

GeoServer, GeoTools and GeoBatch: supporting operational Meteorology and Oceanography

Ing Simone Giannecchini



FOSS4G 2009, Sydney
23rd September 2009



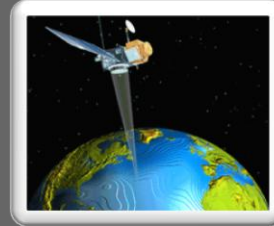
Outline

- **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

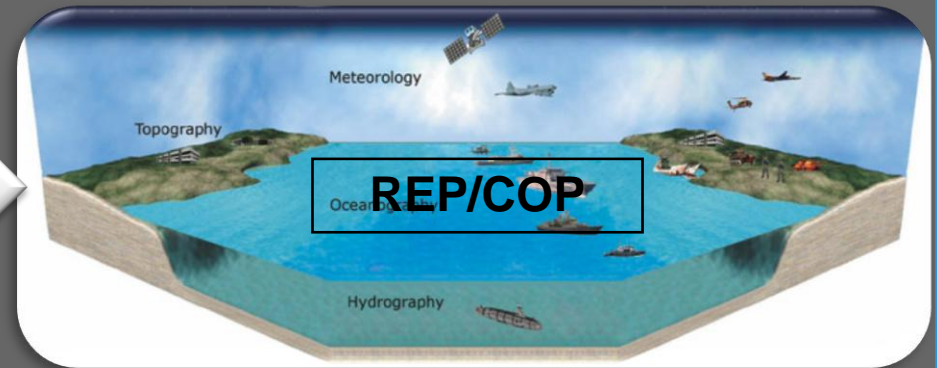
Reference Scenario/Domain



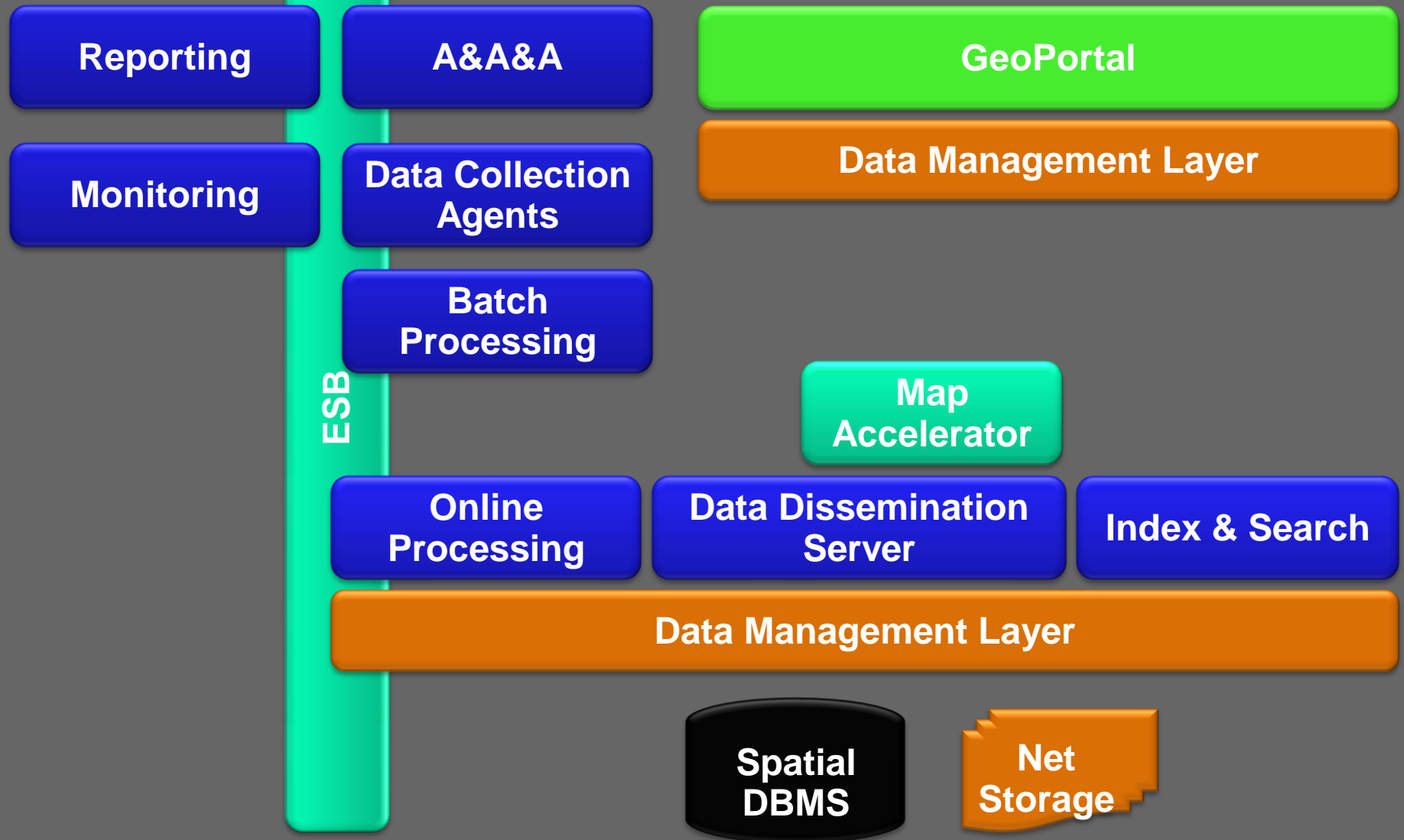
Enterprise Boundaries



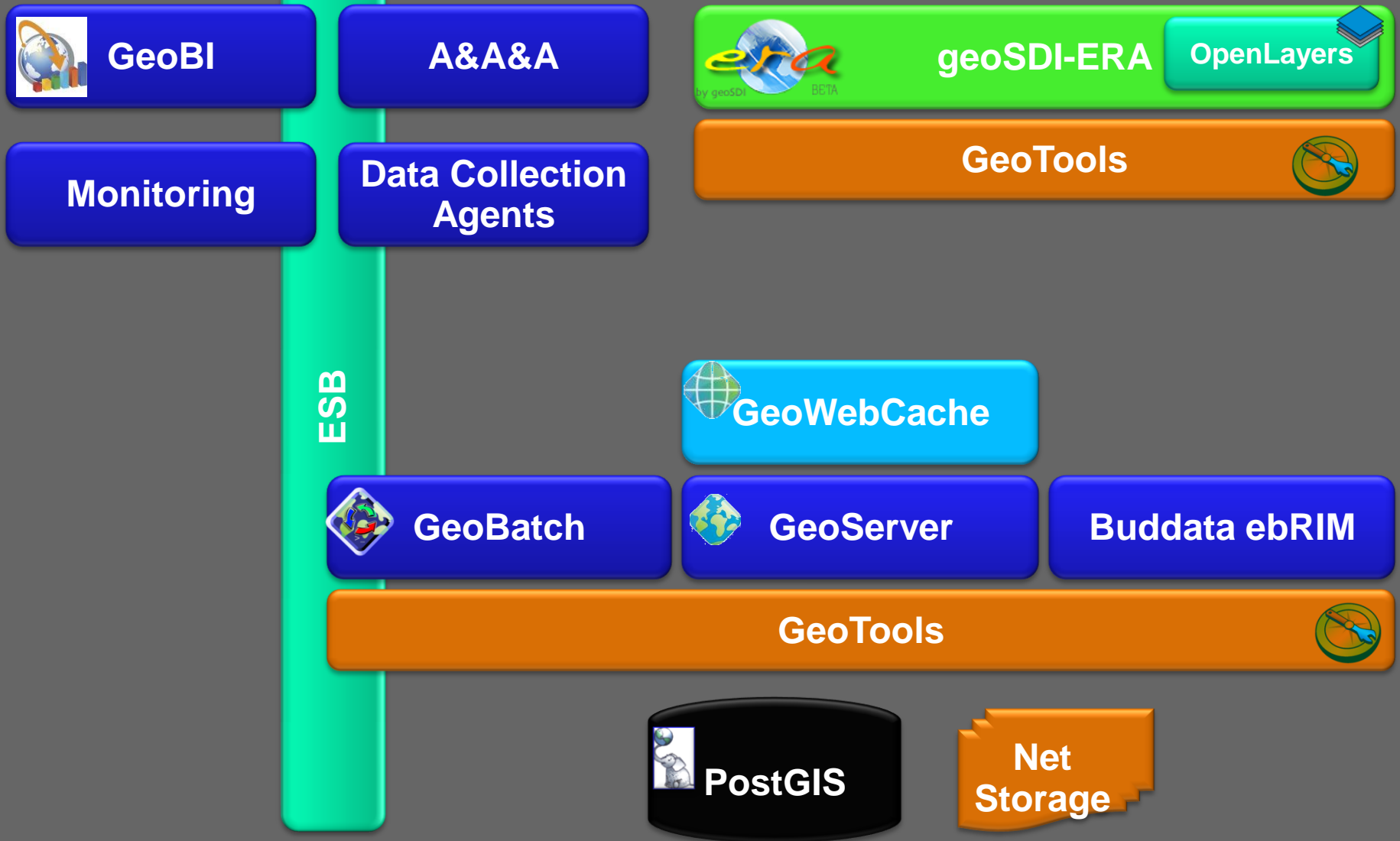
Enterprise
GeoSpatial
Server



Enterprise Spatial Data Infrastructure

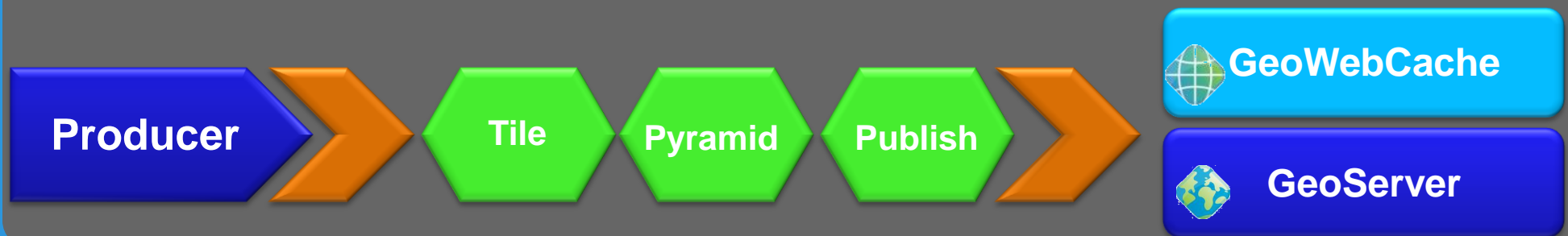


Enterprise Spatial Data Infrastructure



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

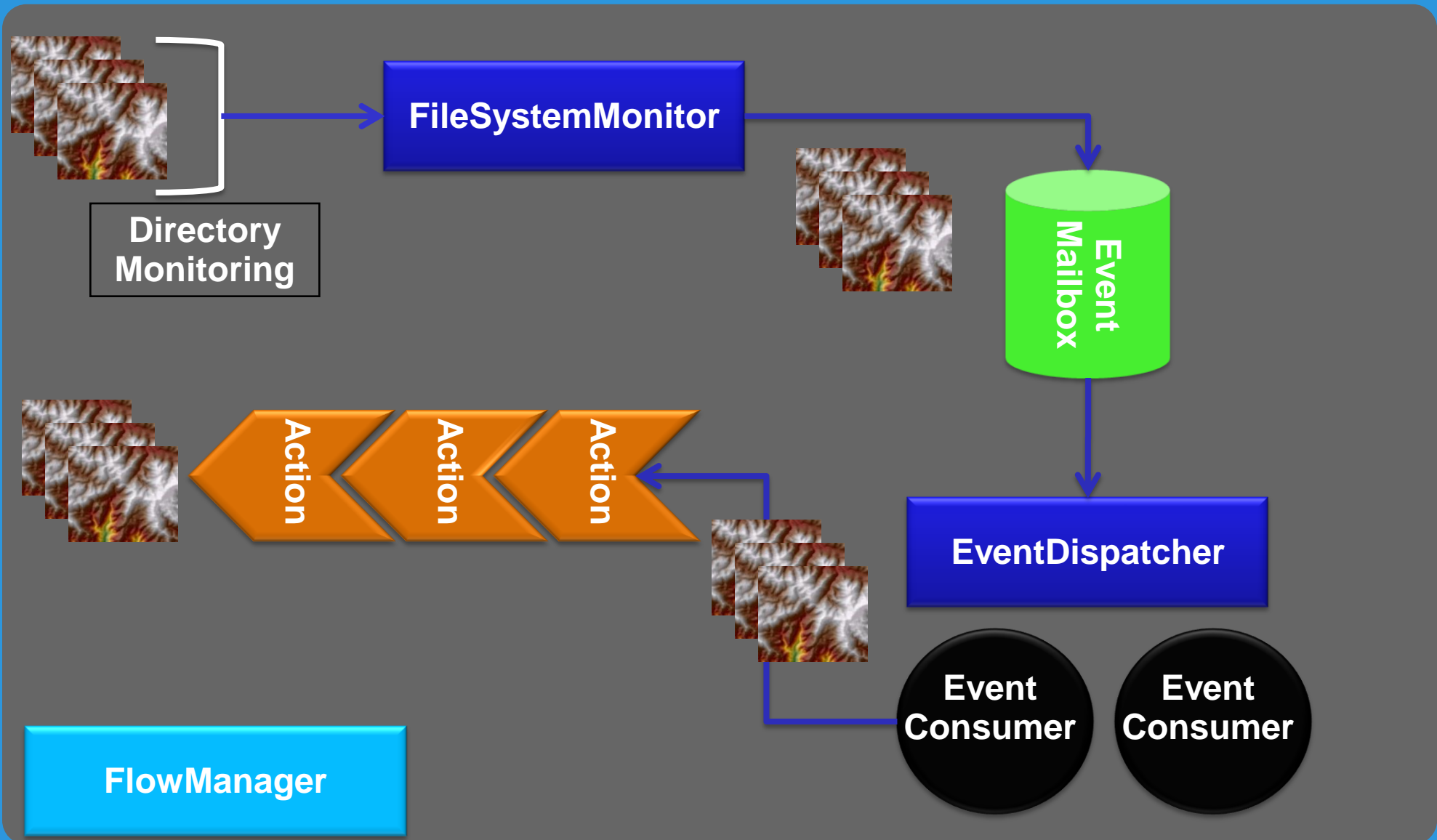


GeoBatch: Basics



- 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



GeoBatch:Sample Flow



```
- <FlowConfiguration>
  <workingDirectory>AISAnomalies</workingDirectory>
  <autorun>true</autorun>
- <EventConsumerConfiguration>
  <workingDirectory>AISAnomalies</workingDirectory>
  <performBackup>true</performBackup>
- <FileEventRule>
  <optional>false</optional>
  <originalOccurrences>1</originalOccurrences>
  <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>
  <styles />
</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>
```

GeoBatch: New UI





GeoBatch

Info
 Geobatch is an event-based geospatial aware batch processing system to ease the development, the deploy, and the management of jobs on streams of geospatial data. Geobatch provides basic components for the collection, processing and publication of data.

Flow Manager
 - A batch job in Geobatch is an XML configuration file called "flow".
 - Each flow consists of three sections: a descriptive part, a second one which is dedicated to data streams monitoring (eventGeneratorConfiguration) and the final one which is devoted to recognition of particular files within a stream, its elaboration and final publication (eventConsumerConfiguration).

Job Execution and Management
 - A simple web application can be used to launch jobs, view the list, check the status, start, stop or dispose
 - Flow can start automatically on application server startup.

 **Manage Flows**

 **Manage FTP Server**

GeoBatch

Manage Flows

ID	DESCRIPTION	INPUT DIR	WORKING DIR	STATUS	ACTION
flowASAnomales	flowASAnomales: testing rejection of ASAnomales shapefiles through Hibernate Spatial	ASAnomales/in	ASAnomales	●	 ●
ASCoverage	flowASAnomales: testing rejection of ASAnomales shapefiles through Hibernate Spatial	ASCoverage/in	ASCoverage	●	 ●
ASforecast	flowASAnomales: testing rejection of ASAnomales shapefiles through Hibernate Spatial	ASForecast/in	ASForecast	●	 ●

GeoBatch

Manage FTP Server

[Add new user](#)

USERID	PASSWORD	HOME DIRECTORY	WRITE PERMISSION	UPLOAD RATE	DOWNLOAD RATE	ACTIONS
admin	admin	admin	true	0	0	

Copyright © 2005 - 2009 GeoSolutions.

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



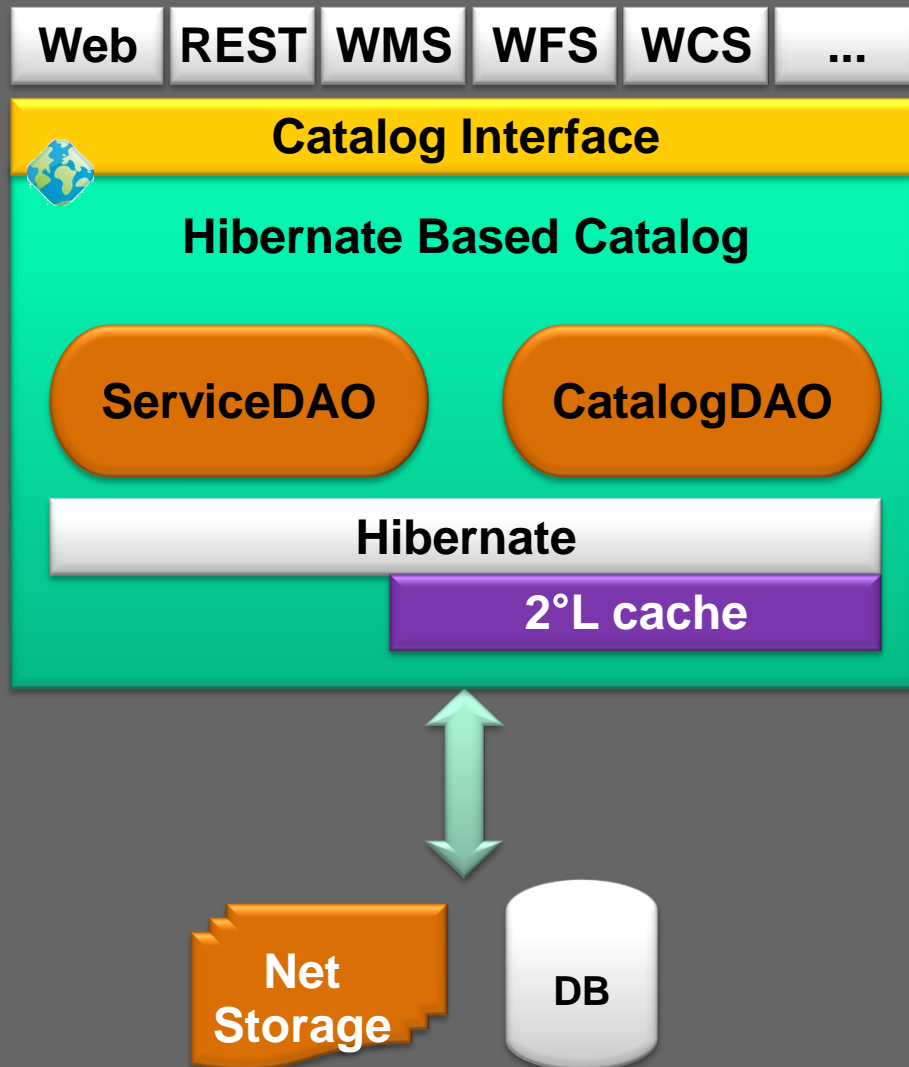
- **GeoSpatial enterprise gateway**
 - Management raster and vector data
- **Standards compliant**
 - OGC WCS 1.0 - 1.1.1 (RI)
 - OGC WFS 1.0 - 1.1 (RI)
 - OGC WMS 1.1.1
 - OGC WPS 1.0
- **Google support**
 - KML, GeoSearch, etc..
- **Enterprise ready (almost!)**

