class 3: Intrusion of incident into one or more lanes, but not all lanes

Class 4: Intrusion of incident into one or more lanes and full closure required for two hours or more

Class 5: Area-wide incident affecting multiple jurisdictions or multiple districts

Class 6: State-wide incident such as hurricane, earthquake

If the incident class is 4 or lower, the local TMC will design the dispatch plan. If the incident class is higher than 4, the local TMC will transmit the incident information to the CDW. Traffic operation agencies in FDOT will be responsible for the design of the incident clearance plan or evacuation plan.

### **Step 2: Determining the nature of the incident, dispatch plan, and rerouting plan**

After the incident is classified, the exact features of the incident will be identified and dispatch plan will be designed. The information includes the detailed information about the incident and the information required for equipment and personnel dispatch, the status of the responder and incident. The status of the responder and incident will be updated in real time. Different algorithms can be used for design of rerouting plan. Please refer to Emergency Management Demo section for an example.

### **Step 3: Sharing data among local agencies and the public**

After the exact nature of the incident and dispatch plan are determined, the information will need to be shared among local EMC, police office, fire and rescue agency, other related

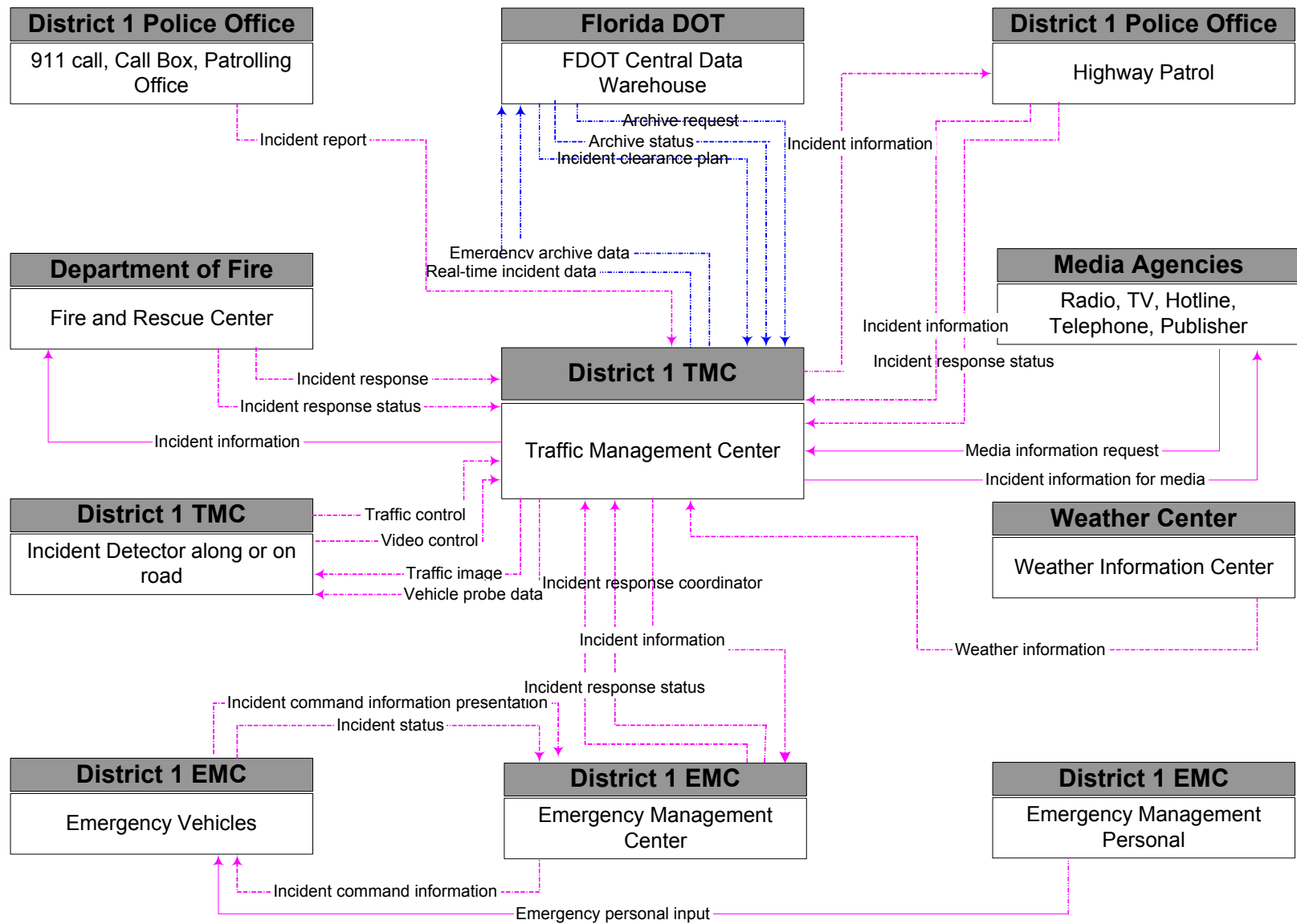
agencies, and the public through the communication link. EMC uses the shared information for prompt response to the incident. And motorists use the information for alternative route selection.

#### **Step 4: Transmitting the data to CDW**

As discussed in Step 2, if the incident Class is 4 or lower, after the incident is removed, the final data about the incident will be transmitted to the CDW for long-term traffic planning analysis. If the incident class is higher than 4, the real-time incident information will be transmitted to the CDW for real-time emergency management. The incident clearance plan or disaster evacuations strategy designed by the operation centers of FDOT will be transmitted to the TMC of each district and then shared with other related agencies of the district-level EMS. Figure 5-1 shows the data flow in the local EMS.

The integrated model provides data communications between emergency response personal (first and secondary responders both at, and away from, the scene of an incident) and all agencies and centers that need to support the emergency response. It can greatly improve the efficiency of emergency management and reduce time to detect, verify, and clear by achieving the following objectives.

- For the first responder
  - Improve incident data to and from scene
  - Improve multi-agency coordinated response
  - Improve post-incident data analyses
- For secondary responders:
  - Reduce incident notification time
  - Reduce incident verification time
  - Dispatch equipment and personnel more effectively
  - Reduce travel time to incident site



**Figure 5-1: Data Flow in the Local-level Emergency Management**

### **5.3.2 State-level EMS**

The state-level EMS mainly includes the CDW, operation centers, weather center, and some media etc. The CDW in the FDOT is the center of the state-level EMS. It not only shares real-time data with existing operation center applications in the entire state, but also stores incident data for long-term analysis. Its main functions in emergency management are as follows.

- The CDW collects the real-time incident-related information (when the incident class is 4 or higher) from the district-level EMS for the operation center application in FDOT. Based on the real-time information, the operation centers in the FDOT designs the incident clearance plan or disaster evacuations strategy that will provide instruction and coordination for district-level EMS.
- The real-time incident-related information collected by the CDW also can be shared among the statewide district-level EMS. Thus, data are collected from one local EMS and then shared with other local EMS through the CDW, especially data on incidents that affected multiple districts.
- The CDW integrates GIS capabilities. Transportation features and other geographically-related data included in the GIS have a great impact on emergency management situations such as fire, hurricane and flood evacuations.
- The CDW provides the incident-related information for the media such as television, highway advisory radio, wireless cell-phone hotlines, and variable message signs, etc. The information disseminated by these media will assist the motorists in travel decision and route choice. The CDW also provides the user with an image of the archived data and facilitates the tools to query, retrieve, and display the data.
- The CDW collects the real-time information about the weather, route changes, and closures from the weather center and other operation ITS centers. This information is shared among the district-level EMS centers.
- The CDW stores incident data for long-term analysis. Long-term analysis contributes to the successful implementation and operation of incident management programs, such as the development of incident detection algorithms based on real time and historical traffic monitoring.

