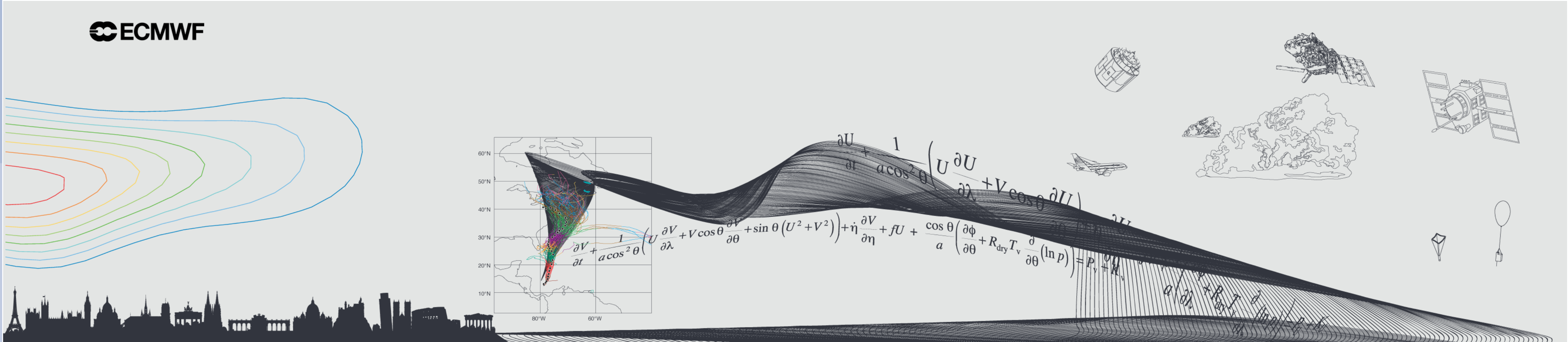
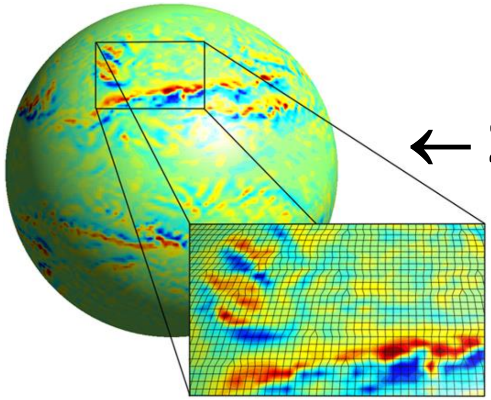
2 years into ECMWF's Scalability Programme: What have we achieved?

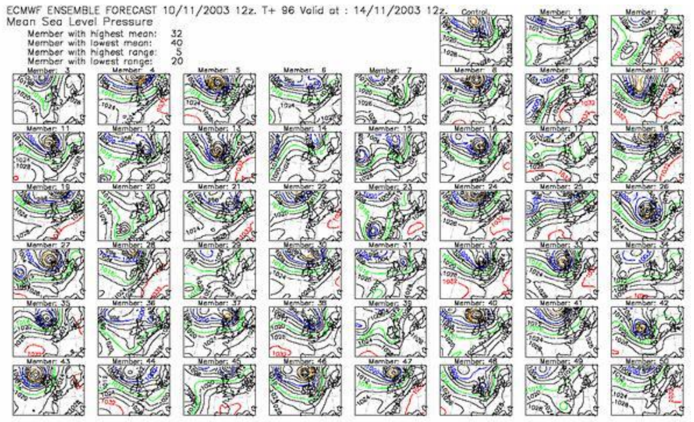
Peter Bauer and many colleagues



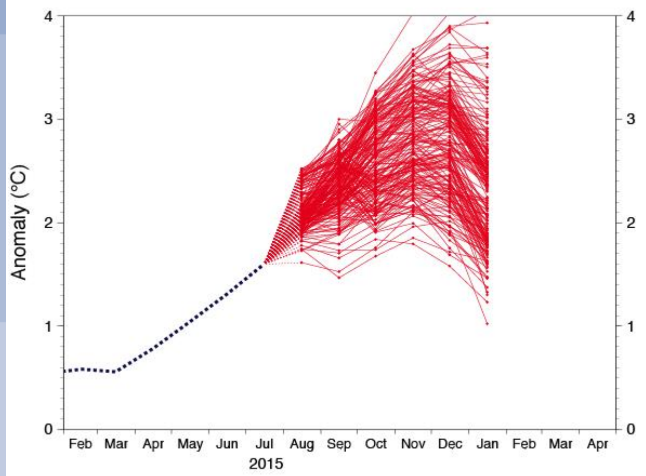
The ECMWF Integrated Forecasting System (IFS)



