National ITS Architecture
Theory Of Operations

Prepared by:

Architecture Development Team
Lockheed Martin
Odetics Intelligent Transportation Systems Division

Prepared for:

Federal Highway Administration
US Department of Transportation
Washington, D. C. 20590

December 1999

Table of Contents

Table of Figures

Table of Tables

1 Introduction

1.1 Purpose

This document presents a high-level, narrative, technical description of the operations of and institutional issues associated with the National Intelligent Transportation System (ITS) Architecture. It is intended to aid those public and private sector infrastructure operators, engineers, designers, etc., who want a better technical and institutional understanding of how a deployed ITS operating under the National ITS Architecture framework would operate.

1.2 Scope

The document is based on and complements the other National ITS Architecture documents: primarily the Logical Architecture, Physical Architecture, Market Package and Standards Requirements.

A deployed ITS is made up of distinct but interoperating physical *subsystems* that also interoperate with the environment and various users. The Theory Of Operations document identifies the options for *interrelationships* between subsystems and the *functional requirements* of those subsystems to implement user services on the fully functional architecture as it evolves to 20-year ("high end state") architecture deployments. Interrelationships in this document are presented as architecture flows (defined in the Physical Architecture) which are information messages between subsystems and between subsystems and users. Information from the environment that is sensed by subsystems is shown in a similar manner, also using architecture flows. Functional requirements define the processing that subsystems do to issue architecture flows based on architecture flows received in specific sequences from other subsystems, users and sensed from the environment.

1.3 Document Structure

The document chapters are structured as follows:

1. Introduction
2. Executive Summary

An overview of the National ITS Architecture definition, operational concepts, and consensus based choices made in its development, as they pertain to this document.

3. Assumptions

This chapter describes the deployments/implementations used in the Theory of Operations.

4. Operational Overview

Organized by user services, this chapter describes the operation of the National ITS Architecture from a combined user/institutional and system perspective. The system perspective includes the subsystem processing and architecture flow communications required to implement the user services. The user perspective identifies the user interactions with the ITS.

5. Operational and Interoperability Issues

This chapter identifies key operational and interoperability issues and outlines the options that were considered in the National ITS Architecture development and the strategies chosen for addressing them. Issues addressed include: infrastructure based route selection; robustness to spatially different deployments; interjurisdictional issues; predictive modeling of link travel-time and congestion; development of public-private partnership models; and deployment of low cost and ubiquitous wireless communication services.

2 Executive Summary

A deployed *Intelligent Transportation System* (ITS) is made up of distinct but interoperating physical *subsystems* that also interoperate with the environment and various users. The Theory Of Operations document for the ITS Architecture identifies the *interrelationships* (messages) between subsystems and between subsystems and users, and the *functional requirements* of those subsystems to implement user services on the fully functional architecture as it evolves to 20-year ("high end state") architecture deployments. Functional Requirements are the processing that subsystems do to issue and receive messages in specific sequences to/from other subsystems and users. This summary covers three main areas: subsystem functionality, communications channels, and some of the key architecture choices.

2.1 Subsystems

The specific choice of 19 subsystems in the ITS architecture is shown in Figure 1, and represents a partitioning of functions that is intended to capture all expected or likely subsystem boundaries for the present to 20-year future. The intersubsystem boundaries identify the range of possible institutional and message interfaces. The subsystems themselves are composed of Equipment Packages with specific functional requirements that represent the smallest units of ITS that can be purchased and deployed. The character of a subsystem deployment is determined by the specific equipment packages chosen. At the same time, subsystems may be deployed individually or in "aggregations" that will vary by geography and time based on local deployment choices.

ITS subsystems communicate with each other using the communication elements and architecture interconnect channels. In Figure 1 the subsystems are shown as white boxes, the communication channels are shown as lines and the communication elements are shown as "sausages."

The subsystems shown as single entities in Figure 1 are representative of multiple instances of the specific subsystem. For example, several Traffic Management subsystems (TMS) in a region, each with their own jurisdiction, may communicate with each other (and each with their many Roadway subsystems) to implement regional ITS policies. Several deployed subsystems of a given type may be individually operated by local, state, federal, or private institutions. The multiplicity expressed for ITS subsystems extends to the wireline and wireless communication elements as well. In the previous example, the Traffic Management subsystems may communicate with each other using a commercial wireline data Communications Service Provider (CSP), but may have their own dedicated wireline communications elements for data communications with their many Roadway subsystems where sensors and signals are located.

The names of the 19 subsystems of the National ITS Architecture are listed again in Table 1 along with the associated acronyms that are used throughout the documentation.

ITS architecture subsystems have been grouped into four classes where they share common communication elements, deployment and institutional characteristics. The following is a summary of the function of each subsystem.

2.1.1 Center Subsystems

These subsystems have no requirement to be on or adjacent to a transportation facility and thus can be located anywhere. To communicate with other subsystems they need access to wireline communications.

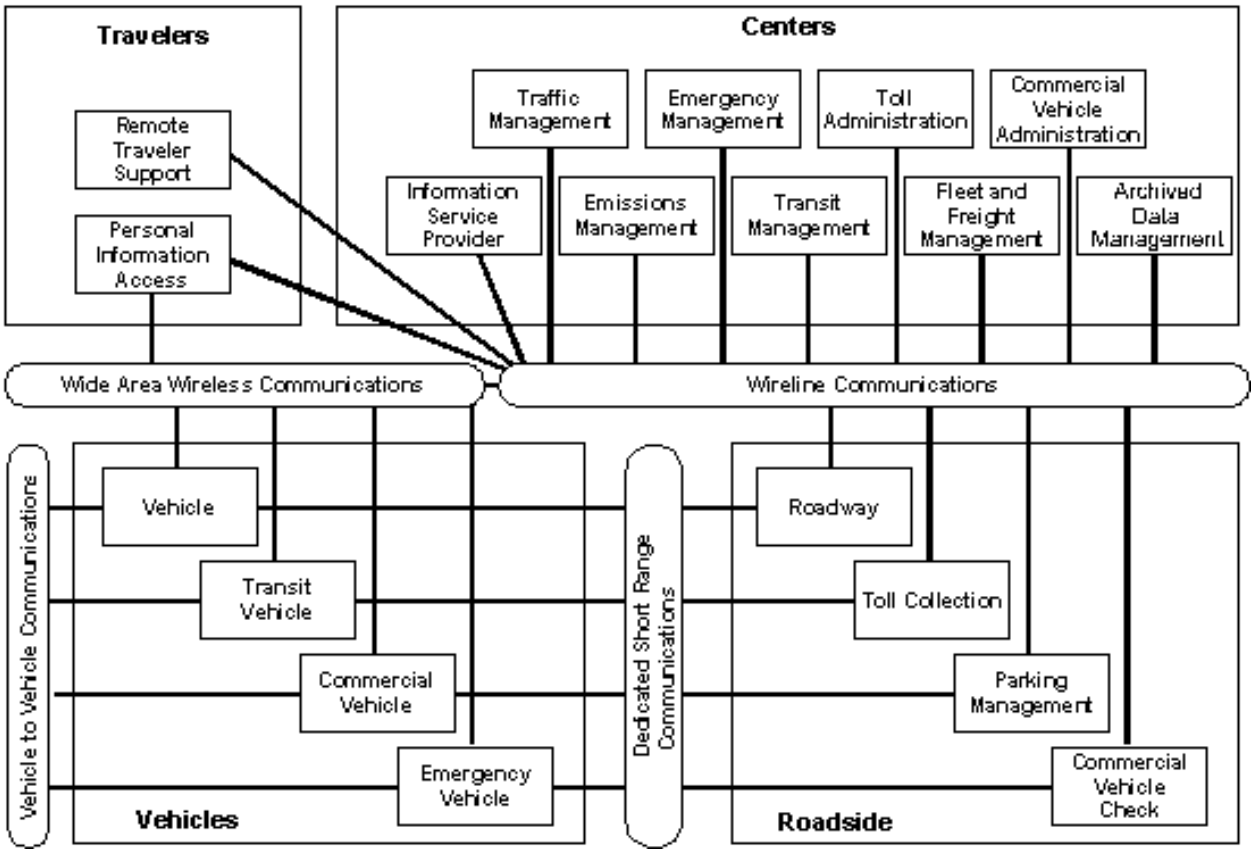

Figure 1. ITS Architecture Subsystems and Communication Elements

Subsystem Acronym	Subsystem Name
ADMS	Archived Data Management Subsystem
CVAS	Commercial Vehicle Administration Subsystem
CVCS	Commercial Vehicle Check Subsystem
CVS	Commercial Vehicle Subsystem
EM	Emergency Management Subsystem
EMMS	Emissions Management Subsystem
EVS	Emergency Vehicle Subsystem
FMS	Fleet and Freight Management Subsystem
ISP	Information Service Provider Subsystem
PIAS	Personal Information Access Subsystem
PMS	Parking Management Subsystem
RS	Roadway Subsystem
RTS	Remote Traveler Support Subsystem
TAS	Toll Administration Subsystem
TCS	Toll Collection Subsystem
TMS	Traffic Management Subsystem
TRMS	Transit Management Subsystem
TRVS	Transit Vehicle Subsystem
VS	Vehicle Subsystem

Table 1. ITS Subsystem Names and Acronyms

1. Commercial Vehicle Administration

Sells credentials and administers taxes, keeps records of safety and credential check data, and participates in information exchange with other commercial vehicle administration subsystems and Commercial Vehicle Operations (CVO) Information Requestors.

2. Fleet and Freight Management

Monitors and coordinates vehicle fleet including coordination with intermodal freight depots/shippers.

3. Toll Administration

Performs back-end operations for Toll Administration.

4. Transit Management

Collects operational data from transit vehicles and performs strategic and tactical planning for drivers and vehicles.

5. Emergency Management

Coordinates response to incidents including Hazardous Materials (HAZMAT).

6. Emissions Management

Collects and processes pollution data and provides demand management input to Traffic Management.

7. Archived Data Management

The Archived Data Management Subsystem collects, archives, manages, and distributes data generated from ITS sources for use in transportation administration, policy evaluation, safety, planning, performance monitoring, program assessment, operations, and research applications.

8. Traffic Management

Processes data and provide basic traffic and incident management services through the Roadside and other subsystems.

9. Information Service Provider (ISP)

ISPs can collect and process transportation data from the aforementioned centers and broadcast general traveler information products (e.g. link times) or deliver personalized information products (e.g. personalized/optimized routing) to individual information requests.

2.1.2 Roadside Subsystems

These subsystems include functions that require convenient access to a roadside location for deployment of sensors, signals, programmable signs, or other interfaces with travelers, vehicles (or possibly freight). Roadside subsystems generally need wireline (or equivalent stationary point-to-point) communications for messages to/from one or more Center subsystems, and possibly also have toll-tag or beacon communications to some or all vehicles passing the specific roadside subsystem deployment.

1. Roadway

Provides traffic management surveillance, signals and signage for traveler information.

2. Toll Collection

Interacts with vehicle toll tags to collect tolls and identify violators.

3. Parking Management

Collect parking fees and manage parking lot occupancy with ISPs.

4. Commercial Vehicle Check

Collects credential and safety data from vehicle tags, determines conformance to requirements, posts results to the driver (and in some safety exception cases the carrier) and records results for the commercial vehicle administration subsystem.

2.1.3 Vehicle Subsystems

These subsystems are installed in a vehicle. The subsystems may support some combination of Vehicle Roadside Communications (e.g. toll-tag or beacon communications) with Roadside subsystems, Wireless Wide Area 2-way or 1-way communications with Center subsystems, and Vehicle-Vehicle communications.

1. Vehicle

Functions that may be common across all vehicle types are located here (e.g. navigation, tolls) so that specific vehicle deployments may include aggregations of this subsystem with one of the following more specialized vehicle subsystems. Includes advanced vehicle safety and operations systems.

2. Transit Vehicle

Provides operational data to the Transit Management Center, receives transit network status, provides enroute traveler information to travelers and provides passenger and driver security functions.

3. Commercial Vehicle

Stores safety data, identification numbers (driver, vehicle and carrier) and last check event data. Provides in vehicle signage for driver pass/pull-in messages.

4. Emergency Vehicle

Provides vehicle and incident status to the Emergency Management subsystem.

2.1.4 Traveler Subsystems

These subsystems represent platforms for ITS functions of interest to travelers or carriers (e.g. commercial vehicle operators) in support of multimodal traveling. They may be fixed (e.g. kiosks or home/office computers using wireline communications) or portable (e.g. a “palm-top” computer using wireless communications) and may be accessed by the public (e.g. kiosks) or by individuals (e.g. personal computers).

1. Remote Traveler Support

These are kiosks for traveler information at public locations (e.g., roadside rest areas, truck stops, transit stops or transit stations) including traveler security functions.

2. Personal Information Access

These are home, office or portable computers for traveler information and emergency requests.

2.2 Communication Element Choices

2.2.1 Wireline and Wireless Communication Elements

The wide area network (WAN) wireline communication element connects center subsystems to roadside,

traveler, and other center subsystems. Augmented by WAN wireless, it may also connect centers to vehicles, or centers to mobile personal computers. These elements can be ITS-dedicated networks or can be privately deployed networks owned and operated by a CSP, where operators of ITS subsystems pay a service fee for connection to and use of the networks for ITS functions and share the network with non-ITS users.

More than one wireline or wireless network used for ITS may coexist in a region, and these networks can be connected (or internetworked) to support ITS message communication between subsystems that are attached to different networks. Thus the "Wireless/Wireline Wide Area Communications" elements shown in Figure 1 may represent networks of interconnected networks. It is expected that the current trend toward ubiquitous internetworking of public and private data networks will continue. This will enable inter-subsystem messaging across local, regional and national ITS subsystems. What the Internet is evolving to (as security, reliability, and performance issues of today's Internet are addressed) has been referred to as the "National Information Infrastructure" or "NII".

In the near term many communication elements will be dedicated, as they primarily are today (e.g. fiber systems for traffic surveillance or Specialized Mobile Radio (SMR) systems for transit/commercial/emergency fleet management). As commercial data networks are deployed, interconnected, and mature, and the cost of access and use of these private data networks drops, we expect more wireline and wireless networks for ITS to be supplied by CSPs. Transitions from private data networks to commercial data networks will vary by region if and when it becomes practical and economical.

Wireless communication elements can also be one-way (e.g. fm-subcarrier) broadcast.

A key concept in the ITS architecture is that only ITS subsystems do ITS processing and communication elements are required only to pass messages from one subsystem to another. Communication elements are thus "commodities" that can be considered separate from ITS subsystems. The benefit of this separation is that the investment in ITS subsystems can be made relatively stable enabling evolution and adaptation to rapidly evolving communication technologies. The architecture supports requirements for standard communication interfaces at several levels (depending on the specific interface) between the communication elements and the ITS subsystems. If these standards are *open* (as opposed to *proprietary*), then multiple vendors will be able to provide functionally equivalent communication modules and the cost of acquiring these communication modules will be contained by competition.

2.2.2 Dedicated Short Range Communications (DSRC)

This communication element represents a direct communication path between vehicles (e.g. toll tags) and roadside equipment (e.g. "beacons").

Toll Tag communication elements use a low-cost tag mounted on the vehicle that communicates either one-way or two-way to a roadside subsystem. One-way systems may simply send the tag id to the roadside. Two-way system can also decrement a tag "cash value" or "purse" by the roadside subsystem; or optionally indicate toll road entry location on the tag which can be read on exit for toll computation. The tag may have a simple user interface to identify values stored in the tag or recent transactions to the traveler.

Beacon short range DSRC elements allow communication between a Vehicle subsystem and either a Roadway subsystem or, along with an associated Wireline communication element, a Center subsystem. The DSRC can be internetworked with the Wireline network in a similar way as the wireless wide area communications network.

Beacon communication capability will be deployed where there is need for a localized exchange of data (e.g. tolling) or where the cost of the beacon communication equipment is less than the cost of using the equivalent Wireless communications infrastructure. Examples of beacon based systems are toll and parking payment operations, CVO checking operations, in-vehicle signing and urban fixed route transit vehicle to transit center communications.

2.2.3 Vehicle-Vehicle Communications

This communication element represents the direct communication path between adjacent vehicles. The deployment of this communication is necessary for some Advanced Vehicle Control Systems (AVCS) services, such as high density platooning of vehicles.

2.3 Operational Options and Architecture Choices

Emergency Notification Location Determination

The location of the emergency calling mobile traveler must be communicated to the Emergency Management subsystem (EM) along with the emergency request message from the traveler. The location could be determined by the travelers' mobile subsystem or can be done by the WAN wireless communication element. The architecture has chosen to support the former, and allow an easy evolution to the latter if the technology matures.

In the architecture the location of the Driver or Traveler sending a data emergency assistance request is determined by the mobile subsystem initiating the message. Although expensive, the technology for locating mobile subsystems with self–contained equipment is currently available and prices are expected to continue to fall. A requirement could have been placed on the communications service provider to determine the caller location (as is done on 9-1-1 systems today for calls from wireline connected telephones). Unfortunately, the technology to locate callers using wireless service providers, although encouraged by the FCC's recent wireless Emergency Telecommunication System rule making, is still in a developing stage, and it is difficult to predict when this technology will mature to a reliable level for this service.

Emergency Notification Data Message Routing

Wireless data emergency notification messages must be efficiently routed to the EM that has jurisdiction at the location where an emergency is taking place. This is particularly challenging for WAN wireless communications providers because it requires them to inspect data messages and determine the message destination where they currently do not inspect messages (they only transmit them) and the sender provides the destination. Two mechanisms have been considered: routing the message by a selected CSP or a selected EM for emergency request message analysis, validation, verification and if necessary routing to an appropriate destination. The architecture supports only the EM method of emergency data message routing because this is the current industry direction. The CSP approach would require many CSPs to adopt a common standard to achieve nationwide interoperability and to date there is no consensus on the need for this.

Map Attribute Referencing for Communication

A common method of referencing transportation links and nodes is essential for many of the ITS services involving cooperative processing between ITS subsystems. A common frame of reference is needed so that these communications between subsystems can be rationally reduced to an unambiguous reference to the same transportation links, ramps and intersections. Two competing general approaches have been

proposed: link-IDs (a national topological network) and coordinates (a national coordinate system using latitude, longitude and altitude). The architecture currently supports both types of representation: with a primary dependence on latitude, longitude and altitude (low technical risk) and an optional dependence on a national link-id database (compact representation for efficient communication).

Integrated Traffic Management, Demand Management and Route Selection

The architecture supports an evolution of traffic management, demand management and autonomous and infrastructure assisted dynamic route selection. This represents a natural evolution from stand-alone in-vehicle systems (today's deployments) to the fully integrated Advanced Traveler Information System (ATIS) and Advanced Traffic Management System (ATMS) deployments. This evolutionary deployment is necessary to maintain traveler benefits as market penetration increases and to allow for increasing benefits as computing and communication costs drop. The rate of evolution can be different in different areas of the country, and regions may support more than one stage of the evolution at the same time.

Electronic Payment Operations

Payment mechanisms were chosen to be consistent with existing and emerging deployments for electronic fare, toll, and parking payment systems, as well as interoperability with the existing financial network for approval and clearance of electronic payments. Consideration is given to always allow a cash payment capability, with payment by financial instrument (e.g., credit card, debit card or stored value (cash) card) as an option.

The current architecture is designed to serve as a starting point for eventual standards on secure electronic transactions that will probably be developed independent of ITS as a part of the Communications layer of the Physical Architecture by CSPs wishing to support emerging models for electronic commerce. Protocols can be implemented for credit card, debit card, and stored value card transactions between customers and service providers using the existing financial network for approval, clearance, and reconciliation. The protocols being developed and supported by the National Architecture are based on public-key cryptography protocols and can be implemented in either software or hardware. Increasing security of these transactions are monotonically related to key management complexity. Deployment can be gradual and incremental.

3 Assumptions

This section describes the deployments or implementations used for the Theory of Operations discussions. It is a high-level overview of the National ITS Architecture definition, and includes a brief discussion of the technological and institutional environment that ITS deployments will need to integrate with, mostly in the area of communications technologies. More detailed analysis of these issues associated with specific choices that were considered and made in the architecture development can be found in Chapter 5.

In general, the ITS Logical Architecture can be deployed in a variety of Physical Architecture implementations. We anticipate that each local community will deploy only those processes that it views as useful for its situation, and specific deployment choices will be made based on institutional considerations, prior deployment choices, financial capability, and anticipated transportation needs.

The Theory of Operations discussion is based on assumptions addressed mostly towards the 20-year most mature, full deployment scenario, so that each user service may be addressed, including those that may only emerge in the 20-year timeframe or beyond. So that the full interoperability and functional operational concepts of each user service may be addressed, we have assumed a deployment involving all the user services. Where operational concepts may evolve over the 20-year timeframe, that evolution is discussed.

3.1 ITS Physical Architecture Subsystems and Terminators

Figure 1 (on page 8) shows at a very high level the ITS subsystems and architecture flows between them. The four large boxes in Figure 1 represent the ITS subsystems in four groups[1]: Center subsystems, *Roadside* subsystems, *Vehicle* subsystems and *Traveler* subsystems. The 19 ITS subsystems that constitute each of these groups are listed in Table 1 and described in more detail in the Subsystems Section of Chapter 2.

3.1.1 Terminators

Terminators, not shown in Figure 1, represent entities outside of ITS that need to communicate or interact with ITS subsystems. The terminators have been grouped into three categories: *Users* (Center Personnel, Roadside Personnel, Vehicle Operator or Traveler), *Systems* (Center System, Roadside System and Vehicle System), and *Environment* which are described as follows:

User Terminators. These are the personnel at ITS Center subsystems and Roadside subsystems as well as Drivers and Traveler who interact with ITS subsystems. These interfaces are *human* interfaces that can be characterized by information flow, but not characterized in terms of data messages.

System Terminators. These are the non-ITS Centers (e.g., Government Agencies that ITS will interact with), Roadside systems (e.g., traditional signals and sensors) and Vehicle Systems (e.g., braking and steering systems) that ITS will interact with using ITS data messages. Also included in this category are *Other Subsystem Terminators.* This is a representation the architecture development team has adopted to

---

1. Readers familiar with *object-oriented* methods may think of these groups of subsystems as *classes* of subsystem *objects*. Objects of a given class tend to have common communication interfaces, data structures, and processes that perform operations on data. As will be discussed later, the subsystems of a given class can be aggregated by designers to form subsystems customized to a particular deployment, but retaining (or *inheriting*) the characteristics of the basic subsystem types.

indicate the interactions between multiple like subsystems. For example, vehicle-to-vehicle messages (Vehicle Subsystem to Other Vehicle Terminator) and Traffic Management Center-to-Traffic Management Center (Traffic Management Subsystem to Other TM Terminator) messages.

In the context of ITS, "the system" is usually meant to be the intelligent transportation system (e.g. the "ITS"), that is, all the subsystems of ITS. The terminators then represent other (non-ITS) "systems" that the ITS shares information with. In many respects, this is a very fuzzy boundary. Certainly, some people will argue that the Construction and Maintenance terminator, for example, should be a part of the ITS. While this terminator is an important source of information for the ITS, at this time it is not included in the ITS. If the User Services that are the basis of the National ITS Architecture, and are expanded to include a significantly larger amount of Construction and Maintenance service requirements, then the Construction and Maintenance terminator may possibly be converted to a subsystem (or part of a subsystem or group of subsystems) with a richer set of architecture functions allocated to it.

Environment Terminators. The environment is sensed by ITS subsystems. Examples are air quality and obstacles. The environment does not *communicate* with ITS subsystems, but rather interacts with or is sensed by ITS subsystems.

The terminators will appear in the physical architecture flow diagrams used to illustrate the ITS Architecture operational concepts in Chapter 4. Table 2 lists all the terminators in the ITS Architecture and identifies their category.

Terminator Name	Terminator Category
Archived Data Administrator	User
Archived Data User Systems	System
Basic Vehicle	System
Commercial Vehicle Driver	User
Commercial Vehicle Manager	User
Commercial Vehicle	System
Construction and Maintenance	System
CVO Information Requestor	User
CVO Inspector	User
DMV	System
Driver	User
Emergency Telecommunications System	System
Emergency System Operator	User
Emergency Personnel	User
Enforcement Agency	System
Environment	Environment
Event Promoters	System
Financial Institution	System
Government Administrators	System
Government Reporting Systems	System
Intermodal Freight Depot	System
Intermodal Freight Shipper	System
ISP Operator	User
Location Data Source	System
Map Update Provider	System
Media	System

Terminator Name	Terminator Category
Multimodal Crossings	System
Multimodal Transportation Service Provider	System
Other Archives	System
Other CVAS	System
Other Data Sources	System
Other EM	System
Other ISP	System
Other Parking	System
Other TRM	System
Other Vehicle	System
Other TM	System
Parking Operator	User
Payment Instrument	System
Pedestrians	User
Potential Obstacles	Environment
Rail Operations	System
Roadway	Environment
Roadway Environment	Environment
Secure Area Environment	Environment
Toll Operator	User
Toll Administrator	User
Traffic	Environment
Traffic Operations Personnel	User
Transit Driver	User
Transit Fleet Manager	User
Transit Maintenance Personnel	User
Transit System Operators	User
Transit User	User
Transit Vehicle	System
Traveler	User
Vehicle Characteristics	Environment
Wayside Equipment	System
Weather Service	System
Yellow Pages Service Providers	System

Table 2. Terminators of the National ITS Architecture and their Category

3.1.2 Subsystem Background

The specific choice of subsystems in Figure 1 represents a partitioning of the ITS architecture that captures all expected or likely subsystem boundaries for the near to 20-year future. In this way, the intersubsystem boundaries identify the likely candidates for institutional and message standard interfaces between subsystems. Thus, the choice of subsystems and the logical process specifications assigned to them are based on identifying the smallest common units of current and anticipated institutional responsibilities and physical deployments from the present through the next 20-years. The identified institutions and general types of responsibilities will continue, but in "aggregations" of subsystems that will vary by geography and time. Allowing the subsystems to be aggregated in different ways enables the architecture to be flexible and adaptable to local deployment choices. Furthermore, choosing subsystems

with small but common units of institutional responsibilities allows each ITS stakeholder to easily identify their functions and interfaces in Figure 1 (discussed in greater detail in Section 0).

An example of an aggregated subsystem deployment would be in regions where the Traffic Management Center and Emissions Management Center subsystems are deployed as a single subsystem. In those cases, the total number of physical subsystems actually deployed would be less than shown in Figure 1.

3.1.3 Subsystem Multiplicity

The subsystems shown as single entities in Figure 1 are representative of multiple instances of the specific subsystem. For example, several TMSs in a region, each with their own jurisdiction, may communicate with each other (and each with their many Roadway subsystems) to implement regional ITS policies.

The multiplicity expressed for ITS subsystems extends to the wireline and wireless communication elements as well. In the previous example, the TMSs may communicate with each other using a commercial wireline data communications service provider, but may have their own dedicated wireline communications elements for data communications with their many Roadway subsystems.

3.1.4 ITS Subsystem Groups

The ITS architecture subsystems of Figure 1 are further grouped where the subsystems may share common communication elements, deployment, and institutional characteristics. ITS subsystems in each group are shown in a common shaded box in Figure 1. The four groups of subsystems are Center Subsystems, Roadside Subsystems, Vehicle Subsystems, and Traveler Subsystems. The communications subsystems and modalities referred to in the following discussion of each group will be further developed in section 3.2 below.

3.1.4.1 Center Subsystems

These subsystems have no requirement to be on or adjacent to a roadway and thus can be located anywhere. To communicate with other subsystems they need access to wireline communications.

3.1.4.2 Roadside Subsystems

These subsystems typically include some function that requires convenient access to a roadside location for deployment of sensors, signals, programmable signs, or some other interface with travelers, vehicles, or freight. Roadside subsystems generally need wireline communications for messages to/from one or more Center subsystems, and possibly also have toll-tag or beacon communications to some or all vehicles passing the specific roadside location where the subsystem is deployed.

3.1.4.3 Vehicle Subsystems

These subsystems are installed in a vehicle. There will be considerable subsystem commonality across the various vehicle types in some areas, e.g., navigation and emergency request functions. In addition to DSRC, vehicles may be equipped with WAN wireless communications equipment to enable data communications with specific Center subsystems. For example, WAN communications between a vehicle and an ISP for a parking reservation prior to arrival at a parking lot, or for one-way (e.g., link-times to the vehicle) or two-way communication with one or more Center subsystems (e.g., Commercial Vehicle to Fleet Management messages). Finally, the vehicle subsystems may be equipped with vehicle-to-vehicle data communications in support of Advanced Vehicle Control Systems (AVCS) services (e.g., high density platooning).

3.1.4.4 Traveler Subsystems

These subsystems represent the "personal" and portable platform for ITS functions of interest to a traveler for support of multimodal traveling. The Personal Information Access Subsystem (PIAS) in this group may have WAN wireless communication capability similar to the capability in vehicles as well as the ability to access the same information services over a wireline communication element (e.g., when the PIAS is "docked" at home, at work, or at a kiosk).

3.2 Physical Architecture Communication Elements and Modalities

The Architecture Interconnect Diagram (AID) of Figure 1 includes communication elements (the "sausages") in addition to the subsystems ("rectangles"). These communication elements are a part of the interconnect channels and do not perform any ITS processing per se. That is, there are no Logical Architecture Process Specifications assigned to them. Their role is to facilitate the transfer of data messages between ITS subsystems where ITS processing does occur.

The four key groups of architecture interconnects identified in Figure 1, and the assumptions to be made in the Theory of Operations to follow, are briefly discussed in the following sections. A more thorough discussion can be found in the communications sections of the ITS Physical Architecture document.

3.2.1 Wireline Communication Elements

The WAN wireline communication element can take a number of forms. Typically it will be a data network of some kind. Physically the network can be fiber, coaxial, twisted pair, or even microwave between two fixed entities. It can be an ITS dedicated or *private network*, such as a communication system installed by a public agency to pass messages between a Traffic Management subsystem and associated Roadway subsystems distributed across a region. Alternatively, it can be a private sector deployed network owned and operated by a communication service provider, where operators of ITS subsystems pay a service fee for connection to and use of the network for ITS functions (called a *public network* or *shared network*). More than one network used for ITS may coexist in a region, and these networks will be connected (or internetworked) to support ITS message communication between subsystems that are attached to different networks.

It is expected that the current trend toward ubiquitous internetworking of public and private data networks, as currently embodied in, for example, the "Internet", will continue. This will enable inter-subsystem messaging across local, regional, and national distances. What the Internet is rapidly evolving to (as security and reliability issues of today's Internet are addressed) has been referred to as the "National Information Infrastructure" or "NII".

In the near term, we expect that many communication elements will be dedicated, as they primarily are today. As commercial data networks are deployed, interconnected, and mature, and the cost of access and use of these private data networks drops, we expect more and more wireline networks for ITS to be supplied from CSPs. The time when the transition from private data networks to commercial data networks becomes practical and economical will vary by region. We expect this transition to be analogous to the transition that was made early in this century from private phone networks to the Public Switched Telephone Network (PSTN). Our expectation is that in the 20-year timeframe most ITS communication will be provided by CSPs. The availability of low cost wireline data communication services (in different timeframes varying by region) will have a public policy implication: the current trend of deployment of dedicated fiber and other wireline infrastructure by public agencies will be replaced by purchase of communication services from CSPs. Public agencies that would otherwise build their own infrastructure might have additional price leverage with CSPs by negotiating long-term purchase agreements.

Table 3 summarizes several examples of current or emerging wireline data protocol technologies, indicating several characteristics of each technology. The table primarily focuses on link layer protocols but includes for comparison a network layer protocol, *IP*. Also included in Table 3 is *ISDN*, a protocol that uses standard phone company lines with ISDN switches to enable digital end-to-end circuit switched connections. Different technologies have different characteristics that will enable different operational concepts to be employed. For example, Asynchronous Transfer Mode (ATM) supports data as well as voice and video transmission. At the same time, it allows bandwidth to be flexibly allocated to different applications, in the same way that the lines and switches in the PSTN are time shared between telephone users. In this way, bandwidth need not be dedicated to a specific application (e.g., sending video from a camera to a TMS continuously), but can be time shared between many sources and destinations of data.

	ATM	**Frame Relay**	**MAN (802.6)**	**FDDI**	**X.25**	**ISDN**	**IP**
Data Rate	45 Mbps to 2.4 Gbps	64kbps to 2Mbps	56 Kbps to 45 Mbps	100 Mbps	9.6 Kbps to 64 Kbps	64Kbps to 1.54Mbps	NA
Information Supported	Data, Voice, Video	Data	Designed for Data, (Voice, Video hooks)	Data (Video with FDDI-II)	Data	Data, Voice, Video	Data
Flexible Bandwidth Allocation	High	Medium	Medium	Medium	Low	Medium	Low
Seamless LAN/WAN	Yes	No (WAN only)	No	No (LAN only)	No (WAN only)	NA	Yes
Latency	Very Low	Low	Low	Low	High	Very Low	NA

Table 3. Example Wireline Data Networking Technologies

For example, in video surveillance, a region may have hundreds of video cameras deployed at Roadway subsystems, but only one operator assigned to view the video images at a TMS. In an ATM data network, only those video camera sources being viewed by the operator at a given time need access to costly network bandwidth, and the video sources may be switched every few seconds. Thus, even though hundreds of video sources are deployed at the Roadway subsystems, only a small amount of network bandwidth (for just a few video signals at a time) need be consumed at any time. This sharing of network bandwidth is analogous to the sharing of telephone line and switch bandwidth on the PSTN, and is a reasonable mechanism to control expenses when wireline network communications are purchased from a data CSP. Future automated incident detection technology that are dependent on analysis of video data could have the video analysis equipment deployed at the Roadway subsystems, so that only the low bandwidth results of the video analysis are sent to the TMS. If an incident is detected automatically in a video image at the Roadway subsystem, then that image (at a higher cost, but for only a brief time) can be sent to the TMS operator for verification, classification, and action.

Data networks using different technologies can be used to interconnect subsystems as well as components within subsystems. We assume that in the 20-year timeframe networks using different technologies will be internetworked into the NII using the specialized components (“bridges”, “switches”, “hubs” and “routers”) at the interfaces to accommodate differences in specific media and protocol. Thus the

requirement on network media and network protocols for ITS wireline communications is that the chosen network technologies be able to be internetworked to the NII. (Application layer protocols are the subject of Chapter 4.)

One significant ongoing standards activity in the wireline area is development of the *National Transportation Communications for ITS Protocol (NTCIP*). This set of standards defines common methods of physically interconnecting ITS control equipment, establishes the protocol and procedures for establishing communications between the components, and defines procedures to develop and register common sets of manageable objects related to controlling and managing the components. The standards are being developed by National Electrical Manufacturers Association (NEMA) with support from the US Department of Transportation (DOT). NTCIP contains a suite of communications protocols, divided into several class profiles, for integrating the various components that may be included in an ITS. The standard defines the elements that allow manufacturer inter-changeability of transportation control equipment. Also, a complete end-to-end data handling procedure is defined, allowing devices to perform tasks associated with communications between traffic management centers and other field equipment.

The NTCIP is designed to support second-by-second, multi-drop, low speed modem signal system, and through modularity, extensibility, and configuration, expand to accommodate modern technology and signal applications. This modularity is achieved by adhering to the International Standards Organization (ISO) Open Systems Interconnect (OSI) 7-layer reference model. By following the ISO-OSI model in defining the NTCIP protocol "stack" (i.e., the layers), modularity is achieved. This permits NTCIP to be compatible with both the current installed signal system infrastructure and advanced applications and technology.

The NTCIP uses this modular approach to define several protocol stack and conformance level "profiles". Each profile is tailored to support the requirements of a particular type of communications link. For example, to support the CalTrans AB-3418 protocol, only layers 1, 2 and a limited subset of one of the layer 7 protocols are required. Full support for an advanced ITS system may require layers 1, 2, 3, 4 and full support of several layer 7 protocols.

To address the goals and topologies described above, the NTCIP uses an industry standard, generalized communications approach. This approach starts with a general model for communications and then adds specific protocols to provide the basis of information transfer. The ISO-OSI model defines a network model consisting of a stack of seven layers and associated protocols and interfaces. By following the model in defining the NTCIP protocol stack layers, modularity is achieved. The various protocols adopted by the NTCIP meet the need for reliable communications. The NTCIP supports the following requirements of ITS:

1. Support for connection-oriented and connectionless services.
2. A mechanism for acknowledged and unacknowledged data transfers.
3. An error detection algorithm scheme that insures the probability of accepting a bad frame is exceedingly small.
4. A structured approach that supports transmission media and data rate independence.
5. An addressing scheme that is extensible to cover existing and future requirements.
6. Support of a message structure that is extensible to account for variability, changeable data content, and varying length structures.

The current definition of NTCIP includes the physical and data link layers (common to all profiles) and the definition of application layer protocols (Simple Network Management Protocol-SNMP and Simple

Transportation Management Framework-STMF). In addition, the following sets of application (device) objects have been defined:

1. Global Object Definitions
2. Actuated Signal controller Objects
3. Variable Message Signs Objects
4. Closed Circuit Television control Objects (CCTV)
5. Ramp Meter controller Objects

The definition scope of NTCIP has also been expanded to cover TMS-to-TMS communications.

3.2.2 Wide Area Wireless Communication Elements

The WAN wireless communication element of Figure 1 is analogous to the wireline communication element in many ways. The communication element can be dedicated to a specific user or agency (and publicly owned or privately owned), or it can be privately owned and operated by a communication service provider who sells access to this data network to many users or agencies for a fee.

A key feature of most wireless communication elements is that they be internetworked to a wireline communication system of some sort. In this way, mobile units can exchange ITS messages with Center or Roadside subsystems. This is shown diagrammatically in Figure 1 by the line connecting the Wireline and Wireless communication elements. We assume and require that in the 20-year timeframe the 2-way ITS wireless communication network will have the necessary coverage for a particular user service application, and that the wireless network will be internetworked to the wireline wide area communications network. The following sections discuss various options that might be deployed in the 20-year timeframe.

3.2.2.1 One-Way and Two-Way Systems

Wireless communication systems can be one-way (broadcast) or two-way. Examples of broadcast systems are FM-subcarrier systems. Two-way systems that are private can be SMR or E-SMR (Enhanced SMR). SMR and E-SMR require licenses from the FCC for operation, and are typically dedicated to a specific service or agency.

3.2.2.2 Trunked and Cellular Two-Way Systems

Wireless systems can be trunked or cellular. Trunked systems have many base stations located in a region, each base station operating on the same frequencies at the same time to allow broad reliable coverage. Cellular systems are more sophisticated, where adjacent base station centered “cells” operate on different frequencies, but non-adjacent “cells” may use the same frequencies. In this way the frequencies can be “reused” across space, making the most efficient use of the available spectrum.

Commercial 2-way wireless data communications providers can be classified as to whether or not they make use in some way of the existing cellular system designed for switched telephony or one of the new systems that are specialized for data. Examples of switched telephony systems adapted for data are Cellular Digital Packet Data (CDPD), Code Division Multiple Access (CDMA) Data or Time Division Multiple Access (TDMA) Data. Examples of commercial systems dedicated to wireless data services are Ardis and RAM.

Table 4 summarizes some of the characteristics of six commercially available or emerging wireless communication systems. They are each briefly discussed in the next two sections. Actual comparative data rates for Table 4 are extremely difficult to get. Effective data rates, which might be useful for comparison (and are rarely available), are computed from raw data rates minus fixed protocol overheads and minus variable contention dependent overheads.

3.2.2.3 Data Switching: Packet and Circuit

"Data switching" refers to how a data message is directed through the wireless and wireline network to its destination.

Circuit switching requires that before any data is sent, the entire connection (or "circuit") from the origin to the destination be setup. This can take around 1 second, and is analogous to the delay experienced at the beginning of a phone call between when the last digit has been dialed and the beginning of the first "ring". This delay can establish a costly minimum expense for short ITS messages that may only take a small fraction of a second to transmit once the circuit is established.

Packet switching is a method by which a generally small packet of data includes a "header" which indicates the destination address for the data. ITS messages may be composed of one or more packets of data. Standard algorithms at the origin subsystem will break messages up into packets for transmission, and analogous algorithms at the destination will reconstruct the original message from received packets. Packets can be sent with very little delay. The nodes of the communication network will interpret the packet header, and based on the location of the destination and the network management protocol will relay the packet to the next node in the network. Eventually the packet will reach the destination. Packet switching is generally more efficient for small messages, because segments of the communication network are not committed to a particular message routing and the message is not waiting for the complete circuit to be established with no actual data being sent as in circuit switching.

	CDPD	**AMPS**	**CDMA**	**TDMA**	**Ardis**	**RAM**
Base Wireless Network	Existing Voice and Data Cellular Duopoly	Existing Voice and Data Cellular Duopoly	Existing Voice and Data Cellular Duopoly and PCS	Existing Voice and Data Cellular Duopoly and PCS	Private Data Network	Private Cellular SMR Data Network
Raw Data Rates	19.2kbps	9.6 Kbps	Circuit = 9.6 Kbps Packet I = 56-64 Kbps Packet 2 = 512 Kbps Packet 3 = 2Mbps	Circuit = 8.0 Kbps Packet = 28.8 Kbps	4.8 Kbps - 19.2 Kbps	8.0 Kbps half-duplex
Estimated Effective Data Throughput Rates	14.3kbps-≈10kbps (See Note 1)	<9.6 Kbps	Circuit = 8.55 Kbps Packet I < 56-64 Kbps Packet 2 < 512 Kbps Packet 3 < 2M bps	Circuit < 8.0 Kbps Packet < 28.8 Kbps	2 Kbps - 8 Kbps	4.0 Kbps half-duplex
Data	Packet	Circuit	Circuit/	Circuit/	Packet	Packet

	CDPD	AMPS	CDMA	TDMA	Ardis	RAM
Switching			Packet	Packet		
Open Standard	Yes	Yes	Circuit: Yes Packet: Future?	Circuit: Yes Packet: Future?	No	Yes
Cost Basis	per packet	per minute	per minute/ per packet	per minute/ per packet	per packet	per packet
Cost for Small Messages	Low	High	Unknown	Unknown	Medium	Medium
Coverage by 1997	≈90% of population	≈90% of population	Unknown	Unknown	≈80% of population	≈80% of population

Note 1: The CDPD downlink (base to subscriber) is broadcast-based and always 19.2kbps (data and overhead). Without overhead the throughput is 14.3 Kbps per channel. The number of channels allocated to CDPD is deployment dependent. The uplink (subscriber to base) is contention-based, and therefore its throughput is generally less than the forward channel. Its maximum net information throughput (i.e., after accounting for the effects of overhead and contention) is approximately 10 Kbps.

Table 4. Wireless Data Networking Technologies (2-way)

Packet switching will have a distinct advantage in ITS applications where the communication channel is shared and where there are numerous short messages to be communicated. For example, ATIS applications based on probe data (short messages from vehicles to ISPs indicating position and time) and for ATMS (e.g., many short messages from Roadway subsystem sensors to a TMS indicating current traffic conditions or short reverse messages updating traffic signal timing.)

3.2.2.4 Two-way Wireless Open Standards

A key concept in the architecture is that communications technology is a "commodity" that can conceptually be considered separately from ITS subsystems and their architecture-defined functions. The benefit of this concept is that the investment in ITS subsystem functions can be made relatively stable and secure, while still allowing rapid evolution and adaptation to evolving communication technologies.

This requires that communication modules in ITS subsystems be replaceable at low cost---so that the benefits of rapidly evolving communication technologies can be incorporated into ITS deployments. An essential element to this concept is communication standards written so that the interface between the communication modules and the ITS modules in an ITS subsystem and across subsystems are well defined. These standards must be *open* (as opposed to *proprietary*), so that multiple vendors will be able to provide functionally equivalent communication modules, and the cost of acquiring these communication modules will be contained by competition.

3.2.2.5 Cellular Telephone Based Systems

CDPD: Cellular Digital Packet Data. CDPD is a separate service operated by the mobile phone companies that makes use of idle channels (and optionally dedicated channels) on the existing cellular mobile phone service. It is expected that all existing mobile phone systems will eventually support CDPD, which is commercially available in many markets. CDPD pricing is expected to fall once additional alternative 2-way wireless data services are deployed and the market for wireless data communication services begins to develop. CDPD has an advantage in that the service providers can leverage the existing mobile phone system infrastructure to "piggyback" the CDPD service onto. No

additional cell sites, antennas, or other key infrastructure components need to be deployed. The additional equipment needed to deploy CDPD includes the equipment necessary to interconnect the wireless communication network to the wireline data network (e.g., "NII") as well as additional data communications equipment in each cell site. Both of these expenses are small marginal expenses compared to establishing a totally new data network.

AMPS: Advanced Mobile Phone System. Data is transmitted using cellular modems attached to mobile phones. Circuit switching is performed using the standard dial-up method, with the attendant cost penalty for short messages. This method has the advantage in that it is ubiquitously available today except in some remote rural areas.

TDMA: Time Division Multiple Access. TDMA was developed as a digital alternative to the analog scheme used for the AMPS mobile phone service. In summary, up to three circuit switched voice calls share a single channel. This is accomplished by digitizing the voice signals and time-sharing the single AMPS channel between the three calls. It has been envisioned that the digital channel used to transmit the voice signal could be used to transmit digital data as well. A TDMA Short Message Service has been defined although open standards have not yet evolved.

