2nd ICA / EuroSDR NMA Symposium
Amsterdam 3/4\textsuperscript{th} December 2015

OSGB Multi-Resolution Data Programme (MRDP)

Derek Howland
Operations Domain Leader, MRDP

Rebecca Walters
Solution Architect, MRDP
MRDP VISION

‘To improve the currency and consistency of current products and allow simpler creation of new derived products that meet changing business and customer needs.’

Development of the generalisation system (‘GenIE’) **started in Jan 2014**, in collaboration with partners 1Spatial. **System release to Production - Jan 2016**
MRDP/GENIE DATA ARCHITECTURE FOR DERIVED PRODUCTS

Resolution - about 1/25,000

Resolution - about 1/10,000
Process of loading data from Core source data (MAIA) and Tactical databases into VMD product store

- MAIA to District Content Store
- District Content Store to VMD (Product Store)
- Publish and Package Products

- Pre-processing
- Generalise Partitions
- Validate
- Automated process
- Read from MAIA
- Build topology
- Dual Carriageway collapse
- Create partitions

- DCS
- Manual Fixing
- Validation
- Automated process
- GEN
- Create Partitions
- VMD Validate
- Automated process
- VMD
- Package
- Publish / Package

UK-OFFICIAL
NATIONAL LOAD: Job Map for September VMD refresh

Source to District Content (298 partitions)
Process of loading data from Core source data (MAIA) and Tactical databases into VMD product store

- MAIA to District Content Store
- District Content Store to VMD (Product Store)
- Publish and Package Products

- Generalise Partitions
- Validate
- Manual Fixing
- Automated process
- Pre-processing
- Automated process

- VMD Validate
- Manual fixing
- Automated process

- VMD:
  - Publish
  - (4 formats) and Package

- DCS:
  - Repeat for DCS to Product store
  - Manual edit – about 300 jobs created automatically
  - about 400 features failing validation rules

- Generalisation: 164 actions in each of 298 partitions
- Parallel processing on 40 nodes
- Validation: each of 298 partitions
- Parallel processing on 40 nodes
- Jobs automatically created and activated
- Manual edit – about 300 jobs created automatically
VECTOR MAP DISTRICT PRODUCT (VMD) — GENERATED THROUGH GENIE [1]
VECTOR MAP DISTRICT PRODUCT VMD—GENERATED THROUGH GENIE [3]
### Performance statistics

<table>
<thead>
<tr>
<th>Generalisation Performance</th>
<th>Source to District Content</th>
<th>3 days</th>
<th>District Content to VMD Product</th>
<th>3 days</th>
<th>6 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual Editing</td>
<td>Source to District Content</td>
<td></td>
<td>District Content to VMD Product</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- about 300 jobs, 400 features</td>
<td></td>
<td>- about 200 jobs, 250 features</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total features edited about 650 (out of 24 ½ million features in the product 🎉)</td>
<td></td>
<td>Total 8 man days effort</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 days elapsed time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publication Performance</td>
<td>Publication in 4 formats:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Raster Full Colour – 2863 tiles (10x10km)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Raster Backdrop – 2863 tiles (10x10km)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Vector – GML – 809 tiles (20x20km)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Vector – Shape – 55 tiles (100x100km)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total 6590 tiles</td>
<td></td>
<td>2 days</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### MRDP team moving forward

**v1.0 dev complete, into production Jan 2015**

<table>
<thead>
<tr>
<th>V1.1 System team</th>
<th>• System enhancements</th>
<th>Team of 8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Component upgrades</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Performance improvements</td>
<td></td>
</tr>
<tr>
<td>Product team 1</td>
<td>3 new Open data Products – District resolution</td>
<td>Team of 8</td>
</tr>
<tr>
<td>Product team 2</td>
<td>Local content store, premium products – Local resolution</td>
<td>Team of 8</td>
</tr>
</tbody>
</table>
SCALE

System is capable of running parallel product generation

Currently only one product generated once every six months

Scale needed to create one product:

- Four development teams for 18 months
- 12 Generalisation Servers (4 Processing nodes each)
- 6 Validation Servers (3 Processing nodes each)
- 37 Servers total in Prod environment
- 272 Servers total across all environments

11 Environments:

- Prototype environment
- 4 Team specific development environments
- CI/CD - Integration including third party integration
- Test - Test script development environment
- Non Functional Test – Integration at scale
- Performance Test – Performance testing
- Pre-Production – Live exact enterprise integration
- Production environment
ARCH / TECH CHALLENGES

1. Building a ‘system’ ie end-to-end automation
2. Understanding the need for change only update
3. Research code – ‘black-boxes’
4. Requires a strong mix of IT, data and generalisation skills, and as a result has quite a steep learning curve
5. Scaling processing for contiguous data
6. Predictability - changing generalisation rules can have a huge effect on performance
7. Scaling for more than one product
8. Performance constrains currency to one month at present
1. Generalisation – ‘*part science, part art*’ – use of validation to ensure product quality

2. Automated generalisation versus Cartographic Quality (Open data versus Premium Products)

3. Currency - at what price? Viability of COU?