← 2x 9-km global high-resolution 10-day forecasts per day



51x 18-km global lower-resolution 15-day forecasts per day… →
… extended to 46 days twice per week at 36 km



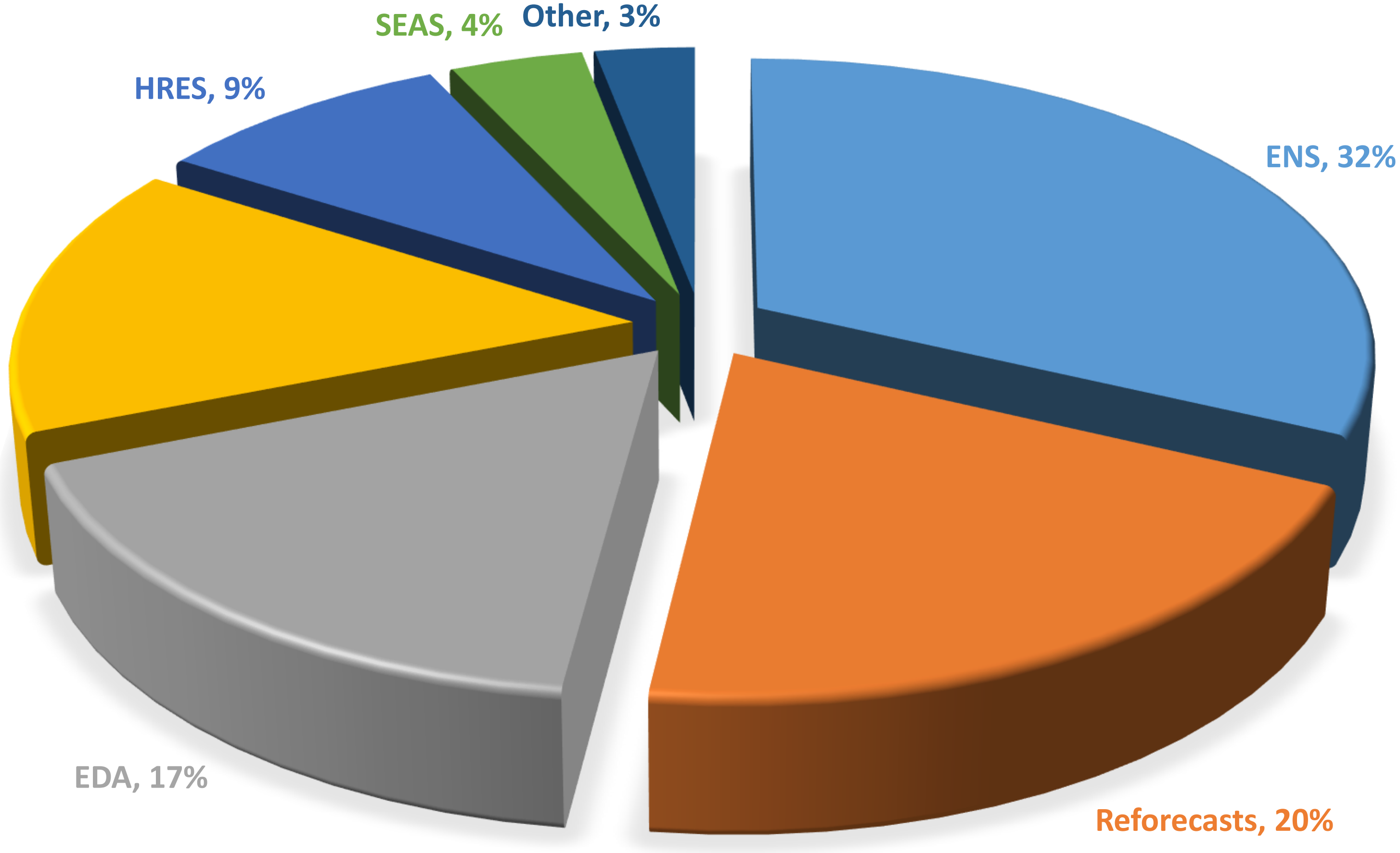
← 51x 64-km global low resolution 7-month forecast per month

