18c ORACLE Database

Spatial and Graph Analytics with Oracle Database 18c

ORACLE WHITE PAPER FEBRUARY 2018

ORACLE

Disclaimer

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Table of Contents

.

Introduction

Oracle provides the industry's leading spatial and graph database management platform. Oracle Spatial and Graph brings advanced features for management and analysis of spatial data, property graphs, and RDF linked data applications to Oracle Database 18c.

Spatial and graph analysis is about understanding relationships. As applications and infrastructure evolve, as new technologies and platforms emerge, we find new ways to incorporate and exploit social and location information into business and analytic workflows. The emergence of Internet of Things, Cloud services, mobile tracking, social media, and real time systems create new challenges to manage the volume of data, but more importantly, to discover patterns, connections, and relationships.

To address these opportunities, Oracle Database 18c includes a wide range of spatial analysis functions and services to evaluate data based on how near or far something is to another, whether something falls within a boundary or region, and to visualize geospatial patterns on maps and imagery. The property graph feature provides graph storage, a SQL-like graph query language, and powerful built-in social graph analytics for making recommendations, finding communities and influencers, pattern matching, and identifying fraud and other anomalies. Property graph also has interoperability with the RDF graph feature that provides linked-data applications with a W3C standards-based triple store, query language and native inferencing.

The property graph feature is a unique, general-purpose graph included in Oracle Spatial and Graph 18c. Current industry offerings tend to focus on providing either analytics or a graph database. Oracle Spatial and Graph is unique in that it has both: a powerful in-memory analyst with over forty built-in analytics, many more than other offerings and a scalable graph database. It also includes a SQL-like declarative graph query language called PGQL that makes it easy to perform graph reversals, and find graph patterns and instances of subgraphs. Property graph support is a game-changing feature because much of the Big Data generated these days contains inherent relationships between the collected data entities. These relationships can be easily structured as a property graph – a set of connected entities, and analyzed to find opportunities and make better business decisions. The property graph models entities as vertices, relationships as edges, and stores associated properties or attributes as key-value pairs for both.

New 18c property graph capabilities for PGQL include querying graph data stored in Oracle Database and finding in-memory subgraph instances that match a given query pattern. New property graph analytics are supported for SQL-based collaborative filtering that can help make recommendations. More in-memory analytics have also been added, including new variants of Pagerank, a Personalized

SALSA for making recommendations, K-Core for finding subgraphs by properties, Diameter, Radius, and Eccentricity to analyze distances in a graph, and PRIM for finding the minimum spanning tree of edges connecting all vertices in an undirected graph. Property graph capabilities also include new support for undirected graphs, Node.js client, Apache Zeppelin and an execution and scheduling manager to better control in-memory analyst tasks and resources.

The geospatial data features are designed to support the most complex requirements found in Geographic Information Systems (GIS), enterprise applications, and location-enabled business and web applications. These features include native support for geocoding, a routing engine, and spatial web services conformant with Open Geospatial Consortium (OGC) and ISO standards. Support for advanced spatial models and types include a network data model, georaster (for geo-referenced imagery and gridded data), topology, 3D, including triangulated irregular networks (TINs) and point clouds (supporting LIDAR data), and linear referencing. These features provide a complete platform for geospatial applications in many domains, including defense, land management retail, insurance, and finance.

With Oracle Database 18c, spatial operations are faster, cloud-ready and more developer friendly. With GeoJSON support, a location tracking service to track millions of objects against thousands of regions, more accelerated spatial functions, and an interactive HTML5 map visualization component, it is the most advanced spatial platform Oracle has ever offered.

New spatial features in 18c include spatial support for distributed and Oracle XA transactions for cloud applications with distributed architectures; expanded spatial JSON support; Oracle REST Data Services support for spatial operations in Oracle Database for modern RESTful development; support for sharded databases with spatial data types; expanded spatial web services features for scalability and usability for WFS, CSW, and WCS; and performance enhancements for operations on spatial point data.

The RDF semantic graph feature included with Oracle Spatial and Graph is a mature, special purpose graph conforming to World Wide Web Consortium standards. It provides parallelized RDF data storage, querying and inferencing that is used in a semantic data integration and linked open data applications. Oracle Spatial and Graph RDF support provides an open and secure RDF database with proven scalability to over a trillion triples. New 18c capabilities include faster operations with support for Oracle Database In-Memory, better SPARQL performance with list-hash composite partitioning,

faster loading of RDF quads having literals greater than 4000 bytes, and native support for Turtle and Trig formats for loading data and creating RDF views.

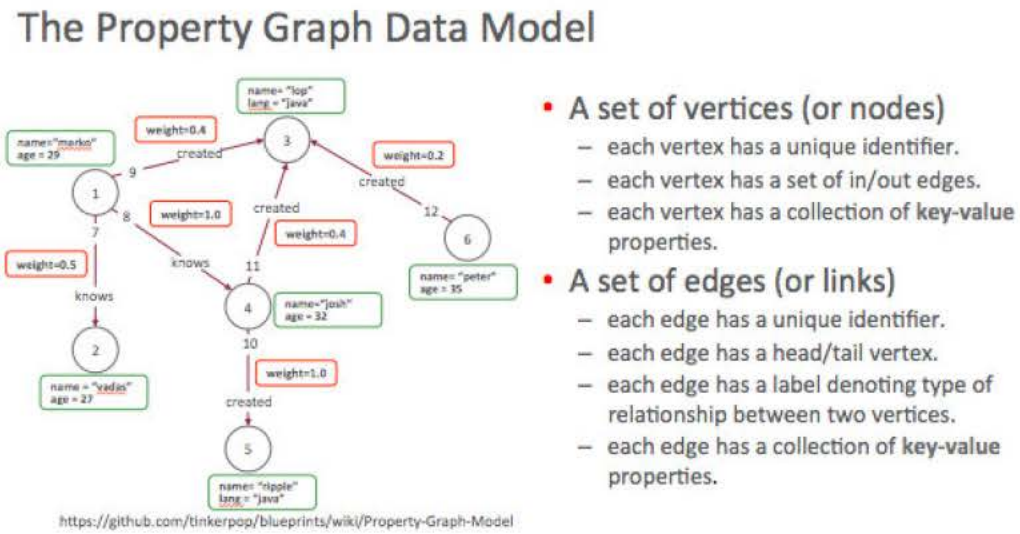
Oracle Spatial and Graph spatial capabilities are part of the database kernel, and geospatial and graph deployments natively harness Oracle Database features for scalability, security, partitioning, and parallelism. They reduce application logic and support real world analysis by moving complex spatial and graph logic into the database. The processing power and bandwidth of Oracle Exadata Database Machine is exploited, realizing extreme performance capabilities that are orders of magnitude over what was previously possible.

This white paper provides an overview of the features in Oracle Spatial and Graph. The appendices list major and new features.

Oracle Spatial and Graph is completely integrated with the performance, scalability, and security of Oracle Database, making it the most advanced spatial and graph database platform available for enterprise class deployments, in the Cloud and on premises.

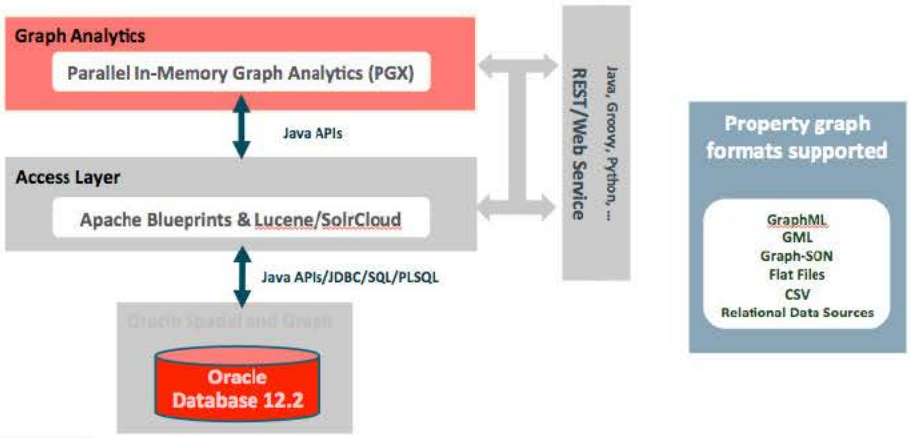
Property Graph Overview

General-purpose property graph support is included in Oracle Spatial and Graph. Much of the Big Data generated these days contains inherent relationships between the collected data entities. These relationships can be easily structured as a property graph – a set of connected entities. The property graph vertices denote entities, the edges denote relationships, and the associated properties or attributes are stored as key-value pairs for both.



Property graph data model

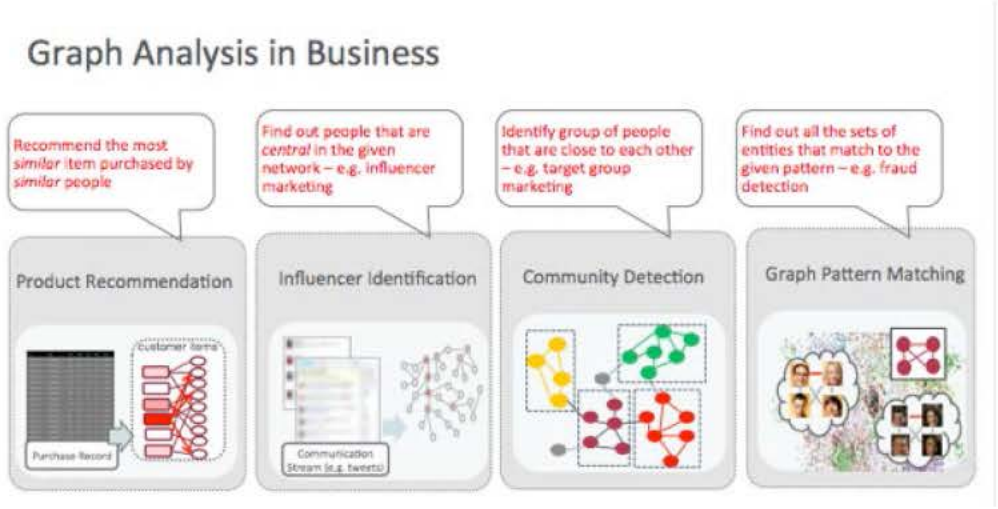
The major capabilities are the in-memory analyst and the data access layer. The in-memory analyst (PGX) is the engine for 40 built-in, powerful, parallel graph analytics. The graph database data access layer includes a Groovy-based console, Java APIs, fast search through text indexing, fast, parallel bulk loading, spatial filtering in graph queries and multi-level security.



