DATA MANAGEMENT TIDE
(Traffic Information of the Durban Environs)

DARRYLL THOMAS[1] and PETER MEANLEY[2]

[1] eThekwini (Durban) Municipality, P O Box 680, Durban 4000
[2] Online Systems, Real Time Control Software Solutions, P O Box 35268, Northway 4065

1. INTRODUCTION : DATA VS INFORMATION
1.1 Intelligent Transportation Systems (ITS)

The purpose of Intelligent Transportation Systems (ITS) is to deliver information about the road network conditions to the appropriate people reliably, accurately and timeously. This simple definition belies the complexity of implementing today's ITS systems. The diversity of information that can be provided is huge and the target audience vast - a control room operator requires timing and incident data, the traveller is only interested in details that would affect his journey, maintenance staff rely on accurate fault reports and inventory records. In fact, many ITS applications can provide varied data as a low / no cost byproduct of their daily operation.

ITS systems are therefore necessarily complex. In an effort to simplify them many are built around a centralised database, with information gathered from multiple sources, which is then accessed through the different applications that process the data into information. While traffic information is obviously the key resource, meaningful information has to be extracted from extensive quantities of diverse, raw data and its presentation is required in a wide range of formats. This data is largely unusable if information cannot be easily extracted and, ideally, applications that require little previous experience and training are required for this purpose. Further, these applications used to tap data sources for information must be simple and intuitive tools, without exposing the complexity of the underlying data access and processing. Applications must ultimately be able to provide user friendly representations of complex data sets - such as thematic or GIS maps which present real time data from field devices as colour coded themes on an intelligent map or touch screen, forms-based displays.

In an ideal world, there would be just a single source of all data, capable of being presented in many different ways to suit all viewing perspectives. Not so in the real world !

1.2 Data

Data is a critical component of any transportation infrastructure and is vital to both the operation of many traffic applications and also the subjective evaluation of ITS projects, having a significant impact on overall integrity and credibility in both areas. Often, as much planning goes into the data management as the actual project itself and careful attention must be given to:

21st Annual South African Transport Conference
'Towards Building Capacity and Accelerating Delivery'
ISBN: 0-620-28855-8
South Africa, 15 - 19 July 2002
Conference organised by: Conference Planners
CD-ROM produced by: Document Transformation Technologies

$ Data availability
The mechanism to collect and record data regarding the performance of an ITS system must be set up in the planning stage of the project. e.g. a freeway incident detection system recording traffic incidents and directing response from emergency services must be designed to have the evaluation in mind so that it contains all the required information, such as time of incident, time reported, departure time of emergency service, time of freeway closure and reopening etc. as well as the more obvious location and traffic condition indicators.

$ Data quality
Often more attention is given to ensuring that the actual data is collected than on its accuracy and the raw data is then not subjected to quality control or error analysis.

$ Data inventory
A system may record various types of data which are technically available for performance assessment, but due to inadequate planning this information it is unusable in practice. e.g. real-time traffic data may be collected at 30 second intervals but is not suitable for traffic control applications due to the lack of an efficient data inventory.

1.2.1 Real-time Data

Online traffic flow data is gained from different sources such as traffic actuated signal detector loops, dedicated traffic monitoring devices and video detection sites. Each source of online data is an external subsystem and delivers data via its particular communications link and data converter interfaced with a central database. Technical standards of detectors, transmission protocols and the quality and level of aggregation of the data may differ. This online data can therefore be classified by type and level of aggregation (counts or speed records, vehicle based or interval based) which is mapped internally in a standard format. Traffic authorities add value by processing the data in various ways, then communicate it to travellers through variety of channels and devices such as in-vehicle monitors, variable message signs, Internet and traffic radio reports.

An ITS system may record various types of data which is technically available for performance assessment, but due to inadequate planning this information it is unusable in practice. For example, real-time traffic data may be collected at 30 second intervals but cannot be used for traffic control applications due to the lack of an efficient data inventory. Also, SCOOT (a real-time Urban Traffic Control System application) generates information messages that, if just one type of message out of dozens generated is collected for a 24-hour period, could exceed 115 Mb. However, this information is probably of no interest to the motorist in alphanumeric format and needs to be superimposed on a GIS map. Consequently deciding which information to collect, and how to deliver it to the user, is of great importance.