GeoServer: Limits/Shortfalls



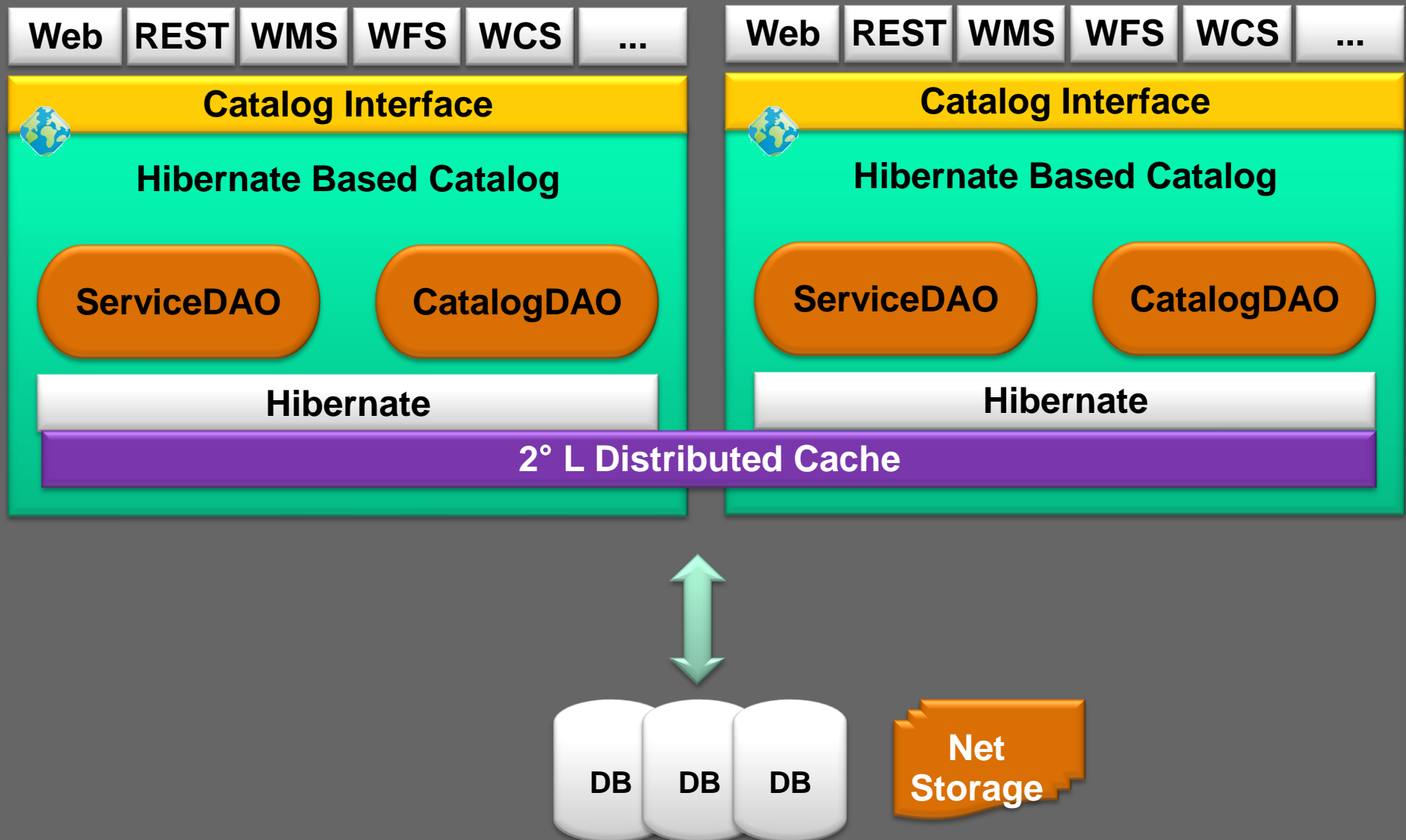
- **In-memory internal catalog**
 - Non transactional, Scale sub-optimally
 - Replication is problematic
- **File Based Configuration**
 - Non transactional, Scale sub-optimally
 - Replications is problematic
- **File based Coverage support (can be overcome easily)**
- **1 CoverageStore → 1 Coverage (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

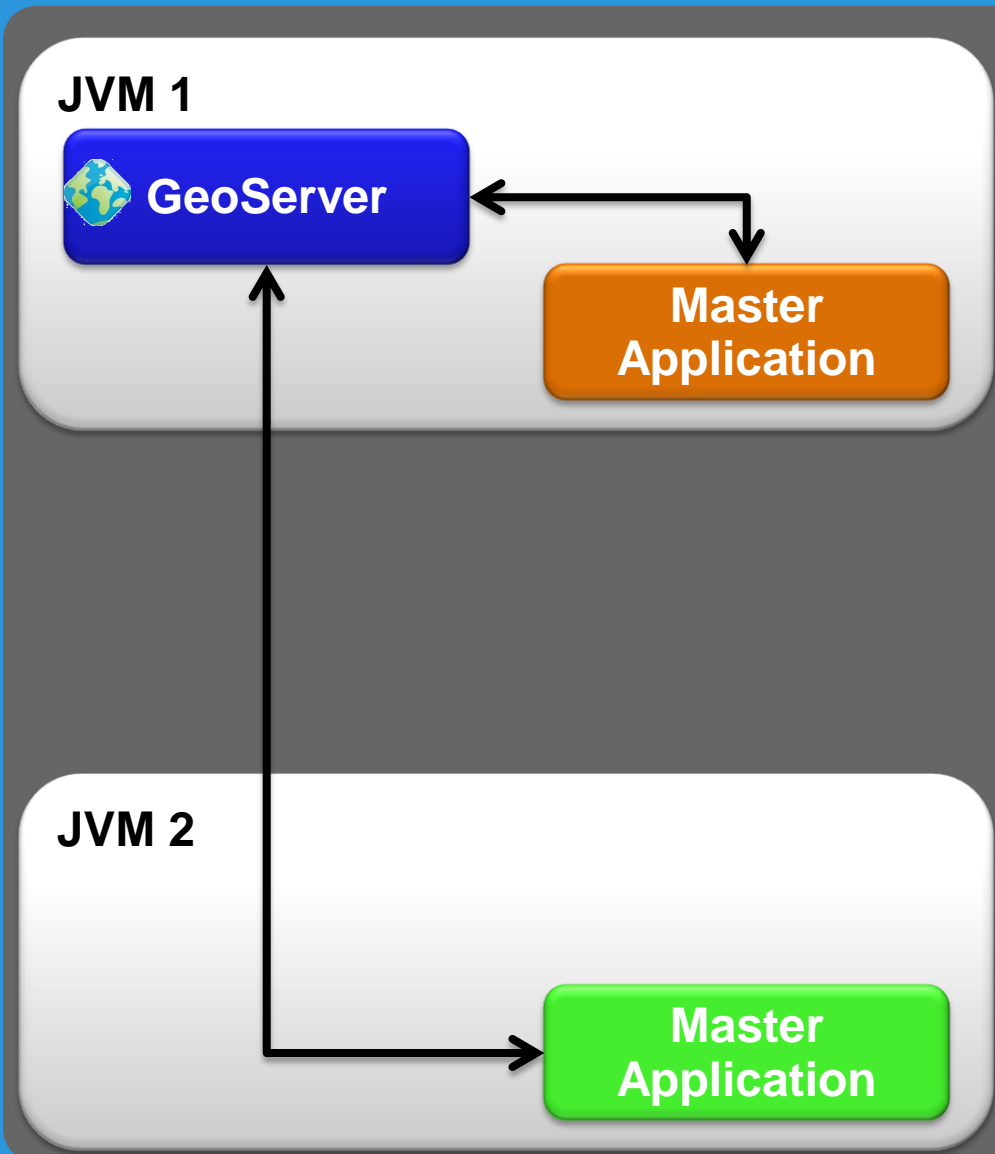


- Configuration stored in DB
 - Hibernate
 - Spring based Dao
- Transactions via Spring/EJB3
- Caching with EhCache
- Same catalog interface, different implementation
- Catalog interface needs tweaking
- GeoServer UI still needs tweaking
- Transaction demarcation needs tweaking

GeoServer: Hibernate catalog

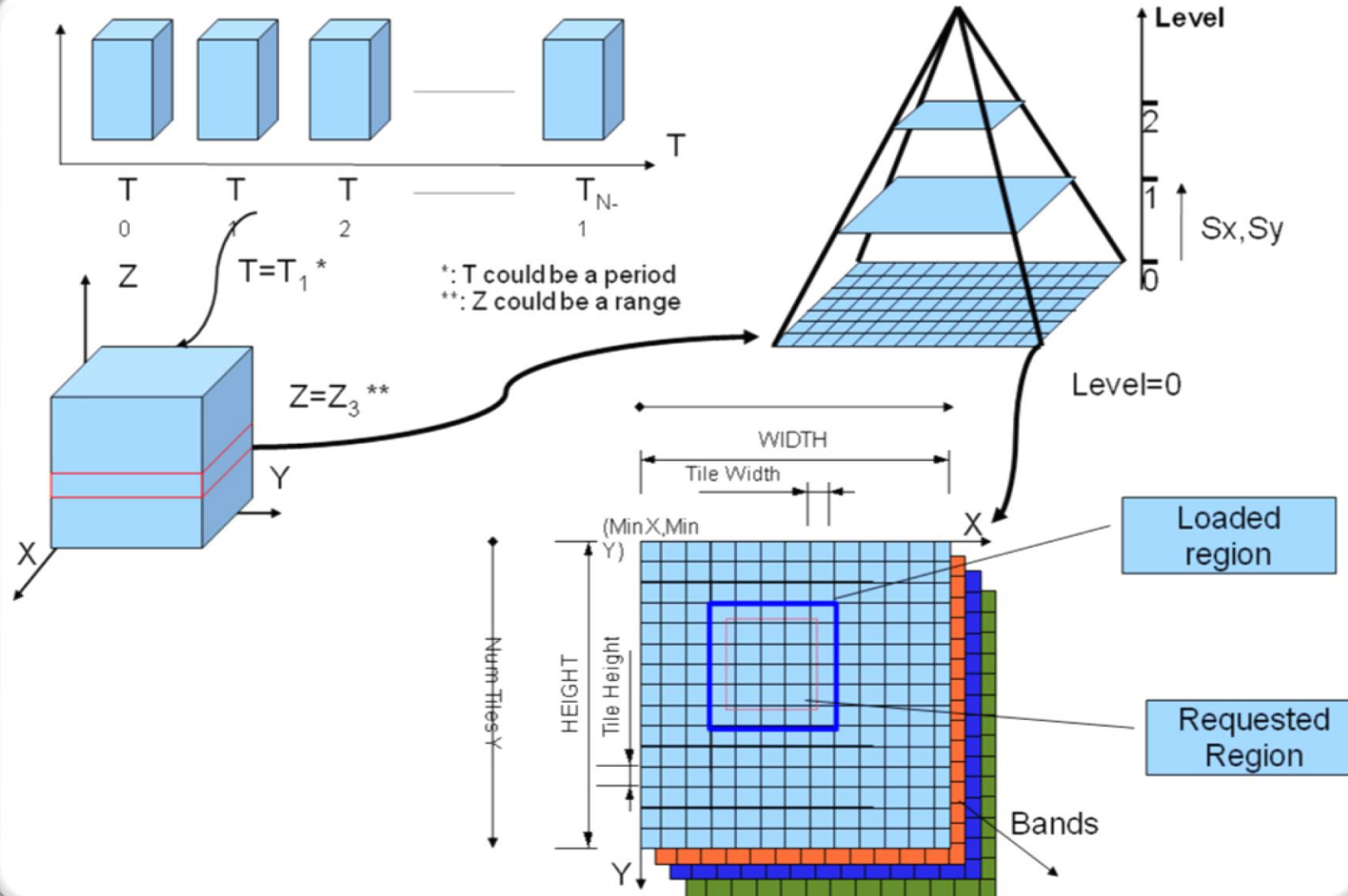


GeoServer: Transactional API



- 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



Web REST WMS WFS WCS ...

Catalog Interface

CoverageStore Interface

RasterLayerDAO

RangeDAO

Hibernate Spatial

2°L cache

Net Storage

DB

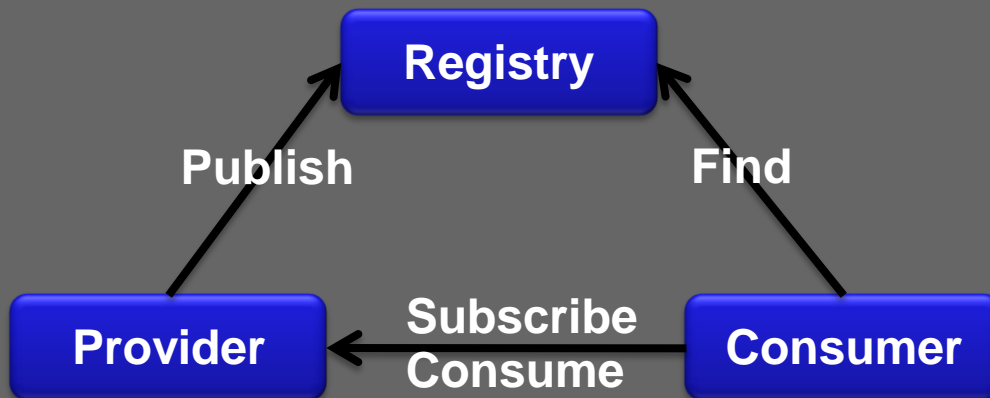
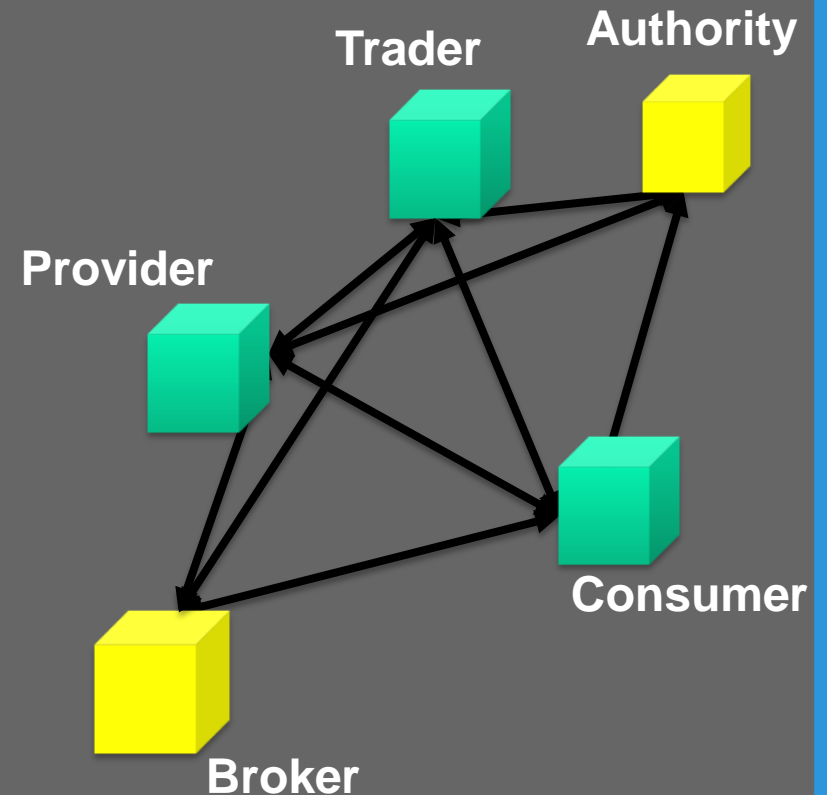
GeoServer

CoverageStore

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

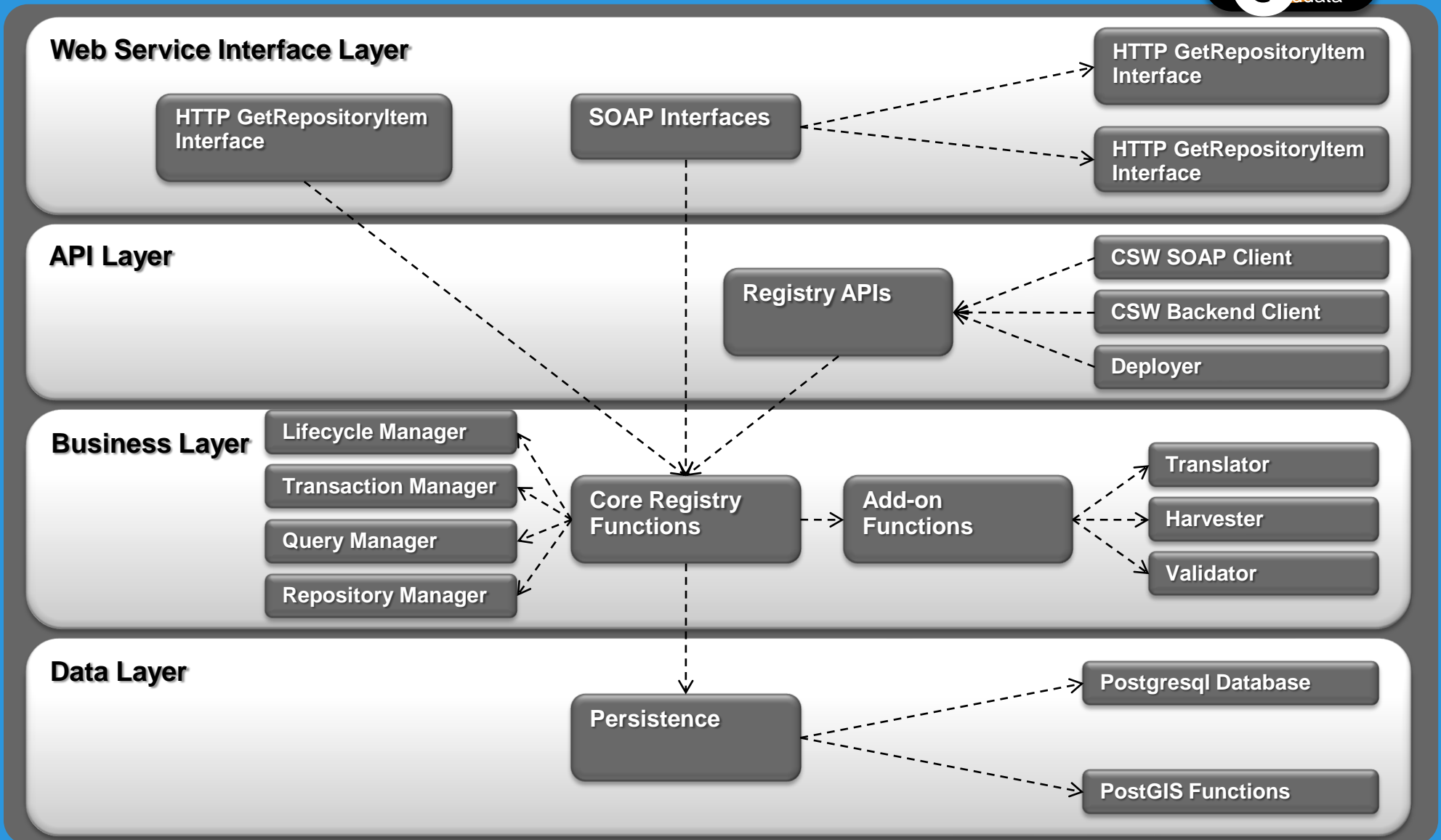


ERGO Buddata EbRR



- 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)