Property graph architecture

Graph analytics

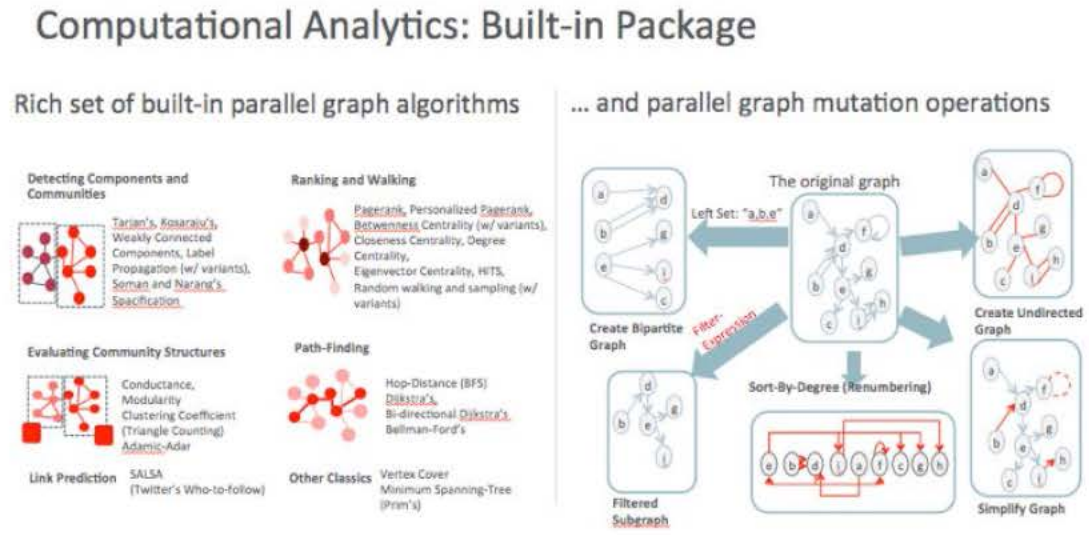
Graph analytics can be performed using the in-memory analyst (PGX) with over forty built-in, powerful, parallel, in-memory analytics, including ranking, centrality, recommendation, community detection, and path finding. Several of the most commonly used property graph analytics can also be executed in-database using SQL. SQL-based analysis can be helpful for large graphs, reducing network traffic and obtaining more up-to-date results.



Example property graph use cases

The in-memory analyst takes advantage of modern server architecture that parallelizes computation using multiple cores and sizeable memory configurations that enable fast non-sequential data access across a larger portion of a graph read into memory. A parallelized filter query on the database reads a subgraph of interest into memory.

The in-memory analytics can either be executed within a Java application or executed in the multi-user, multi-graph in-memory analyst server environment on Oracle WebLogic Server or Apache Tomcat. The output of graph analysis can be another graph, such as a bipartite, filtered, undirected, sorted or simplified edges graph.



Categories of analytics and types of output graphs

Fast and easy to use querying

Querying using the SQL-like PGQL query language

A property graph can be queried using the declarative SQL-like PGQL graph pattern-matching query language. A PGQL query describes a graph pattern with vertices, edges, properties, and their relationships, When the query is evaluated against a property graph, the query engine finds all subgraph instances of the graph that match the specified query pattern. Then the query engine returns the selected data entities from each of the matched subgraph instances.

The following example finds all instances of a given pattern or template in the data graph.

```
SELECT v3.name, v3.age
WHERE
  (v1:Person WITH name = 'Amber') -[:friendOf]-> (v2:Person) -[:knows]-> (v3:Person)
```

PGQL query

PGQL includes support for grouping (GROUP BY), aggregation (e.g. MIN, MAX, AVG), sorting (ORDER BY) and many other familiar SQL constructs.

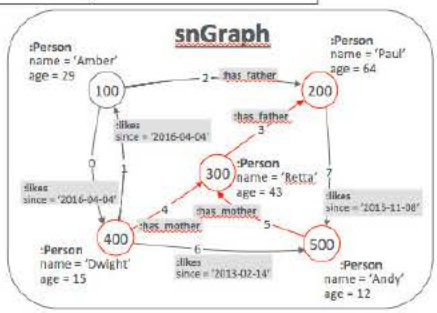
PGQL also supports regular path queries (recursion) for applications such as reachability analysis. A PGQL regular path query is simpler and more concise than a comparable SQL query would be.

The following example matches a pattern repeatedly by defining a PATH pattern at the top of a query, referring to it in the WHERE clause, and using the star symbol (*) for repeated matching.

```
PATH has_parent := (child) -[:has_father|has_mother]-> (parent)
SELECT x.id(), y.id(), ancestor.id()
WHERE
  (x:Person WITH name = 'Andy') -/:has_parent*/-> (ancestor),
  (y) -/:has_parent*/-> (ancestor),
  x != ancestor AND y != ancestor AND x != y
```

Result set

x.id()	y.id()	ancestor.id()
500	300	200
500	400	200
500	400	300



A regular path query

Querying using programmatic APIs

The property graph can be queried using Java APIs that implement Apache Tinkerpop Gremlin APIs. The Java APIs perform parallel scans on vertices and edges. Parallel retrieval takes advantage of the distribution of the data across table partitions, so each partition is queried using a separate database connection.

A property graph can also be queried through SQL. Here is an example of querying for a vertex using a vertex ID.

This example find the vertex with vertex ID 1 in the connections graph.

```
SQL> select vid, k, v, vn, vt
      from connectionsVT$
      where vid=1;
```

The output might be as follows:

```
     1 country     United States
     1 name        Barack Obama
     1 occupation  44th president of United States of America
     ...
```

Example of querying for a vertex using a vertex ID

Ease of development and management

Ease of development and management is facilitated through a Groovy-based console and a set of Java APIs. The built-in Groovy shell provides graph database access and in-memory analyst operations. With this command-line shell interface, you can explore the feature's Java APIs to create and drop property graphs, add and remove vertices and edges, search for vertices and edges using key-value pairs, create text indexes, and perform other manipulations. The Java APIs include an implementation of the Apache TinkerPop interfaces. Support for scripting languages, such as Groovy and Python is included. Groovy scripts allow developers to test Java code snippets more easily without defining objects or compilation.

Fast search with text Indexing

Fast search is enabled through Apache Lucene and optionally Apache SolrCloud. They provide text indexing on properties for fast retrieval of vertices and edges. Native Oracle Text indexing is supported; text queries are automatically translated into SQL SELECT statements with a "contains" clause.

Apache Lucene

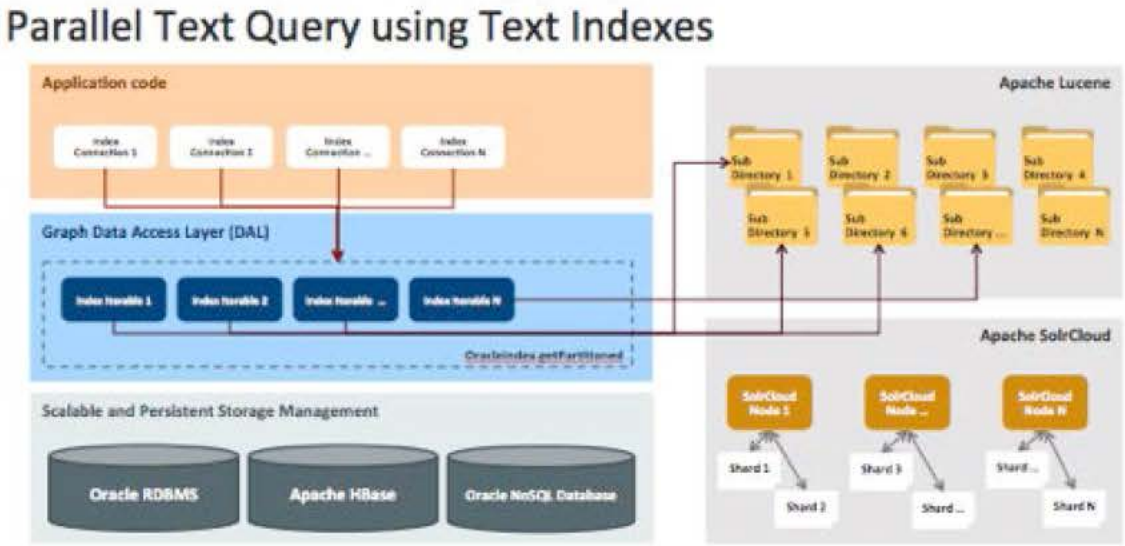
Text searching and indexing is enabled through Apache Lucene for queries on property graph elements. For example, if you have Twitter feeds you can query for tweets containing the word "Oracle" using usual search syntax. Both manual (selective) and auto indexing of graph elements are supported. A limitation of Apache Lucene is that its indexes can't be shared directly among users and applications.