Node-time allocations operational suites

Suite		Nodes	Time [s]	Nodes x Time [h] /day [d]	Comment
EDA		26 x 28 = 728	3200	1300	2/day; in critical path together with 4DV
ENS legA		51 x 20 = 1020	5200	2960	2/day; in critical path together with HRES
Reforecasts		20 x 11 x 10 = 2200	11500	2010	2/week; 20 years done in batches with max. allocation of 500 nodes at once
4DV	LW SW	352 352	3150 1820	615 360	2/day; in critical path together with EDA 2/day; In critical path together with EDA
HRES	LW SW	352 352	800 2800	160 550	2/day; in critical path together with ENS 2/day, in critical path together with ENS

(1 cluster ~3500 nodes
1 electrical group = 360 nodes)

Node-time allocations operational suites



= 25% of the **capacity** (nodes x time), and max. 40% of **capability** (nodes)

ECMWF’s 10-year strategy: 2016-2025

http://www.ecmwf.int/en/about/who-we-are/strategy :

- […] integrated global model of the Earth system to produce forecasts with increasing fidelity on time ranges up to one year ahead [...]

- […] skilful ensemble predictions of high-impact weather up to two weeks ahead. By developing a seamless approach, we also aim to predict large-scale patterns and regime transitions up to four weeks ahead, and global-scale anomalies up to a year ahead.

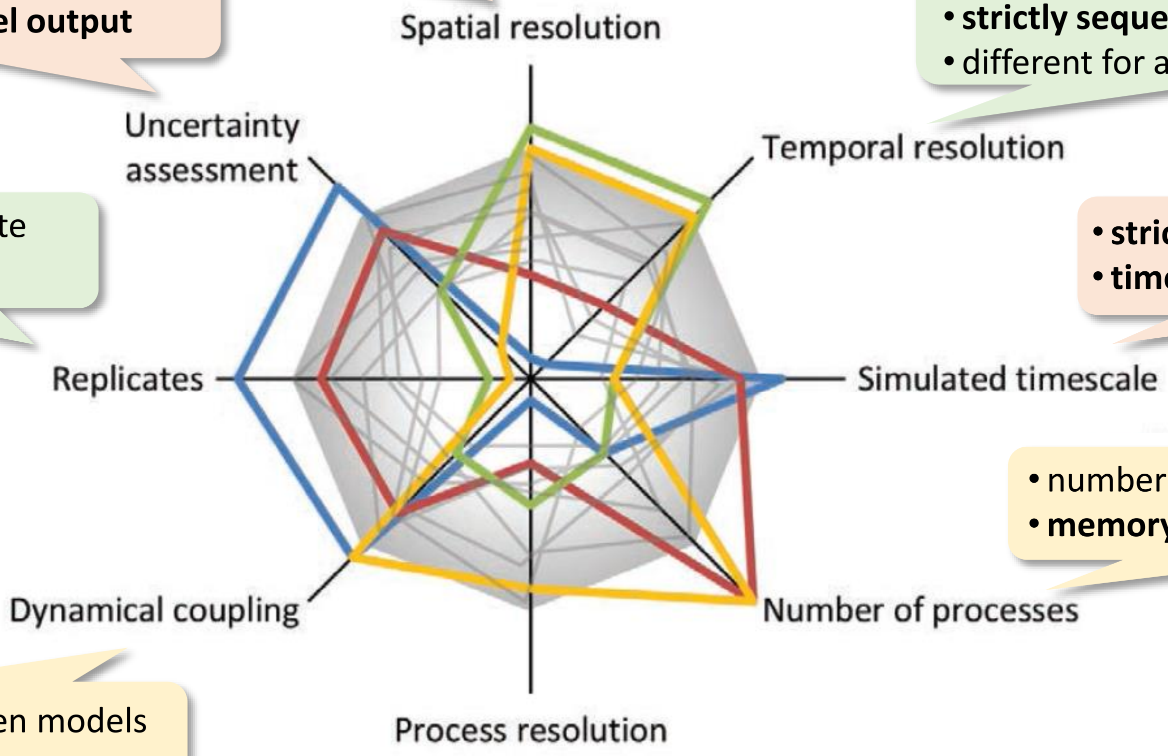
→ **Key quantifiable target: global 5-km ensemble by 2025**

Can’t have it all?