CDMA: Code Division Multiple Access. CDMA, like TDMA, was developed as a way to allow greater numbers of voice transmissions to use the limited number of channels available to the AMPS system. In CDMA, the digitized voice signal is encoded and transmitted using a spread spectrum modulation (as opposed to the narrow-band modulation used by AMPS and TDMA). CDMA, developed after TDMA, promises better signal-to-noise performance as well as being able to support more users of the limited spectrum allocated to the cellular service. Unfortunately, CDMA is not compatible with TDMA, and the industry and market are currently deciding how they will evolve. Like TDMA, it has been proposed that digital data, instead of digitized voice, could be transmitted over CDMA and progress is being made towards open standards for data transmission (a CDMA Short Message Service has been defined and finalized).

3.2.2.6 Non-Cellular Telephone Based Systems

Ardis and RAM: These services each use their own data communication protocols for the wireless communication network, and then interoperate with open or proprietary wireline data communication services through gateways. Ardis uses high-power base stations with a proprietary protocol to gain metropolitan area coverage at the expense of low system bandwidth -- with the consequence of reliable communications, but with potentially high latency when many users are contending for the limited channel capacity. RAM uses the cellular paradigm of frequency reuse with an open protocol, and has invested considerably into its base-station facilities that are dedicated to the RAM network. Both networks use packet switching and are oriented towards short messages.

3.2.2.7 Low Earth Orbit Satellite Enhanced Cellular

One of the limitations of the current cellular deployments is that they are based on terrestrial base stations that are deployed in locations based on expected market. What this means is that locations where customers willing to pay for the wireless service are too sparse will not get the base station deployments and will be left without wireless communication coverage. Thus, some ITS services that require two-way communications will not be possible in the near-term for some rural locations.

Several consortia have been formed to deploy constellations of Low Earth Orbiting (LEO) Satellites for the purpose of providing wireless communication services (voice and data) where terrestrial cellular systems do not reach. As these systems are deployed (as early as 1999 and certainly by 2002) to interoperate with the current cellular terrestrial based infrastructure, cellular coverage will expand to

cover nearly 100% of the road network. Exceptions will still be roads that do not have unobstructed views to the sky such as roads in tunnels, etc. (However note that highly traveled links in tunnels currently do have terrestrial based cellular coverage, e.g., the Lincoln and other Tunnels between New York and New Jersey.)

3.2.3 Dedicated Short Range Communications (DSRC)

This communication channel (not an independent subsystem) represents the direct communication paths between vehicles and roadside equipment.

DSRC (using roadside beacons) are one or two-way communication channels that can either serve between a Vehicle subsystem and either a Roadway subsystem or, along with an associated Wireline communication element, a Center subsystem. Note that the DSRC can be internetworked with the Wireline network in a similar way as the wireless wide area communications network (see Figure 1).

Beacon communication capability may be deployed in Roadway and Vehicle subsystems where the cost of the beacon communication equipment is less than the cost of the equivalent Wireless communications infrastructure. For example, toll and parking payment operations, CVO operations or urban fixed route public transit operations may be most economically served by beacon communications for messaging between the vehicles and the appropriate Roadway or Center subsystems. In these cases, the specific costs and sources of initial capital as well as operations and maintenance expenses must be considered.

In DSRC communication systems, a fundamental difference with the wide area wireless network is that the vehicle that primarily depends on DSRC is out of touch with the infrastructure most of the time, and relies on brief bursts of high bandwidth communications while in the field-of-view of a roadway beacon. This is appropriate for many services, for example, toll and parking payment, or fixed route Transit vehicle operations data up-load to a Transit Management center. However, it may be inappropriate for other services, for example, Vehicle to Emergency Management center requests for emergency service (where the vehicle is likely to have a disabling incident out of the field-of-view of a beacon). In general, services that require “ubiquitous” (or near ubiquitous) coverage, such as the aforementioned “emergency request” service, are inappropriate for DSRC. Also, services that involve vehicle-center two-way transactions will in general be inappropriate for beacon applications. Many applications will typically have communication and processing latencies longer than the fraction of a second it may take a vehicle to transit a beacon field-of-view. But these applications may also require completion of the transaction sooner than the time for the vehicle to reach the next beacon. For example, vehicle navigation services where the center (infrastructure) computes and presents the individual vehicle optimal route will not be appropriate for beacons. Of course, services can be devised to meet these DSRC criteria (e.g. only allow one-way DSRC communication navigation services where the route selection is performed in the vehicle) possibly at the expense of other features.

3.2.4 Vehicle-Vehicle Communications

This communication channel (not an independent subsystem) represents the direct communication path between adjacent vehicles. The deployment of this communication is necessary for some AVCS services, such as high density platooning of vehicles.

3.2.5 Communication/Processing Sharing

As is evident in the AID in Figure 1, there is considerable opportunity for communication network sharing in the ITS architecture. Today, multiple wireless and wireline communication elements are dedicated to specific ITS services. In the 20-year timeframe it is likely that a collection of internetworked

wireline and wireless systems operated by CSPs will support the various ITS services, both public and private (much the same as telephone services are provided today).

Commonality and sharing in ITS processing (in the ITS subsystems) may be more limited. In particular, the expected continued drop in the cost of computer processing will support the distributed processing visible in the AID of Figure 1. Areas where considerable commonality of ITS processing may exist across user services are in the Roadway processing and subsequent distribution of surveillance data. This data is of value both to Traffic Management (TMS), Traveler Information (ISP), and other ITS services, and this data can most easily be broadly collected by sensors located at the Roadway subsystem. (Note: the National ITS Architecture supports using Vehicle subsystems as probes (agents of surveillance), and an effective probe deployment by ISPs through their clients could reduce the utility of sharing the Roadway subsystem processed surveillance data.)

4 Operational Overview

Organized by the User Services, this section describes the operation of the National ITS Architecture from a combined user and system perspective. The system perspective includes the subsystem processing, communications and equipment required to implement the user services. The user perspective identifies the user interactions with the system.

The sequence of architectural processing/messaging events for the implementation of each user service will be presented. In cases of strong synergy, several user services have been combined into a particular architectural mechanism. Where different architecture approaches are allowed for service delivery in the National ITS Architecture, each approach will be discussed.

The Physical architecture is defined by the Architecture Flow Diagrams (AFDs) and Architecture Interconnect Diagrams (AIDs) in the Physical Architecture document. Furthermore, the functional characteristics for a physical subsystem which define how it processes input messages to issue output messages (and thus the message sequences between collections of subsystems) are defined by the logical process specifications assigned to the subsystem, found in the Logical Architecture document. However, the detail in the complete presentation of the architecture can sometimes obscure the concepts of how each user service is supported by the National ITS Architecture.

The following set of simplified physical architecture diagrams (like Figure 2) describe at a high level the subsystems which are used to implement each user service or groups of synergistic user services. Also, the message sequence (numbers in circles) is identified by numeric labels on significant architecture flows. The numbers assigned to the architecture flows correspond to the numerical lists of flow descriptions in the text where important details of the operational message sequencing are described.

For many of the user services there are a variety of evolutionary paths toward increasing functionality. The operational message sequencing given in the chapter describes some, but not all, of these paths. *They are meant to be a guide to how the architecture supports the user services, and are not a meant to be a prescription of the only way the user services can be supported.* The architecture has been designed to have a great degree of flexibility, and can support a wide array of implementations.

Market Packages are an alternative representation of groupings of ITS subsystems to support diverse ITS implementations and can be found in the companion document *Market Packages: A tool for Viewing, Accessing, and Utilizing the National ITS Architecture*. Market Packages address specific sets of users, service levels, regional needs, and incremental deployment scenarios. The Market Packages address the question "how do you *deploy* the ITS National Architecture in specific scenarios." The *Theory of Operations* Operational Overview chapter addresses the question "how does the ITS National Architecture *work* to implement the User Services."

4.1 Travel and Traffic Management

4.1.1 Pre-Trip Travel Information

Figure 2 shows the subsystems and high-level message sequences for the Pre-Trip Travel Information services.

In Figure 2, "clients" requesting information are on the left side of the diagram, the "server" or ISP is in the center, and other subsystems that may be required to assist the ISP in its server function are located on the right side.

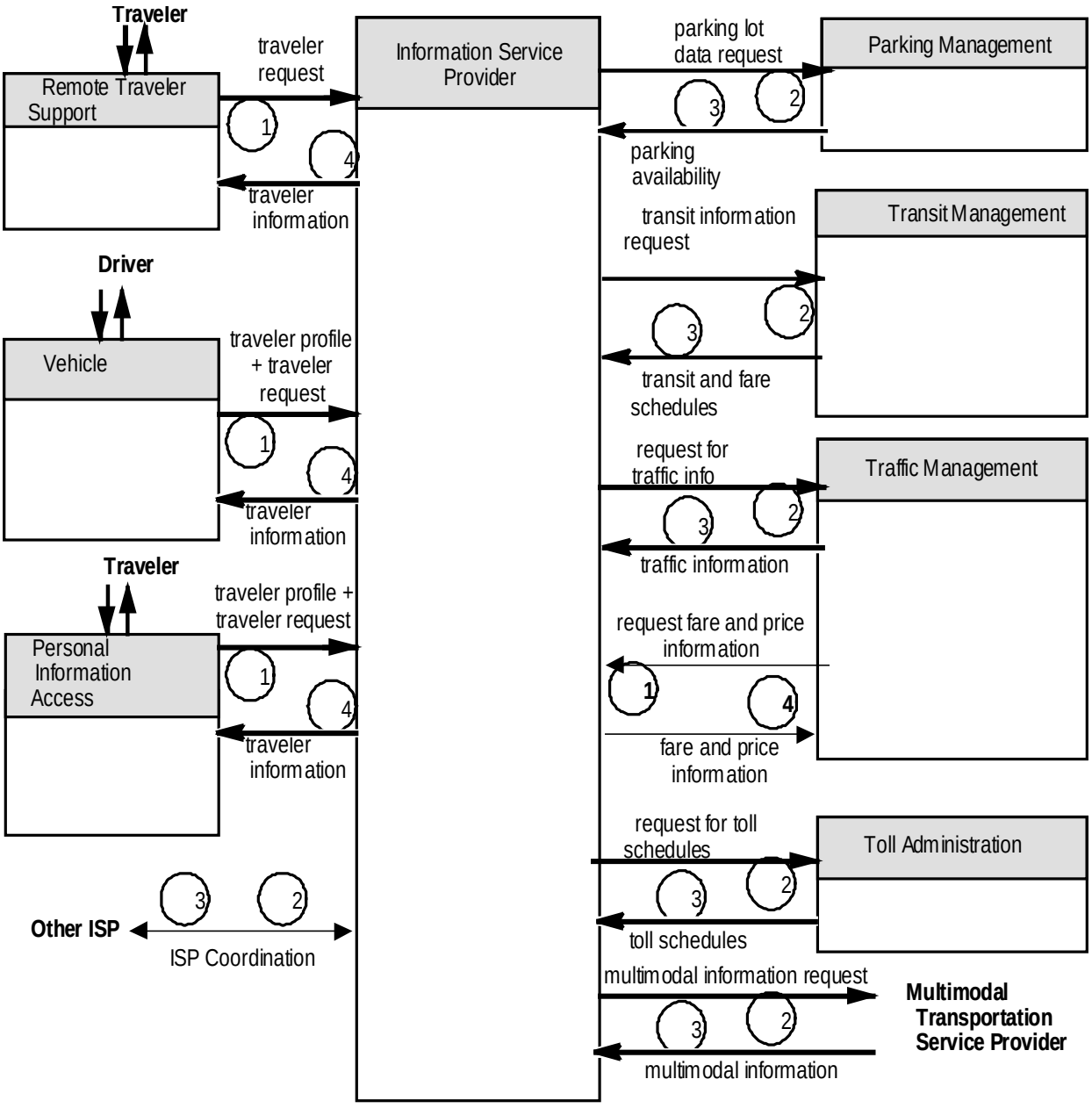

Figure 2. Physical Architecture for Pre-Trip Travel Information

Messages 1 and 4 Figure 2 represent a request/response transaction between client and server, and are required for the advanced pre-trip travel information function described here. It could be possible to deploy a simple pre-trip travel information system with only message 4, which could only contain information of general interest, not customized to a particular travelers plans. The client subsystems would use this general information to respond to specific traveler requests or to make general traveler information displays.

Note that messages 2 and 3 in the figure, which represent an ISP transaction request for additional information, are optional. This is in the sense that it is up to the specific implementation of the ISP to decide when to request and how to use the information requested from the auxiliary information source entities (Parking Management subsystems, Transit Management subsystems, Traffic Management

subsystems, Multimodal Transportation Service Provider terminators and Other ISPs). Ideally, a "just in time" information strategy would have the ISP request specific information from the source on the right as requests need the information. In this way the information is as fresh and timely as possible. Practical considerations such as communication costs and remote request/response latencies may result in the ISP periodically requesting and storing such information and using the stored information for most client requests. These implementation details, and the resultant level of client service and cost, will allow competing ISPs to differentiate themselves in the marketplace.

1. The traveler or driver, either from a Kiosk (Remote Traveler Support subsystem), the Vehicle subsystem, or from a portable computer or home/office computer interface (Personal Information Access subsystem (PIAS)) makes a traveler information request which is sent to the ISP. The traveler information request from the traveler or driver can be for the following types of information:

    - Traffic data on specific route segments
    - Transit deviations on specific route segments
    - Current conditions regarding weather, events and incidents
    - Fare, toll and parking charge information and availability
    - Advisory data based on location and/or transit route or vehicle number.

Note that the architecture does not specify the implementation of subsystems, but rather the functional requirements. For example, the Remote Traveler Support subsystem (RTS) could be implemented as a "kiosk", or as a "human operator" that interfaces with travelers over a phone line. In either case, the human interface-based interaction with the traveler results in a "traveler information request" message being issued from the RTS to the ISP.

Furthermore, the architecture does not specify how the "traveler information request" message is prepared, based on some interaction with the driver or traveler. This is viewed as "implementation" and not "architecture." It could be as simple as filling out a forms screen, or as sophisticated as an intelligent voice synthesis and voice recognition-based interaction with the traveler or driver, making use of information stored from prior interactions to speed and simplify the preparation of the request, as well as default contents that would speed the filing out of the required information. Imagination and cleverness to use technology appropriately by the designers of these interfaces will certainly determine their appeal with travelers and drivers (not to mention marketing savvy as well).

For the messages between the ISP and the Vehicle or PIAS subsystems, which may use a wireless communication interconnection, the size of the messages will need to be minimized in order to minimize communications cost. This can be accomplished by sending only the parts of the messages that have content, with unsent parts of the message assumed to have default values. This is a detail of the specific message protocol that needs to be worked out in the standards setting process.

Finally, the TMS can be a client to request from the ISP "fare and price information" necessary for advanced TMS traffic predictive models.

The ISP identifies which type of service is required and obtains information from the appropriate subsystems, if necessary, in the next two steps:

2. If demand responsive transit (paratransit) is an option, then a transit trip request is sent to the Transit Center and If vehicle travel is an option, then link travel times and other link conditions (e.g. incident locations and status, construction) are requested from the appropriate Traffic Management subsystem(s) and

    If parking is required, then a "parking lot data request" is sent to one or more Parking Management

subsystems and

If toll information is required, then a "request for toll schedules" is sent to the appropriate Toll Administration subsystems.

If other travel service provider modes are an option (e.g., rail, plane, ship, ferry, taxi, shuttle), then a request is sent to the appropriate "Multimodal Transportation Service Provider" external system interface.

Finally, if the needed data is available from another ISP, then a request is sent as appropriate.

3. The response messages for portions of the trip that were just requested (in step 2 above) are sent from their respective subsystems (or terminators) and received at the ISP.
4. The traveler information is sent to the traveler/driver (or TMS) and includes the following information:
    - Current, long term, and/or predictive model traffic data including
        - Incident details
        - Alternate routes
        - Parking lot occupancy
        - Traffic conditions for ramps, occupancy estimates, and HOV lanes
        - Pollution levels
        - Link times and queue delays
        - Transit schedules and deviations
    - Fare, parking and toll charges
    - Various special event and service information (based on yellow-pages type services) e.g. dining, lodging, ticket purchase, cultural information, services information, special events, tourist activities.

4.1.1.1 Autonomous Trip Planning

The Trip Planning function, which is the center of all these operations, could be as simple as the current map plans obtained from travel services such as AAA, or could have the sophistication of a knowledge-based system for computing the optimal multimodal route to minimize trip time (or some other traveler selected criteria) based on predicted travel times, costs, etc. The level of sophistication may evolve over time as ISPs compete by level of service.

The architecture supports the simple case of completely standalone (e.g., in-vehicle) trip planning. This function is termed "simple" because it involves no real-time intersubsystem communications.

4.1.1.2 Broadcast Information for Autonomous Trip Planning

The ISP may "broadcast" information concerning unanticipated changes in link-time (e.g., link travel time) conditions, transit schedule variances and incident locations for the purpose of augmenting an autonomous trip planning function.

4.1.1.3 Traveler Acknowledgement of a Selected Route

It is the responsibility of the traveler/driver to make the final decision on the selection of a route, whether the route is selected ("computed") autonomously or at an ISP. In some cases, it may be necessary to communicate the traveler/driver's selection back to the ISP e.g., to confirm demand responsive transit or other reservations or to update a congestion model used by the infrastructure to compute future routes. (Note: in an advanced system where there is close coordination between ISP and TMS, it is in the traveler's interest to have the Traffic Management congestion model updated to expect their future approximate link occupancies, if the updated model is used to send suggestions to subsequent travelers to use other links, thereby leaving the links the traveler has chosen to travel in a less congested state. Furthermore, the traveler's expected occupancy may be used by the Traffic Management subsystem to better plan the signal timings for the traveler's benefit.)

4.1.1.4 Possible Extensions to Support Intermodal Freight

Optimal routing of freight across different modes and carriers is a common logistical problem in commercial freight transport. Intermodal freight shipping of this kind might be accommodated in the architecture (there are currently no formal user service requirements for this function) by expanding the "2" and "3" messages to the Fleet Management subsystem, which will have a routing process analogous to the Transit Management demand responsive transit function. In this way, Fleet Management would inform the ISP about what freight transport resources were available. In addition, the terminator "Intermodal Freight Shipper" could specify the freight origin, destination, characteristics and shipping preferences (e.g., fastest or most economical) and constraints (e.g., hazardous materiel with specific materiel safety data-sheet ("MSDS") numbers and corresponding quantities). In this way, the ISP could then compute the optimal logistics for the Intermodal Freight Shippers freight.

4.1.1.5 Emergency Management Subsystem Incident Information for the ISP

Ideally, the TMS is tracking incident information from the EM subsystem and incorporating this information into the data sent on request to the ISP (see the Traffic Management architecture shown in Figure 10). In deployments where a TMS does not exist or where the TMS chooses not to make EM information available in its responses to the ISP, the ISP can request this information directly from the EM, as shown in Figure 3.

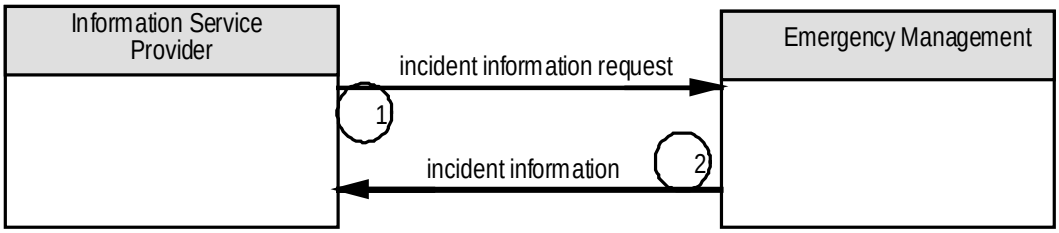

Figure 3. Physical Architecture for Providing Incident Information From EM to ISP

1. The ISP requests incident information from the EM subsystem specifying the type(s) of incidents of interest and how far back in time the EM should search to provide incident information. The time range is useful to avoid subsequent transmission of incident information that was sent in response to a prior request.

2. This message is sent for each incident known to the EM in response to the incident information request message. This message includes incident location, start time, expected duration, type of incident, severity and traffic impact.

4.1.2 Driver Information

4.1.2.1 Driver Advisory and Smart Probes

The basic driver advisory service, shown in Figure 4, is provided by the ISP.

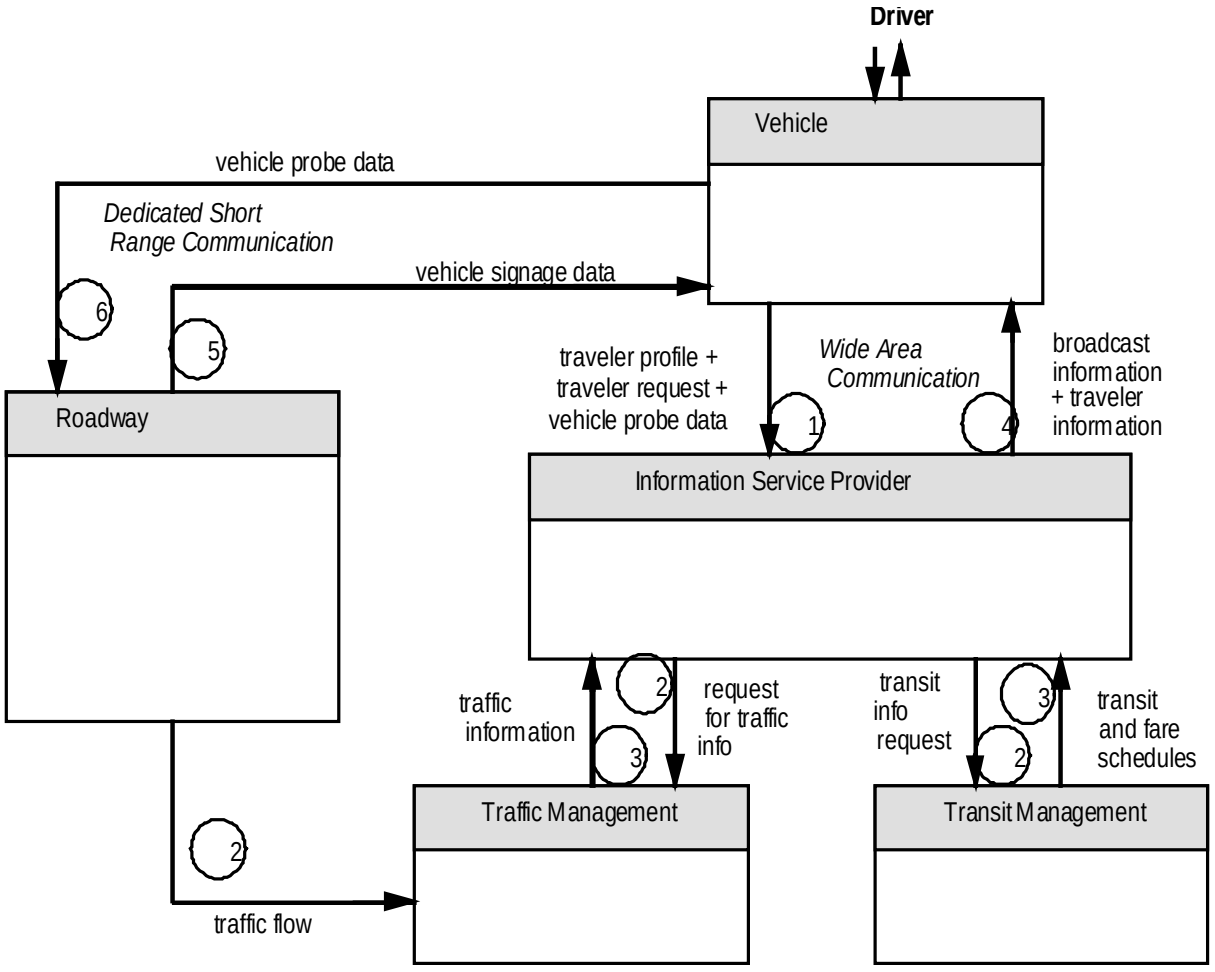

Figure 4. Physical Architecture for Driver Advisory and "Smart Probes"

1. A driver requests advisory information (via the traveler request arch flow) through the driver interface. He also sends the vehicle location and a time-stamp (vehicle probe data), which is the simplest basis for "probe" data, since the ISP can make an estimate of traffic flow conditions based on the time and location differences between these messages from each vehicle. Information that is requested can include

   - toll, fare or parking charges and availability
   - travel advisory for vehicle travel or transit travel
   - Send current location for vehicle travel advisories and transit route for transit advisory

2. The ISP requests advisory data from the TMS subsystem (through the architecture flow "request for traffic information"). At the same time, Transit information is requested from the Transit Management Subsystem (TRMS) via the "transit info request" architecture flow. Note that the ISP may choose to request information from the TRMS and TMS at regular intervals, not related to traveler requests. If the ISP is paying for the TMS and TRMS information, then the ISP will be motivated to carefully manage these information requests.

The Roadway subsystem periodically updates the TMS with surveillance data about roadway conditions. The TMS uses this information to maintain a real-time model of roadway conditions, so that the TMS can perform its own Traffic Management functions, and also rapidly respond to ISP requests.

3. In response to the ISP requests, the TRMS and TMS send "transit and fare schedules" and "traffic information" back to the ISP.

4. Advisory information can be provided to the driver based on the driver request, or can be provided automatically (event driven) once a connection with the ISP has been initiated. In this way, the driver can be notified shortly after an incident that affects his travel plan. The advisory information has two components, a non traveler-specific component (broadcast information) and a traveler-specific component (traveler information).

Broadcast information includes:

- Predicted/current incident descriptions, locations, severity, type, and impact (e.g. predicted end of a sporting event).
- Predicted route segment queue delays, occupancies, volume delays.
- Current traffic data (High Occupancy Vehicle (HOV) lane occupancies, parking availability, pollution state, link travel times, and queue times).
- Transit schedules, deviations.

Traveler information includes information similar in structure to broadcast information but specific to a travelers location or transit routes.

Messages 1 and 4 above represent a request-response mode of operation, where the response is customized to the request. An alternative deployment is to not have message 1 and to broadcast message 4 in an unsolicited one-way mode, using for example FM subcarrier for the wide area communication mode. Transmissions of this kind would necessarily broadcast information of interest to large numbers of travelers in the broadcast market making a wide variety of trips, and thus information may not be as timely or detailed as might be required by individual travelers. More precise, localized information could be transmitted by use of DSRC beacons. It is not practical to cover an entire area with a dense deployment of beacons, but in limited high value areas (such as near a bridge or major highway split) the beacons could provide very valuable local information. Each one-way broadcast beacon might be able to provide timely and complete advisory information in the region of the beacon, and less dense information in regions further from the beacon.

5. The Roadway subsystem may also communicate with vehicles using dedicated short range vehicle-roadside communications ("DSRC" e.g., Beacons) to communicate roadway information relevant to the location of the beacon. Signage information may also be included in this set of DSRC messages.

6. For rural applications where TMSs may not exist and where a wireless WAN infrastructure may also not exist, vehicles equipped to be "smart probes" can collect environmental or road condition sensor data on the vehicle and then send this to the roadside infrastructure as part of probe data. Automated road signing can include a smart beacon which can receive smart probe data and rebroadcast the environmental data to other vehicles passing by. This is intended for rural deployments to provide

real time environmental or road condition information (e.g. bridge icy, fog ahead) to travelers coming from the opposite direction.

4.1.2.2 Alternative In-Vehicle Signage Using Wireless Wide Area Communication

An alternative approach to providing In-Vehicle Signage, shown in Figure 5, from that discussed in the ITS-America Program Plan and shown in Figure 4, is also included in the ITS architecture. This user service may be an outgrowth of the Route Selection provided by the ISP. A navigable database maintained by the ISP can contain information about the location and content of signage on the links. By adding this signage information to route messages, it can be transmitted to the vehicle using the Wireless Wide Area Communications. When combined with a vehicle location system, the route guidance capability in the vehicle can have the ability to provide to the driver the upcoming signage information. The information could be provided either by auditory or visual means through the same driver interface used by the route guidance function.

Figure 5. Alternative Physical Architecture for In-Vehicle Signage

Using this mechanism, signage becomes a service provided by the ISP. It is not constrained to be an extension of the existing infrastructure signage. An active link from roadway signs to vehicles is certainly possible, and can coexist with this alternative, but brings with it added liability on the public sector. Once the public sector provides this service they become responsible for maintaining the capability, and suffer the legal consequences if the infrastructure component fails to work. (The assumption here, which can be debated, is that the private sector may be better prepared to manage the inherent liability of the signage user service.)

ISPs will likely exercise a choice as to what kind of signs they choose to include in routes, possibly taking driver preferences into account.

4.1.3 Route Guidance

The Route Guidance User Service involves two distinct processing operations: route selection and route guidance. The route selection operation involves selecting the route to take based upon the Driver, Traveler or Commercial Vehicle Manager request. The route guidance operation involves presenting the selected route to the driver or traveler in a step by step fashion. All implementations of the Route Guidance user service involve a route guidance process in the vehicle subsystem (VS) or PIAS (e.g., a PDA or Personal Digital Assistant). The location of the route selection process distinguishes different Route Guidance operating modes of the architecture.

The method of providing Route Selection in the fully developed architecture, to either a driver in a vehicle, a traveler with a PIAS or at an RTS, or a Commercial Vehicle Manager at a Fleet Management Center (Fleet and Freight Management Subsystem or FMS), is by the ISP. The architecture supports the autonomous mode of guidance (route selection processing in the VS or PIAS) and also supports the mode

of route selection in the VS or PIAS where the ISP provides link transit times and intersection queuing times to the mobile route selection processes.

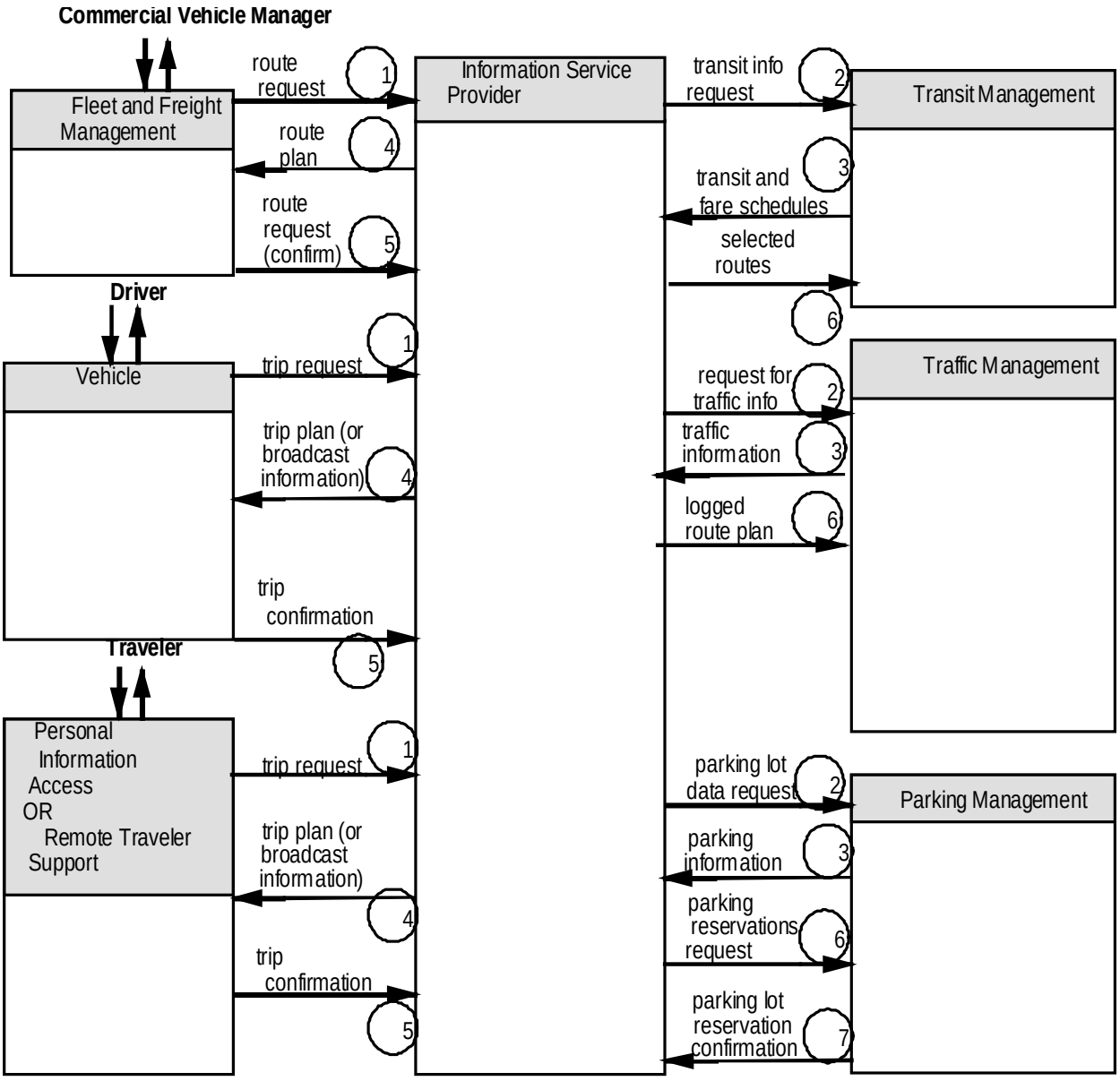

Figure 6. Physical Architecture for Route Guidance

Another key function in providing the Route Guidance user service in the fully developed architecture is the use of a predictive model from which to estimate future transportation link and queue delay times. This process resides in the TMS. Figure 6 shows the high-end state Route Guidance user service where the predictive model resides in the TMS, although subsystem aggregation can be used to include this function of the TMS into an ISP, or a competitive joint venture of ISPs can create a privately operated TMS for the sole purpose of creating a common predictive model. The prediction of link delays (the time to transit a link) and ramp or intersection queue delays (the expected waiting time at a highway on-ramp, off-ramp or an intersection based on a desired turning movement) can be of varying levels of sophistication. In a sophisticated deployment, the link and queue times may be based on the expected statistical occupancy of links and models based on historical data of the relationship between occupancy

and expected (average) link times and queue delays. The expected statistical occupancy of links may be determined by historical time-of-day data, as well as the prior choices of travelers to travel specific routes (routes that have been requested and selected by travelers using an ISP). In Figure 6, these choices are communicated (without any identification of specific vehicles or travelers), by the ISP to the TMS so that the predictive model may be incrementally updated in real-time. In the same way, the expected occupancy of transit vehicles may be updated in real-time through a similar message from the ISP to the Transit Management subsystem.

In Figure 6, "clients" requesting information are on the left side of the diagram, the "server" or ISP is in the center, and other subsystems that may be required to assist the ISP in its server function are located on the right side.

Messages 1, 4 and 5 in Figure 6 represent a request/response/confirmation transaction between client and server, and are required for the advanced route selection function described here. It is possible to deploy a simple route selection system with only message 4, which could contain more general information, not customized to a particular travelers request. The client subsystems would use this general information to compute specific routes for travelers.

Note that messages 2 and 3 in Figure 6, which represent an ISP request for additional information, are optional, in the sense that it is up to the specific implementation of the ISP to decide when to request and how to use the information requested from the auxiliary information source entities (Parking Management subsystems, Transit Management subsystems and Traffic Management subsystems). Ideally, a "just in time" information strategy would have the ISP request specific information from the source on the right as requests need the information. In this way, the information is as fresh and timely as possible. Practical considerations such as communication costs and remote request/response latencies may result in the ISP periodically requesting and storing such information and using the stored information for most client requests. These implementation details, and the resultant level of client service and cost, will allow competing ISPs to differentiate themselves in the marketplace.

Similarly, messages 6 and 7 are optional and only issued when a reservation of some kind is requested by the traveler, e.g.: a paratransit trip reservation (message 6) to the Transit Management subsystem or for a transit vehicle (message 6 for signal priority) or for a parking space reservation (message 6) and confirmation (message 7).

1. Driver, Traveler or Commercial Vehicle Manager interacts with the corresponding subsystem to issue a Trip Request message to the ISP. The trip request can range from simple to complex depending on the requirements of the traveler. The message is a subset of the following. At a minimum, it specifies origin and destination:
    - origin
    - destination
    - departure time
    - desired arrival time
    - constraints
        - number of mode changes
        - acceptable travel time
        - AHS lanes
        - ETA (expected time of arrival) change tolerance.
        - interstate
        - load classification
        - number of transfers
        - special needs

    - urban
    - vehicle type
- preferences
    - modes
    - alternate routes
    - ridesharing options
    - route segments
    - routes
    - transit options
    - weather conditions
- rideshare
    - traveler identity
    - constraints
        - origin
        - destination
        - departure time
        - travel time
        - arrival time
        - acceptable travel time
        - AHS lanes
        - ETA change
        - interstate
        - special needs
        - urban
        - vehicle type
    - preferences
        - alternate routes
        - route segments
        - routes

2. The ISP issues, if necessary, appropriate traffic, parking and/or transit data requests. These requests (and waiting for the corresponding response, message 3 below) may not be necessary if the ISP has recently requested (and stored) this information.

3. The ISP receives current and predicted traffic data (link travel times and ramp or intersection queue delays), parking information (including availability) and/or transit data (available schedules and schedule variances) relevant to the trip request.

Based on the data received in step 3. above, the ISP may make additional requests (step 2. above) to find a better route. These steps are very dependent on the specific implementation of route selection in the ISP. It is expected that there will be considerable innovation in the development of these processes, and the performance of these processes (the quality of the resultant routes and the minimization of the transit/traffic data requests) will determine the cost-performance operating points of competing ISPs.

4. The ISP processes a route and sends it to the requesting subsystem. (For mobile equipment that computes autonomous routes, this flow may not have a route, but simply relevant link travel times or ramp or intersection queue delays that are at variance with the expected values.

   Unless the selected route is reported to the ISP and subsequently to the TMS, there is no possibility in this case to achieve system (near) optimal use of the transportation network.)

Note that this route may be of considerable size, depending on the complexity of the route. The ISP and receiving subsystem may use a data compression/decompression process at both ends to minimize the communications cost of this potentially large message.

The route sent to the traveler can include a route ID number associated with the computed route, so that the traveler equipment can efficiently refer to the selected route in the next step, route confirmation.

The route sent to the traveler includes a list sequence of route segments that the traveler is to follow, with specified start and end locations. This information is sufficient to support turn-by-turn travel instruction to the traveler, thus avoiding the need for an in-vehicle map database and the associated system management and maintenance that a map database requires. This feature is further discussed in Section 5.3.

The trip plan message can include the following items summarized below (considerably more detail can be found in the ITS Logical Architecture document):

- Route ID
- Start time
- List of route segments:
  - segment description
  - segment start and end locations
  - segment estimated travel time
  - segment travel mode
  - segment estimated condition
  - segment predicted weather
  - list of report position locations (locations to report "probe data")
- Current and predicted weather conditions
- Current / predicted incidents with severity, start time, duration
- Paratransit personal schedule
  - Cost
  - Pickup location, destination
  - Service identity
- Rideshare details
  - Selection number
  - Traveler identity

5. The driver, traveler or Commercial Vehicle Manager decides to accept or reject the provided route (for example: making use of a "route preview" process), and may change the route request (going back to step 1 above) or accept the route, issuing the route confirmation message from the receiving subsystem.

Confirming the acceptance of a route may be the basis for real-time, event-driven travel advisories to the driver or traveler, as outlined in Section 4.1.2.1. Because of this, the driver or traveler may be strongly motivated to inform the ISP of the intention to accept a selected route.

6. The ISP sends the selected route to the Transit Management Center and/or the Traffic Management Center (with vehicle classification and occupancy but without traveler or driver/vehicle

identification) so that the appropriate expected statistical occupancy models can be updated, reflecting the incremental congestion and transit time impact that the planned route will have on the transportation network (or transit vehicle occupancy). If a parking reservation is involved, then the parking reservation is requested

7. Parking confirmation is received by the ISP.
8. Although not explicitly shown on Figure 6, the parking confirmation (if present) is passed back to the travelers subsystem.

Note that in the case where there is tight coordination between ISP and TMS, although personal identification of vehicles is not included in the message from the ISP to the TMS (because it is not needed), the *type* of vehicle may be included because different type vehicles will have different impacts on the road network occupancy (i.e. a large truck will occupy more space than a passenger vehicle and will have different acceleration, deceleration, and environmental profiles). Also, vehicle occupancy is communicated in support of traffic control optimization based on movement of people (as opposed to vehicles). Another exception is for commercial vehicles transporting HAZMAT. In these cases, the selected route may also include the hazardous materials manifest, each item identified by the material safety data sheet (MSDS) number (a standard for identifying hazardous materials) and the material quantity. The purpose of this information is for public safety: emergency pre-planning and accelerating incident classification and (appropriate) response.

4.1.3.1 Dynamic Route Selection and Probe Data

A variation of the basic Route Guidance mechanism supported by the architecture involves the mobile subsystem frequently requesting a new route from the ISP at pre-specified waypoints along a route. These additional trip requests can be used to recompute the route for the driver or traveler. If the new route is better than the old route (due to some unanticipated change in link, ramp, or intersection characteristics, e.g., a nonrecurring incident) by some user specified criteria, then the new route is sent to the mobile subsystem as in flow "4" of Figure 6.

The waypoints for the additional trip requests are pre-specified by the ISP in the original route message, and the waypoint route requests need only take the form of a location and a timestamp (the return address of the mobile subsystem is assumed to be included as a part of the message protocol).

Finally, the waypoint specified route locations for updated route request messages serve the ISP as "probe data", which may allow ISPs to construct their own models of the transportation network to be used for route selection. This capability of ISPs to build their own models of the transportation network may be significant in the time period before ISPs and TMSs integrate the ATIS and ATMS functions, or in regions where the TMS chooses not to participate or does not exist, and will allow competing ISPs to differentiate their services. The probe information is determined by computing the timestamp difference between a vehicle's route request message at the current waypoint and the previous waypoint.

4.1.4 Ridematching and Reservations

The Ridematching and Reservations user service is a special type of trip request. It can originate from a Kiosk through the RTS or from the PIAS by computer connections (e.g., a home computer or equivalent appliance) or wireless link (e.g., a PDA). For drivers wishing to participate, the RTS or PIAS function can be aggregated with the Vehicle functions. (Note: although not explicitly supported in the architecture, it is possible that a trip request could be "mediated" for a traveler/driver by an operator who would enter the trip request through a RTS, and report the results from the subsystem to the traveler/driver. In this way, the traveler/driver can gain the benefits of this user service by telephone.

Furthermore, suitable "touch-tone" interfaces could be designed to allow the traveler/driver to interact with the RTS in the same mode as many current inquiry-response systems, e.g., "bank by phone" systems.) The traveler/driver prepares a trip request specifying the rideshare parameters (or specifying demand responsive transit, or any other multimodal form of trip). The ISP subsystem directs the request to the appropriate center if necessary. The ISP provides real time or future matching (if a match is not currently available) for the travelers/drivers request.

The sequence of messages for the Ridematching and Reservations user service is shown in Figure 7 and proceeds as follows.

1. The traveler/driver prepares a trip request (that may include a request for demand responsive transit). This trip request message (contained in the traveler request arch flow) is sent to an ISP. Origins and destinations for the trip request could be specified by phone numbers with the ISP doing a reverse directory search to resolve the numbers to addresses.
2. If the trip request includes demand responsive transit, the ISP prepares a demand responsive transit request to the TRMS. Note that this message step, as well as steps 3, 8 and 9, are dependent on the traveler/driver specifying demand responsive transit for all or a portion of their trip.
3. The TRMS responds to the ISP with a demand responsive transit plan.
4. The ISP processes the complete trip for the traveler/driver, including the computed cost (that the traveler may/must prepay), and sends this trip information in a message to the traveler/drivers subsystem. The ISP may prepare more than one alternative trip information plans for the traveler/driver to choose from.

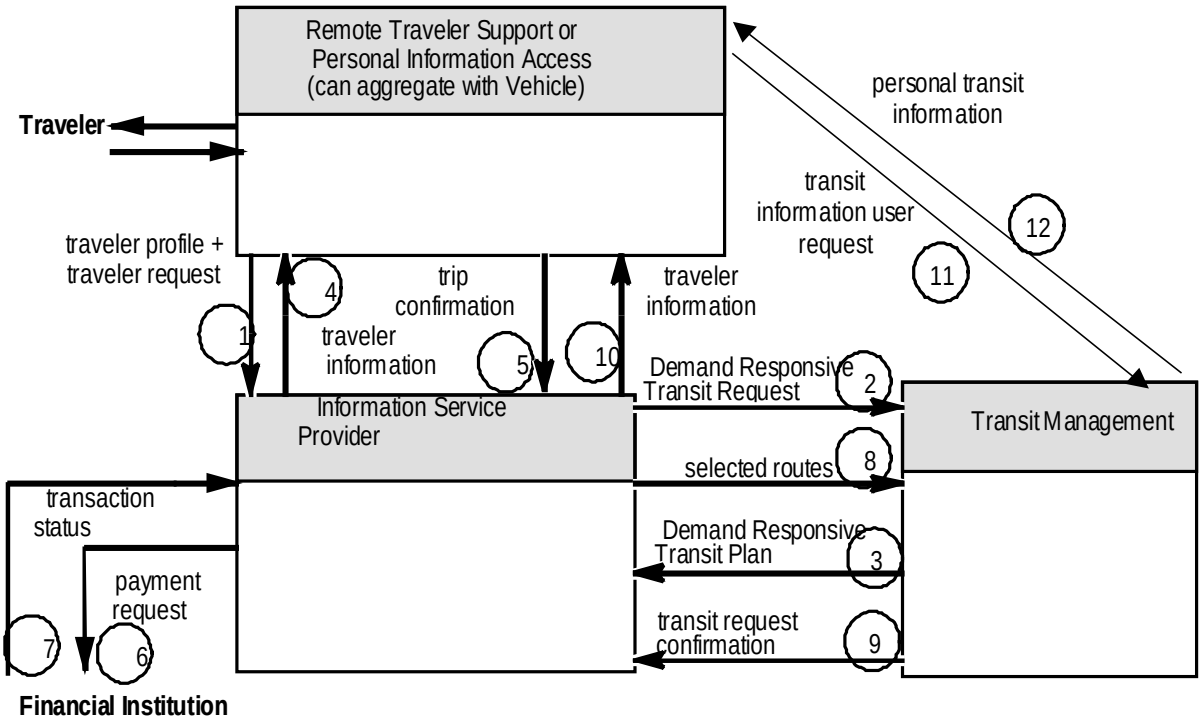

Figure 7. Physical Architecture for Ridematching and Reservation

5. The traveler reviews the trip information, and may change the trip request (i.e. go back to step 1) or select the trip by issuing the Traveler Selection message to the ISP. This message authorizes the ISP to begin processing payment (if any) through an appropriate Financial Institution.