Auto index is easy to use. It is a b-tree index that is created by specifying the index name and what to index, vertices or edges. It is automatically updated as graph elements are changed.

Manual (selective) index is a more flexible manual process. Index content is determined and loaded by the developer. This provides the flexibility to choose what graph elements will be indexed, making the index more focused.

Apache SolrCloud

Faceted search is a powerful capability of SolrCloud. Multiple users and applications can share a SolrCloud index. SolrCloud is implemented with Lucene. The SolrCloud load balancer enables multiple shards on different servers with index replication. Documents are sent to the HTTP interface, sharded, and the index is replicated if that is enabled. The query coordinator transparently queries the shards and assembles the final results.



Using Apache Lucene and Apache SolrCloud

Oracle Text indexing and search

Automatic text indexing using Oracle Text is supported. Oracle Text uses standard SQL to index, search, and analyze text values stored in the property columns of the vertices and edges tables. Oracle Text indexes all the existing K/V pairs in the property graph.

Fast, parallel bulk loading

Fast, parallel bulk loading of very large graphs is accomplished with an easy to use, data type-rich Oracle flat file format and Oracle SQL*Loader. The data can be loaded into multiple database partitions. A utility is provided to easily convert Oracle tables and comma separated values (CSV) files into flat file format. The open source graph file formats GraphSON, GraphML and GML are also supported.

Spatial filtering to enhance graph analysis

Spatial filtering in graph queries can enhance graph analysis. A spatial geometry, such as coordinates for an address can be stored as a property and analyzed; for example, a "within distance" query can determine whether to consider the associated entity in further analysis. Support for point, line and polygon geometries and function-based spatial indexing, and access to spatial analytic functions make this a powerful feature.

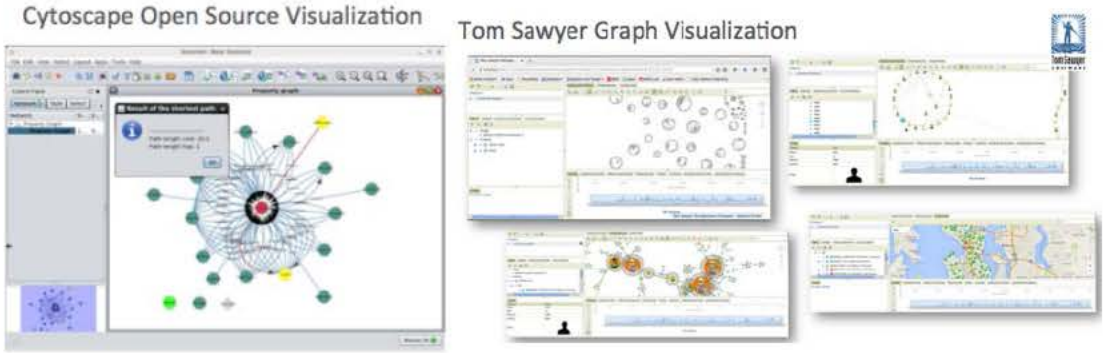
Multi-level security

Multi-level security can be enforced with graph level access control as well as optional use of the Oracle Label Security option for fine-grained access control to individual vertices and edges

Applying Oracle Label Security to graph elements

Graph visualization

The property graph feature supports open source and commercial graph visualization through Cytoscape and Tom Sawyer Perspectives, respectively.



Graph visualization using Cytoscape and Tom Sawyer Perspectives

Spatial Features Overview

Vector Performance Acceleration

Vectors are 2-dimensional and 3-dimensional sets of vertices, for instance latitude, longitude and height that describe geometries, such as points, lines, polygons, surfaces and solids. Geometries often represent real world objects. Vector operations evaluate spatial relationships between geometries, including within-distance, nearest-neighbor, and geometry interactions, such as touch, overlaps, contains, covers, distance, and buffer zone generation around geometries.

Oracle Spatial and Graph vector acceleration capabilities substantially improve the performance of vector operations. Vector acceleration provides enhanced computational algorithms along with CPU and memory enhancements that improve the performance of spatial index creation, geometric computations in functions and spatial operator secondary filter operations. Oracle Spatial and Graph vector performance acceleration builds on general improvements in Locator described the section below "Locator Operator and Function Improvements."

Spatial features support a broad range of applications

Vector Geometry Functions

Oracle Spatial and Graph provides over 400 functions to perform calculations on geometries, such as area of a polygon, length or perimeter. These functions are used, for example, to determine the total area of all counties around a given county, the length of an interstate highway, or the length of a provincial border.

Other functions generate new geometries such as buffers, unions, intersections, and much more. They can be used, for example, to define sales regions by creating a 5 mile buffer around all sales offices, identify the new geometry representing the union of two sales regions, or find the intersection between two sales regions.

Other functions include interior point, concave hull, and generation of triangulated irregular networks through Delaunay triangulation. Cross-endian operations for transportable tablespaces are also supported.

Whole Earth Geometry Model for Geodetic Coordinate Support

The Whole Earth geometry model takes into account the curvature of the Earth's surface when performing calculations on geodetic data. Thus, Oracle Spatial and Graph functions return accurate lengths and areas for both projected and geodetic data. It supports over 30 of the most commonly used distance and area units, including foot/square foot, meter/square meter, kilometer/square kilometer.

Projections and Coordinate Systems

Oracle Spatial and Graph provides comprehensive tools for managing coordinate systems and projections to represent and integrate spatial information effectively and accurately. Over 4000 commonly used mapping coordinate systems are supported; users can also define new coordinate systems. Oracle Spatial and Graph also provides support for implicitly and explicitly transforming data between different coordinate systems – it enables explicit map projection transformations of vector objects from one coordinate system to another. These transformations can be on a geometry-level basis or an entire layer at a time.

Coordinate systems support is based on the European Petroleum Survey Group (EPSG) data model and data set. Although created by the oil and gas industry, this industry model provides benefits of standardization, expanded support, and flexibility for all industries, georaster data vendors, and GIS users in general.

Oracle Spatial and Graph supports 3D coordinate systems, which include height or a "z" coordinate, in addition to longitude and latitude or projected x, y coordinates as appropriate; reprojection of rasters is also supported.

Spatial Aggregates

SQL has long had aggregate functions, which are used to aggregate the results of a SQL query. Oracle Spatial and Graph aggregate functions perform a specified aggregate operation on a set of input geometries, and return a single geometry object. For example, the following statement returns the state boundary of Tennessee generated from all of the counties in Tennessee:

```
select sdo_aggr_union(sdoaggrtype(geom,0.5)) state

from geod_counties

where state_abrv='TN';
```

Other supported aggregate functions include, union, centroid, and convex hull; users can also define other aggregate functions. The use of spatial aggregates improves performance and simplifies coding.

Linear Referencing Support

Oracle Spatial and Graph can store and associate attributes and events with a specified segment on a linear geometry. Attributes and events are stored in tables separately from the geometry, and the geometry does not have to be duplicated in the attribute tables. Linear referencing is often used in the transportation, utilities, and telecommunication industries.

Functions to manipulate linear referenced geometries are also included, such as locating points along a linear feature, clipping a piece of a linear feature (dynamic segmentation), snapping to the closest point of a linear feature of a given point, and conversion between standard and linear referenced geometries. Oracle Spatial and Graph Linear Referencing System functions support 3D geodetic data.

3D Data Type Support

Oracle Spatial and Graph provides native storage, querying, and retrieval for 3-dimensional (3D) data, including points, lines, surfaces, triangulated irregular networks (TINs), an alternative to rasters, and point clouds. Spatial R-tree indexing and SQL operators and analysis functions for 3D data are also provided.

Very large 3D datasets such as urban models, point clouds, and terrain models can be stored and managed in the open Oracle 3D data types, with security, scalability, and high performance. 3D datasets are often found in urban planning and design, government, homeland security, military, oil and gas exploration, transportation engineering, gaming and simulation, geo-engineering, medical applications, business intelligence (for example, real estate and advertising), and LIDAR-based map production.

Oracle Spatial and Graph provides a modeling, visualization and simulation infrastructure for 3D data. A set of metadata tables describes themes, scenes, textures, viewpoints, light sources, non-geographic data, and other elements used to visualize 3D content. This metadata support enables a consistent way to combine all 3D, raster, vector, and non-geometric data into a unified visualization framework. Information may be logically grouped into themes to simplify the development, analysis, use, and maintenance of 3D applications.

Oracle Spatial and Graph has rich 3D and Point Cloud analysis and visualization. 3D parametric (freeform) curves are supported. The Java API in-memory functions support 3D projected and geodetic geometry types. The distance between two 3D points approximates and accounts for the height of those points in the calculation. 3D visualization

and analytical tools can take advantage of metadata views for 3D themes, scenes, and view frames. Pyramiding is supported for Point clouds and TINS to support visualization at different levels of granularity. Contour lines can be generated from point clouds. 2D points can be projected onto a TIN to determine point heights. Linear Referencing System functions support 3D geodetic data.

