Open Source Software as an Efficient Distributed Infrastructure Stack for GIS Data Processing and Dissemination

Przemysław Kulesza and Michał Wójcik

Maritime Institute in Gdańsk

September 23–25, 2013

BalticBottomBase

INTEGRATED INFORMATION PLATFORM ON SOUTH BALTIC ENVIRONMENT

Project Goals

Maritime Institute in Gdańsk has been gathering data on South Baltic environment for many years. The data comes from different sources (e.g. bathymetric, sonar, geological, oceanographic, spatial planning and magnetic field data) in different formats (e.g. vector and raster data including profiles and probes). The main goal of the BBB (BalticBottomBase) project (`http://balticbottombase.eu/info`) is to make these valuable data available to other institutions using contemporary standards for geographic data sharing. In order to make the data available in convenient way, it must be described in the uniform manner. Moreover an infrastructure for storing data should be compatible with SeaDataNet project (distributed Marine Data Management Infrastructure) and INSPIRE (Infrastructure for Spatial Information in the European Community) directive.

Sources of the Maritime Environment Data

The Maritime Institute in Gdańsk collects and processes environmental data on the South Baltic Sea using highly advanced measuring techniques for over 60 years assimilates. Currently, data are gathered by use of our Multipurpose Oceanographic Research Laboratory r/v IMOR, which is equipped with up-to-date specialistic measuring instruments, for the following purposes:

- national and international research projects,
- statutory work,
- R&D activities,
- initiatives, agreements, international associations and existing databases such as:
  - "Brzeg" database,
  - oceanographic and atmosphere models of the Baltic Sea: HIROMB, WAM, COAMPS.

Multipurpose Oceanographic Research Laboratory r/v IMOR – vessel suited for investigations and measurements on the open sea, estuaries and in inland waters.

System Elements

Components of the platform used in the BBB project:

- web browser — all the system functionalities are exposed by the web application as dynamic HTML pages and can be accessed using a web browser,
- web application — single access point responsible for providing dynamic HTML pages and serving data to users, acts as a proxy (authentication and authorization mechanisms) for the MapServer,
- EJB modules — business logic modules which can be used in parallel,
- JMS server — central JMS (Java Message Service) user as an Enterprise Service Bus for system modules and worker service,
- worker — independent Java EE clients which can run on different machines, they can access business logic synchronously by remote EJB interfaces, but tasks submissions to workers are done through asynchronous messages passed through Enterprise Service Bus,
- MapServer — used for publishing spatial data, it serves all the spatial data (both stored in databases and filesystems) to the users through the web browser as a Web Map Service and interactive maps,
- database — storage for the spatial data and all the meta information considering the data,
- Hierarchical Storage Managements (HSM) system which consists of RAID and tape library — storage for the spatial data which can not be stored in database or their access from the filesystem is more efficient.