1.3. Historical Data

A significant chunk of traffic data collected and stored by traffic authorities is historical, providing records of signal timings, accident statistics, classified vehicle turning movements, street furniture / equipment inventories and traffic signal fault records and their corresponding response / repair times. Again this data is useless in its raw form and also requires applications, in the form of traffic simulation models and statistical software packages, to produce practical information such as traffic signal timings, accident blackspots and performance monitoring factors to justify project implementation.

1.4. Sharing Data

With the emergence of the internet and the increasing acceptance of the WWW browser as the interface, there has been an increased use of the web to implement systems on business intranets. It makes sense, all the user needs is a PC and a browser to interact with any other web based system. In essence, by using a browser to interface with the user, system developers are trying to make systems easier to use, or to put it differently, making the user's data (which after all belongs to them) more accessible. This ability to share large amounts of peripheral data generated by transportation systems, in an effort to either maintain and monitor the performance of sophisticated and expensive equipment or offer the best travel advice across all transportation modes at all times, is an ongoing challenge for Urban Traffic Control (UTC) facilities.

Data is expensive to collect, store and maintain, is often duplicated (a source of errors) and stored in dispersed and disconnected physical locations and in different formats. Raw data is undeniably any traffic authority=s greatest asset and must be systematically managed from start to finish in the information generation process - which includes insuring easy access for all the users.

The **TIDE** integrated database is one such management support tool and the system design is such that any future ITS facility can easily be added to the environment, providing an accessible path forward from, simply, traffic control to a flexible travel demand management tool. e.g. advanced algorithms can now yield information on traffic volumes (between sensors) and expected traffic times so with computational enhancement raw information concerning traffic conditions could be translated into accurate estimates of travel time across a network.

2. TIDE

2.1 What is TIDE ?

TIDE refers to Traffic Information of the Durban Environs, the name given to the eThekwini (Durban) Municipality's integrated traffic database, and its design is based on the principle that:

Information can only be used effectively if it can be accessed
and understood by everyone who has an interest in it.

eThekwini (Durban) has a history of traffic data dating back to 1967, as shown in Table 1 below, when accident details were stored on punched cards and magnetic tape at the City Treasurer's office. By 1980 the city's various traffic disciplines had proliferated data on traffic signals, traffic counts, closed circuit TV and accidents which was migrated to a DEC mini-computer in the Traffic Control Centre. The mid-1980's saw possibly the 2 most significant milestones in the evolution of TIDE:

- the creation of a common locations description database, which essentially became the "glue" which would bond all the other TIDE subsystems into a cohesive, integrated database
- soon thereafter a relational database was created by linking the existing traffic modules to the UTC System, which facilitated the automatic updating or cross-referencing of data between the TIDE subsystems.

Table 1 : History of Data Management in eThekwini (Durban)

YEAR	FUNCTION	MILESTONE
1967	Accident Suite	City Treasury (punched card / mag tape) computer
1968	Traffic Volumes	bureau (card based)
1974	ATC System	Honeywell mainframe
1980	Accident Suite	migrated to local DEC mini-computer
1982	Traffic Volumes	migrated to local DEC mini-computer
1984	Traffic Signals	modules integrated (relational database)
1986	Accident locations	moved to relational database
1992	User interfaces	enhanced across all database modules
1997	GIS interface	pilot project for locations database
1999	Migration	from VMS/Oracle to NT/SQL server architecture
2000	TIDE	Phase 1

The migration to NT/SQL server architecture in 1999 set the foundation for the full integration of TIDE in 2000 and for the first time the wealth of historical data (shown in Table 2) was available to all technical staff, instead of just the few select expert users as was previously the case. Microsoft Access applications now made it much easier to collate data from the various database modules to produce technical or performance monitoring reports.