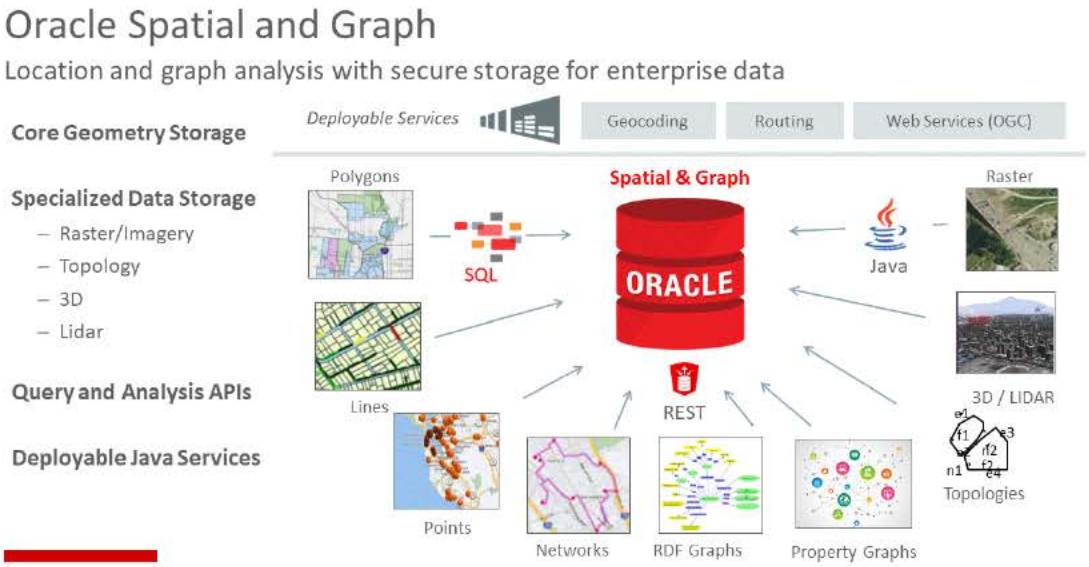
Parametric Curve Support

Oracle Spatial and Graph supports 2-dimension and 3-dimension parametric curves, also called non-uniform rational B-splines (NURBs). The Oracle Database spatial type, SDO_GEOMETRY now supports mathematically precise representation of freeform curves that can be reproduced exactly. NURBs are used to simplify the design and modeling of roads, highways, and rail.

Topology Data Model

Oracle Spatial and Graph includes a data model and schema that persistently stores topology in 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. Topology-based queries typically perform faster than the alternative for relationships such as adjacency, connectivity, and containment. Land management (cadastral) systems and spatial data providers benefit from these capabilities.

Application developers and DBAs can version topologies stored in the Oracle Spatial and Graph topology data model using Workspace Manager, a feature of Oracle Database. Feature level spatial transactions against a persistent topology in the database are supported. A feature insert or update occurs as a single operation, simplifying the process of updating and maintaining topology datasets, and simplify application logic.



Spatial and graph data types and models in Oracle Database

GeoRaster Support

Geometries can be represented by vectors or rasters, or both. Image processing systems typically refer to raster data as images, for instance in satellite imagery and airborne photographs. Raster data used in GIS is normally called gridded data. Oracle Spatial and Graph GeoRaster can store, index, query, analyze, and deliver raster image and gridded data and its associated metadata.

GeoRaster stores multidimensional grid layers and digital images that can be referenced to positions on the Earth's surface or to a local coordinate system. What differentiates GeoRaster is the ability to perform raster analysis on extremely large images and data sets, provide in-place image processing and analysis with no development required, and provide parallelized image processing with simple invocation of PL/SQL procedures.

GeoRaster is used with data from any technology that captures or generates raster data and images, such as remote sensing, photogrammetry, and geospatial thematic mapping. It is used in a wide variety of application areas, including environmental monitoring and assessment, geological engineering and exploration, natural resource management, defense, emergency response, telecommunications, transportation, urban planning, and homeland security.

GeoRaster is designed to deliver enterprise-class data management capability to large image processing and GIS solutions and business applications. Developers can integrate this powerful data management technology with the leading image processing and raster/grid analysis tools and various business applications.

Native Support for Raster File Formats, Compression, Core Operations

GeoRaster loading and native storage is flexible, cost-effective and performant. The file formats GeoTiff, JPEG 2000, and Digital Globe RPC are supported for loading and exporting GeoRaster objects. GeoRaster associates location with the geometries in a raster by assigning location values to a matrix of cells that cover the raster and storing the cells as an array. Native JPEG 2000 compression provides high compression ratios and image quality. JPEG files can be loaded without decompression. Oracle SecureFiles provides transparent lossless compression. Alternatively, GeoRaster-specific industry standard image compression techniques, including JPEG baseline (lossy) and DEFLATE (lossless) can be used and an open plug-in architecture allows additional third party compression techniques. Automatic blocking size optimization chooses a block size for georasters that minimizes storage while optimizing retrieval and processing. GeoRaster supports raster image and raster data at differing sizes and degrees of resolution through pyramiding, and very large images with tiling. Three types of interleaving are supported to optimize access to the data in the raster.

Oracle Spatial and Graph supports relational raster data tables (RDTs) to allow users to specify default alpha channel and pyramid level in its metadata. It includes a resampling algorithm that supports resolution unit specification and parallel processing in many operations, and adds additional loading and exporting capabilities.

Raster Algebra and Analytics

Oracle Spatial and Graph supports raster algebra operations that work on individual raster cells, or pixels to generate new maps from two or more raster layers. Raster algebra operations enable applications to implement sophisticated analytical algorithms, such as a Normalized Difference Vegetation Index (NDVI), and Tasseled Cap Transformation (TCT). Raster operation performance is also substantially faster and can be parallelized to scale up to 100s of times faster for large data sets. Statistical analysis functions can dynamically compute complete statistical values for a GeoRaster object or individual statistical values. Image classification, time series analysis, and raster GIS modeling are supported with the capability to merge multiple bands or layers of different GeoRaster objects into a single GeoRaster object.

Fast, Sophisticated Image Processing

Oracle Spatial and Graph GeoRaster provides advanced image processing and serving capabilities. These include Ground Control Point georeferencing, reprojection (to over 4000 supported coordinate systems), rectification, orthorectification, image scaling, stretching, masking, filtering, image segmentation, NDVI computation, Tasseled Cap Transformation, image appending, bands merging, large-scale advanced image mosaicking, and virtual mosaic. The operations described here are most commonly used to process and serving geospatial images, particularly raw satellite imagery and airborne photographs.

More image processing can be handled in the server instead of the client, and some processing is parallelized. This enables improved performance of image processing on a much larger scale, with larger data sets, which are increasingly being used in government and commercial applications as more raster data becomes available. Customizable memory control further improves performance.

GeoRaster uses industry standard resampling and interpolation methods for image and raster transformations and operations. Transformations between 2D or 3D ground coordinates and 2D cell coordinates, and vice-versa are supported. Non-rectified images (not geometrically corrected for uniform scale) can be georeferenced with GeoRaster's flexible functional fitting polynomial georeferencing model. Irregular polygon-based clipping in queries returns a precise subset of a GeoRaster object. Grid point interpolations infer values at spatial positions between or within cells. Irregularly shaped regions inside an image can be defined with bitmap masks.

Easy and Manageable Administration

GeoRaster provides ease of development, ease of use, and manageability. GeoRaster DML triggers are created and monitored by the system automatically. Users can monitor resource-intensive operations on GeoRaster system data. Partial raster updates are supported. GeoRaster templates are supported to develop GeoRaster applications, such as extraction, transformation, and loading (ETL) tools and image processing systems that work with GeoRaster objects. Raster data versioning with the Workspace Manager feature of Oracle Database, and raster data row-level security with Oracle Label Security are supported.

Java API

A Java API supports query, manipulation, and raster management. It also supports the development of ETL tools, Web applications, and raster processing applications – simplifying the development of Java applications that use, access, and manipulate raster and gridded data sets stored in Oracle Database. It supports features such as ground control point (GCP) storage and manipulation, GCP georeferencing, reprojection, grid interpolations, and getCellValue. The Java API supports all server-side functions and procedures and includes a Virtual Mosaic API.

Loading and Exporting Raster Data

The GeoRaster data type is supported by all leading third party GIS and image processing tool vendors. Geospatial Data Abstraction Layer (GDAL), the leading open source geospatial ETL tool and API available for raster data also supports it. GDAL natively supports importing and exporting over 50 raster formats to and from SDO_GEORASTER. GDAL is a high performance C++ tool that supports large file sizes. It includes C/C++, Java, and Python APIs to access GeoRaster; and utilities to translate raster formats, warp rasters, generate contours from DEM rasters, and many other raster operations.

GeoRaster includes an ETL wizard tool to automate and enable concurrent batch loading and exporting of various image and raster files using GDAL. The tool can load and export large numbers of raster and image files in batches, and do so concurrently. This tool supports all raster formats supported by GDAL.

For more information about GeoRaster, please refer to separate white paper on the Oracle Technology Network, Oracle Spatial and Graph page.

Spatial Analytic Functions

Oracle Spatial supports spatial analysis and mining in Oracle Data Mining (ODM), a component of Oracle Advanced Analytics. ODM allows automatic discovery of knowledge from a database, such as discovering hidden associations between various data attributes, classification of data based on some samples, and clustering to identify intrinsic patterns.

Spatial data can be materialized for inclusion in data mining applications. Data at a specific location is often influenced by data in the neighborhood. The spatial analysis and mining features in Oracle Spatial and Graph let users exploit such spatial correlation in the following ways:

- binning data into regions – determine if southeastern US customers in a certain age or income category are more likely to prefer regular or diet soda
- materializing spatial correlation (neighborhood influence) – examine the values of similar houses in a neighborhood when assessing the value of a house
- colocation mining – determine if locating a pizza restaurant franchise with a video store results in higher sales
- spatial clustering – determine the regions where crime rates are high, decide where to deploy additional police
- location prospecting – identify the best locations for opening new hospitals based on the population of patients who live in each neighborhood

For more information about spatial analytic functions, please refer to the white paper on the Oracle Technology Network, Oracle Spatial and Graph page.