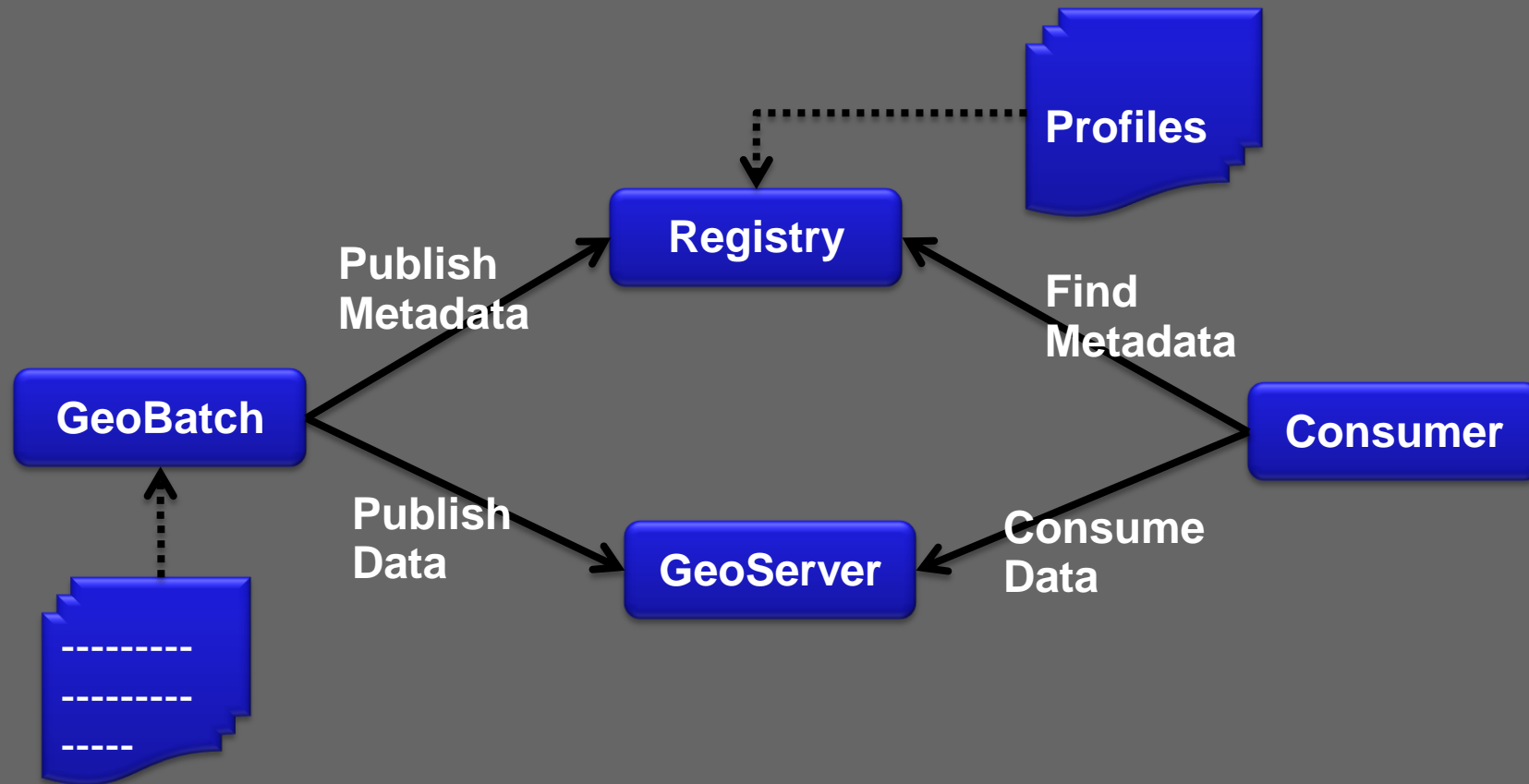
EbRR: Architecture overview



Buddata EbRR @ Work



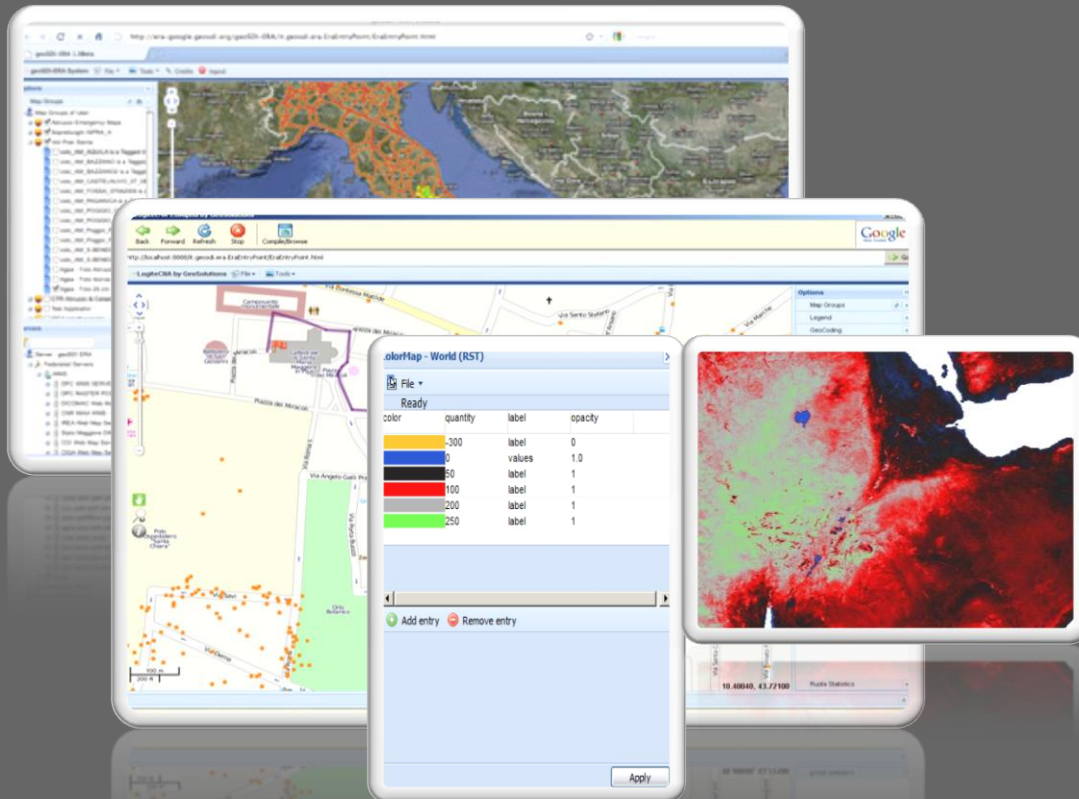
- Metadata published with data contextually
- Use of custom profiles



GeoPortal: geoSDI - ERA

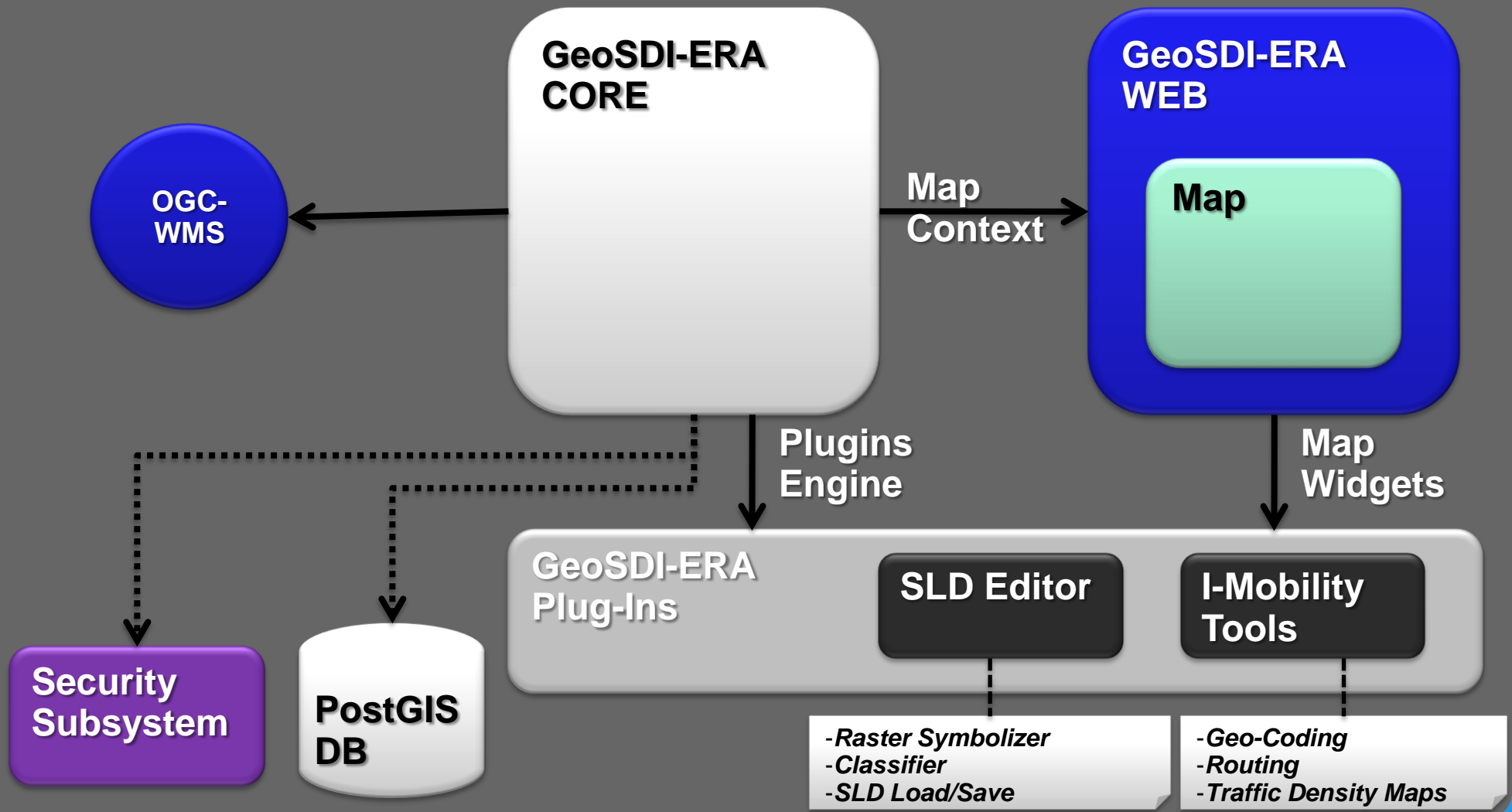


- Modular/Extensible Rapid-mapping Web Application
- EXT-GWT + Openlayers + Hibernate Spatial + Spring
 - Core components to perform basic actions
 - OpenLayers integration
 - Layers Wide search
 - Layer Tree
 - Security subsystem
 - Style editor (ongoing)
 - Plugin Engine to perform specific tasks
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



