May 2012
Oracle Spatial User Conference
Andrew Howles & Jo Shannon
IS Principal Architect
General Manager, Operations
Automated Data Quality Processes in Ordnance Survey’s GDMS
OVERVIEW
- Centrally managed, authoritative database
- Incorporates 5 corporate national datasets
- Supports national data capture and maintenance
- Contains in excess of 500 million features

CHALLENGES / OPPORTUNITIES
- Need to integrate disparate data sets
- Need to extract 32 million features and submit 313,000 changed features per day
- Need to support multiple users in the same geographic area
- Automated data validation to achieve product ready database

ORACLE SOLUTIONS
- Oracle Spatial Database, Partitioning, RAC, Dataguard
- Oracle Workspace Manager
- Oracle Application Server & Weblogic
- Oracle BPEL & Worklist Application
- Oracle Enterprise Service Bus
- Oracle Identity Management

RESULTS
- Consolidation of raster, vector, network data in 5.3TB sized central repository
- Single source database for product derivation
- Seamless working – removal of tile based restrictions
- Resilient production environment achieving 99.5% availability
Agenda

• Introductions
• Ordnance Survey and GDMS business context
• Oracle technology in GDMS
• Radius Studio technology in GDMS
• Challenges
• Next steps
• Conclusions
• Q&A
Introductions

• To the Presenters:
  – Andrew Howles, Ordnance Survey
  – Jo Shannon, 1Spatial

• To the Geospatial Data Management System (GDMS):
  – Enterprise geospatial data factory – end-to-end planning, collection, management, validation, and publication to product stores
  – Multi-vendor service-oriented architecture built on the Oracle technology stack
  – Went live in July 2011
Ordnance Survey business context

- Self-funding government agency; remit covers Great Britain (England, Scotland and Wales) – 243,241 sq km / 93,956 sq miles
- Datasets maintained on a national basis – common specifications to ensure consistency in terms of precision, accuracy and content
- Close collaboration with Land Registry of England and Wales, and Registers of Scotland
- Integrated data capture for field surveyors, HQ-based editors, and external suppliers
Ordnance Survey business context

- Digital geospatial data products over 90% of Ordnance Survey’s business
- Over 450 million unique feature identifiers (TOIDs) introduced by OS MasterMap® in 2001 – hooks to which information can be linked:
  - emphasis on geographic information
  - integration of addresses, boundaries, road centre-lines and network topology with topographic base
GDMS programme context

- Seamless data maintenance solution
- Intelligent and integrated maintenance data model
- Separation of product data
- Automated data validation
- 650+ production users in Ordnance Survey and other organisations
- Platform for new product development
GDMS programme context
GDMS programme context
GDMS programme context
GDMS real-world object model – Geobase-04

• Main unit is the *feature*
• Features are composed of *components* including:
  – identity
  – lifecycle
  – location
  – classification
  – relationships
  – history
• Model prescribes around 150 specification rules
GDMS programme context
Non-functional characteristics

- 500 million features and growing (+ history)
- 4,000 data maintenance jobs per day
- 32 million feature extractions per day
- 313,000 features updates per day
- 650+ users inside and outside Ordnance Survey
- 99.5% availability
• Mainstream architecture patterns:
  – $n$-tier solution
  – service-oriented approach (business logic)
  – workflow layer (orchestration logic)

• Heterogeneous software environment – best of breed COTS components from diverse vendors:
  – 1Spatial, BAE Systems, ESRI, Intergraph, Microsoft, Oracle, Safe Software, Snowflake Software

• Long transaction management, disconnected editing

• Scalable infrastructure on commodity hardware, selective virtualisation
GDMS programme context
Logical architecture

Overview OS production architecture
GDMS programme context
Role of 1Spatial within the Intergraph Consortium

• Member of the Intergraph Consortium along with Snowflake Software
Role Of Oracle in GDMS
Workspace Manager in GDMS
Key to Transaction Integrity
Workspace Manager in GDMS
Acetate Overlay Model
Workspace Manager in GDMS