6. The ISP issues a payment request message to the financial institution.
7. The financial institution issues a transaction status message (indicating either acceptance or denial of the payment).
8. If Demand Responsive transit was indicated in the traveler selected trip, then the ISP issues a Demand Responsive transit Accept message to the TRMS.
9. If Demand Responsive transit was indicated, the TRMS issues a Demand Responsive transit Confirmation message to the ISP.
10. The ISP issues a Trip Confirmation message (as part of the traveler information arch flow) to the RTS or PIAS.
11. As an alternative, a simple request/response is possible between the PIAS and TRMS as shown by the traveler initiating through their PIAS (or RTS) a request…
12. …Followed by the TRMS response.

4.1.5 Traveler Services Information

Traveler Services Information includes getting traffic information, transit information, and yellow pages information (e.g., tourist attractions, lodging, and motorist services). This information can be accessed by travelers at kiosks, in transit vehicles, home or office computers, travelers with PDAs, and by drivers.

The yellow pages service includes a two-way reservation capability, including electronic payment for reservations.

The ISP supports data collection from Transit Management (shown in Figure 15) as well as from Information Brokers (for generic yellow pages information) shown in Figure 8. This data collection can be in response to a specific Information Request or may be collected whenever the Yellow Pages Service Providers issue their specific messages. Note that it will be in the interest of the Yellow Pages Service Providers to issue the travel service information message frequently to assure its accuracy and timeliness, since the providers of the information (the product or service vendors) are probably paying the information brokers to get the broadest possible distribution of their listings.

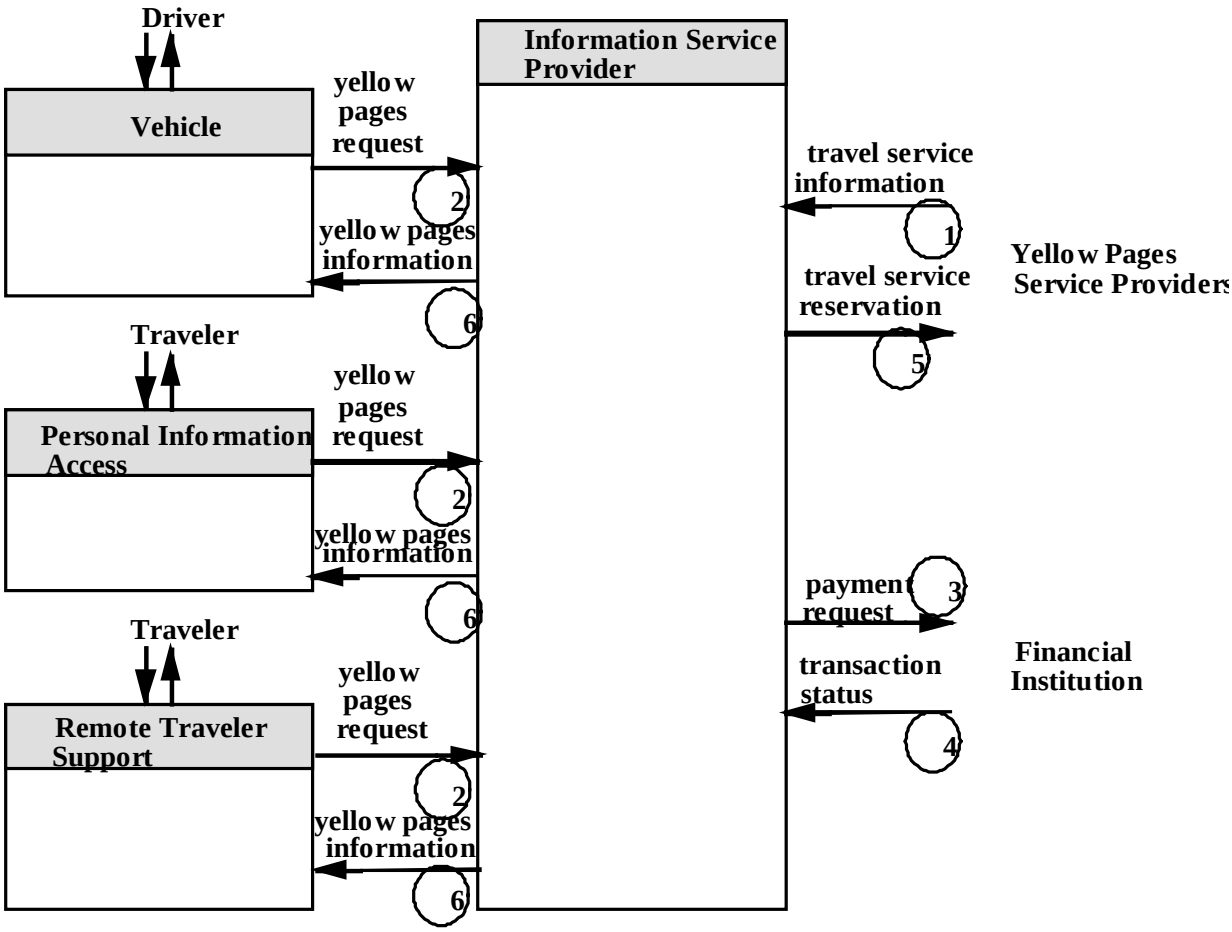

Figure 8. Physical Architecture for Traveler Information Services

Note that the architecture does not specify implementation, so that the RTS of Figure 8 could be implemented with a "kiosk" type electronic interface, or with a human operator implementing the interface. In this way, a traveler could interface with the subsystem through electronic input/output devices, or through the operator, including speaking with the operator over a telephone line.

1. As mentioned before, travel service information from Yellow Pages Service Providers is collected whenever it is made available and then stored in the ISP.
2. Drivers or travelers from the three traveler subsystems shown in Figure 8 prepare and issue a yellow pages request message to the ISP. The request can be only for information, or can request a reservation with optional payment service in support of the reservation.
3. If necessary, the ISP issues a payment request message to the Financial Institution (or Financial Clearinghouse) to transact payment. The payment could be for the information, or (more likely) for reservations for travel services.
4. If a payment request was issued by the ISP, then the financial institution issues an acknowledgement (that payment was transacted or denied).
5. If a travel service reservation was requested, then the ISP issues a message with the reservation and payment confirmation information to the Yellow Pages Service Provider.

6. The ISP returns the yellow pages information message to the requesting subsystem. This message will contain confirmation for the various traveler services requested, if any.

4.1.5.1 Media Information Dissemination

In addition to the Traveler Services Information provided to travelers and drivers shown in Figure 8, the architecture supports the dissemination and collection of traffic, incident, and traveler information by various media. The Physical Architecture for this type of information dissemination is shown in Figure 9.

The "Media" terminator in Figure 9 is a machine interface to the external media's information systems that provide traffic reports, travel conditions, and other transportation -related news services to the traveling public through radio, TV and other media. Media may also be a source for traffic flow information, incident and special event information, and other events that may have implications for the transportation system.

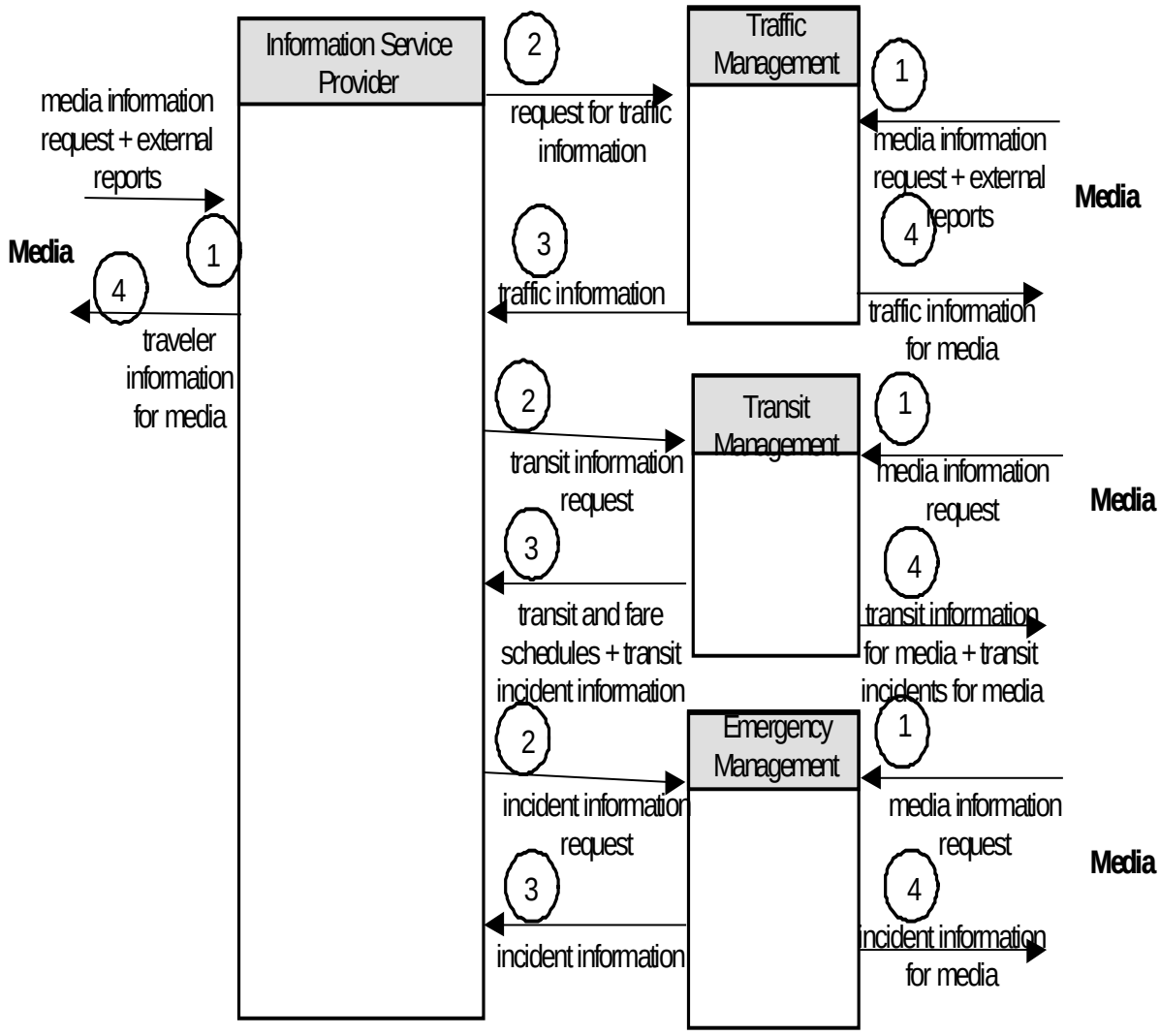

Figure 9. Physical Architecture for Traveler Information Services to the Media

1. In the media support architecture of Figure 9, the Media first specifies and issues a "media information request" message to the ISP, TMS, TRMS or EMS. This message specifies the level of detail and scope of travel information that the Media is interested in being informed about. This message only needs to be specified once, which then allows real-time information for media consistent with the request to be sent to the Media on a regular or event-driven basis by the appropriate subsystem. Also, the Media can send "external reports" based on its own surveillance or other information gathering to the ISP and/or TMS on an event-driven basis. Note that more than one Media may be serviced by the ISP and each may have his own information criteria.
2. The "request for traffic information", "transit information request" and "incident information request" messages, reflecting the geographic scope and depth of detail represented by the union of all "media information request" messages received by the ISP, is sent to the TMS(s), TRMS(s) and EMS(s) and is stored at those subsystems. These messages are used to control search processes that run continuously and issue "traffic information", "transit and fare schedules" and/or "incident information" messages to the ISP on an event-driven basis. Once the ISP has set up the Traffic, Transit and/or Emergency Management subsystems with the information request messages, the ISP will receive relevant notifications of incidents without further requesting.
3. The Traffic, Transit and Emergency Management Center(s) send relevant Traffic, Transit and/or Emergency data to the ISP.
4. The ISP, Traffic, Transit and/or Emergency Management subsystems prepare and distribute customized information to the Media.

4.1.6 Traffic Control

Surveillance data from the roadside (and probe data from vehicles if available) are used to determine the state of the network. Traffic Management and Incident Management equipment packages in the TMS take these inputs and create the signal timing and phasing messages used to control traffic. This control can also include Dynamic Message Sign (DMS), movable lane barriers/markers, or any other roadside traffic control feature (generically referred to as "actuators" in the architecture).

In a highly integrated system, the Transit Center can send requests for signal priority in order to assist specific transit vehicles in returning to schedule. Emergency vehicles can also request signal pre-emption, which is done via the ISP providing route selection for the emergency vehicle. The architecture also supports the current implementations of beacon request for transit priority and emergency vehicle pre-emption.

The ISP for emergency vehicle routing can be a function under the auspices of the Emergency Management agency. An advantage of this architecture deployment is that the TMS, with knowledge of vehicle routes and future expected turning movements, can give a vehicle a class-selective signal priority with minimum disruption to the surrounding traffic. For example, an emergency vehicle can be given a left turn signal, rather than just a green light (or having all signals at an intersection go red, as is sometimes done today). Furthermore, since the TMS can anticipate the future turning movements of the emergency (or transit) vehicle, the signal phases can be adjusted further in advance of the vehicles arrival at an intersection to minimize total system disruption. (Note that some current site specific signal pre-emption systems have the ability to tie in the vehicle's turn signal with the pre-emption message to give turn information to the intersection signal control system. Site specific signal pre-emption systems are useful where centralized control does not exist.)

The Traffic Control architecture is shown in Figure 10, with the following message sequences.

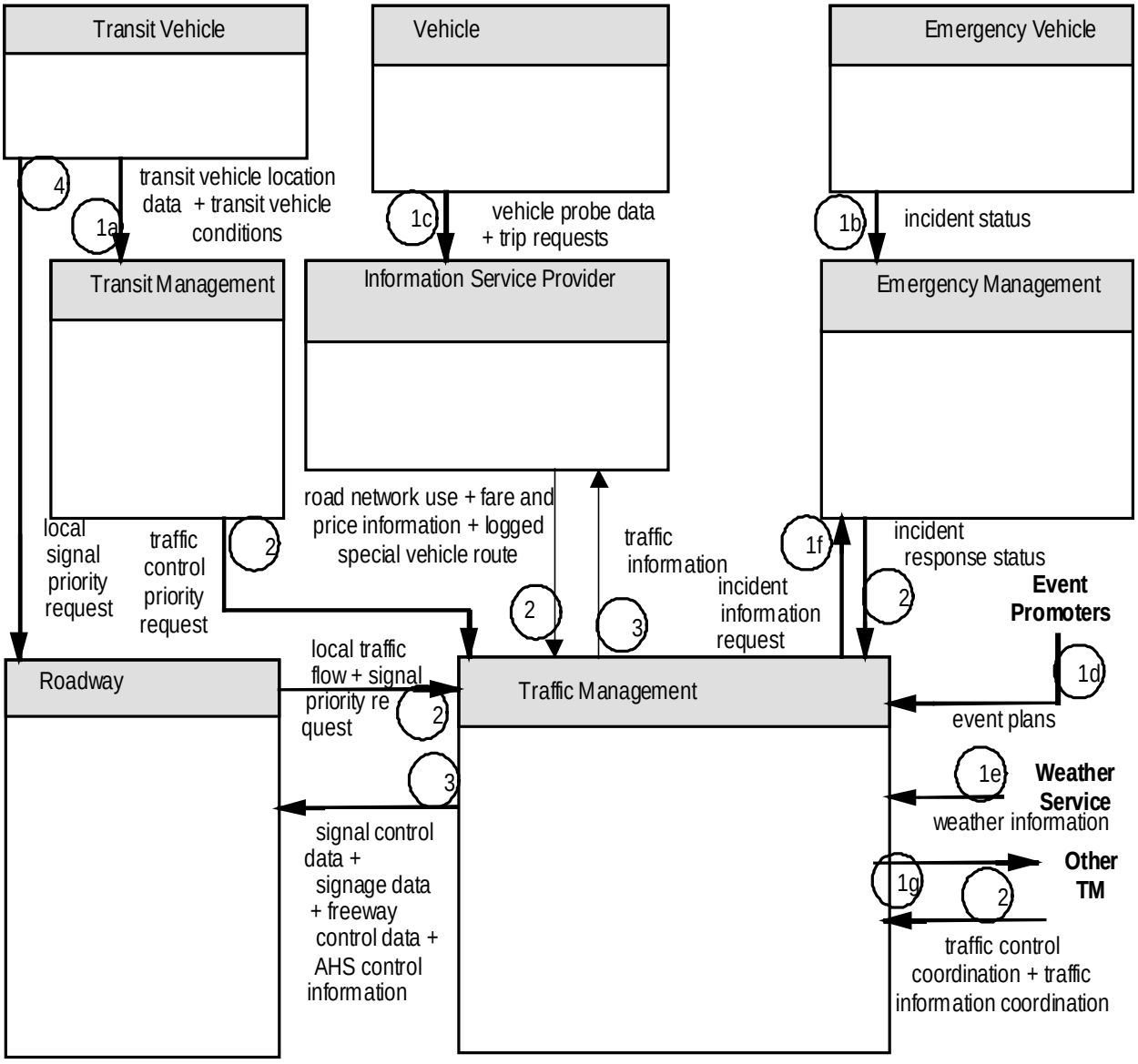

Figure 10. Physical Architecture for Traffic Control

1. Various vehicle traffic management service requirements are collected:

   a. Transit vehicles report their status (including location, a timestamp and occupancy) to the TRMS subsystem. The TRMS processes this data to determine general transit vehicle priority as well an individual transit vehicle priority (e.g., to assist an individual vehicle that is behind schedule to return to schedule).

   b. Similarly, Emergency Vehicles report their current position and selected routes to the EM.

   c. Other vehicles (that opt-in to participate) communicate their selected routes and updated locations to the ISP.

   d. Event Promoters report “Event Info” to the TMS on large traffic generators e.g. sporting events and concerts.

   e. Weather Service “Current and Predicted Weather”.

   f. The TMS can request incident information from the EM.

g. The TMS can communicate with Other TMSs ("Other TM") and request traffic data, predicted incidents and / or current incidents.

2. Roadway subsystem-based surveillance is also sent to the Traffic Management Center subsystem. Included with traditional surveillance data are messages about transit priority and emergency vehicle preemption requests received directly from these vehicles at specific Roadway beacon locations. Also, equipment may be deployed at the Roadway subsystems to read Electronic Toll and Traffic Management (ETTM) tags and use this data as probe information -- which will be forwarded to the TMS as additional surveillance data.

   The EM responds to the TMS request with incident response status.

   In a more advanced scenario, the Other TMSs (the terminator "Other TM" in Figure 10) and ISPs send current position and expected routes and occupancies of vehicles to the TMS. During times of heavy congestion, this may only be for high priority vehicles (Emergency, Transit, HOV, in order of decreasing priority), but in times of low congestion this may include all vehicles opting-in to participate. As communications and processing technology evolve, a larger number of real-time vehicle route schedules can be included in these messages, and the frequency (update rates) of these messages can increase as well. At the same time, the Roadway subsystems are sending surveillance data messages to the TMS.

   The Other TM terminator (representing one or more neighboring TMSs) responds with the "TMC coord" message which contains the "transfer data" logical flow which has the following components:

   - long term data

     This data flow includes the historical patterns for the network covered by the neighboring TMS and includes:

     - parking occupancy data
     - traffic management storage data (control strategies)
     - wide area pollution data
     - processed data (fixed traffic sensor measurements history)
     - o-d matrix
     - current other routes use (scheduled routing for non-vehicular traffic)
     - current stored incident data (current and predicted incidents)
     - traffic flow state (current traffic flow conditions on roads (surface streets), freeways and ramps as well as flows in HOV lanes).

   - cv incidents for other TM

     These are commercial vehicle routes that have been planned and possibly permitted and involve abnormal dimensions or HAZMAT loads. The purpose of this flow is to allow incident preplaning.

   - emergency data for other TMC

     This contains the portion of a strategy that gives priority to emergency vehicles that relates to roads (surface streets) and highways that are outside the area served by the local TMC. This data has been received from other TMC's so the local TMC can implement the requested priority measures to give the emergency vehicles priority throughout their route.

- current data

  This contains data about the current state of traffic on the road (surface street) and freeway network served by the function. It is a sample of the traffic at a single instant in time and is updated periodically
- planned events local data

  This contains information about incidents which have been planned to take place on links in areas served by adjacent traffic management centers (TMC's).
- permit coordination

  This includes permit types, route plans, date/time and special traffic controls.

3. The TMS, using all of the surveillance information available, determines an appropriate signal coordination strategy or signal control output and sends this to the Roadway subsystem. It then updates its predicted traffic model and sends this to the ISP. In a more advanced system, the TMS would simultaneously processes the real-time routes required by emergency, transit, and other participating vehicles to include this in the signal control output.

Coordinated regional traffic management requires that adjacent jurisdictions (which may each have its own TMS) agree on a common traffic management policy. The architecture by itself does not specify a particular policy, since this is a local/regional decision. Agreement and cooperation is a local political process, outside of the National ITS Architecture. If all TMSs in a region cooperate to request and send the "TMC coord" message among themselves, then they will all have exactly the same overall regional traffic database on which to execute traffic control algorithms and policies which are not specified by the architecture, but are implementation dependent. If all TMSs implement identical algorithms and policies (that have been regionally agreed to), then by executing on the common database to control the signals in their individual jurisdictions, they will effectively be executing as if there were one TMS for the entire region.

4. In an additional element of traffic control, which takes place at the roadside, Transit vehicles (or emergency vehicles: the flow exists but is not shown in the figure) may request signal priority service using DSRC beacons directly to the Roadside subsystem. These priority requests may be enabled or blocked by the TMS based on overall transportation network objectives.

4.1.7 Incident Management

Managing the transportation network to minimize the impact of predicted, recurring, and nonrecurring incidents is a primary function of the TMS. The TMS correlates information from a variety of sources: roadway subsystem infrastructure sensors (e.g., loops, video cameras), motorist reports (by cellular or regular phone to 9-1-1 or equivalent agencies (e.g., the older ETS - Emergency Telephone Services) to the EM and then to the TMS), anonymous probe data from instrumented vehicles (including transit vehicles) via ISPs, inputs from outside organizations (such as "weather service" and "event promoters"), and inputs from the EM (which would include 9-1-1 reports).

The TMS can use vehicle probe data, through the ISP, as one indicator of incidents. Probe data is passed from the vehicle to the ISP. The ISP removes the identity of the probe data prior to passing the information to the TMS. The TMS does not need to know whom the data comes from, although the class of vehicle (size, number of axles, or passenger occupancy) could be included. In addition, incidents pertaining to CVO would be passed via ISPs that service commercial vehicles.

Once the data is used by the TMS to detect and classify the incident (possibly including verification, e.g., using video surveillance), an incident response plan is initiated, and information about the incident is shared with the appropriate agencies (e.g., EM, police, fire, HAZMAT response team). As part of the incident response, Actuator Data and DMS data messages are sent to the Roadway subsystem to alter the traffic controls or to present information to drivers and travelers.

Advisory information (including incident information) is communicated to drivers by the ISP and is addressed in the Driver Advisory user service.

There are two related sequences of messages discussed in Sections 4.1.7.1. and 4.1.7.2. that are related to the Incident Management user services, depending on whether the incident is detected and classified by the TMS or by the EM (based on input from either Emergency Telecommunication System or possibly other EMs).

Transit incident security including EM notification is discussed in Section 4.2.4.

Emergency notification to EM of CVO HAZMAT incidents is discussed in Section 4.4.5.

Emergency notification by Vehicle and Personal Information Access subsystems to the EM is discussed in Section 4.5.1.

4.1.7.1 Incident Management Sequence Initiated by TMS

These messages are indicated by the clear circles in Figure 11.

1. The TMS continuously collects data from:
    a. Traffic sensors (e.g., video surveillance in "incident data", flow sensors). In order to collect data, the TMS may control sensors connected to the Roadway subsystem using the flow "surveillance control" (e.g. controlling the pitch, yaw and magnification of a video camera type device).
    b. ISPs. Data on network use derived from probe data.
    c. Weather service data
    d. Event promoter data about upcoming events (e.g., large entertainment events).
    e. Incident alerts from other TMSs.

Based on the collected data, the TMS detects, classifies, and verifies incidents.

2. When an incident is verified, an appropriate incident response is prepared and messages are sent to:
    a. Roadway subsystems to set DMS and actuators (e.g., signal and metering phasing and timing parameters).
    b. EMs to alert them to the incident.
    c. Other TMSs to alert them to the incident.
3. On receipt of a TMS Incident Information message, the EMs will notify:
    a. Other EMs (e.g., separate agency interfaces for police, fire, EMS-Emergency Medical Services) (via incident report arch flow)
    b. Emergency Telecommunication System (via incident notification response arch flow)

4.1.7.2 Incident Management Sequence Initiated by EM

These messages are indicated by the shaded circles in Figure 11.

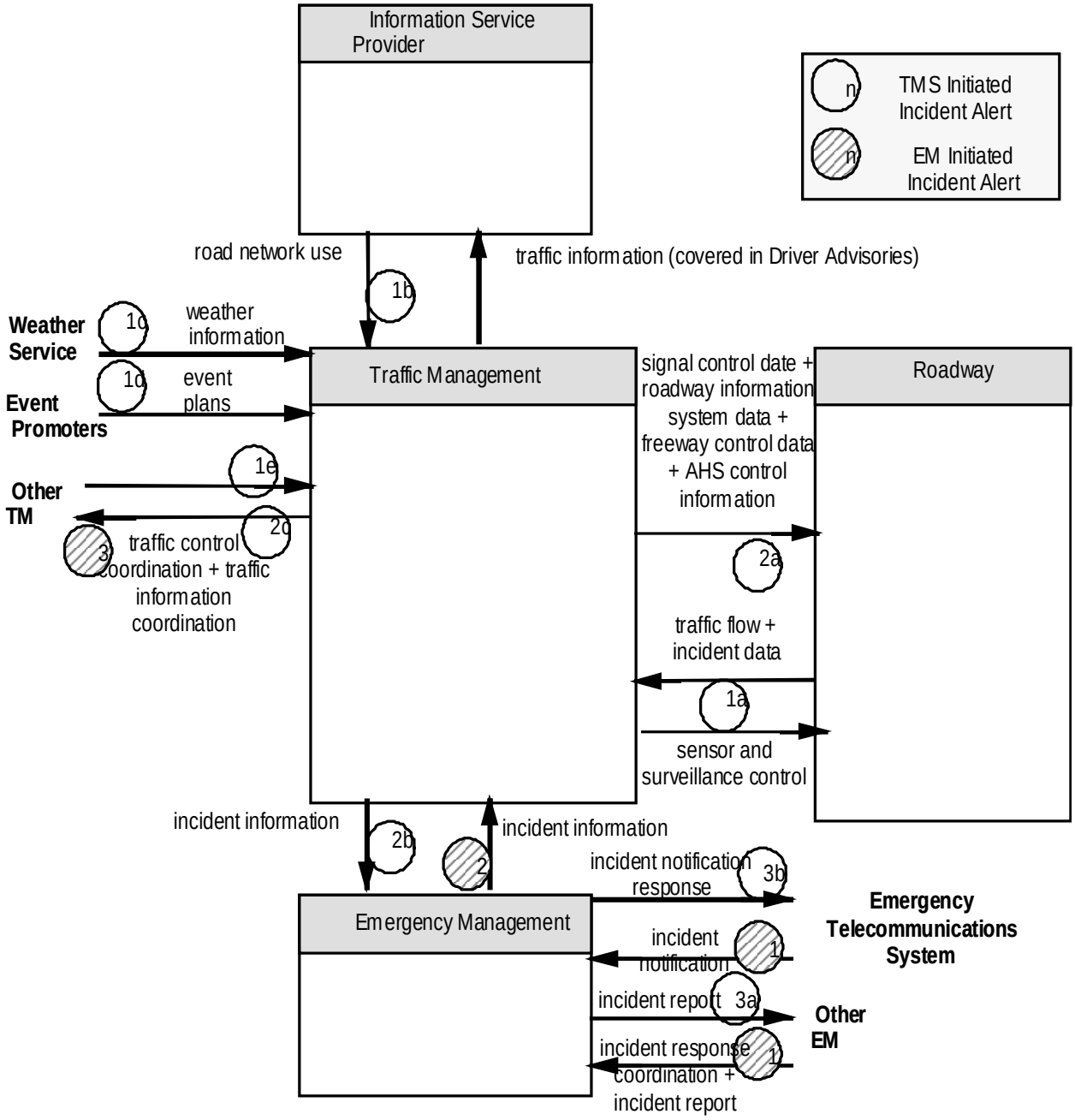

Figure 11. Physical Architecture for Incident Management

1. The EM may receive Incident Notification messages from the Emergency Telecommunication System or Other EMs (via incident report or emergency response coordination flows respectively). This information can be processed and a determination made to forward the incident information message to the TMS.

2. The TMS receives the EM incident information message and, after possibly validating the incident, processes an incident response plan and continues as in 4.1.7.1., number 2. above.

4.1.8 Travel Demand and Emissions Management

Several options are available in the Traffic Management Center Subsystem to set demand management policies:

a. Enforce individual vehicle pollution regulations and HOV occupancy regulations.

b. Adjust DMS, ramp meters, and signals based on demand management strategies. This can be used to give different classes of vehicles different signal priorities with vehicle identification based on sensors and/or toll/ID tags e.g., top signal priority for emergency vehicles, then transit, HOV, etc. (The messages for this process are discussed in the Traffic Control architecture, section 4.1.6.).

c. Adjust prices for parking lots.

d. Congestion Pricing: Using toll/ID tags, the architecture can support toll pricing on links within the transportation infrastructure.

e. Adjust fares/schedules for transit/demand responsive transit vehicles (with Transit Agency approval).

f. Adjust route selection policies. This mode can be used to assist in limiting critical links to their maximum carrying capacity so as to avoid degraded/congested operation of the link.

Transportation demand management processing was assigned to the TMS since the TMS has the broadest overall responsibility for demand management. Demand management comes about for two reasons: congestion and pollution. Both are public agency concerns, and the TMS is generally operated by a public agency.

Although the TMS in the architecture has a role to request pricing action on the part of Toll, Parking and Transit agencies, unless those agencies are under the same public authority as the TMS, there may be difficulties getting institutional cooperation since these price changes may not be in the perceived best interest of these other subsystem operators (and their clients).

The physical architecture for the Travel Demand user services are shown in Figure 12 and have the following message sequences.

The following bulleted transaction sets are each independent of each other and can be deployed independently in a Travel Demand and Emissions Management implementation for a particular region.

- Emissions management surveillance and emissions incident thresholding:

1. Roadway subsystems (RS) sends Pollution Data based on roadway sensor measurements to the Emissions Management subsystem (EMMS). This message can include both general air qualities within each sector of the area and/or monitor the emissions of individual vehicles on the roadway.
2. Based on pollution data, specific pollution reporting criteria are computed at the EMMS and sent to the RSs.

- TMS retrieval of wide area statistical pollution information from the EMMS.

3. The TMS requests wide area statistical pollution information with a pollution state data request to the EMMS.
4. The TMS, ISP and media receive wide area statistical pollution information from the EMMS.

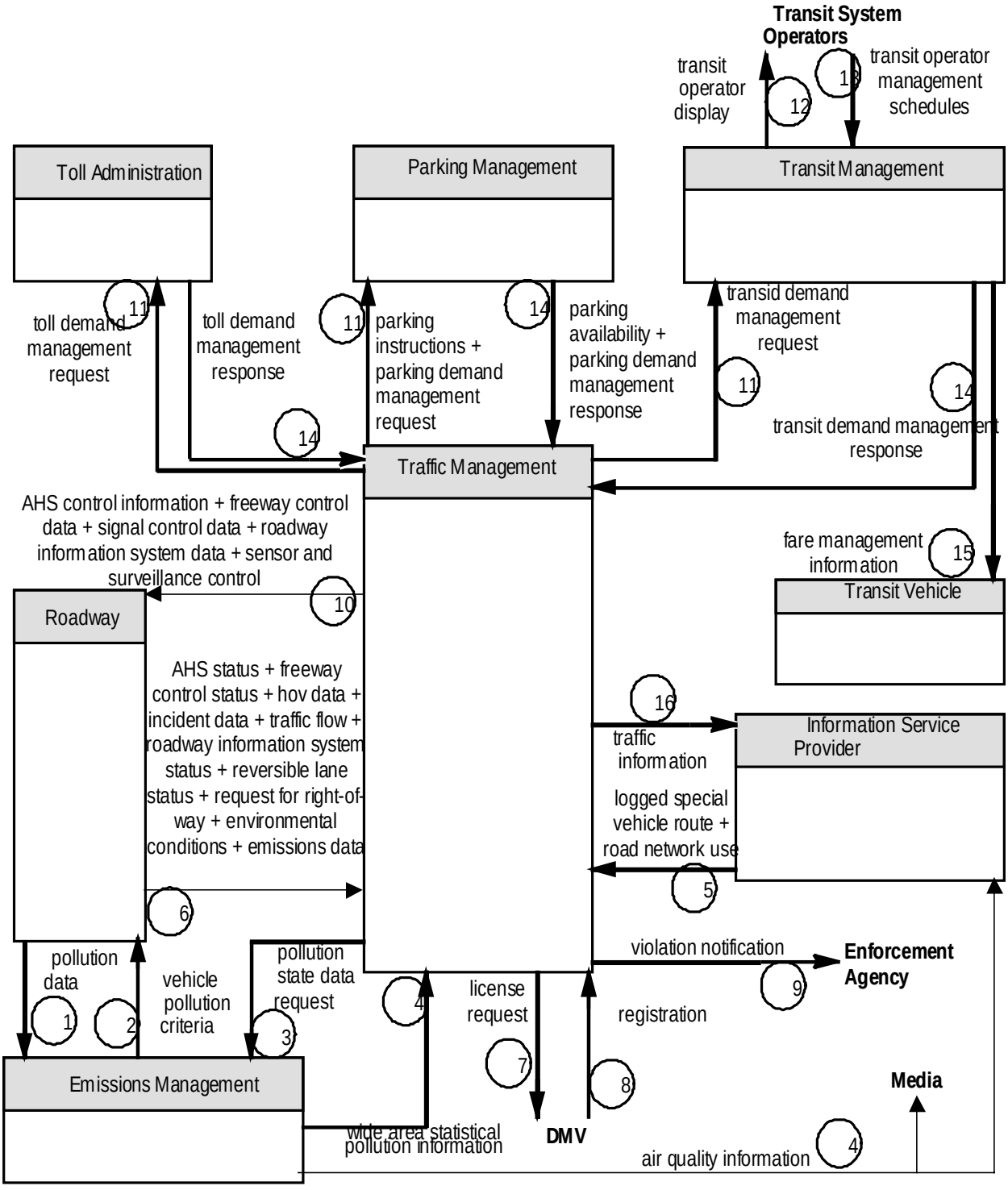

Figure 12. Physical Architecture for Travel Demand and Emissions Management

- As a part of overall surveillance, collect road network use messages from ISPs, and also collect logged route plans to enable anticipation of pollution impacts.

5. The ISPs send logged route plan and road network use messages to the TMS.

- As a part of overall surveillance, collect Roadside information from the RS.

6. The RS subsystem sends Automated Highway System (AHS) status, freeway control status, HOV data (including violation data), incident data (including pollution violation data), local traffic flow and signal priority requests to the TMS.
   - Where emissions violations have been detected, collect registration information and notify appropriate Enforcement Agency.
7. Send license information (vehicle ID) to Department of Motor Vehicles (DMV).
8. Receive owner registration information from DMV.
9. Send violation notification message to appropriate Enforcement Agency.
   - Implement appropriate roadway policies to achieve pollution control objectives.
10. The TMS determines appropriate roadside policies and communicates these to the RS in the messages: AHS control information, freeway control data, signage data, and signal control data.
    - If the TMS determines that demand management by toll, parking, or transit pricing needs to be initiated, then it can request these changes.
11. The TMS communicates updated price data for Tolls, Parking, and Transit to the appropriate subsystems.

Any fare change for transit has to be confirmed by the Transit System Operators using the following two messages.

12. The Transit Management subsystem (TRMS) notifies the Transit System Operator of the TMS fare change request.
13. The Transit System Operator may make changes to the fare schedules in the TRMS.
14. Toll Administration, Parking Management and Transit Management subsystems notify the TMS of their cooperation in the various "…demand management response" messages.
15. The TRMS notifies Transit Vehicle subsystems of any fare changes through the "schedules, fare info request" message.
    - In advanced future deployments where link occupancy is controlled to achieve either stable traffic flow or pollution level management through dynamic traffic assignment (as opposed to controlling occupancy through metering controls at the Roadside as is done today), the following message is used.
16. The TMS can limit critical link occupancy by limiting link capacity in its predictive model. This updated model is sent to ISPs that compute routes for their clients based on available link capacity in the message "traffic information."

4.2 Transit

4.2.1 Transit Management

4.2.1.1 Operations

Transit operations involve collecting information about the locations of vehicles and assessing schedule variances, preparing real-time instructions to drivers to recover from variances in order to maintain schedules, as well as making appropriate requests for customized signal service from the TMS. The

TRMS also collects information about passenger use of the Transit service. The Transit Operations architecture is shown in Figure 13. The architecture delivers passenger loading information to the TMS for the purpose of delivering priority treatment. This may be a consideration since the schedule deviations of a transit vehicle may have impact related to the number of passengers and expected passengers.

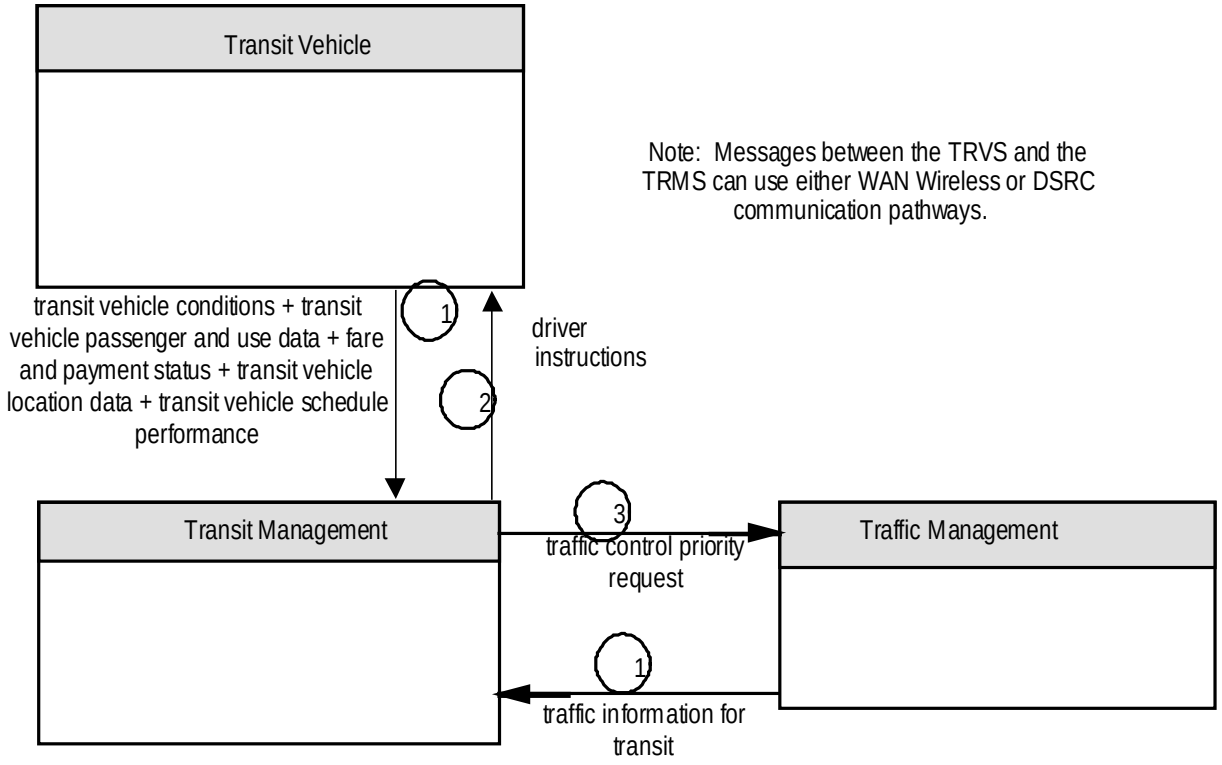

Figure 13. Physical Architecture for Transit Operations

1. Transit Vehicles report to the TRMS: their location, operations data (e.g., occupancy and fares collected), and current running times (see Section 4.1.6., Traffic Control and Section 4.3.3., Transit Fares). The frequency of this message is locally determined, but might reasonably occur at departure from each stop (on an urban fixed route).

   At the same time, the TMS sends Predicted Traffic Data and Link Parameters to the TRMS.

2. Based on the actual vehicle status, the TRMS processes and sends to the driver of the Transit Vehicle subsystem (TRVS): schedule deviation information and (using the Predicted Traffic Data and Link Parameters from the TMS) processes and sends corrective instructions (to reduce or avoid the schedule deviations) on an as-needed basis.

3. The TRMS can send to the TMS two types of messages(contained within the request for transit signal priority arch flow) to request signal priorities or priority for specific transit vehicles (see also Section 4.1.6., Traffic Control, for direct TRVS to TMS signal priority request) on an as-needed basis:

   a. Overall priority: This indicates which classes or routes of vehicles should be given priority, and

b. Vehicle Priorities: this indicates specific vehicles at specific locations needing priority to maintain a schedule.

4.2.1.2 Planning and Scheduling

The TRMS generates schedules off-line for the efficient use of transit resources based on historical and recent ridership data collected on the transit vehicles and reported to the TRMS, as shown in Figure 13 for Transit Operations.

Based on real-time vehicle status, and using the predictive traffic models of the TMS (see Section 4.2.1.1.), the TRMS determines transit vehicle schedule deviations and Expected Times of Arrival (ETAs). As shown in Figure 15, En-Route Transit Information, these are communicated to the users on transit vehicles, at transit information kiosks at transit stops (Remote Traveler Support subsystem), and other locations via the ISP.

4.2.1.3 Transit Personnel Management

The Transit Personnel Management architecture is shown in Figure 14.

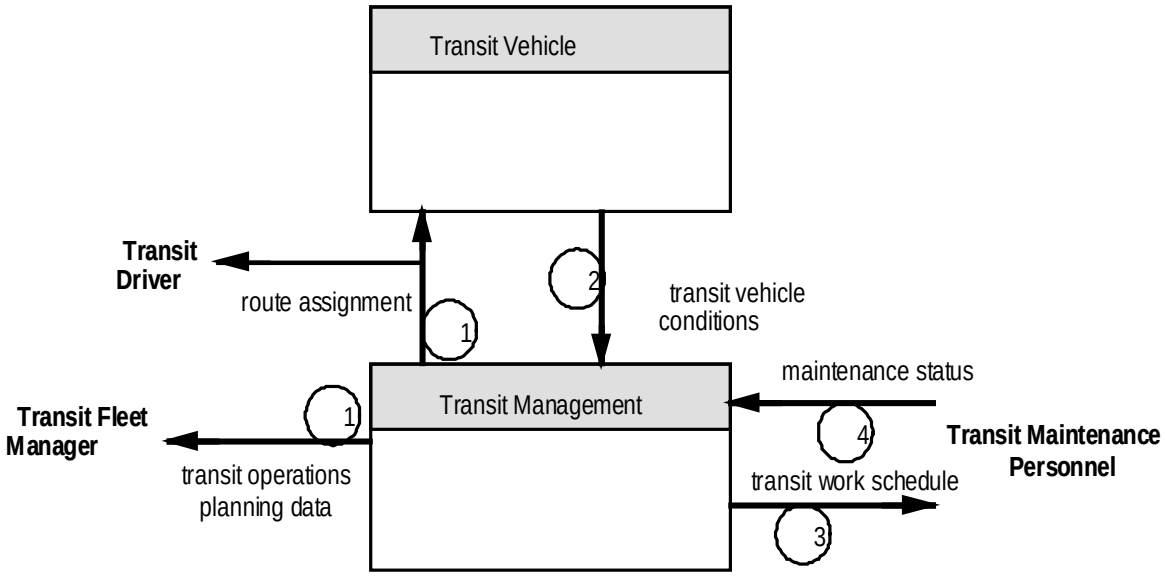

Figure 14. Physical Architecture for Transit Personnel Management

1. The TRMS will generate driver assignment information using the *Vehicle Status* data from the individual transit vehicles to validate driver work hours.
2. During operations, the TRVS will periodically send Transit Vehicle Condition information to the TRMS so that it can plan maintenance.
3. Bus Mileage is tracked through the transit vehicle measures message and this, along with the vehicle conditions data and a database of service completed, can generate maintenance assignments for Transit Maintenance Personnel.
4. On completion of a maintenance assignment, the Transit Maintenance Personnel record the work completed.

