HPC for a weather information company

Some lessons learned and things I wish I had known before

Kim Serradell, Meteomatics

18.05.23, HPCKP23, Barcelona

meteo matics

Agenda

1. Introduction
2. The EURO1k
3. Deployment and lessons learned
4. Next steps
5. Conclusion

Who am I?

- Computer scientist from Facultat d'Informàtica de Barcelona
- 14 years at Barcelona Supercomputing Center. First as support engineer then leading the Computational Earth Sciences Group
- On 2022, left BSC and joined Meteomatics
- Big change, moving from research to industry, from non-remote to remote work
- Regarding HPC, I was no longer only a user…

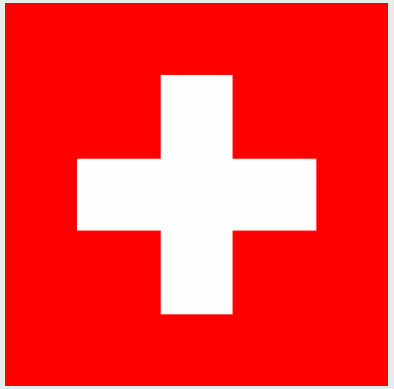
World-Class Meteorology & Data Science

We are the global leader in weather intelligence.

We provide the most accurate weather data for any location at any time, to improve our customers' business.

80

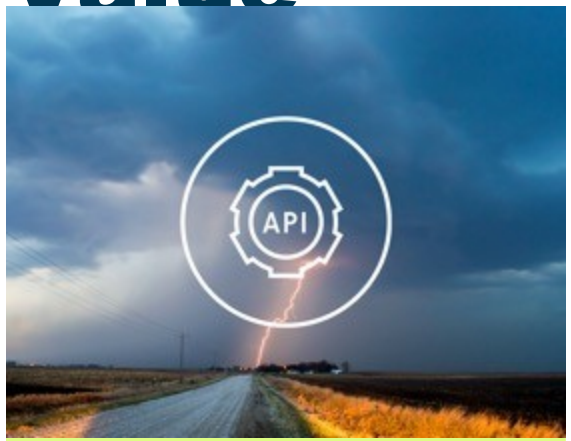
Highly-trained staff across 3 offices



Swiss precision engineering

Over 450 Global clients & partners

Unique Forecasting Solutions that Create Value



Weather API

With our weather API you get continuous access to worldwide high-resolution weather, ocean, environment and climate data as well as historical data, real-time data, forecasts and climate scenarios up to 2100.

