The US National ITS Architecture:
Part 2 - Application to Travel and Traffic Management

Robert S. Jaffe
President, Jaffe Engineering and Development Industries
POB 501, Shenorock, NY 10587
914-248-8466; rsj@jeng.com

Introduction

The first part of this article presented an overview of the US Intelligent Transportation System (ITS) National Architecture (NA). This second part uses the Travel and Traffic Management user services to illustrate some of the operational concepts of the *National Architecture* (NA). For these user services there are a variety of evolutionary paths toward increasing functionality. The operational message sequencing given here describes some, but not all, of these paths. *They are meant to be a guide to how the NA supports the user services, and are not meant to be a prescription of the only way the user services can be supported.* The NA has been designed to have a great degree of flexibility, and can support a wide array of implementations.

Route Guidance and Route Selection

The Route Selection process involves selecting the route to take based upon the Driver, Traveler or Commercial Vehicle Manager request. The Route Guidance process involves presenting the selected route to the driver or traveler in a step by step (sometimes called *turn-by-turn*) fashion. All implementations of the Route Guidance process in the NA are in the Vehicle subsystem or Personal Information Access subsystem (e.g., a PDA or Personal Digital Assistant). The location of the Route Selection process and the source of the Route Selection process information distinguish different Route Selection operating modes of the NA.

The NA supports the autonomous mode of route selection processing in the Vehicle or Personal Information Access subsystems and also supports the mode of route selection where the Information Service Provider (ISP) provides link and queue times to the mobile route selection processes. The most fully functional method of providing route selection in the NA is infrastructure based route selection through a *client-server* request-response between the client traveler subsystem and the server ISP subsystem.

The client-server mode between traveler equipment (Vehicle or PDA) and the ISP is the most fully functional because the traveler is able to make full use of all the data the ISP has systematically developed about the transportation network. With one-way broadcast data, the ISP does not know the travelers destination, and thus must send information of interest to all travelers in the broadcast market. This information must of necessity be more general and not customized to a particular travelers request. The client subsystems use this general information to compute specific routes for individual travelers. Of course, routes computed in the vehicle or in a PDA are completely private, even though potentially less accurate. This tradeoff of privacy for accuracy can be knowingly made by the traveler.

Whether providing travel information in a client-server or broadcast mode, ISPs may themselves be clients requesting information from other ITS subsystems such as Parking Management, Transit Management and Traffic Management. Ideally, a *just in time* information strategy would have the ISP request specific information from these source as the ISP clients make requests that might benefit from 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.

Travelers may modify their trip requests to the ISP to get a variety of multimodal trip plan responses to choose from. For example, a traveler may vary the departure time of a trip request to see the effect on expected time of arrival. The trip request message specifies basic information (e.g. origin, destination, departure or required arrival time), constraints (e.g. modal requirements such as public transit, private vehicle, if a vehicle is equipped for future automated highway lanes, if a vehicle is a commercial vehicle and if a hazardous materiel load is being carried) and preferences (e.g. desired mode(s) of travel, preferred alternate routes, willingness to participate in a rideshare as a driver or a passenger).

The traveler need not indicate to the ISP which trip they plan to take, but in some cases they may gain certain benefits by doing so. For example, if the trip request included a reservation of some kind by the traveler, confirming to the ISP that the traveler will make that trip, may allow the ISP to broker the reservation of some resources that will either enable or substantially reduce the variability of the trip. For example: a paratransit (demand responsive transit) trip reservation to a Transit Management subsystem; a route plan to a Traffic Management subsystem for an emergency vehicle (for signal preemption) or for a transit vehicle (for signal priority); or a parking space reservation for a specific vehicle to a Parking Management subsystem.

The trip plan sent to the traveler can include an ID number associated with the computed route, so that the traveler equipment can efficiently refer to the selected route for route confirmation.

In addition to an ID number, the trip plan message from an ISP to a traveler can include a start time and a list of route segments to follow. Each route segment includes start and end locations (using latitude, longitude, elevation, a short character string identifier and a node ID), travel mode (e.g. vehicle or public transit route identifier), estimated travel time and estimated segment conditions (e.g. ‘‘icy conditions”). The node ID allows reuse of the node specification for later trip plans by compact reference.