geoSDI – ERA Modular

- GeoSDI-ERA Architecture Schema

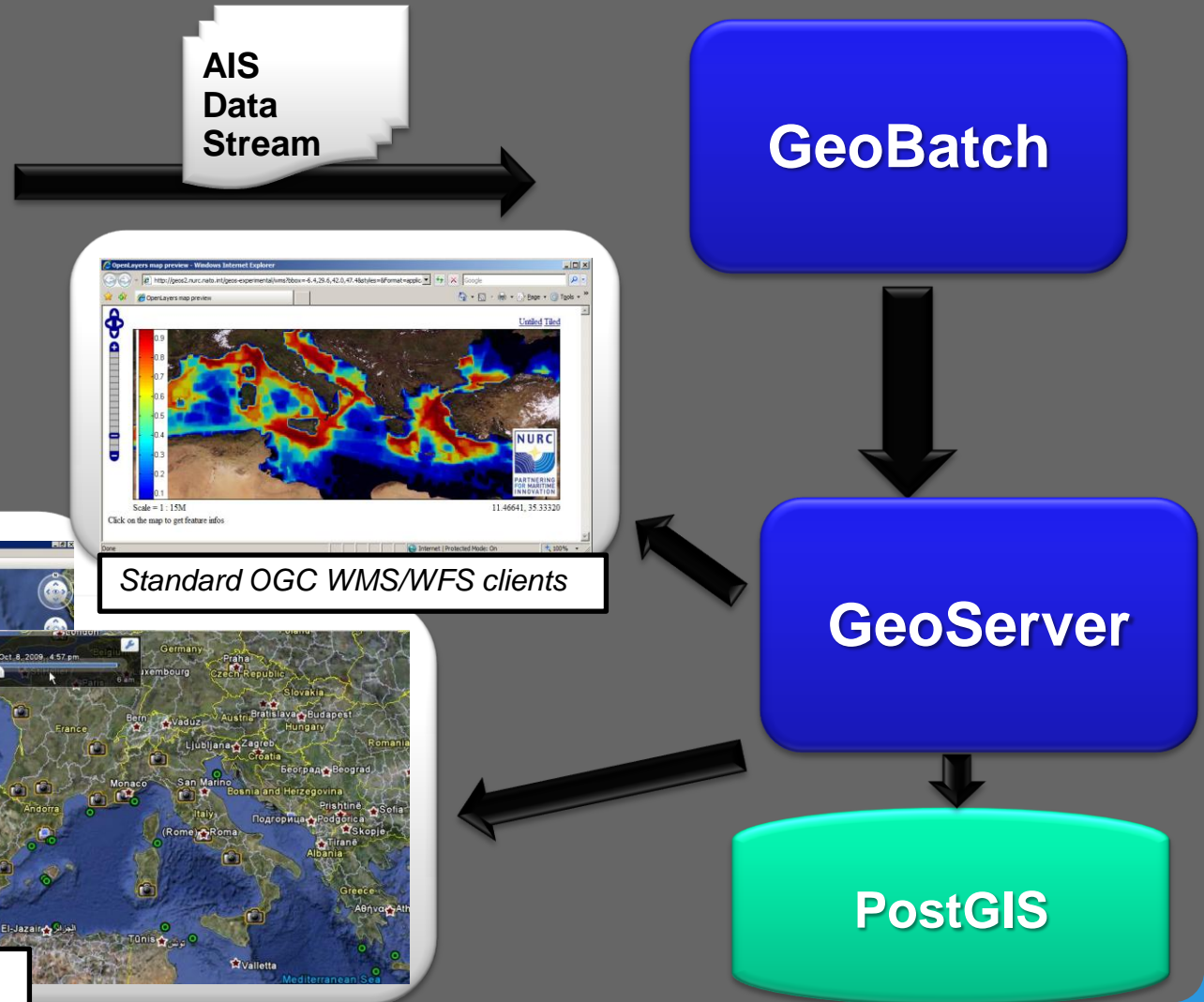


Use Case: AIS data management

- **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*

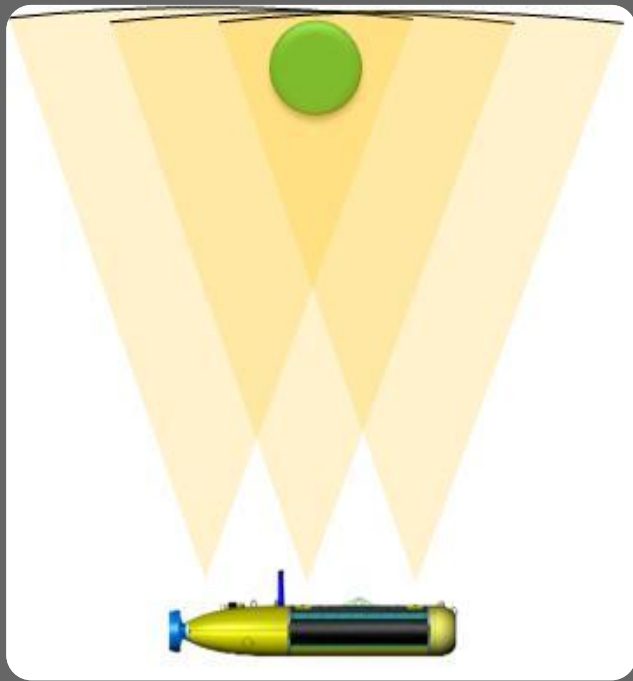
Use Case: AIS data management

- 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



Data and Metadata Management

GeoServer

Registry
ebRIM

WFS

WCS

WMS

Middleware Tier

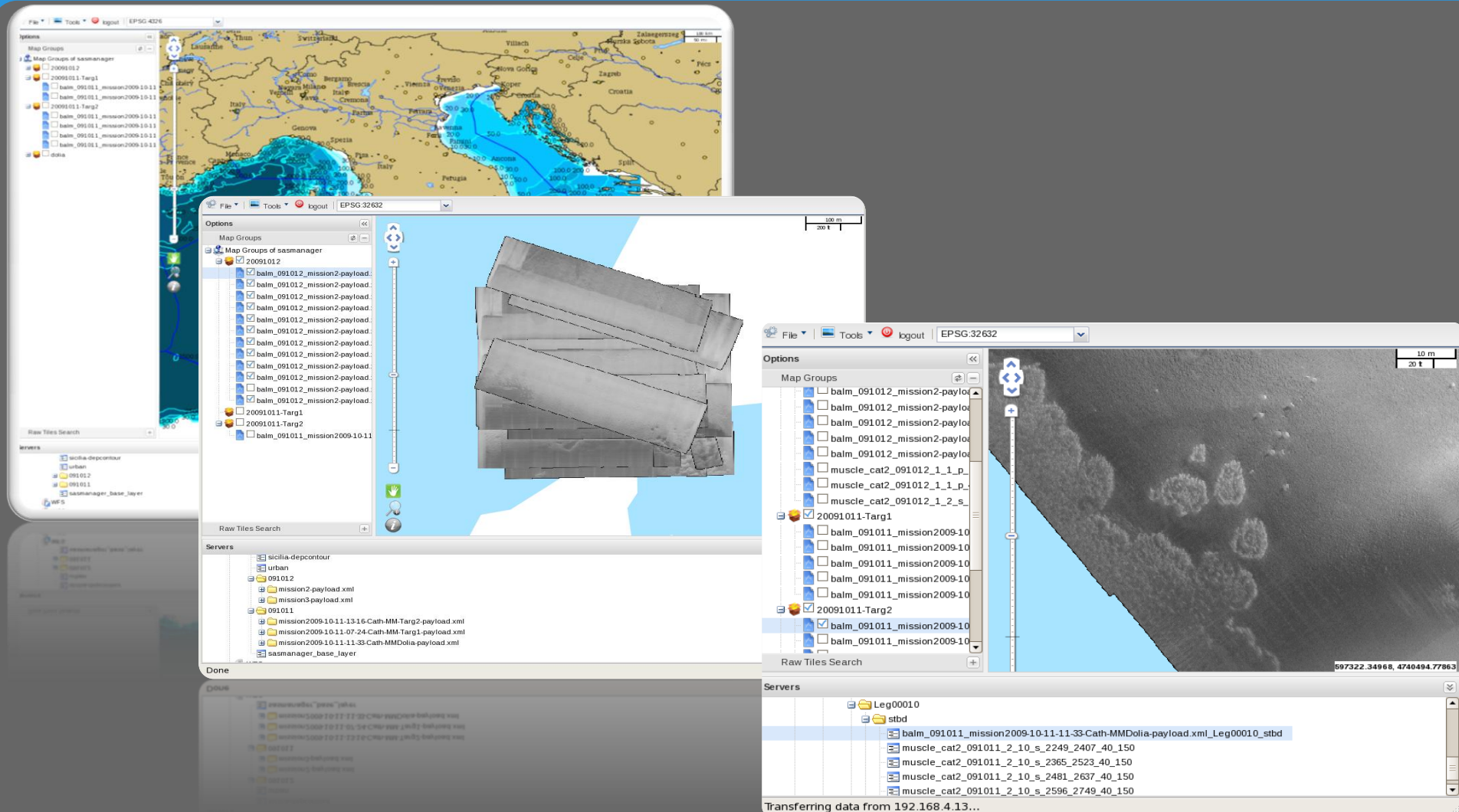
Network
Area
Storage

GeoBatch

Mission Data Package

DataIngestion Tier

Use case: SAS Mission Manager

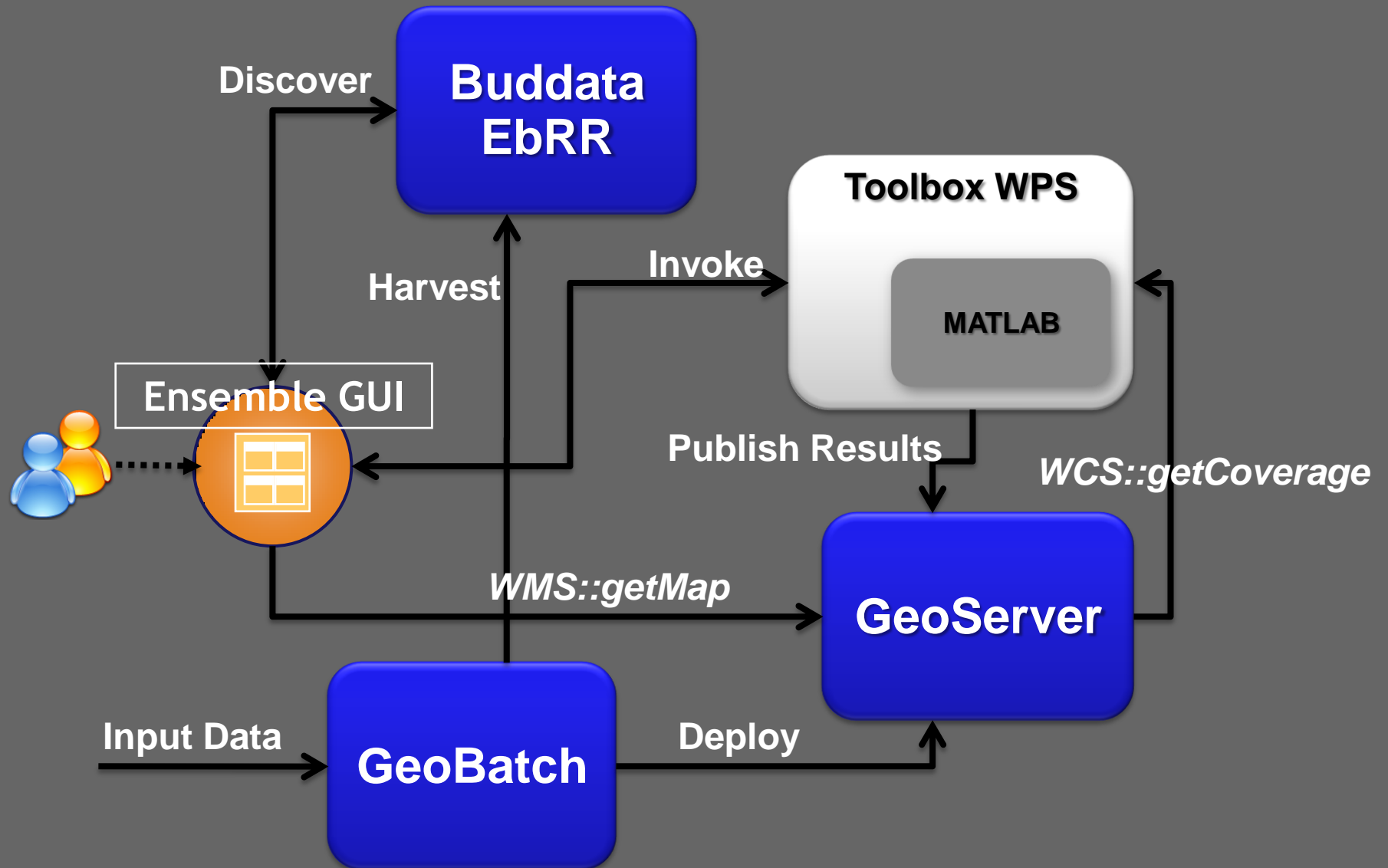


Sonar images are Courtesy of NURC

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

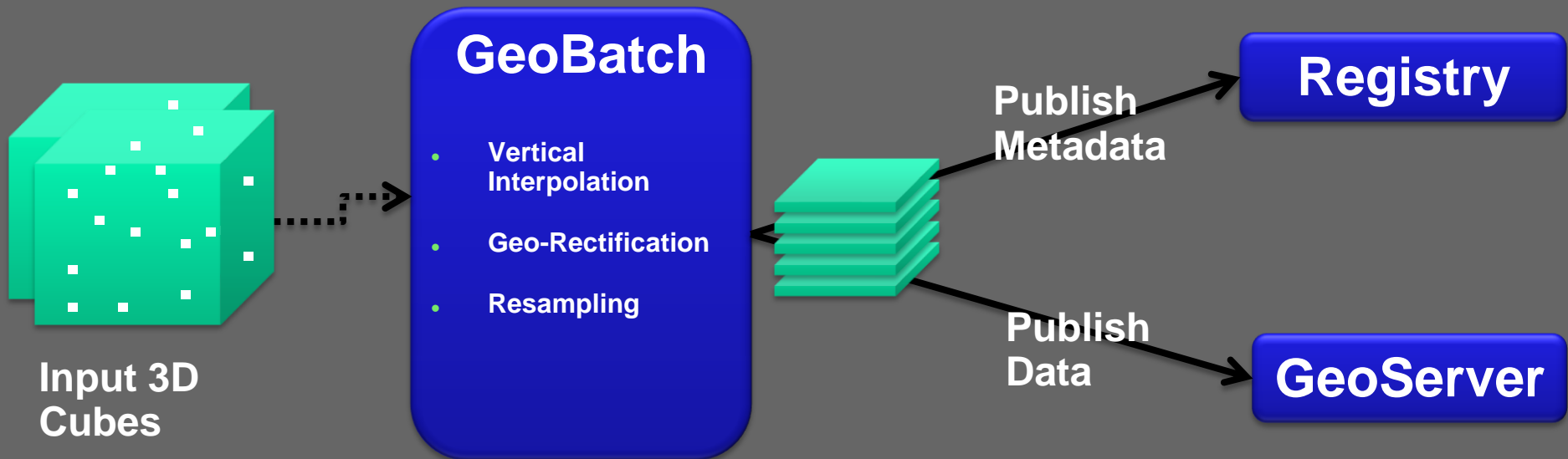


FOSS4G 2009, Sydney
23rd September 2009

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



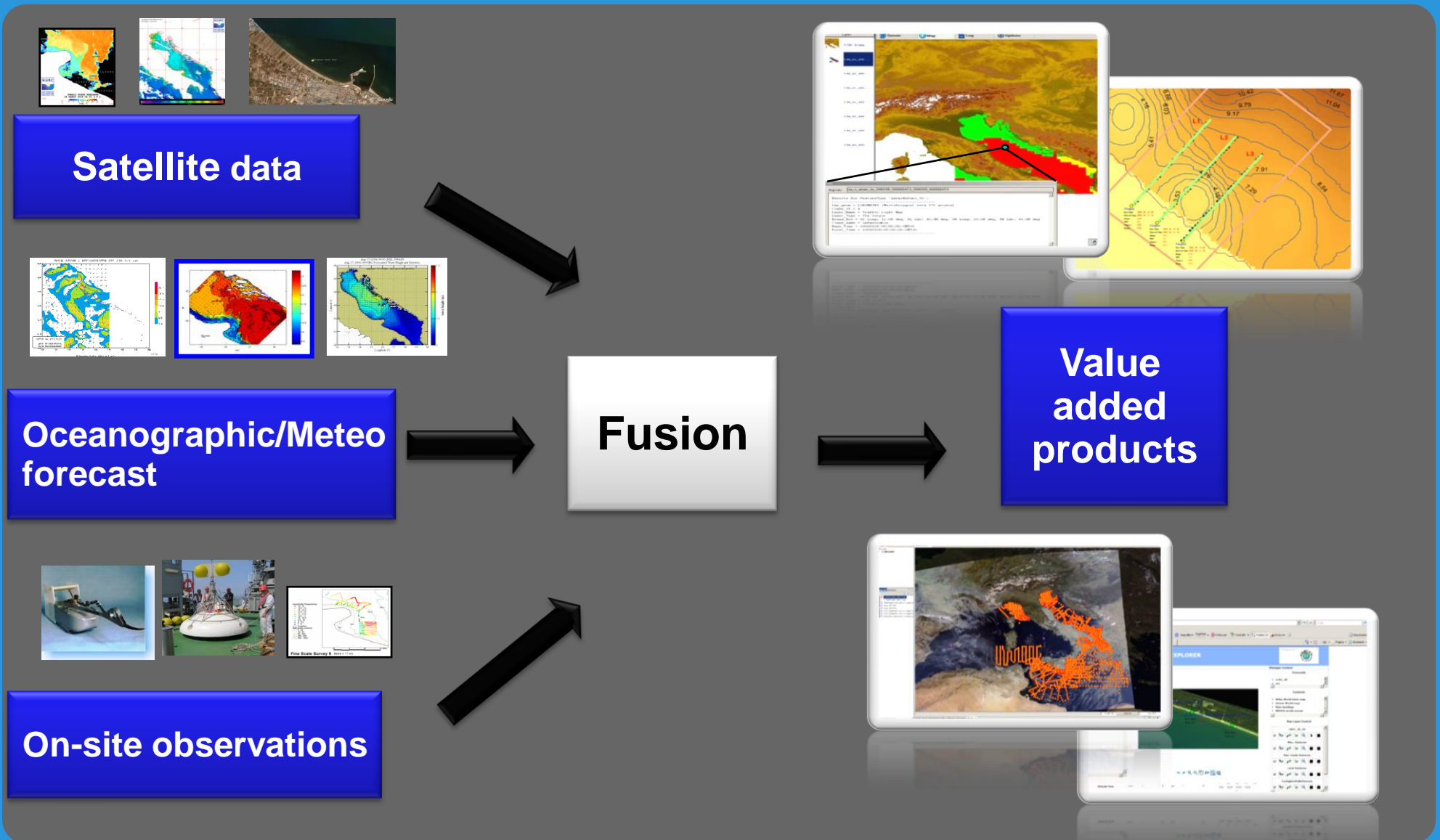
- **Next Step**

- Handle multidim coverages through GeoServer + RasterVault

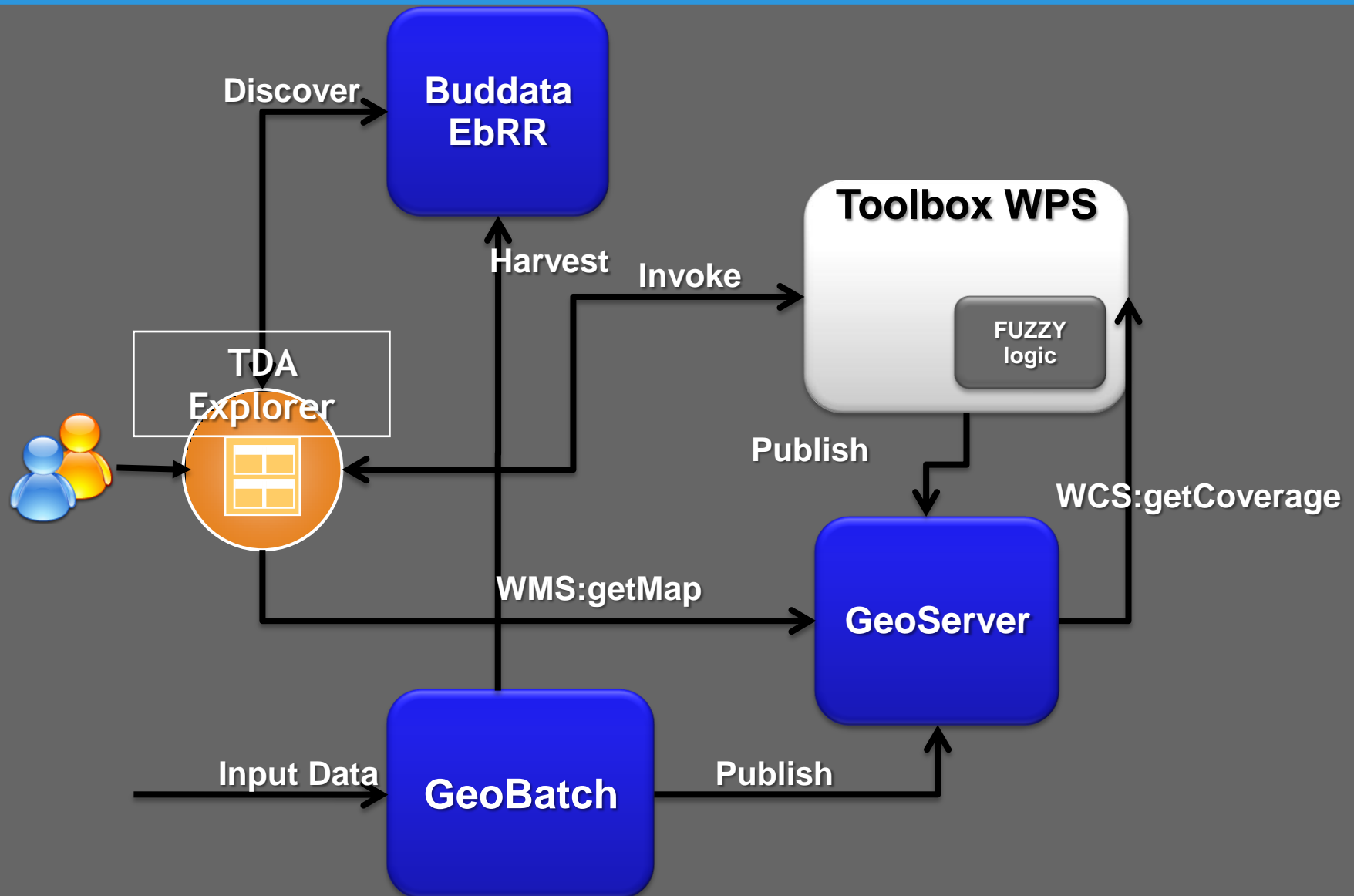
Use Case: Super Ensemble Modeling

The image displays three overlapping screenshots of the Ensemble Modeling software interface. The top-left screenshot shows the 'Parameter Info' tab with 'OCEAN' selected as the discipline and 'mean_wav_dir' as the geophysical parameter. The middle screenshot shows a 'Query Informations' dialog box displaying 'MEAN: 17.0174560546875' for a selected layer. The bottom-right screenshot shows the 'Geographical Filter' tab with a defined area of interest and a 'Grid Coverages List' panel containing model names like 'INGV-MFS_Sys2b-MODEL', 'NRL-NCOM-MODEL', and 'MERCATOR-MODEL'.

Use Case: TDA



Use Case: TDA



Use Case: TDA

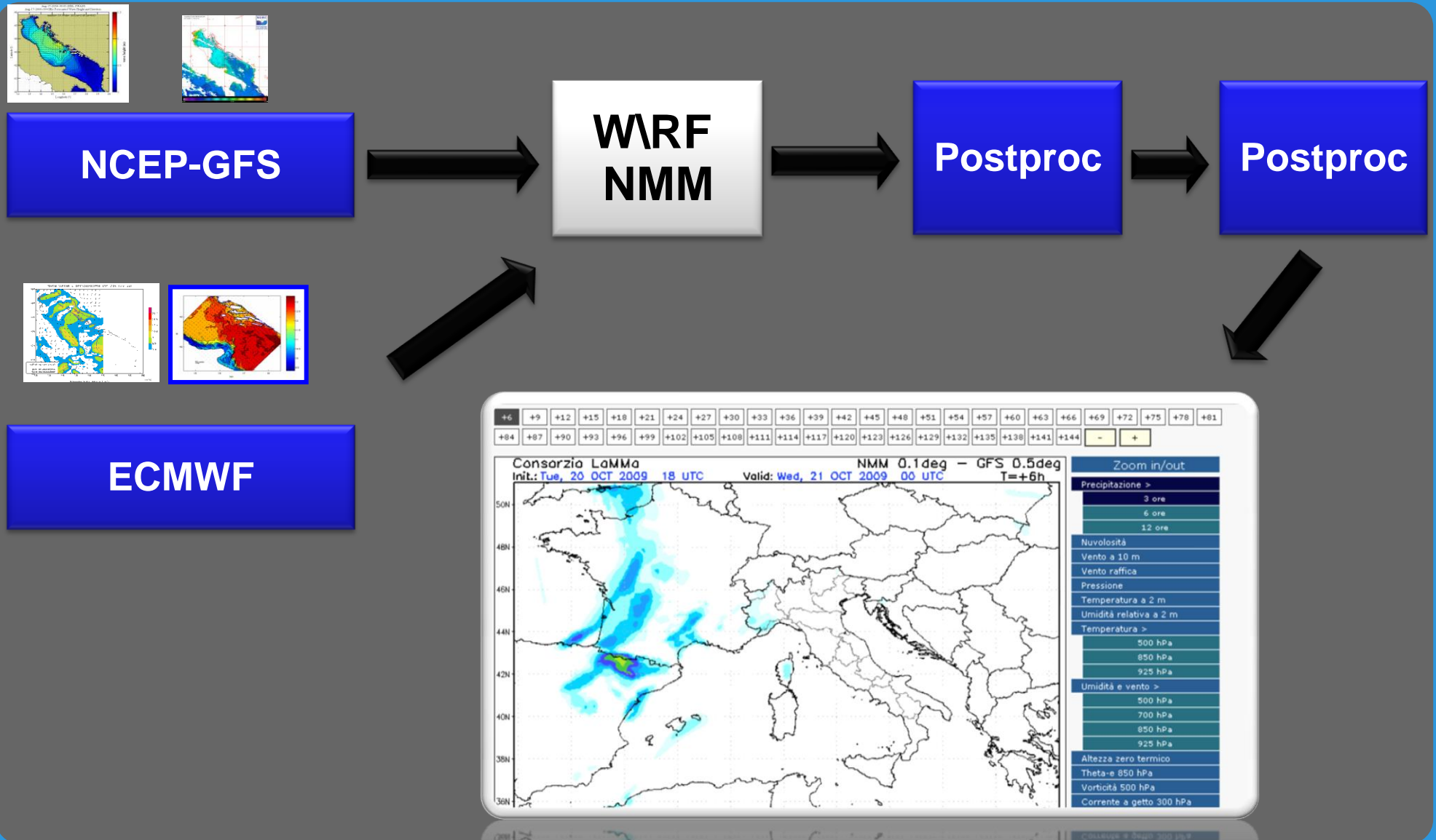
Search page

WPS invocation

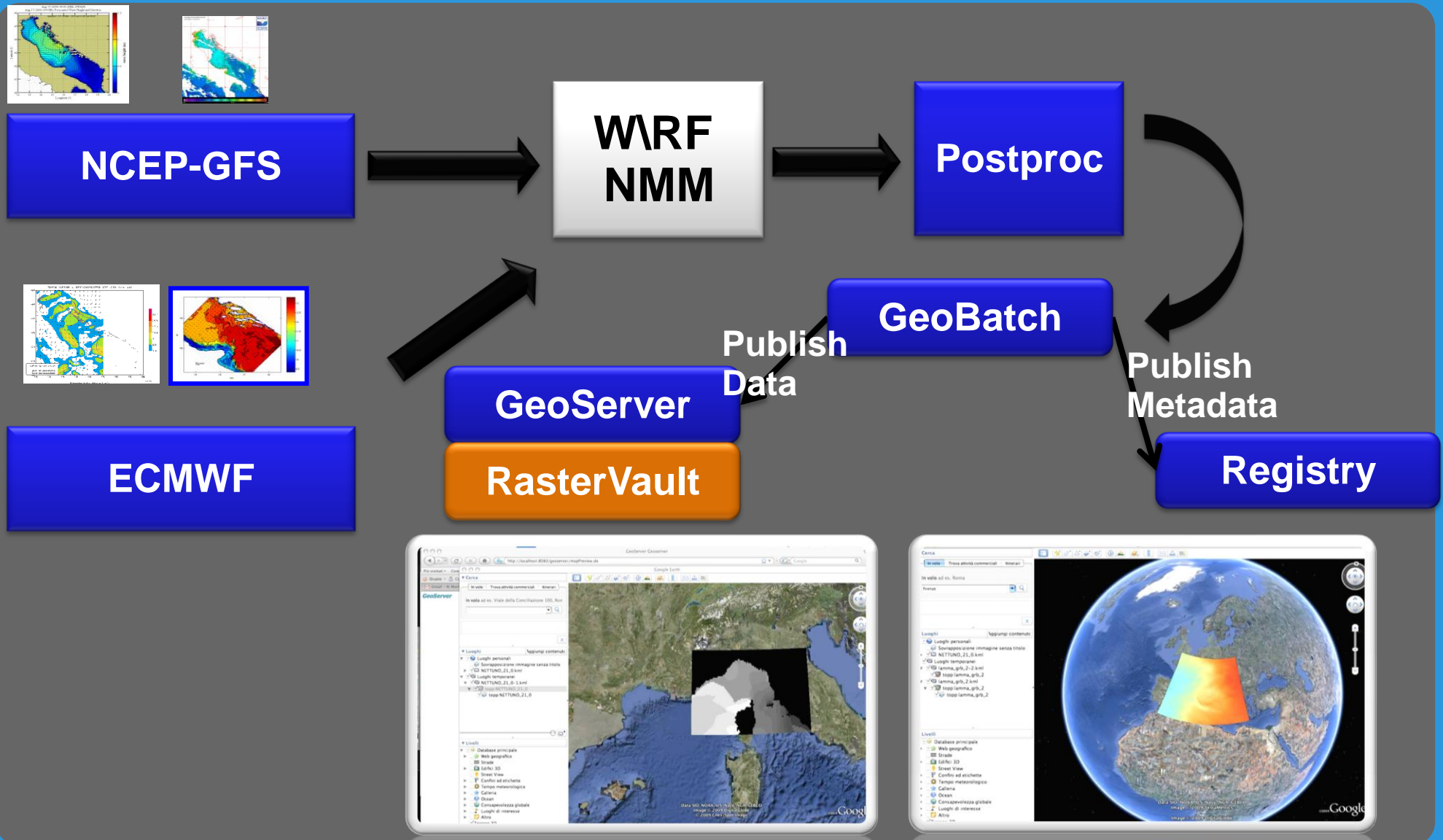
Catalogue Search Results

ProductIdentifier	Platform
EN1-05010109134614-31800 F	Envisat
EN1-05010109163380-31800 F	Envisat
EN1-05010109192147-31800 F	Envisat
EN1-05010110542205-31800 F	Envisat
EN1-05010110595738-31800 F	Envisat
EN1-05010111024504-31800 F	Envisat
EN1-05010112485629-31800 F	Envisat
EN1-05010112514395-31800 F	Envisat
EN1-05010112543162-31800 F	Envisat
EN1-05010112571878-31800 F	Envisat

