

# ECMWF Web Re-Engineering Project

Baudouin Raoult, ECMWF

The European Centre for Medium-Range Forecasts currently provides a vast number of graphical products to users via its website. The users are provided with some limited interactivity, such as basic animation or the ability to create an epsgram at a location of their choice. Due to the success of this facility, ECMWF has started a project to re-engineer its web based services with the aim to provide highly interactive and parameterised graphical products, to forecasters of the ECMWF Member States and commercial users. Users will be allowed to interact directly with the plots and perform actions such as zooming and panning. They will be able to customise some of the products, for example by changing the event threshold of probability maps, and have some control on the graphical attributes (isolines, shading) used to render the plots. Products will also be provided via OGC web map services (WMS) so they can be integrated in the users' own GIS software. ECMWF is investigating technologies that will allow it to implement a solution that is highly available, scalable and fast.

## The web re-engineering project

- **Motivation:**

- Many of our users rely on our graphical web products for their daily work in their forecast offices, and have requested that our web services be continuously available
- At the Annual Users' meetings, we have received requests to create tailored products (e.g. control the event threshold on probability maps)

- **Goals:**

- Redesign the web infrastructure so that the web service is highly available and supported at the same level as the field dissemination
- Provide more interactivity (e.g. zoom, pan, overlay parameters)
- Allow product customisation (e.g. control the event threshold on probability maps)
- Use open (OGC) standards so that ECMWF products can be embedded in users' own software

MOS Workshop November 2009

Slide 2



## The web re-engineering project (cont.)

- **2-year project to implement new ECMWF web service that is:**

- Highly available and operationally supported (same support as current dissemination)
- Aimed at forecasters
- Highly interactive
- Suitable for deployment as computer-to-computer standard 'web services'
- Flexible to meet future requirements

- **Milestones**

- First prototype - November 2009
- Alpha release - February 2010
- Beta release - January 2011
- Operational release - June 2011

MOS Workshop November 2009

Slide 3



## Requirements

- **Highly available – Operationally supported**
  - H/A Hardware
  - H/A Software
  - Operator monitoring
- **Performance**
  - Target: deliver a plot under 1 second
- **Interactivity**
  - Pan, zoom, overlay (à la GoogleMap)
  - Customisation, plots on demands (e.g. changing event probability threshold)
- **Scalability**
  - Support any future user load
  - Extensible: easy addition of new products

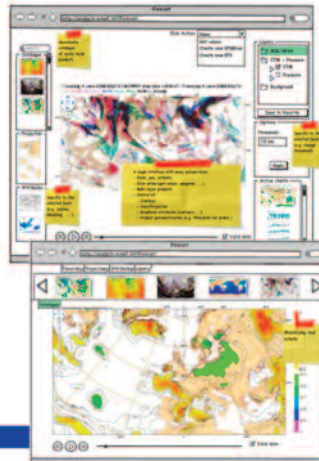
MOS Workshop November 2009

Slide 4



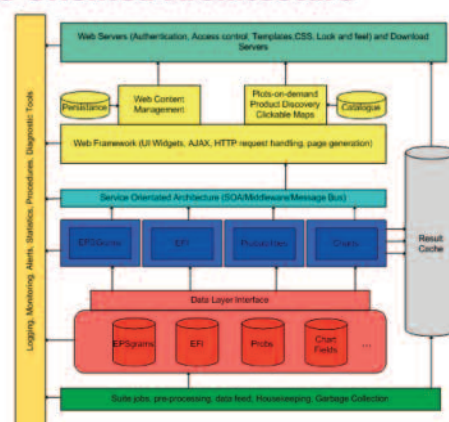
## Gathering of user requirements

- **The project has been presented on several occasions:**
  - ECMWF Forecast Products Users' Meeting, Computer Representatives Meeting
  - Very positive feedback from forecasters
  - Most forecaster requests focused on the desire to be able to create customised products
  - Requests for new products
- **Consultation process will continue throughout the project**



MOS Workshop November 2009

## Service Oriented Architecture



MOS Workshop November 2009

Slide 6



## Hardware



- 2 Foundry Load Balancer ServerIronGT
- 3 servers hosting web servers
- 3 servers hosting web application
- 3 servers hosting several virtual machines
- 6 servers hosting storage, compute and plot services
- HP DL360 G5 Dual 2.5Hz Quad Core Xeon
- OpenSuSE Linux 11.1



- Systems located in different parts of the building, attached to different routers and different power sources