In addition, a trip plan may include a list of locations where the travelers equipment should report *probe data* to the ISP. Probe data at its simplest is a timestamp, indicating when a vehicle passed a specified location. This information can be used by the ISP to compute samples of link travel times, and these link travel times can collectively be used to improve the ISPs model of the transportation network. At the same time, when a traveler reports their current position to the ISP, it gives the ISP an opportunity to recompute the traveler’s best route to their destination, and can then notify the traveler if unexpected congestion or incidents indicate a change in arrival time or that a better route is available. This personalized service benefit may encourage many travelers to participate as probes.

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 Traveler Information and Traffic Management functions, or in regions where the Traffic Management subsystem chooses not to participate or does not exist, and will allow competing ISPs to differentiate their services.

ISPs may aggregate anonymous probe data for different classes of vehicles and share this surveillance data with Traffic Management subsystems in return for Traffic Management fixed surveillance data. Furthermore, the ISP can send selected routes to the Traffic and Transit Management subsystems (with vehicle classification and occupancy but without traveler or driver/vehicle identification) so that the appropriate expected statistical occupancy models for roads can be updated, reflecting the incremental congestion and transit time impacts that the planned route will have on the transportation network (or transit vehicle occupancy).

Where there is close coordination between ISPs and TMSs, vehicle type (but not personal identification of vehicles) is included in the message from the ISP to the TMS. The *type* of vehicle is 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 hazardous materials. In these cases, the selected route may also include the hazardous materials manifest, each item identified by the material safety data sheet number (MSDS, 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.

Traffic Control

Surveillance data from the Roadside subsystems (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 Traffic Management Subsystem (TMS) take these inputs and create the signal timing and phasing messages used to control traffic. This control can also include variable message signs, movable lane barriers/markers, or any other Roadside subsystem traffic control features.

In a highly integrated system, the Transit Management subsystem 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 NA also supports the current implementations of DSRC requests for transit priority and emergency vehicle preemption.) An advantage of this architecture is that the TMS, with knowledge of vehicle routes and expected turning movements, can give selective signal priority to selected classes of vehicles 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).

As part of its overall surveillance activities, the TMS also gets messages from: Event Promoters reporting information on large traffic generators e.g. sporting events and concerts; the Weather Services on current and predicted weather; the Emergency Management subsystems on incident information and finally the TMS can communicate with other TMSs and request traffic data, predicted incidents and / or current incidents.

Roadway subsystem based surveillance equipment may include DSRC toll tag readers that can use this data as probe information which is forwarded to the Traffic Management subsystem as additional surveillance data.

As discussed before, the Transit Management and Information Service Provider subsystems may send current position and expected routes and occupancies of vehicles to the Traffic Management subsystems. During times of heavy congestion, this may only be for high priority vehicles (Emergency, Transit, High Occupancy Vehicles, 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.

Coordinated regional traffic management requires that adjacent jurisdictions (which may each have their own TMS) agree on a common traffic management policy. The NA 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 NA. If each TMS in a region cooperates to request and share coordination data among themselves, then they will each have exactly the same overall regional traffic database on which to execute traffic control algorithms and policies which are not specified by the NA, but are implementation dependent. If each TMS implements 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.

Integrated Traffic Management, Demand Management and Route Selection

Figure 1 shows the high level interactions among traffic management, demand management and ISP based dynamic route selection. Note that Figure 1 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).