4.2.2 En Route Transit Information

En Route Transit Information is available to travelers from a variety of sources. The service can be accessed via the ISP or the TRMSs from the RTSs (e.g., kiosks or automated transit stops) or from the PIAS via a PDA or home/office computers. In addition, the architecture recognizes that TRMSs will have dedicated means of providing en route transit information in the TRVS. Figure 15 below shows a direct interface from the TRVS to the TRMS which can be via the Wireless WAN communications (e.g. SMR or cellular, or via DSRC beacons (via the RS).

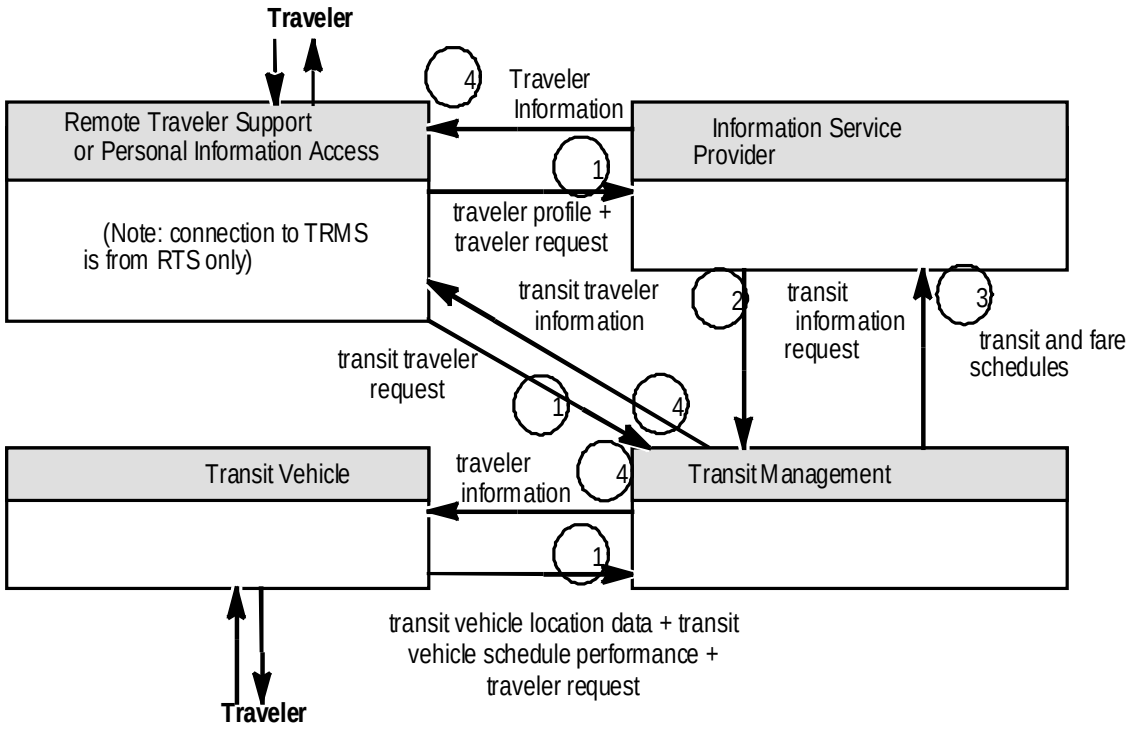

Figure 15. Physical Architecture for En Route Transit Information

1. Travelers interact with one of three subsystems (see Figure 15) to issue a Transit Traveler Request message to the ISP or TRMS. Note that these subsystems may support general traveler displays, and may periodically request update information.

2. The ISP will periodically (or on receipt of a Transit Information Request message) issue a Transit Status Request message to the TRMS. Note that this flow (and flow 3) are not directly tied to the traveler request (flow 1).

3. The TRMS will assemble relevant Transit Status for the ISP request and send this message to the ISP.

4. Travelers receive the requested Transit Information message via the subsystem they are interacting with. (If the traveler is on a Transit Vehicle or is communicating directly with the TRMS from a transit stop, then steps 2 and 3 above are not relevant.)

4.2.3 Personalized Transit

The ISP can serve as a "travel agent" for various modes of transit, including personalized transit. (A Transit System Operator or company could deploy its own "ISP" for this purpose.) The TRMS computes the logistics for individual requests and communicates the personalized schedules to the drivers and travelers via the ISP. This architecture was described in section 4.1.1., Pre-Trip Travel Information.

4.2.4 Transit Security

The Transit Security architecture is shown in Figure 16. The messages associated with this user service are described as follows.

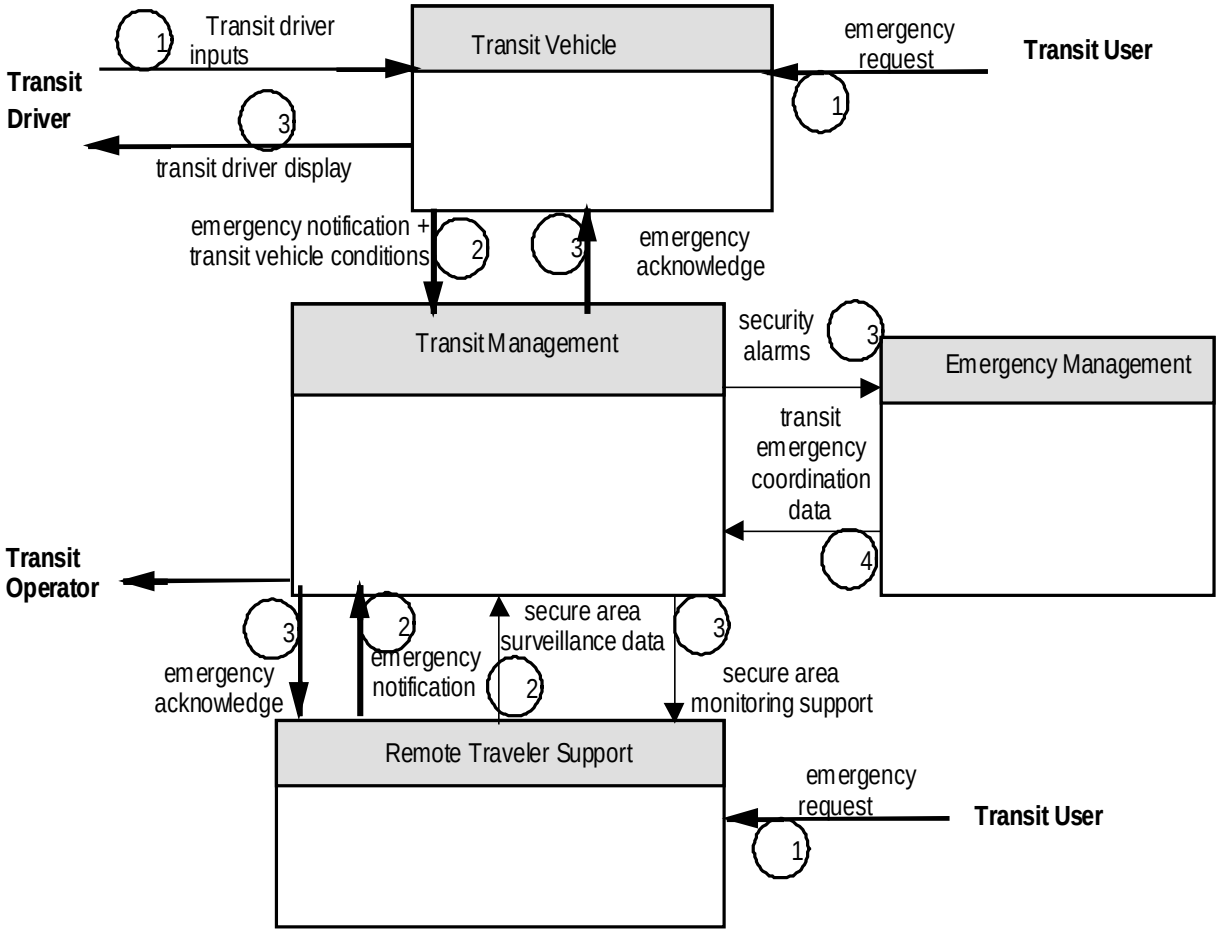

Figure 16. Physical Architecture for Transit Security

1. An emergency message to the TRVS indicating a vehicle emergency can be initiated either by the driver, a passenger or by sensors monitoring the Transit Vehicle Environment.

   Similarly, the RTS, deployed at transit stops, can collect emergency requests by passengers and perform sensor surveillance of the transit stop for incidents.

2. The TRVS or RTS combines all inputs and constructs an emergency request message that is sent to the TRMS. Similarly, the Transit Stop Environment is monitored by the TRMS.

3. The TRMS acknowledges the message from the TRVS with the planned response (which can be presented to the driver "silently") and from the RTS with Transit Stop Response data. In either case, the Transit Management Dispatcher may customize the messages sent to the remote subsystems and may include verbal or data messages. The TRMS may also control surveillance cameras at remote sites. The TRMS sends a transit emergency data message to the EM to coordinate other emergency service response and/or notify Emergency Telecommunication System.
4. The EM will determine the appropriate response and send a transit emergency coordination request message to the TRMS indicating the actions to be taken.

4.3 Electronic Payment Services

The architecture and resultant message sequences for Roadway Tolls, Transit Fares and Parking Payments are very similar, enabling selection of a technology that could serve all ITS payment services.

4.3.1 Features of Payment using Financial Instrument Cards

Three types of financial instrument cards are supported by the architecture.

4.3.1.1 Stored Value Cards

These cards have a an encrypted value stored in memory on the card which is decremented at the point of sale.

Off-line reconciliation may be necessary if the card supports more than one vendor, e.g., for multimodal travel and/or other purchases. (The "vendor" in this discussion can be a toll agency, transit agency, or parking lot operator. For example, if a transit agency supports stored value cards only for its own transportation services, reconciliation is not necessary. On the other hand, if the transit agency subscribes to use a card that can also be used for purchases of goods and/or services elsewhere, then reconciliation is necessary. Reconciliation requires that the point-of-sale equipment reads from the card financial institution and account information so that a back-end electronic money transfer can be effected for the sale. Reconciliation for individual cards may also be used to detect fraud. When reconciliation is used, each card has an individual reconciliation account associated with it, and its value is the amount of money "stored" on the card minus the amount of all transactions. It is up to an individual agency or company to decide whether or not to issue their own stored value cards or to use a "generic" card requiring reconciliation. If issuing their own card, they also have to decide whether or not to support reconciliation. This last consideration may be based on how secure they believe the encryption technology used to store the current value of a stored value card to be.

Cards can be purchased and value added to them with cash, providing total anonymity to the purchaser -- although purchases with credit/debit cards may offer convenience and traceability benefits to the purchaser as well.

4.3.1.2 Debit Cards

This card provides a financial institution identifier and an account number to be immediately debited at the time of the sale.

Debit cards (and credit or charge cards) entail some financial risk on the part of the vendor unless they query the financial institution for the debit transaction at the point of sale. This is necessary to avoid later transactions that are denied due to lack of funds. This real-time financial institution transaction is not practical for toll, transit, or many parking applications. A strategy that has been worked out by the

financial issuers of these cards for these situations is called "preauthorization." In preauthorization, a fixed amount of account balance or credit is put aside for the vendor to charge after the card owner has made one or more purchases. Since the funds are set aside, there is no fund availability risk to the vendor. After a fixed period of time, or after a number of transactions, an off–line charge against the set aside funds can be made to charge for the received services. This "preauthorization" mechanism of "prepayment" also is able to minimize individual transaction charges that card issuers may impose on vendors. These charges can be particularly onerous as a fraction of total charged services for relatively small charges such as some short distance tolls or some transit fares. Using preauthorization for debit, credit, and/or charge cards can enable them to approach the convenience of stored value cards for the vendor, with the benefits of convenience and traceability for the traveler (but without the benefit of total anonymity that a cash purchased stored value card provides).

4.3.1.3 Credit or Charge Cards

Similar to debit cards, except that the initial source of funds is from the card issuer, which is extended as a short term loan to the card user. With a charge card, users pay back the card issuer in full on a monthly basis. With a credit card, users have the option to pay back the card issuer in full similar to a charge card, or can accumulate some portion of the charges to a longer term loan by the card issuer to the card user.

Note that debit, credit, and charge cards are evolving features of stored value cards independent of ITS. The impact is to enable small value purchases while minimizing reconciliation transactions. The features of these cards will include the ability to add value to the cards from any bank ATM terminal, or possibly from personal computers with card proximity read/write capability and a connection to the Internet.

4.3.2 Roadway Tolls

Toll tags can be used to pay tolls using conventional methods. A toll tag is a stored value card combined with DSRC technology. Financial (as well as other transactions in the architecture involving personal or confidential data) are secured by state-of-the-art encryption and authentication processes incorporated in the Physical Architecture Communications Layer.

There is nothing in the architecture that explicitly disallows traditional cash mechanisms for toll payment. As is shown in the details which follow, if a road operator chose to use toll cards exclusively, then the toll cards may still be purchased with cash, maintaining complete anonymity.

The architecture and message sequence for Roadway Tolls is shown in Figure 17.

- The following message is issued by the Toll Administration subsystem (TAS) to prepare the Toll Collections subsystem (TCS) for operations.

1. The Toll Administration subsystem will issue "toll instructions" messages to the Toll Collection subsystems. This message identifies toll rates as well as the serial numbers of lost, stolen, and defective toll tags. Lists of pre-authorized charge, credit, and debit cards may also be included.

    - The following message represents information that is optionally transferred from a travelers Payment Instrument to the toll tag located in the vehicle. This message is optional in that some toll tags will have implicit payment instrument identifiers preprogrammed into it either at the toll tag point-of-sale or during toll tag manufacturing.

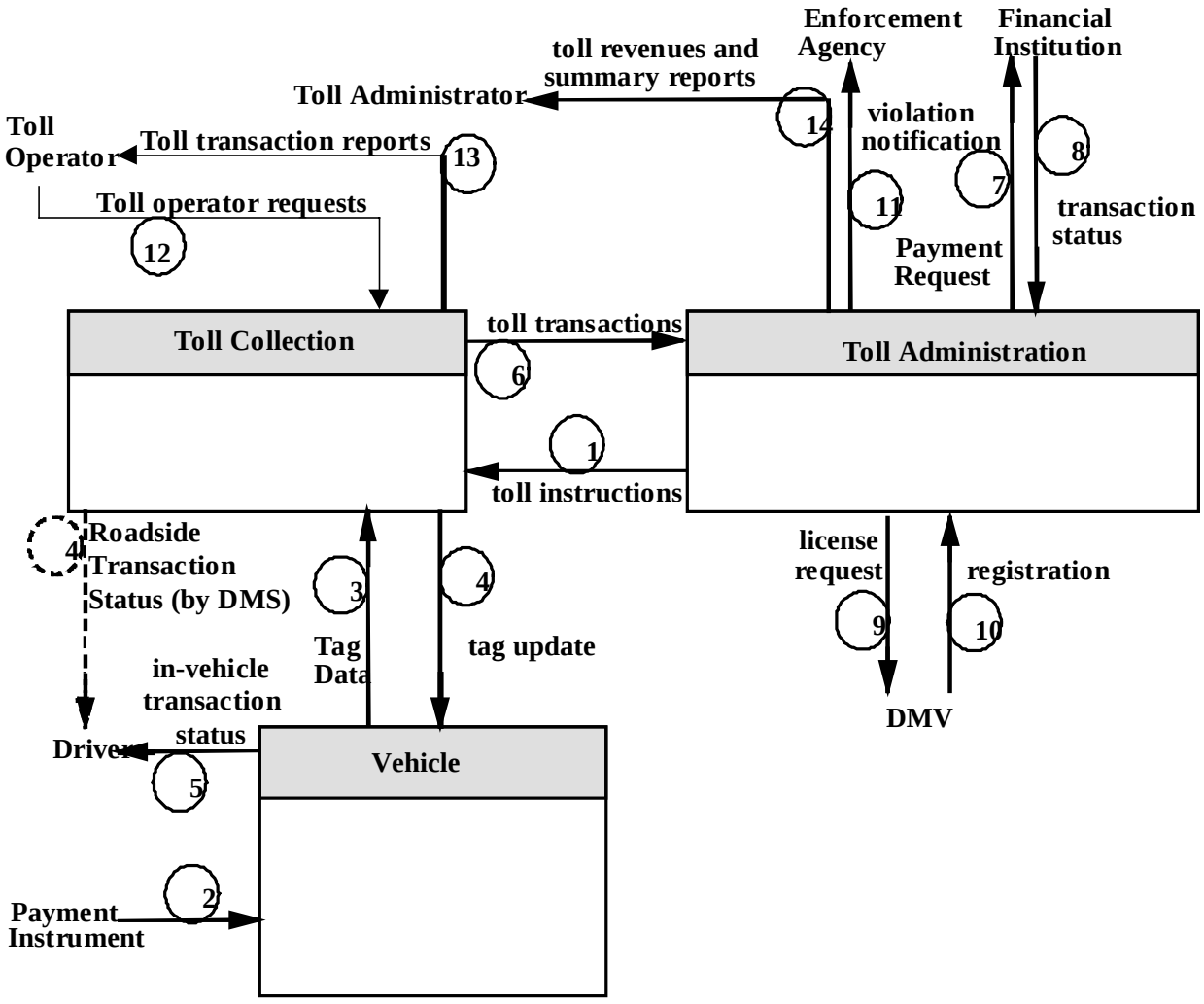

Figure 17. Physical Architecture for Roadway Tolls

2. Financial institution and account information is transferred from a payment instrument to the vehicle toll tag equipment.

   - The following sequence of three messages represent a toll transaction.

3. When a vehicle passes a TCS, a tag data message is sent from the vehicle to the TCS identifying the toll tag serial number and remaining value (positive balance). The technology selected for this interaction must assure that it is "impossible" to counterfeit (or "spoof" as is used in the current encryption-authentication technology jargon) the toll tag message.

4. The TCS rapidly compares the tag serial number with the "Bad Tag List" (lost/stolen/defective tags) previously received from the TAS in the toll instructions message, and confirms that the value stored on the tag is sufficient. The result of this analysis is used to produce a tag update message which is sent back to the vehicle tag. This message is also used to deduct that value of the toll from the tag balance.

   The Roadside Transaction Status may also be communicated to the driver through a two-state DMS e.g., either a PASS (green) message or a PULL-IN (red) message. This optional message is indicated with the dotted line in Figure 17.

If a payment irregularity is detected by the TCS, this information is included in the toll transactions message sent to the TAS.

5. The Transaction Status may also be communicated to the driver through in-vehicle signage equipment.
6. At longer intervals than individual toll transactions, the TCS will issue a toll transactions message to the TAS for "Toll Back-End" reconciliation processing. This message will include violation data.
    - The two messages which follow happen only if the toll system has been designed to accept payment instruments issued by other institutions.
7. The TAS issues a payment request to the financial institution with financial institution and account identifiers chosen by each customer for the customers charges.
8. The Financial Institution uses traditional financial methods (e.g., using financial clearinghouses) to secure the funds. After the clearinghouse transaction is complete, the Financial Institution issues a Payment Acknowledgement message to the TAS indicating that either the payment was secured or denied.
    - Toll violations can be reported to the appropriate enforcement agency as in the following three messages.
9. The TAS reviews violation data from the TCS. For specific vehicles that may have violated the toll requirements, a request is made to the DMV to determine the characteristics of the specific vehicle (e.g. number of axles and commercial or non-commercial status) as well as vehicle registration information. These DMV requests are send to the DMV in the message "license request."
10. The DMV sends the specific vehicle registration information to the TAS in the "registration" message.
11. Toll violation notification messages are sent from the TAS to the appropriate Enforcement Agency (possibly more than one agency).
    - The following messages are each independent of other messages as well as each other and support the general operations of the tolling subsystem.
12. The Toll Operator will have the ability to control the TCS (e.g., lane control, price overrides, violation data overrides) using the toll operator request message.
13. The TCS will update the Toll Operator interface to indicate the operational status of the Toll Collection subsystem using the toll transaction reports message.
14. The TAS will issue periodic "toll revenues and summary reports" to the Toll Administrator.

4.3.3 Transit Fares

Fare collection is handled via a tag, carried with the traveler, which communicates with the transit vehicle for payment.

Tags can be used to pay fares using conventional methods as they are in some transit systems today, i.e., a traveler carried tag is "read" on entering a transit vehicle and may be read again on leaving a transit vehicle (if payment is based on distance traveled). The specific tag technology is, of course, not specified in the architecture. The tag is instantiated with "value" or value added at a sales location (or possibly remotely). Value instantiation is viewed as outside the architecture. Financial (as well as other transactions in the architecture involving personal or confidential data) are protected by state-of-the-art

encryption and authentication processes incorporated in the Physical Architecture Communications Layer.

The architecture and message sequence is shown in Figure 18.

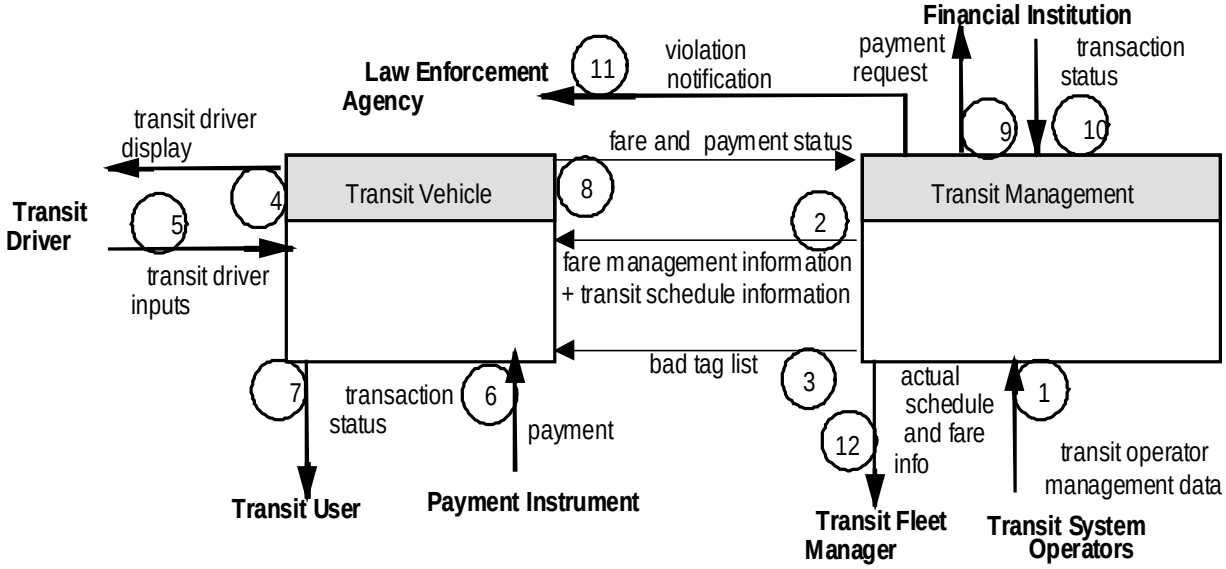

Figure 18. Physical Architecture for Transit Fares

- The following two messages are issued as necessary, initiated by the Transit System Operators.

1. The transit system operators establish the fares and schedules in a message to the TRMS.
2. The TRMS forwards schedules and fare information to the transit vehicles as necessary. In some cases, this may be when a vehicle is put into service at the transit "garage."

- The following message is issued as necessary, initiated by the TRMS.

3. The TRMS will issue "bad tag list" messages to the TRVS. These lists identify lost, stolen and defective payment instruments (to the extent that they can be identified). This message could also be used to identify financial instruments that have an authorized prepayment "hold" in place, so that they can be used for small value fare transactions.

- The following two messages will occur on an event driven basis.

4. The TRVS will update the Transit Driver display to indicate the operational status of the Fare Collection equipment.
5. The Transit Driver will have the ability to control the Fare Collection equipment (e.g., lane control, price overrides, violation data overrides).

- The following two messages will occur in sequence for each passenger paying a fare.

6. When a traveler with a fare tag enters a Transit Vehicle with an electronic payment capability, a tag data message is sent from the tag to the transit vehicle identifying the toll tag serial number and remaining value (positive balance) or in the case of charge, credit, or debit cards, financial institution, and account information. The technology selected for this interaction must assure that it is "impossible" to counterfeit (or "spoof" as is used in the current encryption-authentication technology jargon) the toll tag payment message. Depending on the payment strategy chosen by the

transit system, payment may be made on entry to the transit vehicle (or system), or on a second tag transaction on exiting the transit vehicle or system (so that the charge can be distance based).

7. The TRVS rapidly compares the tag serial number with the "Bad Tag List" (lost/stolen/defective tags) previously received from the TRMS, and also confirms that the value stored on the tag is sufficient. The result of this analysis is used to produce a Transaction Status message which is immediately communicated to the traveler. This message is also used to deduct the value of the fare from the tag balance for positive cash balance instruments. The transit vehicle should have some mechanism for accepting cash to pay the fare directly or to add value to the payment instrument by either cash or financial transaction. If a payment irregularity (possible unpaid fare) is detected, then a message is sent to the TRMS in the "fare and payment status" message where it is recorded and forwarded to the appropriate law enforcement agency for action.
    - The following message is asynchronous to the fare collection sequence of messages. It can occur, for example, on return to the garage after an operational shift.
8. The TRVS will issue a "fare and payment status report" message to the TRMS for fare back-end reconciliation and violation processing.

The following two messages are issued only if payments are supported which use credit, charge, debit or positive cash balance cards which are not issued by the transit management agency.

9. The TRMS issues a payment request to the financial institution.
10. The Financial Institution uses traditional financial methods (e.g., using financial clearinghouses) to secure the authorized funds as specified by the travelers payment instrument. After the clearinghouse transaction is complete, the Financial Institution issues a transaction status message to the TRMS indicating that either the payment was secured or denied.
    - The following two messages are issued as necessary, initiated by the TRMS.
11. A violation notification message is sent to the appropriate Law Enforcement Agency based on in vehicle information from the fare and payment status message and/or other back-end processing.
12. Periodically the TRMS issues "actual schedule and fare info" operations reports to the Transit Fleet Manager.

4.3.4 Parking Payments

Drivers and Travelers, from either their Vehicles or RTS or PIASs, can make Parking reservations in advance with parking lots as part of a Trip Plan, being assured of a space on arrival. This transaction is addressed in the Pre-Trip Planning and Route Guidance sections of this chapter. If the operators of a Parking Management facility wish to allow drivers to transact directly with them (bypassing a third party ISP), this can be accomplished by the Parking Management subsystem (PMS) operator deploying their own ISP (possibly co-located with the PMS) solely for the purpose of supporting a direct Parking Reservation service.

The Parking Payment architecture is shown in Figure 19.

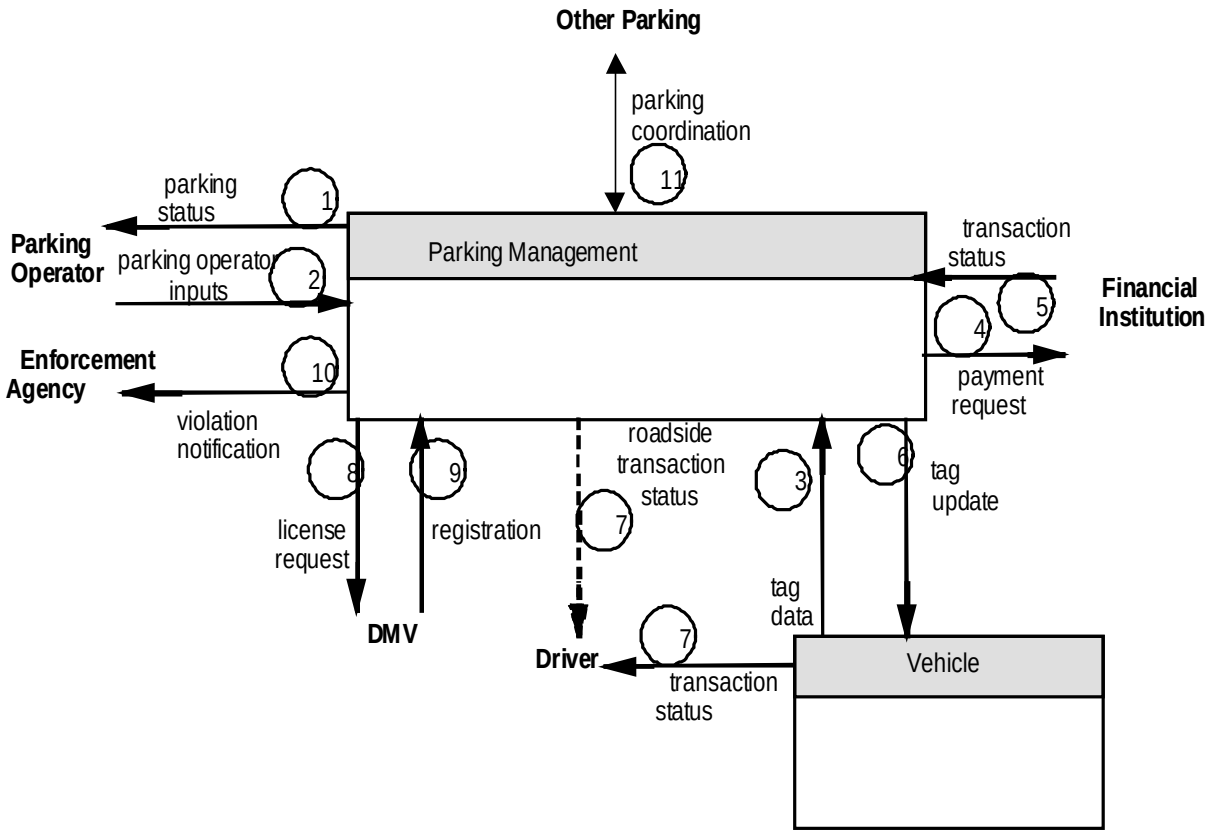

Figure 19. Physical Architecture for Parking Payments

- The following two messages are issued on an event driven basis and independent of other messages.

1. The PMS will update the Parking Operator interface to indicate the operational status of the PMS and to issue periodic operational reports.

2. The Parking Operator will have the ability to issue parking inputs to the PMS (e.g., lane control, price overrides, violation data overrides).

- The following messages 3. through 7. are issued on entrance and/or exit from a parking facility depending on the payment scheme chosen by the parking facility. For example, fixed price parking can be charged on entrance or exit. Time based charges can be made on exit, after reading the tag ID on entrance and recording the time in the PMS.

3. Tag data is read from the vehicle on entrance and/or exit to/from the parking facility. If a non-cash card transaction, the PMS rapidly compares the tag serial number with the "Bad Tag List" (lost/stolen/defective tags) that is internally maintained, and also confirms that the value stored on the tag, if it is a cash balance card, is sufficient for payment.

- Messages 4 and 5 which follows are issued only if the driver is using a financial instrument not issued by the Parking Management subsystem or associated agency.

4. The PMS issues a payment request to the financial institution that they have chosen.

5. The Financial Institution uses traditional financial methods (e.g., using financial clearinghouses) to secure the authorized funds as specified by the vehicle tag. After the clearinghouse transaction is

complete, the Financial Institution issues a transaction status message to the PMS indicating that either the payment was secured or denied.

6. If payment was secured, the PMS issues a tag update message to decrease the stored value of the tag as per the drivers original instructions.

7. The result of the tag transaction is issued to the driver (tag value deduction and/or other transaction information e.g. entry time, total elapsed parking time. This message can be communicated to the driver by in vehicle signage associated with the tag or directly by the PMS e.g. by DMS and/or printed receipt.

    - If a payment irregularity is detected by the PMS, then the following three messages are processed by the PMS.

8. A license request is issued to the DMV.

9. The "vehicle characteristics" message is received from the DMV.

10. A "violation notification" message is sent to the appropriate Enforcement Agency.

    - The following message is issued on an event driven basis independent of other messages.

11. Representing another parking facility, system or subsystem, the Other Parking terminator provides a source and destination for information that may be exchanged between peer parking systems in a region. This terminator is a reciprocal Parking Management Subsystem. This architecture flow exchange enables parking management activities to be coordinated between different parking operators or systems in a region.

4.4 Commercial Vehicle Operations

Commercial Vehicle Operations (CVO) in the ITS National Architecture are based on the following key concepts:

1. *Vehicle-to-roadside communications via a transponder (beacon DSRC tag) on the vehicle that is read from and written to by a roadside reader*. At a minimum, the DSRC tag supplies screening data that includes identifiers for carrier, driver, and vehicle. The DSRC tag supplies clearance event data (Commercial Vehicle Check station ID, date and time when passed, measurements obtained, bypass status) from the last Commercial Vehicle Check station which wrote to the DSRC tag. The DSRC tag also supports international border crossing and safety inspection.

2. *Commercial Vehicle Drivers and Fleet managers purchase credentials and pay taxes electronically to state (or regional) Commercial Vehicle Administration subsystems (CVAS).* Records of these electronic transactions are made available to roadway Commercial Vehicle Check subsystems (CVCS) in the region covered by the CVAS. Similarly, results of clearance, safety and border checks made at the CVCSs are reported electronically to the CVASs.

3. *The CVASs collectively form a virtual network of a "national CVAS."* Carrier, vehicle and driver snapshots containing credential and safety identification and status information are assembled and exchanged between CVASs so that a CVAS can have as little or as much data as is available about carriers, vehicles and drivers at its respective CVCSs for screening. Furthermore, on request, it can make detailed carrier, vehicle and driver profiles available to CVCSs (or other inquirers) rapidly.

Participation in ITS for CVO will be voluntary. The scenarios in the sections that follow describe how the system will operate for those stakeholders who are fully equipped. Unless mandated by State or Federal Agencies, there will continue to be a mix of the current operations where a vehicle automatically pulls into the roadside station and automated operations as described below.

4.4.1 Commercial Vehicle Electronic Clearance

Electronic Clearance involves two distinct but interacting operations: registration (electronic credential and tax filing) of drivers, vehicles and carriers, and vehicle clearance (pass/need to stop determination). Registration is addressed in Section 4.4.4.

Electronic Clearance involves screening the approaching vehicle and determining whether or not to perform a more detailed verification or an inspection. Commercial vehicles are equipped with an electronic DSRC tag that provides identifying and status information to the roadside commercial vehicle check station. The roadside station reads data from the DSRC tag and determines whether to pull the vehicle in or let it pass. Domestic Electronic Clearance is extended by adding customs and immigration activities to accomplish International Electronic Clearance.

In most cases, screening will result in clearing a safe and legal carrier/driver/vehicle to proceed. In some cases, screening may lead to conducting a safety inspection. In other cases, further verification and perhaps inspection may lead to issuance of a citation.

An evolution of mechanisms has been chosen by the architecture team for commercial DSRC and identification involving a tag in the vehicle that is readable and writable by the roadside equipment at a CVCS. There is an evolution of how much data is accessible through the tag to allow increasingly complete checks and screens of vehicles at mainline speeds.

a. Limited changeable data on each tag: "Driver ID, Vehicle ID, Carrier ID, and cargo-type flag" plus "last clearance event information." This mechanism uses what are referred to as "Type II tags."

   This is the simplest/least costly approach per vehicle. The driver re-programs the tag any time there is a change in carrier or driver. In addition, roadside CVCSs can write a "clearance event record" on the tag for reading by subsequent CVCSs.

   IDs may be encrypted by the issuing agency, rendering it difficult to forge new IDs, but relatively easy to copy. If agencies perform occasional audits to identify duplicate drivers and vehicles (e.g., a driver or vehicle ID in two places at the same time), and deploy vigorous enforcement, counterfeiting of IDs may be relatively infrequent. Cargo type information should not be encrypted, and the integrity of this data must be enforced with occasional cargo inspections.

b. Tag is used as an input/output device to vehicle data: Vehicle Safety information accessible to the roadway through the tag. This mechanism uses what are referred to as "Type III tags."

   The tag allows the CVCS to access safety data about the vehicle, driver and/or cargo. This process is described in section 4.4.2.

The architecture and messages for the electronic clearance of vehicles is shown in Figure 20.

Note that this figure includes both the Commercial Vehicle Terminator (that represents the non-ITS part of a commercial vehicle) and the Commercial Vehicle Subsystem (that contains the ITS part of a commercial vehicle).

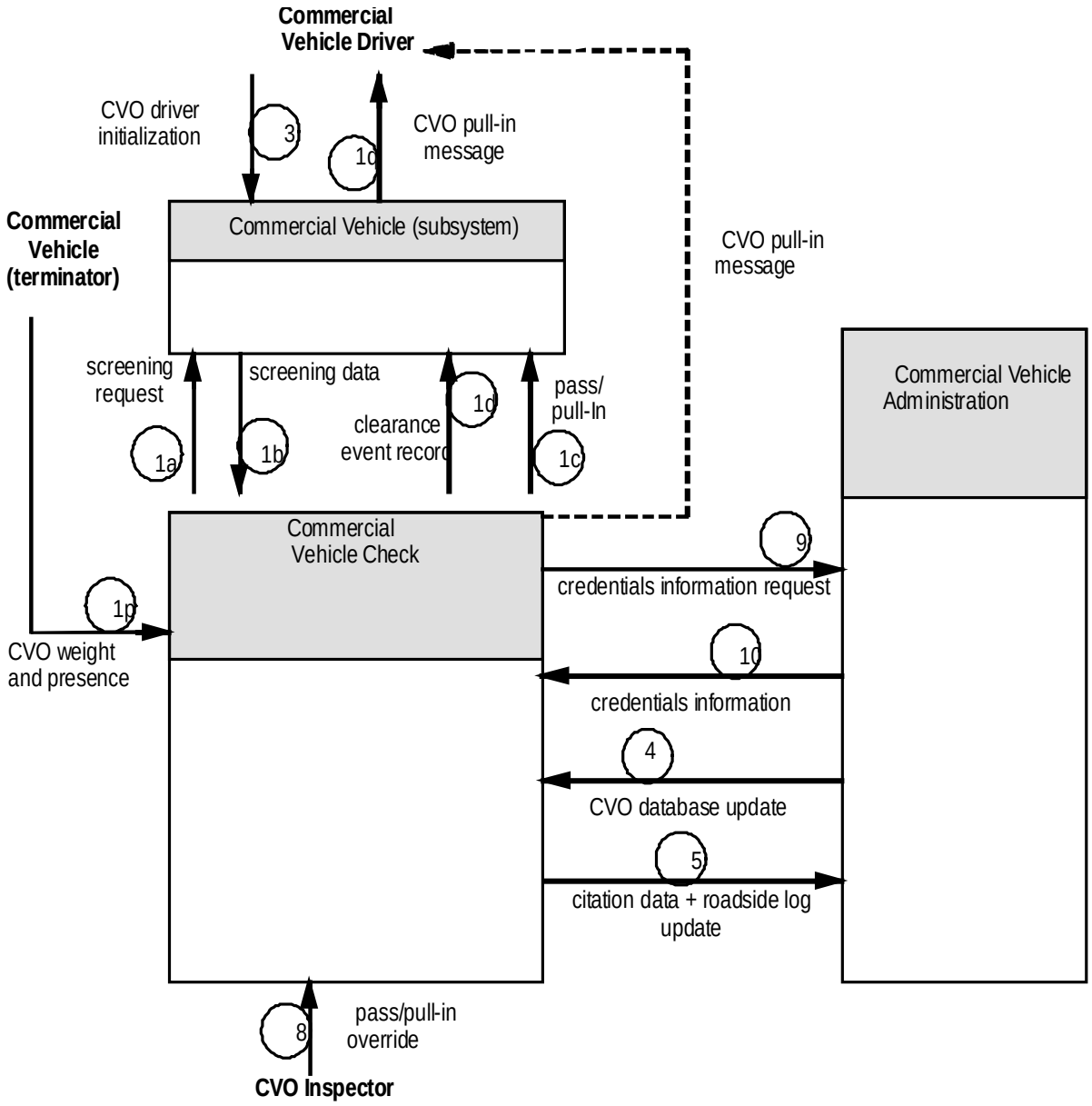

Figure 20. Physical Architecture for Commercial Vehicle Electronic Clearance

The first set of architecture flows describes the transactions between the commercial vehicle and the CVCS.

During operational shifts, the CVCS sends a screening request (flow 1a) to the Commercial Vehicle Subsystem (CVS). The DRSC tag in the vehicle in response sends the flow screening data (1b) to the CVCS. This flow contains commercial vehicle tag IDs for vehicle, carrier and driver as well as the *Specially Regulated Load Flag* and *Last Clearance Event info*. In addition, the CVCS may use sensors to measure physical characteristics of the commercial vehicle such as presence, weight, weight distribution and number of axles (architecture flow 1p). The CVCS compares the screening data to its database of cleared vehicles, makes a pass or pull in decision and communicates this to the Commercial Vehicle (flow 1c).

The pass/pull-in message is communicated to the driver one of two ways:

a. If the tag on the vehicle has in-vehicle signage capability, then the driver can be notified via the DSRC beacon. This data flow (1c) would be presented to the driver as either an aural or visual message (flow 1q). This method may be preferred by infrastructure designers concerned about accurately communicating the decision to the driver over a range of vehicle speeds, vehicle headways, and CVCS processing latencies.

b. Using a simple Dynamic Message Sign: PASS (green) or PULL-IN (red). This method may be preferred by truckers favoring minimum-cost in-vehicle equipment.

The architecture can currently support either of these physical roadside-to-vehicle communication mechanisms.

Finally, to complete the electronic clearance process, the CVCS writes to the tag the clearance event record (flow 1d).

Performing electronic clearance at an international border follows a similar transaction sequence as described above. A depiction of the dataflows between the CVS and the CVCS is given in Figure 21. As with the interstate electronic clearance process, physical characteristics measured by the CVCS may be used. One additional feature of the international border is the need to verify cargo. This can be performed with an electronic lock. This feature involves a second tag on the vehicle attached to the cargo doors which is set when the cargo is originally loaded, or when inspected prior to getting to the border. The process is a simple lock request (flow 2a in Figure 22) followed by a lock data response (flow 2b). The pass/pull-in could be communicated to the driver in either of the ways mentioned above.

In addition to the CVCS/CVS interfaces discussed above, the following actions support the overall electronic clearance process.

Prior to arriving at the CVCS, the tag on the vehicle must be initialized. The tag ID information for vehicle, driver and carrier and also the "specially regulated load flag" (e.g. HAZMAT) must be entered on the tag prior to beginning a trip. This can occur in a variety of ways. If the driver has a data entry mechanism in the vehicle (i.e. he has a type III tag with a means of entering data), then the data can be entered as shown in dataflow 3 in Figure 20. Alternate methods of entering the data are via a tag reader at the FMS or a handheld tag reader used by the driver (these possible methods of data entry are not explicitly shown in Figure 20).

Periodically (e.g. prior to beginning operations for a shift or day) the CVCS on the roadside must update its local store of credentials for vehicles, carriers and drivers from the CVAS (flow 4 in Figure 20). In some deployments, this store may contain only problem (e.g. expired) credentials and out-of-service records, and in other deployments could be *all* credentials. If the CVCS is a border crossing, the CVAS may also download additional International Border Crossing Data such as specific driver, cargo and vehicle clearances associated with a specific vehicle (see Figure 21).

