

ORACLE DATABASE 10G: A SPATIAL VLDB CASE STUDY

By: *Xavier R. Lopez, Oracle Corporation*

Contributing authors: Brigitte Colombo, Oracle Corporation, Ed Parsons, Ordnance Survey, Berik Davies, Ordnance Survey

EXECUTIVE OVERVIEW

All organizations have location data they are not taking advantage of, such as customer addresses, locations of assets, and sales territories. They often need to answer questions based on location relationships such as sales by territory, location of service personnel relative to a maintenance site, or proximity of a first responder to an emergency call. By incorporating location capabilities into core e-business applications, organizations can make better decisions and respond to customers more effectively.

Oracle Database 10g with Oracle Locator or Oracle Spatial provides the foundation for deploying enterprise-wide spatial information systems and wireless location-based services. With Oracle Locator, every Oracle database is a spatial database, enabling application developers to quickly and easily incorporate location information into Internet e-Business and wireless applications and services. In addition, Oracle Application Server MapViewer 10g is a Java component used for rendering maps and viewing geospatial data managed by Oracle Spatial or Locator.

IDC has found that Oracle holds an 80-90% share of the spatial database management market¹ (Dec. 2002). This unprecedented market leadership is due to Oracle's high performance, scalability and ease of use – essential for deployment of enterprise-class GIS, location-based services, and location-enabled e-Business solutions. With Oracle Database 10g, Oracle Locator and Oracle Spatial advance to new levels of performance and include a rich feature set that makes it easy for any application developer to use Oracle for the deployment of GIS, eBusiness applications and wireless location-based services. New functionality in Oracle Spatial specifically addresses the challenging requirements arising from the defense, critical infrastructure and homeland security domains.

This paper highlights how Oracle Locator 10g and Oracle Spatial 10g meet the complex business and technical requirements in deploying location-based services. New Oracle Spatial 10g features such as topology management, georaster support, and spatial analysis are highlighted. Finally, a case study for Britain's Ordnance Survey, which manages one of the world's largest, most sophisticated civilian mapping operations, shows how Oracle Spatial saves time and money as the foundation of their nearly one-terabyte cadastral database.

LOCATION-ENABLED BUSINESS AT THE SPEED OF THE INTERNET

The Internet has redefined the ways that location information is collected, managed, queried and disseminated. Location-based services are now deployed on large, professionally managed servers and accessed via Internet browsers and mobile interfaces. Smaller numbers of large servers are being used to consolidate databases, resulting in lower costs, improved reliability, heightened security, and dramatic improvement in quality of services. The only way to truly make location applications scale to meet the requirements of Internet and wireless deployments is to embed spatial capability natively into the database, which can then be accessed by virtually any device. This is what Oracle has done with Oracle Locator and Oracle Spatial.

Only Oracle offers a completely pre-integrated location services infrastructure, which includes the Oracle Database 10g with Oracle Locator, Oracle Spatial, and Oracle Application Server 10g. Oracle Locator, a standard feature in the

¹ IDC, *Spatial Information Management: Competitive Analysis, 2002*, David Sonnen, Henry D. Morris, Dan Vesset, December 2002 (report # 28348)

Oracle Database (Standard and Enterprise Editions), offers unique, market-leading native spatial capabilities, for most customer applications and partner solutions. Oracle Spatial Option (available for Enterprise Edition) extends Locator with advanced capabilities for complex GIS applications. Organizations that use Oracle Locator or Spatial can take better advantage of their location-based data to improve the effectiveness of their decision-making and service delivery.

Oracle's location technologies support the requirements of many different application areas including:

- eBusiness customers who wish to add useful spatial queries to CRM, data warehousing, and business intelligence applications
- Geographic information systems (GIS) which need to store, retrieve, and manage very large spatial databases
- Internet and wireless location-based applications
- In-vehicle telematics and fleet management services

Oracle Database 10g with Oracle Locator or Oracle Spatial solves key “*How to*” problems for developers of applications, such as how to:

Efficiently store, retrieve, and manage location data and attribute data from a single, open database

- Improve performance for very large databases containing terabytes of location data
- Maintain versions of spatial database tables transparently, and identify any conflicts in table values
- Store location data once and make it accessible to users from multiple, heterogeneous GIS tools and e-Business applications without any modification via open, standards-based APIs.