Weather Models

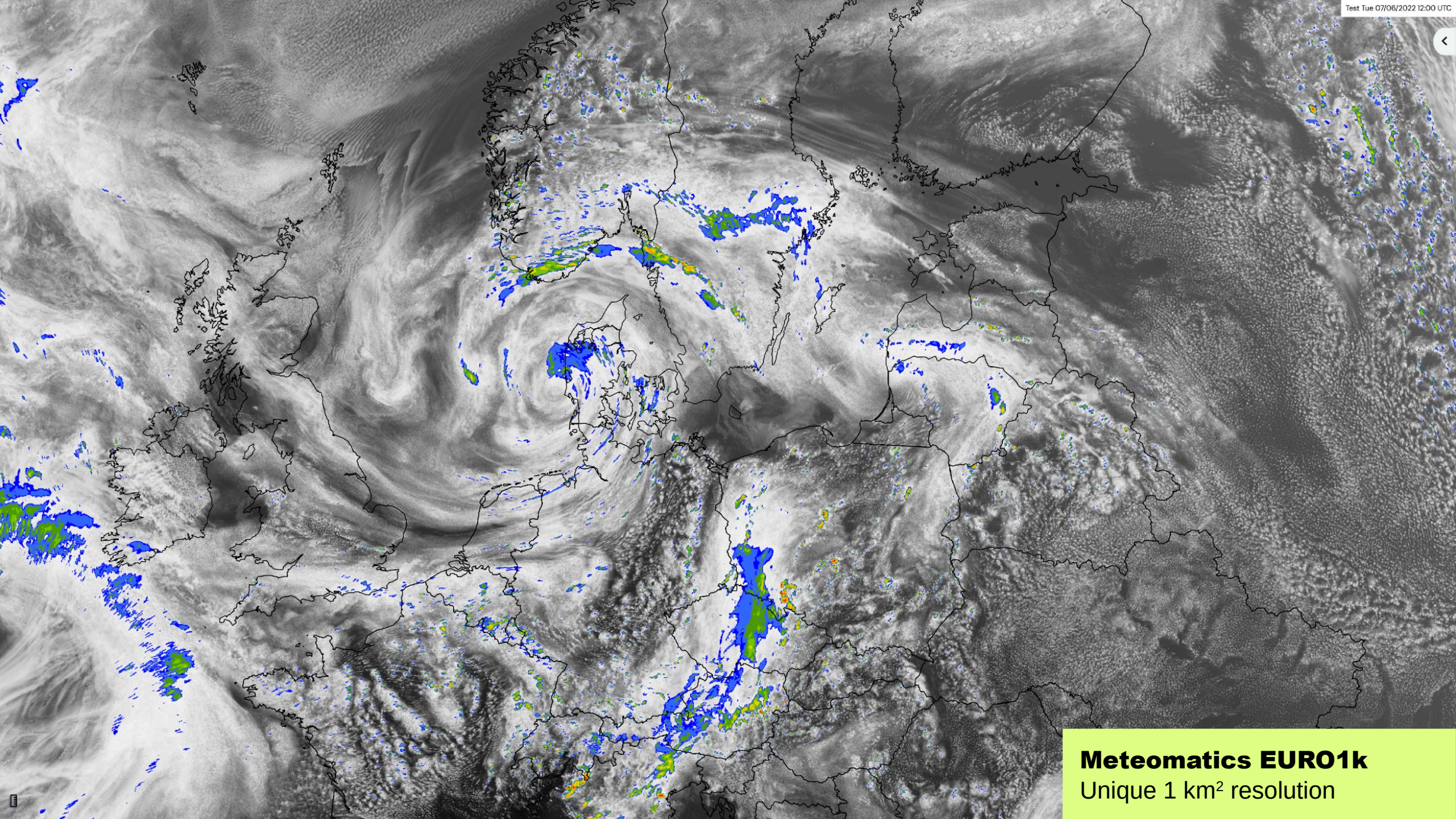
European weather model at 1 km resolution, other areas upon request: get accurate weather forecasts across Europe and beyond to make better decisions.