Table 2 : Historical Data on Record (2001)

TIDE STATISTICS
Accident Suite
➢ 76,000 Unique locations on 6,900 roads
➢ 241 Areas
➢ 381,500 Accidents
Traffic Volumes
➢ 1,400 Unique locations
➢ 11,882 15 minute classified traffic counts
Traffic Signals
➢ 620 Signals & 25,174 Inventory items
➢ 102,026 Signal faults & annual maintenance costs
➢ 23,857 Site Inspections

In 1997 Durban's TIDE officially rolled in with the start of a phased implementation of a formal data management system. An outline of the proposed development of TIDE can be represented as block diagrams (see Diagram 1 below) showing Phase 1 and Phases 2 & 3.

Phase 1 (completion deadline 2002)
Link existing traffic databases
INTRANET enabled (LAN)
GIS enabled
Training

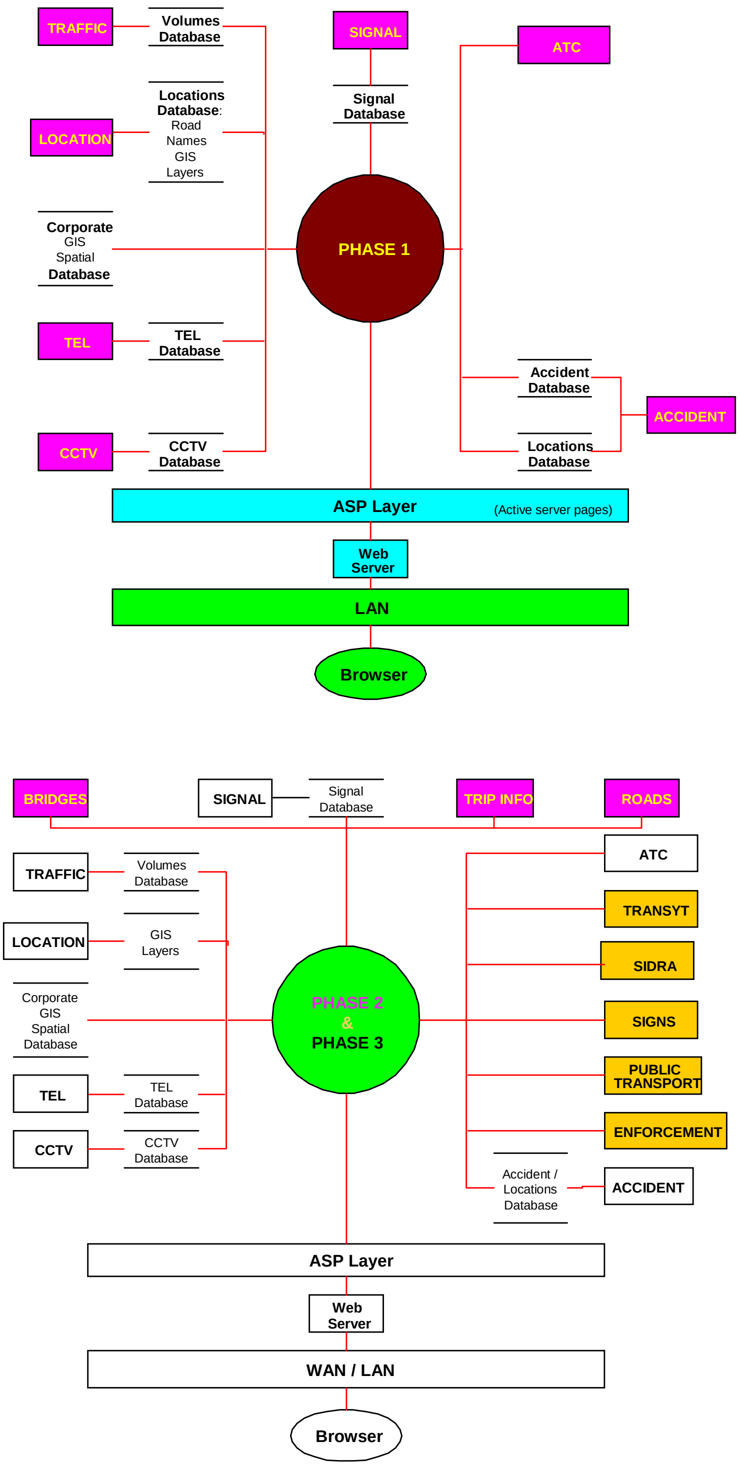
The 7 systems shown (pink boxes in diagram below) already exist as live systems. Phase 1 involves grafting a web interface onto these disparate interfaces, providing the user with a single interface at browser level. The ASP layer consists of a number of active server pages and supporting software developed using Microsoft's Visual Studio. Each (coloured) block represents an entire system which in some cases consists of thousands of lines of code written in Java or Visual Basic.