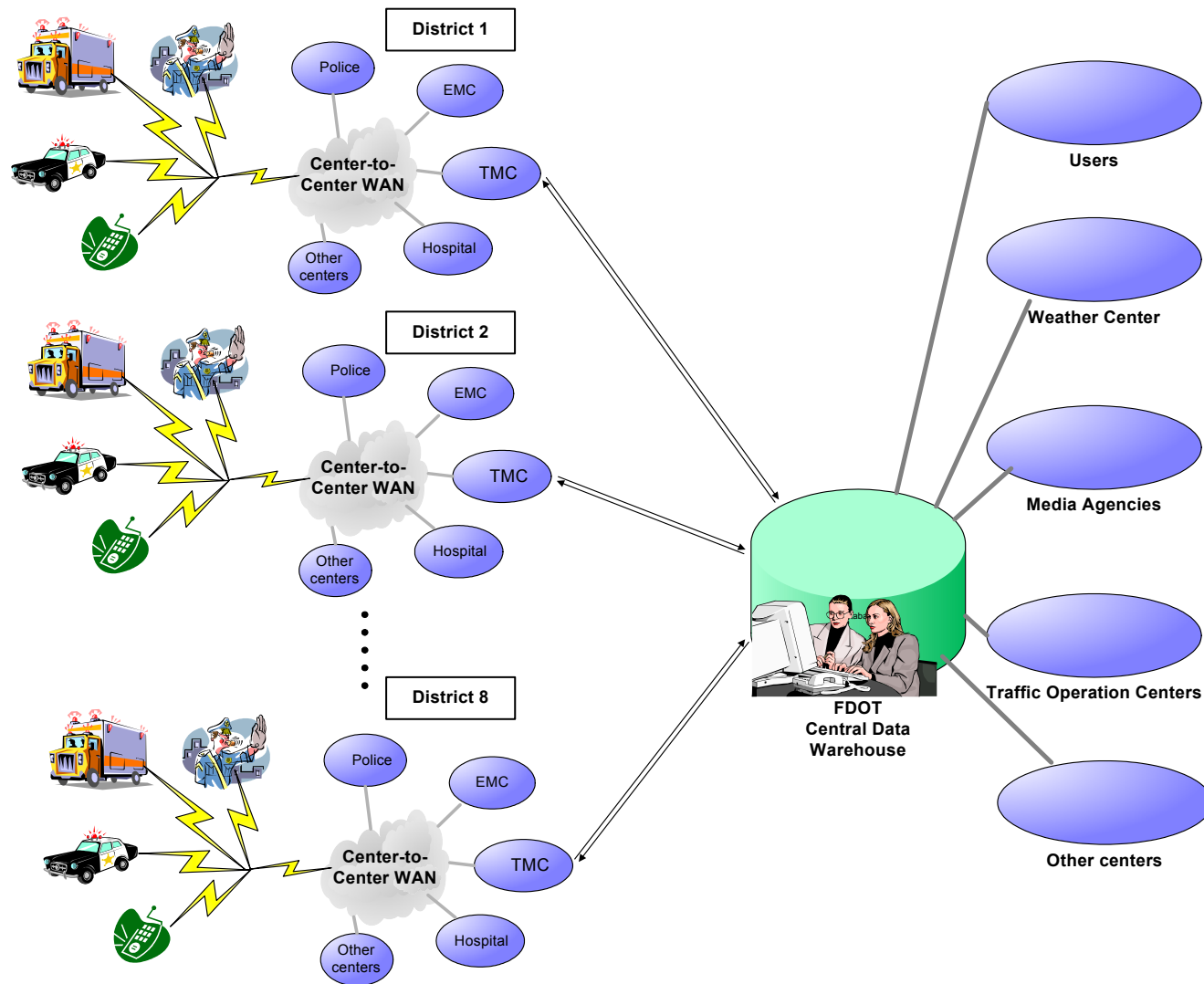


Figure 5 - 2: Data Flow of State-level Emergency Management Syst

- The longer-term data can be used for planning, research, performance monitoring, and policy purposes. For example, most incident management activity depends heavily on the operations and maintenance budgets of the state DOT. The historical incident data provide better reference for developing annual funding plans.

The data flow of a state-level EMS is shown in Figure 5-2. The format used in the Washington State Department of Transportation's (WSDOT) incident response program [5] provides a good model for other states. This data format includes enough information for qualitative and quantitative analysis. Data elements that are needed in the CDW are identified in Table 5-1.

In general, a statewide, integrated EMS is an effective emergency management model. The district-level EMS enhances the communication of incident data among incident managers at operations centers and incident response personnel at the incident scene. In the state-level EMS, the CDW not only implements the data sharing throughout the state but also provides macroscopic instruction to the district-level EMS.

#### **5.4 An Application of a CDW for Emergency Management Services**

In this section, we describe the demo design, which shows the proposed organization of the data warehouse, its query functionality, and an application of the data warehouse in supporting emergency management operations. The purpose of this demo is to illustrate the application of data warehouse concepts and the web-based server-client architecture. While this application has the potential for expansion is a proof of concept study, it is not capable of performing any useful work in its present stage of development.

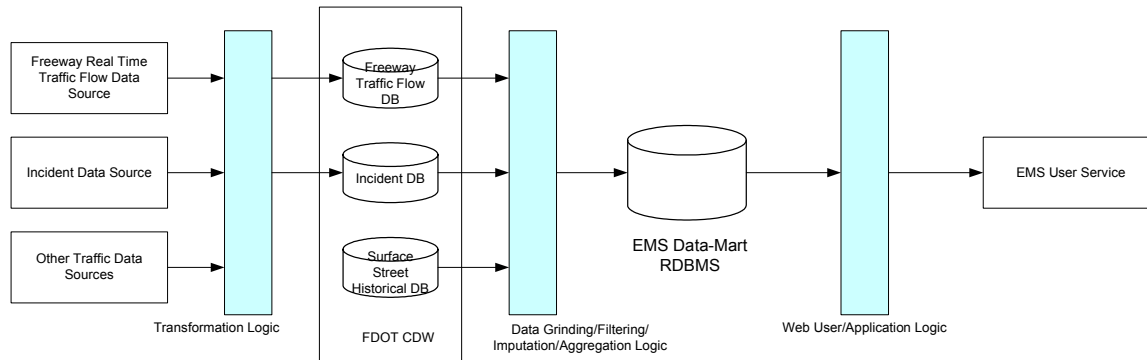
Table 5-1: Incident Information Database Elements

| CATEGORY                   | INFORMATION                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|----------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Location (City and County) | <ul style="list-style-type: none"> <li>• Nearest city</li> <li>• County in which the incident occurred, by code</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| General Information        | <ul style="list-style-type: none"> <li>• Name of person preparing the report</li> <li>• Date of the report</li> <li>• Date of the incident</li> <li>• Time of the incident</li> <li>• Time the IRT received the call to respond to the incident</li> <li>• Time the first IRT member arrived at the scene</li> <li>• Date and time incident ended</li> <li>• Time last IRT member left the scene</li> <li>• Repair notes</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| WSDOT Personnel            | <ul style="list-style-type: none"> <li>• Number of employees involved</li> <li>• Number of hours each was at the incident site</li> </ul> Location<br><ul style="list-style-type: none"> <li>• Regional Maintenance Area number</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Highway/Route Information  | <ul style="list-style-type: none"> <li>• State route number and nearest milepost number</li> <li>• Description of the intersection state route and milepost number not available</li> <li>• Travel direction of affected lanes</li> <li>• Lanes closed (i.e. Ramp, single lane, multiple lanes, all lanes in one direction, or all lanes in both directions)</li> <li>• Roadway surface</li> <li>• Reason for road closure (i.e., single-vehicle accidents, multiple-vehicle accidents, fatal accidents, hazardous and non-hazardous material spills)</li> </ul> Travel Conditions<br><ul style="list-style-type: none"> <li>• Weather conditions (i.e., rain, snow, fog, wind, calm and clear)</li> <li>• Road conditions (i.e., dry, wet, ice-covered, snow-covered)</li> <li>• Light conditions (i.e., day, dawn, dusk, or night—night with street lights on, night with street lights off, or night with no street lights at all)</li> </ul> Agency Participation<br><ul style="list-style-type: none"> <li>• Agencies present at the incident site</li> </ul> |
| Equipment                  | <ul style="list-style-type: none"> <li>• equipment used</li> <li>• Incident Response Team equipment used</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Materials and Maintenance  | <ul style="list-style-type: none"> <li>• IRT vehicle materials used</li> <li>• Follow-up maintenance</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Clean-up                   | <ul style="list-style-type: none"> <li>• Delayed cleanup until off-peak time</li> <li>• Conditions at the incident scene (i.e., presence of hazardous materials, non-hazardous materials, fuel spillage, fire, flammable liquid, corrosive material,</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |

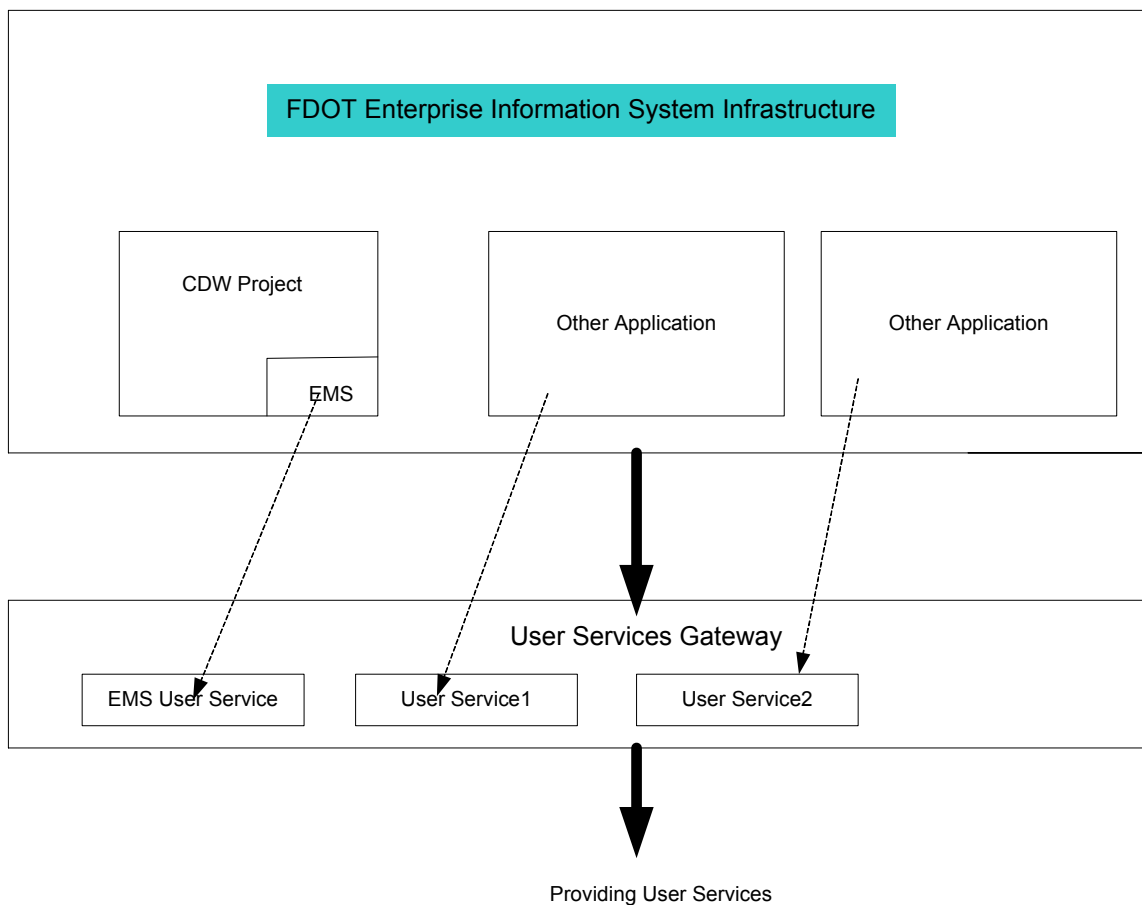
|                                   |                                                                                                                                                                                                                                                                                                                                                                                                       |
|-----------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                   | explosive material, radioactive material, or toxic materials)<br>• Agency responsible for cleanup                                                                                                                                                                                                                                                                                                     |
| Traffic Control                   | • Lane where the incident originated<br>• Detour route, if applicable<br>• Occurrence of incident in construction zone<br>• When lanes opened                                                                                                                                                                                                                                                         |
| Investigation                     | • Method of Washington State Patrol investigation (i.e., tape, total station equipment)<br>• WSP accident and case number (if applicable)<br>• Lead investigating agency<br>• Number of vehicles involve in the incident<br>• Number of injuries,<br>• Number of fatalities                                                                                                                           |
| Number of Vehicles                | • Number of vehicles by type (e.g., one bus, two passenger cars, and a taxi)                                                                                                                                                                                                                                                                                                                          |
| Causing Party's Vehicle Type      | • Type of vehicle the causing party was driving                                                                                                                                                                                                                                                                                                                                                       |
| Driver and Vehicle Identification | • Driver's last name, first name and middle initial<br>• Driver's license number<br>• State or province in which the licenses was issues<br>• Vehicle license number of the party at fault<br>• Vehicle year, make, model, and vehicle identification number<br>• State or province that issued the vehicle license of the party at fault<br>• Insurance of the party at fault<br>• Insurance company |
| Comments                          | • Description of cargo that was cleared from incident, how it was disposed of, or whether it was stored, etc.<br>• Or other information/comments                                                                                                                                                                                                                                                      |