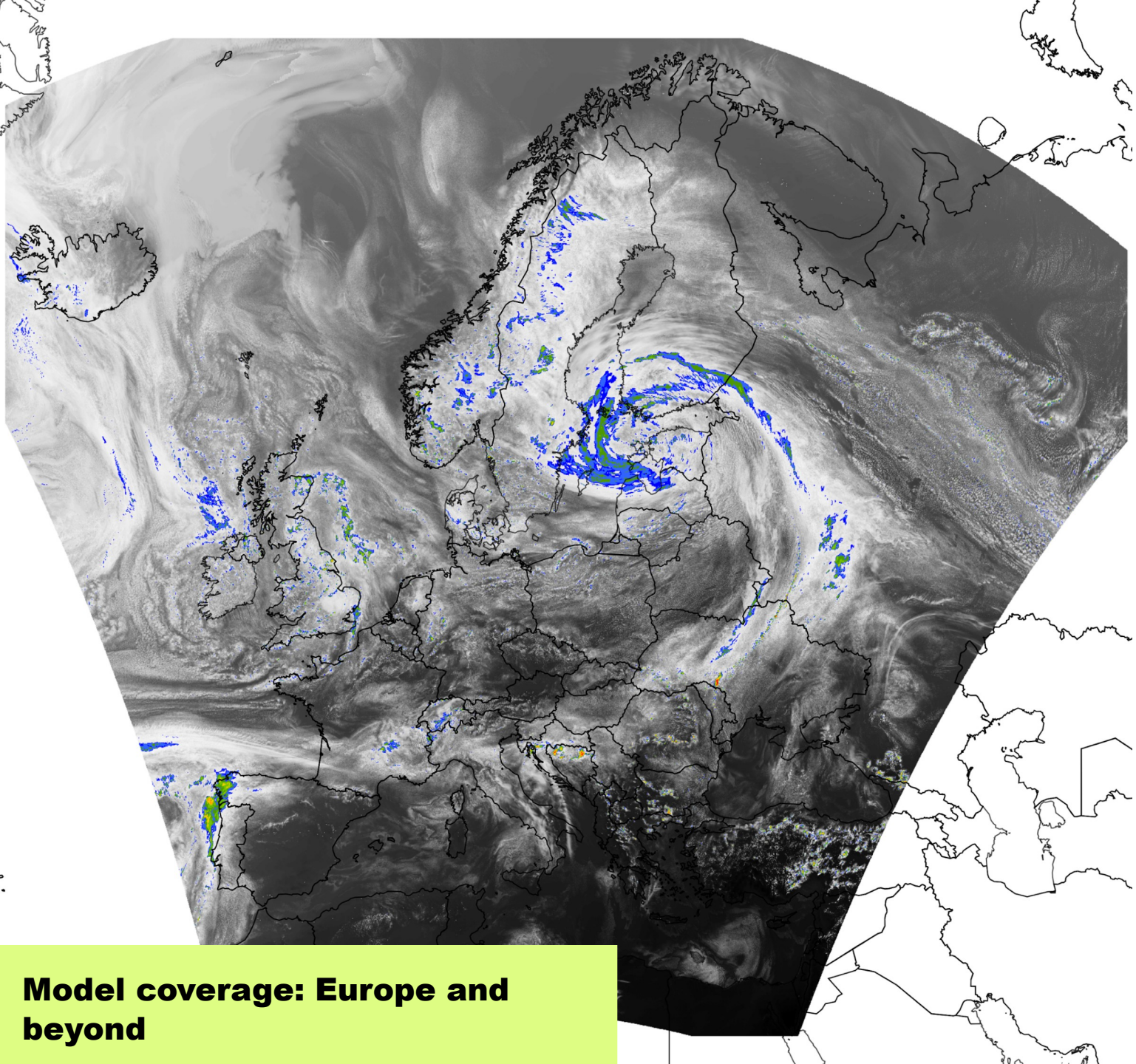
Maps and Visualization

Visualize and analyze all weather events in a high-resolution map view - with our web-based weather map tool MetX based on our Weather API.

Meteodrones

Close the meteorological data gap in the Earth's lower and mid atmosphere with Meteodrones. Designed and assembled in Switzerland.


Test Tue 07/06/2022 12:00 UTC
Meteomatics EURO1k
Unique 1 km² resolution



Model coverage: Europe and beyond

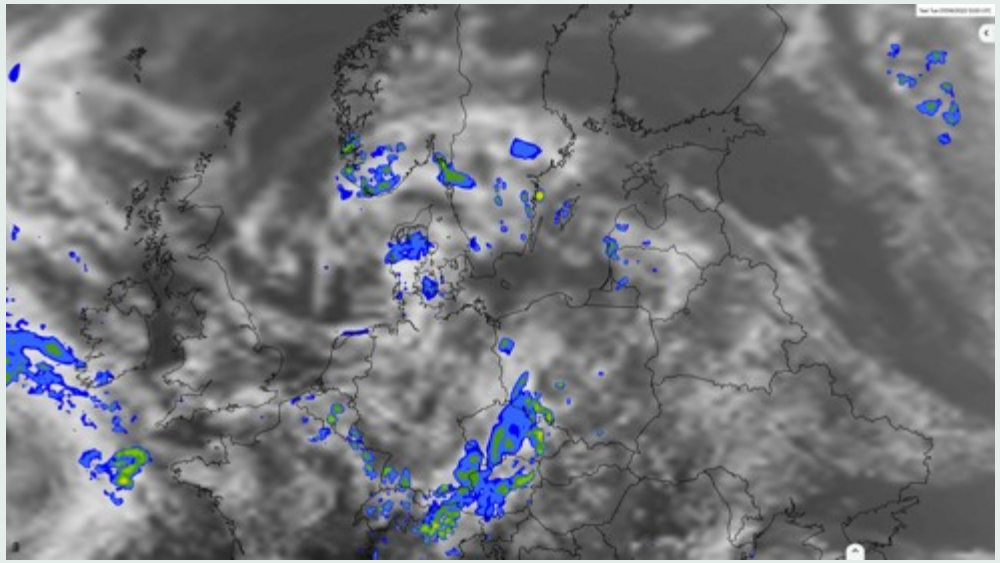
Meteomatics EURO1k Key Details

- Full European domain + North Africa
- 24 runs per day
- Exclusive integration of Meteodrone data
- 1 km^2 native spatial resolution
- Approximately 20 million grid points
- 20 minute native temporal resolution
- 80 vertical levels in 10 m to 100 m steps
- ECMWF-IFS used for boundary conditions