Oracle Database with Locator or Spatial provides a reliable platform that is extensible, secure, and standards-based resulting in lower training fees, improved learning curves, fewer design and programming iterations, and more efficient data administration.

APPLICATIONS

Oracle Locator and Spatial can be used by information systems (IS) professionals and software developers for a variety of applications. For example:

- *General Location-Based Services*: Enable Internet and wireless service providers to enhance their content offerings with the delivery of geocoded information through a variety of location aware devices. New types of services include looking up addresses, online geocoding, finding travel directions, and identifying the nearest hotels, gas stations, and other places of interest.
- *Customer Relationship Management (CRM)*: Enable organizations to understand, anticipate, and respond to their customer needs, in a cost-effective manner. An Internet-centric business model can use electronic storefronts and self-service to expand service delivery, shorten response time, improve efficiencies, and reduce costs for the fastest ROI.
- *Data Warehousing/ Business Intelligence*: Analyze all transactions being collected in ERP systems (customer purchasing, sales, asset characteristics by time and place) to derive insight and enhance decision making.
- *Supply Chain Management*: Optimize the flow of goods through the supply chain (product mix, inventory, distribution, warehousing, and shipment routes). Add a location dimension to a supply chain so that suppliers can directly review and take action on information that affects them.
- *Geographic Information Systems (GIS)*: Query topographic data for flood plain and basic land management. Combine with population data in a single database to support disaster preparation and relief.
- *Utilities Infrastructure*: Maintain spatial database of the entire network including individual poles, lines, hydrants, or distribution centers. Overlay road and housing data for “dig safe” queries. Manage “long transactions” through version management.
- *Energy Exploration and Distribution*: Maintain virtual maps of underground oil or gas deposits. Determine where

to locate drill sites, refineries, or storage facilities.

- *In-vehicle Telematics*: Enhance and/or enable the delivery of diverse multimedia services to a car or fleet. These services include driving directions, real time weather and traffic (E511), and a range of personalized concierge services enhanced by location data. Intermodal transportation and Intelligent Transportation Systems are relevant, related areas.
- *Personnel Information Management (PIM)*: Integrate location data with typical PIM information in phone book, contact and calendar applications to increase efficiency in scheduling, appointments, business travel and a range of day to day business activities.

BENEFITS OF MANAGING LOCATION DATA IN ORACLE

By effectively managing spatial and attribute data in a single system—with common storage, indexing, query optimization, security, and user management—Oracle reduces processing overhead and eliminates the complexity of coordinating and synchronizing disparate sets of location and attribute data. Other benefits include:

- Spatial data and all types of attribute data can be stored in a single server
- Seamless geographic coverages—no tiling
- Better management of spatial data—SQL access
- Elimination of hybrid GIS architectures and schemas
- Greatly reduced complexity of systems management
- Standard spatial types: proprietary data structures are avoided
- Scalability, data security, replication, partitioning, bulk load utilities
- Support by leading GIS and location technology tool vendors
- Breaking the size barriers—support for tens of terabytes of data
- Support for both short and long transactions (version management)
- OGC conformance certification
- Ready support for Geographic Markup Language (GML)

SUMMARY OF LOCATION FEATURES IN ORACLE

Oracle has recognized the importance of location information to the enterprise for many years. Beginning with Oracle9i location and the ability to manage location information became an integral, infrastructure component of the database server. Baseline capabilities are briefly summarized below.

GEOMETRIES AND SPATIAL REFERENCING

Oracle Spatial supports three basic geometric forms that in turn can be used to represent features such as roads, administrative boundaries, utilities, etc. that typically comprise spatial databases. These spatial primitives include:

- *Points and point clusters*: Points can represent locations such as buildings, fire hydrants, utility poles, oil rigs, boxcars, or roaming vehicles.
- *Lines and line strings*: Lines can represent roads, railroad lines, utility lines, or fault lines.
- *Polygons and complex polygons with holes*: Polygons can represent outlines of cities, districts, flood plains, or oil and gas fields. A polygon with a hole might represent a parcel of land surrounding a patch of wetlands.