- quadratic # grid points
- **global communication**
- **memory limited**

- **multiplies entire model compute**
- **multiplies entire model output**

- depends on spatial resolution
- **strictly sequential, time stepping scheme**
- different for atmosphere, ocean



- multiplies entire model compute
- multiplies entire model output

- **strictly sequential**
- **time to solution** (weather vs climate)

- number of prognostic variables
- **memory limited**

- communication of data between models
- **latency between models**

- number of prognostic variables
- **memory limited**

[Smith et al. 2014, BAMS]

The 5-year challenge= ½ way

- a global N-member ensemble at 9 km resolution (up to day 15 in critical path),
- that is coupled to a land, ¼ degree ocean and a sea-ice model,
- that includes prognostic atmospheric composition,
- and that is initialized with a N-member hybrid variational/ensemble analysis with 9 km resolution, land, sea-ice and ocean model coupling and atmospheric composition.

With N=51 the cost increase towards the above target configuration would be:

- **Ensemble analyses**:
  hor. resolution x5, coupling x1.2, ensemble size x2, atmos. composition x1.2
- **Ensemble forecasts**:
  hor. resolution x5, vert. resolution x1.5, coupling x1.2, atmos. composition x1.5
- **Reforecasts**:
  ensemble size x1.6, hor. resolution x5, vert. resolution x1.5, coupling x1.2, atmos. composition x1.2

~ x15

→ Just for the ensemble forecasts ~x4.5 one XC-40 cluster

ECMWF Scalability Programme

Governance: ECMWF, Member states, Regional consortia

Projects:

Talk by Peter Lean

Observation processing:
- Lean workflow in critical path
- Object based data store
- Screening/bias correction

Data assimilation:
- Flexible algorithms (C++)
- IFS integration
- Coupling with ocean and sea-ice
- Parallel minimization

Talk by Nils Wedi

Talk by Andreas Müller

Numerical methods:
- Numerical methods
- h/v/t-discretization, multiple grids
- Prognostic variables, coupling

Talk by Tiago Quintino

Model output processing:
- Broker-worker workflow
- Near-memory processing
- Data compression