ECMWF vs. EURO1k – Key Details

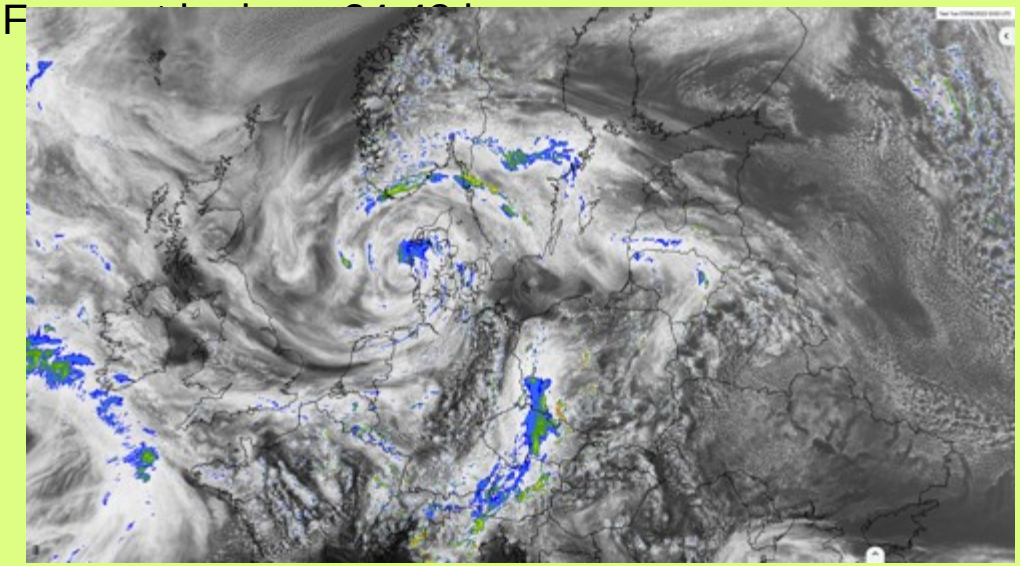
ECMWF – Current Best Practice

- Spatial resolution: 7-9 km
- Temporal resolution: 1 h
- Update rate: every 6 h
- Forecast horizon: 240 h



EURO1k – The Future of Forecasting

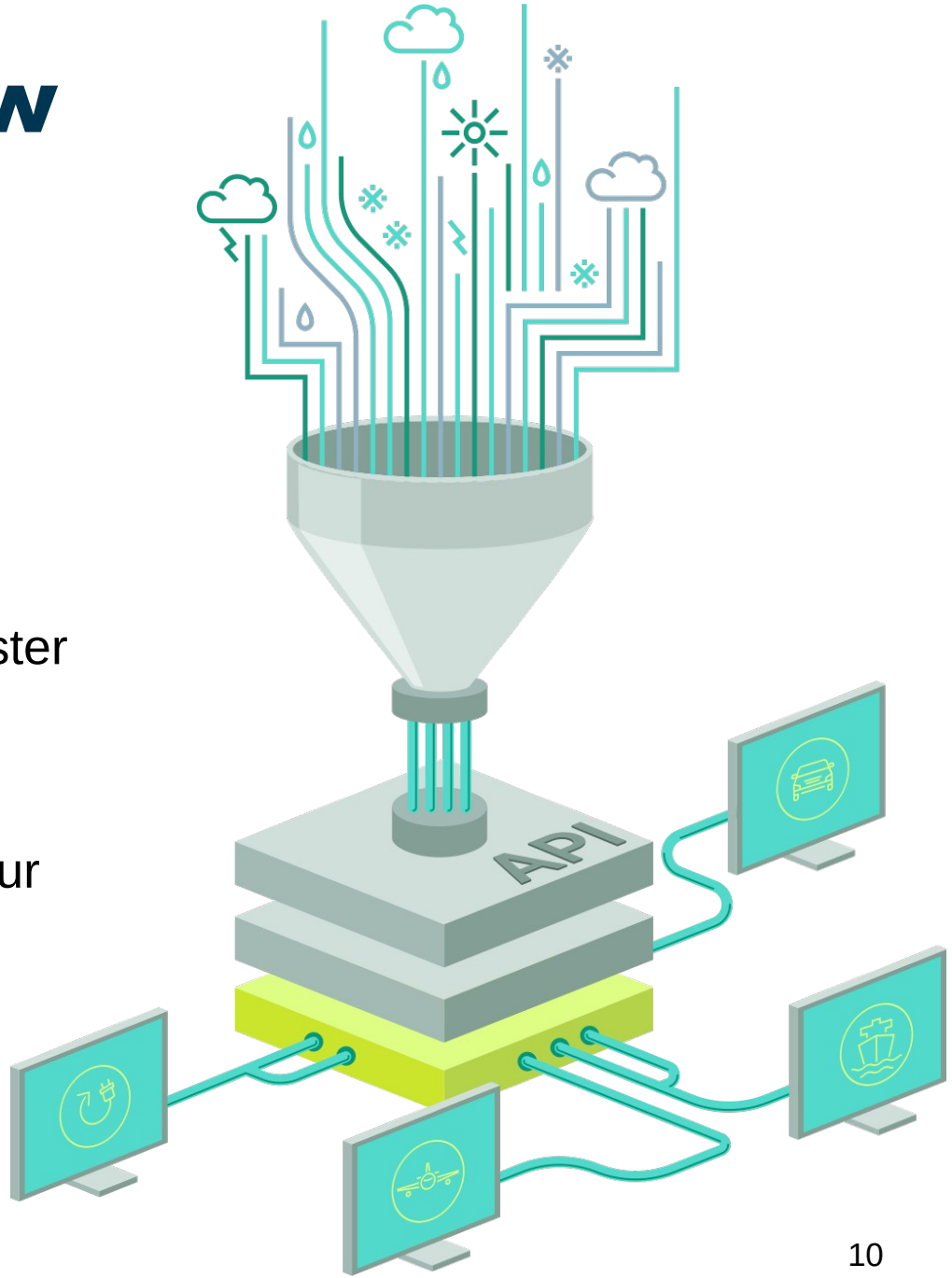
- Spatial resolution: 1 km
- Temporal resolution: 20 min
- Update rate: every hour
- F