To integrate and represent spatial information effectively, Oracle Locator and Spatial provide comprehensive tools for managing coordinate systems and projections. Over 950 commonly used mapping coordinate systems are supported; users can also define new coordinate systems. Oracle Spatial also provides support for converting data freely between different coordinate systems. These transformations can be on a geometry-level basis or an entire layer at a time.

SPATIAL INDEXING: R-TREES

To optimize performance of spatial queries, Oracle Locator provides spatial indexing for relational databases. Oracle Locator provides R-tree indexing, which generates and stores approximations of geometries for the index. R-tree indexes are easy to create and use – almost no tuning is required for optimal performance. R-tree indexes can be created on two, three, or four dimensions of spatial data.

SPATIAL OPERATORS

The interaction of various geometric features can be determined through the use of comparison operators, such as *contains*, *covers*, and *anyinteract* (any interaction). This permits answers to such requests as “list all the school zones crossed by this railroad line,” or “find all pizza parlors within this city.”

FUNCTION-BASED INDEX SUPPORT

Spatial queries can now be performed on relational attributes without explicitly storing location information as a column of type SDO_GEOMETRY. Users can create spatial indexes on spatial data stored in relational columns (e.g., longitude and latitude). Users can then invoke spatial operators on these relational columns without the need to create an SDO_GEOMETRY column.

GEODETIC COORDINATE SUPPORT

Spatial functions, operators, and utilities provide correct results regardless of coordinate system, relating to earth surface measurements. Distance, area, and angular units are fully supported.

PARTITIONING SUPPORT FOR SPATIAL INDEXES

Spatial indexes can be partitioned in association with partitioned tables. Partitioning spatial data and using partitioned local indexes can provide performance gains for queries on large datasets, and concurrent queries and updates. It also makes index maintenance easier.

PARALLEL SPATIAL INDEX CREATION

Index creation can be subdivided into smaller tasks that can be performed in parallel, making use of unused hardware (CPU) resources. For certain spatial data sets and index types and parameters, parallel index creation can substantially increase index build performance and provide a significant time savings.

LINEAR REFERENCING SUPPORT (ORACLE SPATIAL ONLY)

Oracle Spatial stores "measure" information as part of an Oracle Spatial linear geometry. This feature is key to linear networking and dynamic segmentation applications common in Internet street routing, transportation, utility and telecommunications networks, and pipeline management.

SPATIAL AGGREGATES (ORACLE SPATIAL ONLY)

Spatial aggregate functions summarize the results of SQL queries involving geometry objects. Spatial aggregate functions return a geometry object of type SDO_GEOMETRY. The use of spatial aggregates improves performance and reduces the complexity of the underlying code base.

LOCATION IN ORACLE 10G

Oracle Locator and Spatial take advantage of enhanced Oracle Database 10g features and capabilities, including RAC, High Availability, improved manageability, expanded database size limits, high-performance VLDB maintenance utilities, replication, caching, and faster backup and recovery. These capabilities in conjunction with new enterprise Grid features introduced with 10g enable customers with demanding application requirements to reduce cost while improving performance, reliability and their ability to deliver service. The benefits, features and capabilities of the location in Oracle 10g are discussed in the sections that follow.

ORGANIZATIONAL BENEFITS

Handling location data within Oracle Database 10g provides significant organizational benefits – increasing operational efficiency while reducing cost:

Using Oracle Spatial lowers cost by 90% per transaction compared to using separate, isolated, proprietary systems. For the State of Arkansas, its GeoStor system has produced savings of more than \$2 million dollars in its first 18 months of operation, and will continue to generate dramatic savings by cutting the costs of spatial data publishing and other transactions by as much as 90%.

In addition to this validation there are a number of other key business and organizational benefits realized via 10g. These include:

- **SAVINGS:** Provide complete, low-cost, entry-level location services capability using Oracle Locator and MapViewer for enterprise apps. Developers can integrate spatial features directly into business and location-based applications at relatively low cost and with minimal training
- **CONSOLIDATION:** Consolidate all geospatial data types (vector, raster, grid, imagery, network, topology) in a single, open, standards-based data management environment. This in turn reduces cost, complexity and overhead of managing separate, isolated, proprietary systems.
- **SIMPLIFICATION:** Reduce training, software, support and application integration costs inherent in multiple stove-piped geospatial and enterprise systems
- **RISK REDUCTION:** In four separate surveys since 1999, IDC has found that Oracle holds an 80-90% share of the overall geospatial database management market.