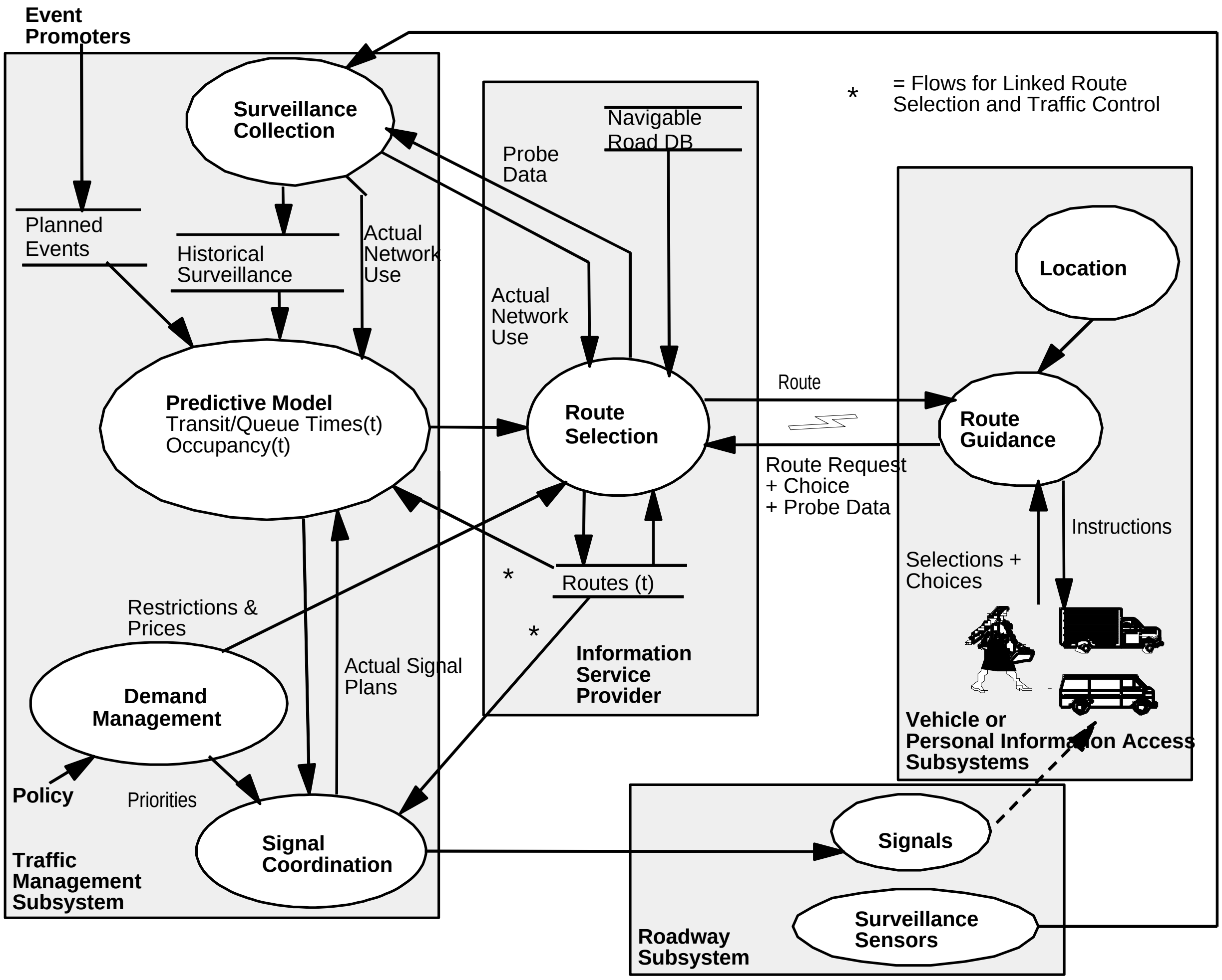

Figure 1. High-End State Traffic Management, Demand Management and Dynamic Route Selection

- 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 functions as well as TMS functions.
- 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.

Probe Data Reporting to ISPs and ISP Updates to TMSs

As shown in Figure 1, 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 traveler 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 step-by-step instructions to the traveler. Of course the traveler 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.

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.

Public-Private Partnerships

The predictive model process that resides in the TMS (in Figure 1) 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.

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 in a *competitive joint venture* in the ownership, operation and maintenance of this Predictive Model only TMS. The point of this venture is to improve the routing that they each 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 Traveler Information / Traffic Management architecture in the high end state of ITS deployment.

TMS Demand Management

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

TMS Prioritized Routing by Vehicle Class

As shown in Figure 1, the routes of vehicles participating in route selection are sent 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.

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., variable 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.

Route Selection

As shown in Figure 2, the NA currently supports a continuum of modes of route selection from simple in-vehicle autonomous to fully integrated Traffic Management - Traveler Information route selection.