Use Case: LAMMA

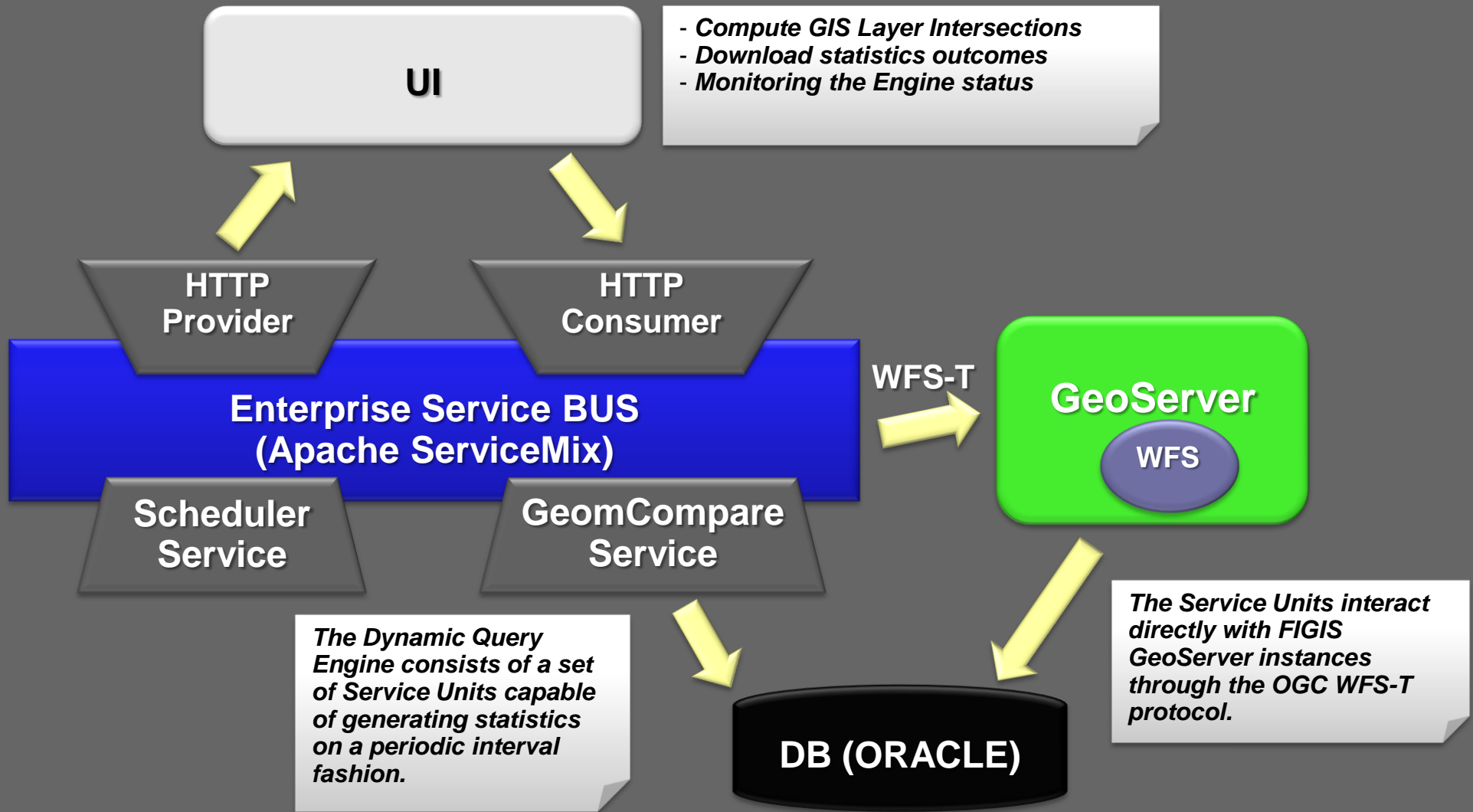


Use Case: LAMMA



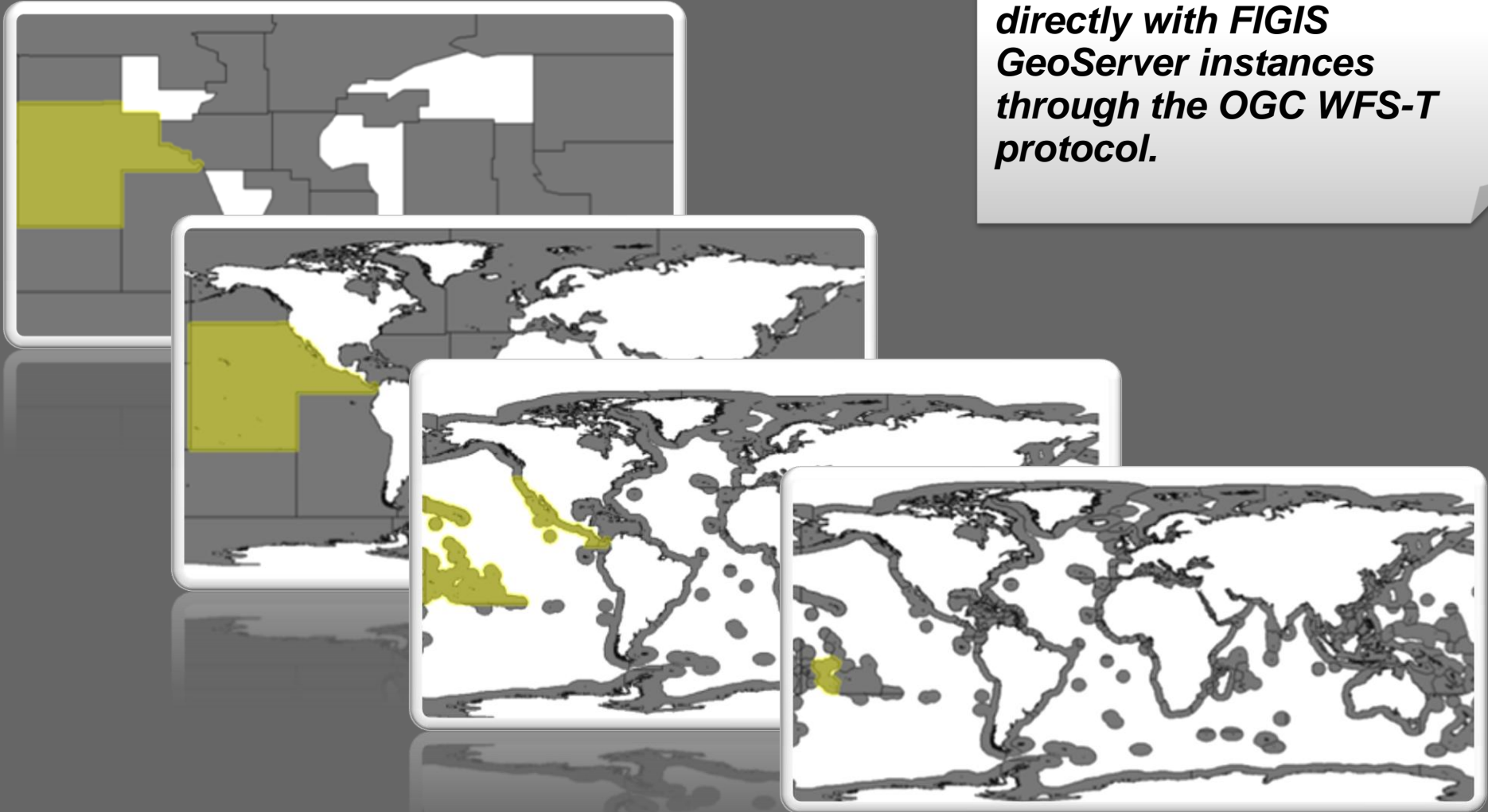
Use Case: FAO FIGIS

- Dynamic Statistical Query Engine



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

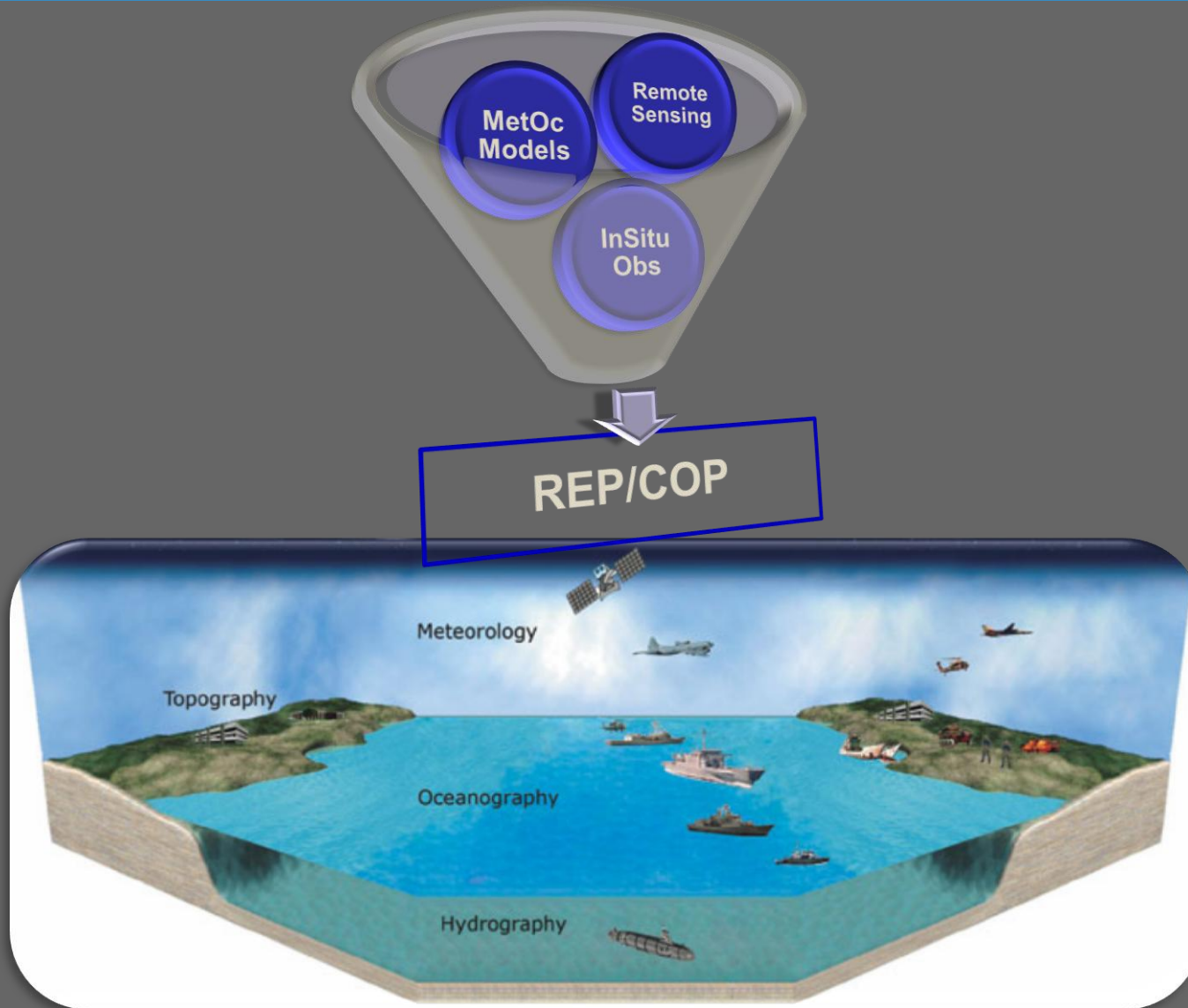
Ing Simone Giannecchini



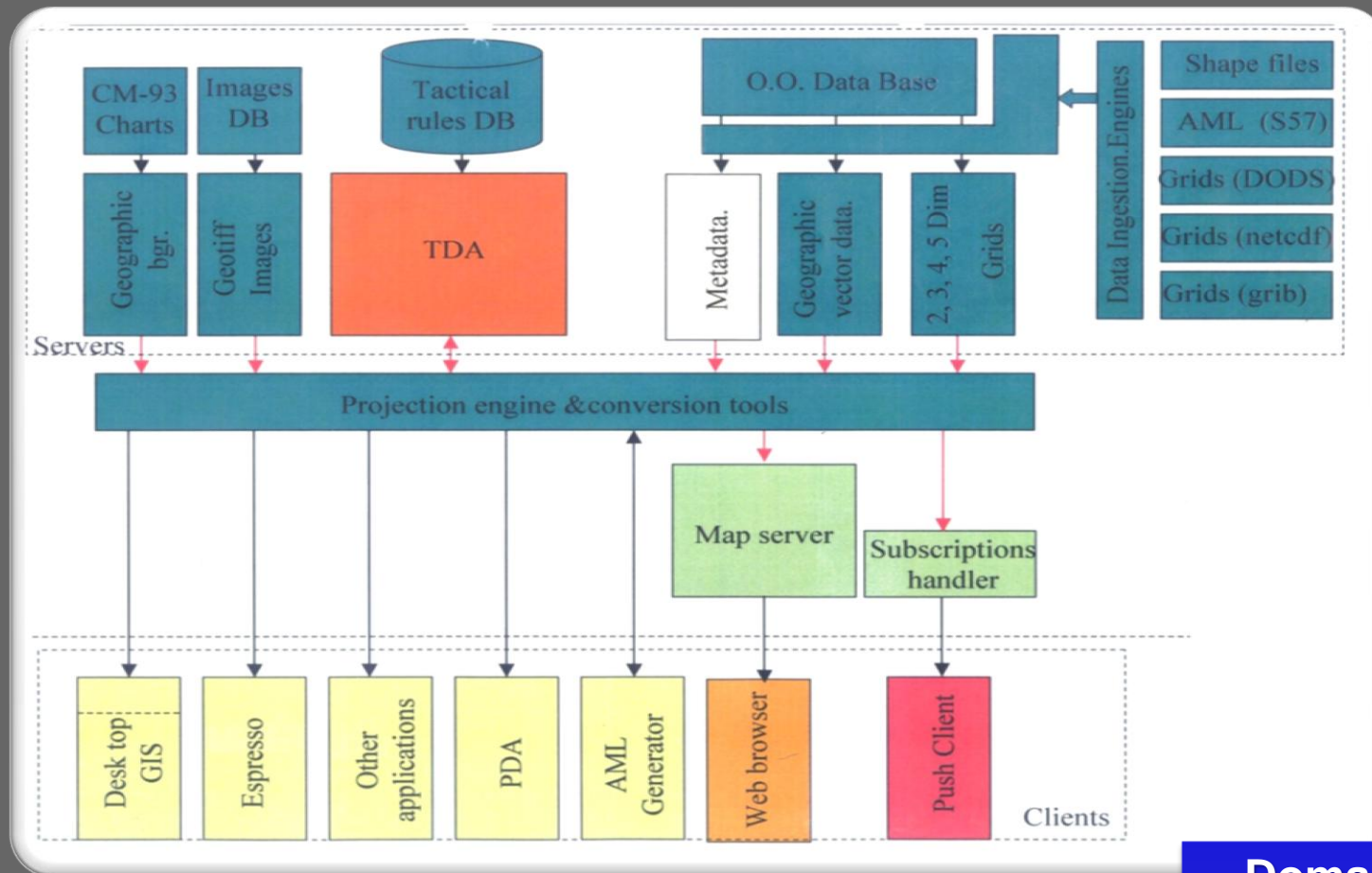
FOSS4G 2009, Sydney
23rd September 2009



Reference Scenario/Domain



Generic Enterprise Infrastructure



Domain Expert
View
2001

GeoBatch: RoadMap

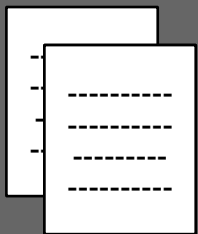


- **Improved Event Hierarchy**
 - Generic Event Subsystem
 - Time Based Events
- **Improved UI**
 - Visual Builder for flows
- **Refactor Messaging Subsystem**
 - JMS
 - ESB (ServiceMix)
- **Actions for GeoWebCache**
- **SOAP based Event generator**
- **Integrate with GeoServer ?**



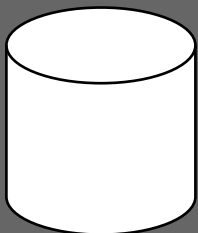
GeoServer

Shapefile



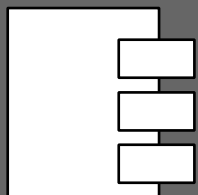
Vector files

PostGIS
DB2
Oracle
MySQL



DBMS

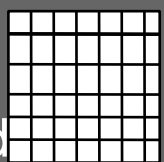
ArcSDE
WFS



Servers

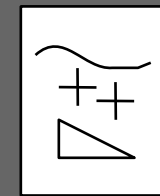
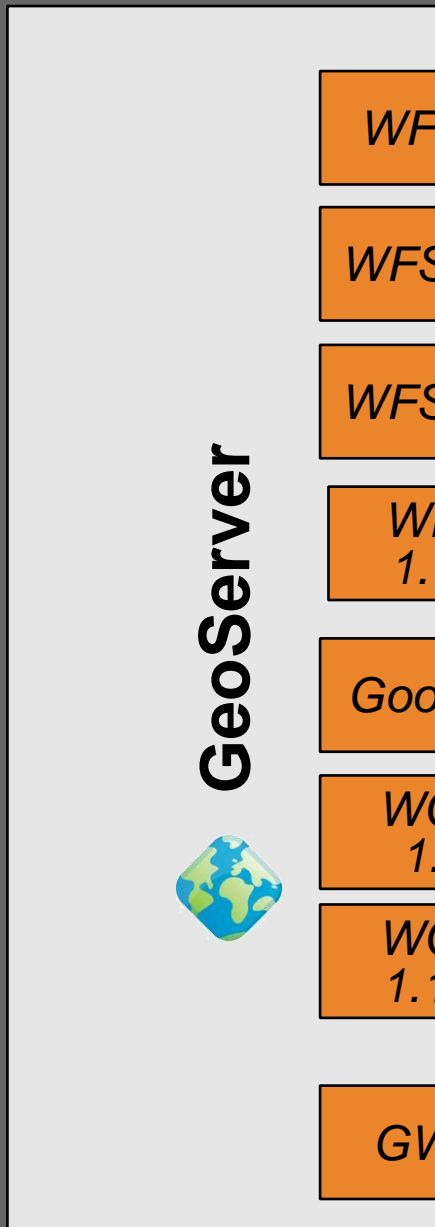
GeoTIFF

ArcGrid
GTopo30
Img+world
Mosaic
MrSID
JPEG 2000



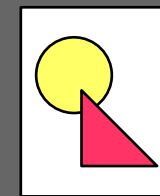
Raster files

ECW, Pyramid



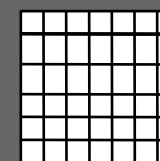
Raw vector data

Shapefile.zip
GML2
GML3
GeoRSS
GeoJSON
CSV/XLS



Styled maps

PNG, GIF
JPEG
TIFF,
GeoTIFF
SVG, PDF
KML/KMZ

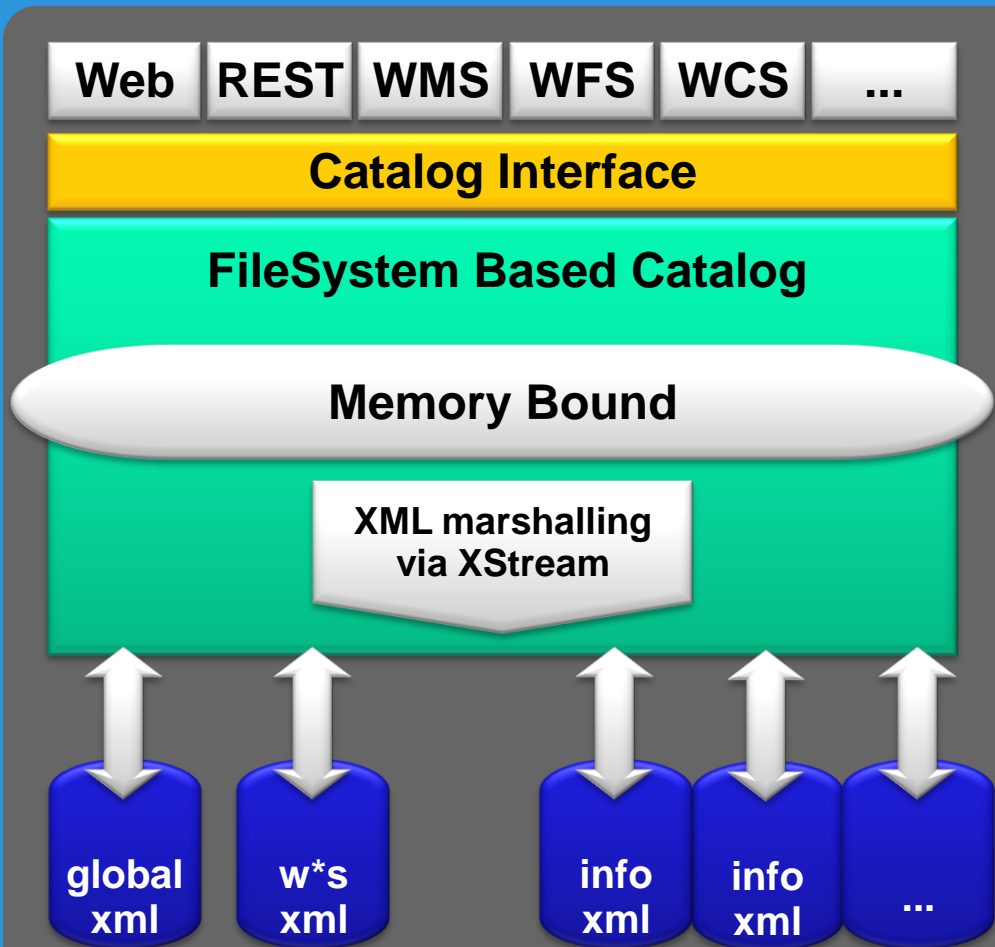


Raw raster data

GeoTIFF
ArcGrid
GTopo30
Img+World

KML superoverlays
Google maps tiles

Old configuration: Filesystem catalog

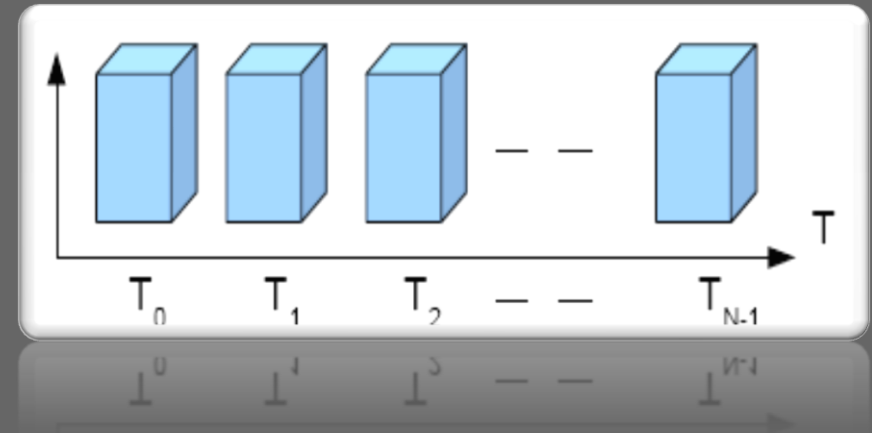


- All the catalog objects are kept in memory: cannot deal with a huge amount of data.
- Filesystem persistence is not transactional
- External applications have to pass through GS to get information.