NEW LOCATION FEATURES IN ORACLE 10G

There are a range of new features in 10g. Some of these apply to both Locator and Spatial and some are available only in Oracle Spatial. Each of these are discussed in turn in the sections below.

PERFORMANCE ENHANCEMENTS

Performance enhancements improve both Locator and Spatial. Applications ranging from location-based services to GIS asset management and land management must frequently update and query location data — and demand high performance. With every new release, Oracle Locator and Spatial have provided leaps in performance to address this requirement. Oracle Locator and Spatial 10g continue to do so – providing performance improvements of several orders of magnitude over release 9i for location queries and index updates.

Oracle Locator 10g has up to 1000% performance improvement for location (distance) queries over Oracle9i Locator. Additional gains of 400-600% are realized, when running spatial queries in parallel on partitioned spatial indexes. Internal benchmarks show (1) Up to 1000% improvement for all distance-based queries (performance critical for location based services apps); (2) 20-40% improvement for spatial “relate” queries, which calculate exact location relationships; (3) 400-600% improvement for queries accessing multiple partitions.

Other Oracle Locator features can be used to boost performance. Building spatial R-tree indexes in parallel can dramatically reduce index creation time for very large non-point spatial datasets. Spatial aggregate functions speed retrieval of large sets of SDO_GEOMETRY objects.

NEW ORACLE LOCATOR 10G FEATURES

PARALLEL SPATIAL QUERIES

Spatial queries can now run in parallel on partitioned spatial indexes, improving the performance of "within distance", "nearest neighbor", and "relate" queries. Performance scales with the number of CPUs used to execute a query. This helps location service and land management applications, which need to execute high volumes of spatial queries quickly.

NEW ORACLE SPATIAL 10G FEATURES

As traditional GIS technologies hit their own scalability and reliability limits, users are increasingly turning to database centric spatial computing. Spatial databases like Oracle Spatial, move spatial processes and operations directly into the database kernel, thereby increasing performance and security. Oracle Spatial 10g supports new spatial features, which extend the range and productivity of application developers, enabling a broader range of applications and improving performance. The focus of the new Oracle Spatial 10g features is to address the requirements of enterprise GIS customers by providing GeoRaster and network and topology data models, geocoding, and routing capability.

NETWORK DATA MODEL

A data model is provided to store network (graph) structure in Oracle Database 10g. It explicitly stores and maintains connectivity of link-node networks and provides network analysis capability such as shortest path, connectivity analysis. Applications requiring network solutions include transportation, transit, utilities and life sciences (biochemical pathway analysis).

For transportation applications, the network data model also supports a routing feature. Oracle introduces a scalable routing engine that provides driving distances, times, and directions between addresses (or locations that have been geocoded in advance). It is provided as a Java client library to the network data model that can be easily deployed in either Oracle Application Server or standalone OC4J environments. Other features include: preference for either fastest or shortest routes, returning summary or detailed driving directions, and returning the time and distance along a street network from a single location to multiple destinations.

TOPOLOGY DATA MODEL

Oracle Spatial includes a data model and schema that persistently store topology in the Oracle Database. This is useful when there is a high degree of feature editing and a strong requirement for data integrity across maps and map layers. Another benefit is that topology-based queries typically perform faster for queries involving relationships such as adjacency, connectivity, and containment. Land management (cadastral) systems and spatial data providers benefit from these capabilities.

GEORASTER

A new data type natively manages georeferenced raster imagery (satellite imagery, remotely sensed data, gridded data) in Oracle Database 10g. Oracle Spatial's GeoRaster feature provides georeferencing of imagery, XML schema for metadata management and basic operations like pyramiding, tiling, and interleaving. Applications in environmental management, defense/homeland security, energy exploration, and satellite image portals will benefit.

SPATIAL ANALYTIC FUNCTIONS

New server-based spatial analysis capabilities include classification, binning, association, and spatial correlation – essential for business intelligence applications.

