GeoServer, GeoTools and GeoBatch: supporting operational Meteorology and Oceanography

Ing Simone Giannecchini

FOSS4G 2009, Sydney
23rd September 2009
Outline

- GeoSolutions Facts
- Reference Scenario/Domain
- Enterprise SDI
- SDI Building Blocks explained
  - GeoBatch
  - GeoServer
  - BUDDATA ebRIM
  - geoSDI-ERA
- Relevant Use cases
GeoSolutions - Facts

- Funded mid 2006 in Italy
- Staff of 7 engineers (2 consultants)
- Expertise
  - Image Processing, GeoSpatial Data Fusion
  - Java, Java Enterprise, C++, Python
  - JPEG2000, JPIP, Advanced 2D visualization
- Supporting/Developing FOSS4G projects
  - GeoTools, GeoServer, GeoBatch, ImageIO-Ext, uDig
- Focus on
  - Consultancy (agencies, large private companies, etc…)
  - Building stack of geospatial Open Source products

GeoSolutions

FOSS4G 2009, Sydney
23rd September 2009
Reference Scenario/Domain

Enterprise Boundaries

Enterprise GeoSpatial Server

REP/COP

GeoSolutions

FOSS4G 2009, Sydney
23rd September 2009
Enterprise Spatial Data Infrastructure

ESB

Reporting
Monitoring

A&A&A
Data Collection Agents
Batch Processing

Online Processing

Data Dissemination Server
Index & Search

GeoPortal

Data Management Layer

Map Accelerator

Data Management Layer

Spatial DBMS
Net Storage

GeoSolutions

FOSS4G 2009, Sydney
23rd September 2009
Ingestion/Preprocessing

- Real-time produced data
  - Remote Sensing, In situ, Models, etc..
- Recognizable, isolate data streams
- Flow
  - Incoming series of geospatial data
  - Specific format
  - Defined rules for preprocessing
  - Defined rules for ingestion and/or exploitation
  - Automatic Ingestion and preprocessing → No manual intervention
- Example, flow of EO raster

[Diagram showing the flow of data from Producer to GeoWebCache and GeoServer]
GeoBatch: Basics

- Event-based geospatial-aware batch ingestion/processing system
- Basic tools for automatic collection, processing and publication of data.
- Open Source (leverages on GeoTools, Apache FTP, Spring, Xstream, etc…)
- Embedded FTP Server
- Publish to GeoServer
- Publish to GeoWebCache (ongoing)
- Web based
- XML configuration
- http://docs.codehaus.org/display/GEOBATCH
GeoBatch: FlowManager

FileSystemMonitor

EventDispatcher

Event Mailbox

Event Consumer

Event Consumer

Directory Monitoring

FlowManager
GeoBatch: Sample Flow

```xml
<FlowConfiguration>
  <workingDirectory>AISAnomalies</workingDirectory>
  <autorun>true</autorun>
</FlowConfiguration>

<EventConsumerConfiguration>
  <workingDirectory>AISAnomalies</workingDirectory>
  <performBackup>true</performBackup>
</EventConsumerConfiguration>

<FileEventRule>
  <optional>false</optional>
  <originalOccurrencies>1</originalOccurrencies>
  <regex>.*\.shp</regex>
  <id>r1</id>
  <description>description</description>
  <name>test</name>
</FileEventRule>

...<GeoServerActionConfiguration>
  <crs>EPSG:4326</crs>
  <envelope/>
  <dataTransferMethod>URL</dataTransferMethod>
  <geoserverPWD>geoserver</geoserverPWD>
  <geoserverUID>admin</geoserverUID>
  <geoserverURL>http://localhost:8080/geoserver</geoserverURL>
  <wmsPath>/</wmsPath>
  <defaultStyle>polygon</defaultStyle>
</GeoServerActionConfiguration>
```
GeoBatch: Sample Flow