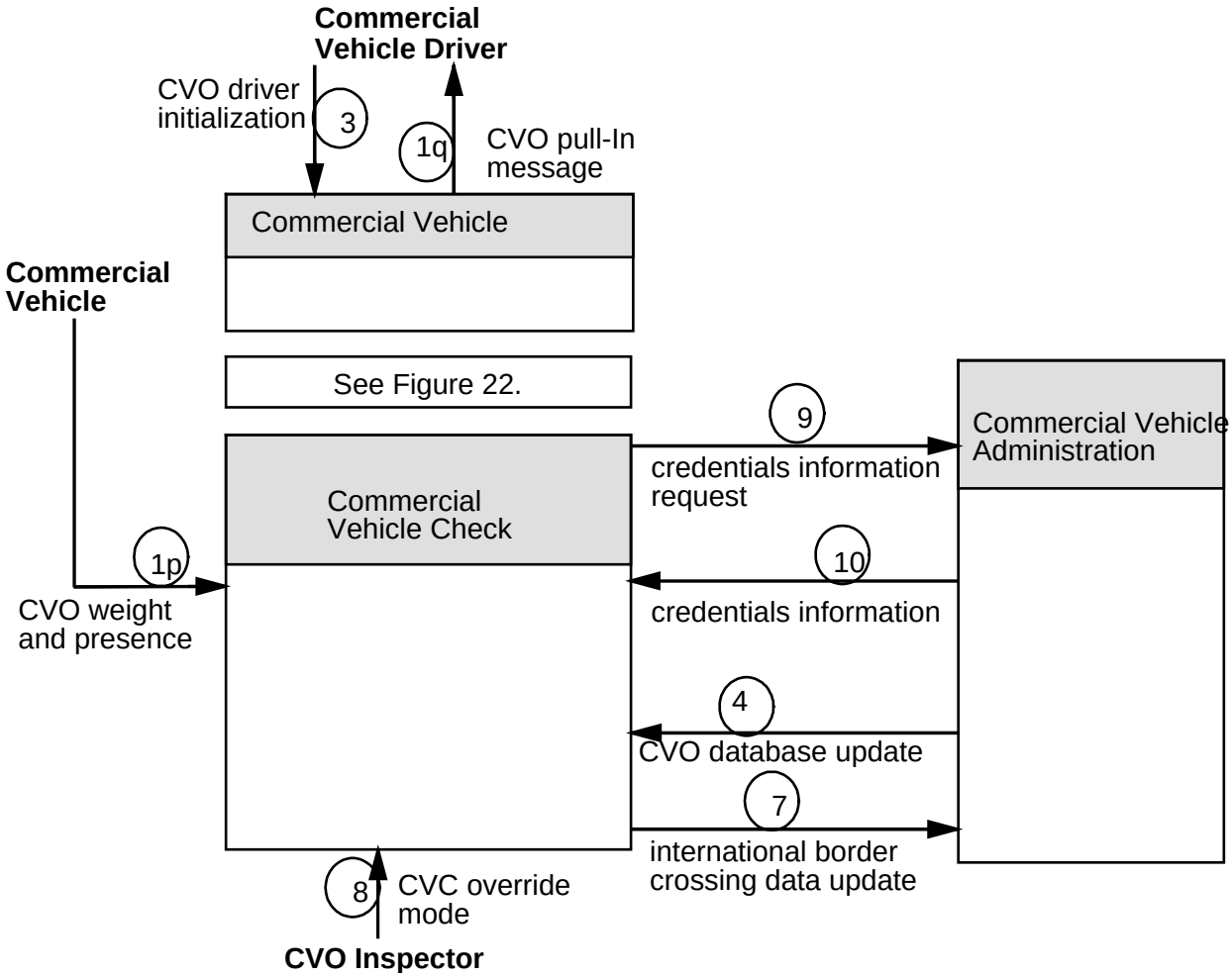

Figure 21. Physical Architecture for Electronic International Border Clearance

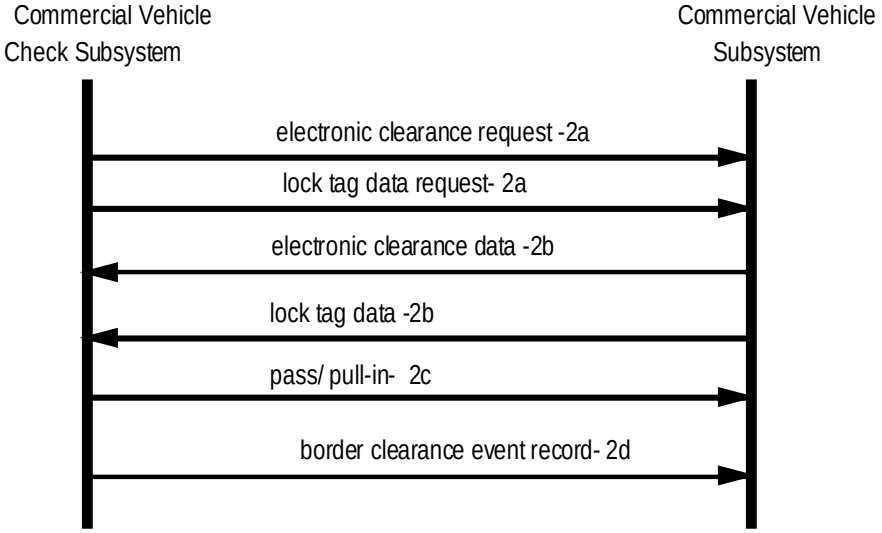

Figure 22. Physical Architecture for Electronic International Border Clearance: CVS/CVCS Interface

In some cases CVCSs will be "mobile," that is, the location of their operation will change -- but during operation they will be fixed. CVCSs operating in this way may use wireless WAN communications to exchange messages with their CVAS.

Also, on a periodic (or as needed) basis, the CVCS will upload the results of its activities to the CVAS. There are three types of updates which can be provided: clearance event data (flow 5), inspection report data (discussed in Section 4.4.2.), or border crossing data (flow 7 in Figure 21).

The CVO Inspector can adjust thresholds for Pass/Pull-In determination as well as set manual or random override flags for Pass or Pull-In determination (flow 8 in Figure 20 or Figure 21).

The following steps allow a CVCS to access the full and most recent CVAS stored profile about a particular vehicle, driver or carrier that has been pulled-in. In some deployments where latency of this request-response transaction is sufficiently fast, these steps can be executed prior to making a pass/pull-in determination. For example, some roadside CVCS deployments may use two beacons separated by sufficient space on the roadway so that vehicles passing the first beacon where the tag information is read will take sufficient time before passing the second beacon where the pass/pull-in message is sent, even at mainline speeds, for all data requests and processing to take place. For example, if the beacons are separated by one-mile, then vehicles traveling at 60-miles per hour will allow one-minute for the data access and pass/pull-in determination processing. Different spacings can be determined to allow for different maximum vehicle speeds and different worst case data access and processing latencies.

1. The CVCS uses the IDs read from the vehicle tag to access the respective Carrier, Vehicle and/or Driver snapshots. If the snapshots are not available locally, or if additional information (i.e., carrier, vehicle, driver profile, or a copy of a specific credential) is needed, a request may be issued to the CVAS for the appropriate snapshots, profiles, or credentials (flow 9). International border crossing information may also be requested from the CVAS (flow 9 in Figure 21). If the carrier, vehicle or driver snapshots, profiles, or credentials are not available from the CVAS that the CVCS is connected to, then the CVAS will attempt to retrieve the information from the appropriate "home" CVAS. If the enforcement personnel are going to issue a citation, they will request current information (snapshot, profile, or credential) from the authoritative source.
2. The requested snapshots or profiles are sent from the CVAS to the CVCS (flow 10). This can be for electronic clearance or for international border clearance. Any citations generated (violation notifications) can be sent to the CVAS via the dataflows clearance event data (flow 5 in Figure 20) or border crossing data (flow 7 in Figure 21).

4.4.2 Automated Roadside Safety Inspection

The architecture for automated roadside safety inspection could be configured to support several different evolutionary variations that may arise in meeting this user service requirements. The options are based on assumptions about the amount of on-board equipment and data storage that commercial vehicles will support as well as the amount of processing performed at the roadside:

a. In the lowest on-board equipment scenario, the vehicle supports only the capability of Identification Numbers + Flags + Last Clearance Event Info, as described in Section 4.4.1. However, this information is not utilized in realtime to make the pass/pullin decision. In this scenario, a Safety Inspector may measure and collect safety data after the vehicle is stopped at the CVCS.

b. In the next scenario, the vehicle contains on-board safety monitoring capability and this information is transferred to a Safety Inspector via a handheld, or fixed beacon after the vehicle

has pulled off the highway. Again the pass/pull-in decision is not based upon any real–time electronic data from the vehicle.

c. In a third scenario, the CVCS reads the electronic identification data, compares the information to a database, and uses safety status from the database as part of a real time pass/pull-in decision process. Once pulled in, the safety inspection can proceed as in scenario a or b.

d. In the most advanced scenario, the vehicle supports on-board safety monitoring equipment, and the communications capability to convey this information to the roadside in real-time. In this scenario, some or all of the data that might be collected by a Safety Inspector can be collected and recorded via the tag at mainline speeds mitigating the need to stop some vehicles by applying more sophisticated pass/pull-in processing based on more safety data. The choice to implement safety inspections at mainline speed is a local one that must consider both the technical feasibility as well as the related liability issues for the public agency.

Broad deployment of standardized inspection selection algorithms for CVCS deployments is important to achieving safer and more efficient commercial vehicle operations.

If a roadside CVCS is equipped for both clearance and inspection activities, then clearance and safety inspection screening (making the pass/pull-in decision) should be integrated so that the driver gets one message about whether they must pull over.

The Automated Roadside Safety Inspection architecture is shown in Figure 23.

The first set of data flows describe the transactions between the commercial vehicle and the CVCS when the vehicle is pulled in for an inspection. The Safety inspector, either with a handheld device, or from a fixed tag reader issues an on-board safety request (flow 1a). The vehicle responds with on-board safety data (flow 1b). The inspection may have varying levels of automated or manual procedures. Summary results of the inspection are transmitted to the vehicle tag (for use by future inspection efforts) in the flow "inspection record" (flow 1c). The architecture allows the possibility that some of the data sent in flow 1c could be entered by the CVO Inspector (flow 1p). As part of the inspection physical characteristics of the vehicle may also be collected using roadside sensors (flow 2).

In the most advanced scenario, where a real time decision is made to pull-in the vehicle for a safety inspection, the transaction begins in the same manner with flows 1a and 1b. Utilizing physical measurements of the vehicle (flow 2) and allowing for an override function by the CVO Inspector, (flow 3a) the pull-in decision is made and transmitted to vehicle or driver as discussed in section 3.4.1.(flow 3b in Figure 23).

The following data flows describe the transactions between the CVAS and the CVCS.

The CVAS prepares Carrier, Vehicle and Driver Safety Snapshot data and sends this periodically (e.g., each shift or once daily) to the CVCSs (flow 4- CVO database update). Real-time updates of the CVO database update can also be sent from the CVAS to the CVCS to announce changes such as an out-of-service condition obtained in real time from another CVCS.

The following steps allow a CVCS to access the full and most recent CVAS stored profile about a particular vehicle, driver, or carrier that has been pulled-in. In some deployments where latency of this request-response transaction is sufficiently fast, these steps can be executed prior to making a pass/pull-in determination. For example, some roadside CVCS deployments may use two beacons separated by sufficient space on the roadway so that vehicles passing the first beacon where the tag information is read will take sufficient time before passing the second beacon where the pass/pull-in message is sent, even at mainline speeds, for all data requests and processing to take place. For example, if the beacons are separated by one-mile, then vehicles traveling at 60-miles per hour will allow one-minute for the data

access and pass/pull-in determination processing. Different spacings can be determined to allow for different maximum vehicle speeds and different worst case data access and processing latencies.

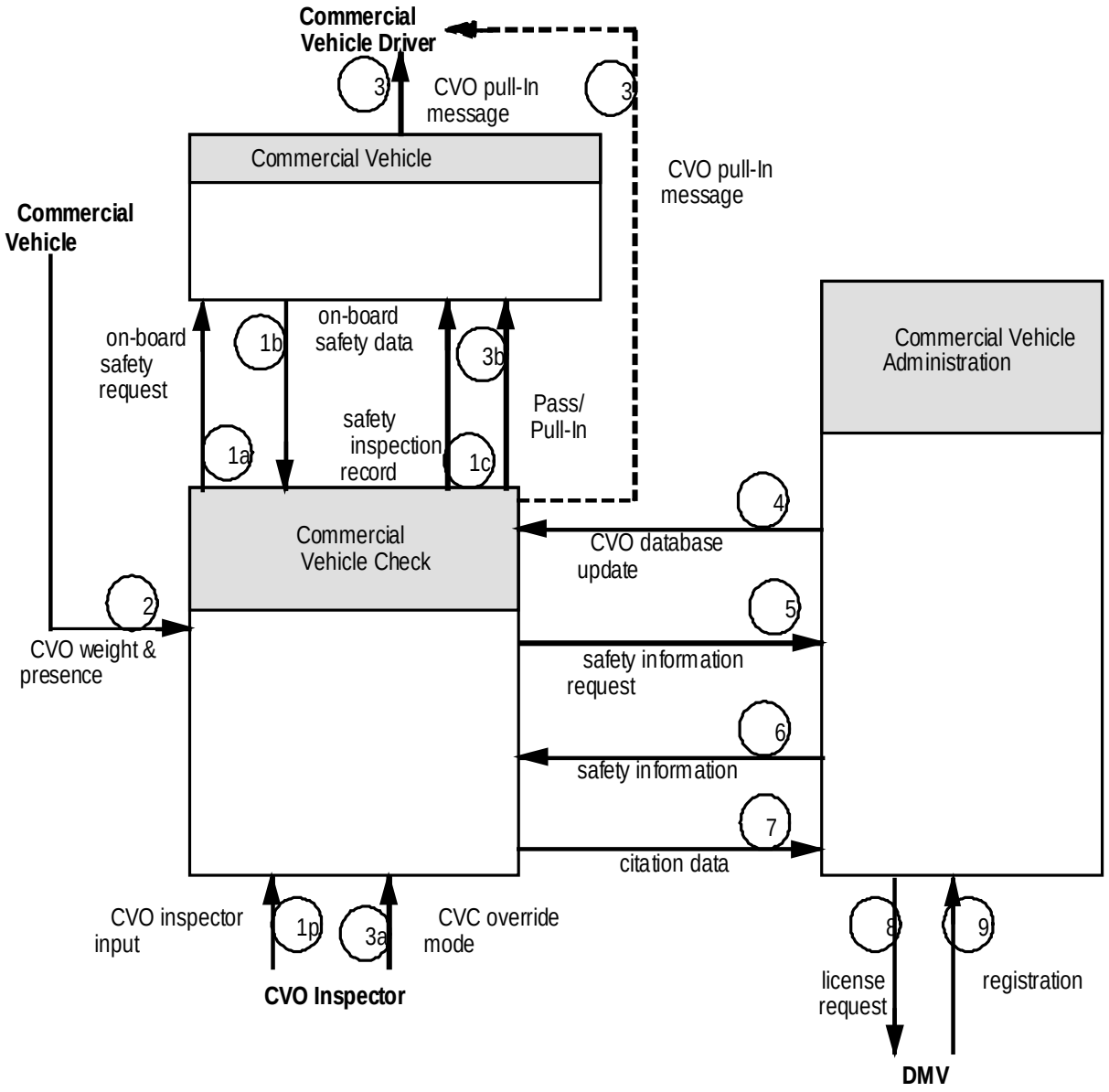

Figure 23. Physical Architecture for Automated Roadside Safety Inspection

1. After Pull-In, or if using two beacons after having read the Identification Numbers at the first beacon, a request to the CVAS for prior inspection reports or the full data profile of the carrier, vehicle, and/or driver can be issued by the CVCS (flow 5 in Figure 23).

2. The CVCS receives all requested prior inspection reports or profiles from the CVAS (flow 6).

Results of the inspection activity are sent from the CVCS to the CVAS on a periodic basis. (Flow 7 in Figure 23). If the safety assessment involves significant changes from the current snapshot data or Profile data, (e.g. a change in out-of-service status) then an operational update can be sent in real time to the CVAS. Otherwise, safety inspection results are periodically (e.g. once per shift or daily) sent from the CVCS to the CVAS.

In the event of safety violations or accidents, the CVAS will process a violation and in doing so may request registration information from the DMV, flows 8 and 9 in Figure 23.

4.4.3 On-Board Safety Monitoring

The on-board safety monitoring architecture is shown in Figure 24. The sequence of messages is as follows:

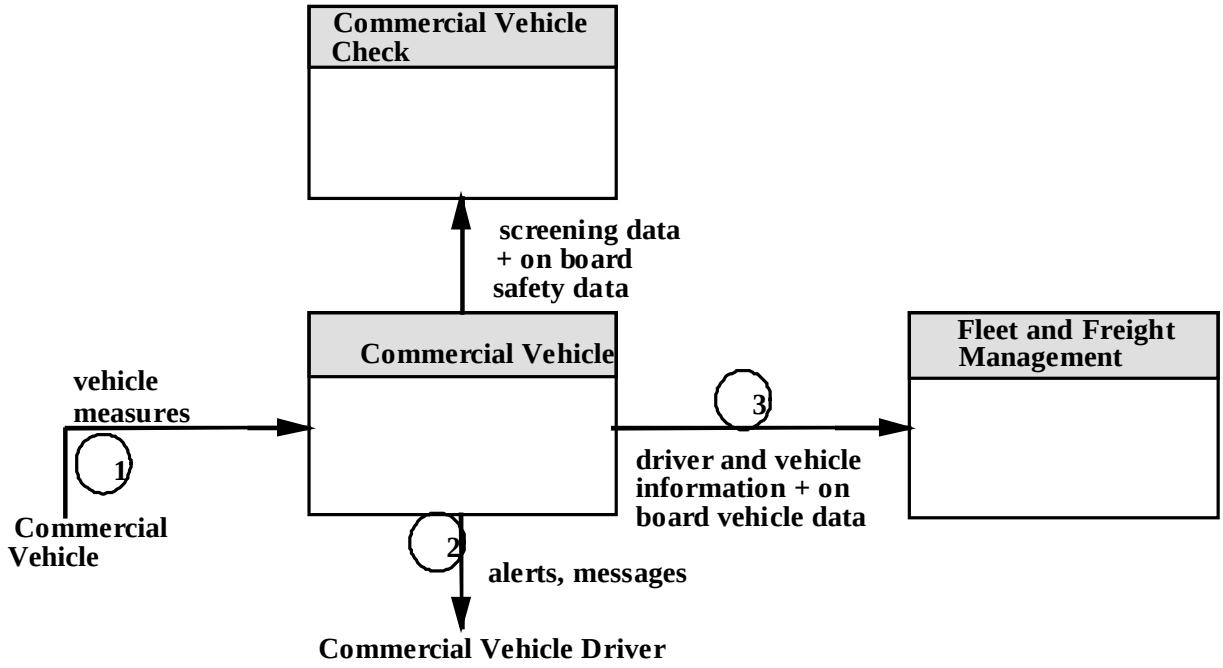

Figure 24. Physical Architecture for On-Board Safety Monitoring

1. The Commercial Vehicle collects on-board sensor data of vehicle conditions for the vehicle, cargo/trailer, and driver safety as well as the condition of the braking system. Data is primarily collected by exception; that is, data is stored when an abnormal condition is sensed.
2. When a safety exception is processed, the driver is alerted immediately by a safety alarm message.
3. The exceptional safety data may also be forwarded to the FMS. This communication could be accomplished either in real-time by using a WAN communication channel, or using a beacon-to-WAN communication capability.

The on-board safety monitoring capability communicates stored exception data to the CVCS on the roadside. This message is addressed in Section 4.4.2 in the Automated Roadside Safety Inspection user service.

4.4.4 Commercial Vehicle Administrative Process

The Commercial Vehicle Administrative user services have been decomposed into Credentialing of Vehicles and Carriers (purchase of credentials and tax payments) and distribution of CVO profiles.

4.4.4.1 Electronic Purchase of Credentials and Enrollment for Electronic Clearance

Enrollment of commercial vehicles, carriers, and drivers is implied by the electronic credentialing process.

Credentialing can be performed by:

1. The Commercial Fleet Manager from the Fleet and Freight Management Center or
2. The Commercial Vehicle Driver from
    a. The CVS e.g., an in-vehicle unit).
    b. The RTS e.g., a kiosk at either truck stops or CVCS Roadside facilities.
    c. The PIAS e.g., home/office/laptop computer.

The latter is an example of aggregating the FMS with the CVS, RTS, or PIAS subsystems.

The physical architecture for electronic purchase of credentials is shown in Figure 25. The sequence of messages are as follows:

1. The CVAS requests the tax/credential fee schedule from the appropriate Government Administrator(s). These may be local, regional, state, or national agencies, and as such, multiples of these requests may be issued, one to each relevant agency.

    The architecture allows for a future national agency information clearinghouse for tax/credential fee schedules and also for CVAS names and network addresses. This information allows a CVAS to act as an agent for credential purchases and tax payments for jurisdictions covered by other CVASs that comply with the ITS National Architecture. **Note that this creates an implied responsibility for some entity (probably a government organization) to take responsibility for this "name server" maintenance function.** Alternatively, CVAS jurisdictions and network addresses could be maintained at each CVAS on an Ad Hoc basis, or could be maintained by an association of state and regional Commercial Vehicle agencies. This request message may serve also to add the requesting CVAS to the "name server" database of CVASs (if this is the first time they are requesting this information).

    The information returned includes appropriate financial account information for deposit of collected fees and taxes. This request will be issued before operations of the CVAS and at regular (e.g., monthly or quarterly) intervals thereafter. The Government Administrators may also choose to send this information to CVASs at regular intervals.

2. The government agencies send responses to the previous request.

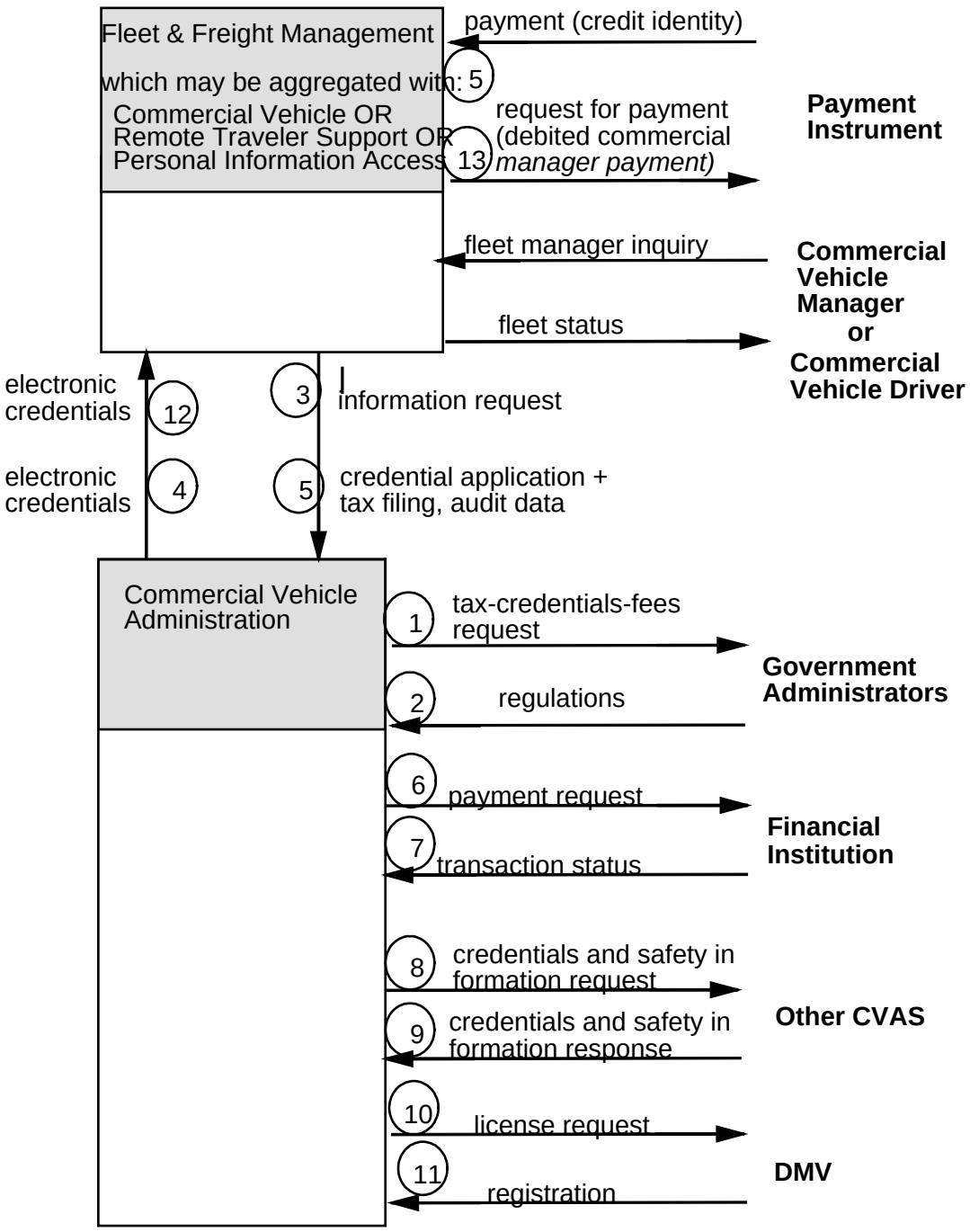

Figure 25. Physical Architecture for Electronic Purchase of Credentials

As indicated in Figure 25, a Commercial Vehicle Driver (at a CVS, RTS or PIAS) or Commercial Fleet and Freight Manager at an FMS interacts with his respective subsystem using a "human interface" to issue and receive messages as follows. The four aforementioned subsystems will be referred to here as the "User" subsystem. Note that the architecture has assigned the processes to support the User subsystem to the FMS. In the case where the User is a Commercial Vehicle Driver, then a FMS (equipped only with the Fleet Credentials and Taxes Management and Reporting Equipment Package) is aggregated with the respective user subsystem.

3. The User subsystem issues an Information Request message to the CVAS.

   The message contains all the relevant information necessary to participate in the CVO ITS program including regions to be traveled and borders to be crossed, class of vehicle, cargo class or manifest as necessary, weight class, etc. The purpose of this message is to get information regarding the required taxes, duties and credentials, and cost to purchase for a particular plan of operations. The credentials can include: annual or temporary credentials and specific or multiple permits.

   The reporting message can include quarterly reports, vehicle logs, or fuel purchase data.

4. The CVAS responds with the requested Electronic Credential and Tax Filing information in a message to the User subsystem.

5. Similar to message 3. above, the Credential Application message also contains

   a. Payment information (credit identity), possibly originating from a "Payment Instrument" (e.g. credit card, debit card, or cash card") to initiate a purchase of credentials or tax payment and

   b. Driver, carrier and/or vehicle identification numbers as necessary for the automated processing of credential purchase and tax payment applications.

6. The CVAS processes the request to determine the total payment required for credentials, taxes, and duties in its jurisdiction. It issues the appropriate payment request to the financial institution that the CVAS operator has chosen.

7. The Financial Institution issues a Payment Confirmation message indicating whether the financial transaction has completed or been rejected.

8. If the payments were accepted, the CVAS processes the application to determine if any taxes or credentials need to have been purchased or paid from any other CVAS(s). If so, then the appropriate Electronic Credential information request message is issued.

9. The CVAS waits and receives information response from the other CVAS(s).

10. The CVAS requests registration information from the DMV based on licensing information as necessary to process credentials.

11. The CVAS receives the requested DMV registration information.

12. The issued tax and credential receipts (including HAZMAT clearance) are combined into a message and sent to the User subsystem.

13. The User subsystem issues a payment confirmation (debited commercial manager payment part of the "request for payment" physical flow) to the Payment Instrument. This message allows a cash card to decrement its internally stored value.

4.4.4.2 Distribution of CVO Snapshots and Profiles

Figure 26 shows the architecture for the electronic distribution of Driver, Carrier or Vehicle Snapshots to Other CVASs and Profiles (current credentials, tax reports, violation history/status) by the network of CVASs to CVO Information Requestors (e.g. roadway enforcement officers, insurance companies, drivers, carriers, fleet managers).

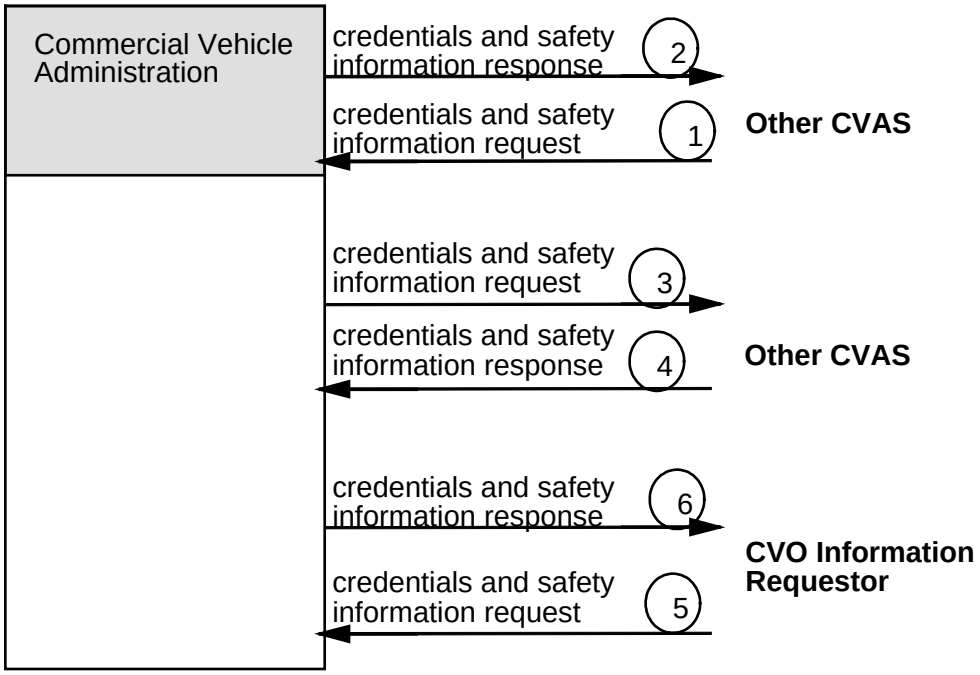

Figure 26. Physical Architecture for Commercial Vehicle Electronic Data Sharing

The snapshots and profiles are assembled as follows:

1. An other CVAS requests a snapshot or profile by sending the request to the appropriate CVAS.
2. The CVAS responds by sending the requested snapshot or profile.
3. Similarly, this CVAS can request snapshot or profile data from other CVAS,
4. and receive the information to assemble complete snapshot or profile data for its own CVCSs. The distribution of snapshots from the CVAS to the CVCSs is discussed in Sections 4.4.1. and 4.4.2.

The profiles and snapshots are distributed to CVO Information Requestors as follows:

5. A CVO Information Requestor requests a profile or snapshot by sending the request to the appropriate CVAS. If the information is stored at another CVAS, the requested information can be retrieved using flows 3 and 4 above.
6. The CVAS sends the requested snapshot or profile to the CVO Information Requestor.

4.4.5 Hazardous Material Incident Response

The EM coordinates HAZMAT incident response. The architecture for HAZMAT response is shown in Figure 27, and the sequence of messages is as follows:

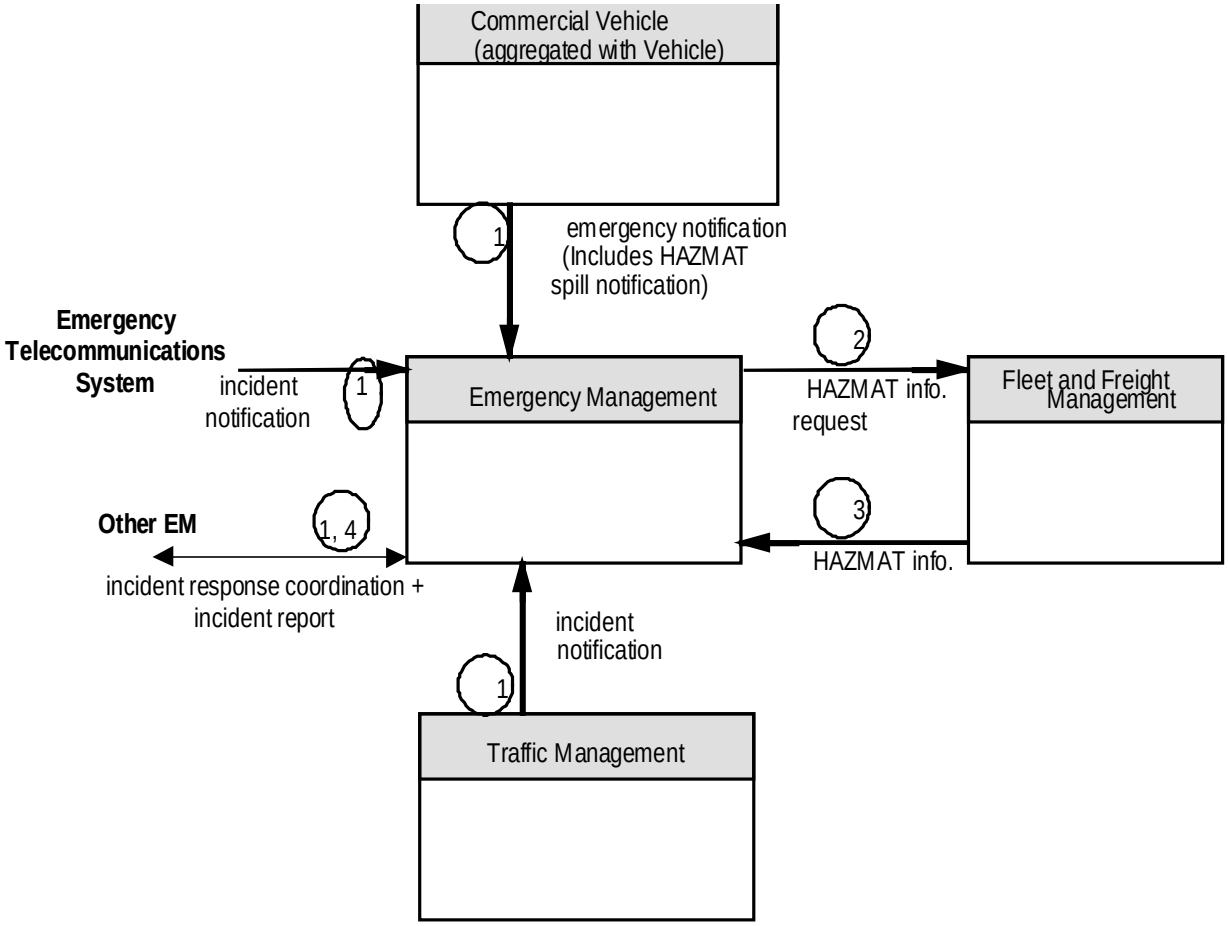

Figure 27. Physical Architecture for Hazardous Material (HAZMAT) Incident Response

1. Notification to the EM of a HAZMAT incident may occur in several ways:

   a. The EM can receive an Emergency Request Details message notification of the incident from the Vehicle (see Section 4.5.1.2.). This message can include the HAZMAT manifest, including the standard MSDS numbers for hazardous materiels and the quantity of each materiel.

   b. An incident notification from the Emergency Telecommunication System interface or an incident report from an Other EM subsystem.

   c. An Incident Alert message from the TMS can also include incident requests for EM service involving HAZMAT. This message can be sent prior to a HAZMAT containing vehicle transiting an EM's jurisdiction for the purpose of emergency management incident preplaning.

   d. A dispatched emergency vehicle on the scene may carry portable beacon equipment able to read a tag on the disabled vehicle. This equipment would be functionally identical to the beacon readers located at CVCSs located on the roadway. By reading the Vehicle and Carrier IDs, the Emergency Vehicle could communicate with the EM (e.g. by voice) to identify the disabled vehicle, with the expectation of receiving instructions from the EM as part of the coordinated response (step 4. below). This source is not shown in Figure 27 because it represents the aggregation of the Emergency Vehicle subsystem with the CVC subsystem for this special application.

2. The EM may ask the FMS for more information about the HAZMAT load (and to notify the FMS about the incident).

3. The FMS is responsible for supplying HAZMAT load information to the EM on request.

4. The EM determines a coordinated response and issues appropriate request messages to other EMs.

HAZMAT response is assisted during the route selection process (see Section 4.1.3.) since commercial vehicles carrying HAZMAT cargoes will note the cargoes in the Trip Request. Once a route is selected, the route is forwarded to the TMS for reference in preparing an Incident Alert message to the EM should an incident occur involving that vehicle.

4.4.6 Commercial Fleet Management

Commercial Fleet Management subsystems will be able to do commercial vehicle route planning through ISPs as discussed in section 4.1.3., Route Guidance. It is likely that some large trucking companies may deploy their own "captive" ISP for routing/logistics purposes, where smaller trucking operations may buy the logistics services of a larger operator or of an ISP.

Vehicles that install the wireless data communication subsystem will be able to communicate data messages (e-mail) to/from their Commercial Vehicle Managers and the Commercial Vehicle Managers will similarly be able to be in electronic contact with multimodal transportation providers. Similarly, voice communications requirements will continue to be served by the existing and emerging mobile phone services which are internetworked with the PSTN standard.

Since the ITS will use the NII for its wide area data communications, a side effect is that users of the NII (including ITS users) will have pervasive electronic communication capability with each other.

The architecture for Commercial Fleet Management is shown in Figure 28, and the sequence of messages is as follows:

1. The Commercial Vehicle supplies vehicle data such as status and location to the Fleet and Freight Management Center Subsystem.

2. The Fleet and Freight Management Center uses shipping and schedule information as inputs to their own business processes.

3. Messages to drivers are "free-form," and may include information about problem areas, schedule changes, and cargo availability.

4.5 Emergency Management

4.5.1 Emergency Notification and Personal Security

These Mayday services augment those described in Section 4.2.4., Transit Security. Two key issues in this service required to support seamless operation over geography, are: the location determination of the mobile emergency requester and the mechanism by which the initial emergency message is routed to the appropriate EM.

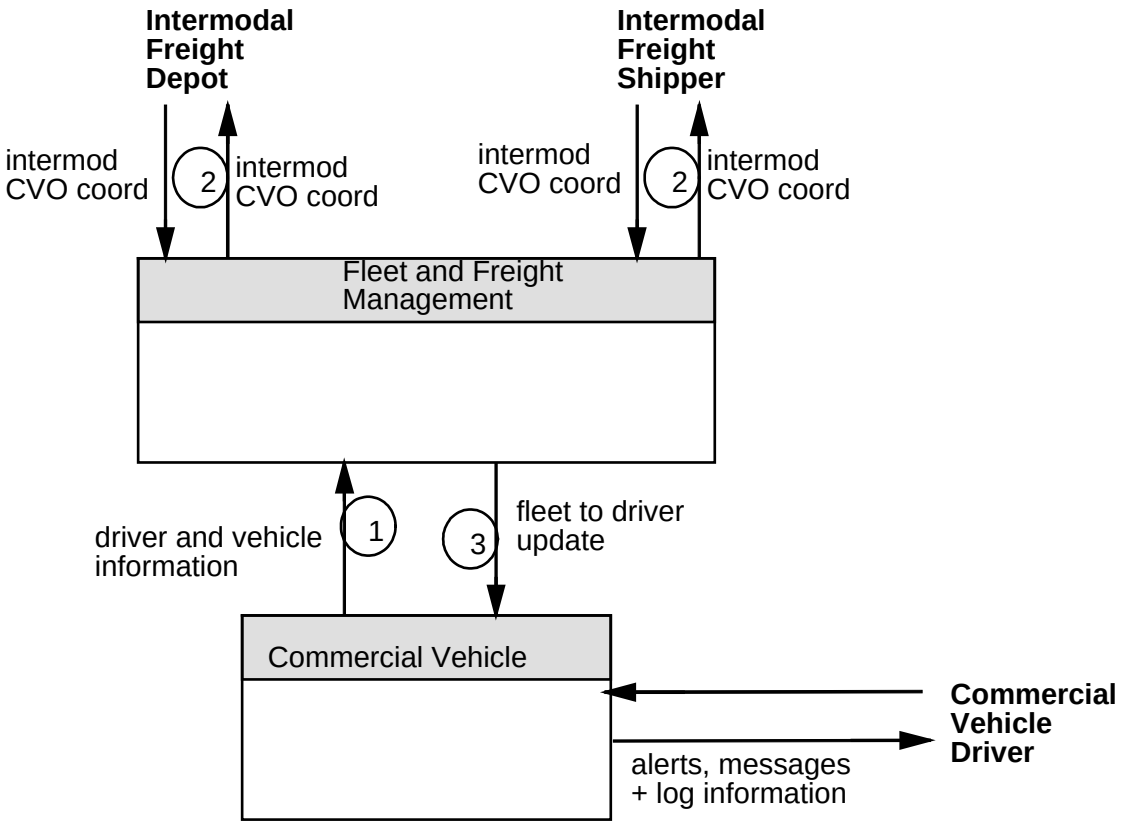

Figure 28. Physical Architecture for Commercial Fleet Management

4.5.1.1 Emergency Notification Location Determination

The user service requires that the location of the emergency caller be available to the infrastructure based emergency service provider (the EM in the National Architecture). A choice has been made to use location determination equipment incorporated into each mobile subsystem that might issue an emergency notification. However, location determination by the communication infrastructure may be reasonable at some time in the future, and the National Architecture can then be enhanced (in this case simplified) to incorporate that location determination mechanism.

The location of the Driver or Traveler calling for assistance is determined by the mobile subsystem that they are interacting with. An alternative is to put a requirement on the communications service provider to provide the caller location (as is done on Emergency Telecommunication Systems today for 9-1-1 calls from wireline connected telephones). Unfortunately, the technology to locate callers using wireless service providers is still in a developing stage, and it is difficult to predict if this technology will mature to a useable level for this service. Although expensive (but getting less expensive very quickly), the technology for locating mobile subsystems with self–contained equipment is currently available and the current prices are expected to continue to fall.

4.5.1.2 Emergency Notification Message Routing

Wireless emergency notification messages must be efficiently routed to the appropriate EM for the type of assistance required and the origin of the request. This is particularly challenging for wireless WAN communications. Two mechanisms have been proposed, each with its own advantages and disadvantages. The architecture described could be readily adapted to any of these mechanisms.

1. Wireless Communication Service Provider Routing

   The mobile subsystem always sends emergency messages to a standard IP (Internet Protocol) address. This special EM address is detected by the CSP equipment. The Communication Service Provider (CSP) equipment will read the location and type of assistance needed in the standard Emergency Request message, and based on the location and assistance request details will forward the message to the appropriate EM.

   This mechanism requires an extension to current communication standards as well as special equipment to be deployed by wireless carriers to implement that standard.

2. Emergency Management Center Subsystem (EM) Routing

   Emergency Request messages will be routed to an EM. The EM will interpret the Emergency Request Details message, and based on the location and type of emergency assistance, forward the message to the appropriate EM.

   Users of this service will be required to subscribe to an EM. Of course, the EM could be provided for this service to all ITS wireless mobile users at public expense (or at the expense of wireless service subscribers through a "tax" as is currently done to support 9-1-1 service for cellular phone subscribers in some states).

The Emergency Notification and Personal Security architecture is shown in Figure 29. The messages are as follows.

1. VSs or PIASs with Emergency Request equipment packages notify the EM that an emergency has happened either autonomously (triggered by processing of vehicle sensor data, e.g., the same sensor data that deploys an air-bag) or by the driver/traveler activation. The mechanism for driver or traveler activation can be a "panic" button, as well as other interfaces that will provide a variety of pre-programmed messages to be sent (e.g., panic, collision, personal security, roadside assistance needed). The emergency request message contains location, direction, and speed of travel as well as the emergency data from the driver/traveler and the sensors. There is also an area in the Emergency Request Message where a Vehicle can indicate HAZMAT loads. This is done using the MSDS numbers and quantity of each hazardous material on the vehicle.

2. The EM acknowledges receipt of the emergency request indicating the response plan. The EM may also request additional information from the traveler as well as administer some control functions (e.g. unlock a door or turn off an alarm system). A similar path exists from the Personal Traveler Guidance subsystem to the EM.

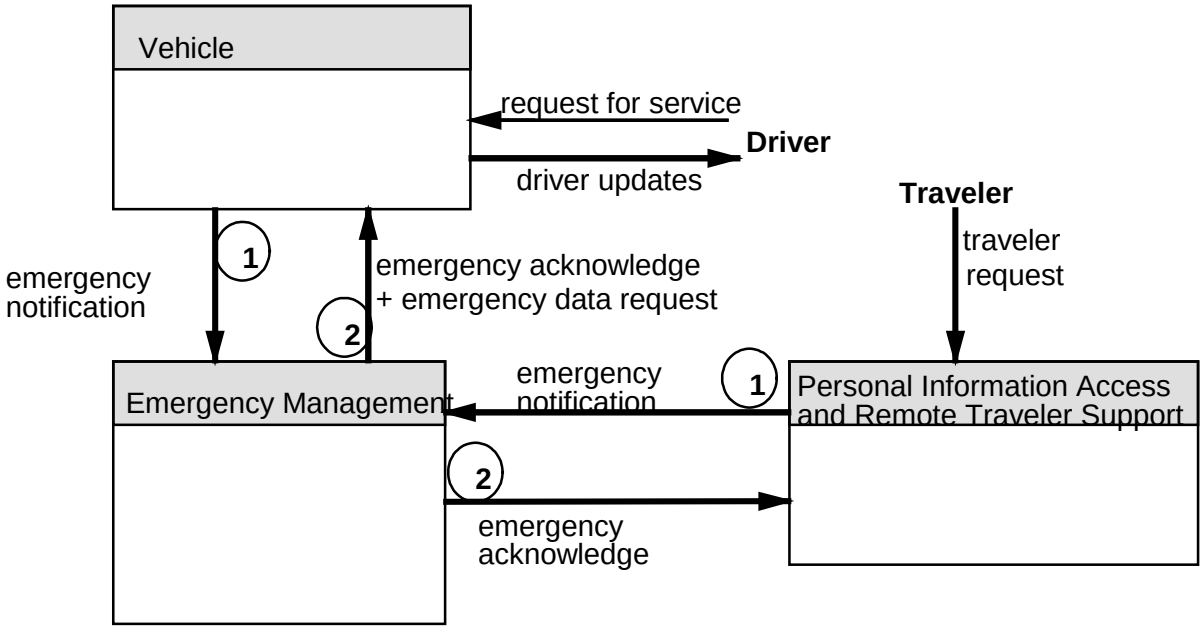

Figure 29. Physical Architecture for Emergency Notification and Personal Security

4.5.2 Emergency Vehicle Management

When responding to an emergency, the EM subsystem may dispatch emergency vehicles as shown in Figure 30. Several variations of this architecture can be supported.