GEOCODER

Geocoding is the process of associating geographic references, such as addresses and postal codes, with location coordinates (longitude and latitude). With Oracle Spatial 10g, a fully functional Java geocoding engine is provided. It provides international address standardization, geocoding and POI matching by querying geocoded data stored in Oracle database. Its unique unparsed address support adds great flexibility and convenience to customer applications. The Oracle Spatial geocoder is implemented as Java stored procedures inside the Oracle database server. A PL/SQL API for geocoding is provided.

MAPVIEWER IN ORACLE APPLICATION SERVER 10G

Application Server 10g includes a component called MapViewer. MapViewer is a JAVA-based visualization tool that uses location information from either Locator or Spatial to build and display maps in either a browser or in the context of a specific application. MapViewer can be used to:

- Create customized maps that show features such as roads, city areas, waterways and other transportation networks and

- Display map themes such as national, state and local boundaries
- Visualize business data (e.g. population demographics, psycho-demographics, sales metrics etc.), to portray and explore relationships that can often best be expressed graphically as geographic maps
- Complement an applications workflow, providing interaction with mapped data
- Deliver custom maps over the Internet as a component of JDeveloper or as a standalone tool.

Application Server 10G MapViewer is a J2EE component available with JDeveloper and OC4J. This component was developed to simplify the creation of applications that render and present location data as part of internet and wireless business applications. MapViewer provides a Java Client API and JSP Tag Libraries for application developers to embed map-rendering capability directly into existing e-business applications. Support for session control and basic viewing functions (map legend, select theme to view, etc.) are provided via the JSP Tag Libraries. More control of the image and the application interaction with the map is afforded through the Java Client API's that support map interaction such as pan, zoom, locate, re-center etc.

MapViewer **saves money** by reducing development time via tight integration with the JDeveloper environment. MapViewer **increases the value** of location data in the Oracle database by providing a tightly coupled tool to visualize these data. MapViewer **improves applications** summarizing complex business and geographic data and relationships in an easily understood, universally recognized format – a color map.

PROGRESS THROUGH PARTNERSHIP

Oracle Locator and Spatial are directly integrated with the leading GIS mapping and location services technology vendors. Through the combination of Oracle Database 10g and partner tools, it is possible for developers to rapidly deploy scalable and secure enterprise GIS and location service solutions. For a complete list of partners, please see the Oracle Technology Network at <http://otn.oracle.com/products/spatial>.

OPEN STANDARDS

Oracle consistently works to influence and support the latest open standards, particularly in the area of geographic information and location-based services. Oracle Spatial conforms to the OpenGIS Consortium Simple Features guidelines. Oracle is also committed to supporting the new OGC Geographic Markup Language (GML), as well as Open Location Service and Web Map Server interfaces. The object-relational model used for geometry storage by Oracle Locator and Spatial also conforms to the specifications associated with SQL92 representation of points, lines, and polygons.

CASE STUDY: ORDNANCE SURVEY

Britain's Ordnance Survey manages one of the world's largest, most sophisticated civilian mapping operations. The internationally recognized market leader in its field with over \$150 million in annual revenue, Ordnance Survey supplies map data to every major commercial enterprise in Britain. Oracle Spatial saves Ordnance Survey time and money, while provide a foundation to their nearly one-terabyte cadastral database.

Ordnance Survey's flagship product, OS MasterMap™, is an intelligent digital map, with comprehensive geographic/topographic information covering Britain and other UK countries. MasterMap includes over 430 million topographic features, and is currently 600 Gigabytes in size.

BUSINESS OBJECTIVES

Ordnance Survey has the goal of supplying public bodies and commercial businesses easier and faster access to more up-to-date, accurate, consistent, and useful geographic information. It needs flexible IT systems that are responsive to changing customer requirements.

Historically, IT systems at Ordnance Survey were largely based upon departmental, project driven implementations – resulting in many incompatible systems and databases. These systems were difficult to integrate, costly to maintain, and complex to adapt or extend. OS needed systems to support seamless integration, unified data management, robust geoprocessing, and flexible application configuration.

The key business drivers are summarised in the table below.