#### 5.4.1 EMS as A Candidate Component of CDW within FDOT EIS

EMS could be implemented as an active component of the CDW project and behave as a user service within the FDOT EIS. The EMS will function as an independent DSS itself and will have its own data-mart, which contains the data repository for its direct query purpose. Figure 5-3 shows the architecture [6] view of the relationship between the EMS and CDW. Figure 5-4 shows the status of the EMS within the FDOT EIS organization.



**Figure 5 - 3 Architectural Relation of EMS and CDW**



**Figure 5 - 4 EMS Status within FDOT EIS Organization**

### **5.4.2 Data Organization**

To optimize the database query, a special data organization has been adopted. Since the majority of the EMS database activities will be online query, the database entity relations within the DBMS are organized as a flat pattern so that the querying time through the entities is minimized. This type of DBMS data organization is also called Star Schema design. An Oracle 9i DBMS has been adopted as the data warehouse infrastructure for this demo, which is running on a Windows XP server.

#### **Meta Data Strategy [7]**

The star schema is composed of one single fact table and several dimensional tables associated with it. The fact table contains the actual querying content named fact values and the foreign keys of the dimensional tables. The fact table is highly normalized. The fact table behaves as the central connector for the associated dimensional tables. The dimensional table contains the primary keys associated to the central fact table and the feature attributes for the various querying point views. The dimensional table is normally denormalized.

#### **Entity-Relation Diagram (Star Schema Design)**

Two star schema diagrams have been defined for the EMS data warehouse data organization. One is for the traffic flow data and the other is for the traffic incident data.

Within EMS traffic flow star schema, there is one fact table and six associated dimensional tables. The six dimensional tables represent District, City, County, Freeway, Surface Street and Time. The District, City, County tables provide the location dimension information. The Time table provides the date dimension. The Freeway and Surface Street tables provide the various highway dimension views. The fact table contains the traffic flow volumes and the

foreign keys for the other six dimensional tables. Figure 5-5 provides details of the design. The real-time freeway traffic flow data from loop detectors in the highway will be updated in this schema. The EMS retrieves and updates the real-time traffic flow in the Traffic Flow Data table. The anticipated traffic flow for the surface network will also be stored in this schema. The surface street traffic flow will be derived from the corresponding data stored in CDW data repository.

Within EMS traffic incident star schema, there is one fact table (Incident Data) and five associated dimensional tables. The five dimensional tables represent District, City, County, Freeway and Time, which have the same definition as those in traffic flow star schema. The fact table contains the traffic incident data and the foreign keys for the other five dimensional tables. Figure 5-6 demonstrates the details.

### **Schema Data Dictionary**

The entities of the star schema and the relationships associated with it are described in Tables 5.2.

#### **5.4.3 Incident Rerouting Plan**

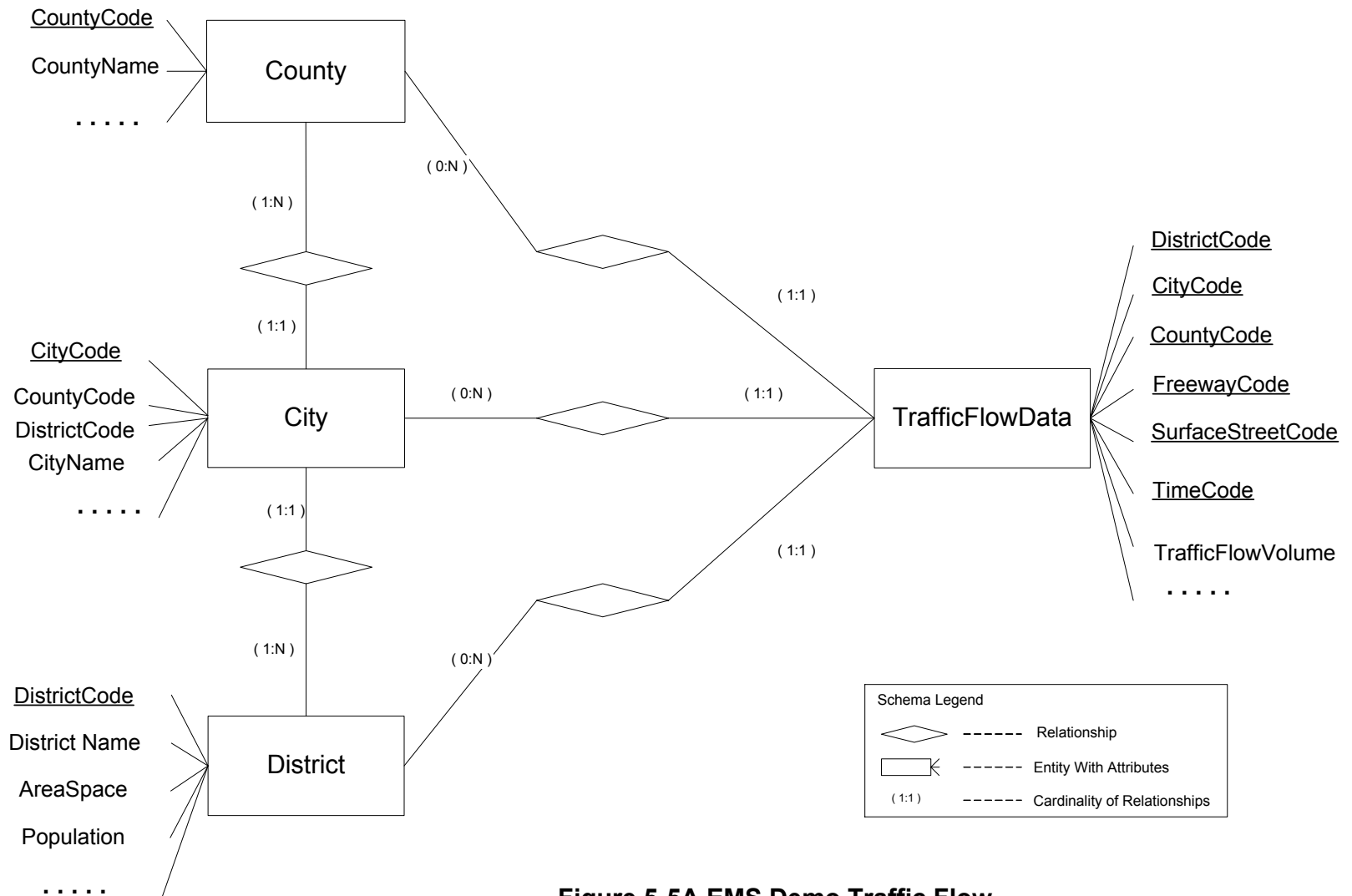
When an incident happens on a freeway, the consequent congestion would occur around the incident location. A rerouting plan provides the traffic with alternative path(s) to take (usually using major arterial streets) and the goal is to minimize the traffic congestion caused by the incident. The rerouting plan will be devised by district TMCs and distributed to the travelers.