In most cases today the wireless communication modality used by the Emergency Vehicle will be a private licensed radio medium that only allows communication between a fleet of emergency vehicles and the EM (e.g., an SMR or Special Mobile Radio license).

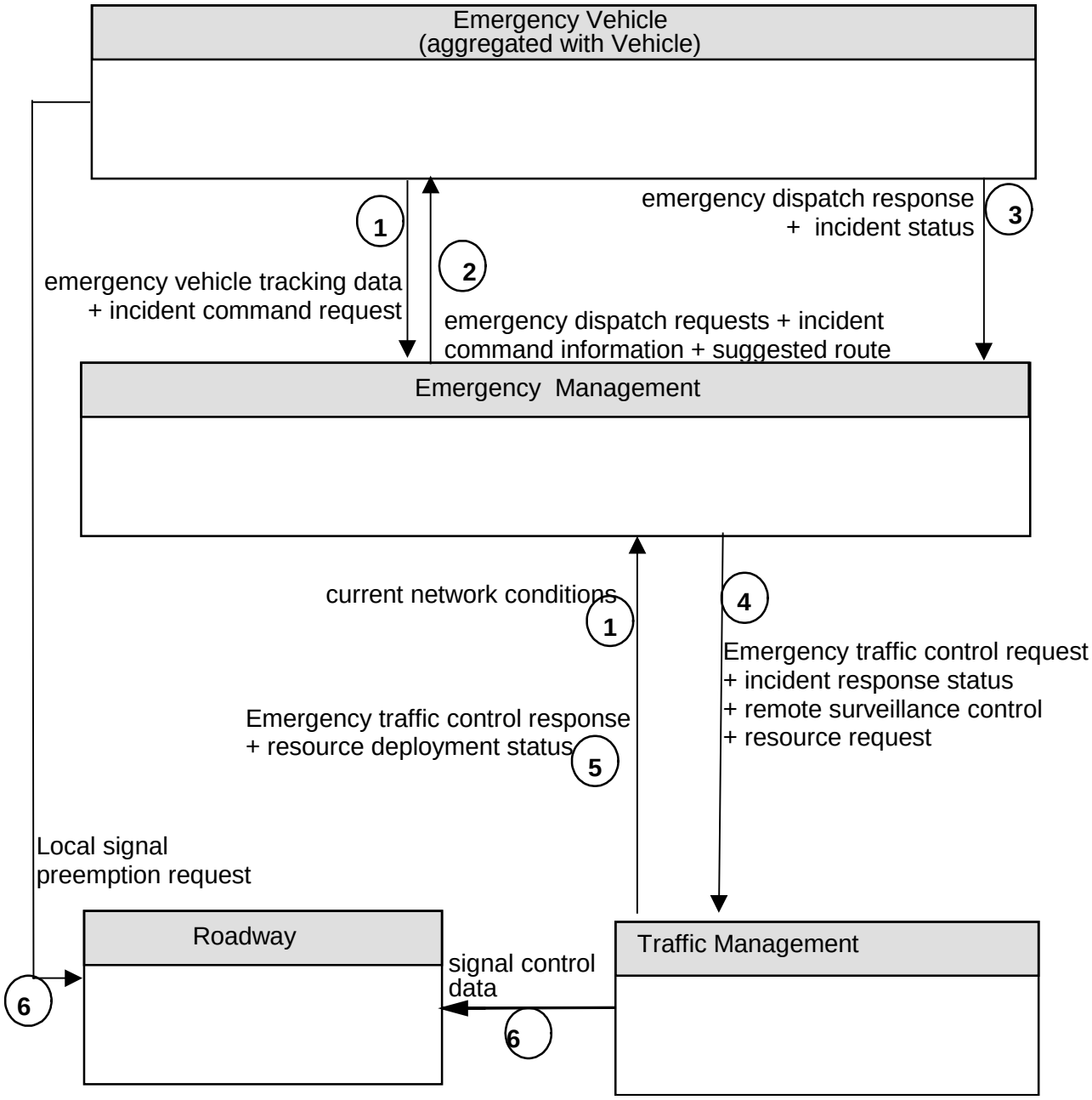

Figure 30. Physical Architecture for Emergency Vehicle Management

1. The EM monitors the location and status (e.g., available, busy, out-of-service) of a "pool" of emergency response vehicles. Emergency Vehicles continue to periodically send status and command request messages to the EM while enroute and after arriving on the scene to provide a Drivers assessment of the incident. At the same time, the EM is monitoring the state of the transportation network from the TMS.

2. The EM selects the vehicle(s) best located to respond and sends the emergency dispatch order message(s) to the vehicle. This may include a suggested route message. (Note that the ISP interacts with the TMS(s) and the Roadway to get preferential signal service for the emergency vehicles as per

the Route Guidance user service shown in Figure 6, if the TMS(s) are able to provide that service and if the EVS is aggregated with a vehicle that is getting ISP route guidance service.)

3. The selected emergency vehicle responds, indicating to the EM that it is proceeding to the incident, and when the EVS arrives at the incident, provides incident status to the EM.

4. The EM may directly request traffic controls and/or resources from the TMS in order to facilitate incident response for the EVS or to mitigate the traffic consequences of the incident (the "resource request" message). Also, the EM will inform the TMS of the incident status, and may also control TMS surveillance resources (e.g. video surveillance) to monitor and assess an incident.

5. The TMS can respond to the EM as to actual traffic control response to the prior request and deployment of resources to assist with the incident.

6. If the emergency vehicle has a vehicle mounted vehicle to Roadside broadcasting system for signal pre-emption, this message can be used to trigger signal pre-emption based on the approach of an emergency vehicle to a signal. Otherwise pre-emption, if available, is controlled by the TMS directly.

4.6 Advanced Vehicle Safety Systems

The Advance Vehicle Safety Systems user services are not fully addressed by the architecture: the functions can be implemented by in-vehicle equipment or can be augmented with interactions with equipment in the Roadway Subsystem. This is particularly true in the case of the Intersection Collision Avoidance user service, where communication with sensors and possibly some control functions in the Roadway subsystem may be necessary. Operational concepts for these user services, including Intersection Collision Avoidance, are still being explored at this writing by other DOT programs. As part of the ITS National Architecture we have tentatively defined the functionality required within the vehicle but have not specified the detailed physical architecture of the vehicle itself, believing that this is the purview of the vehicle manufacturers.

Figure 31 shows the physical architecture for Intersection Collision Avoidance, and the data flows between the vehicle and Roadway Subsystem, as well as Roadway to TMC data flows. Little detail is provided for these data flows in the Logical and Physical architecture, since at this time is is very difficult to identify with confidence the operational concept, as discussed in the previous paragraph.

1. The Roadway subsystem collects vehicle status data from vehicles (using DSRC). This might include anticipatory vehicle trajectory status. At the same time, the Roadway incorporates surveillance information from unequipped vehicles, pedestrians, and others that may be using the intersection (e.g. bicycles). The Roadway uses this information to estimate the intersection "state" (a collection of individual object state vectors) and to estimate future states for possible collisions. *The Roadway subsystem is more likely to have good surveillance of the entire intersection than any vehicle operating independently.*

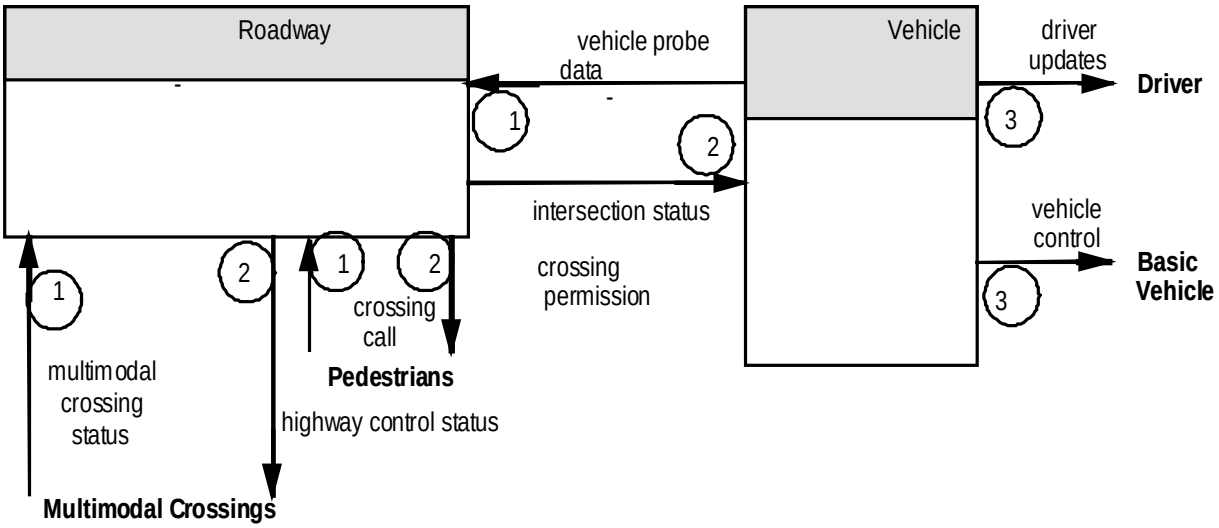

Figure 31. Physical Architecture for Intersection Collision Avoidance

2. If a collision is predicted by the Roadway subsystem, then the Roadway Subsystem can determine the appropriate countermeasure: advisory, warning or control intervention, and communicate this to the equipped vehicle(s) or to signage/controls at the Roadside environment.

3. If a vehicle receives an intersection advisory or warning, these are communicated to the driver through appropriate in-vehicle signage. Passive or active vehicle controls will be communicated to vehicle actuators (e.g. brake, throttle, transmission, steering) immediately. Any vehicle control actions could be under closed loop control using the Vehicle Data flow above.

4.7 Automated Vehicle Operation (Automated Highway Systems)

The current Automated Highway System (AHS) architecture is shown in Figure 32.

The current version of the ITS Architecture for Automated Vehicle Operation (AVO) focuses on the vehicle-to-vehicle interactions and limited vehicle-to-roadside messages. The more sophisticated connections to roadside infrastructure which are considered in many AHS implementations have not been included. We think that it is premature to pick a specific operational concept for vehicle-roadside interaction for vehicle operations such as AHS, and will add this depending upon the results of the AHS Precursor System Architecture (PSA) study efforts to define what the best infrastructure component should be for AHS implementation.

1. AHS Control Information is sent from the TMS to the Roadway subsystems. This is "slow changing" control information and as such does not participate directly in vehicle control except to set control parameters.

2. The vehicle requests to check-in to the AHS and sends vehicle data to the Roadway subsystem.

3. A check-in interrogation test is applied to the vehicle data. If the vehicle is determined to be ready for AHS, and if the Roadway determines that the AHS Roadway is ready for the vehicle, then the vehicle gets the "Pass" message, otherwise the Fail message. If the vehicle is passed, then the Roadway passes control parameter information to vehicles prior to entry onto the AHS.

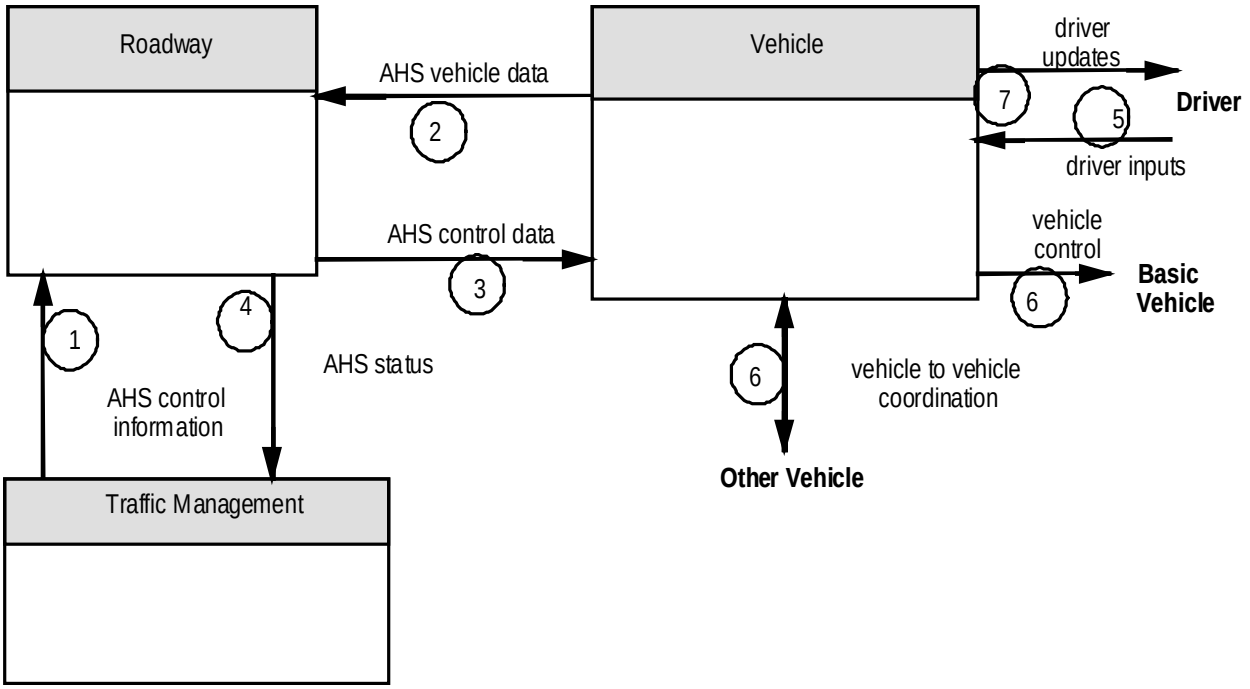

Figure 32. Physical Architecture for Automated Vehicle Operations

4. Check-in and Check-out operations and test results are sent from the Roadway to the Traffic Management Center.

5. Drivers exert vehicle operation selection prior to entry and exit of AHS lanes. This begins and ends transition sequences where the AHS equipment on the vehicle begins and ends its interactions with vehicle actuators (e.g., throttle, transmission, brakes, steering).

6. Vehicles will communicate status and movement intentions while AHS equipment is activated to adjacent vehicles. AHS entry and exit intentions are part of the status. This will allow vehicles to effect necessary separations prior to entry and exit maneuvers. Vehicles that are platooning communicate their acceleration/deceleration profiles to adjacent vehicles so that engine throttling can "lead" and variations in headway can be minimized without requiring high engine output force to vehicle mass ratios.

7. The vehicle constantly processes sensor inputs about the roadway environment and compares these measurements with the received status messages from adjacent vehicles. If a vehicle is sensed that is not properly communicating status, then an AHS Alarm message is sent to the driver indicating a possible "fault" with the equipment of one of the vehicles or with a driver having taken his vehicle into the AHS lane without the necessary equipment.

4.8 Highway Rail Intersection Operation

In this section we define the operational concepts needed to accomplish the safety and coordination aspects of the Highway Rail Intersection (HRI) user service. The following subsections discuss the key operational interactions of the Roadway and Traffic Management Subsystems with each other and with the Rail Operations terminator and the Wayside Equipment terminator. Reference is included to interfaces with Traffic (sensors), as will as signage terminators for Pedestrians and Drivers (including in-vehicle signage).

4.8.1 HRI Safety at Standard and High Speed Railroad Grade Crossings

The HRI User Service is made up of Standard Speed Rail and High Speed Rail Subservices. Both subservices have the goal of improving safety at the intersection through an integration of Highway and Rail operations. The interfaces shown in the physical architecture of Figure 33 between the Rail Operations terminator (RO), the Traffic Management Subsystem (TMS), the Roadway Subsystem (RS) and the Wayside Equipment terminator (WE) support the user service requirements for safety at High Speed Rail Subservice HRIs. The simplified diagram of Figure 34, with several flows removed, represents the Physical Architecture for the less stringent requirements of the Standard Speed Rail Subservice HRIs, which do not require safety processes associated with detection, classification, and response to intersection blockage. Note that some of the flows needed to support these user service requirements are discussed in detail elsewhere in this document. In particular, see Sections 4.1.6. and 4.1.7. for discussions of incident notification, traffic data video surveillance (including control of cameras) and traffic control for Pedestrians and Drivers.

Note in Figure 33 the redundancy of the "intersection blockage notification" message and in both Figure 33 and Figure 34 the redundancy of the "hri status" message. These messages originate at the HRI supported by the RS and are received at both the TMS and the WE. Although the full deployment of the user service requires participation of a TMS, this redundancy enables beneficial deployment where there is no TMS, such as at rural HRIs. Where there is no TMS, the RS can still notify the WE of conditions that might indicate safety problems, and the WE can process or forward the messages to the Train (to indicate an emergency stop), RO or other appropriate emergency agency.

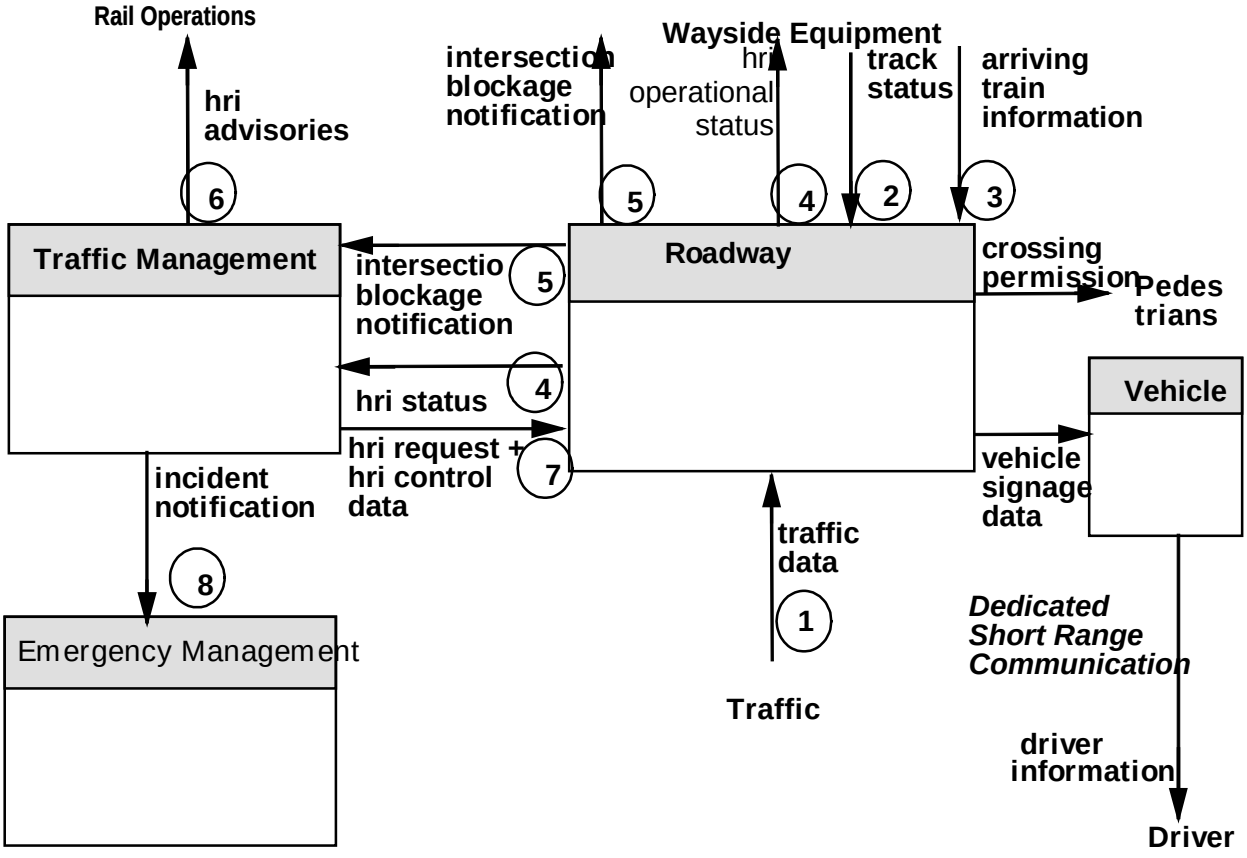

Figure 33. Physical Architecture for High Speed Railroad Grade Crossings

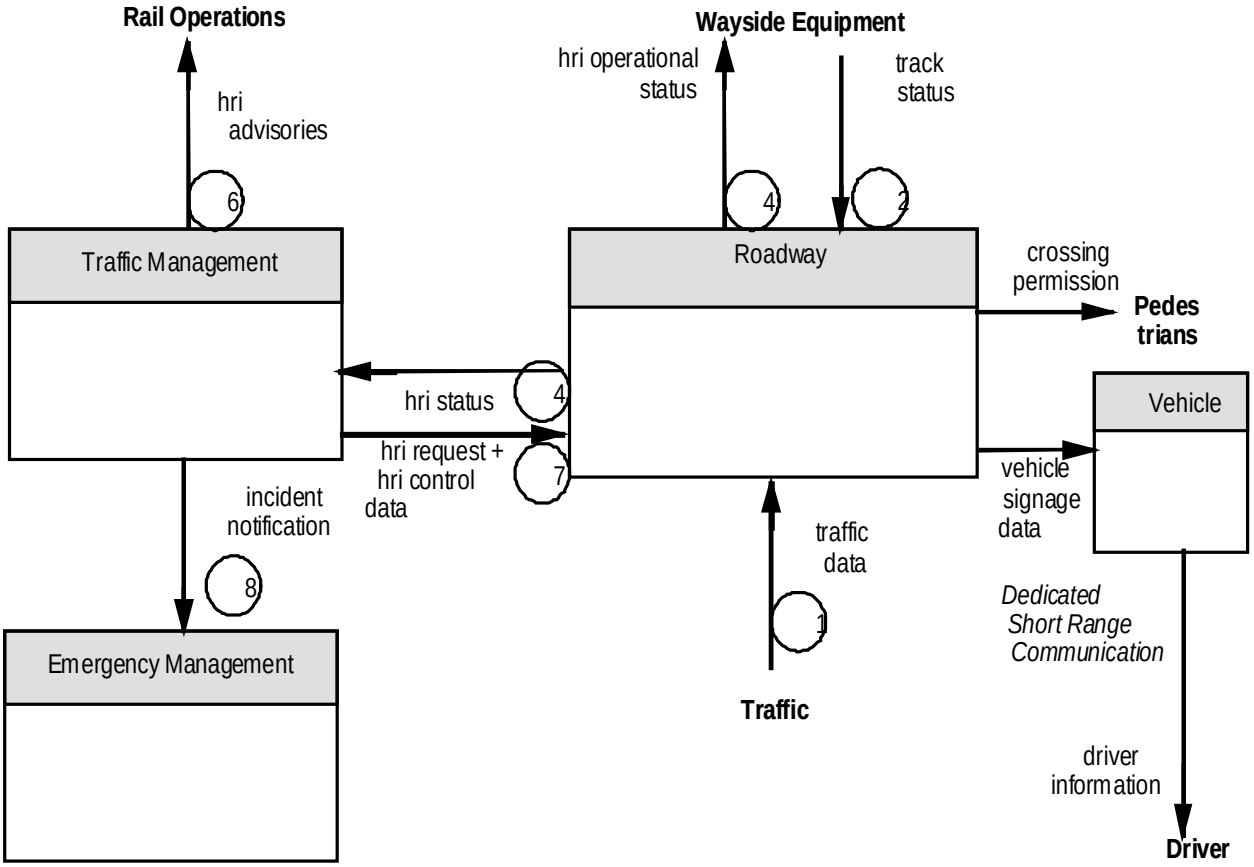

Figure 34. Physical Architecture for Standard Speed Railroad Grade Crossings

The following describes the operation of both types of crossings, highlighting where differences occur between standard speed and high speed HRI deployments.

1. The Roadway subsystem determines the status of the traffic at or near the HRI. This typically involves loops or other detection devices where the "traffic data" flow represents an input to the surveillance process. In the Standard Speed Rail Crossing Equipment Package this surveillance may be limited to nearby road sensors (e.g. loop detectors), but in the High Speed Rail Crossing Equipment Package the surveillance includes some way of identifying that the HRI is blocked with a vehicle or other object. This could be from visual or infrared sensors, or whatever other technology is appropriate.

2. The Wayside Equipment provides a real time indication of its operational status via the track status message. This would alert the roadside equipment to possible failures or problems in the wayside equipment. It also provides a simple indication of the arrival of a train (i.e. the track circuit indication currently used for interconnecting active warning devices to nearby traffic controller).

3. For the high speed railroad grade crossing requirements, the Wayside Equipment provides real time information on the approach of a train. This is a more advanced set of information than the track status above (which could be the binary indication from a simple track circuit). The wayside equipment would provide expected time of arrival and length of closure from, for example, either advanced track circuits or advanced positive train separation systems. Note that these systems which develop real time information on the approach of a train will vary with specific rail operator deployments. By using the Wayside Equipment terminator, the architecture is not specifying how this information must be developed.

4. The “hri status” message is issued to the TMS and to the Wayside Equipment terminator. This message is made up of several components: information about the crossing itself (confirmation that the highway grade crossing is closed and that trains may proceed at full authorized speed), information about the traffic in the neighborhood of the crossing, information about the expected closure time and duration (obtained from the wayside equipment), and information which can be displayed via DMS or beacon (for in-vehicle signing).

5. If the deployment is for a High Speed Rail Crossing Equipment Package, an “intersection blockage notification” message is sent from the Roadway to the Traffic Management subsystem and to the Wayside Interface Equipment terminator to provide an indication if a blockage of the HRI exists. This supports the unique requirement on the high speed crossings to provide a blocked intersection warning which could be sent directly to the train crew (either from the Wayside Equipment directly, or via Rail Operations). In the event that a detected vehicle is entrapped in a four quadrant gate installation, the Roadway subsystem process will lift the appropriate gate(s) so that the entrapped vehicle may exit the intersection.

6. The TMS would notify the RO in near real time about equipment failure, intersection blockage, or other incident information (e.g., nearby HAZMAT spill). The TMS would also send information about planned maintenance activities which are occurring at or near the grade crossing and which could impact the railroad right of way.

7. The TMS will communicate with the Roadway with two messages: “hri control data” and “hri request.” The hri control data message includes three components: “indicator sign control data” which is used within the Manage Traffic function and contains the actual data from which instructions to the driver and traveler can be produced by indicators, dynamic message (DMS), advisory beacons, and other types of signs on the roads (surface streets) in the vicinity of railroad grade crossings; “rail operations advisories” which contains advisory information for HRI vehicular traffic that has been derived from information received from Rail Operations (see Section 4.8.2.) and forwarded by the TMS; and “rail operations device command” which contains HRI device commands that have been derived from information received from Rail Operations and provides for Rail Operations preemption capability. The hri request data message includes three components: “hri traffic surveillance” which represents the various traffic sensor inputs to HRI from the Traffic Surveillance processes; “rail operations requests” which is generated in response to a need for HRI status for Rail Operations; and “TMS requests” which is generated in response to a need for HRI status for Traffic Management.

8. In either subservice, if a collision should occur between train and vehicle the TMS would receive the collision notification as part of the HRI status message and could forward this information to the Emergency Management Subsystem as a part of its incident management process.

4.8.2 Rail Operations and Traffic Management Coordination

The second aspect of the HRI User Service is the provision for coordination between Rail Operations and Traffic Management. In the architecture implementation of this user service the Traffic Management subsystem first collects information from Rail Operations about scheduled and anticipated HRI closings as well as actual roadway conditions at the HRI intersections. This combined information is used to prepare signal control plans and ATIS messages which are communicated to the Roadway for travelers at or approaching the HRI, and also communicated to the ISP for travel planning by the customers of the ISP. The following discussion contains some redundancy with the former discussion on safety features, because the coordination aspects of the HRI are related and share functionality with the safety improvement features.

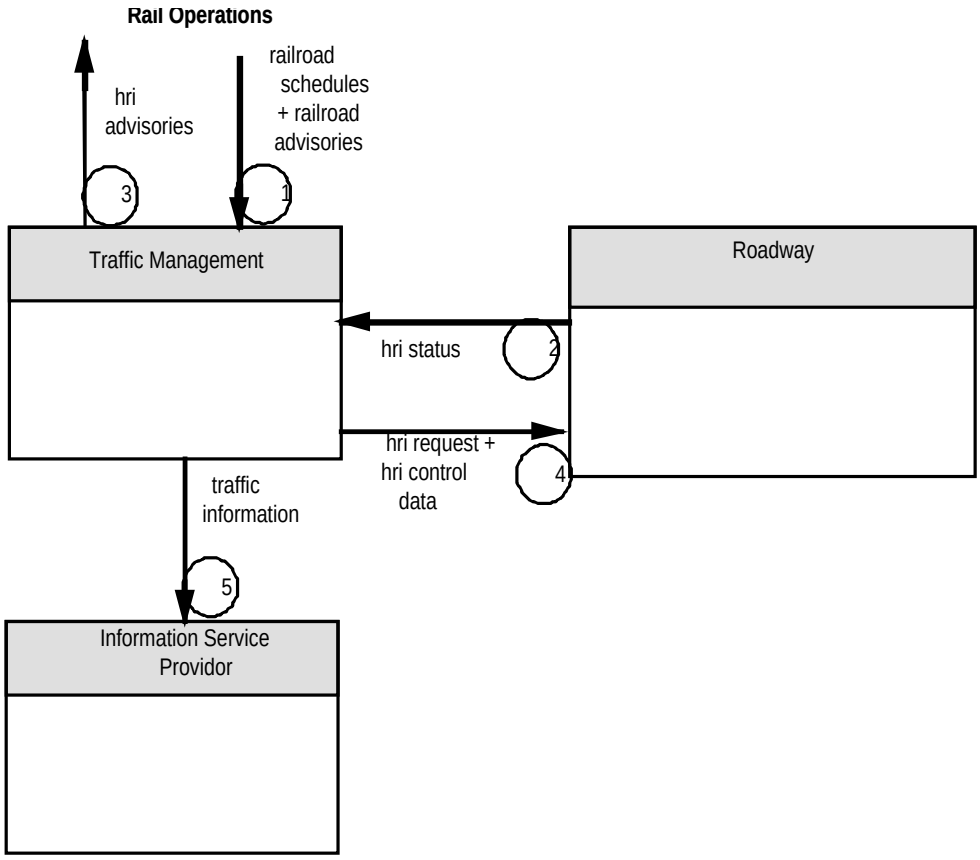

Figure 35. Physical Architecture for Rail Operations and Traffic Management Coordination

1. The Rail Operations function will send information to the TMS to support forecasting of HRI closures. This includes train schedules (that would be pertinent to HRIs) and maintenance schedules (which would affect HRIs) in the message "railroad schedules." In addition the RO will send the message "railroad advisories" to the TMS including information about rail incidents which may impact vehicle traffic. This latter message would be sent in near real time, while the "railroad schedules" message would be provided on a periodic basis (e.g., daily).

2. The "hri status" message is issued to the TMS by the Roadside subsystem. This message is made up of several components: information about the crossing itself (confirmation that the highway grade crossing is open or closed), information about the traffic in the neighborhood of the crossing, information about the expected closure time and duration (obtained from the wayside equipment), and information which should be displayed via DMS or beacon (for in-vehicle signing).

3. The TMS would notify the RO in near real time about equipment failure, intersection blockage, or other incident information (e.g., nearby HAZMAT spill). The TMS would also send information about planned maintenance activities which are occurring at or near the grade crossing and which could impact the railroad right of way.

4. The TMS will communicate with the Roadway with two messages: "hri control data" and "hri request." The hri control data message includes three components: "indicator sign control data" which is used within the Manage Traffic function and contains the actual data from which instructions to the driver and traveler can be produced by indicators, dynamic message (dms),

advisory beacons, and other types of signs on the roads (surface streets) in the vicinity of railroad grade crossings; "rail operations advisories" which contains advisory information for HRI vehicular traffic that has been derived from information received from Rail Operations (see Section 4.8.2.) and forwarded by the TMS; and "rail operations device command" which contains HRI device commands that have been derived from information received from Rail Operations and provides for Rail Operations preemption capability. The hri request data message includes three components: "hri traffic surveillance" which represents the various traffic sensor inputs to HRI from the Traffic Surveillance processes; "rail operations requests" which is generated in response to a need for HRI status for Rail Operations; and "TMS requests" which is generated in response to a need for HRI status for Traffic Management.

5. Finally the TMS communicates anticipated and scheduled HRI closures to vehicles to the ISP for broader dissemination and use in travel planning. This information is in the "traffic information" message described in Section 4.1.6.

4.9 Managing Archived Data

The purpose of this service is to collect ITS and related data, archive it, and make it available to other users. The service centers on the Archived Data Management Subsystem (ADMS), with its architecture flow connections to other subsystems and terminators shown in Figure 36.

The Archived Data Management Subsystem collects, archives, manages, and distributes data generated from ITS sources for use in transportation administration, policy evaluation, safety, planning, performance monitoring, program assessment, operations, and research applications. The data received is formatted, tagged with attributes that define the data source, conditions under which it was collected, data transformations, and other information (i.e. meta data) necessary to interpret the data. The subsystem can fuse ITS generated data with data from non-ITS sources and other archives to generate information products utilizing data from multiple functional areas, modes, and jurisdictions. The subsystem prepares data products that can serve as inputs to Federal, State, and local data reporting systems. This subsystem may be implemented in many different ways. It may reside within an operational center and provide focused access to a particular agency's data archives. Alternatively, it may operate as a distinct center that collects data from multiple agencies and sources and provides a general data warehouse service for a region.

Figure 36 only illustrates an example source of archive data (the Traffic Management Subsystem). Any one (or more) of the following subsystem or terminator ITS data sources shown in Table 5 can be a source for archive data.

Library Analogy

An ADMS can be viewed as a library, with the Archived Data Administrator terminator viewed as the librarian. Referring to Figure 36, the ADMS, like a library, has users (the Archived Data User Systems and Government Reporting Systems), publishers that supply content (Architecture data sources, Other Data Sources and Map Update Providers) as well as an interface to the content of other libraries (Other Archives). Finally, user payments for data and associated analysis services are facilitated with an interface to Financial Institutions (enabling them to pay by credit/debit/stored-value instruments).

The ADMS can be operated either privately or publicly. Parameters such as what data is stored locally, the length of data retention, sources of data, analytical services, charges (if any) and user access are determined by the individual operator. As with other subsytems, the ADMS can be aggregated with other physical architecture entities. For example, an ADMS may be aggregated in some regions with the

Traffic Management Subsystem, and operated by the same agency. Institutions or entities that are unable to access an Archive that serves their needs may choose to operate their own archive.

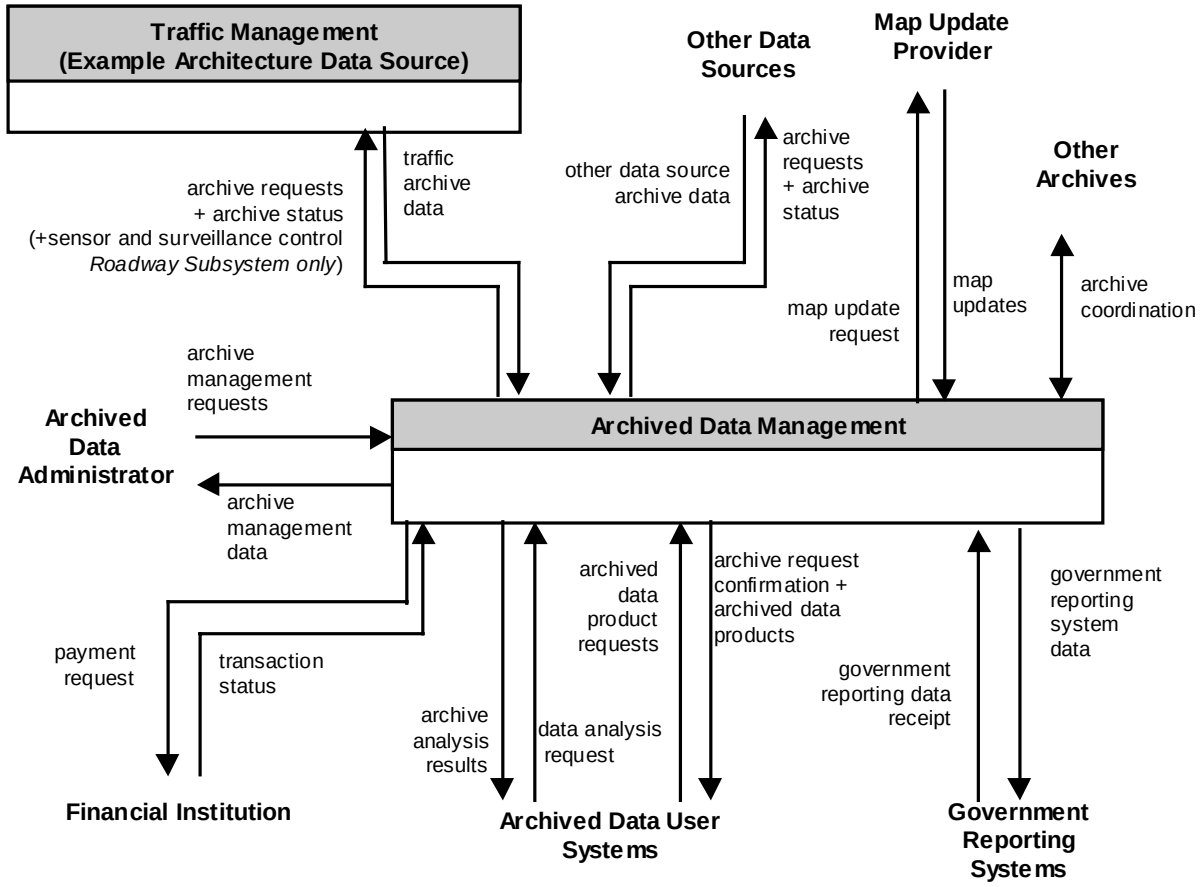

Figure 36. Physical Architecture for Managing Archived Data

Internally, an ADMS can include the following elements:

- Archive Data Administrator Interface. This is the human interface to this terminator.
- The local archive. The local archive represents the data stored locally in the ADMS. It is up to the Archive Data Administrator to specify what data is stored, how long it is stored, how data is formatted and checked prior to storage, what information is not stored/cleansed (e.g. some personal information), and permissions for users (or other systems) to access the data.

  The local archive has a schema (organization) that is specified by the Archive Data Administrator. The schema corresponds roughly to a library catalog. It incorporates rules for how data is stored and related to other data. While standards for ITS Archive Schema do not exist today, broad usage of a common schema could enable easy access to data across many archives. It is expected that data which is regularly "rolled up" on a statewide or national level (such as some government reporting

data e.g. Highway Performance Monitoring System measurements) will benefit from a standardized schema.

Subsystems	Architecture Flow Containing Data for Archiving
Commercial Vehicle Administration	commercial vehicle archive data
Emergency Management	emergency archive data
Emissions Management	emission archive data
Toll Administration	toll archive data
Parking Management	parking archive data
Roadway Subsystem	roadside archive data
Traffic Management	traffic archive data
Transit Management	transit archive data
Information Service Provider	traveler archive data
Terminators	
Intermodal Freight Depot	intermodal freight archive data
Multimodal Transportation Service Provider	multimodal archive data
Construction and Maintenance	construction and maintenance archive data
Other Data Sources	other data source archive data
Map Update Provider	map updates
Weather Service	weather information

Table 5 Architecture Data Sources and their corresponding Architecture Flow containing Data for Archiving

- Machine interface to Archive Data User Systems. The interface to this terminator will almost certainly use the World Wide Web for many implementations of the ADMS. As such, it might be expected that the interface could use Hypertext Markup Language (HTML) or the evolving Extensible Markup Language (XML) to present graphical or tabular data or analytical results to the users. Similarly, File Transfer Protocol (FTP) could be used to deliver large amounts of data to user systems for subsequent analysis. Of course, other protocols and methods are also possible, and will ultimately be decided by individual deployers and standards committees.

  Users will in some ADMS archives be able to select either data products, or analysis products. Data products are datasets that the user has selected from the archive catalog. Analysis products are the results of specific operations on data that the archive supports (e.g. data fusion, data mining, and aggregation).

- Interface to Other Archives. The ADMS can access the schema and data of other archives (that it has permission to access). It can exchange schemas (archive organization rules) with the other archives, and it can exchange catalogs. The catalog indicates to a remote archive what data is currently stored in the archive. Collectively, the local and remote schemas and catalogs can be presented to the user

to indicate what data is available locally and remotely. This is necessary to allow the management and selection of large datasets while minimizing transmission and storage costs.

- Machine interface to Architecture and Other Data Sources. The ADMS will access data and catalogs from National ITS Architecture Subsystems and Terminators as necessary. This interface will generally be wireline, and will use standard message sets and protocols to be determined by standards development organizations.

The ADMS has the following major functions:

1. Collecting and storing data in the archive. This can include collecting data from other archives or simply collecting information about what is in another archive (its schema and catalogue).

2. Managing the archive. This involves monitoring the operation of the archive by the Archive Data Administrator.

3. Enabling users to access the archive. Here users can either request data, request the results of analysis on archive data, or request special analysis results or data for specific government reporting system requirements.

Figure 37 illustrates the sequence of architecture flows (time proceeds from top to bottom) that make up transactions involving the Archive Data Management Subsystem to implement the major functions stated above. The vertical bars in Figure 37 represent the subsystems and terminators in Figure 36, and the horizontal arrows represent the architecture flows. The architecture flows begin and end at the their respective origin and destination. Finally, the labels on the architecture flows include a numbering scheme that illustrates the approximate sequence of events. Architecture flows that have the same number can happen at the same time or are alternative flows, as discussed in the following text.

1) Initiation of Archive Activity. Archive activity can be initiated by either the Archive Data Administrator or an Archive Data User System. (Note, activity can also be initiated by a remote archive, see 5c and 6c below).

   a) The Archive Data Administrator can issue an Archive Management Request message. This contains requests and inputs from the archive data administrator terminator to request that data be archived and the parameters needed to control the import of the data. This data flow also contains the security permission data necessary to ensure the archive data is secure. This data flow contains the requests sent to the Manage Archive function to administer the archive database.

   b) The Archived Data User Systems can request either a data product or an analysis product by issuing the appropriate message. If the ADMS requires payment to service the request, either of these messages can include payment instrument identification information.

2) If payment is required to service an Archived Data User System request for data or analysis products, then a Payment Request message is issued from the ADMS to the Financial Institution terminator.

3) If the Payment Request message was sent to the Financial Institution, then it will respond with a Transaction Status message to the ADMS, indicating payment is confirmed or rejected.

4) This architecture flow from the Manage Archived Data function to the Archive Data User System terminator contains the confirmation of whether the requested data will be imported into the archive and how the data will be identified.

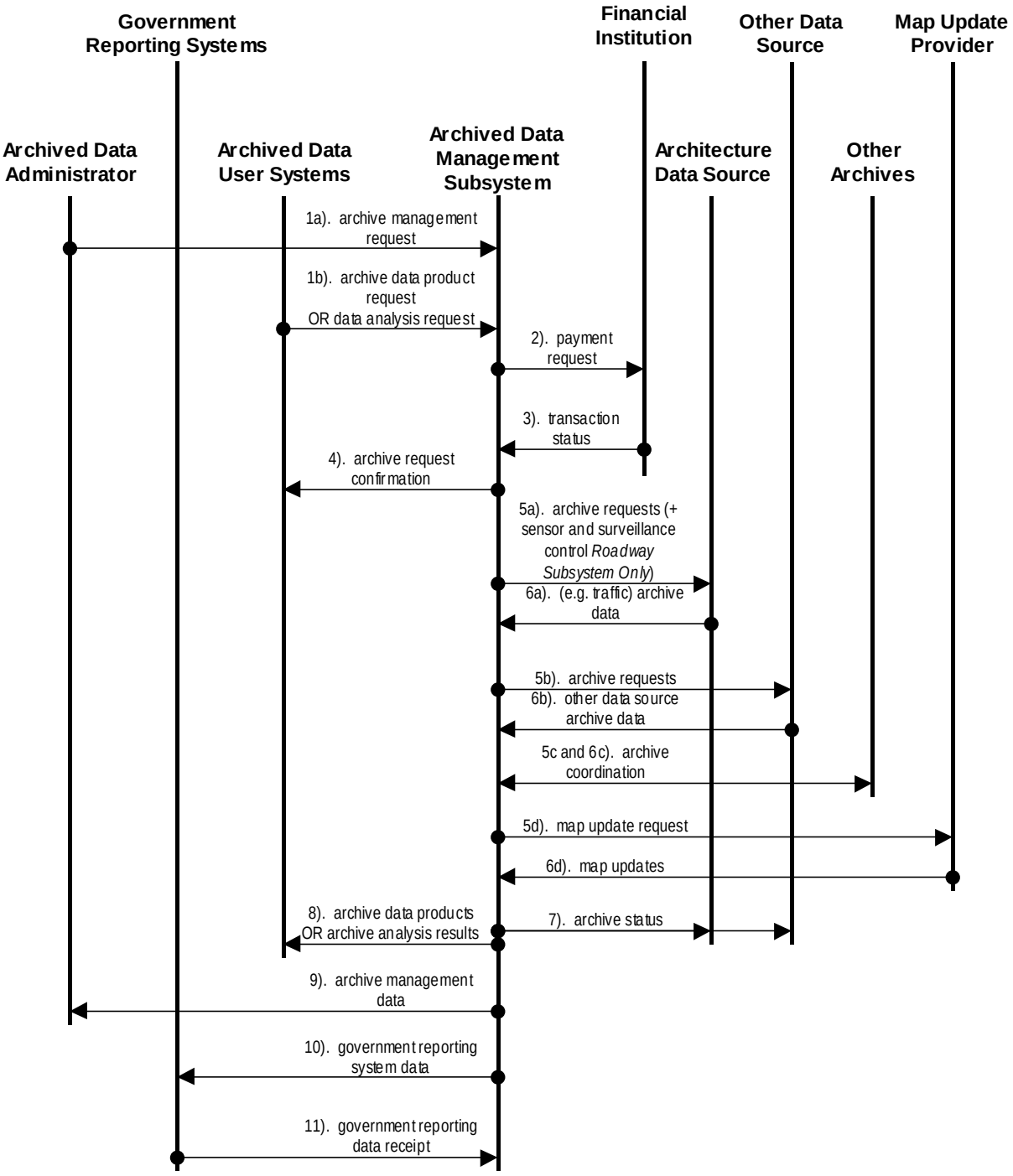

Figure 37 Architecture Flow Sequencing for All Services Involving the Archive Data Management Subsystem

5) This group of messages is from the ADMS to the sources of data for the ADMS. The ADMS is essentially a client accessing data from various servers to respond to a Archived Data User Systems request or a request from the Archived Data Administrator. Each of the request messages here has a corresponding response message in the next group (e.g. request 5a is followed by response 6a, 5b by

6b, etc.) The requests/responses in this and the next group can proceed sequentially or in parallel. The particular strategy chosen is an implementation (not an architectural) issue.

a) Archive Request to Architecture Data Source. The architecture data source is one of the subsystems or terminators that can source archive data to the ADMS. This request can be for a catalog of available data or for specific data itself (identified through a previous access of the catalog).

b) Archive Requests to Other Data Sources. This data flow contains the request for data collected and stored by sources external to ITS that may be of interest to archived data users systems that is not included in data from sources within the ITS subsystems and terminators. This data flow includes request for a catalog of the information available as well as the request for the data itself.

c) This architecture flow represents the data that is to be shared between different Archive systems. Information included on this interface can include the requests for data that is located in other systems. This data flow also represents the flow of data from the local archive to the other archive system.

d) The ADMS may use maps provided by a map update provider. If so, this interface allows maps to be requested.