Phase 2 (2003)
Automated enforcement
Other traffic systems
INTERNET enabled (WAN)
Thematic maps

Phase 3 (2005)
Appoint co-ordinator
Include roads & bridges
Public access
Pre-trip & In-trip information

Additional systems (some of which are under development) will be added during these last two phases, extending TIDE access to the public via the Wide Area Network (WAN).

Diagram 1 : Implementation Phases



2.2 Architecture

The TIDE system architecture uses the three tier client / server concept that any task accomplished by computers can be organised in 3 layers:

1. Data layer facilitates storing, retrieving, searching and maintaining the integrity of information created by the system and can range in complexity from simple direct access files to indexed files to private relational databases and corporate data warehouses.
2. Presentation (user interface) layer controls all interaction between the hardware and the user.
3. Application layer is the software that applies business logic and functions (based on input from the presentation layer) to the data.

This architecture provides the benefit of:

- Flexibility -even though the same base data is available for different purposes to a variety of users, any part of the system can be changed without affecting operation of the rest of the system.
- Modularity - information systems developed to support the transportation infrastructure are highly complex, which is a problem that software developers must overcome if such systems are to be reliable and easily expanded to cope with future requirements. The modular construction of the software components of a database enables the system reliability to be tested in small, discrete manageable portions and also respond to external events, servicing requests from client processes.
- Transparency - the user is only aware of the interface that is being accessed, without being exposed to the complexity of the underlying data access and processing.

2.2.1 Distributed database

TIDE is a distributed database with a central server holding the current database and client systems providing the user interfaces. This enables each database node to be at an appropriate location, if necessary, with its own copy of the data and synchronised to the other databases. However, each node is a separate entity and needs only hold the data that it requires. One node may deal with incident data, another statistical data, maintenance records and so on. The major advantages are scalability and expansion where nodes can be added as required to reduce load on existing nodes if there is increased demand from a user. Likewise new ITS applications can be implemented simply by adding new nodes capable of storing the required data.

2.2.2 Interface

The interface between man and computer has been going through an evolutionary process. From punched paper tape and cards for input with impact printers and typewriters for output, through to keyboards and screens with pointing devices today. Along with the hardware changes have come changes in the software. In the early 1980's the Lotus 123 spreadsheet package with its drop-down menus revolutionized the way users could interact with the computer.

Since then, nothing much has changed with the hardware interface, we're still stuck with screens and keyboards, but all sorts of innovations have been introduced in the way the computer screen is used to interface with the user. Some of these innovations have been really inovative, but (and here's the rub) just about every single piece of software has its own unique interface. The net result is that user must learn the new interface and this becomes yet another barrier to the average man in the street using any computer.

2.2.3 Browsers
A growing medium for the distribution of transportation information is the Internet - a technical term for a network of computers linked together by routers. It started as a scientific / academic network in the USA in the 1960s but has grown rapidly as a global framework for interconnecting diverse computer networks throughout the world. Based on transmission control protocol / Internet protocol (TCP/IP) packet switching protocols the Internet provides the framework for the distributed multimedia client / server system called the world wide web (WWW). Users around the world with access to computers or kiosks can interact with the Internet to obtain specific, updated, traffic information. Future development of wireless connection will improve the availability of the WWW for en-route traveller information as well.