Business Objectives	
Reduce deployment costs	Store & maintain once, use many times
Reduce maintenance costs	Facilitate new product opportunities
Improve data integration	Increase business agility
Improve access to data	Standards and COTS based
Improve systems integration	

TECHNICAL STRATEGY

Ordnance Survey developed a new technical strategy based on these business requirements. In accordance with a government mandate to make all public services available electronically by 2005, Ordnance Survey adopted an eStrategy to deliver its products via the Internet. For OS MasterMap, this means that customers can receive quotations, place orders, and download data online. Additionally, Ordnance Survey decided to move away from custom, proprietary solutions, in favor of standards-based, commercial off-the-shelf software (COTS). A COTS based system would reduce long term cost of ownership, and increase speed and flexibility of new system development. Ordnance Survey also wanted to move away from proprietary GIS systems to software based on open standards.

This approach also supports the deployment of enterprise-wide solutions. An example of this architecture is a single spatial database providing data integration, with a suite of integrated capture, editing, and product generation applications that implement business logic and fulfill operational requirements. The advantages include the technical and operational benefits of “store it once, use it many times”, and the replacement of the multitude of existing legacy systems with integrated, lower cost solutions that are easier to maintain and more flexible.

A BUSINESS CASE FOR ORACLE SPATIAL

Ordnance Survey chose Oracle Spatial as the platform for its “Maia” maintenance database. The Maia database system will store the seamless digital mapping base of Great Britain that is the foundation for current and future products in the OS MasterMap line. Maia must be highly scalable to support future expansion to store high volumes of complex data, and deliver high functionality.

A business case for Oracle Spatial
<ul style="list-style-type: none"> • An “open” repository for geospatial data • Industry standard for spatial data in Relational Database Management Systems • Robustness, reliability, scalability, availability • Part of the whole Oracle product family • Strategic enterprise software within Ordnance Survey has adopted Oracle as its enterprise database platform and has the technical infrastructure and skills to implement this software

As the de facto industry standard, Oracle Spatial was seen as advantageous for the development of the next generation editing system (the Field Object Editor) as well as new initiatives that will act on the database, such as automated generalisation. Oracle Spatial supports OpenGIS® and SQL standards, with out of the box integration offered by the major GIS vendors. Thus, multiple vendor GIS applications can act as clients to the Oracle Spatial database.

An important selection criterion was the ability to store and manage very large volumes of data. There are more than

430 million features in Ordnance Survey's large-scale topographic data set. Oracle Spatial stores and spatially indexes these features in a database of about 600 Gigabytes. Managing this volume of data requires the use of advanced spatial features such as partitioned spatial data and indexes, and parallel spatial index creation for improved performance.

Oracle Spatial comes with the high availability, scalability, reliability, and performance that are strong features of the Oracle RDBMS. As its digital data business evolves, these features will become increasingly critical to Ordnance Survey.

Another important feature was the ability to support a dynamic data model. With these high data volumes, changes to the database structures will be time consuming to process, but operations such as changing a table definition are fully supported and can be completed online, without taking the database out of production. These types of manageability features and the productivity tools provided with Oracle (for example, Oracle Enterprise Manager) increase system availability and contribute to the maintainability of the database system.

SUMMARY

Oracle Locator 10g and Oracle Spatial 10g offer the following advantages, which are driving the creation of a new class of spatial applications:

- **Low Cost of Ownership:** Applications are deployed on a corporate level and spatial data stored centrally, lowering ownership costs. Internet/intranet architecture eliminates the need to install and maintain client-side software on the desktop and store and manage data separately, outside of the corporate database. All that is needed on a client machine is a standard Internet browser.
- **Low Risk:** Spatial information is integrated directly into Oracle Database 10g. This results in scalable, secure, and high-performance applications. Developers can choose to deploy on any server platform, including Solaris, Unix, Linux, Windows NT, and Windows 2000, using existing IT resources to manage these applications.
- **High Value:** Using the Internet, much larger numbers of users can access the application at virtually no additional costs to the organization. This means that users can access mission critical information 24 hours a day, 365 days a year.
- **Increased Usability:** Oracle provides graphical interfaces and tuning wizards which simplify the creation and management of spatial tables and indexes.

Oracle Database 10g with Locator or Spatial provides the foundation for deploying enterprise-wide spatial information systems and wireless location-based services. It becomes both faster and easier to incorporate location information into Internet and wireless applications and services. This new release makes a leap in performance, feature availability, and ease of use to meet the unique, demanding requirements of enterprise geographic information systems, Internet map portals, and wireless carriers.