6) Responses to requests for data in the previous group. In general, responses can be single, continuous/periodic, or event driven as specifically requested.

a) This represents the catalogs or data that the requested architecture source provides in response to the corresponding request in 5a above. For each National ITS Architecture subsystem, the architecture flow is made up of data flows representing operational data that may be of interest to an archive. The actual data provided will be determined by the operator of the subsystem or terminator in each instance. For example, the Commercial Vehicle Administration subsystem provides in its architecture flow commercial vehicle archive data:

- cv_credentials_enrollment_attributes
- cv_credentials_enrollment_data
- cv_daily_logs
- cv_daily_logs_attributes

The component data flows can be identified for each of the architecture flows identified in Table 5 using the physical and logical architecture documentation or (more easily) using the hypertext documentation of the National ITS Architecture.

b) The data flow from Other Data Sources to the Manage Archived Data function contains a catalog and details of data. This represents data and associated meta-data that may be of interest to the archive data users systems that cannot be obtained directly from ITS subsystems or terminators. The specific meta-data for this data is listed below. Detailed definitions for the metadata can be found in the Logical Architecture Data Dictionary for these meta data attributes:

- authorization_to_use
- collection_conditions
- collection_equipment
- data_aggregation
- data_concept_identifier
- data_reductions
- data_revision

- data_version
- date_archived
- date_created
- date_published
- equipment_status
- error_handling
- methods_applied
- owner_entities
- perishability_date
- personal_identification_status
- quality_control_attribute
- record_size
- security
- standard_data_attribute
- standard_message_attribute

c) See 5c above.

d) Requested maps are delivered on this architecture flow.

7) The ADMS sends notification to the source of the data that the data provided contains erroneous, missing, or suspicious data or verification that the data provided appears valid. If an error has been detected, the offending data and the nature of the potential problem are identified.

8) If architecture flow 1b above (archive data product request or data analysis request from the Archived Data User Systems) was used to start the transaction, then archive data and/or catalogs are delivered to the Archived Data User Systems.

9) If architecture flow 1a above (archive management request from the Archived Data Administrator) was used to start the transaction, then "archive management data" is sent to the Archived Data Administrator in response.

10) If the Archived Data Administrator initiated the transaction to deliver "government reporting system data" to the Government Reporting Systems, then this architecture flow is sent.

11) The Government Reporting System acknowledges receipt of government reporting system data with this architecture flow back to the ADMS.

5 Operational and Interoperability Issues

The process of defining and building an architecture leads to many crossroads, where fundamental decisions must be made about the approach to be used. The architecture that finally evolves is the sum of these decisions. Clearly, an understanding of the architectural and non-architectural implications of different decisions is critical to the quality of the ultimate product.

This chapter identifies key operational and interoperability issues, and frames the architecture options and chosen strategies for addressing them.

5.1 Communications Systems

5.1.1 Wireline Communications

ITS Wireline Communications were discussed in Section 3.2.1. The issues for wireline communications as discussed there as well. By separating the ITS architecture into a communications and transportation layer (as well as an institutional layer), and by modularizing the communications interfaces on the ITS subsystems, the choice of which particular communications technology is used is a local decision, and does not have an impact on the ITS system. As long as the chosen communication technology meets some basic requirements, summarized in section 5.1.4., the choice is not critical. Similarly, the choice as to whether to deploy a dedicated system for a particular application, or to buy communication services, is not critical to the architecture.

5.1.2 Wide Area Wireless

ITS Wireless Communications were discussed in Section 3.2.2. The wireless WAN communication elements of ITS are analogous to the wireline communication elements in many ways. The communication element can be dedicated to a specific user or agency (and publicly owned or privately owned), or it can be privately owned and operated by a communication service provider who sells access to this data network to many users or agencies for a fee. As long as the chosen communication technology meets some basic requirements, summarized in section 5.1.4., the choice is not critical.

Wireless communication systems can be one-way (broadcast) or two-way. For broadcast systems, an example is FM-subcarrier or paging systems. Two-way systems that are private can be SMR or E-SMR. SMR and E-SMR require licenses from the FCC for operation, and are typically dedicated to a specific service or agency. The issues of one-way vs. two-way broadcast are summarized in Table 6. Both broadcast modalities and two-way are supported in the architecture for various appropriate applications. This approach allows attractive early deployments using mature and emerging one-way data services (e.g., pager technology and FM subcarrier), yet also supports use or evolution towards more functionally rich two-way modalities.

5.1.3 DSRC (Dedicated Short Range Communication) Beacons

5.1.3.1 DSRC and Wireless Wide Area Network Characteristics

In Section 3.2.3., some of the characteristics of DSRC were discussed. Table 7 that follows comparatively summarizes these characteristics of DSRC Beacons and Wide Area Wireless communications.

	One-way Data	Two-Way Data
Cost	Lower to Medium.	Higher
Coverage of User Services	Partial. Some services require two-way transmissions.	All services covered.
Payment for ATIS Services	Billing model is not clear for FM-subcarrier (unless publicly funded).	Billing model is clear.
Leverage existing facilities	Can make use of existing commercial broadcast/pager facilities.	Some systems build on existing mobile telephone services.
Geographic Coverage	Regional for FM-subcarrier. National for some pager systems.	National.
Geographic Interoperability	Requires that FM Broadcasters in different regions agree to broadcast using the same standards. Likely, especially if broadcasters are affiliated with a commercial network. Some national pager networks with national or emerging national coverage.	Many 2-way communication services will offer national coverage, the mobile user only needs to select a single carrier/provider**.**
Rural Coverage	FM: No (or at risk). Pagers: Yes.	Yes (in the 10-year timeframe, due to Low Earth Orbit satellite coverage).

Table 6. Implications of One-Way and Two-Way Communications

Beacon communication capability may be deployed in Roadway and Vehicle subsystems where its special roadway location characteristics give it an advantage or where the cost (considering capitalization and operations / maintenance) of the beacon communication equipment is less than the cost of the equivalent Wireless communications infrastructure. For example, toll and parking payment operations, CVO operations or urban fixed route public transit operations may be more economically served by beacon communications for messaging between the vehicles and the appropriate Roadway or Center subsystems. In these cases, the specific sources and availability of deployment capital as well as operations and maintenance expenses must be considered.

	DSRC Beacon	Wide Area Wireless
Access to right-of-way	Required. Multi-jurisdictional access negotiations may be necessary for applications crossing jurisdictional boundaries.	No access required.
Equity	Public funding of beacons risks concerns of direct benefits exclusively to travelers with expensive in vehicle equipment (except for signage only beacons). Private funding of beacons requires a new "franchise" strategy to equitably give limited public right-of-way access to private concerns.	Exclusively private expense for private individual benefit. Uses established FCC strategy to allocation of public airwaves for commercial use and special access for public agency use e.g. SMR (Special Mobile Radio) licenses.
Breadth of Coverage	Limited to where beacons are deployed. Deployment first on main corridors. Unlikely deployment on rural, feeder and residential roads.	Limited to where cell sites are deployed, moving towards ubiquitous coverage.
Cost of deployment to ITS	The cost of deployment is entirely an ITS expense, since DSRC is only used for ITS.	The cost of deployment is shared with non-ITS applications.
Operating and Maintenance Charges	As an ITS deployment, only the direct costs of maintenance apply (although some public agencies may resist the new operations and maintenance burden that beacons on a public right-of-way imply).	Air time charges by private cellular carriers.
Ubiquitous coverage	Coverage limited to a small area ("field-of-view") near beacons located on the roadway right-of-way.	Ubiquitous coverage in urban areas with growing coverage on interurban and rural areas.
Data Throughput	Linearly related to number of beacons.	Linearly related to number of base stations.
Latency Tolerance	Messages must be transferred while the vehicle is in the field of view of the beacon---or wait until in the field of view of the next beacon. For dense deployments of beacons with broadcast messages (e.g. traveler advisory), the beacon has a "just in time" characteristic that is very desirable.	Data can be queued until access is available.
Inherent Vehicle Location Determination	Yes. Location is determined at the site of the beacon. Accuracy is directly related to the beacon field-of-view. This must be augmented with another in-vehicle location mechanism to determine vehicle position between beacons.	Difficult today. In general, requires a separate vehicle location determination equipment package in the vehicle. Vehicle location could be determined by the cellular provider, and then communicated to the vehicle. Accuracy is currently poor for most ITS services, and improvements are still in

		development which is driven by a need to locate cellular subscribers making 9-1-1 calls with location information going to the 9-1-1 PSAP (Public Safety Answering Point), not the subscriber.
Vehicle to Center messages	Yes. "Uplink" messages can be routed from the beacon through the wireline network to a specific center. "Downlink" is much more difficult.	Yes.
Center to Vehicle Broadcast	Yes. All vehicles at a beacon can get the same message.	Yes.
Center to (single) Vehicle messages	Difficult. In general, center subsystems will not know which beacon to send a personalized message to.	Yes. This is a feature of the "roaming" capability of cellular systems. It is based on the overlap between adjacent cells in the system and control messages to/from subscribers so that at any time the cellular system knows with which cell site each subscriber is in contact.
Standards	No open standards in use for the airlink -- therefore all subscriber (mobile) units are currently proprietary.	Both open (e.g. CDPD and GSM) and proprietary (e.g. RAM and ARDIS) interfaces in commercial use.

Table 7. Characteristics of Beacon DSRC and Cellular Communications

5.1.3.2 DSRC and Wireless Wide Area Network Applicability by Application

Table 8 identifies the choices made in the architecture between DSRC Beacons and Wide Area Wireless communications and rationale for specific ITS services. There is a tension in the architecture definition effort to not limit technology choices so as to accommodate all options and let the market make decisions, and at the same time make prescriptive technology choices that will guide investments (especially when public assets are involved) to reduce risk (of selecting a "Betamax"), assure interoperability, and accelerate deployment of ITS. This presumes that the architecture team, using a consensus process, can themselves avoid prescriptions for "Betamax" like technologies in the architecture. Both approaches have well known advantages and disadvantages. The choices indicated in Table 8 reflect this tension, especially in the area of Vehicle Navigation.

	DSRC Beacon	**Wide Area Wireless**
Toll Payment	Yes. Positive vehicle location and payment transaction with low cost in-vehicle equipment.	No.
Parking Payment	Yes. Positive vehicle location and payment transaction at the parking location with low cost in-vehicle equipment.	Yes - if vehicle is prepaid / confirmed (e.g. by cellular) then vehicle can be verified at parking lot by reading license plate by parking management subsystem.
Parking Pre Payment	No. Prepayment / reservation is difficult with beacon alone because it is hard to assure that the "confirmation" (that a space is available) downlink message can be sent to the vehicle	Yes.

	via beacon.	
Commercial Vehicle Checking Operations	Yes. Note that processing latencies between the initial vehicle tag read by beacon and the "pass / pull-in" message by beacon may be tolerated if a pair of beacons are used on a stretch of road prior to the Commercial Vehicle Check subsystem, separated by enough space such that vehicles traveling at maximum mainline speeds will reach the second beacon for the "pass / pull-in" message after the worst case processing latency time.	No.
Transit Vehicle Operations	Yes, for uplink of operations data and downlink messages that go to all transit vehicles.	Yes.
Transit Vehicle Signal Priority and Emergency Vehicle Signal Preemption	Yes. Direct vehicle to roadway signal.	Yes, via Traffic Management Center subsystem. Allows for signal coordination that can anticipate requesting vehicle turning movements and minimize overall system disruption. (E.g. GPS in the vehicle with differential correction at the TRMS or ISP for location accuracy < 10 meters.)
In Vehicle Signage	Yes.	Yes. Downloaded as part of a "route plan."
Vehicle Navigation (Route Selection in Vehicle)	No. Either link time variances or suggested routes can be broadcast at each beacon site. Because of the limited time a vehicle is in the field of the beacon, some significant compromises must be made to select the information sent to the vehicle e.g. complete information local to the beacon, and sparse information remote to the beacon. Vehicles must delay getting real-time information updates until they reach a beacon - when the information may have become "stale". Dense deployments of beacons to counter this effect may be prohibitively expensive.	Yes. A critical and desirable feature is that trip selection can be done and traveler information accessed at the very beginning of a trip, prior to making any directional choices, mode choices or beginning to drive to a beacon.
Vehicle Navigation (Route Selection in Infrastructure and optionally coordinated with ATMS)	No. The worst case request / response latency will generally be longer than the time it takes a vehicle at mainline speeds to transit the beacon field of view. Furthermore, traveler requests for real time information or routes will have to wait for responses until the vehicle passes a beacon. This may not give satisfactory perceptions of service.	Yes. A critical and desirable feature is that trip selection can be done and traveler information accessed at the very beginning of a trip, prior to making any directional choices, mode choices or beginning to drive to a beacon.
Emergency Request	No.	Yes.

Table 8. Applicability of Beacon DSRC and Wide Area Cellular Communications to Specific ITS Services

5.1.3.3 Expected DSRC and Wide Area Wireless Coverage

The previous sections have identified some key themes regarding the use of beacons and wide-area network wireless communications. Wide area wireless communication is best suited for services benefiting from near ubiquitous coverage e.g.:

- Traveler Information
- Route Guidance
- CVO/Fleet Management
- Emergency Response

DSRC communication, on the other hand, is best suited for applications where services will benefit from the location specific nature of each beacon installation e.g.:

- Parking Systems
- Highway/Rail Crossings
- Toll Systems
- Transit systems (e.g. fixed urban route)
- Traffic Probes (using ETTM tags)
- Intersection Collision Avoidance
- In-vehicle signing. Note that beacons for in-vehicle signage are generally viewed as very simple, low cost and physically robust modules that simply broadcast fixed signage information. For example. these beacons could be powered by solar cells and would not require a wireline communication interface.

5.1.4 Dedicated and Shared Communications

We think it likely that in the 20-year timeframe, WAN data communications will have become largely a commodity with many competing suppliers of wireline and wireless services. As a commodity, the ITS Subsystems (not a commodity) will use whatever WAN communication service is available that meets the operators needs at the lowest price. For this reason, the ITS Architecture has been decomposed into three layers: the Institutional, Transportation and Communication layers. (The institutional layer is addressed in the Physical Architecture document and elsewhere.) The Transportation layer is composed of the ITS subsystems. The Communications layer is composed of several communications "elements" of which the WAN wireless and wireline elements are those which we believe to be strong candidates for becoming commodities. The communication elements can be dedicated or shared. Because of the expected "commoditization" of communications, and the cost and performance benefits to be had at that time, the following three requirements can be stated for any WAN communication technology to participate as an ITS communication element:

1. The interfaces use open standards.

   This guarantees that ITS subsystem equipment from many competing manufacturers can be used to connect using the communications element. The cost of the data communication module (e.g., modem, transceiver) should be small relative to the ITS subsystem.

2. The communication element be internetworked with other communication elements.

   The communication element provider must participate in the open internetworking standards that enable messaging between users of different communication element technologies.

3. The communication element be nearly ubiquitous to the nation or at least a region.

   This enables users to "roam" over a substantial area of user interest and have seamless access to ITS services. The roaming capability is supported by the communication service provider.

Today, since the commoditization of WAN data communications is still in its early stages, many ITS early deployments and Field Operational Tests (FOTs) properly use dedicated WAN systems because shared WAN service providers offering adequate commercial service may not be available. Taking a 20-year perspective, the implications of the different approaches (and the risks) are summarized in Table 9. The architecture supports both private and shared deployments as appropriate. This will allow early deployments using mature dedicated data services (e.g. SONET fiber, SMR and E-SMR) and evolve to new deployments to shared technologies as they continue to emerge and make sense locally.

	Dedicated ITS	**Shared Use**
Capitalization	Federal funds can be used if publicly owned.	Private capital available for national-scale infrastructures.
Operations and Maintenance Funds	Lowest in the near term.	Possibly lowest in the 20-year timeframe (because operations and maintenance is shared).
Wireless Spectrum Use Efficiency		Operators highly motivated to use spectrum efficiently.
Cost of Access	Cost of O&M	Competitive usage fees.
Public Safety Access	Special FCC allocated frequencies for Police, Fire, EMS.	Regulations may be needed to guarantee low-cost access and guaranteed performance (e.g., guaranteed maximum latency) for public safety agencies (that currently depend on dedicated systems.
Efficient use of Spectrum	Motivated in some cases (FCC pressure). Many systems use the less efficient "trunked" mode.	Highly motivated to serve demand with limited spectrum. Most systems use the more efficient "cellular" mode.
Technology Deployment	Infrequently refreshed.	Frequently refreshed in competitive markets.
Long Term Risks	Demand for spectrum may force more efficient use of available spectrum and unexpected capital expenditures.	The market may not become large enough to allow the development of a competitive market.
	Maintenance of dedicated systems may exceed cost of purchasing communication services from competitive/regulated providers.	The technology may not become sufficiently mature to allow the development of a competitive market.
	Future public capital may not be available for the deployment of large systems.	Future private capital may not be available for the deployment of large systems.

Table 9. WAN Data Communications Dedicated vs Shared Implications

5.1.5 Use of Electronic Data Interchange (EDI) Standards

EDI is a standard mechanism, embodied in the X12 and more recent EDIFACT protocols, for exchanging messages between computers. EDI is used in many commercial inter-enterprise business processes in North America, and in particular, is very heavily used in the commercial shipping industries, as well as

elsewhere. It is expected that the CVO messages defined in this architecture will be communicated between fixed business and government subsystems using these protocols. (Messages to/from vehicles over dedicated short range wireless links may benefit from a simpler protocol.) The choice as to specific protocols is of course, ultimately up to the consensus standards committees that will use the ITS National Architecture as input for their standards development role.

5.2 Emergency Notification and Personal Security

Two key issues in this service required to support seamless operation over geography are: the location determination of the mobile emergency requester and the mechanism by which the initial emergency message is routed to the appropriate EM.

5.2.1 Emergency Notification Location Determination

The location of the emergency calling mobile traveler must be communicated to the EM along with the emergency request message from the traveler. The location could be determined by the travelers mobile subsystem or can be done by the WAN wireless communication element. The options are compared in Table 10.

	In Vehicle Location Determination	**Wireless WAN Communications Based Location Determination**
Technical Risk	Lower	Higher
Compatibility with other User Services Requiring Location	Good.	Poor. Would require communications interactions where not otherwise necessary.
Compatibility with Mayday Only Deployments	Higher cost, but marginal cost for in-vehicle location capability is expected to fall.	Lower mobile cost. Higher communications infrastructure cost. Lower overall system cost.
Accuracy (guess of 20-year capability)	Good (1 meter?)	Fair to Good (10-150 meters?)
Current Plans and Deployments	Ford in 1996 (RESCU - Remote Emergency Satellite Cellular Unit) will have a proprietary national Emergency Management Center located in Irving, TX to receive vehicle emergency requests.	Proposals for AMPS Emergency Telecommunication System enhancement which might have some applicability to data emergency request messages.

Table 10. Implications of Alternative Traveler Location Mechanisms

In the architecture the location of the Driver or Traveler sending a data emergency assistance request is determined by the mobile subsystem initiating the message. Although expensive, the technology for locating mobile subsystems with self–contained equipment is currently available and prices are expected to continue to fall.

A requirement is placed on the communications service provider to determine the caller location (as is done on Emergency Telecommunication systems today for 9-1-1 calls from wireline connected telephones). Unfortunately, the technology to locate callers using wireless service providers is still in a developing stage, and it is difficult to predict if this technology will mature to a useable level for this service over the next 20-years.

5.2.2 Emergency Notification Data Message Routing

Wireless data emergency notification messages must be efficiently routed to the EM that has jurisdiction at the location where an emergency is taking place. This is particularly challenging for WAN wireless communications. Two mechanisms have been considered, each with its own advantages and disadvantages.

1. Communication Service Provider

   The mobile subsystem always sends emergency messages (which include location information) to a standard IP address. This special Emergency Management address is detected by the CSP. The CSP will read the location in the standard *Emergency Request* message, and based on the location will forward the message to the appropriate Emergency Management center.

   This mechanism requires the CSPs to implement this standard.

2. Emergency Management Subsystem (EM)

   Emergency Request messages will be routed to an EM pre-selected by the traveler. The EM will interpret the location and other details about the emergency and forward the message to the appropriate EM.

   Users of this service will be required to subscribe to an ISP. Of course, the ISP could be provided for this service to all ITS wireless mobile users at public expense (or at the expense of wireless service subscribers through a "tax" as is currently done to support 9-1-1 service for cellular phone subscribers in some states).

The architecture supports the EM method of emergency data message routing.

	Communication Service Provider (CSP)	**Emergency Management Subsystem**
Where Special Emergency Messages are Interpreted (type/location) and then forwarded.	Communications Layer	Emergency Management center Subsystem
Requires New Communications Standards	Yes.	No. Incorporated into ITS standards.
How Funded	Wireless data service "tax" (similar to AMPS 9-1-1 in some states)?	Service to subscribers for a fee (or tax if publicly operated).
Current Plans / Deployments	None	Yes, Ford in 1996 (RESCU) with EMC in Texas for full US coverage.
Emergency Message Security	Good. Communication provider to interpret message.	Excellent. Only private (although could be public) EMC needs to interpret messages.

Table 11. Comparison of Emergency Request Routing Mechanisms

5.3 Map Attribute Referencing for Communication

A common method of referencing transportation links and nodes and their attributes is essential for many of the ITS services involving cooperative processing between ITS subsystems. For example, adjacent TMSs may share data regarding congestion attributes of links (current and predicted occupancy, speed or incident locations), ISPs may request current and predicted congestion attributes of links, ramps, or intersections from TMSs, and ISPs will send sequences of links and nodes (i.e., a route) with specified attributes (e.g., signage, expected conditions, waypoint locations to report probe information) to their mobile clients.

A common frame of reference is needed so that these communications between subsystems can be rationally reduced to an unambiguous reference to the same transportation links, ramps, and intersections. Two competing general approaches have been proposed: link-IDs (a national topological network) and coordinates (a national coordinate system using latitude, longitude and elevation). Both have their advantages and disadvantages, summarized in Table 12. The architecture currently supports both types of representation. For interoperability, all deployments should support the coordinate system, and optionally support the link-ID system.

	National Coordinate System based on DGPS	**National Link-ID Database based on the NHPN**
Rationalization with DGPS coordinates and Vendor Map Databases.	Provided by each map database vendor for links in their products.	Provided by map database vendors for their products.
Pre-existing?	Most map database vendors use latitude, longitude and optionally elevation.	Partial. Could be built upon the National Highway Planning Network (NHPN).
Compact Representation for Efficient Communication	No. Transmission of latitude, longitude, altitude to desired accuracy, for each point.	Yes. Transmission of link-IDs, node IDs. IDs used to look-up referenced link or node at receiving subsystem.

Table 12. Comparison of Link-ID and Coordinate Map Data Transfer Methods

The National Highway Planning Network (NHPN), a geographically-based analytic network representation of major highways in the United States jointly sponsored by the Federal Highway Administration and the U.S. Army Forces Command referred to in Table 12, has the potential to be an important basis for a future national transportation network attribute message standard for ITS. The NHPN currently covers the National Highway System, and thus needs a mechanism for extension to cover the remaining large number of links and nodes.

One possibly preferred mechanism to extend the NHPN is to use the proprietary map databases available from various vendors. By using the NHPN as an open frame-of-reference, subsystem designers can use whichever vendor map database serves their needs best; and as long as the vendor is able to rationalize their product with the NHPN database, then they can still exchange location information with other ITS subsystem that may use the NHPN with other proprietary map databases.

The NHPN network is composed of lists of nodes and links, nodes being points representing junctions or traffic generators and links representing sections of roadway between them. Each node's entry contains its latitude and longitude (elevation is currently not included) and, frequently, a nearby place name. Each link's entry contains its endpoint nodes, attribute information about the road (its sign route number or name, length, administrative and functional class, access control, number of lanes, and various other

characteristics important for simulating traffic operations and performance), and a digitized chain of latitude / longitude coordinates depicting the road's alignment.

The NHPN has the following characteristics (from "White Paper: *Location Referencing for ITS*", Cecil W. H. Goodwin, University of Tennessee, cecil_goodwin@ece.engr.utk.edu, 11/8/95):

- *The NHPN, and systems to maintain it are already in existence, which greatly reduces costs for generation of the ITS Datum and time to deployment.*
- *The NHPN contains the most important roads for inter-regional travel, and in many cases those roads most susceptible to congestion for commutes in urban areas; thus it supports immediate product development.*
- *The NHPN is public domain.*
- *The NHPN contains geographic coordinates of nodes in all urban areas at a sufficient density to support coordinate-based referencing.*
- *The NHPN contains network topology and attributes across the country sufficient to serve as the datum for linear and link-based referencing.*
- *The NHPN dataset is small enough for universal use and practical maintenance as a datum.*

If the NHPN is adopted as the basis for an emerging national link-ID standard, it could be in the interest of an industry group to develop and evolve the NHPN without federal government support. NHPN nodes are currently accurate to 80-meters, but could be improved to a few meters by a national DGPS survey (as suggested by Goodwin). Version 2 release 0 of the NHPN, dated September 15, 1994, is available on a CD-ROM, prepared by the Federal Highway Administration, Office of Environment and Planning (HEP-11), Washington DC. This database is currently updated annually (Version 2.2 is available from the USDOT website). (The NHPN may also be available from The National Energy Software Center, 9700 S. Cass Ave., Argonne, IL 60439. Telephone: 708-252-7250).

A useful organization of the NHPN data is by node (as pointed out by Goodwin) where for each node exists:

- Node ID
- node latitude
- node longitude
- node elevation (altitude)
- a list of IDs of nodes that can be navigated to from this node with associated ground distances and street names.

This organization enables using the NHPN for navigation on its own, but also allows relatively easy use as a common frame of reference for registration of proprietary vendor databases to the NHPN nodes (or NHPN *datum*s). Note that the street names may enable registration augmented by map matching, since the accuracy of the NHPN and/or the vendor map databases may not be sufficient to alone guarantee unambiguous registration.

A proposed ("strawman") mechanism for standard link and node referencing in ITS communications has been proposed by Cecil Goodwin after some communications with the ITS Architecture development team and various stakeholders in the ITS community. The Location Reference Messaging Specification (LRMS) is reproduced here, in Table 13.

The LRMS can be used to identify links in terms of link end node IDs, arbitrary points along a link defined in this way, arbitrary points with respect to a node, and arbitrary links with respect to a node. The

nodes can be NHPN datums OR could be defined in a special node attribute record. The advantage of the node attribute record is that a subsystem (e.g. a vehicle) may not carry any map database at all, but can receive from an infrastructure subsystem (e.g. an ISP) the location definitions of a limited subset of nodes relevant to a particular trip. These nodes can then be used as the basis for defining a set of links that the traveler should be guided by in-vehicle equipment to follow. In this way, the mobile subsystem does not need to carry and maintain a map database, but receives the essential map data necessary for a selected trip just prior to selecting a route. The nodes defined with IDs in this way can be NHPN nodes (the *National link-ID* approach) or could be nodes specifically defined for the travelers route using Global Positioning System (GPS) or Differential GPS (DGPS) coordinates (the *National coordinates* approach).

Finally, the usefulness and generality of the LRMS approach is not limited to ISP-Vehicle (or more generally infrastructure-mobile) interfaces, but can be used between fixed centers as well e.g. TMS-TMS, TMS-ISP etc.

Bits	**Content**	**Values/Range**
0	Pad	
1-3	Type:	000=ITS Datum link
		001=ITS Datum link with Offsets
		010=Georeferenced Point
		011=Georeferenced Link
		100=Node Attribute Record
		101-111=expansion

Case Type=ITS Datum link (6 byte record)

Bits	**Content**	**Values/Range**
4-23	Start node ID	1-1024K
24-27	Pad	
28-47	End node ID	1-1024K

Case Type=Point along ITS Datum link (12 byte record)

Bits	**Content**	**Values/Range**
4-23	Start node ID	1-1024K
24-27	Pad	
28-47	End node ID	1-1024K
48	Pad	
49-71	Offset1 from Start Node	8192K decimeters
72	Side (optional)	Binary 0=right-hand; 1=left-hand
73-95	Offset2 (optional)	8192K decimeters

Case Type=Georeferenced Point (15 byte record)

Bits	**Content**	**Values/Range**
4-23	Node ID	1-1024K
24-43	Delta X to point	+/- 512K decimeters
44-63	Delta Y to point	+/- 512K decimeters
64-79	Delta Z to point	+/- 32K decimeters
80-119	Street Name	First 5 characters of name

Case Type=Georeferenced link (22 byte record)

Bits	**Content**	**Values/Range**
4-23	Node ID	1-1024K
24-43	Delta X to start point	+/- 512K decimeters
44-63	Delta Y to start point	+/- 512K decimeters
64-79	Delta Z to start point	+/- 512K decimeters
80-99	Delta X to end point	+/- 512K decimeters

100-119	Delta Y to end point	+/- 512K decimeters
120-135	Delta Z to end point	+/- 512K decimeters
136-175	Street Name	First 5 characters of name

Case Type=node attribute (14 byte record)

Bits	**Content**	**Values/Range**
4-23	Node ID	1-1024K
24-55	Node Longitude	+/- 180,000,000 microdegrees (per ISO 6709)
56-87	Node Latitude	+/- 90,000,000 microdegrees (per ISO 6709)
88-91	Pad	
92-111	Node Altitude	+/- 512K decimeters

Table 13. Strawman ITS Location Reference Messaging Protocol (LRMS)

5.4 Mobile Subsystem Location Systems

Mobile location mechanisms can be split into four broad categories. They are:

1. *Mobile based*, e.g. combinations of GPS, compass, inertial guidance, odometry, map matching.
2. *Mobile based with WAN assistance*, e.g., DGPS, communication service provider assisted location determination.
3. *Mobile based with roadway assistance* by reading some encoding deployed on the transportation infrastructure, e.g. beacon interaction or machine readable symbology on the road surface to label the nodes.
4. *Infrastructure based*, e.g. triangulation by cellular base stations.

Note that basic GPS is considered "Mobile Stand-alone" since we assume that the satellite based GPS signals will always be present, independent of ITS. The alternative categories are compared in Table 14.

The architecture supports Mobile Based Stand-Alone, Mobile Based WAN Assisted and Infrastructure based deployments.

	Mobile Based: Stand-alone	**Mobile Based: WAN Assisted**	**Mobile Based: Roadway Assisted**	**WAN Based**
Example	GPS, inertial, magnetic, odometry, map-matching.	DGPS	Beacons, location encoding on the roadway.	Cellular provider location determination.
Accuracy	Good	Excellent. Degraded accuracy when out of WAN coverage, e.g. as with DGPS.	Good-Excellent	Poor-Good? Currently experimental.

Reliability	High. If GPS alone is considered, augmentation is necessary for *urban canyon* effects.	High.	High.	Medium? Currently experimental.
Cost Vehicle	Low	Medium	Low	Medium.
Cost Infrastructure	N/A	Low	Medium	Medium.
Institutional Concerns	None.	Low. Provision of Differential Correction data possible from ISP, Communication Service Provider or special public or private service provider.	High. Jurisdictions must cooperate to deploy a standard beacon or symbology on the roadway. If publicly funded, equity concern and possible public liability concern.	High. Cellular providers must agree on standards for this service (which will go beyond FCC rules for mobile 9-1-1 call origination location).

Table 14. Comparison of Mobile Location Determination Approaches

5.5 Integrated Traffic Management, Demand Management and Route Selection

This section discusses key subsystem operations and interactions for the evolution of traffic management, demand management and autonomous and infrastructure assisted dynamic route selection.

Figure 37 shows the high level interactions between traffic management, demand management, and dynamic route selection. Key interactions will be discussed in the following sections. Note that Figure 37 does not show all the detail, but rather key and representative examples (e.g., "weather data" input to the predictive model is not shown). A few observations about the groupings of functions:

- Mobile subsystems interact only with the ISP for navigation, not the TMS.

  This was done in response to public agency personnel concerns to providing "personalized" information, believing that except for public safety vehicles, this was the role for the private sector. Note that the ISP may be operated by a public or private sector entity.

  Also, travelers were found to feel more comfortable getting personal guidance information from a private company than from a government agency (for privacy reasons).

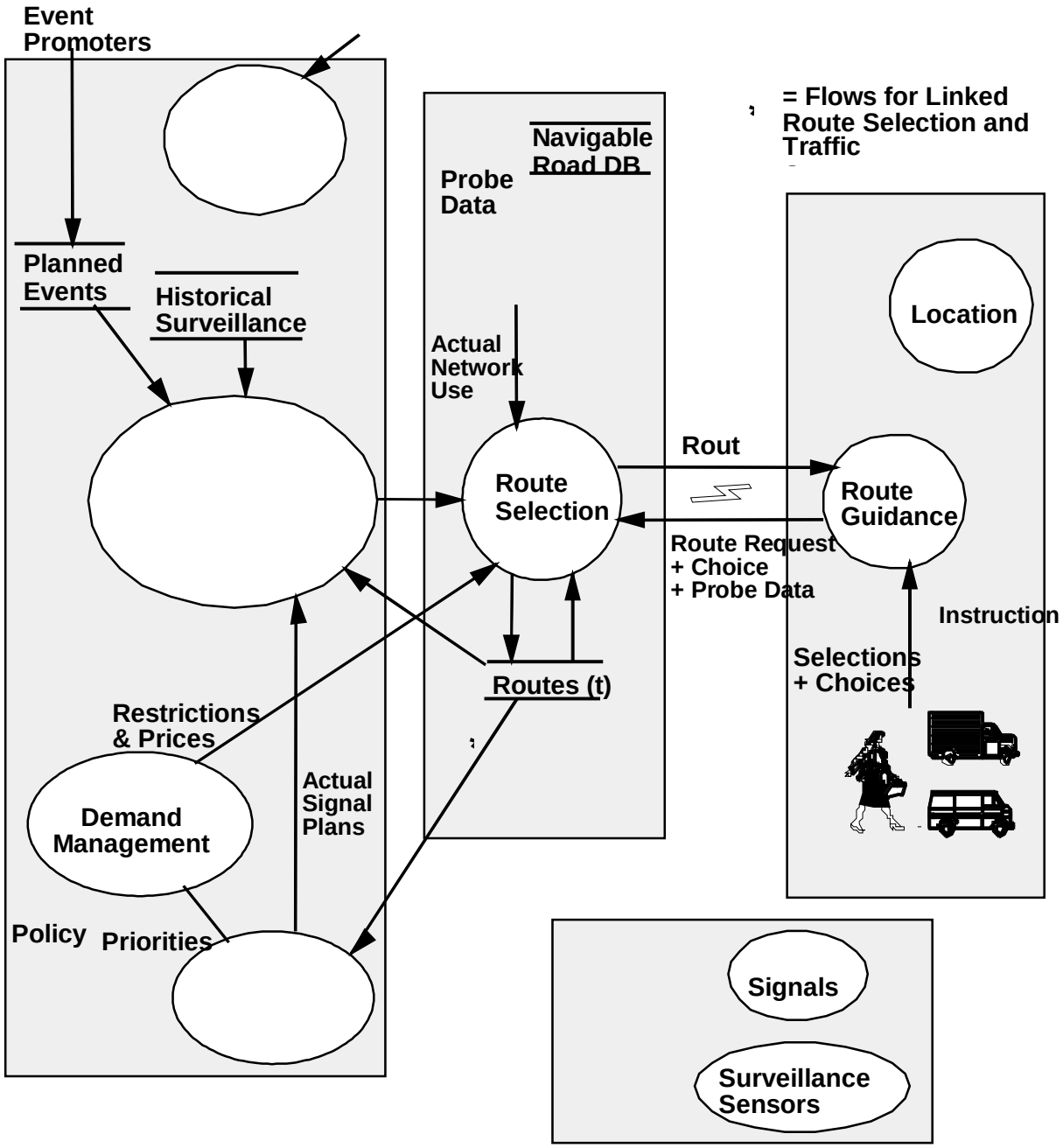

Figure 38. High-End State Traffic Management, Demand Management and Dynamic Route Selection

- The ISP and TMS potentially exchange considerable data

  TMS to ISP: Predictive Model of link, ramp, and intersection traffic conditions; Actual Network Use surveillance and link restrictions and pricing.

  ISP to TMS: Probe data (vehicle types but not identities); vehicle routes (showing vehicle type but not identity).

- Roadway surveillance data is used for both ISP/ATIS functions as well as TMS/ATMS functions. Revenue flows are not shown.
- TMS and Roadway subsystems are generally deployed to cover all the transportation infrastructure in a non-overlapping way. ISP subsystems may have considerable overlap in the markets that they serve. Mobile subsystems will usually be interacting with only a single ISP at a time.

5.5.1 Probe Data Reporting to ISPs and ISP Updates to TMSs

As shown in Figure 37, mobile travelers submit a route request to the TMS, receive one or more route options, and choose a route. Note that for liability reasons, it is always important that the travelers have the final choice of a route, because they have primary responsibility for their safety while traveling. The selected route is stored at the ISP for the duration of the trip, and is also stored at the mobile Route Guidance process, which provides instructions to the traveler. Of course the travelers can at any time request an updated route or change their route request.

As the traveler progresses, the Route Guidance process will at ISP-determined waypoints, provide probe data to the ISP. Potentially, on receipt of each probe data, the ISP recomputes the travelers best route, and if an alternative route is better, (due perhaps to a non-recurring incident) then the better route can be offered to the traveler.

The probe data, without the driver identity, is also sent to the surveillance collection process at the TMS, where it is used to estimate congestion parameters on links that may not be fully instrumented with roadway surveillance sensors.

5.5.2 TMS Predictive Model and Open ISP Access/Updates to that Model

The route that is computed at the ISP is based in part on a predictive model stored at the TMS. This model can be used by the TMS for traffic management, as well as being provided (possibly for a charge) to ISPs. The model is based on statistical occupancy of links. The occupancy is based on historical surveillance, as well as actual expected occupancy provided by route inputs from the ISPs. In this way, many ISPs can use the same predictive model, and as travelers select routes, these selected routes are sent to the TMS to incrementally update the predictive model, thus allowing a balanced allocation of travelers to the transportation links, and avoiding overcongestion of any one link when better alternatives are available.

The prediction of link delays (the time to transit a link) and ramp or intersection queue delays (the time to transit a highway on-ramp, off-ramp or an intersection based on a desired turning movement) can be of varying levels of sophistication. In a sophisticated deployment, the link and queue times may be based on the expected statistical occupancy of links, and models based on historical data of the relationship between occupancy and expected (average) link times and queue delays. The expected statistical occupancy of links may be determined by historical time-of-day data, as well as the prior choices of travelers to travel specific routes.

Although not shown here, the expected occupancy of transit vehicles may be updated in real-time through a similar message from the ISP to the Transit Management subsystem.

5.5.2.1 TMS-ISP Public-Private Partnerships

The predictive model process that resides in the TMS (in Figure 37) can be either publicly or privately operated. In cases where this model is maintained by a public agency, data can be exchanged between the

TMS and one or more ISPs to both maintain the model and use the model to compute optimum routes for the ISP clients.

5.5.2.2 TMS-ISP Private-Private Partnerships

In cases where there is no TMS, or the TMS has opted not to participate, then each ISP can either build their own predictive model (i.e. aggregate their ISP subsystem with a TMS subsystem that only has a Predictive Model Equipment Package) or they can join with other ISPs to create a stand alone TMS that also only has a Predictive Model Equipment Package. In this later case, the ISPs may join into a "competitive joint venture" in the ownership, operation, and maintenance of this Predictive Model TMS. The point of this venture is to improve the routing that they are able to provide to their clients by using a common model for the links that their clients are sharing. In this case where a public agency has chosen to opt out of the Predictive Model function, it will be more difficult to result in a tightly integrated ATIS/ATMS architecture in the high end state of ITS deployment.

5.5.3 TMS Demand Management

The TMS executes demand management policy in two ways: through priority signal coordination and by issuing restrictions and road pricing.

5.5.3.1 TMS Prioritized Routing by Vehicle Class

As shown in Figure 37, the routes of vehicles participating in route selection are send from the ISP to the TMS Signal Coordination process. When possible, priority will be given to vehicles based on the Demand Management Policy. For example, Emergency vehicles (e.g., fire trucks and ambulances) may be given the highest priority, then Transit Vehicles, HOV vehicles, etc. The actual signal plans are fed to the predictive model, so that changes in the plans can be used to figure the expected link-times and ramp and intersection queue delays of the model. At the same time, the predictive model is used by the Signal Coordination process to set the signal plans. Clearly, in the 20-year timeframe, these processes will be very closely coupled.

5.5.3.2 Demand Management by Restrictions and Pricing

The Demand Management process also implements demand management policy as it pertains to lane restriction and prices. These are communicated to all travelers through signage in the traditional way (e.g., (dynamic message) signs indicating HOV-n lanes during certain hours of operation). In addition, the restrictions are communicated to the ISP Route Selection process, so that these restrictions and prices can be taken into account when processing the route requests from their clients.

5.5.4 Route Selection

As shown in Figure 38, the architecture currently supports a continuum of modes of route selection from simple in-vehicle "Autonomous Route Guidance" to fully integrated ATMS-ATIS route selection "Integrated TM/Route Guidance". Issues associated with route guidance modes are summarized in Table 15.