Data Conflict
Workspace Manager in GDMS
Logical Inconsistency
Role of Radius Studio

What is Radius Studio

- Spatial data integration platform
- Knowledge Store
- Rules based spatial analysis
- Web application & web services
- GRID processing
- Extensible – built-ins, datastores
Role of Radius Studio

Rules based processing paradigm

Data → Rule → Pattern → Rule → Action → Report → Reconcile
Role of Radius Studio

Examples of Rules

**FACT**

Rule condition for **Railway Points** objects:

- There is at least 1 **Rail** object for which ...
  - **Rail**.geometry is within a distance of 0.8 of **Railway Points**.geometry

**PATTERN**

Railway points must be within 0.8 metres of at least 1 rail object

**ACTION**
Challenges

1. Physical realisation of the logical model

- Schema design choices can cause physical data conflict to arise in instances where there is no logical conflict
Challenges

1. Physical realisation of the logical model
Challenges

1. Physical realisation of the logical model
Challenges

1. Physical realisation of the logical model

• Schema design choices can cause physical data conflict to arise in instances where there is no logical conflict

• Our advice: When optimising your schema, don’t just think about performance but consider conflict avoidance as well
Challenges

2. Conflict resolution

- Complex data requirements mean simple A versus B type decisions are not suitable for conflict resolution:
  - logical semantics of spatial data – e.g. topological structuring
  - version metadata
Challenges

2. Conflict resolution
Challenges

2. Conflict resolution
Challenges

2. Conflict resolution
Challenges

2. Conflict resolution
Challenges

2. Conflict resolution

Conflict resolution options in Oracle Workspace Manager:
Challenges

2. Conflict resolution
Challenges
2. Conflict resolution
Challenges

2. Conflict resolution
Challenges

2. Conflict resolution

Conflict resolution options in Oracle Workspace Manager:

- Live
- Workspace 2
- Base

The solution we require for supply to customers
Challenges

2. Conflict resolution

• Complex data requirements mean simple A versus B type decisions are not suitable for conflict resolution:
  – logical semantics of spatial data – e.g. topological structuring
  – version metadata

• *Our advice*: Understand the implications of your data model and schema for the process of conflict resolution
Challenges

3. Work planning

• Planning work to avoid conflict, where possible, is better than being in a conflict resolution situation
Challenges

3. Work planning
Challenges

3. Work planning
Challenges

3. Work planning

• Planning work to avoid conflict, where possible, is better than being in a conflict resolution situation

• Our advice: Investigate and understand the implications of work planning processes with respect to susceptibility to conflict, and agree an appropriate approach with users of the system
Challenges

4. Charting new territory

• Building innovative systems with cutting edge technologies and designs, means you will be charting new territory and encountering unexpected challenges.

• *Our advice:* Develop a trusted relationship with suppliers to ensure a collaborative approach to resolving issues as they arise.
GDMS next steps

• Data quality improvement for existing products
• Strategic data enhancements to support new products:
  – additional intelligence (lifecycle, attribution, cross-referencing)
  – increasing 3D capabilities
  – additional networks
• Operational improvements:
  – Oracle Database upgrade to 11.2.0.3
  – Mid-tier upgrade to WebLogic
Conclusions

- Cutting edge development requires a strong architectural vision and well managed requirements
- Core technologies must support the non-functional requirements out of the box
- Getting the business processes right is as important as building the right systems
- Understanding the logical data model’s physical implications is critical
- Collaborative supplier relationships are key to success
“If successful the Ordnance Survey's new geospatial database and data management system will define a best practice for the collection, distribution and use of geospatial data.”

International Data Corporation, 2007

- GDMS operational since July 2011
- 1.2 million Ordnance Survey units of change committed to date
- Data, system and process enhancements now in progress for delivery of new products
Q&A