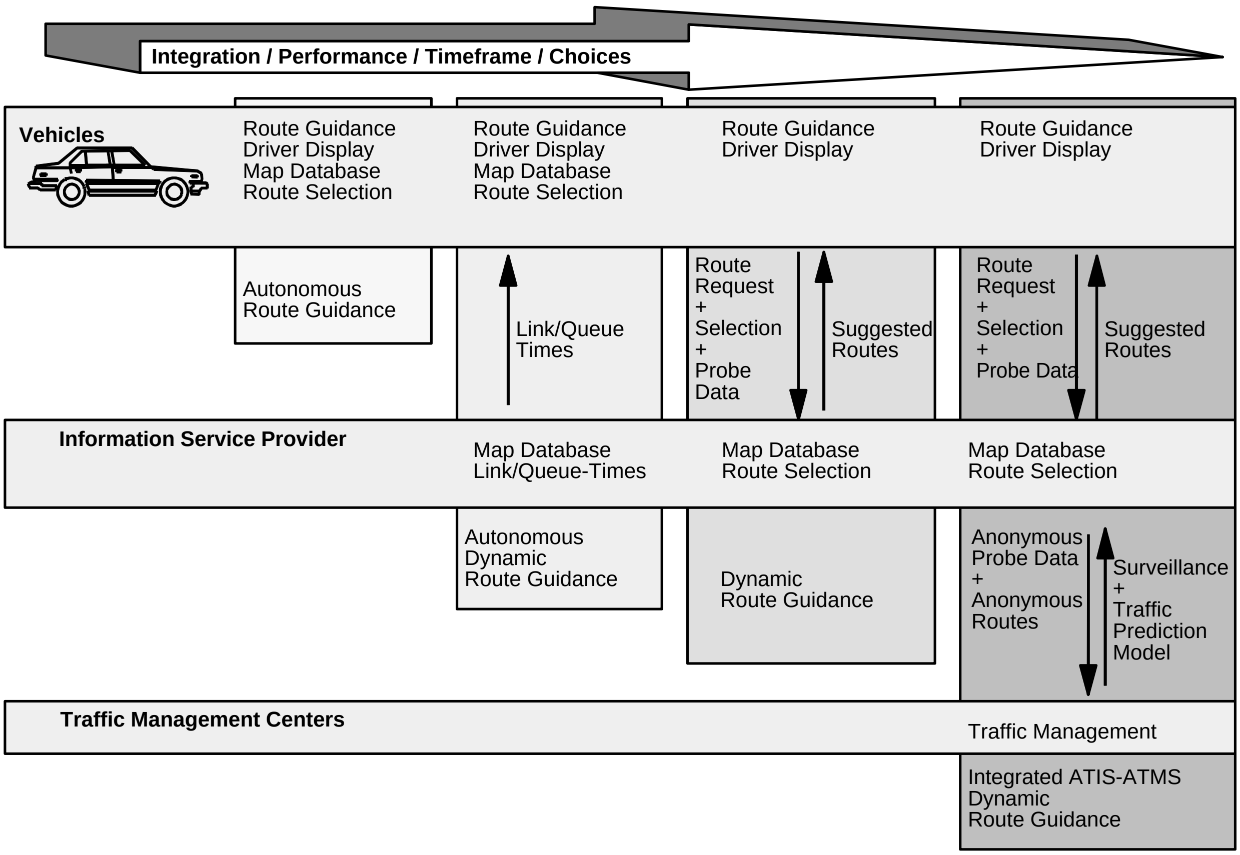

Figure 2. Route Selection Alternatives

Advertising as a Revenue Source for Advanced Traveler Information Systems (ATIS)

Signage information can be added by an ISP 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 NA, is to deploy roadside DSRC beacons at or near sign locations, and in-vehicle equipment receives and repeats the DSRC messages.

A key feature of the ISP based in-vehicle signage mechanism is that it may provide a 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.

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 in the NA for this data sharing is shown in Figure 3. Here each TMS has a basic ITS map database with a single attribute for each link datum: the data communications ID (e.g. an *Internet Protocol* or *IP* address) of the TMS that maintains the attributes for the link.

TMS-ISP Communications

Figure 3 also shows that the same mechanism used for TMS-TMS communications is used for TMS-ISP communications.

Figure 3 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).