The WWW browser is therefore an integral component of TIDE. The proposed development for TIDE is made possible because an easily extensible architecture (web applications and a browser interface) has been chosen. Just to emphasis that point: because TIDE has been developed for the web from the beginning, it is easy to extend its availability to a wider network of users.

Phase 1 will see TIDE implemented on the Traffic & Transportation Department's LAN supporting only Microsoft's browser - Internet Explorer Version 5 or later, since this is the eThekwini (Durban) Municipality's Corporate standard. User authentication will be done at operating system level, i.e. if a user has an account on the LAN, they will at least have read access to TIDE. Permission to change TIDE data will be controlled by grouping the users, again at operating system level i.e. users will be permitted to change only their specific data. The beauty of this scheme, is that it is relatively easy to maintain by system management staff.

Phase 2 of TIDE will see more data being made available along with implementation on the WAN. Authentication of users on the WAN has not yet been specified, but it would probably involve some registration of users outside Traffic & Transportation, and on a read only basis. The third phase of TIDE envisages WWW access, again this has not yet been specified, but it will be possible to permit free read only access to some data while collecting online payment for other data that the public are currently charged for, e.g. accident records and traffic counts.

2.2.4 Platform
The Internet facilitates easy access (with appropriate access privileges) to the data. It also provides a cost-effective way to send and receive information between different locations and was chosen as the publishing medium for TIDE because the data transmission network was already in place.

Browser-based distributions are system independent and eliminate concern about deployment hardware (Apple Mac, Pentium I to IV) or operating system (Windows, Mac, UNIX). In essence applications would also be platform independent, as the same application could run on any PC hardware, and with this inherent interoperability, cross-platform compatibility is achieved using standardised Internet components.

Providing documentation in HTML permits electronic distribution and online help with hyperlinks is yet another advantage. HTML essentially provides computer browsers and Web sites with a common language. Any Web browser can access any place on the Web and read data because it is all formatted in HTML and the browser understands that language. SQL is a standard database description language that is more complex than HTML but is essentially the same concept. A decision was made to migrate TIDE from VMS/Oracle to NT/SQL server platform because of the inordinate disparity in licence fee and easier integration with existing PC word processors and spreadsheets. Windows/SQL Server platform has subsequently been chosen as the Corporate standard for eThekwini Municipality.

2.4 Benefits

The following benefits offered by TIDE are a direct result of the many enhancements offered, compared to the previous system of separate databases, as detailed in Table 3 below.

1. Easier access to Departmental traffic information
- Data is available via a common interface (also to other traffic authorities & police)
- Integrated sub-systems
- Easier to ensure systems are kept updated
- Less staff training required (in the use of various systems)
- Increasing use of the available data

2 . Potential improvements in worker productivity
- Quicker response from staff
- Performance tracking
- Simple reports

Table 3 : Comparison with Previous Database

TIDE	Previous Database Modules
➢ Graphical colour displays	➢ Character based mono displays
➢ Windows NT, SQL, Visual Basic, MsAccess, Internet Server, IE Browser, HTML	➢ Vax / VMS, RdB, Fortran, Rally, DecForms software
➢ Client-server architecture using multiple Intel Pentium 32 bit servers	➢ Single server DEC Alpha chip based 64 bit hardware
➢ 32 bit OS and database software	➢ 64 bit OS and database software
➢ Homogeneous user interface	➢ Different user interfaces
➢ GIS data display	➢ None
➢ Integration with other PC packages	➢ None
➢ Wider access to traffic data on LAN	➢ None

3. DATA MANAGEMENT

3.1 Data structure

Data used is grouped into data objects, which are logical groupings of data attributes, the fundamental, lowest-level data items. The data structure is a set of rules setting out the relationship between the ITS functions, the data items and the data attributes.