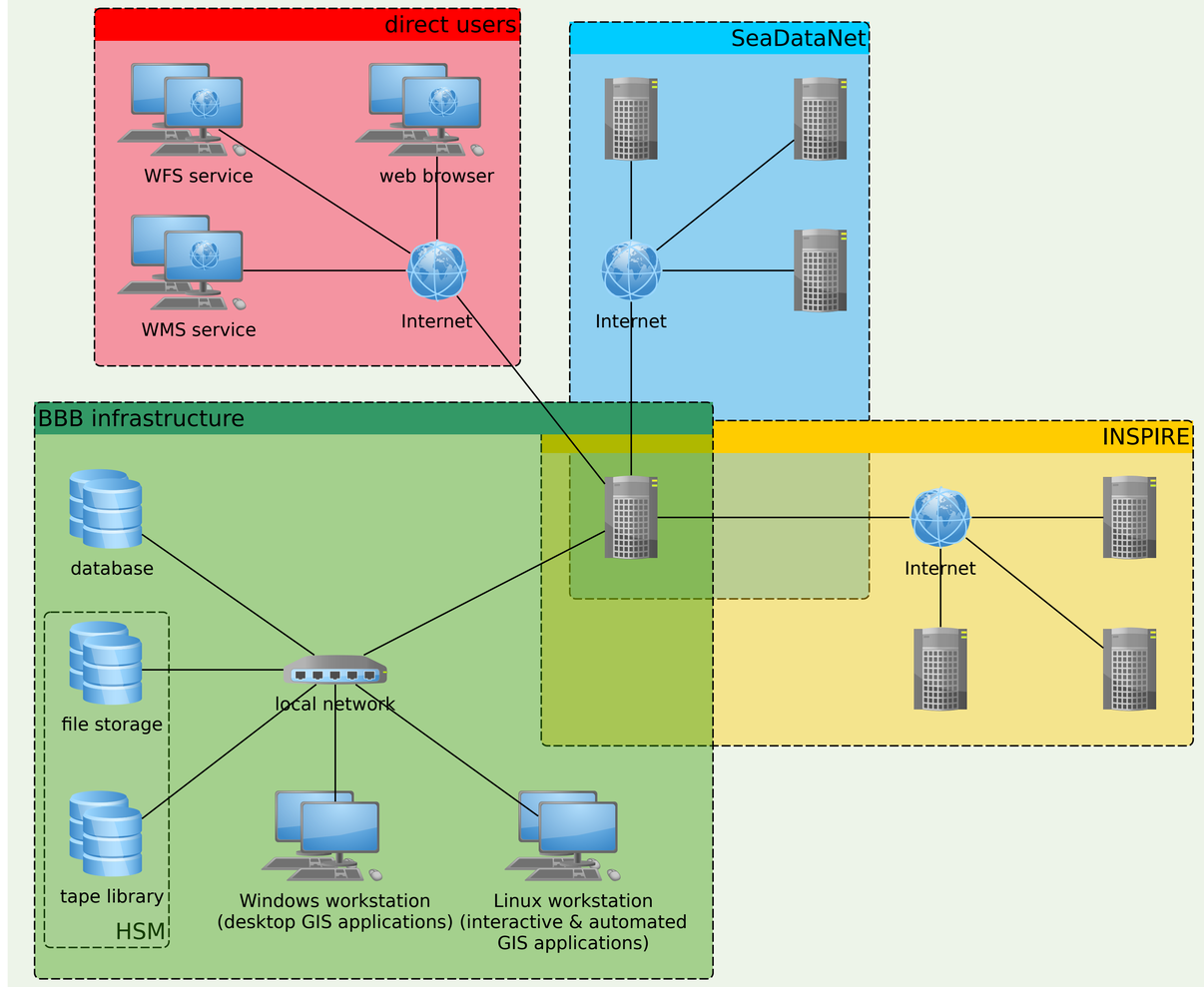
Web application, EJB modules and JMS server are all deployed on Java EE GlassFish application servers. Modular design of the platform will allow to expose remote interfaces compatible with SeaDataNet project and INSPIRE directive.

Publishing the Data

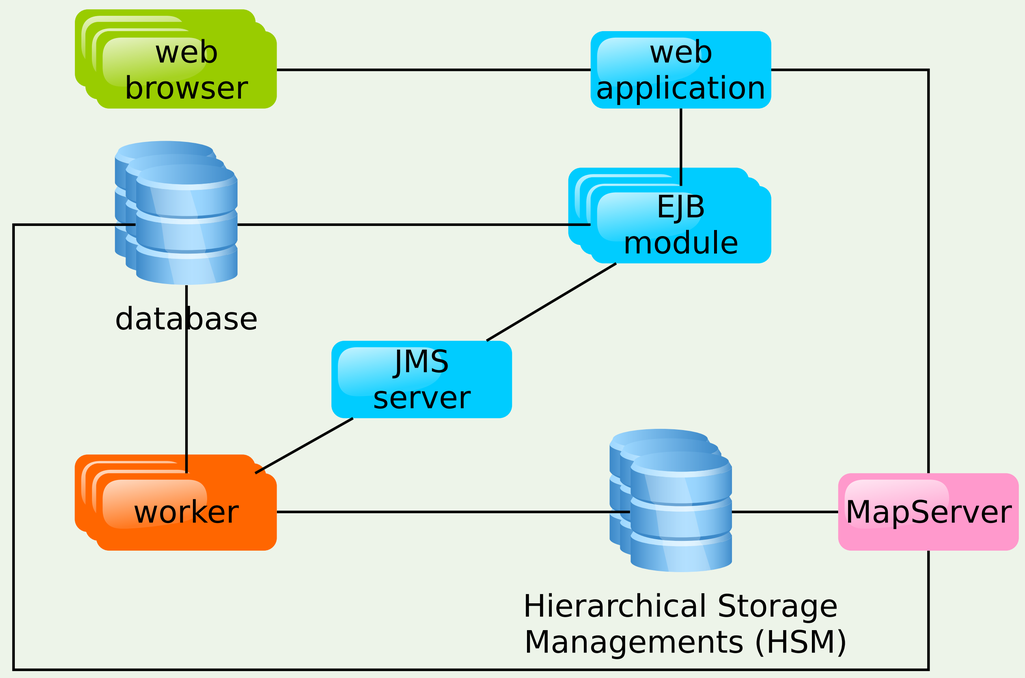
Process of publishing data in the BBB platform:

1. user provides the type of the data content and files using the web interface,
2. preliminary data processing for the purpose of providing detailed data description is done by the worker with the lowest load using GDAL (Geospatial Data Abstraction Library),
3. user optionally modifies the data details and accepts it,
4. the actual processing is done by the worker with the lowest load:
   1. if request consists of a number of raster files, those are merged into one with a *gdal_merge.py* tool,
   2. the new raster with lower resolution (5 % of input raster size) is created with a *gdal_translate* tool,
   3. from the low resolution raster the simplified outline is created with a *gdal_trace_outline* tool,
   4. the outline in WKT format is stored in the spatial database through the JDBC,
   5. source raster is processed by *gdal_translate* tool which is responsible for converting it to GeoTIFF format with defined tiles :
      1. if raster is small (threshold is set in worker's configuration) then it is stored without modifications,
      2. if raster is big, the optimized version is created by adding overviews with *gdalddo* tool;
   6. object registration in the MapServer is done by adding appropriate layers in two mapfiles responsible for exposing contours and raw data, to style those layers, corresponding SLD (Styled Layer Descriptor) files are created by GeoTools library,
   7. temporary files clean up,
   8. object registration in business layer through remote EJB interfaces.

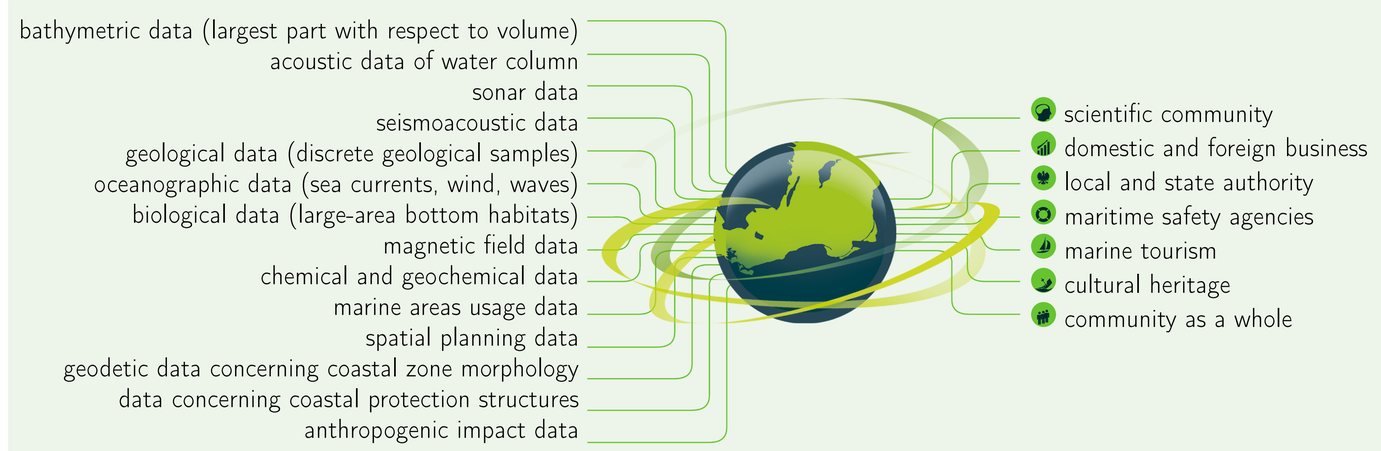
System Infrastructure



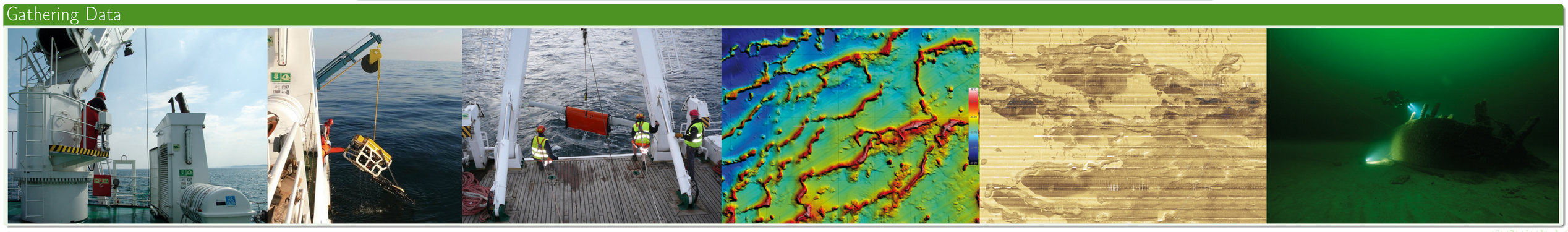
System Architecture Overview



BalticBottomBase Database Content



Gathering Data



INNOVATIVE ECONOMY
NATIONAL COHESION STRATEGY

MARITIME INSTITUTE IN GDAŃSK

BalticBottomBase
INTEGRATED INFORMATION PLATFORM ON SOUTH BALTIC ENVIRONMENT

EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND

Project co-financed by the European Regional Development Fund under the Operational Programme Innovative Economy.