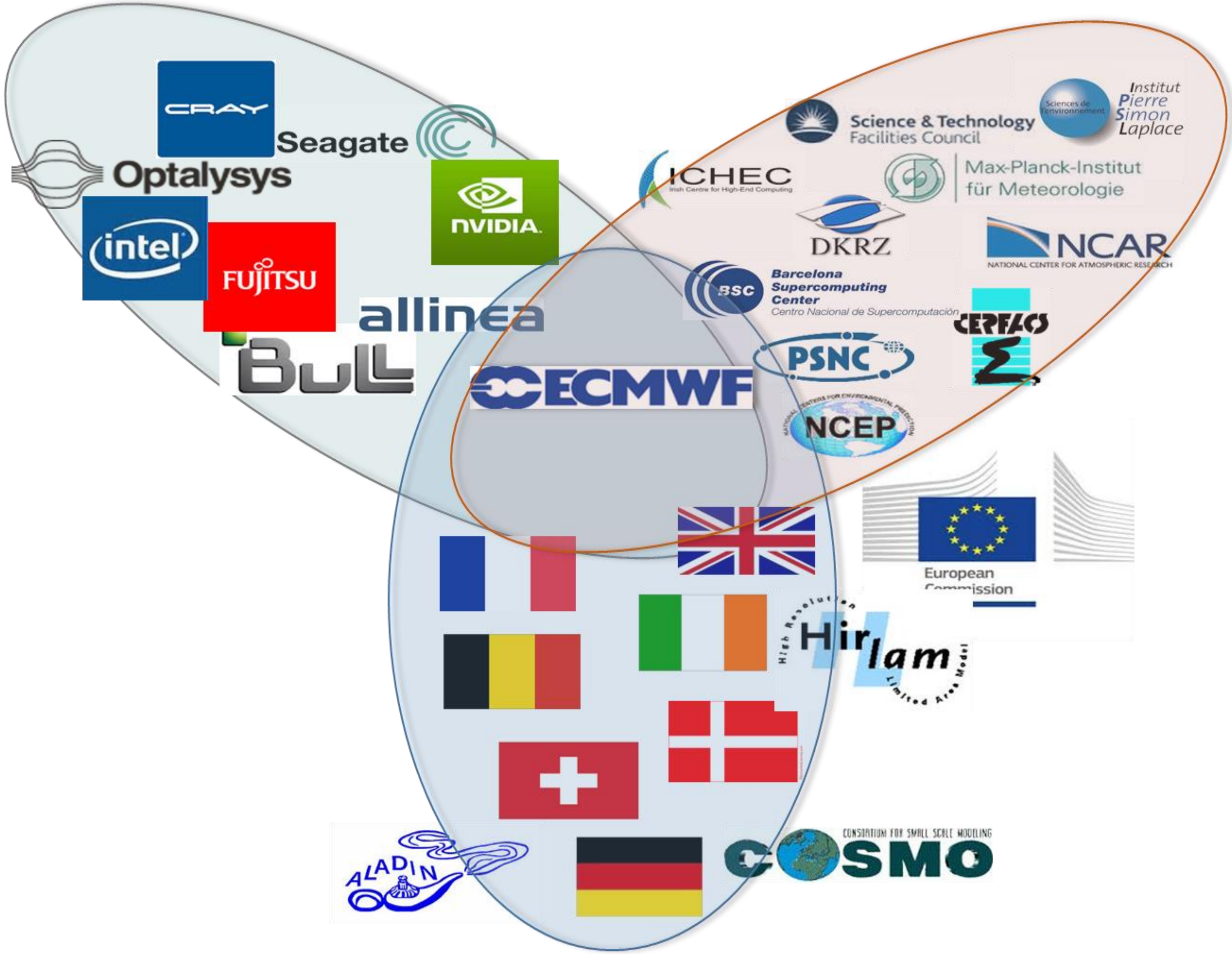
Talk by George Mozdzynski

Talk by Deborah Salmond and Sami Saarinen

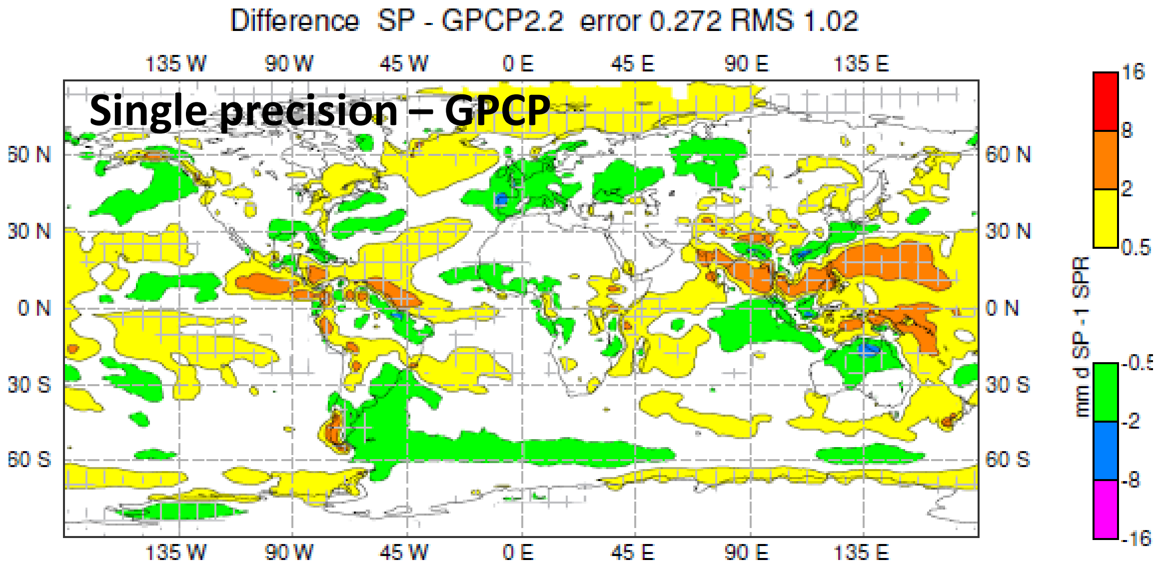
Code adaptation:
- Be
- Vectorization
- Programming models
- Precision
- DSL/libraries
- Transition to operations