High-Performance Computing for EURO1k

EURO1k Worflow

- Production workflow
  - HPC
    - EURO1k runs
    - Post-processing
  - Fetchers
    - Retrieving data from cluster
    - Ingesting data to API
  - API
    - Delivering EURO1K to our customers

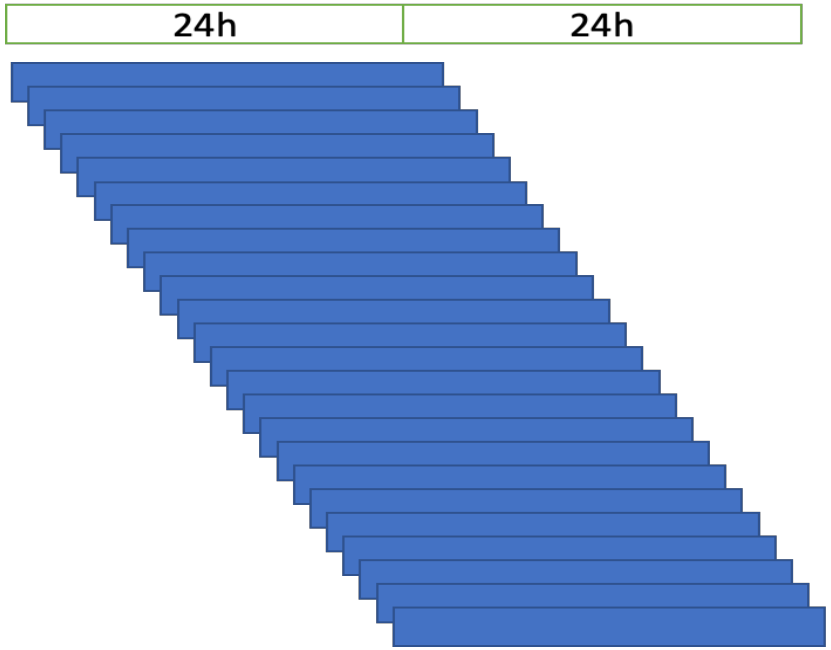


EURO1k technical requirements

- EURO1k is a weather forecasting code written in Fortran
- MPI and MPI+OpenMP parallelization
- Running on ~ 37k cores
- Output
  - 65O GB per run
  - 24 runs a day
- Strong time constraints
  - runs should last less than 50 minutes

EURO1k Worflow

- Production workflow
  - 24 runs a day produced
  - Cycling based on rapid refresh
  - 26h of lead time produced
  - <2h delay from initialization to delivery