Traffic control is an ITS function and its *data attributes* can :
- convey information about a traffic situation as a phrase
- be a coded value
- be a numerical value
- control traffic signal equipment

Traffic control *data objects* could include :
- traffic signal timing plans
- traffic signal design plans
- traffic volume counts and flows
- lane configuration
- traffic equipment inventory

For example the traffic signal timing plans *data object* includes the following *data attributes*:
- traffic signal not working (a phrase)
- signal status (0 = unknown, 1 = red, etc)
- cycle time (numerical value)
- ISOL (UTC system command to isolate a particular traffic signal from the UTC system)

3.2 Data Dictionary

The data dictionary defines the data used by a system and includes a set of rules to allow interpretation by the user. Raw data must be linked to a data dictionary which contains meta-data (descriptive information about the data and the system , e.g. the data comes from loop sensors and is in units of measurement) and content information (such as location of the sensor, its units of measurement, how the data is stored). Together the meta and content data provide enough details to even duplicate the database at another site. A data dictionary must be compatible with that of other standard protocols (such as NTCIP) and the naming convention used for the National Architecture so that the data, or information reports, can be easily exchanged between databases.

Problems commonly found with sharing raw data:
- data must be in the correct format, understandable for its intended use - the more entities that receive the data the more complicated this process becomes
- if an organisation distributes data, it has to be able to explain to recipients what the data is and this requires a dynamically changing data dictionary (e.g. if new loops are installed that now record classified vehicle counts everyone who can access the data needs be aware of this change or the new data will be meaningless).

3.3 Integration

Integration is the key to truly successful application of ITS technology but the limited success of initiatives to change the way in which we travel in our town centres is an indication of the difficulty of integration. The problems of congestion are well known, public transport is underutilised but creating an effective multi-modal solution to relieve congestion through greater use of public transport has only been achieved in a handful of countries. Isolated fixes only scratch the surface of the problem - an integrated plan is therefore essential. An integrated database such as TIDE is one link in the chain of possible solutions - by making the relevant information readily available to all (eg. static & moving traffic violations via CCTV to enforcement authority, updated public transport timetables to passengers, routes effected by delays or roadworks to motorists).

There are a host of technologies currently available for a huge range of applications, from licence plate recognition to GPS-located public transport. Information systems can be used in the management and operation of transportation infrastructure to integrate these separate ITS technologies into a single, more effective system. The common denominator in all aspects of integration is the sharing of information through access to an integrated database, which must address three essential issues:

3.3.1 Functional integration
ITS uses emerging technology to revolutionise the exchange of information. It will integrate and build on many existing or emerging physical components, including communication networks (cable, telephone, mobile, radio, Internet and satellite systems), computers and information appliances and information in the form of databases, information services, bulletin

boards and libraries. Each ITS component is usually developed in isolation for a specific function, by specialist companies with their own unique solutions, but the information generated is only useful in combination with that from other functions.

Effective functional integration should consider how the following information can be exchanged:

< weather conditions (real time)
< traffic conditions (real time)
< status of traffic devices (camera, VMS)
< highway maintenance projects - current & planned
< reported accidents (real time & historical)
< public transport schedules
< location of public transport (real time)

3.3.2 Institutional integration - sharing information between different organisations

It is important to remember that organisational arrangements (hierarchies, departments, functions etc.) are as much a part of the transportation system as technology and the blame for lack of progress usually lies with institutional problems and red tape, as opposed to technical hurdles. Problems can be either explicit (such as one Department claiming or avoiding responsibility for specific functions) or implicit (if one Department cannot use or understand information provided by another). Either way, there will be frustrating problems ranging from error-prone duplication of data to more serious situations where essential tasks cannot be carried out. Institutional and organisational arrangements often hinder both access and understanding, both key factors for information to be used effectively.

Through the South African Society for ITS (SASITS), moves are afoot to standardise on the database structure and data elements in order that information and reports can be transferred between databases of traffic authorities around the country.

3.3.3 Data integration - being able to obtain meaningful information from diverse data sources.

A successful transportation system is judged by its ability to move people and goods safely and efficiently and this requires the support of a fully integrated infrastructure. Data, and the information gleaned from that data, are key to that integration so this implies that standardisation of information systems that handle the data and present the information are critical to the effective operation of our transportation infrastructure.