GeoServer: Advanced Raster

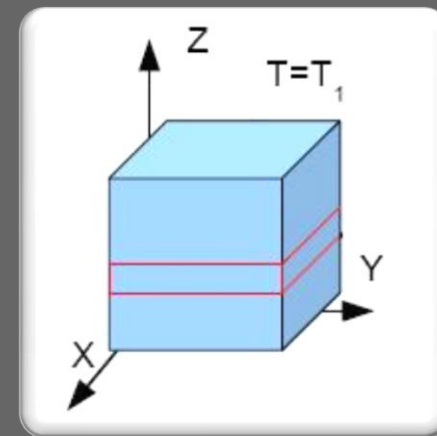
- **Time Management**

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



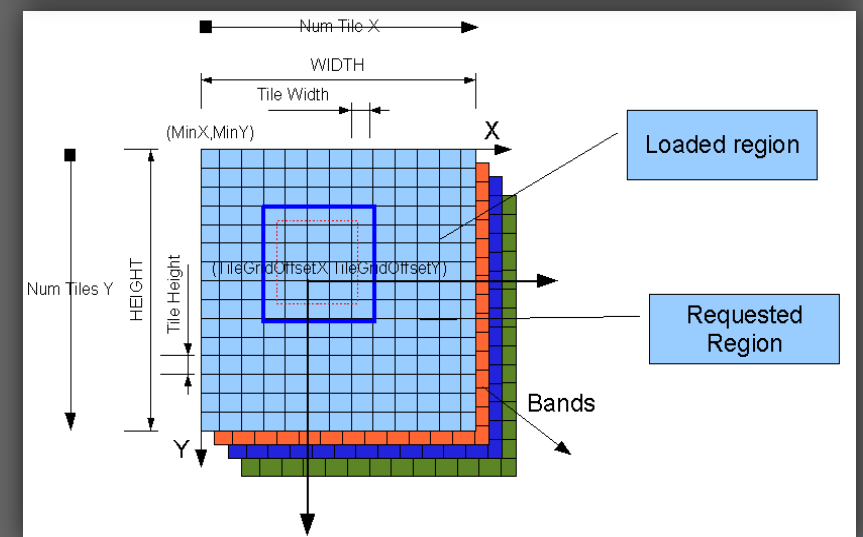
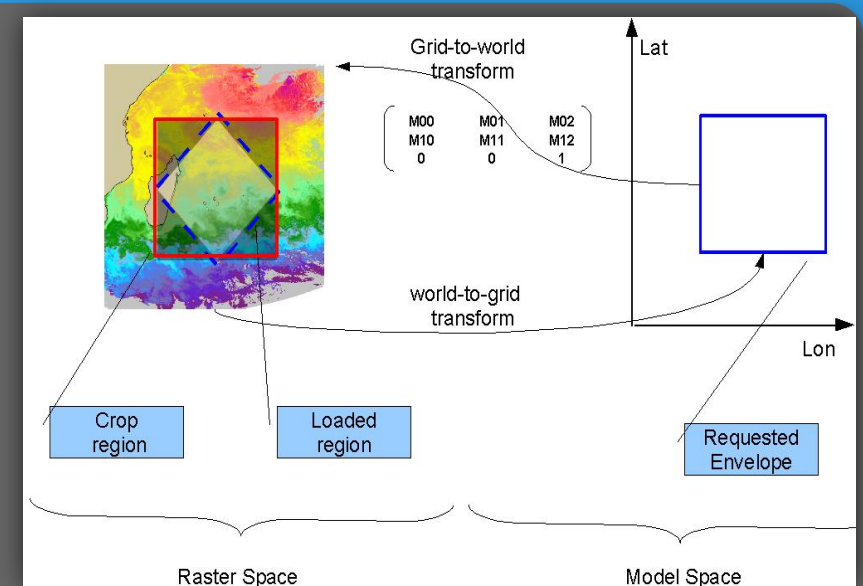
- **Elevation Management**

- Sequence of 2D layers
- Positions values with Nearest Neighbor interpolation
- B-Trees (logarithmic access) to improve performances

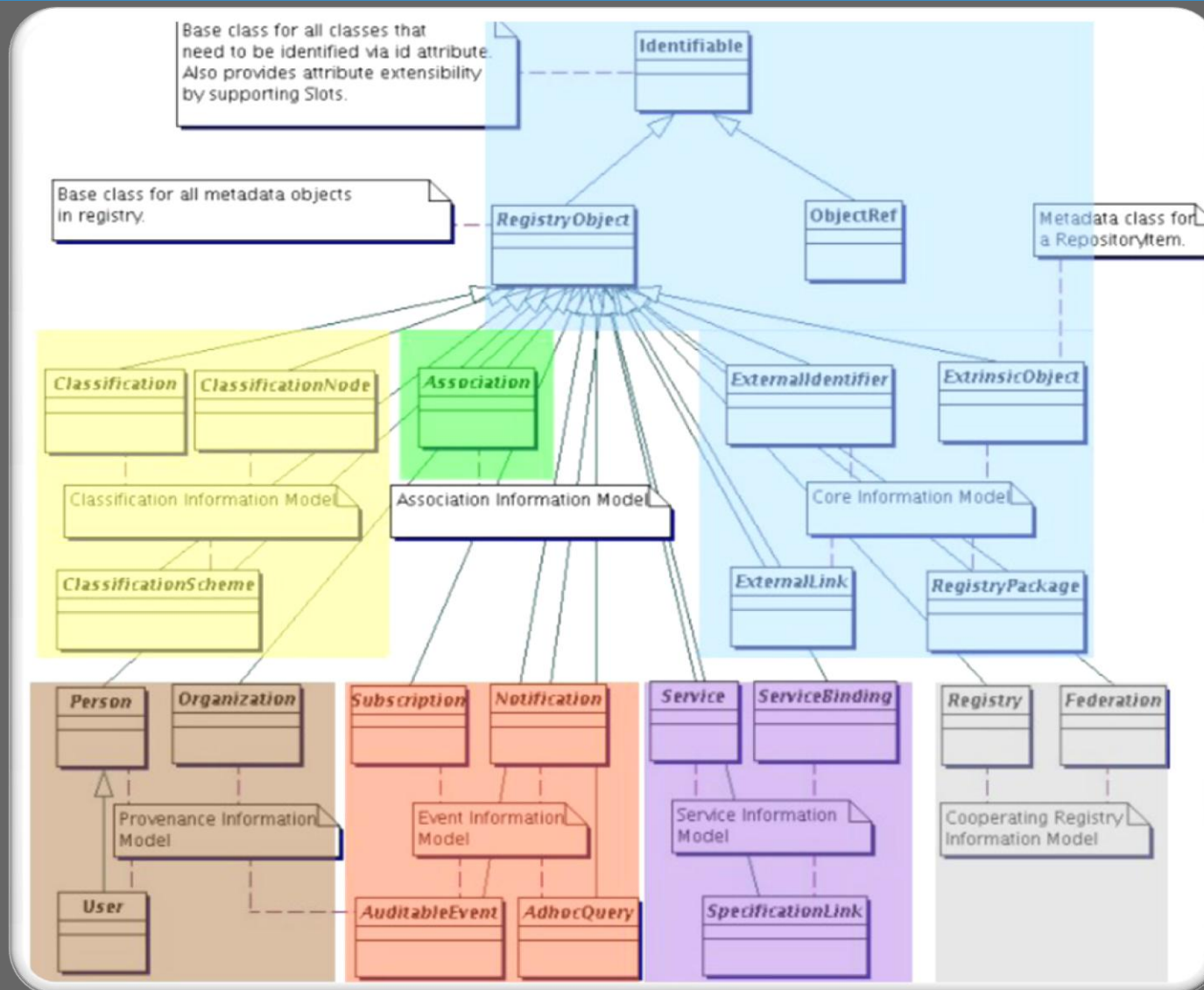


GeoServer: Advanced Raster

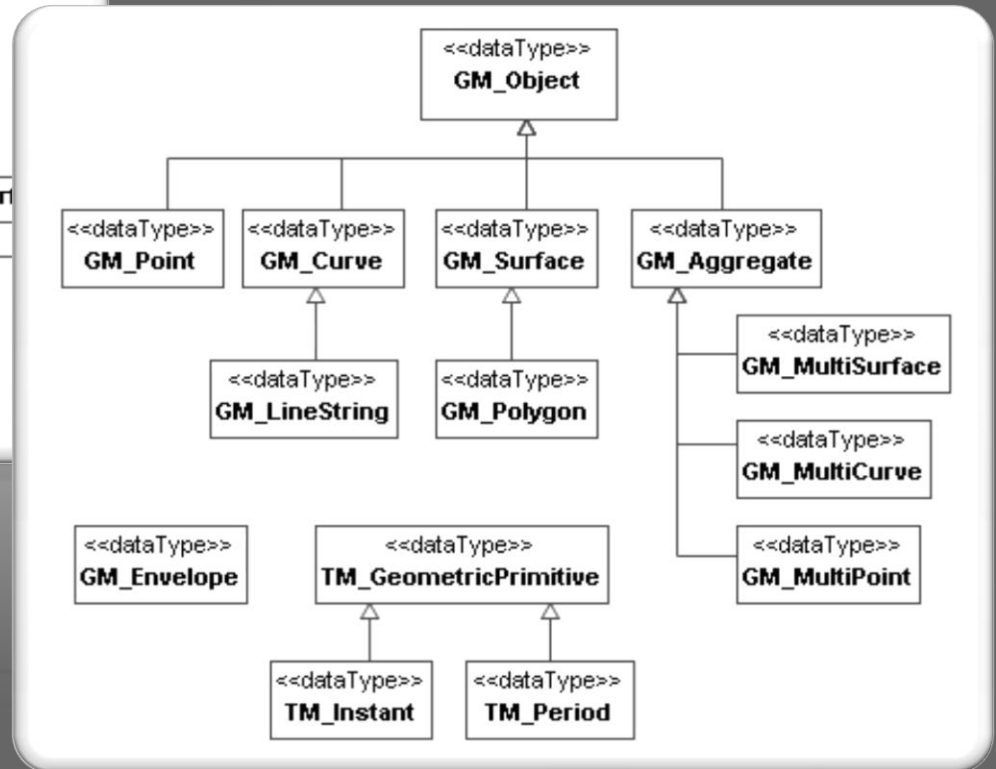
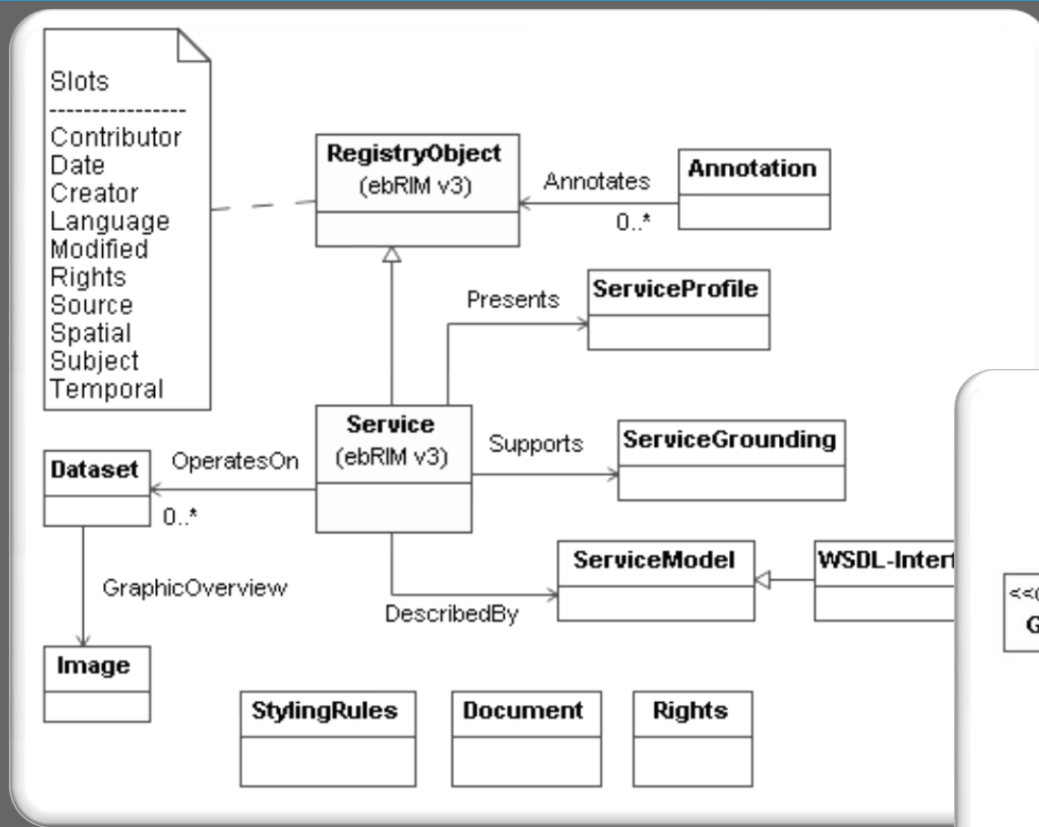
- Create a regular tessellation of original data
 - Work in raster space
 - Save georeferencing
 - Save original translation factors
- Tile integer indexing
 - Optimize most frequent case (TMS like approach)
 - Simplified tile lookup (no DBMS query)
 - Integer arithmetic + Recomposition by simple translation
 - Tile Caching highly simplified
- Multiband Tiles
- Consider MetaTiling
 - Reduce file access