HPC on the cloud

- Meteomatics had a first successful experience in the Cloud with a “reduced number” of cores
- Then in Summer 2022, first attempt to move production to the cloud
- Not able to deploy the rapid update workflow
- Lots of issues to ensure availability of such high volumes of nodes
- Results
  - Degraded workflow in
  - Decision: buy our own HPC
- Tender open
  - NEC won the tender
  - and I joined Meteomatics
  - Objective: deploy an HPC

Deploy the HPC

Always been a proficient user

and now deploying a cluster from the customer side interacting with vendors

First thing learned

I was totally unaware of many things

- Delivery times
- Networking and security
- Burning tests
- Deployment tools and time
- And many more…

Each of those points, can have a great impact on the schedule

Hardware deployment started in January, mid February first access to the cloud

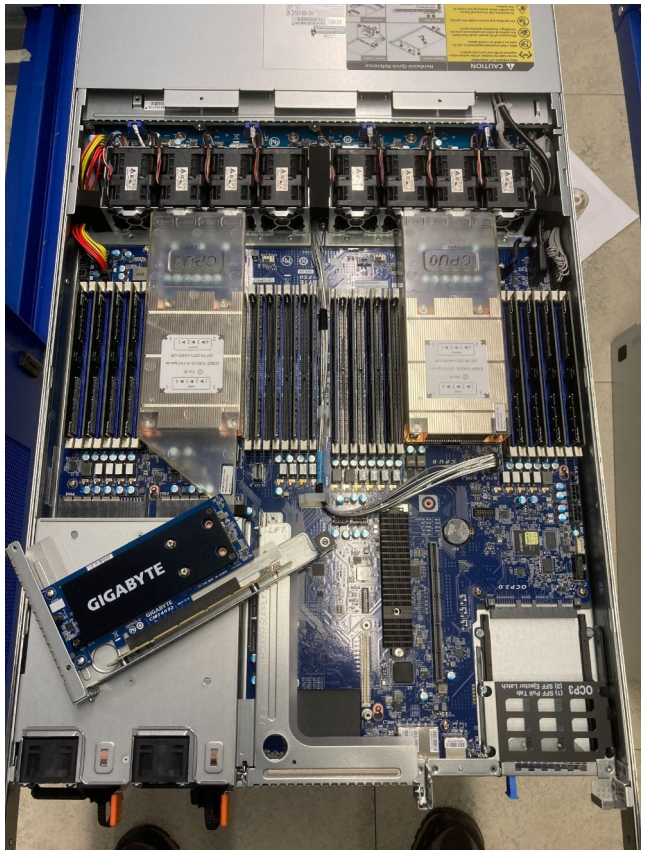
Cluster description

The High-Performance cluster is made of:

- login nodes (connection, software deployment)
- management nodes (to administer the cluster)
- storage nodes (to deliver 240 Tb of parallel disk)
- 300 compute nodes
- a high-performance network

Occupying 14 racks (max. consumption 290 kW)

Deployed by NEC Germany.



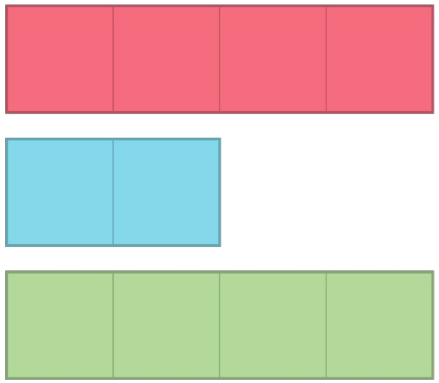
Details of a compute node

EasyBuild

Special slide for this amazing tool and team

With Easybuild the time to start running your code is strongly reduced

- combination of toolchains
- most of usual dependencies are there
- modules automatically generated
- recreate the software stack if needed



~~Domain scientists shouldn't deploy software stacks~~ (never said that)

Scientists should focus in compiling and testing their applications

EASYBUILD.io
building software with ease

Second thing learned

Automatize as many as possible (probably you'll need to repeat tasks)

Workflow manager

- Complex workflows can’t be maintained using a simple bash/python script
- Use existing tools for your domain (or create your own in the worst case)
- Workflow management tools
  - reduces deployment time
  - simplify the production operation
  - enable the operation to non HPC/domain experts
- Invest and add a strong autorecover feature (not easy but strongly recommended). Take advantage of some SLURM features (i.e. node ranges)

Acceptance tests

Prepare acceptance conditions in advance

- Synthetic benchmarks (HPL, STREAM, OSU, …)
- Application
  - Vendors probably don’t know your application
  - Spend time to define clear uses cases with expected performance and results

Third learning

Distribute a precise benchmark to the vendor

Thanks to EasyBuild, the complexity of the task to test different toolchains including Compilers, Libraries and MPI implementations is extremely reduced.