MOS Workshop November 2009

Slide 7



## Software

- We investigated technologies used by the “big players” (e.g Google, Yahoo, Amazon, Facebook, Wikipedia ...):
  - Memcached (Very fast distributed memory)
  - Tokyo Tyrant (Scalable, distributed persistent space)
  - Hadoop (High availability and redundant distributed data)
  - Xen (Virtualisation)
  - DRDB (Network RAID)
  - Ganeti (H/A Cluster management)
  - Nagios (Alerts system)
  - Scribe (Distributed logging)

MOS Workshop November 2009

Slide 8



## Software (cont.)

- ...
  - Ganglia (Distributed monitoring)
  - Django (Python based Web framework, server side)
  - jQuery (JavaScript based web framework, client side)
  - OpenLayers (JavaScript based OGC WMS-client)
  - Apache 2.2 (Web server)
  - MySQL (Database)
- And of course:
  - Magics++
  - grib\_api
  - Mars
  - Metview ...

MOS Workshop November 2009

Slide 9



## About Hadoop

- A framework that supports data intensive distributed applications
- Inspired by Google's MapReduce and Google File System (GFS) white papers.
- Yahoo, Amazon, IBM, Facebook, AOL, Fox, Last.fm, Microsoft, ...
- Hadoop – HDFS - Distributed storage, with a filesystem like API (HDFS)
  - Data nodes hold blocks of data. Each node uses local storage
  - Name node holds the file names and the blocks location (single point of failure)
  - Each file is spread of several data nodes
  - Each block has several copies distributed over the cluster
  - Designed for large blocks (64 MB)
- MapReduce facility to be investigated

MOS Workshop November 2009

Slide 10



## About Ganeti (H/A Pairs)



- Ganeti is a cluster virtual server management software tool built on top of existing virtualization technologies (Google)
- Xen virtual machines (Hardware assisted virtualization: 3% overhead)
- DRBD (Distributed Replicated Block Device)
  - "Network RAID1" (20% overhead write, 0% read)
- Live migration
  - Two passes memory migration: 10s for 12GB memory (Stoppage of around 60-300 ms is required to perform final synchronization)
  - No interruption of service: IP connections not broken (MAC address move)
  - Fail over : restart VM on backup machine.
  - Command line tools: can be done by operators

MOS Workshop November 2009

Slide 11



## Service Oriented Architecture

- Multi-tier architecture, deployed on a series of Linux clusters:
  - Web frontend (Web server)
  - Web backend (Dynamic page generation)
  - Services (Plotting, probability computations, EPSgrams, ...)
  - Data layer (Raw fields)
- Cluster approach provides built-in scalability, redundancy and load balancing
- Critical components run on virtual machines that can be redeployed dynamically

MOS Workshop November 2009

Slide 12



## Deployment

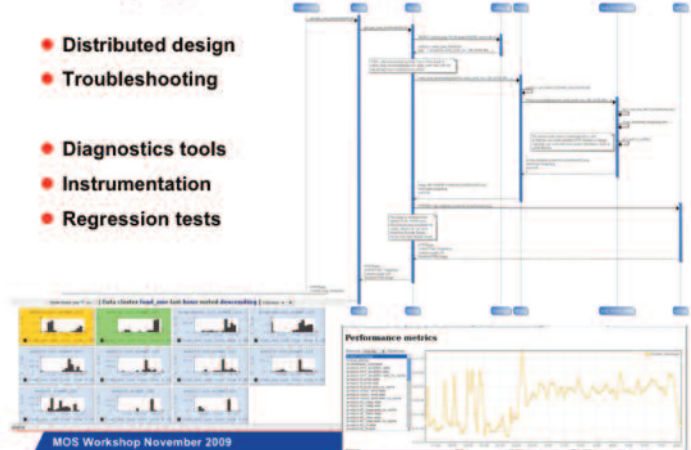
- Virtual machines for critical components and single points of failure
  - Hadoop name node
  - SOA Broker
  - Spot database
  - Catalogue (MySQL)
  - All virtual machines sized in such a way that they can fit in a smaller number of nodes if necessary
- Physical machines for components with built-in redundancy
  - Hadoop data nodes
  - Memcached servers
  - Services (plot, retrieve, probabilities, epsgrams, ...)