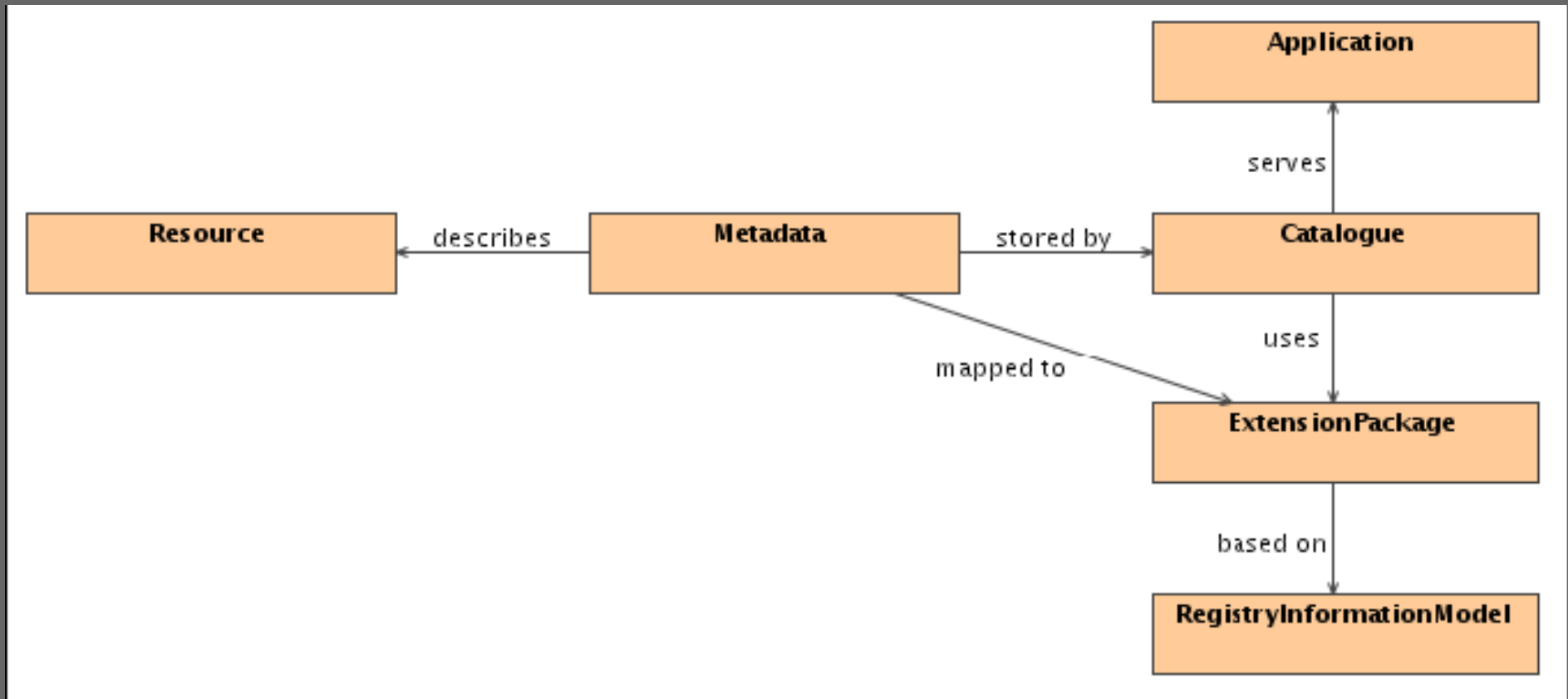
ebRIM: Registry Information Model 3.0



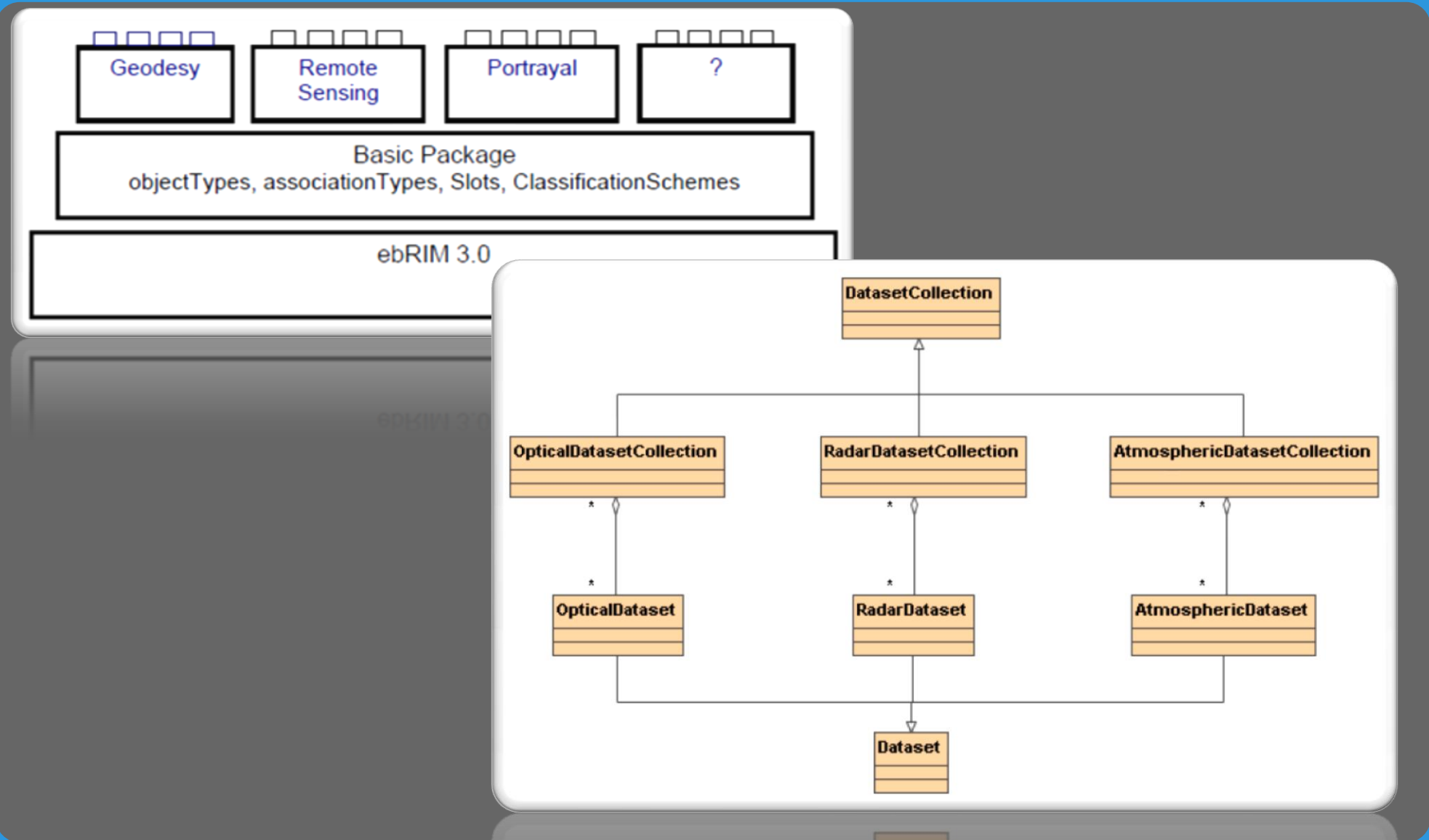
ebRIM: OGC Basic package



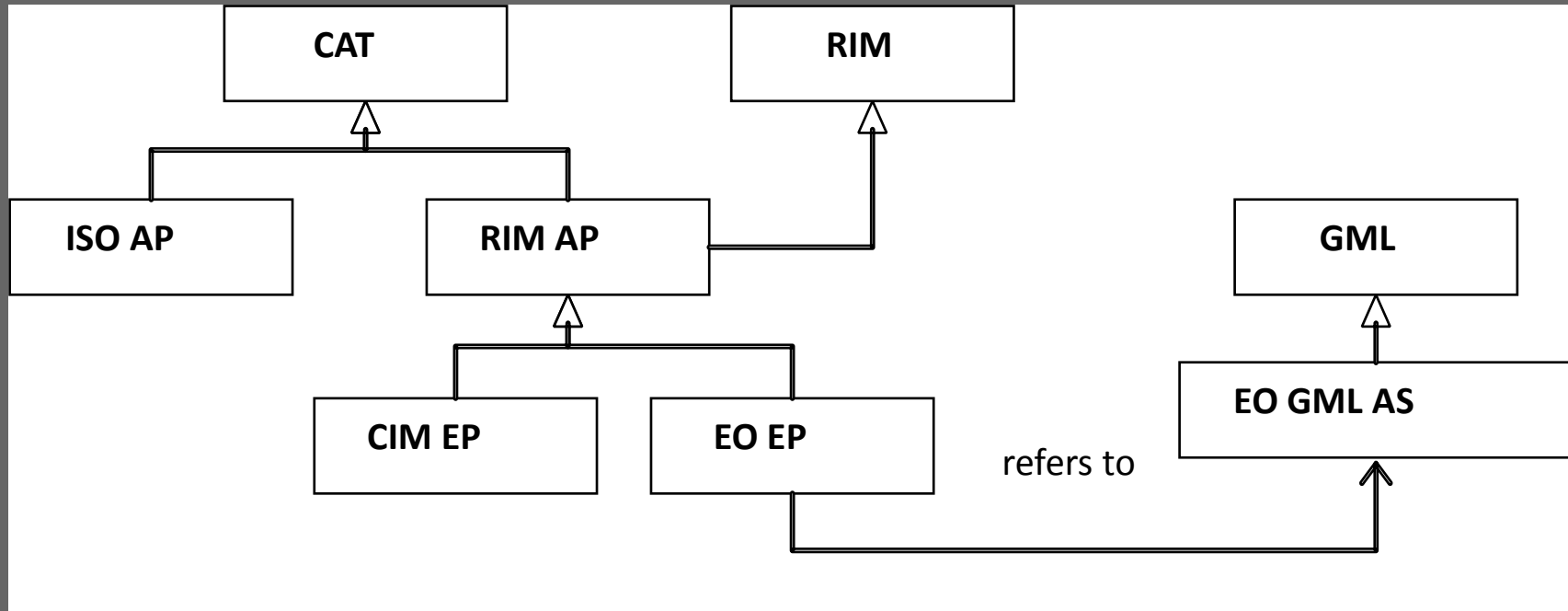
ebRIM: infrastructure



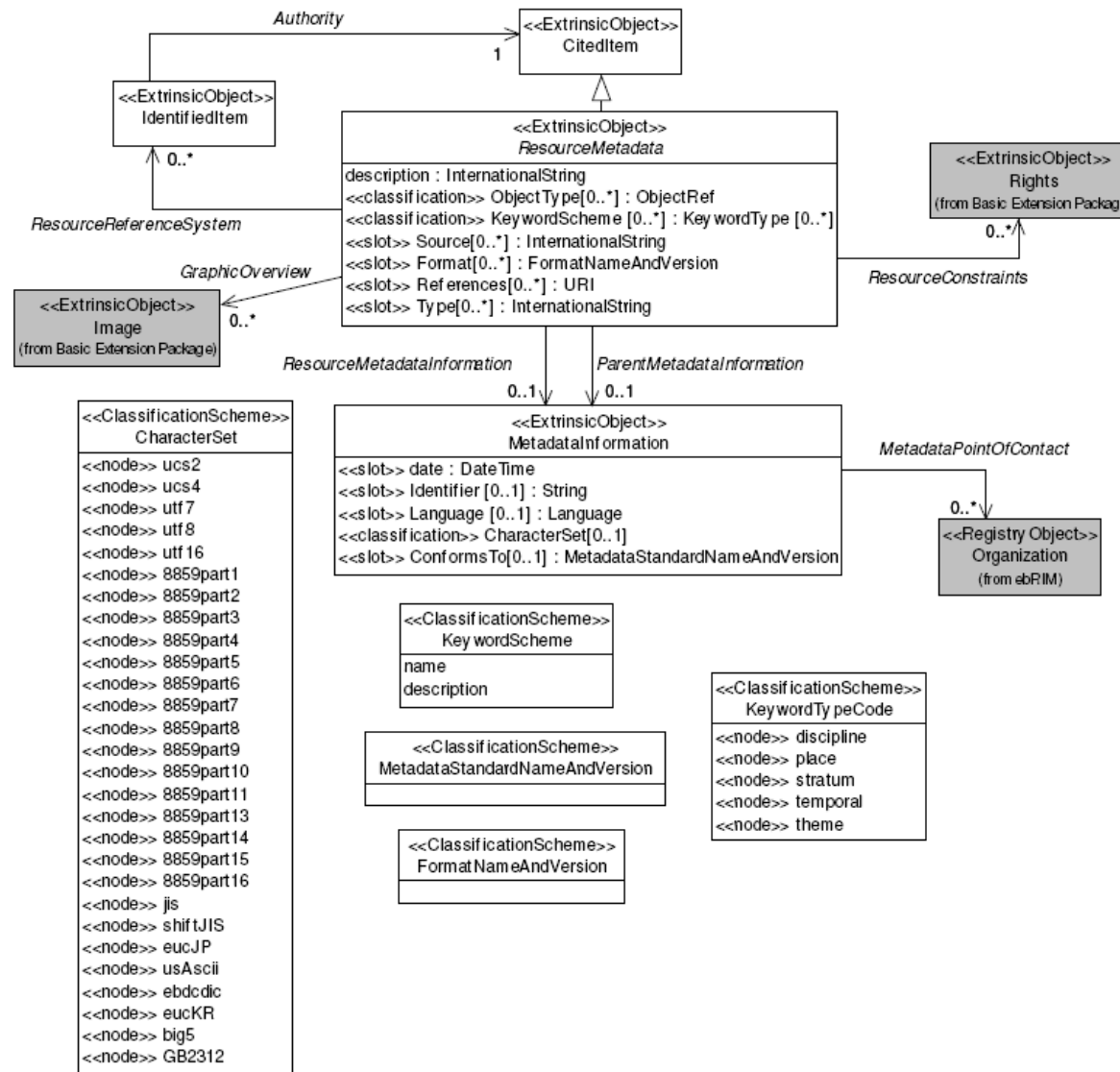
ISO Data types defined in the Basic package



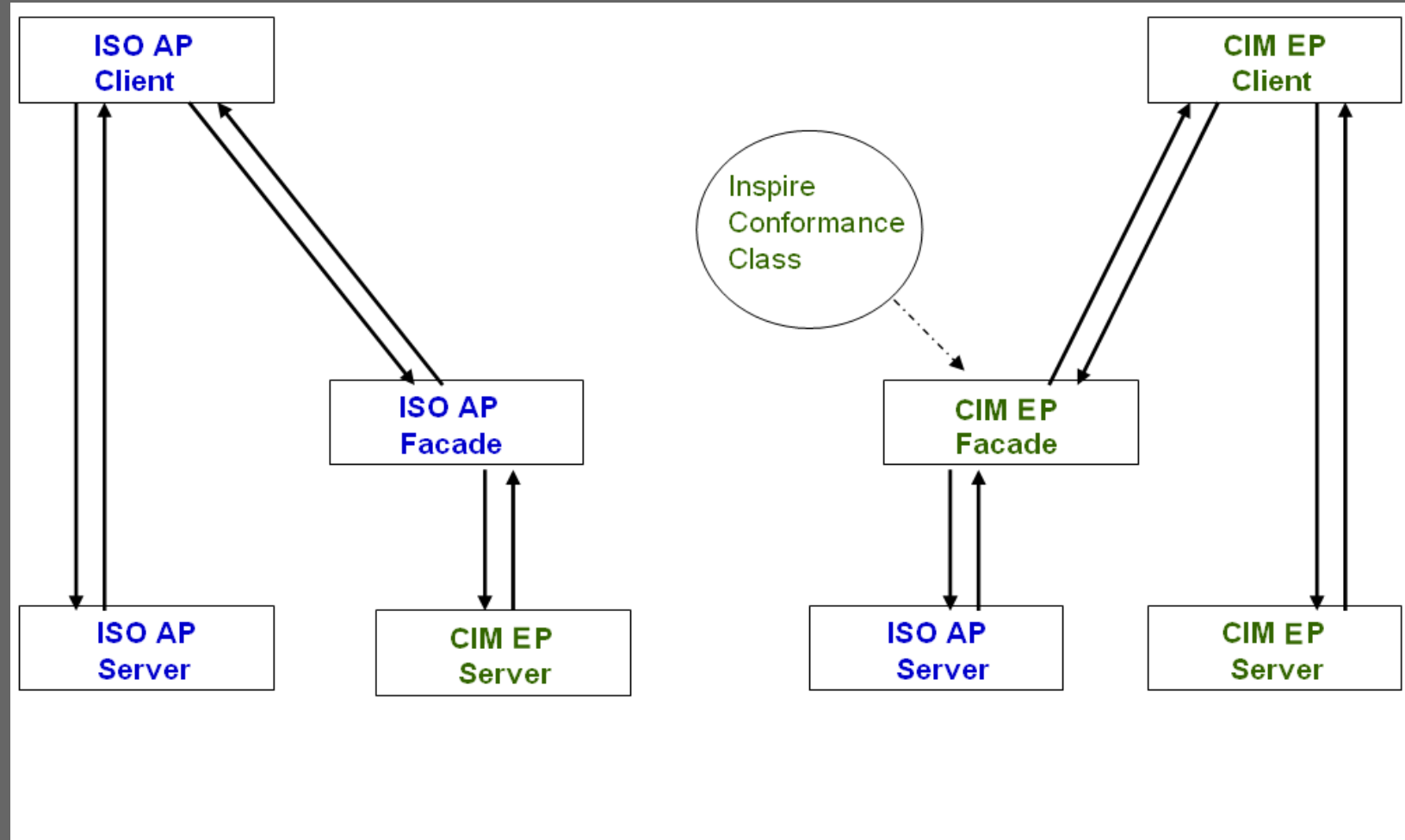
CIM EP and ISO AP in the OGC CAT Protocol Space



Some CIM classes and how they extend some ebRIM classes



Interoperability between CIM EP and ISO AP



EbRR @ Work: ESA portal



The screenshot displays the EbRR portal interface with several key components:

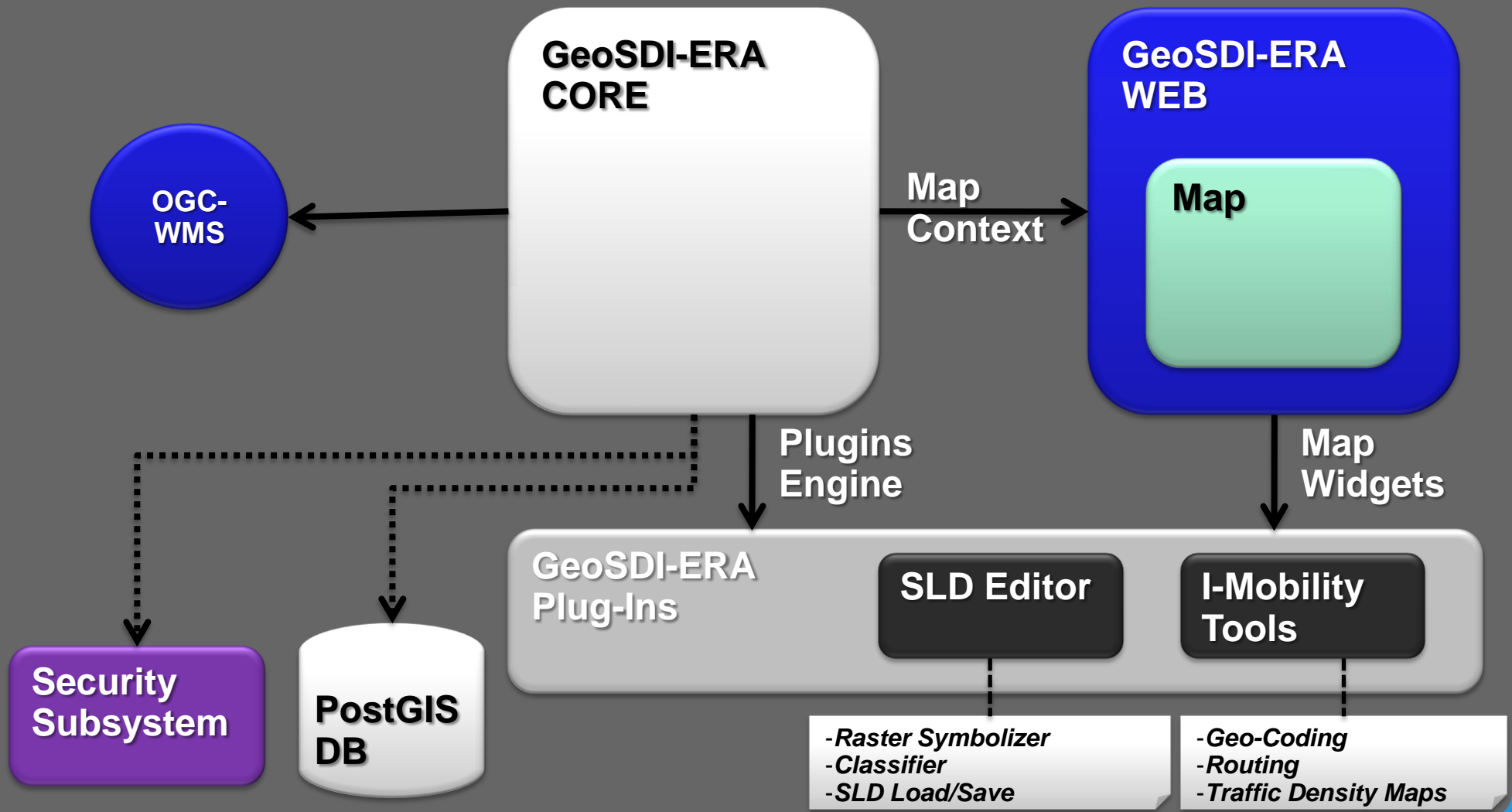
- Search Parameters:**
 - Collection: ESA.EECF.ENVIS
 - Spatial Coverage: Envelope with coordinates (NORTH: 53.7890625, WEST: -1.58203125, EAST: 27.5097856, SOUTH: 32.6953125)
 - Temporal Coverage: From 12/01/2008 to 12/02/2008
 - Presentation: full
- Map:** A satellite map of Europe and North Africa with a red bounding box indicating the search area.
- Response Table:**

ProductIdentifier	Platform	AcquisitionDateTime
EN1-08120108332012-31600.IF	Envisat	2008-12-09T10:28:46Z
EN1-08120110110833-31600.IF	Envisat	-
EN1-08120110135990-31600.IF	Envisat	-
- Result Details:**

(EN1-08120108332012-31600.IF) Result Details	
MdContactRPOrgName:	
MdContactRole:	
AcquisitionDate Time:	2008-12-09T10:28:46Z
IdentCode:	MER_FR
Platform:	Envisat
InstShNm:	MERIS
SatDomOrbit:	35317
SatDomOrbitDirection:	1
SatDomPerisyFrame:	3500
SatDomPerisyTrack:	193
SatDomSwathM:	
SatDomPassCoverageStart:	2333128
SatDomPassCoverageStop:	2529128
ProductIdentifier:	EN1-08120108332012-31600.IF
MIABs:	Envisat MERIS product
MIStatus:	7
TempExtTempBeginEndBegin:	2008-12-01T08:33:20.122
TempExtTempBeginEndEnd:	2008-12-01T08:36:36.122
DataExtGeoExtCoordinates:	41.28,22.51 38.84,35.14 35.68,33.87 32.19,32.58 28.69,31.40 27.45,30.99 29.72,19.98 30.98,20.25 34.52,21.01 38.06,21.79
DataExtGeoExtCenterCoordinates:	34.48,27.15