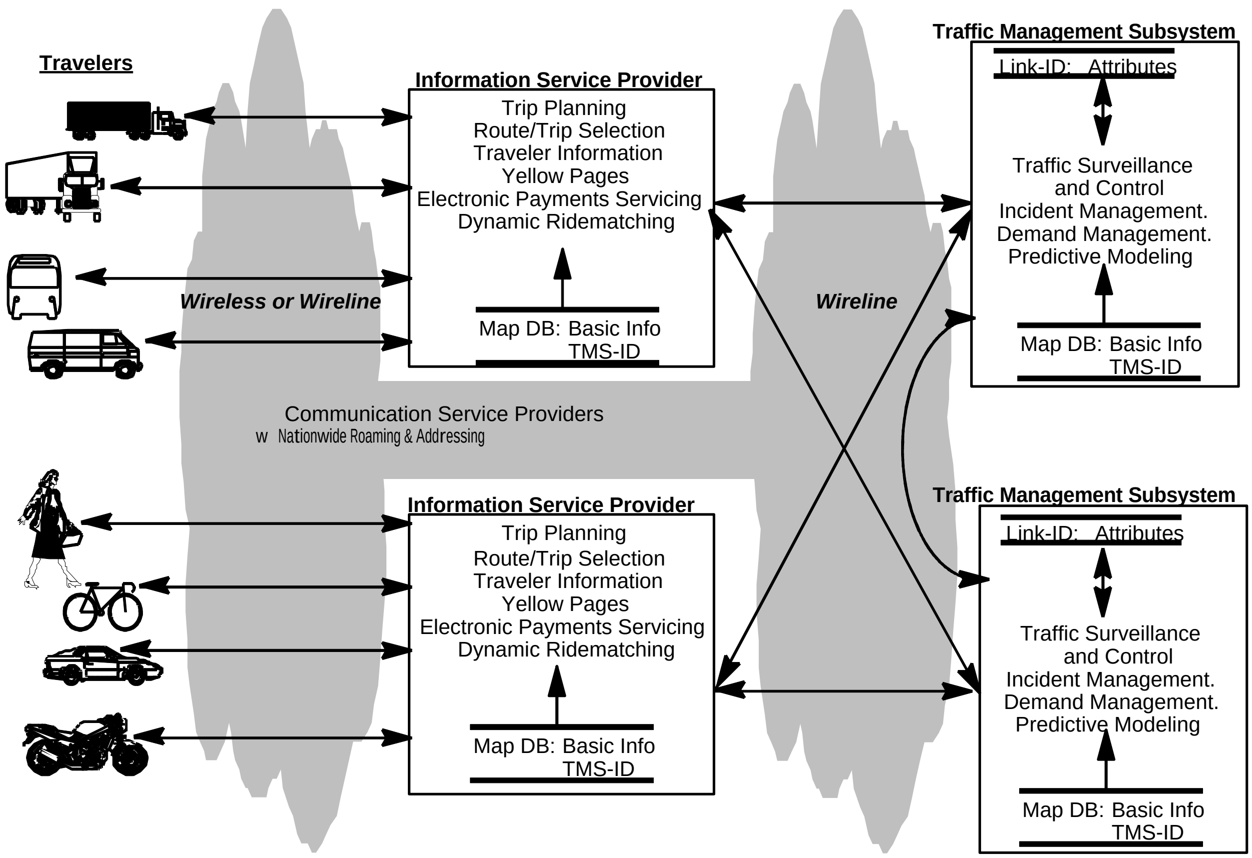

Figure 3. Open National Compatibility

Conclusions

A variety of NA supported paths toward increasing functionality and integration for the Travel and Traffic Management user services was shown. This flexibility to support a wide array of implementations is a key characteristic of the NA.

For different deployments of the NA the specific benefits, tradeoffs and implications from the perspectives of the various stakeholders were discussed. Examples were chosen that demonstrated how the NA supported various institutional partnerships (public-public, public-private, and private-private) to allow each stakeholder entity to achieve individual benefits (e.g. just-in-time or minimized time planning ability) with simultaneous achievement of broader societal ITS goals (e.g. signal control and demand management for safe and efficient traveler, freight or public safety utilization of public infrastructure).