Geocoding

Geocoding is the process of associating geographic references, such as addresses and postal codes, with location coordinates (longitude and latitude). Oracle Spatial and Graph offers full geocoding capabilities. It provides international address standardization, geocoding points of interest (POI) and matching them to geocoded data stored in Oracle Database, reverse geocoding, batch geocoding, and other geocoding capabilities. Its unique unparsed address support adds great flexibility and convenience to customer applications. SQL, Java, and XML APIs for geocoding are provided, and it can be deployed either at the middle tier (Oracle Fusion Middleware) or at the database server tier.

Sample data is available online. Data sets in the format supporting Oracle Spatial and Graph are also available from leading data providers. For more information, please visit the Oracle Technology Network for Oracle Spatial and Graph and navigate to the Partners tab.

Oracle Spatial and Graph geocoding supports standard address geocoding based on interpolation and point-based geocoding where data sets include the exact location of addresses, intersections, and points of interest. Point-based geocoding is becoming increasingly popular because it allows for more accurate results and can be used in situations where interpolation is not possible.

Oracle Spatial and Graph geocoding includes point address geocoding support for countries that don't have address ranges and language support for countries that have addresses in multiple languages. Reverse geocoding can be performed without specifying a country code.

Routing Engine

A routing engine provides driving distances, times, and directions between addresses or between locations that have been geocoded in advance. The Oracle Spatial and Graph routing engine is provided as a Java client library

that can be easily deployed in JEE servlet containers. It supports preference for either fastest or shortest routes, and provides summary and detailed driving directions, and the time and distance along a street network from a location to multiple destinations. It also provides driving distances, times, and directions between addresses for over a dozen Western European countries, including Germany, United Kingdom, and France to support logistics, transportation, and location-based services applications.

Sample data is available online. Data sets in the format supporting Oracle Spatial and Graph are also available from leading data providers. For more information, please visit the Oracle Technology Network, Oracle Spatial and Graph page and navigate to the Partners tab.

The routing engine provides driving directions in Western European languages, including German, French, Spanish, and Italian. Turn-specific geometries can be generated for location services applications that require turn-specific point of interest data. Computed routes can be returned as a set of relationships between points that can be used for further analysis.

The Oracle Spatial and Graph routing engine supports restrictions and conditions that are required for advanced routing applications, such as truck-specific routing. It can provide truck-specific routing based on roads, weight, height, time of day, and other conditions applied to commercial and logistics applications and logical turn restrictions. It can compute the drive times based on truck speed limits, which often differ from car speed limits. It can also provide information on truck services, such as weigh stations and truck stops along a route. Finally, it can handle logical turn restrictions involving more than two way points in the route geometry. These enhancements yield more accurate results for logistics and truck routing applications.

Routing Engine Enhancements Introduced in 18c

A routing engine servlet supports lightweight location-based queries related to speed limit and traffic speed.

Spatial Web Services[JSI1]

Oracle Spatial and Graph provides a web services platform to access, incorporate, publish, and deploy geospatial services, including services for geocoding, routing, mapping, business directory, catalog, and geospatial feature transactions. Oracle Spatial and Graph is tightly integrated with Oracle Database and Oracle Fusion Middleware to provide a transactional service-oriented architecture platform with enterprise-class security. They provide security, including authorization, authentication, and transport confidentiality and integrity.

Oracle Spatial and Graph has supported the Open Geospatial Consortium (OGC) and ISO TC211 standards for many releases. It supports these XML-based geospatial web services standards: OGC OpenLS 1.1, Web Feature Service – Transactional (WFS-T) 1.0, Web Feature Service 1.1.0, Catalog Service for the Web (CSW) 2.0.2, and Web Coverage Service (WCS) 2.0.1 on a variety of client technologies and platforms.

Web Feature Services (WFS) enable query and retrieval of geographic feature information in vector format, encoded in GML. Oracle Spatial and Graph includes full support for database transactions on WFS-T feature tables through SQL without restriction. It also supports Workspace Manager versioning and WFS feature tables. Java and PL/SQL client APIs are provided.

Catalog Services (CSW) are used to locate, manage, and maintain distributed geospatial data applications, and services. Web Coverage Service (WCS) support enables retrieval of coverages, or raster data such as satellite imagery or DEMs.

Oracle Spatial and Graph provides a unified framework and Web-based administrative console for WFS, WCS, and CSW, for easy deployment, administration, and diagnosing. This menu-driven user interface simplifies registration of spatial layers. It allows users to browse existing spatial layers, eliminating the need for a DBA to run PL/SQL scripts

to publish spatial layers. It includes a tutorial on configuring and using WFS, and includes a sample request and response page for WFS queries. The user interface can also be used as a client to other WFS servers.

Web Services Enhancements Introduced in 18c

WFS, WCS, [JSI2] and CSW can support reading and publishing data from multiple data sources in the same WebLogic Server instance, making it easy to scale web services applications. The administration console user interface allows users to manage multiple data sources for WFS, WCS, and CSW.

A new metadata application profile defined by ISO for OGC CSW 2.0.2 is now supported. This enables easy exchange of spatial data and metadata among applications and organizations. Oracle Spatial and Graph CSW services can interoperate with other services and query languages that also implement the ISO profile.

XML full text indexes can be created on metadata profile data, enabling XQuery full text queries that search on XML documents efficiently by combining text and structured search. This improves CSW query performance significantly.

Network Data Model Graph Features Overview

Network Data Model (NDM) stores physical and logical network data structures commonly used in transportation, utilities, and oil and gas. It explicitly stores and maintains network connectivity and provides network analysis capability, including shortest path, nearest neighbors, within cost, and reachability.

NDM has a PL/SQL API for managing network data in the database, and a Java API for performing network analysis and creating and applying network constraints.

Oracle Spatial and Graph NDM users can benefit from the speed of in-memory analytics with networks that are larger than available memory. NDM supports partitioning large networks into manageable sub-networks and can automatically load network partitions into memory as needed for efficient in-memory analysis. Partitioning utilities are also available.

NDM is integrated with Oracle Spatial and Graph geocoding and the routing engine; applications using these features can perform analysis using NDM functions. NDM supports commercial street network data from Nokia (Navteq), in Oracle Delivery Format (ODF).

NDM includes a number of modeling and analysis features that meet the requirements for utility networks, logistics, transportation, and other network-based applications.

Modeling features

- Model and represent any point along a link for all analysis functions, such as specific addresses in street networks with any number of properties on the nodes and links.
- Model partial-link paths (subpaths).
- Customize link and node properties: e.g., costs.
- Perform path analysis with multiple link and node properties (e.g., distance/time/hops costs).
- Perform partitioning of logical networks, for example social and biochemical pathway networks, based on metrics appropriate to the application.

Network analysis features

- Compute the shortest route connecting a given set of nodes.
- Generate a polygon representing the region that can be reached from a given node with a specific cost. A typical application is the generation of drive-time and drive-distance polygons.
- Generate the shortest path on a hierarchical network, where links are prioritized by property (e.g., highways, local roads), to support queries such as finding the route between two addresses that favors highways over local roads as much as possible.

» Compute a buffer based on network cost; the buffer representation contains coverage and cost information.

» Compute K shortest paths between two nodes.

NDM includes an example with JSP and Java files for application developers to quickly and easily deploy routing and other network analysis using data stored in NDM. Users can visualize analysis results in a web browser. The example works with Nokia ODF network data, and uses the NDM load on demand API, Oracle Fusion Middleware MapViewer, and the Oracle Spatial and Graph geocoding engine. Download the example from the Oracle Technology Network at oracle.com/technetwork/database/options/spatialandgraph/spatial.

Oracle Spatial and Graph NDM feature and time modeling enhancements simplify application development and support real-world analysis by moving complex spatial logic into the database.

Feature Modeling, Analysis, and Editing

Oracle Spatial and Graph NDM simplifies feature editing and analysis by providing a feature analysis function that associates feature representations with network elements. Feature modeling bridges the gap between concrete objects of interest in real world and abstract network elements.

Feature modeling simplifies application development by associating real world objects with network elements. For example, if a utility network application needs to find affected households when a substation experiences a power failure, it is necessary to associate the application features (substations, power lines and transformers) with network elements (links and nodes). Feature modeling maintains these relationships through feature metadata, simplifying application development and maintenance.

Network feature editing (NFE) lets you create and manage an NFE model. An NFE model extends the feature modeling capabilities by enabling you to visualize and manipulate features using Java Swing components and a PL/SQL API. You can also define features on the top of an existing network.

Network Modeling with Time; Multimodal Transportation Routing

Oracle Spatial and Graph adds support for modeling networks having a time dimension. Users may associate time properties with nodes and links, and specify temporal inputs in network analysis queries.

Most real-world networks have a time element. Travel times on road segments vary with the time of day. Utility networks experience different demand loads based on seasonal demand and the time of day. Analytic and planning applications can benefit from more accurate representation of real-world conditions. NDM supports queries such as finding the fastest travel route for a specified time of day. NDM supports modeling and analysis of multimodal transportation networks, and computing the fastest paths on multimodal transportation networks.For more information about Oracle Spatial and Graph Network Data Model, please refer to the white paper on the Oracle Technology Network, Oracle Spatial and Graph page.

GeoJSON Support

Oracle Spatial and Graph supports the use of GeoJSON objects to store, index, and manage geographic data that is in JSON (JavaScript Object Notation) format, directly in Oracle Database. GeoJSON is an open standard format designed for representing simple geographical features, with their non-spatial attributes. It is based on JSON, a lightweight data interchange format that has become a standard for reading and publishing data in web, Big Data, and Internet of Things environments.