Input and Output

Fourth learning

Think on I/O. Not everything is computation

GPFS (and I extend to other parallel filesystems) is complex to operate and get full performance

In fact, we easily reach good computing times but struggled to get performance in IO

Then you start to learn things as MPI-IO, ROMIO, PnetCDF, Asynchronous I/O … until you get the expected performance.

- IO is hard to profile.
- Tested tools like darshan but hard to use it
- We used benchmarks: BTIO, IOR, …
- Ended up modifying EURO1k source code
- It took two weeks to find the right configuration

Bonus track: transferring so much data with traditional rsync/scp is not efficient… We had to deploy other tools…

What’s next?

Production

Finally, 19th March moved to production

But work is not done, still many tasks to do

- prepare your HPC for operational activity
- deploy monitor tools, alerts, sanity checks, testing, (Reframe?), write procedure for duty officers, …
- deploy CI/CD strategy

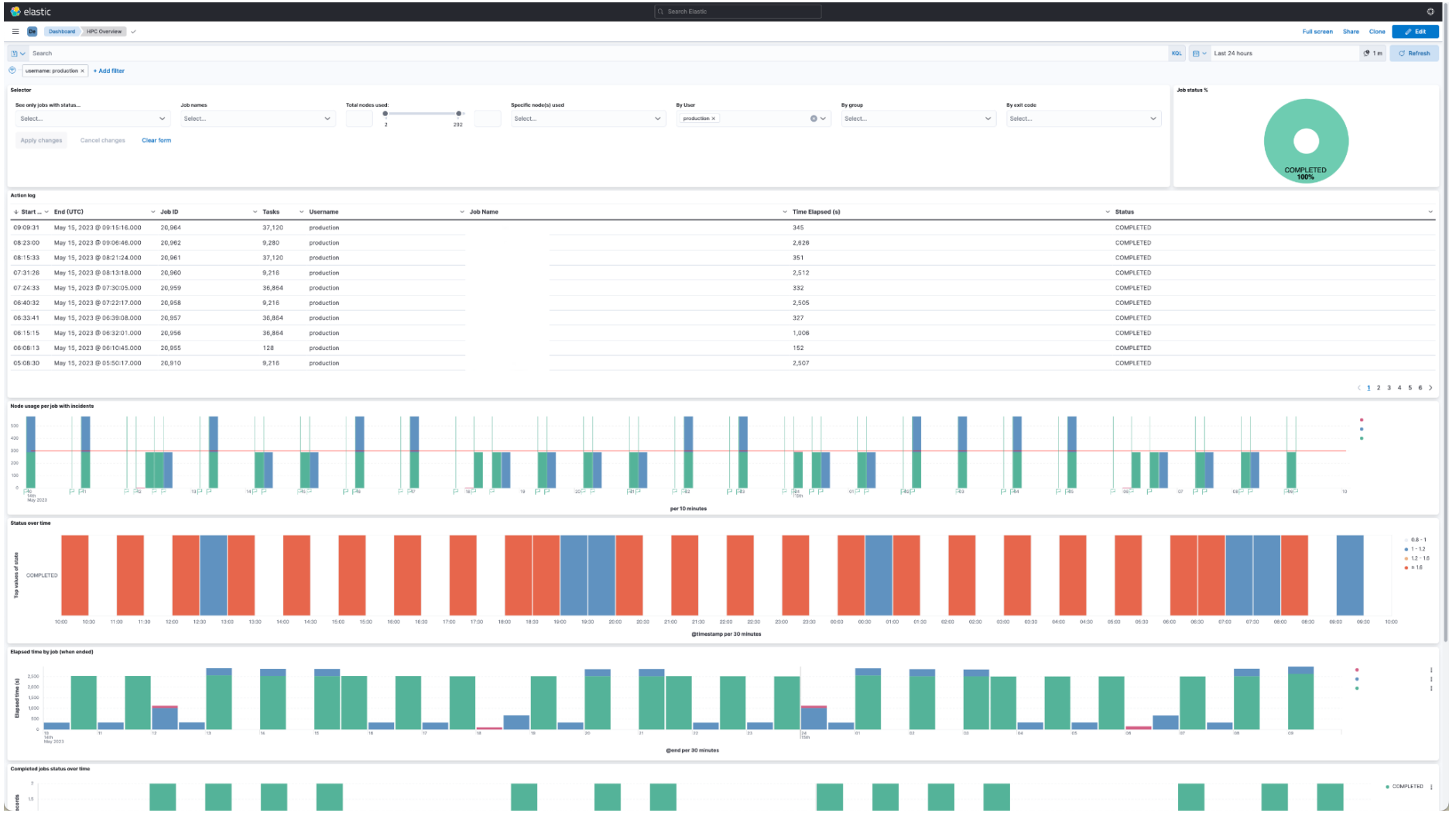
Fifth learning

Nodes have regular [minor] problems, more than expected

Tools like Lawrence Berkeley National Laboratory Node Health Check (NHC) are extremely helpful to maintain operational activity

Monitoring

Invest in monitoring tools (i.e. Elastic and SLURM integration)