Table 5-2: Entities of The Schema

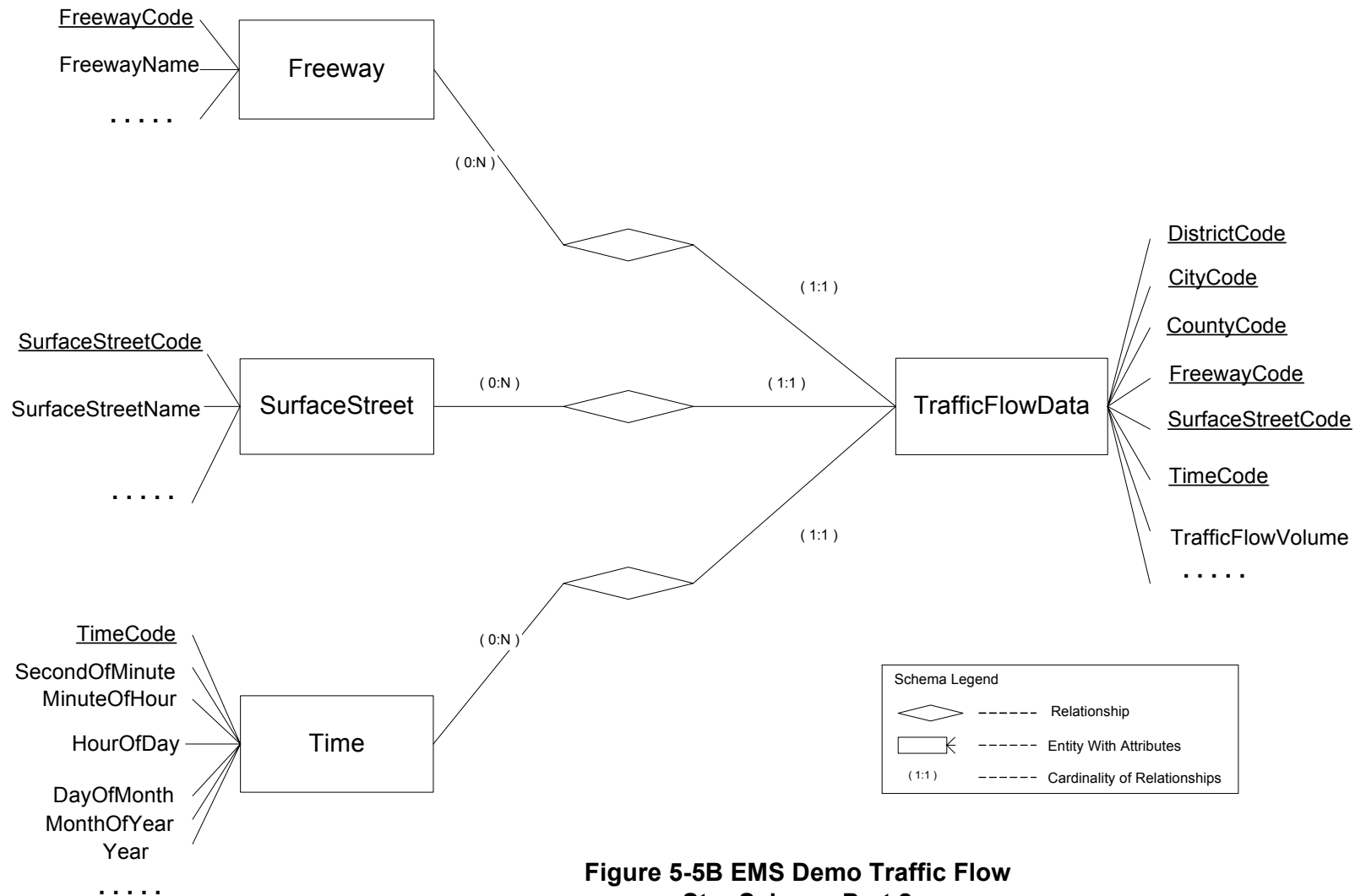
| Entity            | Description                                                                  | Attributes                                                                                                                                                                  | Identifier                                                                                      |
|-------------------|------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| District          | Dimensional information for traffic location with district perspective       | <u>District Code</u> , District Name                                                                                                                                        | District Code                                                                                   |
| City              | Dimensional information for traffic location with city perspective           | <u>City Code</u> , City Name                                                                                                                                                | City Code                                                                                       |
| County            | Dimensional information for traffic location with county perspective         | <u>County Code</u> , County Name                                                                                                                                            | County Code                                                                                     |
| Freeway           | Dimensional information for traffic location with freeway perspective        | <u>Freeway Code</u> , Freeway Name                                                                                                                                          | Freeway Code                                                                                    |
| Surface Street    | Dimensional information for traffic location with surface street perspective | <u>SurfaceStreet Code</u> , SurfaceStreet Name                                                                                                                              | SurfaceStreet Code                                                                              |
| Time              | Dimensional information for time perspective                                 | <u>Time Code</u> , SecondOfMinute, MinuteOfHour, HourOfDay, DayOfMonth, MonthOfYear, Year                                                                                   | Time Code                                                                                       |
| Traffic Flow Data | The traffic flow volume for every potential perspective views                | <u>District Code</u> ,<br><u>City Code</u> ,<br><u>County Code</u> ,<br><u>Freeway Code</u> ,<br><u>SurfaceStreet Code</u> ,<br><u>Time Code</u> ,<br><br>TrafficFlowVolume | Composite of District Code, City Code, County Code, Freeway Code, SurfaceStreet Code, Time Code |
| Incident Data     | The incident information for every potential perspective views               | <u>District Code</u> ,<br><u>City Code</u> ,<br><u>County Code</u> ,<br><u>Freeway Code</u> ,<br><u>Time Code</u> ,<br><br>IncidentLevel<br>IncidentDescription             | Composite of District Code, City Code, County Code, Freeway Code, Time Code                     |

A severe incident on a freeway could cause dramatic capacity reduction and thus traffic congestion, which will slow down the clearance process. Thus, it would be beneficial to re-route the traffic that is heading toward the incident site to alleviate the congestion as much as possible. For example, the following rerouting strategy could be applied:

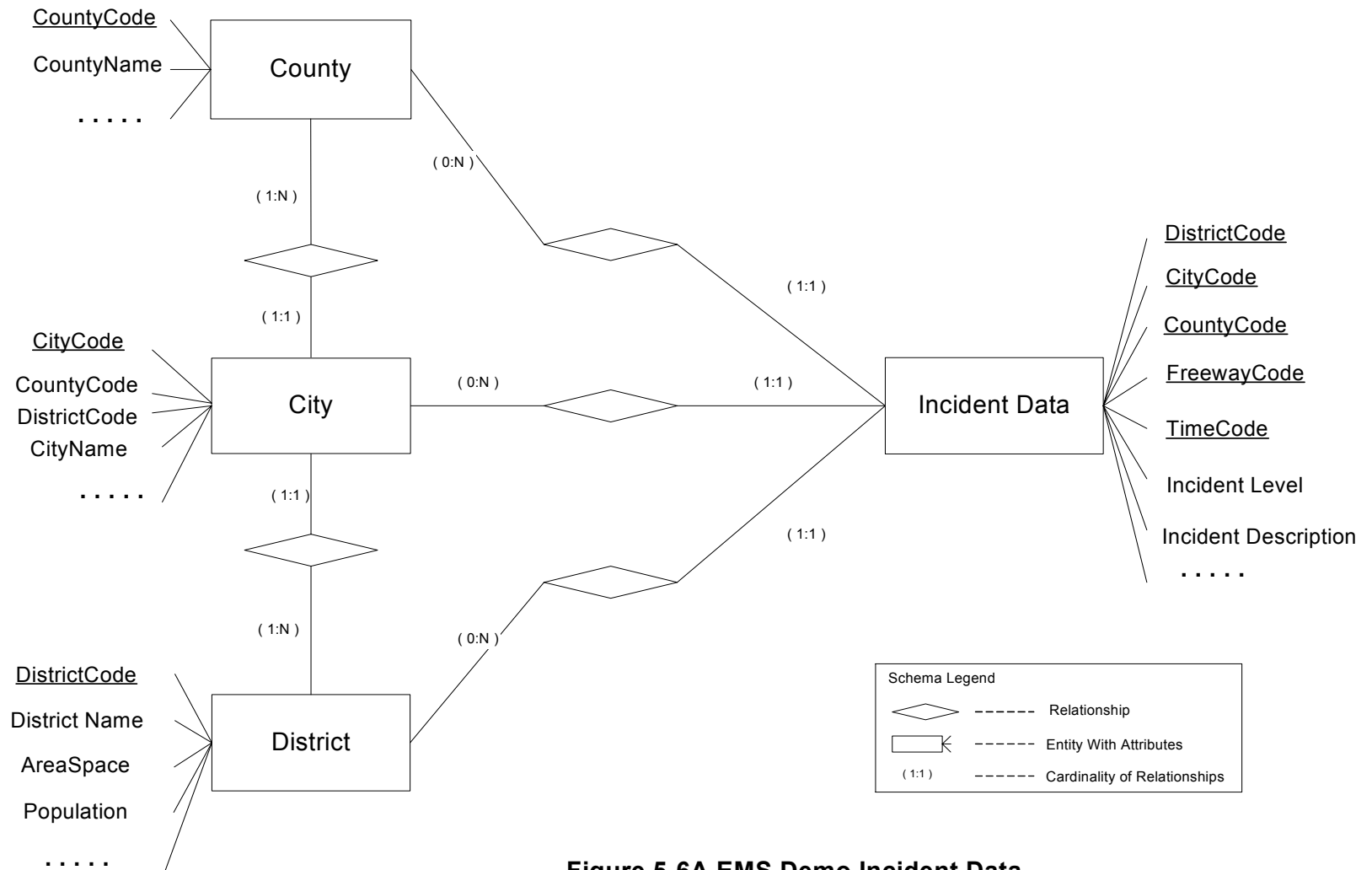
1. The immediate upstream on-ramp from the incident site will be closed. The upstream traffic flow will be advised to go through the surface streets to the next on-ramp.