You can convert data from GeoJSON objects to Oracle Spatial SDO_GEOMETRY objects, and from SDO_GEOMETRY objects to GeoJSON objects. You can use spatial operators, functions, and a special SDO_GEOMETRY method to work with GeoJSON data. If the application requires a JSON data store, GeoJSON data can be embedded into that JSON store; these JSON documents can then be spatially indexed and used in

spatial queries. The GeoJSON features leverage capabilities in Oracle Database to store, manage, and index JSON documents. Developers can access JSON through REST services or other APIs, and use SQL to query JSON documents, providing flexible application development and powerful SQL analytics for modern development environments.

Enhancement for 18c: JSON support is expanded to support a larger range of geometry types and coordinate systems beyond those included in the GeoJSON standard, including 2D and 3D, solid, surface, and LRS geometries. You can convert these additional geometry types between JSON and SDO_GEOMETRY formats.

Oracle REST Data Services (ORDS), a Java application that allows developers to easily create RESTful services with Oracle Database, now supports spatial operations. See **"New Spatial Features in Oracle Database 18c"** section for a description.

Location Data Enrichment

Oracle Spatial and Graph includes a place name data set, with hierarchical geographical data from HERE, that you can load into the database and search using the new SDO_UTIL.GEO_SEARCH function. The data set includes commonly used textual location data such as place names, addresses and partial addresses, and latitude and longitude information.

Location tags are extracted from customer text data, matched with well known place names using Oracle Text, and enhanced with other geographic information associated with the well known place names. The results can be stored as additional attributes with the original data. This feature enables you to process less structured geographic and location data so that the information can be categorized, compared, filtered, and associated with other data. For example, data with only partial names can be enriched to include city, county, state, and country, allowing it to be joined or analyzed with other data sets that may have state level information. This is especially useful when comparing Big Data results with structured information in operational systems and data warehouses.

Spatial Index Improvements

Spatial indexes support all partitioning methods, including list, interval, and hash partitioning; previously, this was restricted to range partitioning. This enables improved performance.

Additionally, Spatial indexes can be system-managed. The main benefit is simplified spatial index management. This is most beneficial in cases of partitioning, because this new index type eliminates the need for most, if not all, index partitioning management operations. Full support is provided for almost all Oracle Database base table partitioning models. Customers are strongly encouraged to use this index type for all new spatial indexes created, regardless of whether the spatial table or the spatial index is partitioned.

Index Type Optimized for Point Data

Point only data sets commonly used for location services, such as moving objects, point of interest, and yellow page data, can be massive, and applications often require fast update and query performance. Performance using R-tree indexes can be challenging when large numbers of concurrent DMLs are performed on spatial tables.

An alternative index type, the composite B-tree index, can significantly improve the performance of spatial index creation and DML operations for large volumes of updates for point only data sets, with 20 – 30 times faster index creation, and 10 times faster updates. **Enhancement for 18c:** Performance for DML operations and queries using composite B-tree indexes is 3 – 4 times faster compared to the 12.2 release.

Map Visualization of Geographic Data

The spatial visualization feature is an HTML5-based mash up component that allows developers who are familiar with SQL and JavaScript to incorporate a wide variety of map styles and spatial analysis into business applications. Its mapping engine visualizes data in Oracle Spatial and Graph, and allows developers to combine this data with external web services such as WMS, WFS, GeoRSS streams, and WMTS. Developers can define their own third party map tile layer using a plug-in interface, and there is support for third party GeoJSON files. Dynamic tile layers enable large data sets to be visualized in an interactive manner, with the ability to explore individual features if needed.

These capabilities, previously available in Oracle Fusion Middleware MapViewer, are now also available to Oracle Spatial and Graph users. The map visualization feature must be deployed in a JEE container or in Oracle Java Cloud Service.

Location Tracking Server

The Oracle Spatial and Graph location tracking server enables you to define regions of interest, track the movement of objects into or out of those regions, and receive notifications when certain movements occur. As location becomes an increasingly important aspect of our lives, and as location-sensing devices become ubiquitous, there is an increasing demand for applications to be able to monitor subscriber location data continuously. The monitoring of the location data may translate into alerts being generated in the system.

For example, a trucking company may want to monitor its network of 10,000 trucks as they move along their specified routes towards their destinations. They may want to track the movement of trucks within a specified range of the route and expect notifications to be generated to detect undesirable deviations the vehicles from their desired routes. Proactive location-based services (LBSs) generalize such applications that track locations of subscribers inside or outside a specified region of interest for various purposes, such as location-based advertising and notifications about friends nearby.

The Oracle Spatial and Graph location tracking server provides a simple framework for setting up a location tracking network within the database through a PL/SQL interface. An API is provided for continuous location monitoring of objects within a tracking network. A queuing mechanism handles incoming location updates and tracking requests and for outgoing relevant notifications, using Oracle Advanced Queuing. This delivers efficient, continuous location monitoring for thousands of relevant objects within the database.

New Spatial Features in Oracle Database 18c

Spatial Support for Distributed Transactions

In a distributed database, portions of the database are stored in multiple physical locations. Distributed transactions update data among multiple database nodes. The use of spatial R-tree indexes is now supported in distributed and Oracle XA transactions, which will be useful for several cloud-based applications that use distributed web architectures.

Oracle REST Data Services Support for Spatial Operations

REST is today's dominant software architectural style for creating modern, scalable web services for cloud environments. Oracle REST Data Services (ORDS) is a middle tier Java application that allows developers to easily access data in Oracle Database through REST, and turn them into RESTful services. ORDS maps HTTP uniform resource identifiers to a query or object in Oracle Database, runs the appropriate SQL, and returns output as JSON.

ORDS supports Oracle Spatial types and operations, so using REST, SQL queries can be issued on SDO_GEOMETRY data, which is automatically converted to JSON.

ORDS is included with both Oracle Database and Oracle SQL Developer installations. It is supported in WebLogic, Tomcat, Glassfish, and also as a standalone application running Jetty in embedded mode.

JSON support for spatial data has also been enhanced in 18c. See "**JSON Support**" section.

Spatial Index Now Optional for Spatial Operations

Creating and using a spatial index is no longer required for the use of any spatial operators. Spatial indexes are generally recommended, and optimize performance. However, in some cases queries may need to be written against views where the geometry data is constructed on the fly, and spatial indexes cannot be easily built. Spatial SQL queries will work against such data even without a spatial index.

Other new spatial features in 18c include expanded JSON support, spatial support for sharded databases (see "**Enterprise Features Supported in Oracle Database 18c**" section), spatial web services enhancements, and more. See Appendix 2 for a full list of new 18c spatial features.

RDF Semantic Graph Features Overview

The RDF Semantic Graph feature of Oracle Spatial and Graph is a special purpose graph for linked data and semantic web applications conforming to World Wide Web Consortium standards. RDF and OWL are standards for representing and defining complex, semantically related data and SPARQL is a query language designed specifically for graph analysis. Application developers benefit from the industry's leading open, scalable graph data platform integrated with Oracle Database for scalability, security, performance and high availability.

Storing, Loading, and Data Manipulation

RDF Semantic Graph has proven scalability to over one trillion triples (LUBM benchmark). It supports all standard database loading, storing, and data manipulation operations on RDF/OWL models. Each RDF model contains a set of subject – object – relationship triples organized as a graph of directed, labeled edges. The edge is the link (or relationship) that connects a subject node to an object node and is labeled by a predicate (property). Space-efficient storage saves up to 60% disk space for scalable and performant loading, querying, and inferencing.

Native Inferencing

RDF Semantic Graph has native, forward-chaining, persistent inferencing using any combination of RDF, RDFS, and OWL 2 RL and EL profiles, as well as user-defined rules for specialized inference capabilities. It also provides a plug-in framework to support third party specialty reasoners. Optimizations include optimized large owl:sameAs sets, incremental inference to update entailments after triple inserts, and parallel inference on multi-core or multi-CPU architectures.

Oracle Spatial and Graph supports ladder-based inferencing that ensures newly inferred triples are labeled with the proper security and user-defined inferencing has been added.

Querying RDF Graphs in Oracle Database

RDF Semantic Graph provides SPARQL 1.1 endpoint web services and Java APIs via open source Apache Jena / Joseki and Sesame. RDF/OWL data can also be queried using SQL; the Oracle SQL SEM_MATCH table function embeds SPARQL graph pattern queries in a SQL query. A virtual model capability provides a view-like feature to combine models for querying.

The SEM_MATCH table function in Oracle Spatial and Graph supports SPARQL 1.1. The OGC GeoSPARQL standard is supported for storing and querying spatial data in an RDF graph.

Viewing Relational Data as RDF Triples

RDF views can be created on relational tables, views and SQL query results in Oracle Spatial and Graph . The W3C specifications for automatic mapping (called Direct Mapping), custom mapping (using the W3C R2RML language) and materializing views are supported. RDF views present relational data in RDF triple format so it can be queried using SPARQL and connected with other linked data and RDF graphs to relate and facilitate enterprise data integration.

RDF, XML, JXON and Relational Interoperability

Ontology-assisted SQL querying allows SQL queries to extract more semantically complete results from table data by associating relational data with ontologies that organize the domain knowledge of the data.

Oracle Spatial and Graph provides a SPARQL Gateway feature that presents SPARQL query results in XML format for visualization tools that support XML data sources, such as Oracle Business Intelligence (OBIEE). RDF query results can be returned in the JXON interoperability format.

Fine-grained Security

Model-level access control is the default for RDF graph data. Triple-level security using the Oracle Label Security Option is also supported for the most stringent security levels. Sensitivity labels can be defined on individual triples and users to conditionally restrict a user's access to individual triples stored in an RDF model.