- `<GeoServerActionConfiguration>`
  - `<crs>EPSG:4326</crs>`
  - `<envelope />`
  - `<dataTransferMethod>URL</dataTransferMethod>`
  - `<geoserverPWD>geoserver</geoserverPWD>`
  - `<geoserverUID>admin</geoserverUID>`
  - `<geoserverURL>http://localhost:8080/geoserver</geoserverURL>`
  - `<wmsPath>/</wmsPath>`
  - `<defaultStyle>polygon</defaultStyle>`
  - `<serviceID>aisAnomaliesGeoServerGeneratorService</serviceID>`
  - `<styles />
</GeoServerActionConfiguration>`

`<EventConsumerConfiguration>`

`<eventGeneratorConfiguration class="FsEventGeneratorConfiguration">`
  - `<wildCard>*.*</wildCard>`
  - `<watchDirectory>AISAnomalies/in</watchDirectory>`
  - `<keepFiles>true</keepFiles>`
  - `<osType>OS_UNDEFINED</osType>`
  - `<eventType>FILE_ADDED</eventType>`
`</eventGeneratorConfiguration>`

`<id>flowAISAnomalies</id>`
`<description>flowAISAnomalies: testing injection of AISAnomalies shapefiles</description>`
`<name>flowAISAnomalies</name>`
`</FlowConfiguration>`
Management/Dissemination/Exploitation

- Raw Data Storage and Dissemination
- Metadata Storage and Dissemination
  - Index & Search Capabilities for data
- Service Discovery
  - Index & Search Capabilities for services
- Raw Data Portrayal
  - Coverage → Maps
  - Feature → Maps
- Processing
GeoServer

- **GeoSpatial enterprise gateway**
  - Management raster and vector data
- **Standards compliant**
  - OGC WCS 1.0 - 1.1.1 (RI)
  - OGC WFS 1.0 - 1.1 (RI)
  - OGC WMS 1.1.1
  - OGC WPS 1.0
- **Google support**
  - KML, GeoSearch, etc..
- **Enterprise ready (almost!)**
GeoServer: Limits/Shortfalls

- In-memory internal catalog
  - Non transactional, Scale sub-optimally
  - Replication is problematic
- File Based Configuration
  - Non transactional, Scale sub-optimally
  - Replications is problematic
- File based Coverage support (can be overcome easily)
- 1 CoverageStore → 1 Coverage (can be overcome, not easily)
- Embedding/Controlling GeoServer can be difficult
  - REST uses HTTP always
  - REST is non transactional
- No Time and Height/Depth support for raster
- No CSW or ebRIM support (yet!)
- WCS support needs to be improved
GeoServer: Hibernate catalog

- Configuration stored in DB
  - Hibernate
  - Spring based Dao
- Transactions via Spring/EJB3
- Caching with EhCache
- Same catalog interface, different implementation
- Catalog interface needs tweaking
- GeoServer UI still needs tweaking
- Transaction demarcation needs tweaking
GeoServer: Hibernate catalog

Catalog Interface

Hibernate Based Catalog

ServiceDAO

CatalogDAO

Hibernate

2° L Distributed Cache

DB

DB

DB

Net Storage

FOSS4G 2009, Sydney
23rd September 2009
GeoServer: Transactional API

- In-JVM communication
  - Transactional
  - No HTTP, No Serialization
  - Spring, EJB3, JMX

- Intra JVM
  - Transactional
  - Serialization is involved
  - Separate mechanism for data upload
  - HTTP, TCP, etc...
  - Spring Remoting, EJB3, JMX
GeoServer: Advanced Raster
GeoServer: RasterVault
Index & Search Capabilities: ebRIM

- Registry versus Repository
- Flexible Information Model
- Object Lifecycle Management
- Content Management Service
- Event Notification
- Protocol Bindings
- Publish-Find-Subscribe(Consume)
- Federation

Diagram:
- Registry
- Provider
- Consumer
- Broker
- Trader
- Authority

Geo Solutions

FOSS4G 2009, Sydney
23rd September 2009
Open Source ebXML Registry/Repository (or ebRR in short)