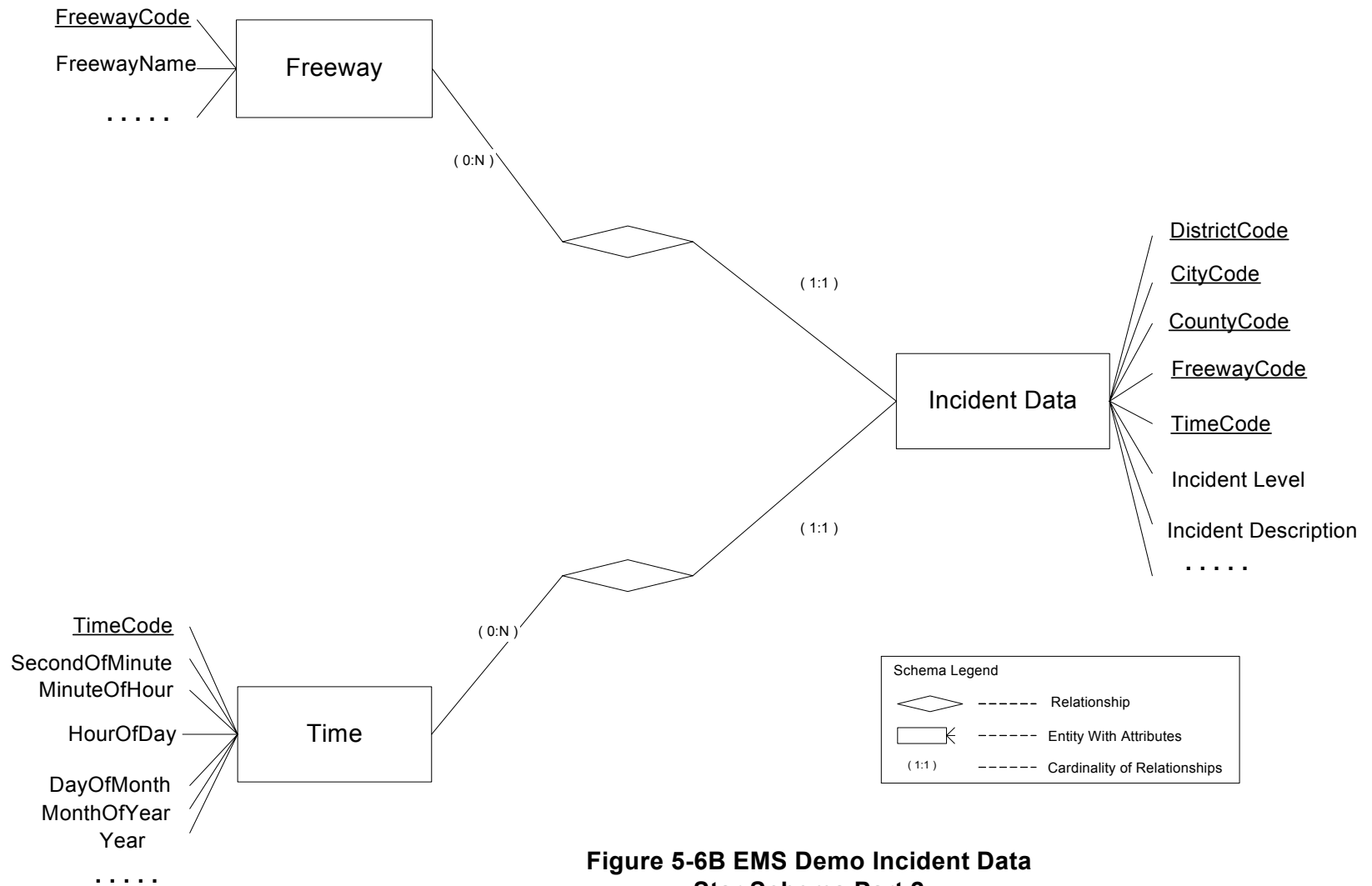
**Figure 5-5A EMS Demo Traffic Flow  
Star Schema Part 1**



**Figure 5-5B EMS Demo Traffic Flow  
Star Schema Part 2**



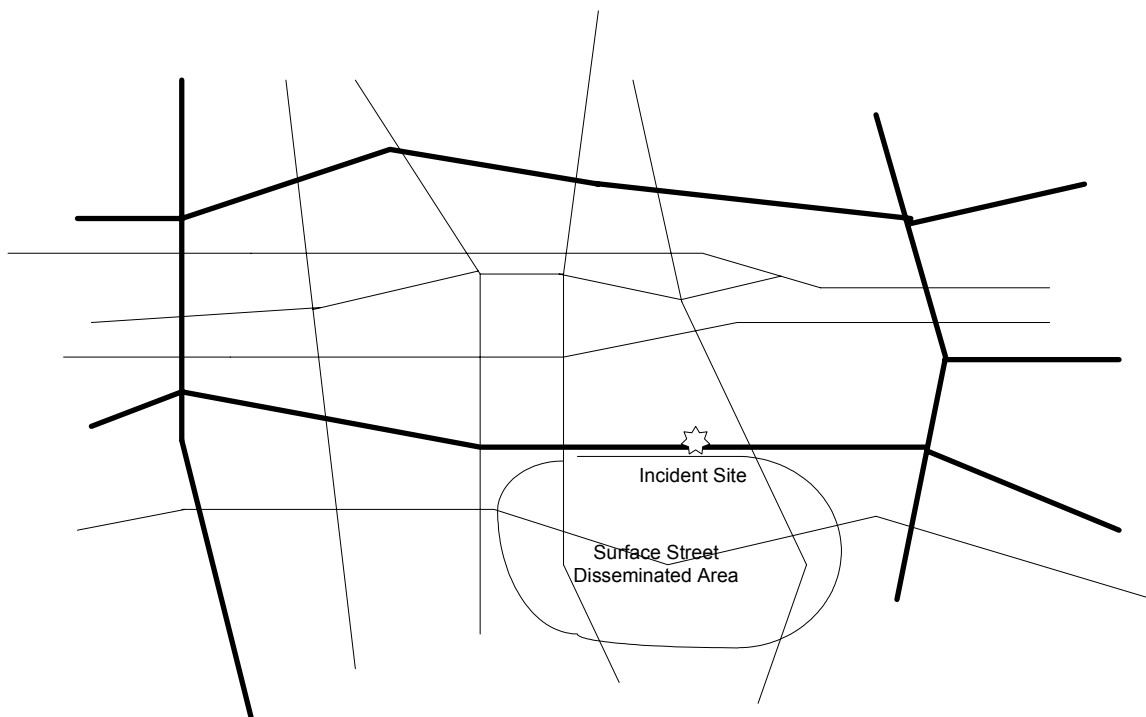
**Figure 5-6A EMS Demo Incident Data  
Star Schema Part 1**



**Figure 5-6B EMS Demo Incident Data  
Star Schema Part 2**

2. The route that is suggested to the drivers will be the one that will provide the best system performance.

In this demo project, our focus is to investigate the proper format to store the traffic data in the CDW, the data organization of the stored data in the CDW, the inquiry functionality of the CDW, and the communication of the rerouting plan to the TWC. For this purpose, a standard System-Optimal traffic assignment model is used to find the rerouting path. A more advanced model that accounts for the traffic dynamics will be a good candidate project for future research. Figure 5-7 shows the affected surface street area applied for the redistribution strategy.



**Figure 5 - 7 Traffic Flow Rerouting Plan Strategy**

## **System Components**

To accomplish the tasks of CDW design concept as described in Section 3, the following components (also see Figure 5-8) are necessary in an incident management system:

- Identify an incident and its location;
- Obtain real-time highway traffic data from loop detectors in a particular area;
- Obtain historical traffic data for surface streets in the area around the incident site;
- Extract other relevant data from the database, for instance, the network structure, the network capacity, etc.
- Execute a program that computes the rerouting plan based on the information obtained; and
- Communicate the rerouting plan to field devices (such as a DMS).

Each of these components is described in detail in the following. Note that the demo is an initial simple implementation of these components, whose functionalities should be enhanced in future implementations.

## **Report An Incident**

The Report Incident tab is used to report the location and severity of an incident. The user (TMC) needs to enter and submit the necessary information, the next screen will show graphically the incident location and provide the user with choice of running re-routing algorithm.