Meaningful information has to be extracted from raw data from thousands of sources relating to the maintenance and operation of the transportation infrastructure. However, this data is useless unless it can be easily extracted and accessed. The key to effective use of data is being able to :

- access ALL the data that has a bearing on a specific issue
- view that data in a format or display that makes it readily understood
- compare related data sets easily (eg. accident records, traffic signal timings, maintenance records and road conditions)

This allows useful information to be obtained (and understood) from diverse data sources and a number of factors contribute to making this process easier:

- emerging standards such as a National Architecture
- communication standards (NTCIP) to ensure every item of field equipment can communicate effectively with each other, and the traffic control centre, using common protocols.
- adoption of open platforms for information systems

4. CONCLUSION

4.1 Perspective

While the concept of ITS is not new in South Africa, there is currently no coordinated effort between the various road traffic authorities. This has unfortunately resulted in numerous ad hoc ITS type projects being implemented with no thought of overall standardisation or sharing of resources. A classic example of this is the different VMS systems on the N1 highway operated by Midrand and the M1 highway operated by Johannesburg - essentially on the same road for a huge volume of traffic.

Sophisticated Area Traffic Control systems have been operated in our major cities for well over twenty years. Variable message sign systems have been utilized on freeways in major metropolitan areas and in tunnel control applications for a number of years. Closed circuit traffic surveillance systems are operated extensively in the major metropolitan areas.

Generally, the main problems associated with existing ITS applications are lack of political awareness and funding, particularly for the on-going maintenance and operation of the systems. It is this prevailing lack of support at the national level that make successful projects such as TIDE and AtraMS (the Cape Town integrated database) all the more noteworthy. Without a National ITS Architecture or established standards to provide the necessary technical guidance, the development and implementation of both these integrated traffic and transportation databases have relied on the co-operation between eThekwini and Cape Town Metro Transportation authorities, the innovation of a few key traffic technical staff and world class expertise of the computer consultants involved.

4.2 Bottom Line

You do not buy a database, you build one ! .... but build it methodically, after carefully considering issues such as :

I. Who is going to use the data and how will it be used ?
II. Have you established the correct criteria to monitor with this data ?
III. Are you ready (and able) to act on whatever this information tells you ?
IV. Do you have the mechanisms in place to ensure the long term maintenance of the database?
V. Do those involved understand the process and it's impact ?
VI. Have the correct criteria to be monitored been identified ?

Above all, data is extremely expensive to collect, store and maintain so the crucial question that must be answered is: Will the benefits justify the cost ?

5. REFERENCES

1. ITS Handbook [PIARC Committee on Intelligent Transport, edited - K.Chen & J.C.Miles]
2. Intelligent Transportation Primer - Institute of Transportation Engineers

DATA MANAGEMENT TIDE
(Traffic Information of the Durban Environs)

DARRYLL THOMAS[1] **and PETER MEANLEY**[2]

[1] eThekwini (Durban) Municipality, P O Box 680, Durban 4000
[2] Online Systems, Real Time Control Software Solutions, P O Box 35268, Northway 4065

c.v.: Darryll Howard Thomas
Manager : Traffic Management Branch
eThekwini (Durban) Municipality
South Africa

EDUCATIONAL QUALIFICATIONS

1.MSc. Engineering (Transportation)	: 1993,	Natal University, Durban, South Africa
2.Diploma in Datametrics	: 1985,	University of South Africa (U.N.I.S.A.)
3.Registered as Professional Engineer	: 1981	
4.BSc. Engineering (Civil)	: 1975,	Natal University, Durban, South Africa

COMMITTEES
1.Director on Board of SASITS (SA Society of Intelligent Transportation Systems)
2.National ITS Committee of South Africa
3.Advanced Traffic Management Systems (ATMS) Workgroup

PRESENTATIONS
1.Internet CCTV : Live Traffic Broadcast (ITS World Congress, Sydney, 2001)
2.Expanding Infrastructure : The ITS Option (SATC 2001)
3.Review of Urban Traffic Control Systems (SATC 2000)
4.New Road Construction : The ITS Option (SARF 2000)
5.Emerging ITS : South Africa (ITS World Congress, Toronto, 1999)