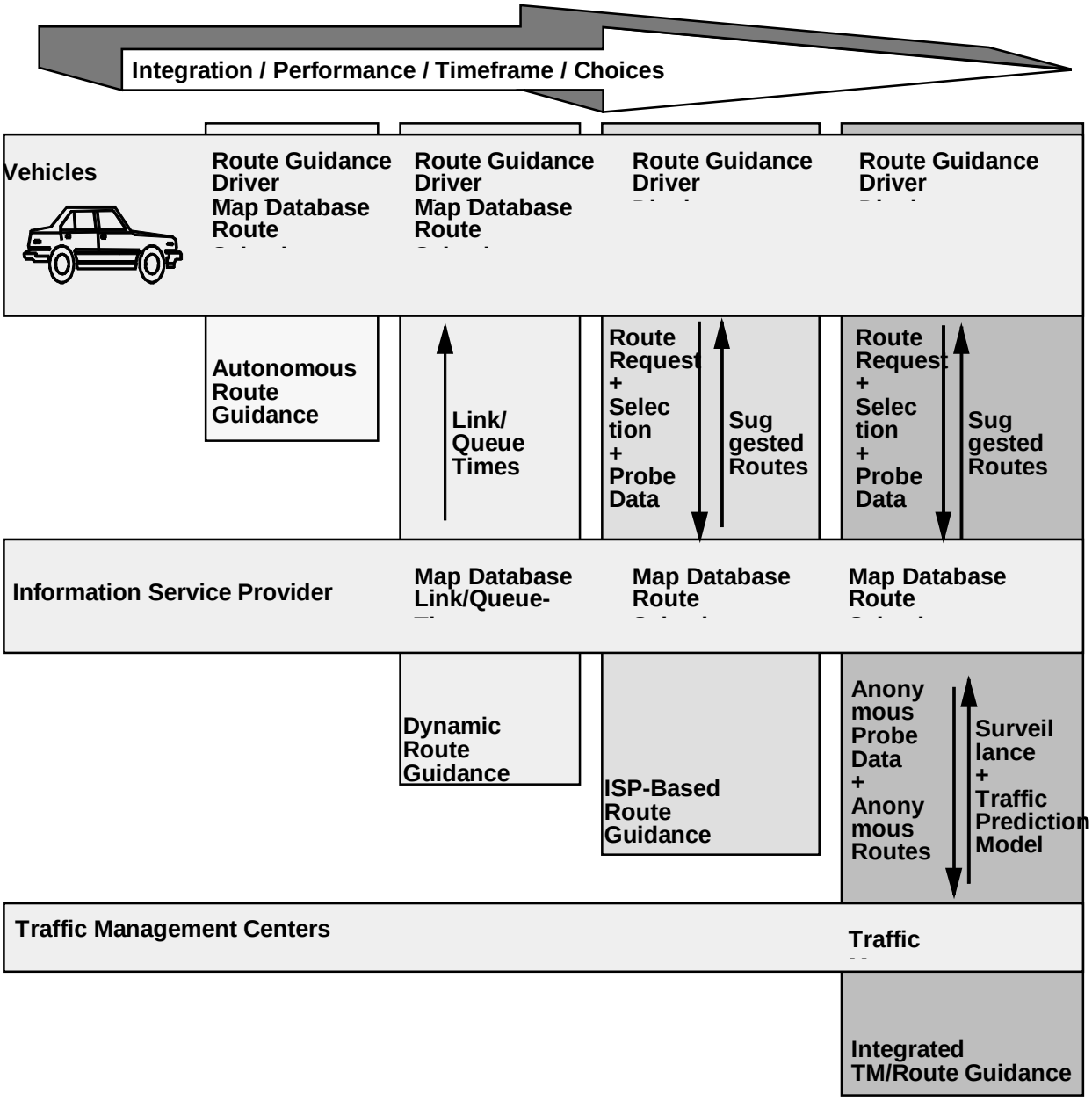

Figure 39. Route Selection Alternatives

By supporting a continuum of Route selection modes, the architecture allows the market to decide which will be dominant (if any), at the risk of diluting investment (and speed of development) as well as risking the advanced traffic and demand management benefits of the fully integrated ISP-TMS based solution (Integrated TM/Route Guidance) shown in Figure 37.

The autonomous route guidance, or dynamic route guidance are fairly well understood, and the former is currently operational. The two route selection approaches shown on the right side of Figure 38, where the route computation process is in the ISP, and the vehicle has no map database, are a little more complex to visualize. A number of questions may be posed about these methods:

- How is the trip begun?

After requesting and receiving a trip plan, the driver can proceed. In some cases, the driver could proceed before receiving a trip plan where there is little risk of deviating from the plan (or the driver is willing to accept the risk of making a mistake and needing to get a corrective trip plan).

- What happens if the driver gets lost (deviates from the route)?

  Either the in-vehicle navigation portion of the process can detect that the traveler is "off-course" and request a new plan from the ISP based on current position, or the ISP can detect that the vehicle is off course based on a "probe message" from the vehicle and process a new trip plan.

- What happens if the ISP learns of a major incident along the route?

  The ISP specifies locations for the vehicle to report probe data along the route. This is sent to the vehicle as part of the route plan. Based on the last probe transmission from the vehicle, and the stored route plan that the ISP has for the vehicle, the ISP can interpolate the approximate position of the vehicle. When an incident is detected by the ISP, it can recompute the best routes for all its en-route clients, and send an advisory message to all its affected clients and send a new route to those clients that are both affected by the incident and have an alternative route.

- How is the trip ended?

  The last probe reporting point should be at (or near) the destination. This would "close" the trip plan from the ISPs perspective.

- What are the requirements on the vehicle location function to specify the origin and to sense when the vehicle has deviated from the route?

  The vehicle and ISP must be able to locate the origin with enough specificity to determine where on the transportation network the vehicle is located. The architecture does not need to specify a particular technology, but, for example, GPS at the vehicle, and differentially corrected at the ISP, should be more than adequate (<10meters accuracy). GPS alone with simple dead-reckoning augmentation and map-matching (route matching) has been demonstrated to be adequate under most conditions for route following. Other location technologies should not be precluded and may in fact be superior in either cost, performance, or both.

- What if the driver starts after submitting a route request but before receiving a route plan? What are the requirements on end-to-end latency?

  The longer the latency, the higher the risk of leaving before receiving a route and departing from the route before receiving the route! Of course, a patient driver will not have this problem. How patient the market will be is undetermined. Also undetermined is the actual network delay for enough of the route plan message to get started. Experience today using the "World Wide Web" on the Internet may indicate that network delays of a few seconds are likely, but that several minutes are unlikely. Of course, using value added network services can result in considerably less delay (at a cost). The market will ultimately decide if this approach is feasible.

- What information must be included in the ISP route message (more than link IDs)?

  As discussed in Section 5.3., links and route segments (which may make up a multiplicity of links) are defined by nodes that can be defined with respect to NHPN coordinates or defined in

terms of arbitrary latitude, longitude and elevation. Since the proposed LRMS mechanism allows the nodes to themselves be defined to the vehicle by the ISP, there is no NHPN or other map database requirement for the vehicle.

- Would the route have to be tied to a precise coordinate/direction system?

  No. As stated in the previous system, the ISP can define the nodes and then use the defined nodes to specify a sequence of route segments that the vehicle should navigate based on latitude, longitude and elevation coordinates.

- In the high end ISP architecture (right most deployment in Figure 38), the TMS and TRMS are providing information to the ISP. Why would they do this?

  For different reasons in different regions:

  1). They have a common objective: the TMS to reduce congestion by allocating vehicles out of the most congested links to less congested links, and the TRMS to get information out about their services so that they can serve more customers. Everyone benefits by reducing the uncertainty of travel, and thus raising the value of travel.

  2). If the ISP is generating revenues from the distribution of their information products, they may be willing to pay the ISP and/or TRMS for information that allows them to distribute products that their clients want.

  3). If TMSs are non-cooperative (or don't exist), then the ISPs may find a way to get the surveillance data that they need on their own. For example, ISPs can individually collect surveillance data using client probe data, and competing ISPs can pool their probe data in a "TMS" competitive joint venture (see Section 5.5.2.2.).

A discussion of the features of the different route guidance methods is given in Table 15.

	Autonomous Route Guidance	**Dynamic Route Guidance**	**ISP Based Route Guidance**	**Integrated TM/Route Guidance**
Current Deployments	Yes	Field Operational Tests	Field Operational Tests	Field Operational Tests
Wireless Communication		One-way Broadcast	Two-way	Two-way
Privacy	Absolute	Absolute	Probe data to ISP	Probe data to ISP
Real-Time Advisories		General	Detailed	Detailed
Traveler Fees		Yes (unless broadcasts publicly funded).	Yes	Yes
Integrated ITS and non-ITS personalized services			Yes	Yes
Mobile Map Database	Yes	Yes		

TMS benefits				Fixed surveillance data for revenue; Probe data for ATMS.
Traveler Paid Revenue Spread		Private or public information broadcasters	ISP, Communication Service Provider	ISP, TMS, Communication Service Provider
Advertising Revenue			Yes	Yes
Efficient Use of Transportation Network with High Market Penetration	Poor: no awareness of incidents.	Good: unstable allocation of alternative routes	Very Good: Depending on market penetration of ISP	Excellent
Integration with Traffic Management and Demand Management				Yes: Vehicle class based signal priority and statistical link occupancy management.

Table 15. Route Selection Choices and Features

5.5.5 ISP Operation: Public, Private and Public-Private

It is anticipated that by the 20-year timeframe, ISPs serving private sector travelers and vehicles will be primarily privately operated. In the near term, a mixture of public, private and public-private partnership operated ISPs are expected to serve this market, especially if there is a near term focus on providing intermodal ISP services using a mixture of local public sector, federal public sector and private sector funds for deployment (e.g. the "Intelligent Transportation Infrastructure" program proposed by the US DOT).

The ISP can be an example of a client-server architecture, for-profit service. The public currently does not have much experience with this model, and there are many questions and opinions as to the viability of this business model. A few examples taken from today's "World Wide Web" may provide some insight.

- The World Wide Web itself: Paradigm shift or passing fad?

The World Wide Web (WWW), based on open standards (e.g. "html-hypertext markup language)" and TCP/IP message transport is currently enjoying tremendous growth. It is estimated that 30 million people in the US access the WWW (these numbers are difficult to ascertain). Similarly, the number of web sites to choose from is about 5 million world wide and has been growing by about a factor of 2 each year through the 1990s. Although security, reliability and performance are real issues on the WWW today, solutions might be expected. A fantastic feature of the web is still the potential for broad connectivity and access from client (user) to server (web server) that is achieved. The use of "value added networks" for a fee solve the security, reliability, and performance problems; but at a higher cost and a limitation on the content (web sites) accessible (you can only access content on the value-added network). Many value added network providers offer their clients access to the WWW as well.

In addition to the problems just mentioned, the array of content on the web is difficult to access for some. Others have learned to use freely accessible search tools (web sites themselves) to sort through and find content of interest and value. The WWW, and the underlying Internet, have already become a basic and essential business tool for many individuals.

Finally, there is not a consensus that after the initial fascination, individuals will either move away from the WWW because of its technical problems or because the value is not perceived, or if they will become regular users. Different individuals often take polarized views on this question, and only time will tell for sure.

- Web site business models

Many web sites have appeared at some obvious cost to the web owners. How will these web sites survive financially? One model that has been tried with some past success are subscription content provider networks e.g. AOL, Compuserve and Prodigy. These operators have either created content themselves or purchased content. Many of these content provider services (who also each have their own communication networks) are rapidly moving to provide WWW only access services in response to the market.

Other web sites are valued for the advertising they provide directly to the sponsor (e.g. "corporate web sites"), or for the revenue the owner can receive by selling space on their sites for advertisements. The advertising rates are presumably related to the number of accesses made to the site.

Finally, some web sites appear to require subscriptions to access their services, requiring user-IDs and passwords to access content, with subsequent billing (either based on usage or a flat rate) for services. The success of failure of this model is still to be determined.

- Case study: "Easy Sabre" airline flight schedule service.

An example of a very popular transportation service is commercial airline flight scheduling. American Airlines has the service "Easy Sabre" that provides a flight scheduling information or ticket purchase service on-line. Users can specify origin, destination, travel time, airline, and cost preferences and be offered in response the available flights and flight combinations. This service is not currently available over the WWW, but is available on most subscription networks.

- Will ISPs need to contract or partner with certain CSPs?

Based on the discussion above, it is difficult to say. ISPs may be able to charge for service to a mass of customers using low cost service providers for access to a common internetworked communication medium. Alternatively, a CSP partnership may give them immediate access to an established CSP client base. The biggest unknown here is how the wireless communications industry will evolve. Will the new deregulation and competition in wireless service providers cause rates to drop quickly and access to grow exponentially? This is certainly difficult to predict.

There is a spectrum of views as to if and when a sufficient market will emerge to support a competitive privately operated ATIS market. Factors influencing these views exist on both the demand side and the supply side. On the demand side, how much are people willing to spend for various types and quality levels of ATIS services. On the supply side, how soon will the cost of computing and communications technologies fall to reach a point where ATIS services can be supplied at a price people will pay. "Traffic Master" is an example of a relatively advanced private sector ATIS system that appears financially viable which has been deployed in the UK and for which licenses have been purchased to operate in the US. This proprietary (non-open) system uses a wireless paging network to distribute link congestion information to mobile clients. A viable market for this or similar types of ATIS deployments in the US is still unproven. A comparison of the issues associated with public vs. private ISP operation is summarized in Table 16.

Operation of ISPs for Emergency and some Transit Vehicle fleets will be predominantly by public agencies in the near to mid term. By the 20-year timeframe these ISP services might be "outsourced" to the private sector if and when the private sector is able to provide the level of service and reliability that these fleets, especially emergency vehicles, require.

5.5.6 Advertising as a Revenue Source for ATIS

As was discussed in Section 4.1.2.2., an alternative approach to providing In-Vehicle Signage from that discussed in the ITS-America Program Plan is included in the ITS architecture. Signage information is added to route messages and is communicated to the traveler using the Wireless WAN Communications. The traditional approach to in-vehicle signage, also supported in the architecture, is to deploy roadside beacons at or near sign locations, and in vehicle equipment receives and repeats the beacon messages.

A key feature of the ISP based in-vehicle signage mechanism is that it may provide a key revenue source for competitive ISPs that would encourage them to deploy ITS ATIS services at a lower price to travelers. This is because they would have the opportunity to get a revenue stream from advertisers based on market size. In this model, customized advertising messages are included with ATIS information. Table 17 summarizes the features of the two mechanisms for the in-vehicle signage user service.

5.6 Electronic Payment Operations

Payment mechanisms were chosen to be consistent with existing and emerging deployments for electronic fare, toll, and parking payment systems, as well as interoperability with the existing financial network for approval and clearance of electronic payments. Consideration is given to always allow a cash payment capability, with payment by financial instrument (e.g., credit, debit card or cash card) as an option.

5.6.1 Privacy and Traceability Impacts of Payment Instrument Choice

The inherent characteristic of credit/debit cards and cash cards (or negotiations directly in currency) have impacts on privacy and traceability in ITS.

With a debit or credit card, users can expect to get an accounting on a periodic basis of where and when charges to their accounts are made. With cash, or with cash cards, there is no such accounting or traceability possible (outside the records that may be stored on the card itself).

Privacy, on the other hand, is at some risk with credit and debit cards, since the financial institution servicing the card usually keeps a detailed record of each of the users financial transaction in support of traceability. No such record exist (outside of the card itself) with cash cards.

	Public Operated ISP	**Private Operated ISP**
Geographic Scope	Determined by jurisdictions (or groups of jurisdictions)	Determined by markets, ISP niches.
Institutional Dependencies	Requires cooperation of jurisdictions for an area to be covered.	Requires sufficient market to attract investment to cover an area.
Partnership Agreements	None.	Public-Private for integrated ATIS-ATMS, although ATIS Private-Private would be possible where TMSs don't cooperate or don't exist.
Crossover Investment	If private ISPs are operational, reluctance to deploy public funds to compete.	If public ISPs are operational, reluctance to deploy private funds to compete.

Impact		
Support of Geographic Roaming	Requires multiple jurisdictions deploying ISPs to have a national or regional standard interface for mobile subsystems.	Requires mobile subsystems to comply with one interface chosen by the ISP that the traveler chooses.
Beacon DSRC (also see Section 5.1.3.)	Direct access to right-of-way, but many jurisdictions must cooperate. Some user services not supported (e.g., emergency service request).	Must negotiate access to right-or-ways with many jurisdictions. Some user services not supported (e.g., Mayday).
Wireless RF WAN DSRC	Competitive, internetworked, open interfaced, broad coverage options.	Competitive, internetworked, open interfaced, broad coverage options.
One-way Broadcast	Equitable mechanism for funding is clear.	Mechanism for billing, personalized advertising is not clear.
Maintenance Costs and Skills	Some jurisdictions concerned about cost and skills to maintain a sophisticated DSRC system.	
Liability	Some jurisdictions concerned about public liability associated with operation of an ATIS system.	Better mechanisms available to private companies for managing liability (e.g., incorporation, bankruptcy).
Privacy	Some individuals reluctant to use some ATIS user services requiring personal information from public agencies.	Some individuals reluctant to use some ATIS user services requiring personal information from private companies.
ISP Competition	Generally, there is only one agency to provide ATIS services.	Competition possible where the market will support.
Technology Deployment	Driven by public expenditure priorities.	Driven by competition.
Adherence to Standards	Likely if stimulated by federal funds.	Market driven.
Early Deployment	Likely if stimulated by federal funds.	Market driven.
Public Safety Vehicle Support	Yes.	May require legislation to guarantee adequate access and response.

Table 16. Implications of Public and Private ISP Operations

	Beacon Based	**Wide Area Communications Based**
Liability	Public Sector	Private Sector
Infrastructure	Public Right-Of-Way	WAN Communications
Advertising Revenue	Beacon “Billboards”	Route Plan “Billboards”
Equity	Advertising possible where beacons deployed.	Advertising revenue can lower the cost of personalized ATIS services in a competitive market and therefore increase the accessibility of these services and

		accelerate the deployment of ITS.
Operations and Maintenance Funding Sector	Public or Public-Private Partnership	Private
Needs in vehicle Location Capability	No	Yes.
In-vehicle equipment cost	Low	Medium. Marginal cost is very low if already have in-vehicle navigation.
Public Sector Infrastructure cost	High	None.

Table 17. Comparison of In-Vehicle Signage Techniques

This inherent choice/tradeoff of traceability for privacy is left to the user in the ITS architecture, since both classes of Payment Instrument are supported.

Today, phone calls can be placed from payphones using "cash" with total anonymity. In principle, a "cash card" could be used for 2-way wireless data communications with the same privacy/traceability tradeoff features.

5.6.2 Secure Electronic Payment Over the NII

The current architecture is designed with the intention of serving as a starting point for eventual standards on secure electronic payment that will be developed independent of ITS as a part of the Communications layer of the Physical Architecture by Communication Service Providers.

Protocols can be implemented for credit card and debit card transactions between customers and service providers using the existing financial network for approval and clearance. The protocols are based on public-key cryptography and can be implemented in either software or hardware. Increasing security of these transactions are monotonically related to key management complexity. Deployment can be gradual and incremental.

5.7 Commercial Vehicle Operations

5.7.1 Commercial Vehicle Check

Three mechanisms have been under consideration by the architecture team for commercial vehicle-roadside communication and identification. Each scheme uses a tag in the vehicle that is readable by the roadside equipment. They differ in how much data is stored on the tag, whether the data on the tag may be programmed by the driver or Commercial Vehicle Check roadway subsystem, and the impact on the infrastructure. The three schemes are summarized as follows:

1. Serial number only scheme.

   This is the simples/cheapest scheme per vehicle but puts a burden on the driver to re-enroll the vehicle any time there is a change in carrier or driver. This will generally not be a burden for small owner-operators. In addition, there is a burden on the infrastructure to maintain a current database of "problem" drivers, carriers and vehicles at each roadside Commercial Vehicle Check subsystem.

2. Three number scheme: driver, vehicle, and carrier.

This scheme avoids the re-enrollment problem by allowing a simple tag reprogramming for the three identification numbers.

3. Fully programmable credentials on the tag scheme: full credential information for driver, vehicle, and carrier.

   This tag holds considerably more data and can be written to by the infrastructure. Because all the data for a vehicle, driver and carrier is stored on the tag, the database maintenance and distribution task for the infrastructure is substantially reduced. Data integrity on the tag is a technological concern.

The three approaches are summarized in Table 18. The architecture supports the second scheme with an evolution to the third scheme.

	Serial Number Only	**Driver, Vehicle and Carrier Numbers**	**Fully Programmable Credentials and/or Safety Inspection Data from the Tag**
Per Vehicle Cost	Low	Medium	High
Re-enrollment:	Any change in driver or carrier on a vehicle	Reprogram new Driver or Carrier number on vehicle tag	Reprogram new Driver or Carrier number on vehicle tag
Cost of Roadside Commercial Vehicle Inspection Station Database	High	High	Low
Favored by:	Small Owner-Operators	Large Fleet Operators	Regulators

Table 18. Impact of Different CVO Tag Storage Levels

5.7.2 Automated Roadside Safety Inspection

The architecture for automated roadside safety inspection could be configured to support several different evolutionary variations that may arise in meeting this user service requirement. The options are based on assumptions about the amount of on-board equipment and data storage that commercial vehicles will support:

1. In the lowest on-board equipment scenario, the vehicle supports only the capability of electronic identification, using the ID capability of the tag service package.

2. In the most equipped scenario, the vehicle supports on-board safety monitoring equipment, the capability to record and report previous inspection history, and the communications capability to convey this information to the roadside in real-time. This has the benefit of putting the smallest burden (for real-time database maintenance) on the infrastructure. This approach has the concern for security of the on-board data.

The architecture described in this document is for the less equipped scenario. The summarized impacts of these different approaches are shown in Table 18.

5.8 Location of Intelligence for Intersection Collision Avoidance and Automated Highway Systems

In the advanced vehicle safety services as well as in the automated highway system services, a choice can be made as to whether to place "intelligence" processes in the vehicle, in the infrastructure or a combination of both. For the advanced safety features of the architecture, the choice was made to place most safety processes in the vehicle with the exception of intersection collision avoidance. In intersection collision avoidance, as well as in the AHS services, a balance of process intelligence between the vehicle and the infrastructure was chosen.

5.8.1 Intersection Collision Avoidance

The tradeoffs between an infrastructure based and a vehicle based intersection collision avoidance system are shown in Table 19.

Function	Infrastructure Based	Vehicle Based
Intersection State - presence of signal control devices, their state, timing and data on road surface conditions.	Only realistic source for information.	NA.
Sensing - detect and locate all vehicles within a given radius of the intersection.	Unobstructed view of entire intersection and its approaches.	Limited effectivity due to sensor design limitations and field-of-view obstructions.
Determine present and future vehicle states - position vs. time.	Vehicle state vectors calculated from sensor data.	Could be calculated from own position data transmitted by vehicles(s); requires properly equipped vehicles.
Identify potential collisions.	Calculated from vehicle state vectors.	Calculated from vehicle state vectors.
Determine countermeasure response: advisory, warning or control intervention.	Based on decision-making algorithms.	Based on decision-making algorithms.
Communications.	Primarily vehicle-to-roadside Dedicated Short Range Communication (DSRC).	Primarily vehicle-to-vehicle.
Actuation.	NA.	Throttle control, automatic soft, moderate, or hard braking and/or lateral steering control.

Table 19. Infrastructure vs Vehicle based Intersection Collision Avoidance Tradeoffs

5.8.2 Automated Highway Systems (AHS)

The tradeoffs between an infrastructure based and a vehicle based AHS are shown in Table 20. The architecture currently supports an infrastructure/vehicle based AHS:

1. Infrastructure controls entry/exit to dedicated AHS lanes.
2. Architecture supports advisories and signing through infrastructure.

3. Longitudinal/Lateral displacement is vehicle based.

4. Lane change/merging/platooning is based on Vehicle-Vehicle cooperation.

Function	Infrastructure Based	Vehicle Based
AHS State/Advisories.	Only realistic source for information; requires vehicle-infrastructure Dedicated Short Range Communication (DSRC).	NA.
Entry/Exit.	Requires DSRC; provides most secure access control.	Based solely on in-vehicle sensors.
Longitudinal Displacement.	Requires sensors, processing and DSRC interface throughout AHS network.	In-vehicle sensors (e.g. millimeter wave or laser radar, machine vision).
Lateral Displacement.	Vehicle based sensors or signal pickups track lane boundaries/center markings.	Vehicle based sensors or signal pickups track lane boundaries/center markings.
Lane Change.	Requires sensors, processing and DSRC throughout AHS network.	Requires lateral/longitudinal sensors, proximity sensors, and vehicle-vehicle communications.
Platooning.	Requires sensors, processing and DSRC throughout AHS network.	Requires lateral/longitudinal sensors, proximity sensors, and vehicle-vehicle communications.
Communications.	Primarily DSRC.	Some DSRC and vehicle-to-vehicle.
Collision Avoidance.	Requires sensors, processing and DSRC throughout AHS network.	Requires lateral/longitudinal sensors, proximity sensors, and vehicle-vehicle communications.
Roadway Instrumentation.	Required.	Considerable function on non-instrumented roadways.

Table 20. Infrastructure vs Vehicle based Automated Highway System Tradeoffs

5.9 Architecture Robustness to Spatially Different Deployments

Section 3.1.2. included a discussion describing the specific decomposition of the ITS into subsystems to support many variations of deployment by location and over time. This section will discuss the effect of the communications layer on architecture robustness to spatially different deployments.

5.9.1 Using the NII for Inter-Subsystem Communications

This section discusses how different deployments (Urban, Interurban and Rural) of the architecture will interoperate. Specific focus on the strategy for mobile subsystems will be discussed.

Figure 39 shows how a large number of CSPs both internetwork themselves together and connect to their clients, thus provide point-to-point messaging for their clients to and from any other clients on the set of internetworked CSPs. This is, at a very high level, how the Internet functions today, and how the NII will function in the future.

The key feature of this system is that the clients can connect to their chosen CSP with any technology, wireline or wireless, and send messages to any other client of any other CSP, independent of the technology that they have chosen to connect with.

Mobile clients will be able to "roam", that is move geographically anywhere that their service provider offers service, and be able at any time to send or receive data messages. This capability already exists for the Analog Mobile Phone System (AMPS), and will be extended for the emerging data services based on the same cell-based communication technologies. Mobile data service providers can be expected to make arrangements with providers out of their area to reciprocally give their clients roaming privileges outside of their "home" areas, thus extending mobile connectivity across the country.

Institutionally, the clients have a choice of CSPs to choose from. Competition for clients will result in aggressive pricing and technology deployments by the CSPs, leading the commoditization of wireline and wireless communications referred to earlier.

Figure 40 shows how the NII will connect different regions of the country. In this way, ITS can develop by regions (or "islands of ITS"), which as they grow will merge, using the NII as the communications backbone between the clusters of ITS deployment.

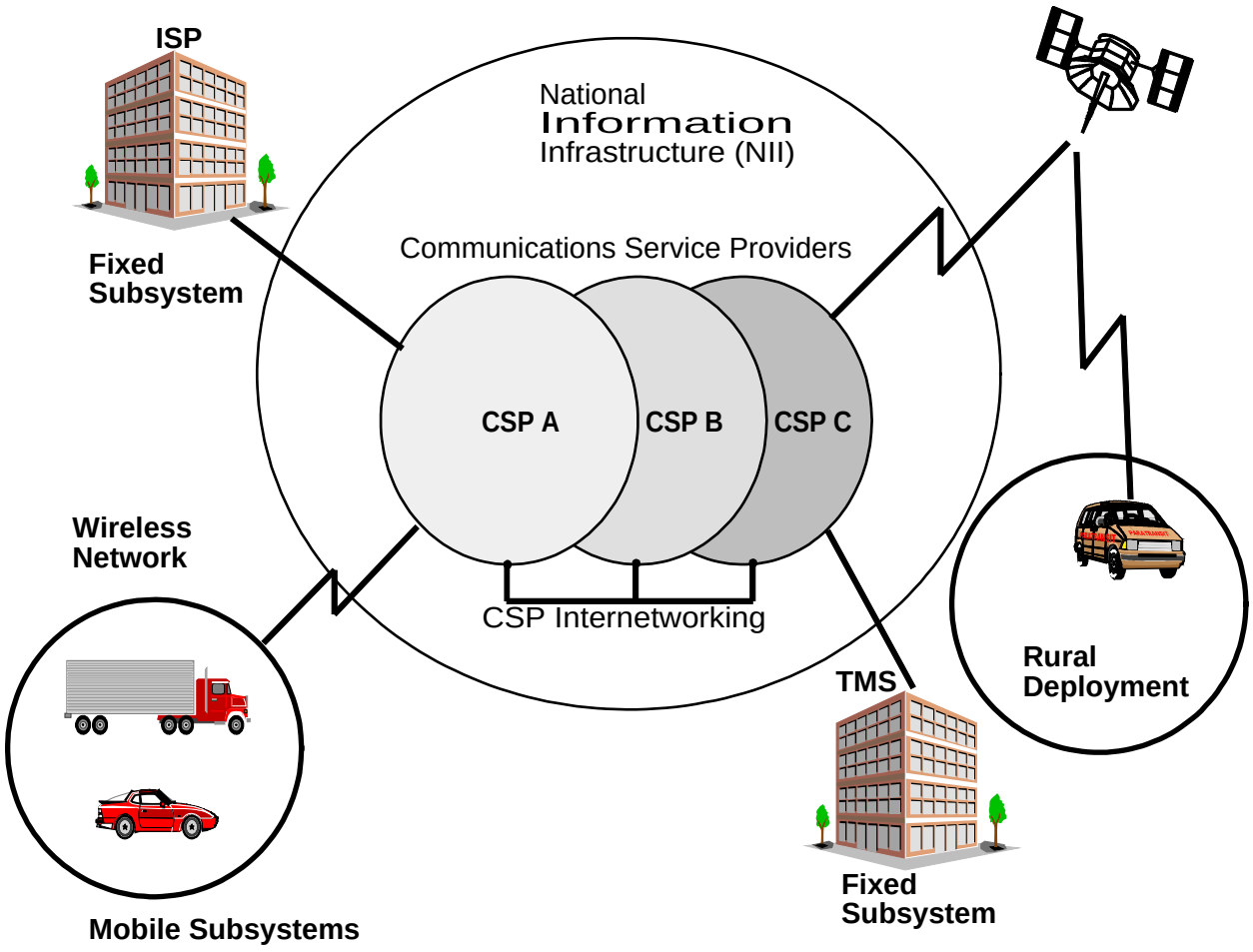

Figure 40. ITS Communications Network Topology

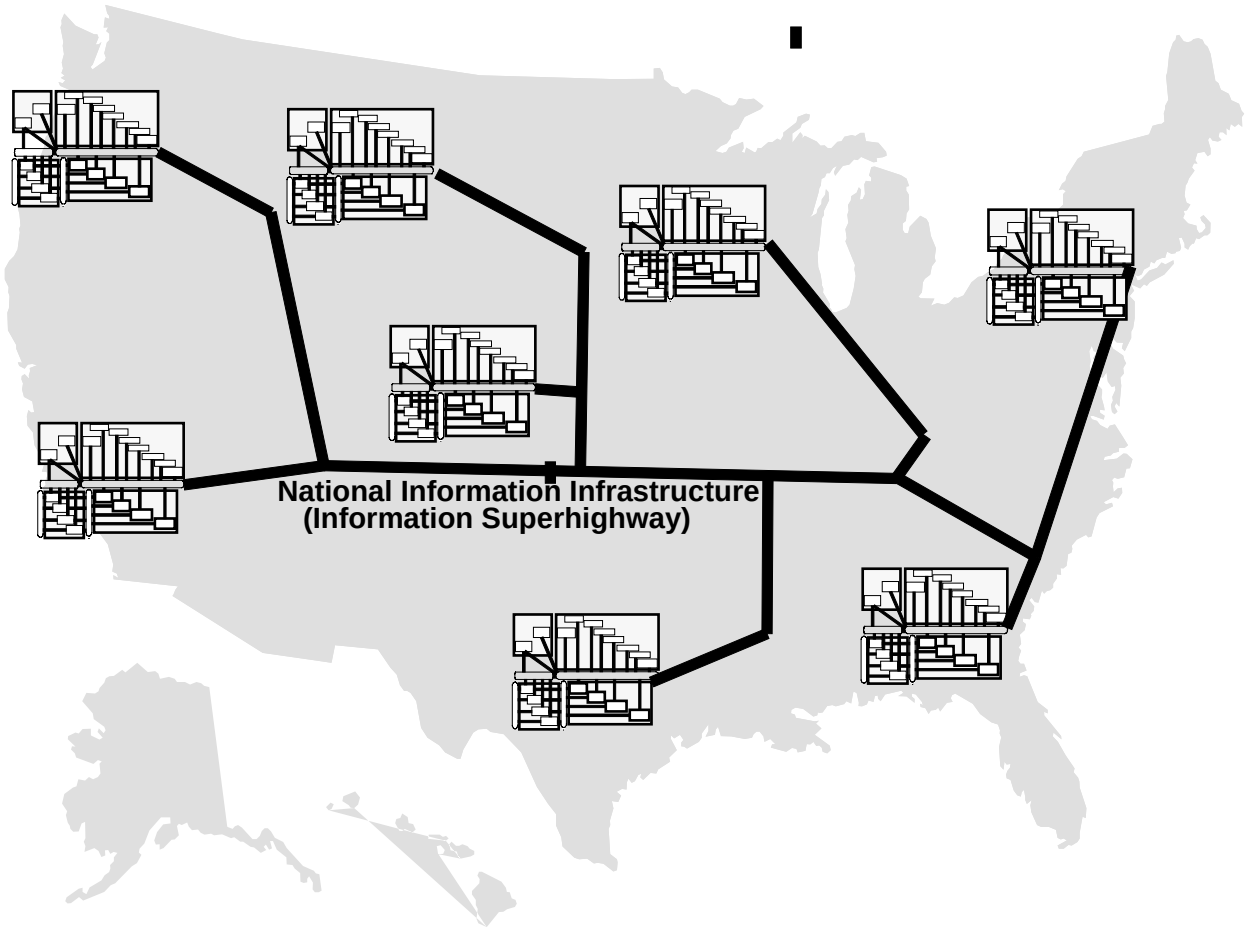

Figure 41. National ITS System Interoperation on the NII

5.9.2 TMS-TMS Communications

TMSs in a region may choose to institute a regional traffic management strategy. If they do, they can each program the same "policy" for demand management and signal coordination, in their respective TMSs, and then by sharing data they can collectively implement a regional traffic management strategy.

Technically, sharing the link and intersection attribute data, mostly at the boundaries between the TMSs, requires that each TMS requesting data from an adjacent TMS know which TMS to contact. This becomes more complicated as the links of interest get further away, and it may be difficult to easily achieve this on an ad hoc basis. A mechanism proposed for this data sharing is shown in Figure 41. Here each TMS has a basic ITS map database with a single attribute for each link datum: communications ID of the TMS (if any) that maintains the attributes for the link.

5.9.3 TMS-ISP Communications

Figure 41 also shows that the same mechanism used for TMS-TMS communications is used for TMS-ISP communications.

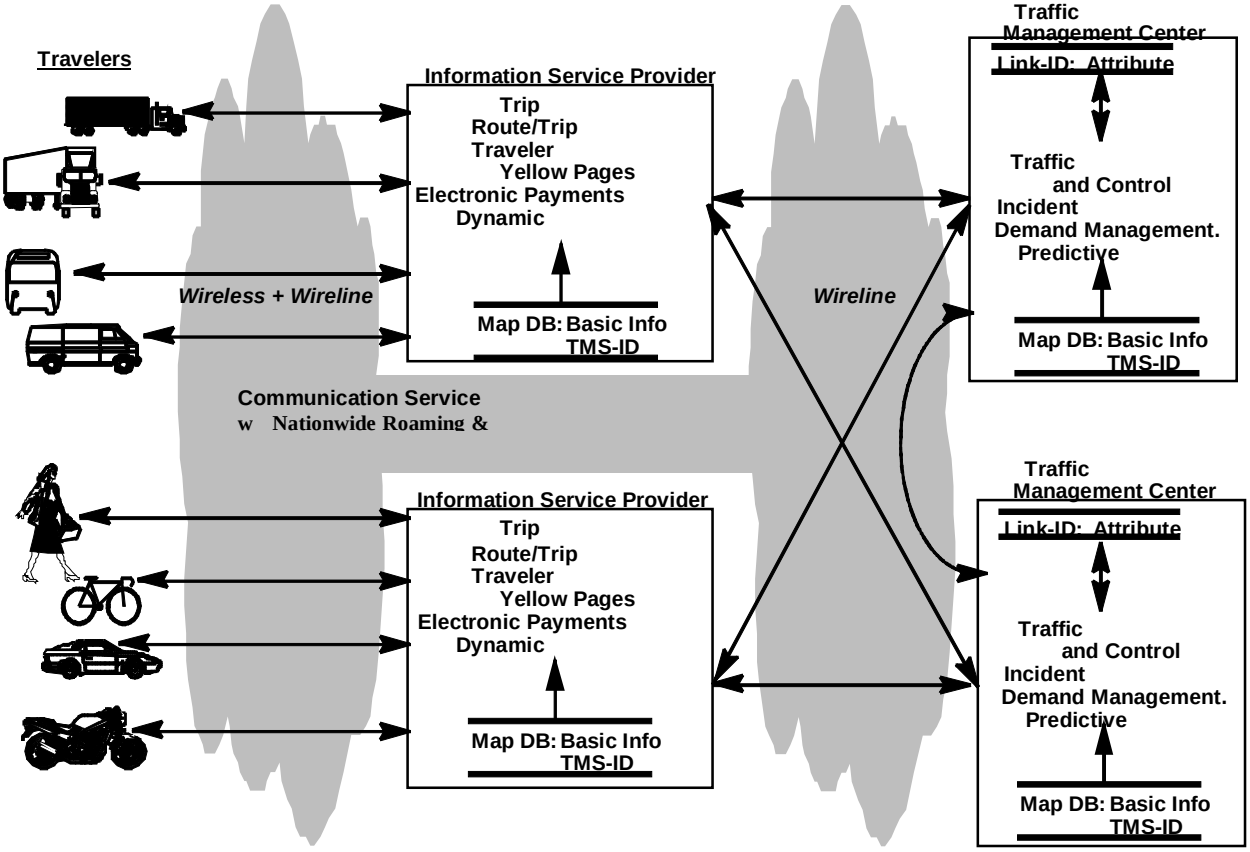

Figure 42. Open National Compatibility

5.9.4 ISPs and End Users

Figure 41 also shows that ISPs use wireline and wireless communications to provide services to their clients (including Transit Vehicles and Emergency Vehicles for Publicly operated ISPs). Figure 42 combines the previous discussion regarding Communication Service Providers, internetworked to effect the NII, and the connection to an ISP, to show how a client can choose an ISP and be in communications with the selected ISP no matter where they are in the country. At the same time, the ISP is using the NII to access TMSs across the country to service the requests of the client.

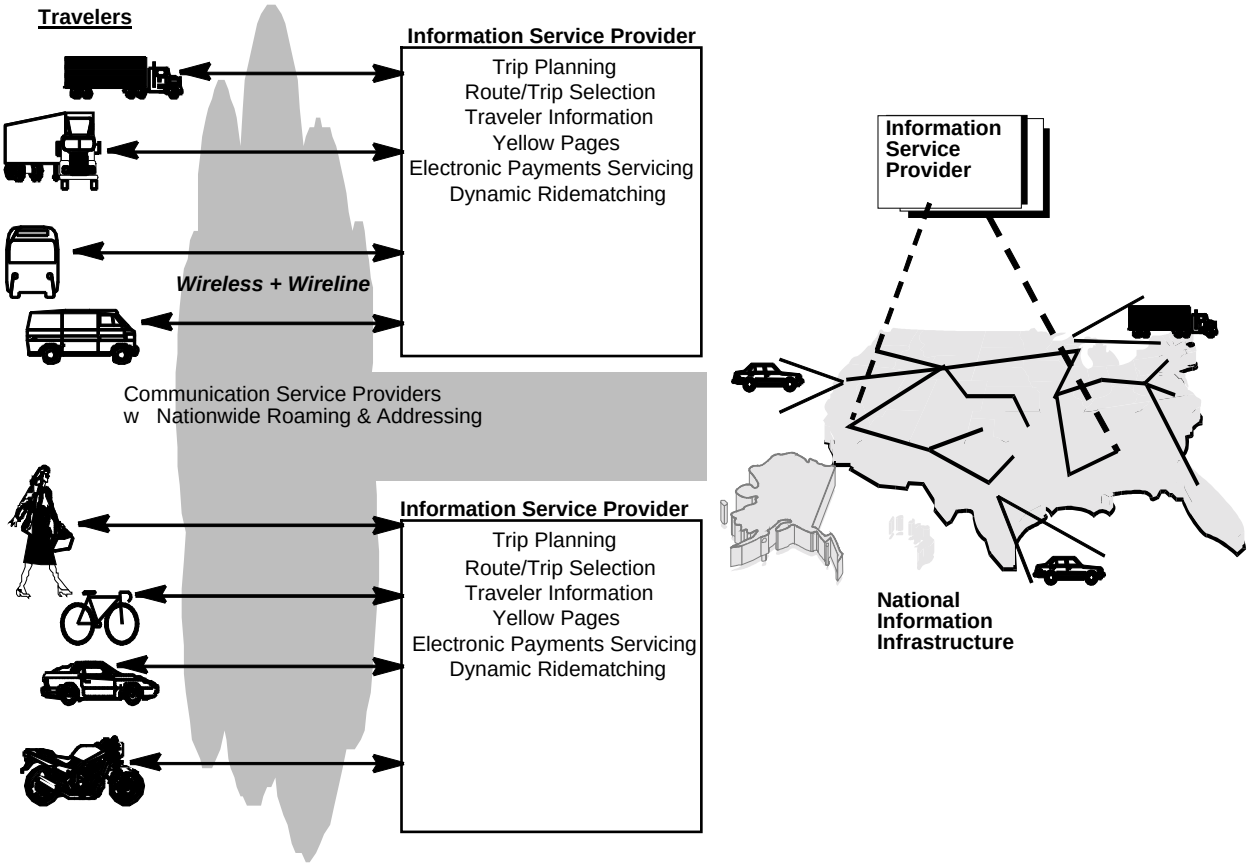

Figure 43. Traveler to ISP National Interoperability

5.10 Communications in the Highway Rail Intersection (HRI) Architecture

A critical consideration in developing the HRI architecture is to accommodate the substantial investments that have been made and are being made in communications with trains and train crews. In the US, there are more than one mechanism for communicating vital safety information to the train regarding forward intersection safety. For example, some systems use or are planning to use communication schemes that go directly from the intersection to the train, using some form of Dedicated Short Range Communications (DSRC) (implemented with beacons or spread spectrum radio). Other systems are planned to use a broadly centralized scheme where wayside equipment and trains each communicate by radio to base stations spaced along the right-of-way, and the base stations relay the data to a centralized train control facility. Vital wayside to train messaging is accomplished in this case via the centralized train control facility.

To accommodate these diverse operational and institutional requirements, the National Architecture has placed the Rail Operations to Wayside Equipment interface and the interfaces to the Train outside of the architecture. The National Architecture communicaties with the Rail Operations and Wayside Equipment terminators, which then communicate messages to trains by some mechanism, determined locally by the rail operator. This architecture is shown at a high level in Figure 43.

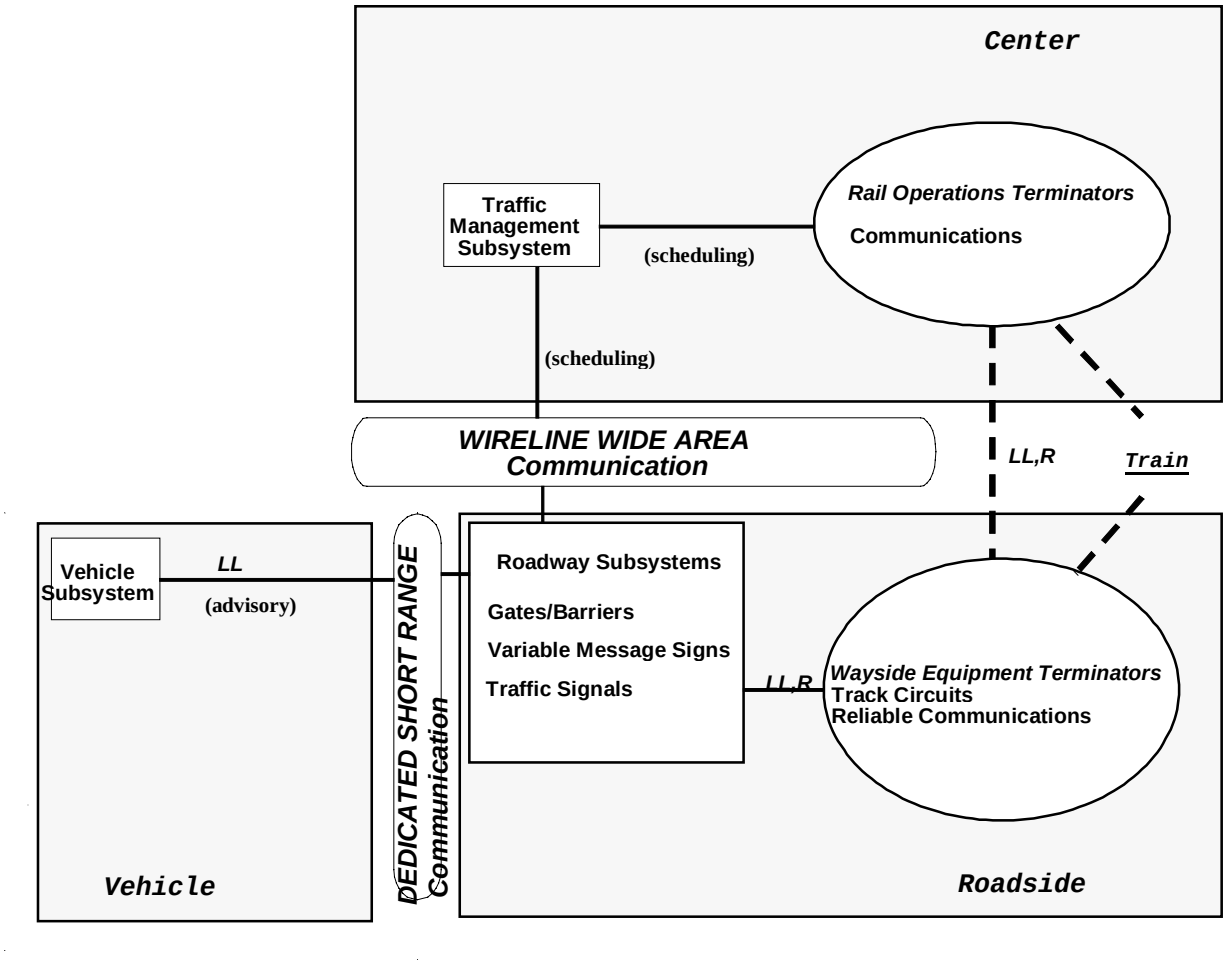

Figure 44. Communications, Subsystems and Terminators in the Highway Rail Intersection (HRI) Architecture