Graph Analytics

Oracle Spatial and Graph supports SPARQL 1.1 property path expressions to find graph patterns across any length path. Results from graph queries can be used with Oracle Advanced Analytics Data Mining and Oracle R Enterprise.

Semantic Indexing for Documents

Semantic indexing for documents provides an index type to semantically index information in unstructured documents, table data and URLs extracted by third party natural language processors and annotators. Semantically indexed documents can be searched using SEM_ CONTAINS operator within a standard SQL query. The search criteria for these documents are expressed using SPARQL query patterns that operate on the information extracted from the documents.

Advanced Performance and Scalability

RDF Semantic Graph supports parallelism, compression, partitioning, Oracle Real Applications Clusters (RAC) and Oracle Exadata Database Machine for enterprise-level performance and scalability. Required licensing for this functionality is: Oracle Database Enterprise Edition, Oracle Spatial and Graph option, and the Oracle Partitioning option.

Other RDF Semantic Graph Features

You can perform SPARQL Update operations on a semantic model. The W3C SPARQL 1.1 Update is supported in Oracle Database semantic technologies through the SEM_APIS.UPDATE_MODEL procedure. Integration with property graphs is enabled by RDF views on property graph data for SPARQL queries and property graph views on RDF data for property graph social network analytics. Queries on RDF data that use SPARQL ORDER BY semantics can use the ORDER_TYPE, ORDER_NUM, and ORDER_DATE columns in the MDSYS.RDF_VALUE$

metadata table for more efficient operations. Specifying the DISABLE_ORDER_COL option in the SEM_MATCH query can disable this capability. A SEM_APIS.SPARQL_TO_SQL function enables SQL translation and execution of a SPARQL query. SPARQL Query functions are provided: orardf:like, orardf:sameCanonTerm, orardf:textScore. Integration with Oracle Flashback Query is supported.

Oracle Exadata Database Machine

Engineered systems provide high performance, high bandwidth, and massive parallelism with enormous capacity to address the challenges faced by high volume workloads. Combining Oracle Spatial and Graph advanced analysis with Oracle Exadata Database Machine performance and scalability delivers an ideal platform for the most demanding applications.

Oracle Spatial and Graph fully uses the balanced hardware and highly parallelized architecture of Oracle Exadata. It can achieve performance results over 150x faster than other spatial database machines and solutions. Real customer scenarios, test results, and strategies have maximized performance for massive spatial and graph computations and data ingest.

Oracle Spatial and Graph features are engineered to natively leverage the parallelism, partitioning, indexing and the scalability features of Oracle Exadata without application changes. The fully-parallelized joins and aggregations of Oracle Exadata coupled with the extreme I/O bandwidth and high performance of Exadata Storage server provide Oracle Spatial and Graph with the processing power needed for server-based geo-processing and graph applications. OLTP index compression improves query performance by compressing and increasing memory residence for spatial and graph indexes. Exadata Hybrid Columnar Compression increases memory residence for large spatial and graph data sets as well as the rule sets used in inferencing.

For more information, including customer scenarios, test results, and strategies with Oracle Exadata and Oracle Spatial and Graph, please refer to Oracle Technology Network, Oracle Spatial and Graph page.

Enterprise Features Supported in Oracle Database 18c

Oracle Database 18c provides powerful, reliable support for an organization's mission-critical applications. These enterprise features enrich Oracle Spatial and Graph capabilities via a flexible Internet deployment architecture, object capabilities, and robust data management utilities that ensure data integrity, data recovery, and data security. This level of support can only exist in the homogenous environment of an enterprise database solution, and cannot be effectively replicated in a hybrid solution that marries an external location-based solution with a traditional enterprise solution, no matter how tightly integrated the two components may appear.

Oracle Spatial and Graph takes full advantage of expanded database size limits, high-performance VLDB maintenance, utilities, replication, versioning of geospatial data (Workspace Manager), faster backup and recovery, and partitioning. Only users of the native geospatial and graph data types in Oracle Database can take full advantage of features such as partitioning, replication, parallel spatial index builds and queries, and geospatial and graph-driven multi-level security. The full range of Oracle utilities (e.g. SQL*Loader) are also available to ease migration and help upgrade applications that use the spatial features. Some of these key enterprise features are described below.

Partitioning Support for Spatial Indexes

The Oracle Database architecture includes partitioning, in which a single logical table and its indexes are broken up into one or more physical tables, each with its own index. Spatial indexes associated with partitioned tables can be

partitioned; range partitioning is the partitioning scheme supported for spatial indexes. Starting with the 12.2 release, spatial indexes also support list, interval, and hash partitioning. A graph in Oracle Database is subdivided into models that reside in separate partitions.

Partitioning offers significant performance, scalability, and manageability benefits, including the following:

- Reduced response times for long-running queries; partitioning can reduce disk I/O operations.
- Reduced response times for concurrent queries; I/O operations run concurrently on each partition.
- Easier index maintenance because of partition-level create and rebuild operations.
- Ability to rebuild indexes on partitions without affecting the queries on other partitions.
- Ability to change storage parameters for each local index independent of other partitions.
- Partitions that can be split, merged, and exchanged.

Parallel Index Creation

Spatial and graph indexes and index partitions can be created in parallel. Geospatial R-tree index and graph B-tree 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. Large non-point datasets (commonly used in many standard GIS applications) can show dramatic performance improvements.

Parallel Load, Query and Inference

Spatial queries can 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. RDF graph data loading, graph queries and inferencing operations are also fully parallelised.

Replication

Oracle GoldenGate can replicate data in the native geospatial data types. It supports SDO_GEOMETRY, SDO_GEORASTER, SDO_TOPO_GEOMETRY. Distributed systems that involve geographically dispersed yet logically replicated web sites, can take advantage of synchronized replication of spatial data objects across multiple databases.

Database Workspaces and Long Transactions

Workspace Manager, a feature of Oracle Database, provides a virtual environment (workspaces) that allows current, proposed and historical spatial data values to be managed in the same database. Workspaces can be shared and used to: isolate a collection of changes to production data until they are approved and merged into production; keep a long term history of changes to data; and create multiple data scenarios based on a common data set for "what if" analysis. Workspace Manager is supported by most GIS vendors.

Sharded Database Support **(New in 18c)**

Database tables with Oracle Spatial SDO_GEOMETRY columns and spatial indexes, operators, and functions can now be used with Oracle Sharding, introduced in Oracle Database in the 12.2 release. Spatial tables can be sharded across different databases, and spatial indexes can be created on sharded tables, so that spatial queries get routed to the correct shard.

This allows customers who desire the horizontal scalability and geographic distribution of data enabled with Oracle Sharding to include spatial data types. Applications that run on a sharded database architecture can achieve linear

scalability, extreme data availability and geographic data distribution. Sharding is an architectural pattern popularized by Internet and online companies that need very high scalability and absolute availability.

Support for Oracle Multitenant

Oracle Multitenant is an Oracle Database technology that enables database consolidation without changes to applications. Designed for the Cloud, it allows many databases to be managed as one, while retaining the isolation and resource prioritization of separate databases. The multitenant architecture consolidates multiple Oracle Databases (each referred to as a pluggable database) to run under a single occurrence of Oracle Database software (referred to as a multitenant container database). Architectural separation is enforced between each pluggable database (user data and metadata) and its multitenant container database (Oracle metadata). Pluggable databases are compatible with traditional Oracle Databases not in a multitenant container database.

Oracle Spatial and Graph functions transparently in a multitenant architecture; spatial and graph applications benefit from the efficient administration of one multitenant container database, and the separation and resource prioritization allowed by multiple pluggable databases.

Open Standards

Oracle consistently works to help shape, drive, implement and support the latest open standards in the spatial, location services, and graph database areas. Oracle is a founding and Principal Member of the Open Geospatial Consortium (OGC). Oracle is a World Wide Web Consortium (W3C) member and active contributor and/or editor in various technical working groups, such as the W3C RDF, SPARQL, OWL and RDB2RDF working groups and the OGC GeoSPARQL Working Group.

Multiple versions of Oracle Spatial and Graph comply with the OGC Simple Features Specification for SQL, Revision 1.1, Types and Functions Alternative; OGC OpenLS 1.1, Web Feature Service – Transactional (WFS-T) 1.0, Web Feature Service (WFS) 1.1.0, Web Coverage Service (WCS) 2.0.1, and Catalog Service for the Web (CSW) 2.0.2. Oracle Spatial and Graph also supports the SQL/MM types and operators, as specified in ISO 13249-3, Information technology - Database languages - SQL Multimedia and Application Packages - Part 3: Spatial. Oracle Spatial and Graph operators corresponding to those defined in this standard, as well as the SDO_NN and SDO_WITHIN_DISTANCE operators, can be used on data stored in the SQL Multimedia root type.

Property Graph supports multiple open source Apache projects, including Tinkerpop, Lucene, SolrCloud, Groovy and Tomcat, as well as other open source, such as Eclipse Jetty, Python, and file formats GraphSON, GraphML and GML.

RDF Semantic Graph supports the W3C semantic and OGC GeoSPARQL standards. Supported W3C specifications include: RDF, RDF Schema (RDFS), SPARQL 1.1 query language, OWL 2 (RL and EL profiles) knowledge representation languages for authoring ontologies, the Simple Knowledge Organization System (SKOS) and the RDB2RDF specifications for creating RDF views on relational tables - Direct Mapping (DM) and Mapping Language (R2RML). The OGC GeoSPARQL 1.0 – A Geographic Query Language for RDF Data standard is supported for queries and analysis of geospatial data stored in an RDF graph.