MOS Workshop November 2009

## Developing in an SOA environment (is hard)

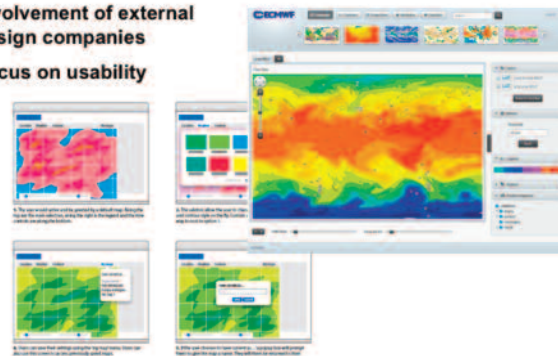
- Distributed design
- Troubleshooting
- Diagnostics tools
- Instrumentation
- Regression tests



MOS Workshop November 2009

## Web user interface

- Involvement of external design companies
- Focus on usability



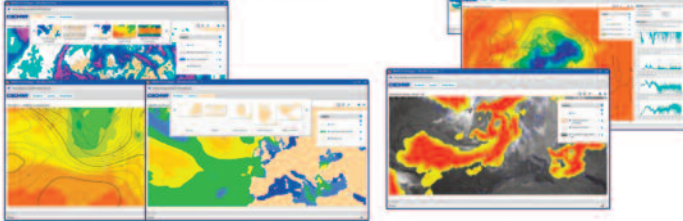
MOS Workshop November 2009

Slide 15



## Prototype: Forecasting tool

- Interactivity: zooming, panning, ...
- Customisation:
  - Probabilities threshold, ...
  - Show/hide, add/remove layers
- Related products: Epsgrams

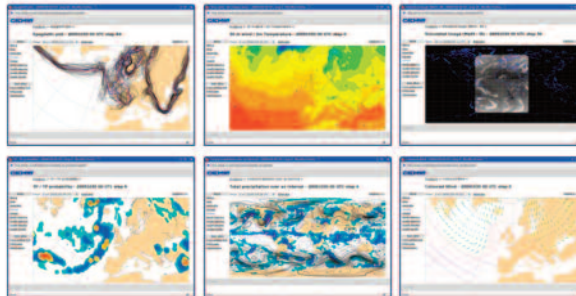


MOS Workshop November 2009

Slide 16



## Prototype: Catalogue browsing



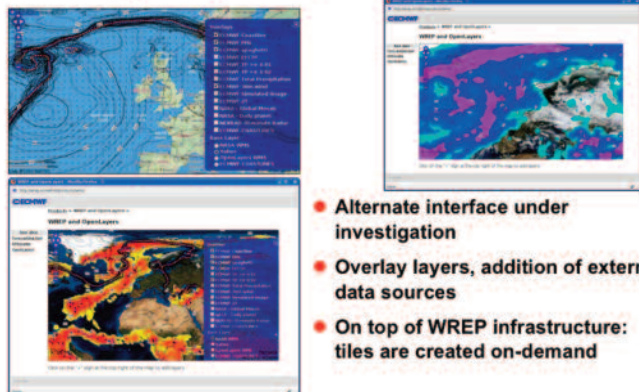
- Browsable catalogue – Link to Forecaster tool
- Limited interactivity – Preset number of projections, animation
- Similar to current web catalogue, but use the WREP infrastructure

MOS Workshop November 2009

Slide 30



## Prototype: OpenLayers integration



- Alternate interface under investigation
- Overlay layers, addition of external data sources
- On top of WREP infrastructure: tiles are created on-demand

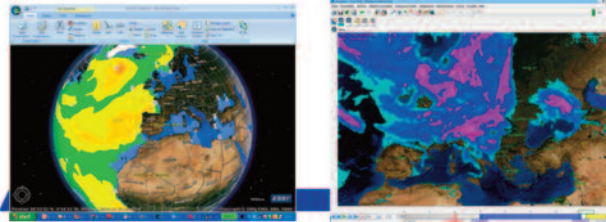
MOS Workshop November 2009

Slide 31



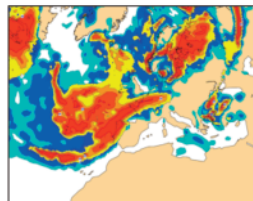
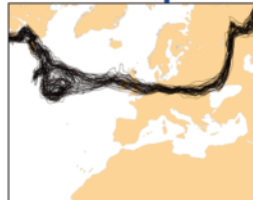
## Prototype: OGC Web Map Services

- Aim: to make it possible to embed ECMWF products directly in the forecasters' workstations
- On top of WREP infrastructure:
  - "GetCapabilities document" build dynamically from product catalogue content
  - Layers are created on-demand
- Challenges: access control, time dimension, customisation



## Conclusion: fully functional proof of concept

- All products created "on-demand" (2D maps, EPSgrams)
- Zoom, pan, overlay
- Customisation: setting of probability thresholds, contouring
- Browsable catalogue
- Initial user interface
- OGC Web Map Service (WMS)



MOS Workshop November 2009

Slide 33



## Future work

- Persistence
- Security and access control
- Monitoring, alerts and service statistics
- Management tools
- Performance tuning
- Develop further WMS aspect
- More products
- User testing

MOS Workshop November 2009

Slide 34