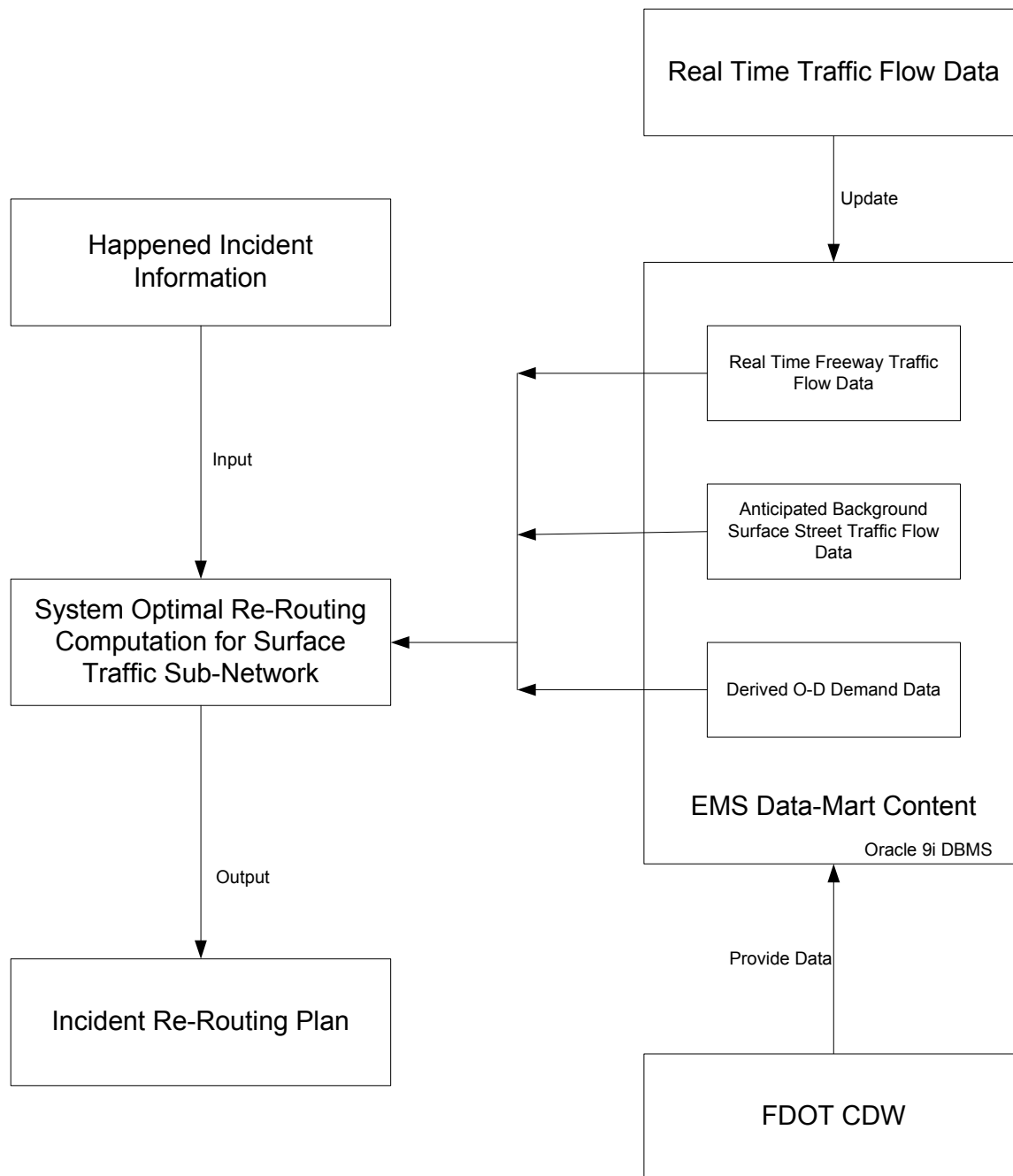
## **System Optimal Re-Routing**

System Optimal (SO) traffic assignment is used to find the alternative paths. The traffic pattern under SO provides us with the lowest total travel time in the area. This assumption usually is not used for transportation planning, since under normal circumstance, travelers will have enough information and knowledge about the network and they would choose the routes

that are best for them (and usually not the best for the system). However, for the case of incident management, when an incident occurs, it causes network congestion and travelers in the affected area usually do not have enough information about the network condition to choose the best routes for themselves. Furthermore, under this scenario, the traffic management agency are more concerned with trying to get the best system performance so that the capacity of the network can be best utilized and the traffic condition brought back to normal as quickly as possible, thus SO becomes a reasonable assumption.

The System-Optimal logic is applied for devising the re-routing strategy. Involved in the computation are real-time traffic data from the upstream freeway link, which is the amount of traffic to be re-routed, the surrounding surface street network, and the existing flow on the surface street links, which is used as background traffic in the assignment process. These data need to be extracted by database inquiries.

The rerouting planning service will be executed as Oracle DBMS server application. A pre-compiled C++ program for the SO computation [8] is used and merged to the Java based EMS web application.



**Figure 5.8 Incident Rerouting Plan Data Flow**

### **Other Functions Related to Incident Management**

In addition to the basic functionality, the following functions for query incident information are also implemented in the demo.

- Query incidents by time:
- Indicate the time period you are interested in, and the query returns the detailed information (location, severity, time) of incidents that happened during that time period. Incident information summary by location:
- Indicate the criterion you would like to use (for instance freeway), the query returns the total number of incidents that happened in each of the locations that belongs to that category (i.e., each freeway). The user can then click on the number associated with each entry and get detailed information about those incident.
- The same functionality is available for query with regard to city, county, and district.

### **EMS Demo Architecture**

The server-client communication architecture mode has been adopted to build the EMS demo system. Figure 5-9 shows the EMS demo architecture.

On the server side, a Dell Dimension 2500 Dual CPU machine has been used as the platform for the Oracle DB server, demo application server and web server. Oracle 9i has been adopted as the DBMS infrastructure. The JDBC API [6] has been adopted as the SQL communication layer between the application server and the Oracle DBMS. The application server is built based on Java Servlet Engine via JSP [9]. The web server is the Apache HTTP server.

On the client side, the user interface based on HTML has been adopted and the general web browser is the running environment.

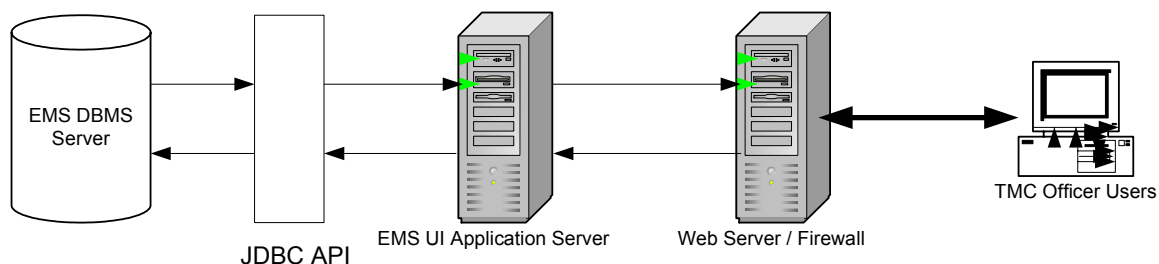
#### **5.4.4 User Interface for EMS Demo**

Four user interfaces (UI's) are developed:

- Current Freeway Network Traffic Flow Frame
- Traffic Flow Update Frame
- Incident Report Frame
- Incident Enquiry Frame

The traffic network UI will show the freeway network and the real time traffic flow associated with each path. The traffic flow update UI will allow the user to change the freeway traffic flow volume manually. The modified traffic flow value will be updated in the Oracle DBMS in real time. The incident report UI will let the user report the incident to the EMS. The EMS will records the incident in the Oracle DBMS and will

provide a rerouting plan if available. The incident enquiry UI will let the user query the



**Figure 5 - 9 EMS DEMO Architecture**

historical incident data based on various analysis purpose. The user can select the time period of the incident happening and the analysis criteria. More details of the demo are provided in Appendix B.

## 5.5 Future Research Tasks

The EMS scenario was chosen as an example to illustrate a prototype CDW function because of the interest of the FDOT in the EMS aspect of freeway operations. Before the

prototype can be developed further the project team must gain a much deeper understanding of EMS operations through consultation with appropriate FDOT and EMS staff.

A further research effort is needed to investigate the efficient way to link the traffic location data with GIS information naturally.

### **Chapter 5 Reference:**

- [1]: Traffic Incident Management Enhancement Program, “TIME Blueprint and Evaluation Version 2.0-Release”, <http://www.uwm.edu/Dept/CUTS/time/>, Access in March 2003.
- [2]: Integrated Incident Management Systems (IIMS) of New York State Department of Transportation, <http://www.dot.state.ny.us/reg/r11/iims/index.html>, Access in March 2003.
- [3]: PB Farradyne, “Traffic Incident Management Handbook”, Federal Highway Administration Office of Travel Management, [http://www.itsdocs.fhwa.dot.gov/jpodocs/rept\\_mis/@9201!.pdf](http://www.itsdocs.fhwa.dot.gov/jpodocs/rept_mis/@9201!.pdf), Access in March 2003.
- [4]: Randolph W. Hall, Yatri H. Mehta, “Incident Management: Process Analysis and Improvement”, Center for Advanced Transportation Technologies, Dept. of Industrial and Systems Engineering, and University of Southern California.  
<http://www.path.berkeley.edu/PATH/Research/presentations/hall.pdf>, access in March 2003
- [5]: “Database will Improve Management of Incident Response Data,” TRAC Research Review, August 1994, Washington State Transportation Center, University of Washington.
- [6]: Todd M. Thomas, ( 2002 ), “Java Data Access: JDBC, JNDI, and JAXP”, M&T Books, ISBN 0-7645-4864-8
- [7]: *Ulka Rodgers*, ( 1999 ), “Oracle: A Database Developer’s Guide”, Prentice-Hall, Inc., ISBN 0-13-841420-3
- [8]: Datex Definition, <http://www.vdot-itsarch.com/nova/docs/comm.pdf>, Access in March 2003
- [9]: Jason Hunter William Crawford, (January 15, 2001) “Java Servlet Programming, 2nd Edition” O'Reilly & Associates, ISBN

## **Appendix A**

### **RESEARCH NEEDS STATEMENT**

#### **Development of a Central Data Warehouse for ITS Data**

#### **Phase II: Proof of Concept**

## RESEARCH NEEDS STATEMENT

### I. PROJECT TRACKING NUMBER

*To be assigned by FDOT research staff.*

### II. PROBLEM TITLE

Development of a Central Data Warehouse for ITS Data, Phase II: Proof of Concept

### III. CLASSIFICATION

*Using the attached lists, please classify this proposed work by PRIMARY Mode and Function.*

#### Mode

Highway

#### Functional Areas

Management

Information Technology

Planning

Policy

Systems

Statistics

Operations

Traffic Operations

Intelligent Transportation Systems

Emergency Preparedness / Disasters

### IV. RESEARCH NEED STATEMENT

*Describe why, this research is needed. Be explicit: What is the intended product? How will the product be used and by whom? Using the classification list (provided), identify all other areas that will be affected by this research.*

The Statewide Intelligent Transportation Systems (ITS) Architecture contains a central data warehouse (CDW) component in its Statewide Concept of Operations. The principal function of a CDW is to collect and store statewide information from many data sources and make it available for use by different applications. The CDW will operate in the ITS center interacting with other traffic control centers, transportation management centers, emergency management centers, weather centers, disaster management centers, motor carrier control centers and so on. The CDW will allow data from these separate and disparate data sources to be aggregated.