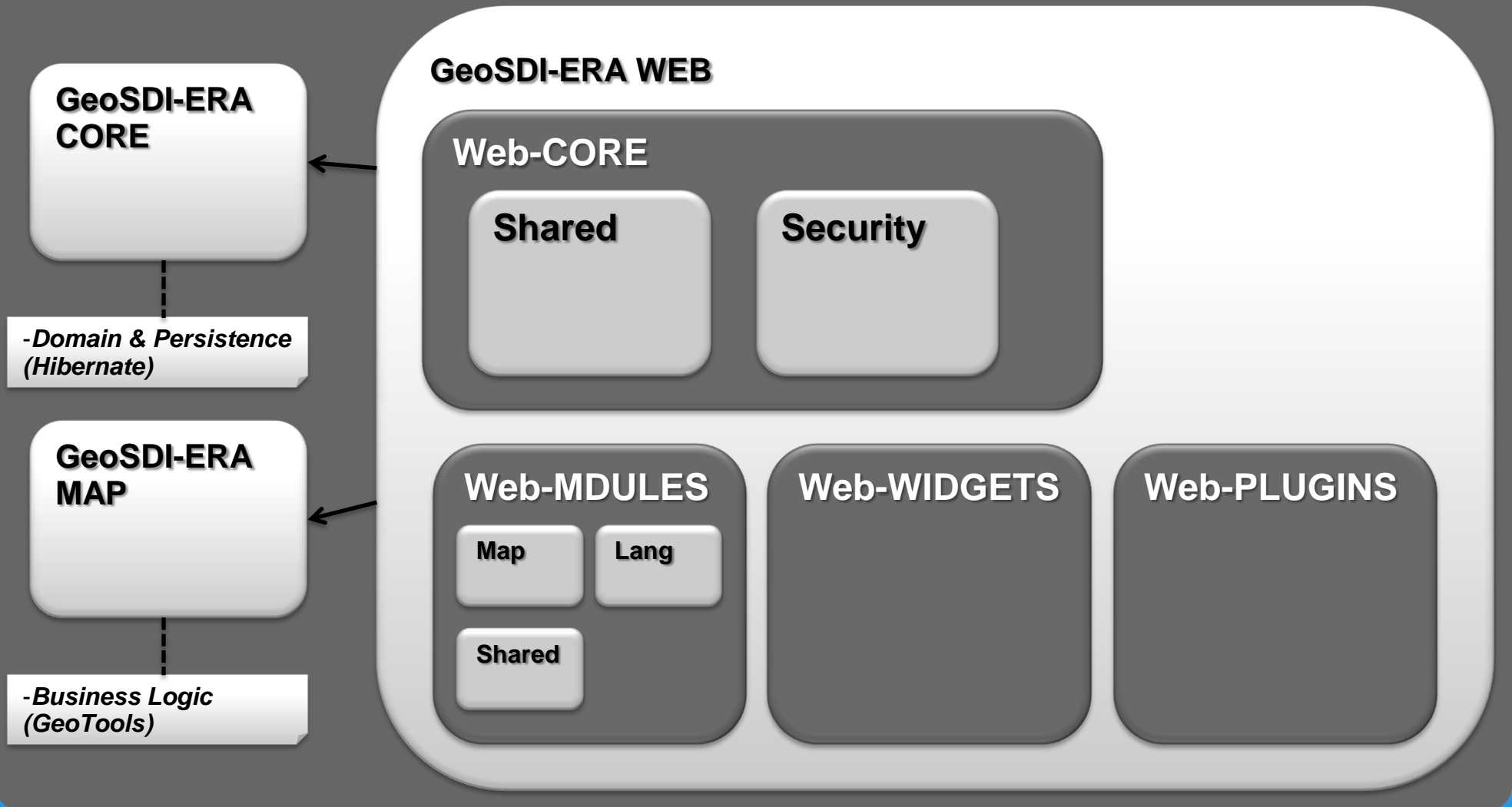
geoSDI – ERA Modular

- GeoSDI-ERA Architecture Schema



geoSDI – ERA Modular

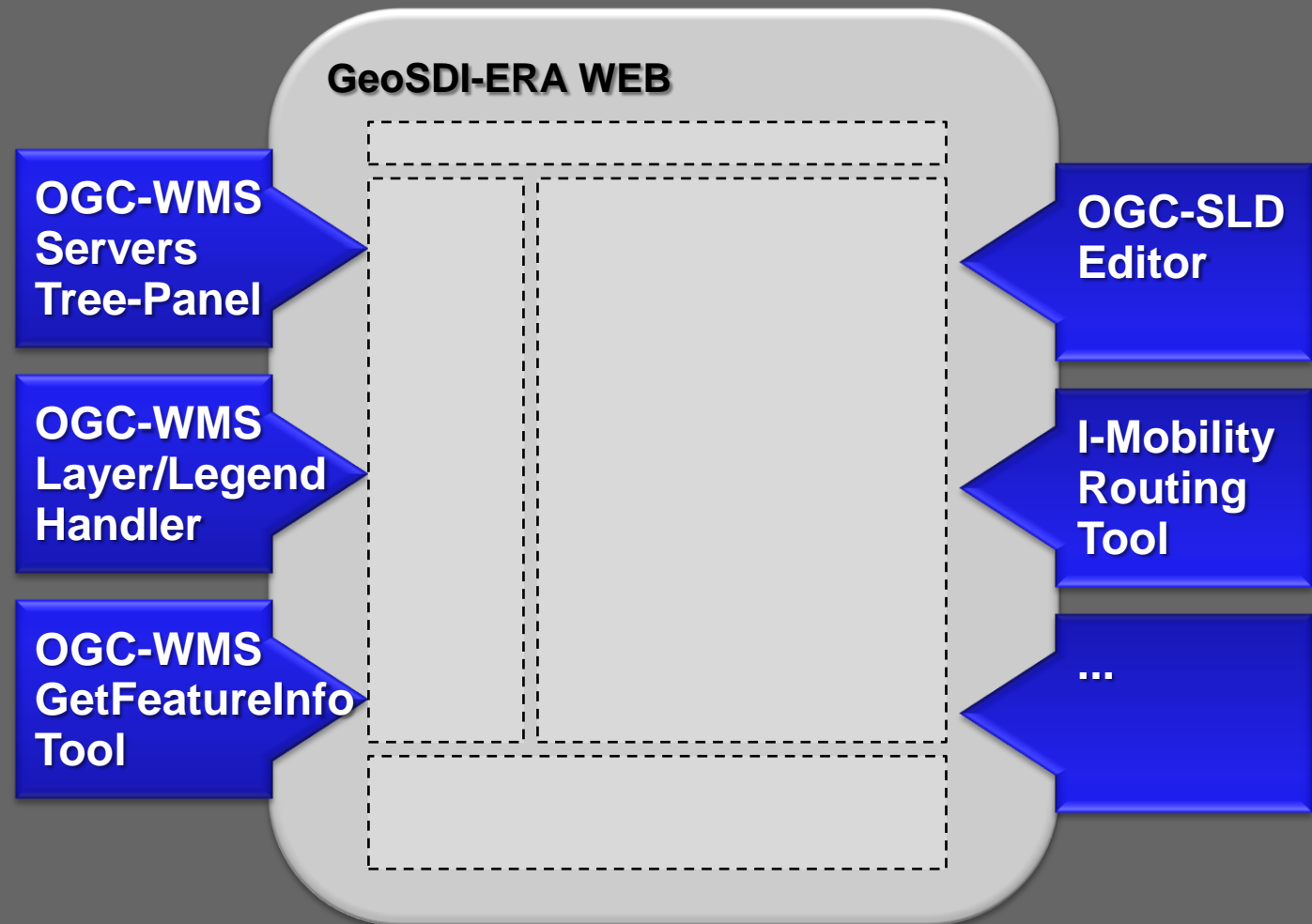
- GeoSDI-ERA Architecture Schema



geoSDI – ERA Modular

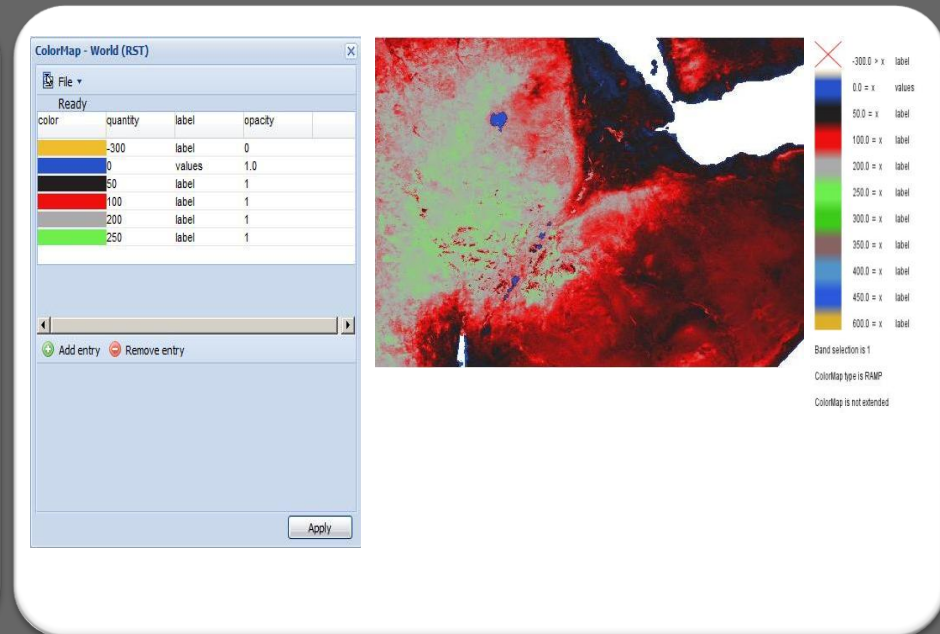
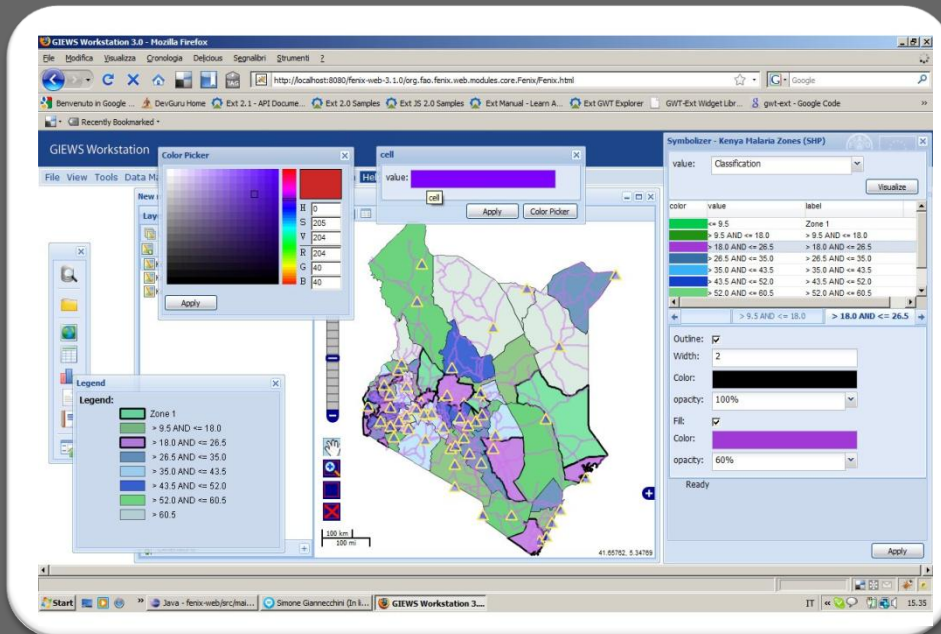
- GeoSDI-ERA Portal Concept

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



geoSDI – ERA 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

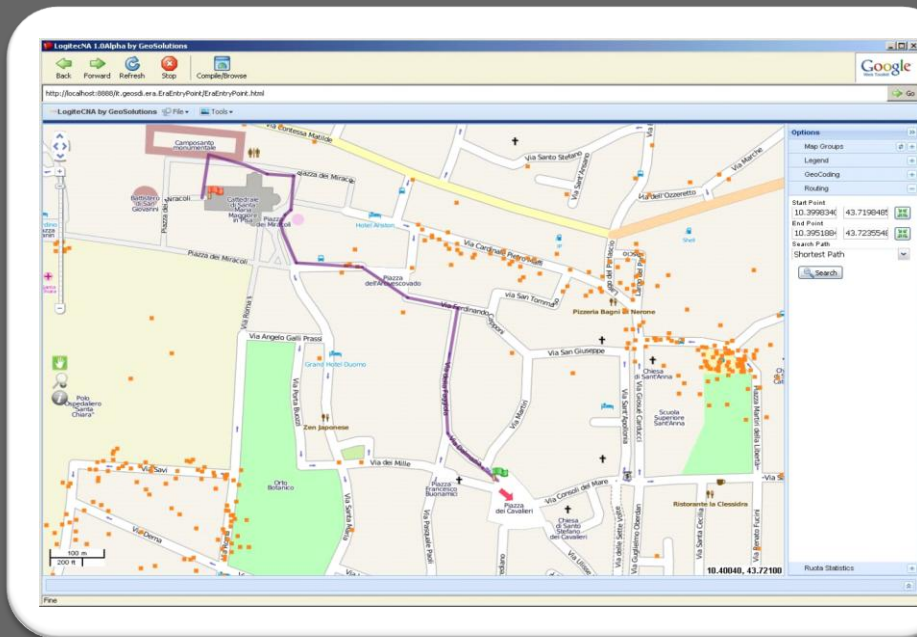


Polygon Symbolizer with Classification

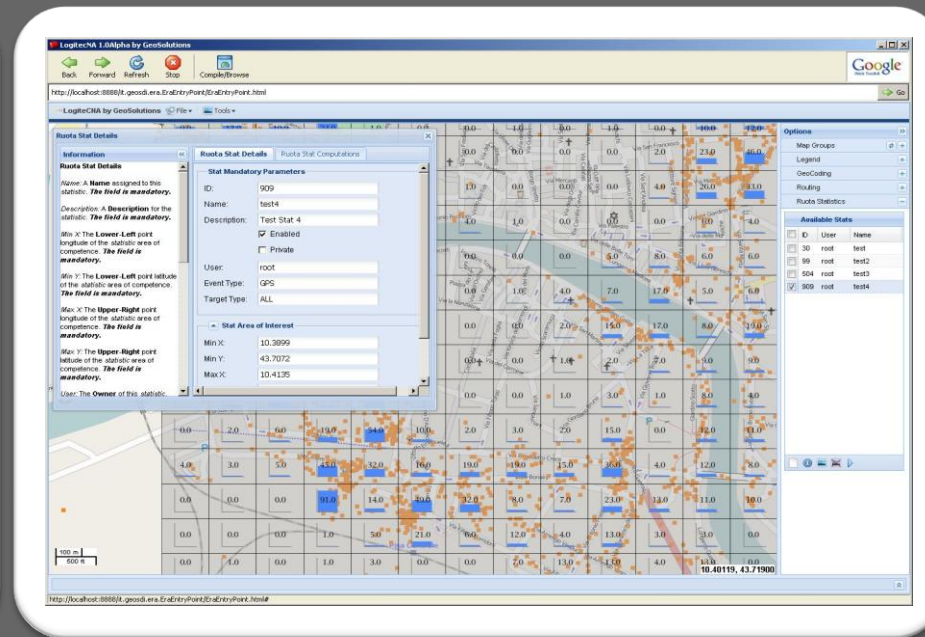
Raster Symbolizer Editor

geoSDI – ERA Modular

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



Shortest Path Routing example



Generated Density Maps with Graphs

Use Case: AIS data management

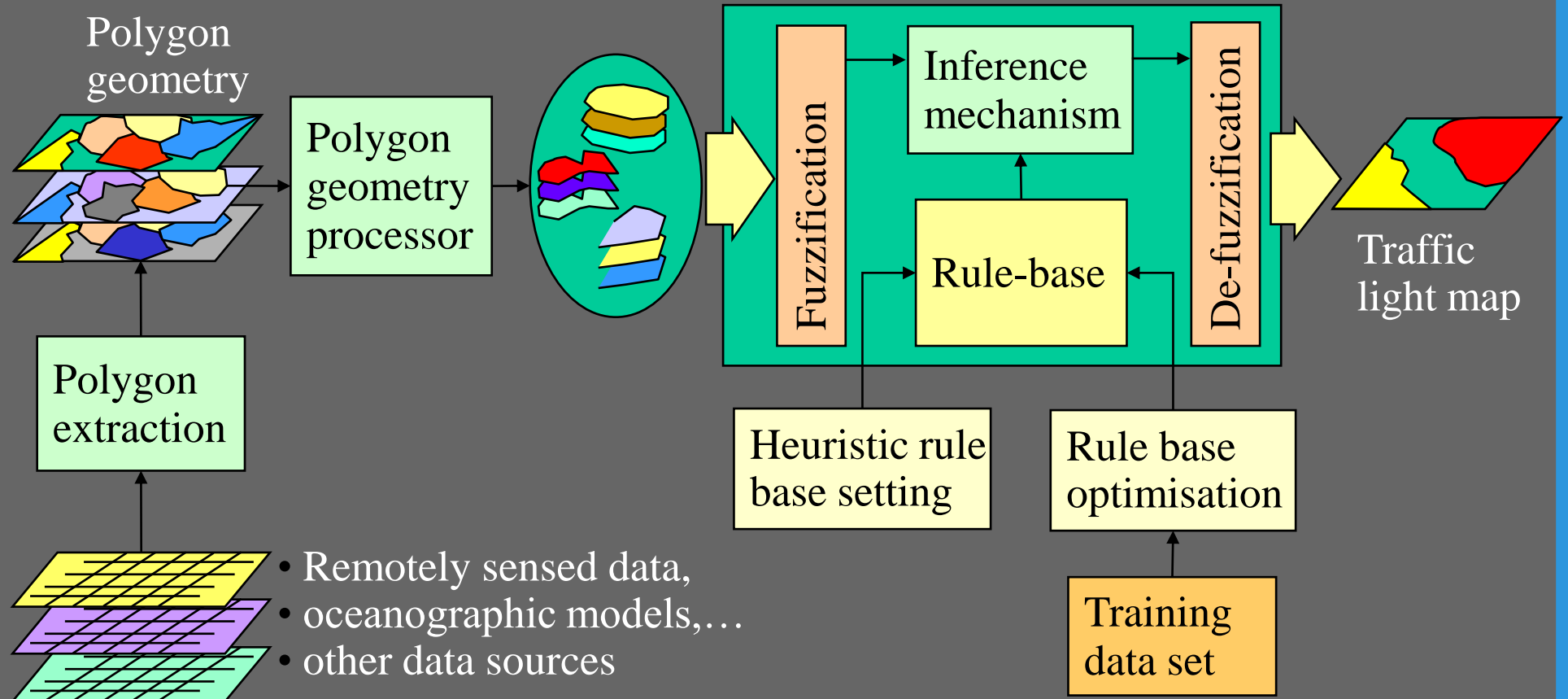
- Fully automatic AIS coverage and anomaly detection ingestion and publishing process.
- A control GUI allows to monitoring and stop/resume GeoBatch Data Flows

Ingestion Engine Control Panel

Flow Managers

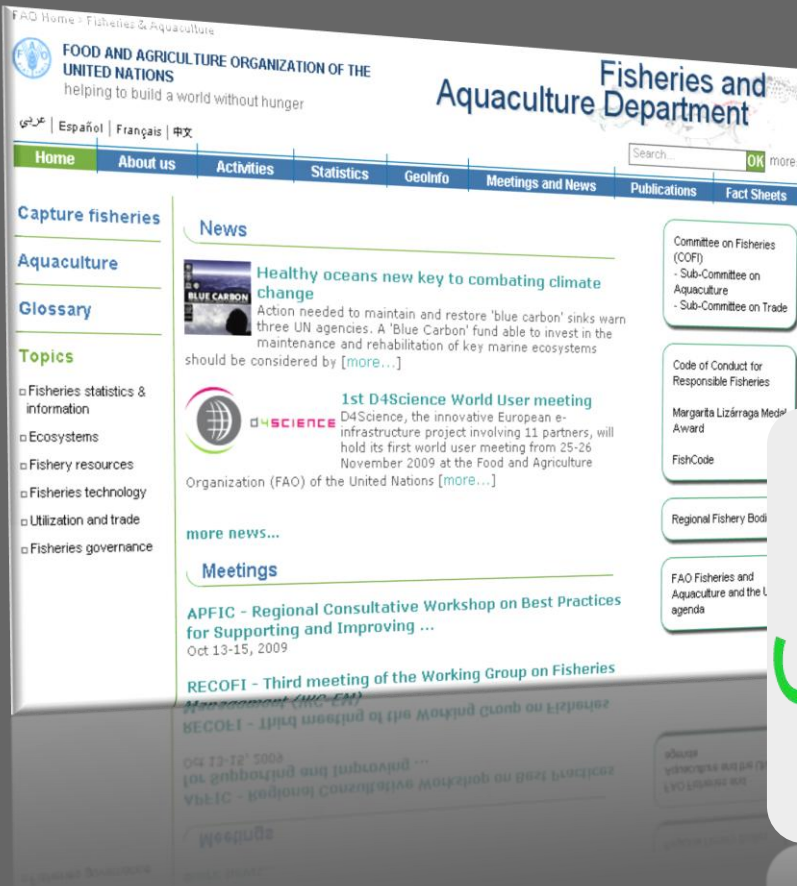
ID	DESCRIPTION	INPUT DIR	OUTPUT DIR	STATUS	ACTION
AISCOGAnomalies_fm_01	Flow Manager for the ingestion of AIS-Anomalies Files.	D:\WWW\AIS\AisONOFFAnomalies	D:\WWW\AIS\AisONOFFAnomalies\out		
AISSOGAnomalies_fm_01	Flow Manager for the ingestion of AIS-Anomalies Files.	D:\WWW\AIS\AisONOFFAnomalies	D:\WWW\AIS\AisONOFFAnomalies\out		
AISEmissionAnomalies_fm_01	Flow Manager for the ingestion of AIS-Anomalies Files.	D:\WWW\AIS\AisONOFFAnomalies	D:\WWW\AIS\AisONOFFAnomalies\out		
AISStopAnomalies_fm_01	Flow Manager for the ingestion of AIS-Anomalies Files.	D:\WWW\AIS\AisONOFFAnomalies	D:\WWW\AIS\AisONOFFAnomalies\out		
AISCoverage_fm_01	Flow Manager for the ingestion of ArcGRID Files.	D:\WWW\AIS\AISCoverage	D:\WWW\AIS\AISCoverage\out		
AISCoverage-NDays_fm_01	Flow Manager for the ingestion of ArcGRID Files.	D:\WWW\AIS\AisCoverageNDays	D:\WWW\AIS\AisCoverageNDays\out		
AISForecast_fm_01	Flow Manager for the ingestion of ArcGRID Files.	D:\WWW\AIS\AISForecast	D:\WWW\AIS\AISForecast\out		

Use Case: TDA



Use Case: FAO FIGIS

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

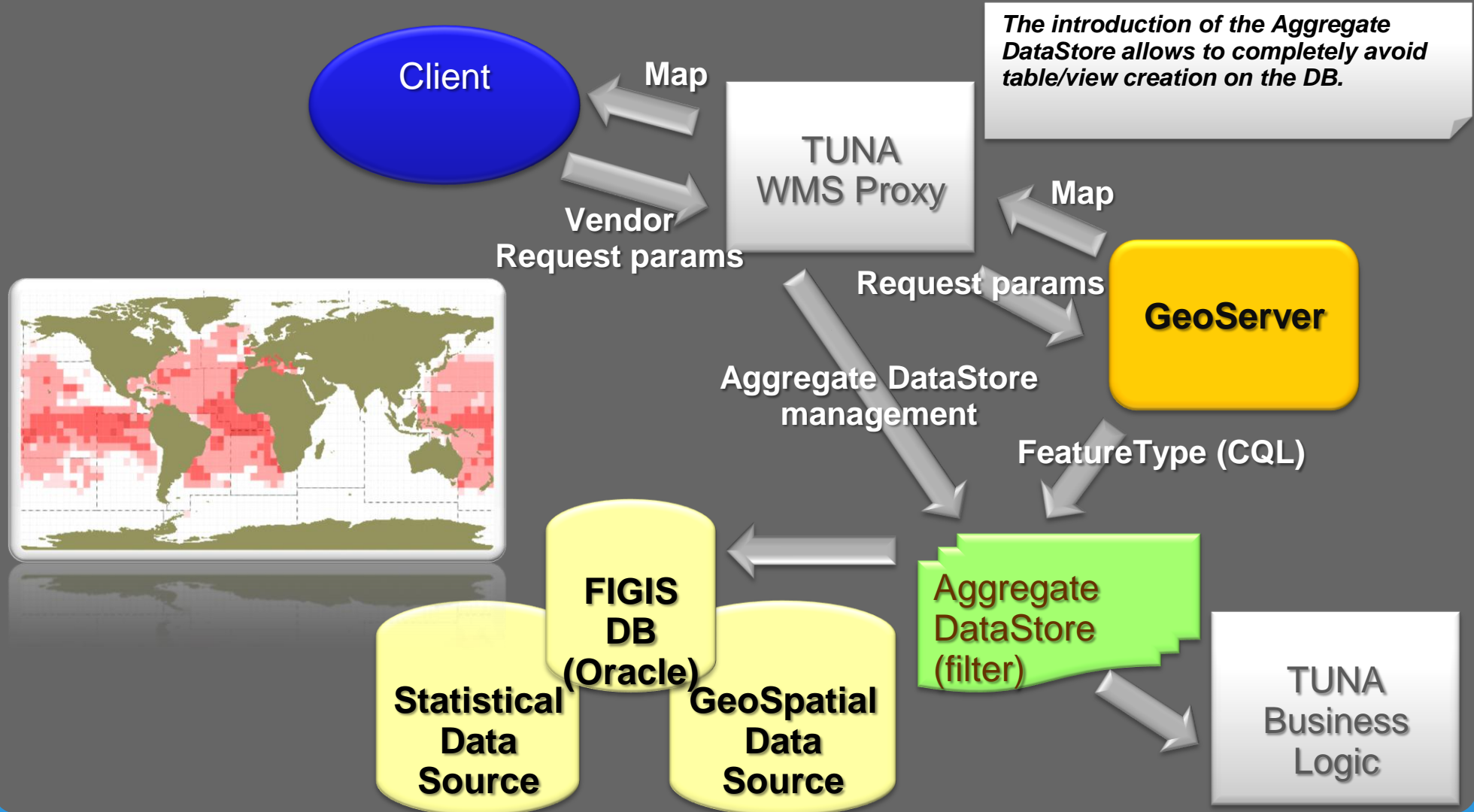


The mission of the Fisheries and Aquaculture Department of FAO is to facilitate and secure the long-term sustainable development and utilization of the world's fisheries and aquaculture.

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



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