ESA founded ERGO project lead by Intecs

OGC specifications implemented

- 07-110r4: CSW-ebRIM Registry Service - Part 1: ebRIM profile of CSW (1.0.1)
- 07-144r4: CSW-ebRIM Registry Service – Part 2: Basic extension package (1.0.1)
- 07-038: OGC Cataloguing of ISO Metadata (CIM) using the ebRIM profile of CS-W (0.1.11)
- 06-131r6: EO Products Extension Package for ebRIM (ISO/TS 15000-3) Profile of CSW 2.0

Code on Google Code (http://code.google.com/p/buddata-ebxml-registry/)

Wiki ERGO project (ESA HMA wiki)
Buddata EbRR @ Work

- Metadata published with data contextually
- Use of custom profiles
GeoPortal: geoSDI - ERA

- Modular/Extensible Rapid-mapping Web Application
- EXT-GWT + Openlayers + Hibernate Spatial + Spring
  - Core components to perform basic actions
    - OpenLayers integration
    - Layers Wide search
    - Layer Tree
    - Security subsystem
    - Style editor (ongoing)
  - Plugin Engine to perform specific tasks
    - Routing, geocoding, etc..
  - GEOGWT?
    - GEOEXT + EXT-GWT
geoSDI – ERA Modular

- GeoSDI-ERA Architecture Schema

- GeoSDI-ERA CORE
- GeoSDI-ERA Plug-Ins

- OGC-WMS
- PostGIS DB
- Security Subsystem

- Map Context
- Plugins Engine

- Map
- Map Widgets

- SLD Editor
- I-Mobility Tools

- Raster Symbolizer
- Classifier
- SLD Load/Save
- Geo-Coding
- Routing
- Traffic Density Maps

Geo Solutions

FOSS4G 2009, Sydney
23rd September 2009
Use Case: AIS data management

- Automated Identification System (AIS)
  - WHF emitter/receiver
  - GPS
- Continuously recording AIS traffic
  - Mediterranean Sea
  - Black Sea
- Providing support for storage and publishing of the results of the AIS processing
  - AIS Coverage
  - AIS Forecasts
  - AIS Anomalies
  - More in the near future, like Radar contacts and Fused Tracks
Use Case: AIS data management

- Fully automatic AIS coverage and anomaly detection ingestion and publishing process.
Use case: SAS Mission Manager

- SAS: Synthetic Aperture Sonar
- Sea bed analysis and detection of features/artifacts

- Automatic Ingestion of SAS data and metadata
- SAS Data preprocessing
- SAS mission data management and access
- SAS mission data visualization
Use case: SAS Mission Manager

SAS Mission Manager Web interface

Presentation Tier

GeoServer

WFS
WCS
WMS

Middleware Tier

Registry ebRIM

Data and Metadata Management

Mission Data Package

Network Area Storage

GeoBatch

Data and Metadata Management

Mission Data Package

Network Area Storage

GeoBatch

GeoServer

WFS
WCS
WMS

Presentation Tier

Registry ebRIM

GeoSolutions

FOSS4G 2009, Sydney
23rd September 2009
Use case: SAS Mission Manager

Sonar images are Curtesy of NURC
Use Case: Super Ensemble Modeling

- Optimal MetOc models mixing
  - Same model, different boundary conditions
  - Different model same boundary conditions
- System training (Kalman filter) → historical data
- Output → Series of weight for optimal MetOc model mixing
- Result → optimal MetOc model as fusion of weighted models
Use Case: Super Ensemble Modeling

Discover → Buddata EbRR

Harvest → Toolbox WPS

Invoke → MATLAB

Publish Results → WCS::getCoverage

WMS::getMap → GeoServer

Input Data → GeoBatch

Deploy → Ensemble GUI

GeoBatch

GeoServer

Toolbox WPS

MATLAB

Buddata EbRR

FOSS4G 2009, Sydney
23rd September 2009
Use Case: Super Ensemble Modeling

- Current solution
  - Split 3D hyper-cubes of data into a set of 2D slices
  - Use registry for registering T and Z