Standards compliance testing for Oracle Spatial and Graph is ongoing, and compliance with more recent versions of standards or with new standards might be announced at any time. For current information about compliance with standards, visit the Oracle Technology Network, Oracle Spatial and Graph page.

Oracle Spatial and Graph Partners

Oracle builds and maintains active partnerships with the leading data providers; systems integrators; and geospatial and graph tool, application, and service providers. Oracle's longstanding commitment to depth and breadth of partnerships provides users flexibility and the widest possible choices. Developers and IT managers can select best of breed tools and applications to meet their industry and organization-specific requirements and rapidly deploy scalable, secure enterprise geospatial and location service solutions. Support from leading data providers and systems integrators, both from the geospatial and enterprise IT domains, provides customers with choices for fast deployment of customized solutions to meet their needs.

All of the most widely used GIS software technologies available in the industry support Oracle Spatial and Graph. The leading geospatial data vendors provide data products in Oracle's spatial format with worldwide coverage, and integrators around the globe provide expertise and experience at delivering Oracle-based spatial solutions.

RDF Semantic Graph features are directly integrated with the leading third party graph tools and applications as well as Oracle Business Intelligence and Oracle Advanced Analytics products. RDF Semantic Graph also supports the leading open source application development framework, Apache Jena, and associated open source tools as well as linked open data ontologies.

For a complete list of partners and links to sample data, free downloads, and other resources, visit the Oracle Technology Network, Oracle Spatial and Graph page.

Conclusion

Spatial and graph analysis is about understanding relationships. The emergence of Internet of Things, Cloud services, mobile tracking, social media, and real time systems creates new challenges to manage the volume of data, but more importantly, to discover patterns, connections, and relationships. To address these opportunities, Oracle Spatial and Graph in Oracle Database 18c includes a wide range of spatial analysis functions and services to evaluate data based on distance, and proximity, and to visualize geospatial patterns on maps and imagery. Also, a powerful general purpose property graph feature enables analysis of inherent relationships between collected data entities for a variety of graph use cases, including making recommendations, finding communities and influencers, pattern matching, and identifying fraud and other anomalies.

Oracle Spatial and Graph provides advanced capabilities for Oracle Cloud and Oracle Database 18c. It addresses the business-critical needs of customers in traditional geospatial domains such as defense and intelligence, homeland security, land management, transportation, as well as a broad range of business domains requiring graph and location technologies, including finance, retail, life sciences, publishing and media companies. Features from native 3D point cloud, raster, topology, and network models to geocoding, routing, web services, JSON and REST interfaces, and map visualization make this a complete, advanced geospatial platform. Oracle Spatial and Graph users include several of the largest organizations in the world that use mission-critical, location-enabled enterprise systems. Customers and partners rely on Oracle to deliver performance, scalability, security, and ease of use for their spatial and graph applications. Oracle Spatial and Graph is supported by all the leading geospatial and location services vendors and systems integrators.

The special purpose RDF Semantic Graph model enables creation of a unified metadata layer for disparate applications that facilitates identification, integration,and discovery. It has been adopted for linked data and social network applications by leading organizations worldwide in the healthcare and life sciences, finance, media and intelligence communities.

With every release since its introduction twenty years ago, Oracle Spatial and Graph has delivered the most advanced spatial and graph data management capabilities to database management systems. With the unmatched enterprise data management capabilities of Oracle Database 18c, it continues to be the world's leading database management platform for enterprise spatial and graph systems, on premises and in the cloud.

Appendix 1: Oracle Spatial and Graph Features

Property Graph Features

- Optimized schema for storage of vertices, edges and k/v properties
- PGQL, a SQL-like Graph Query language
- 40+ powerful parallel, in-memory analytics for social network analysis
- Ease of development with Java and Tinkerpop APIs, Groovy and Python scripting, SQL queries on graph data
- Analytics execute in Java or in a multi-user, multi-graph in-memory analyst server on WebLogic or Tomcat, or in the database using SQL
- Analysis results can flow into a bipartite, filtered, undirected, sorted or simplified edge output graph
- Fast retrieval of vertices and edges using text indexing of properties with Oracle Text, Apache Lucene, or optionally SolrCloud
- Spatial filtering, indexing and analytic functions on spatial data in properties
- Property graph analysis of RDF graphs using property graph views
- Parallel bulk load with optimized Oracle flat file format, rich data types and from relational and CSV data
- Open source formats GraphSON, GraphML and GML supported
- Groovy management console
- Security for graphs and graph elements

Spatial Features

- Vector Performance Acceleration: from 5-900 times faster execution of spatial operations and more efficient use of CPU, memory, and partitioning
- Vector geometry type native to Oracle Database
- Vector operators and functions: 400+ functions to analyze and generate geometries through SQL access
- R-tree indexing
- Projections and coordinate systems: support for 4000+ commonly used coordinate systems, implicit and explicit transformations by layer or geometry, geodetic model, user-defined coordinate systems
- GeoRaster support for imagery and raster data: Virtual Mosaic, raster algebra analytical algorithms and operations, image processing (parallel for high performance), Java and PL/SQL APIs, GDAL loader support with ETL Wizard
- 3D data model – native support for 3D geometries, surfaces and point clouds (LIDAR data).
- Spatial Web Services: support for OGC and ISO geospatial standards – Web Feature Service (WFS) 1.1.0, Web Map Service (WMS), Catalog Service for the Web (CSW) 2.0.2, Web Coverage Service (WCS) 2.0.1, Open Location Services 1.1, with unified administration console/UI
- Geocoding engine
- Routing engine
- Spatial analysis and mining functions
- Topology data model
- Linear referencing system

- Network Data Model Graph – a storage model to represent graphs and networks in link and node tables, including feature, time, and multimodal transporation modeling
- Parametric Curves (NURBs): mathematically precise representation of free form curves that can be reproduced exactly for 2D and 3D data
- GeoJSON support and REST APIs: spatial data access for modern development environments
- Location data enrichment services: API with bundled geographic hierarchy and place names data set, to search and categorize large collections of unstructured text records with location tags for analysis
- Map visualization: HTML5-based mash up component for spatial visualization of geographic data in Oracle Spatial and Graph or from external services. (Must be deployed in a JEE container or in Oracle Java Cloud Service)
- Spatial index optimization for point data: composite B-tree indexes for improved performance on index update and creation for point only data
- Spatial indexes: support for all index partitioning methods and system managed indexes for improved performance and simplified management
- Location tracking server tracks the movement of objects in and out of regions of interest on a massive scale for logistics and IOT

RDF Semantic Graph Features

- RDF views: present relational data in RDF format for SPARQL queries and connect with other linked data and RDF graphs to relate and facilitate enterprise data integration.
- SPARQL 1.1 in the Database: SPARQL 1.1 supported by SEM_MATCH table function
- Store/query spatial data in RDF: The OGC GeoSPARQL standard is supported
- Security on inferencing: Ladder-based inferencing ensures inferred triples are labeled with the proper triple-level security.
- More inferencing support: User-defined inferencing and full OWL 2 EL profile support
- Graph analytics: SPARQL 1.1 property path expressions find graph patterns across any length path
- In-memory graph analytics: shortest path, reachability, within-cost, and nearest-neighbor.
- Graph mining and statistics: Oracle Advanced Analytics Data Mining and Oracle R Enterprise are supported
- SPARQL Gateway: SPARQL query results can be presented for visualization in tools that support XML data sources, such as Oracle Business Intelligence (OBIEE)

Appendix 2: New Features in Oracle Database 18c

Property Graph New Features

- PGQL queries on graph data stored in Oracle Database
- PGQL queries to find in-memory subgraph instances that match a given query pattern
- SQL-based collaborative filtering
- More in-memory analytics: new variants of Pagerank, Personalized SALSA, K-Core, Diameter, Radius, and Eccentricity, and PRIM
- Support for undirected graphs
- Node.js client
- Apache Zeppelin support
- An execution and scheduling manager to better control in-memory analyst tasks and resources

Spatial New Features

- Support for R-Tree spatial indexes in distributed and Oracle XA transactions
- JSON support expanded: ability to convert data between JSON and all the types supported by SDO_GEOMETRY, support for larger range of geometries including 2D, 3D, solid, surface, and LRS geometries
- Oracle REST Data Services support for spatial operations on Oracle Spatial and Graph data
- Support for sharded databases with spatial data types
- Spatial Web Services: administration console enhancements
- Spatial Web Services: WFS, CS-W, WCS can support reading data from multiple data sources in the same WebLogic Server instance
- Spatial Web Services: Support for CSW 2.0.2 - ISO Metadata Application Profile
- Spatial Index now optional for spatial operations
- Performance enhancements for spatial indexes for point data (composite B-tree)
- Routing Engine: new servlet for lightweight location-based queries related to speed limit and traffic speed
- New Utility Functions: SDO_UTIL.GETFIRSTVERTEX and SDO_UTIL.GETLASTVERTEX

Enhanced concave hull supportRDF Semantic Graph New Features

- Oracle Database In-Memory support
- List-hash composite partitioning support
- Faster loading of RDF quads having literals greater than 4000 bytes
- Native support for Turtle and Trig formats.

ORACLE

Oracle Corporation, World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065, USA

Worldwide Inquiries
Phone: +1.650 506.7000
Fax: +1.650.506.7200

CONNECT WITH US

blogs.oracle.com/oracle
facebook.com/oracle
twitter.com/oracle
oracle.com

Integrated Cloud Applications & Platform Services

Spatial and Graph Analytics with Oracle Database 18c
February 2018
Authors: William Beauregard, Siva Ravada, and Jean hm