The main reasons for the development of a CDW are:

- To provide more and better information in managing and operating the transportation system: Archived operations data can be used to predict when and where problems may occur again, as well as helping to evaluate alternative strategies for preventing or mitigating the problem.

- To maximize cost-effectiveness of the data collection infrastructure: Data archiving permits transportation agencies to optimize their investments in data collection infrastructure by re-using the same data for numerous transportation planning, design, operations and research needs.
- To reduce data acquisition costs: Data archiving is significantly less expensive than having a planning or design group re-collect even a small percentage of the data using manual methods or special studies.
- To conform to established business practice in other industries: The retention and analysis of operational data is an established practice in most competitive industries that use data to manage their business activities. For example, the retail sales industry uses data warehouses full of customer transactions and inventories to better understand the basics of supply and demand in numerous markets around the world.

Recognizing the necessity for the development of this ITS component in Florida, the FDOT initiated a research project in 2001 (Contract BC-354-61) to perform an ITS central data warehouse feasibility study. The final report for this project has provided a high level conceptual framework as a roadmap to guide future development efforts.

This research needs statement describes a “proof of concept” study to constitute Phase II of the proposed CDW development. Because of the potential widespread use of the CDW, virtually all operational offices of the Department, both Central Office and Districts will be affected, certainly at least the ITS Program Office along with Traffic Operations; Construction; Maintenance; Information Resources; and possibly Planning, Legal and Contracts as well. This clearly will impact policy issues, so the Executive Committee will ultimately be involved.

Other state and local agencies will be affected as well, particularly the State Technology Office; Florida Highway Patrol; Metropolitan Planning Organizations; local public works/traffic, public service, and emergency management, and possibly others.

## **V. LITERATURE SEARCH SUMMARY**

*The submitter is strongly encouraged to conduct a literature search in order to avoid duplicating with other current or past research. At a minimum, searches should be conducted on TRIS online (<http://ntl.bts.gov/tris>) and the Research in Progress database (<http://rip.trb.org/search>).*

Because of the importance of archived data to statewide ITS applications, several states have embarked on the development of their own CDW systems. Because of differences in the ITS applications and communication resources among states, there is minimal duplication of effort involved. Most states are following the national ITS architecture, as will the project proposed in this document.

The two principal keywords for searching the literature on this topic are “data warehouse” and “ADUS” (Archived Data User Services). Of the 40-50 citations produced by a search, about a dozen entries in the literature apply directly to this project. The majority of the

citations describe case studies of data warehouse activities in other states. Some references present principles and guidelines for warehousing traffic data. A standard for ADUS has recently been developed by an ASTM committee. While all of the literature in this category will be useful in developing Florida's ITS data warehouse, none of the results of the literature search suggest any duplication of the actual research effort proposed for this project.

Some coordination will be required between the proposed project and a parallel effort now underway in Florida to develop a data warehouse for a specific Traffic Management Center in the Orlando area. The Department views the Central Florida project as an integral part of the plan to establish a statewide facility, and not as a competing or duplicative effort.

## **VI. RESEARCH OBJECTIVE**

*A statement of the specific research objective, defined in terms of the expected final product, that relates to the Research Need Statement in IV above. Define specific tasks necessary to achieve the objective.*

The primary objective of Phase II will be to implement a prototype of the system described in the Phase I report that will demonstrate the potential of a fully deployed CDW. The scope of the project is described in terms of issues to be addressed and tasks to be carried out.

### **Issues to be Addressed**

There are two categories of issues to be addressed: data warehouse issues and transportation domain issues. The data warehouse issues include:

- Timely access to dynamic information,
- Maintenance of warehoused data,
- Automatic detection of relevant updates in the underlying data sources,
- Development of support tools for efficient source adapter generation and
- Efficient support for "ad-hoc" user queries.

The transportation domain issues include:

- Storage of Data: . The amount of dynamic traffic data (volume, speed, occupancy, etc.) that are collected everyday is overwhelming and the space requirement is almost prohibitive if all the data are going to be stored. Research issues here include
  - what and how much historical data need to be stored;
  - For the data that are not stored, what is the best way to keep the equivalent information
- Use of the data warehouse in traffic analysis and transportation planning: Several concepts need further research including,
  - Ability to support the "what-if" analysis
  - Use of real-time data for planning applications,

- Time-dependent origin-destination (OD) matrix estimation,
- Calibration and testing of simulation models,
- Prediction of travel time,
- Time-dependent shortest path search using a travel-time curve, and
- Route guidance.

### **Tasks to be Carried Out**

The following tasks should be included in the definition of the project scope:

- Task 1. Determine the functional requirements of the CDW prototype: The Phase I report suggested several items that could be included in the functional requirements. These items must be discussed with FDOT district and central office personnel and with local agency personnel to arrive at a final set of functional requirements. The *ASTM Standard Guide for Archiving and Retrieving ITS Generated Data* should be followed in the development of the functional requirements. This document sets forth eleven guiding principles should that be adhered to in the development of a CDW:
  1. Reliance on user needs and requirements processes
  2. Get archived data from other centers
  3. Integrate selected other transportation data including roadside data collection
  4. Manage the archive to account for data quality
  5. Provide security for the ADMS
  6. Specify and maintain metadata to support the ADMS
  7. Manage the Interfaces of the Archive Data Administrator
  8. Interact with other archives and monitor other standards
  9. Process user requests for data
  10. Support analysis of the archived data
  11. Prepare data for government reporting systems
  12. Interface with statewide TMC Software Library System.
- Task 2. Establish the institutional framework for the system: Several key questions about the relationships between the CDW and the FDOT Enterprise Information System must be addressed in this task. Issues of who will have access and at what level with they have it are also of concern. One of the biggest concerns for the data sources will be the amount of effort required on their part to participate in the CDW.
- Task 3. Develop the hardware and software configuration of the prototype: The configuration details will depend heavily on the results of Tasks 1 and 2. For example, if it is determined in Task 2 that the CDW should reside physically within the EIS, the hardware configuration will be totally different than it would be if the prototype were to be developed separately with the intent to integrate the functions at the end of Phase II.
- Task 4. Procure the hardware required for the prototype.

- Task 5. Procure or develop the software for data communications and management.
- Task 6. Implement the prototype on an operational environment to demonstrate its value to the ITS community
- Task 7. Document the results of the project with specific recommendations for a future course of action to expand the CDW both functionally and geographically. The results of this final task should lead directly to the operational implementation phase

## **VII. ESTIMATE OF FUNDING, RESEARCH PERIOD and EQUIPMENT**

### **Recommended Funding:**

The estimated budget requirement is \$250,000

### **Research Period:**

*An estimate of the number of months of research effort, including three months for preparation of a draft final report, necessary to the accomplishment of the objectives in Section VI above.*

All work should be completed within a 24 month period.

### **Equipment:**

*The Department will purchase, lease or rent equipment that is specific to the subject research. Justification must be provided. The Department will NOT purchase, lease or rent "common" equipment such as personal computer workstations, laptops, digital cameras, etc.*

As a proof of concept study, this project will require some equipment. The detailed equipment needs, cannot be determined until the final equipment configuration is decided in Task 3. All equipment required to conduct the project will be for implementation by FDOT and others, not the researchers.

## **VIII. URGENCY, PAYOFF POTENTIAL, AND IMPLEMENTATION**

*Statements about the urgency of this research and the potential payoff (couched in benefit/cost terms if at all possible) from achievement of project objectives should be given. A statement should be included that further describes the anticipated product(s) of the research. The anticipated steps necessary for implementation of the research product should also be delineated.*

As a component of the FDOT Enterprise Information System, Archived transportation system data will play an important part in the operation of all ITS centers in Florida.

The research aspect of this undertaking will end with the proposed Phase II project. The intent is that the Phase II results will ultimately lead to one or more consultant contract(s) to undertake the formal architecture development, systems engineering and implementation tasks to deploy a fully-functional CDW.

There are many benefits to be derived from systems that archive and retrieve ITS generated data, including: (a) providing information to better manage and operate the transportation system, (b) maximizing cost-effectiveness of the data collection infrastructure, (c) being less expensive than typical data collection, and (d) supporting governmental reporting systems. Since the FDOT is heavily involved in the development and operation of centers with ITS capabilities, it will be the main recipient of all of these benefits.