- Next Step
  - Handle multidim coverages through GeoServer + RasterVault

GeoBatch:
- Vertical Interpolation
- Geo-Rectification
- Resampling

Input 3D Cubes

Registry

GeoServer

GeoBatch

Publish Metadata

Publish Data

FOSS4G 2009, Sydney
23rd September 2009
Use Case: Super Ensemble Modeling
Use Case: TDA

Satellite data

Oceanographic/Meteo forecast

On-site observations

Fusion

Value added products
Use Case: TDA

- **Discover**
  - Buddata
  - EbRR

- **TDA Explorer**

- **GeoBatch**
  - Input Data
  - Publish

- **GeoServer**
  - Publish
  - WCS:getCoverage
  - WMS:getMap

- **Toolbox WPS**
  - FUZZY logic
  - Invoke
  - Harvest

- **WMS:getMap**

- **FOSS4G 2009, Sydney**
  - 23rd September 2009

**Geo Solutions**
Use Case: TDA

Search page

WPS invocation

FOSS4G 2009, Sydney
23rd September 2009
Use Case: LAMMA

NCEP-GFS → WRF → NMM → Postproc → Postproc

ECMWF

GeoSolutions

FOSS4G 2009, Sydney
23rd September 2009
Use Case: FAO FIGIS

- Dynamic Statistical Query Engine
  - Compute GIS Layer Intersections
  - Download statistics outcomes
  - Monitoring the Engine status

The Dynamic Query Engine consists of a set of Service Units capable of generating statistics on a periodic interval fashion.

The Service Units interact directly with FIGIS GeoServer instances through the OGC WFS-T protocol.

Enterprise Service BUS (Apache ServiceMix)

UI

HTTP Provider

HTTP Consumer

Scheduler Service

GeomCompare Service

WFS-T

GeoServer

WFS

DB (ORACLE)
Use Case: FAO FIGIS

The Service Units interact directly with FIGIS GeoServer instances through the OGC WFS-T protocol.
GeoServer, GeoTools and GeoBatch: supporting operational Meteorology and Oceanography

Ing Simone Giannecchini
Reference Scenario/Domain

[Diagram showing the relationship between MetOc Models, Remote Sensing, InSitu Obs, and REP/COP.]
GeoBatch: RoadMap

- Improved Event Hierarchy
  - Generic Event Subsystem
  - Time Based Events
- Improved UI
  - Visual Builder for flows
- Refactor Messaging Subsystem
  - JMS
  - ESB (ServiceMix)
- Actions for GeoWebCache
- SOAP based Event generator
- Integrate with GeoServer?
GeoServer

Shapefile
PostGIS
DB2
Oracle
MySql
ArcSDE
WFS

GeoTIFF
ArcGrid
GTopo30
Img+world
Mosaic
MrSID
JPEG 2000
ECW, Pyramid

Vector files
DBMS
Servers
Raster files

WFS 1.0
WFS 1.1
WMS 1.1.1
Google*
WCS 1.0
WCS 1.1.1
GWC

GeoServer

PNG, GIF
JPEG
TIFF,
GeoTIFF
SVG, PDF
KML/KMZ
Shapefile.zip
GML2
GML3
GeoRSS
GeoJSON
CSV/XLS

Raw vector data
GeoTIFF
ArcGrid
GTopo30
Img+World
GWC
KML superoverlays
Google maps tiles

FOSS4G 2009, Sydney
23rd September 2009
Old configuration: Filesystem catalog

- All the catalog objects are kept in memory: cannot deal with a huge amount of data.
- Filesystem persistence is not transactional
- External applications have to pass through GS to get information.
GeoServer: Advanced Raster

- **Time Management**
  - Sequence of 2.5D snapshots
  - Temporal ranges \([T_1, T_2]\) with Nearest Neighbor interpolation
  - B-Trees (logarithmic access) to improve performances

- **Elevation Management**
  - Sequence of 2D layers
  - Positions values with Nearest Neighbor interpolation
  - B-Trees (logarithmic access) to improve performances
GeoServer: Advanced Raster

- Create a regular tessellation of original data
  - Work in raster space
  - Save georeferencing
  - Save original translation factors