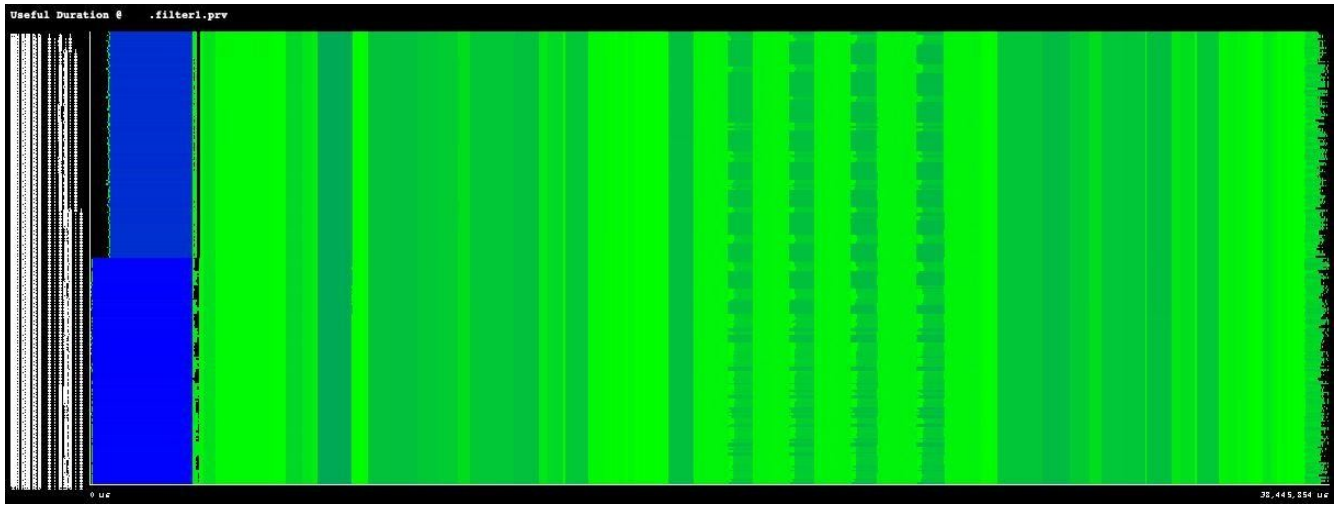
Improve the performance

Not an easy task

We don’t have a full scale development environment

Some ideas:

- Performance analysis and profiling
  - Performance tools are not prepared for this scale (EXTRAE has issues to deal with a production run)
- Test other compilers, libraries, fabrics (OPX)…
- Improve IO (test novel approaches)



Useful duration trace of a EURO1k execution using 4 nodes

Join us

To do such tasks, we need people!

Visit Meteomatics careers (https://careers.meteomatics.com/)

- HPC Engineer (f/m/d)
- IT Support Engineer (f/m/d)
- Site Reliability Engineer (f/m/d)

Last learning (not new)

It’s good to invest in machines but machines need humans to work. Invest in talent.

Lessons learned

- I was totally unaware of many things
- Automatize as many as possible (possibly you'll need to repeat tasks)
- Distribute a precise benchmark to the vendor
- Think on I/O. Not everything is computation
- Nodes have regular problems, more than expected
- It's good to invest in machines but machines need humans to work. Invest in talent.

Questions

Disclaimer: I will take all your questions, except the ones related with drought in Spain.

kserradell@meteomatics.com

Thanks

meteo matics