## **IX. PERSON(S) DEVELOPING THE PROBLEM**

*A statement providing the specifics (name, title, affiliation, address, telephone number, e-mail address) of the person(s) who developed the problem in all its detail.*

Kenneth G Courage, Professor Emeritus  
University of Florida Transportation Research Center  
365 Weil Hall  
Gainesville, FL 32611  
352-392-7575, ext 1452  
kcourage@ufl.edu

Liang Y. Hsia, P.E., ITS Engineer Administrator  
Suite 327 F, Rhyne Building, 2740 Centerview Dr.  
Mail Address: 605 Suwannee St. MS 90  
Tallahassee, Florida 32399-0450  
Phone: (850) 410-5615 Suncom 210-5615  
Fax: (850) 410-5502 Suncom 210-5502  
E-mail: [liang.hsia@dot.state.fl.us](mailto:liang.hsia@dot.state.fl.us)

## **X. PROBLEM MONITOR**

*A statement of the specifics (name, title, affiliation, address, telephone number, e-mail address) of the person who will be assigned by the Administrator or Committee submitting this problem to monitor the research, if programmed, from inception to completion. A statement must also be included on how others affected by this research will be included in final Scope development and project progress monitoring.*

Liang Y. Hsia, P.E., ITS Engineer Administrator  
Suite 327 F, Rhyne Building, 2740 Centerview Dr.  
Mail Address: 605 Suwannee St. MS 90  
Tallahassee, Florida 32399-0450  
Phone: (850) 410-5615 Suncom 210-5615  
Fax: (850) 410-5502 Suncom 210-5502  
E-mail: [liang.hsia@dot.state.fl.us](mailto:liang.hsia@dot.state.fl.us)

Because of the broad impact of this project, it is anticipated that an *ad hoc* advisory group will be configured to oversee the project and to evaluate its progress.

## **XI. PROPOSED RESEARCHER**

*If the submitter recommends a researcher for this project, the proposed researcher's name, title, affiliation, address, telephone number, and e-mail address should be included in this section.*

The recommended Principal Investigator is:

Kenneth G Courage, Professor Emeritus  
University of Florida Transportation Research Center  
365 Weil Hall  
Gainesville, FL 32611  
352-392-7575, ext 1452  
kcourage@ufl.edu

Prof. Courage was the Principal Investigator for the Phase I feasibility study.

## **XII. DATE AND SUBMITTED BY**

Liang Y. Hsia, P.E., ITS Engineer Administrator  
Street Address: Suite 327 F, Rhyne Building, 2740 Centerview Dr.  
Mail Address: 605 Suwannee St. MS 90  
Tallahassee, Florida 32399-0450  
Phone: (850) 410-5615 Suncom 210-5615  
Fax: (850) 410-5502 Suncom 210-5502  
E-mail: [liang.hsia@dot.state.fl.us](mailto:liang.hsia@dot.state.fl.us)

## **Appendix B**

### **Description of the CDW Demonstration CD-ROM**

## **Appendix B**

### **Description of the CDW Demonstration CD-ROM**

#### **Purpose**

In this section, we will talk about the CD-Demo, which is contained in the Appendix B CD-ROM delivered with the report. The reader could use the MS Internet Explorer or Netscape for an animated view of the content of the Demo. This CD-Demo will show the basic concept of the CDW Emergency Management Service (EMS), demonstrate the operation process of the EMS demo, show a hypothetical example traffic rerouting plan and represent the incident analysis by querying historical incident data.

The run-time, internet-based client-server EMS Demo is built in the UF campus internet domain with the constrained access. Because of the limitation of the functionality that the local HTML files and JavaScript could provide, only static querying samples of the EMS demo can be presented on the CD-ROM.

#### **File Hierarchy Structure of the CD-ROM**

The CD-Demo only uses the HTML files stored in the attached CD. No additional on-line resources or plug-in components are needed in order to run the Demo. The general internet web browser is necessary to have the full functional view of the Demo.

The root directory in CD is Web\_UI with the main entry file of “index.html”. Clicking this file will launch the Demo. The Demo will show up in the web browser and you can navigate through the demo by clicking on any of the functional buttons and hyperlinks within the demo. Figure B-1 shows the file structure of the CD-ROM.

## **Navigating the Demo**

There are four main frames in the Demo

- Current Traffic Network,
- Traffic Flow Update,
- Incident Report and
- Incident Enquiry.

These frames are accessed by clicking the corresponding hyper-links at the top of the screen.

For each frame, a pop-up description window will be shown to describe the detailed functionality of the triggered frame. Also a description hyper-link in the upper-right corner of each frame can be clicked and a similar description window will be launched. The description windows will describe the functionality of the frame step by step and will indicate the next stage of the operation. The general functionality of each frame is described as follow:

- The *Current Traffic Network* frame will show up the traffic volume for each link of the demo freeway network. The link is identified by the unique starting node number and ending node number.
- The *Traffic Flow Update* frame will let users to modify the traffic flow volume of the specific traffic link by typing in the new volume in the corresponding text box.
- The *Incident Report* frame will let user to report the incident by typing in the incident information. The incident site map will be shown up according to the user input. A rerouting plan will be computed and shown up with the surface street network map.
- The *Incident Enquiry* frame will support incident analysis by querying historical incident data. The time period must be selected first from the drop-down menu showing the starting and ending times. Then the analysis query is made by selecting one of the five analysis criteria The details of the five query criteria are as follows:
  - *Based on District*---The incidents that have happened in each district for the specific time period will show up. More detailed incident information can be obtained by clicking the hyper-link of the incident. More detailed district information can be obtained by clicking the hyper-link of the district.
  - *Based on City*---The incidents that have happened in each city for the specific time period will show up. More detailed incident information can be obtained by clicking the hyper-link of the incident. More detailed city information can be obtained by clicking the hyper-link of the city.

- *Based on County*---The incidents that have happened in each county for the specific time period will show up. More detailed incident information can be obtained by clicking the hyper-link of the incident. More detailed county information can be obtained by clicking the hyper-link of the county.
- *Based on Freeway*---The incidents that have happened in each freeway for the specific time period will show up. More detailed incident information can be obtained by clicking the hyper-link of the incident. More detailed freeway information can be obtained by clicking the hyper-link of the freeway.
- *Based on Detail*---The detailed information about all the incidents that have happened for the specific time period will show up.

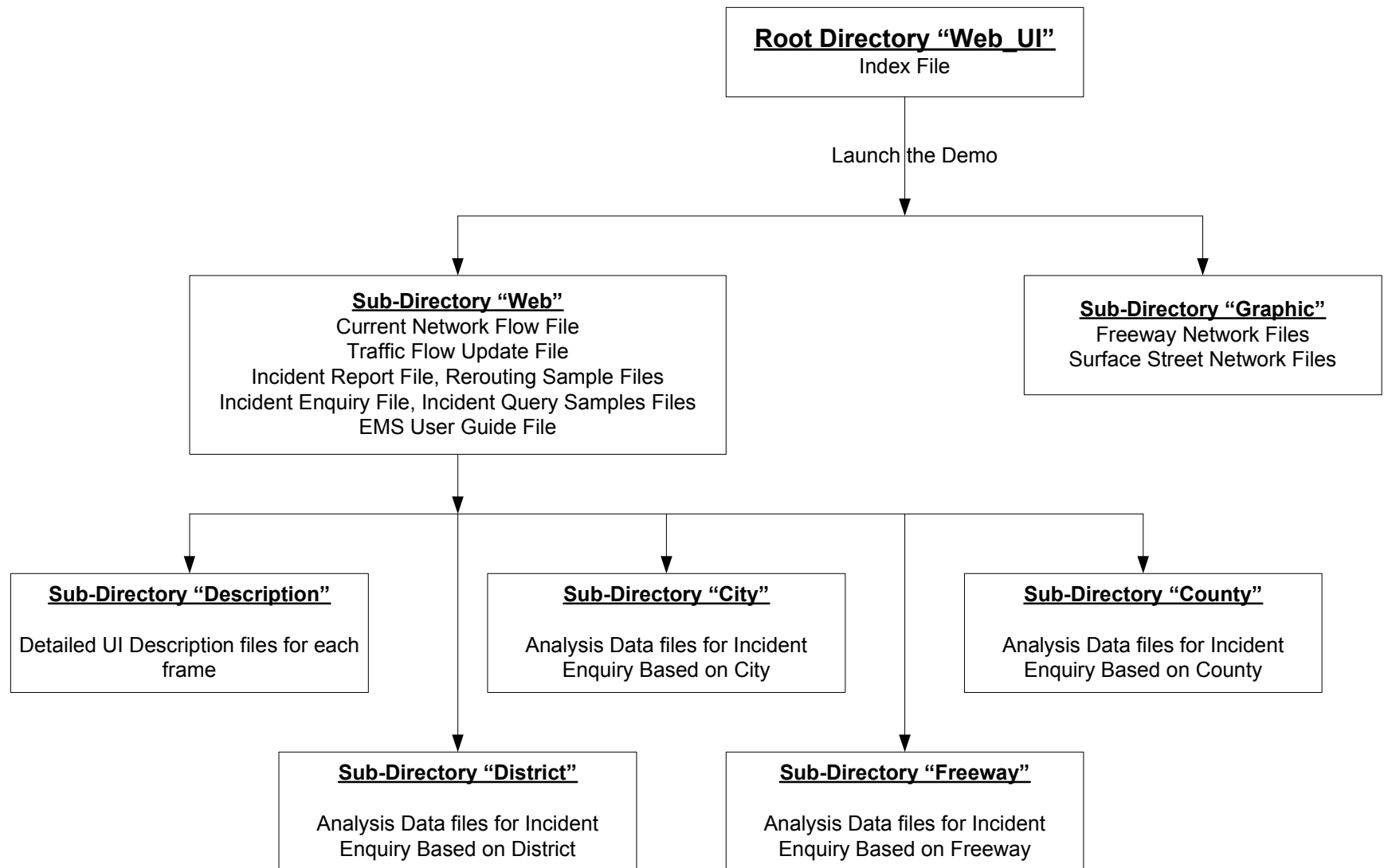


Figure B-1 File Hierarchy of the Demo CD