- Tile integer indexing
  - Optimize most frequent case (TMS like approach)
  - Simplified tile lookup (no DBMS query)
  - Integer arithmetic + Recomposition by simple translation
  - Tile Caching highly simplified

- Multiband Tiles

- Consider MetaTiling
  - Reduce file access
ebRIM: Registry Information Model 3.0
ebRIM: OGC Basic package
ebRIM: infrastructure
ISO Data types defined in the Basic package
Some CIM classes and how they extend some ebRIM classes
Interoperability between CIM EP and ISO AP
EbRR @ Work: ESA portal

FOSS4G 2009, Sydney
23rd September 2009
geoSDI – ERA Modular

- GeoSDI-ERA Architecture Schema

- GeoSDI-ERA CORE
- GeoSDI-ERA WEB
- Map
- Map Context
- Plugins Engine
- GeoSDI-ERA Plug-Ins
- SLD Editor
- I-Mobility Tools
- Security Subsystem
- PostGIS DB
- OGC-WMS

- Security Subsystem
- PostGIS DB
- OGC-WMS

- Raster Symbolizer
- Classifier
- SLD Load/Save
- Geo-Coding
- Routing
- Traffic Density Maps

FOSS4G 2009, Sydney
23rd September 2009
GeoSDI – ERA Modular

- GeoSDI-ERA Portal Concept
- Enriched by Pluggable Map Widgets
- The Panels can be customized and moved
- Layout adapted to the specific Application requirements

GeoSDI-ERA WEB

OGC-WMS Servers Tree-Panel
OGC-WMS Layer/Legend Handler
OGC-WMS GetFeatureInfo Tool

OGC-SLD Editor
I-Mobility Routing Tool...

Geo Solutions

FOSS4G 2009, Sydney
23rd September 2009
geoSDI – ERA Modular

- GeoSDI-ERA OGC-SLD Editor Widget
- Able to recognize the rules of an SLD and dynamically creates the editor panel
- Able to interact with the GeoServer/GeoTools Classifier

Polygon Symbolizer with Classification

Raster Symbolizer Editor
geoSDI – ERA Modular

- GeoSDI-ERA I-Mobility Tools & Widgets
- Fully Pluggable I-Mobility tools
- GeoCoding, Reverse GeoCoding, Routing, Traffic Density Maps

Shorest Path Routing example
Generated Density Maps with Graphs
Use Case: AIS data management

- Fully automatic AIS coverage and anomaly detection ingestion and publishing process.
- A control GUI allows to monitoring and stop/resume GeoBatch Data Flows
Polygon geometry extraction

Polygon geometry processor

Fuzzification

Rule-base

Inference mechanism

Heuristic rule base setting

Rule base optimisation

Training data set

Traffic light map

• Remotely sensed data,
• oceanographic models,…
• other data sources
Use Case: FAO FIGIS

- GeoSolutions is providing support to Fisheries and Aquaculture Department of FAO

The mission of the Fisheries and Aquaculture Department of FAO is to facilitate and secure the long-term sustainable development and utilization of the world’s fisheries and aquaculture.

The Department promotes policies and strategies aiming at sustainable and responsible development of fisheries and aquaculture in inland and marine waters.
Use Case: FAO FIGIS

- Interactive Querying and Displaying of Tuna and Billfish Catches

The introduction of the Aggregate DataStore allows to completely avoid table/view creation on the DB.

Client

Map
Vendor Request params

TUNA WMS Proxy

Map
Request params

GeoServer

Aggregate DataStore management

FeatureType (CQL)

FIGIS DB (Oracle)

Statistical Data Source

GeoSpatial Data Source

Aggregate DataStore (filter)

TUNA Business Logic

Statistical Data Source

GeoSpatial Data Source

FIGIS DB (Oracle)
Use Case: FAO FIGIS

- **Aggregate DataStore**
  - GeoTools module/plugin
  - Performs custom aggregate SQL-like functions above standard DB DataStores
  - Builds Aggregated FeatureTypes on the fly