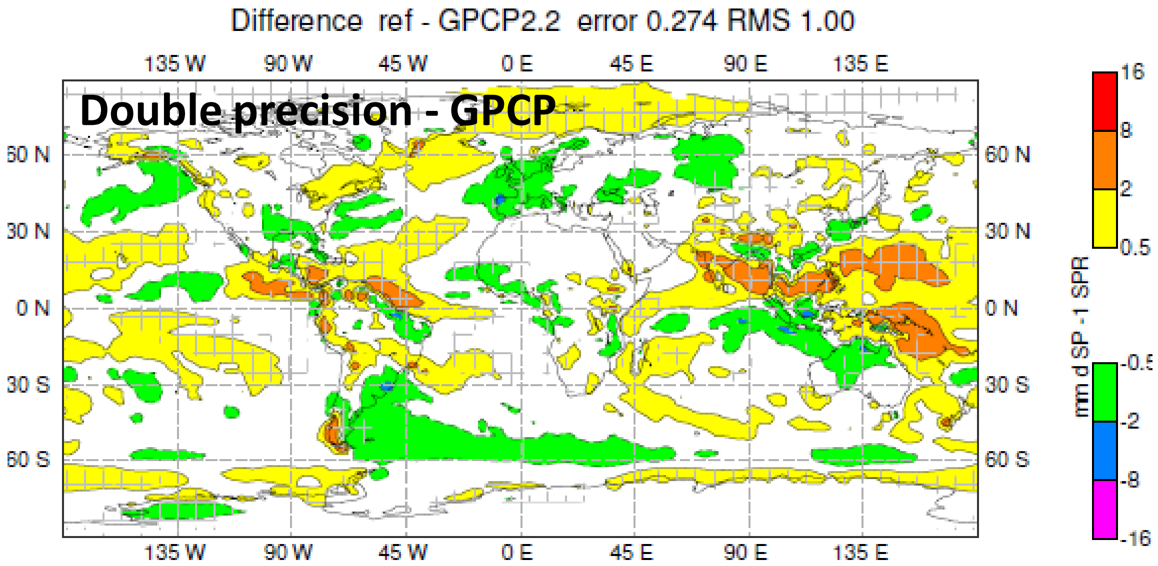
Computer architecture support :
- Benchmarking
- Novel architectures
- Compilers
- Vendor support
- Transition to operations

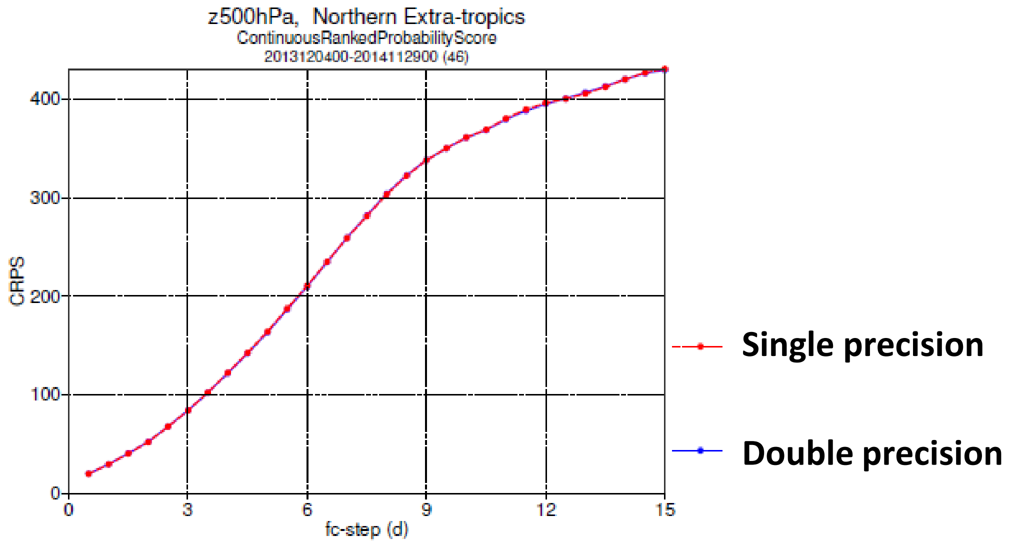
ECMWF Scalability Programme Partnership



Low hanging fruit: Single precision IFS





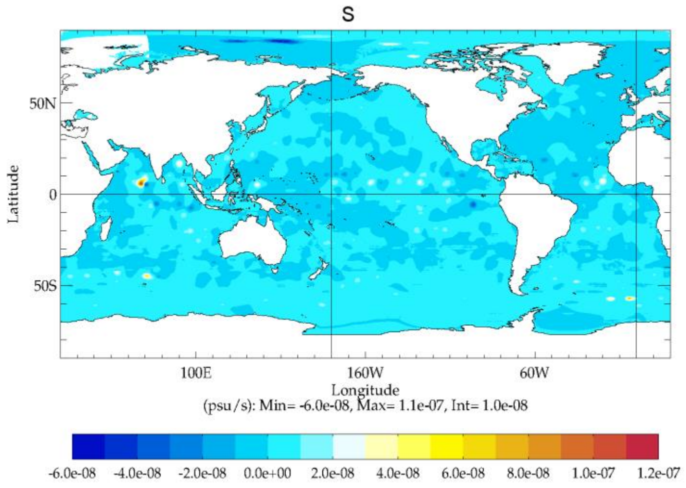


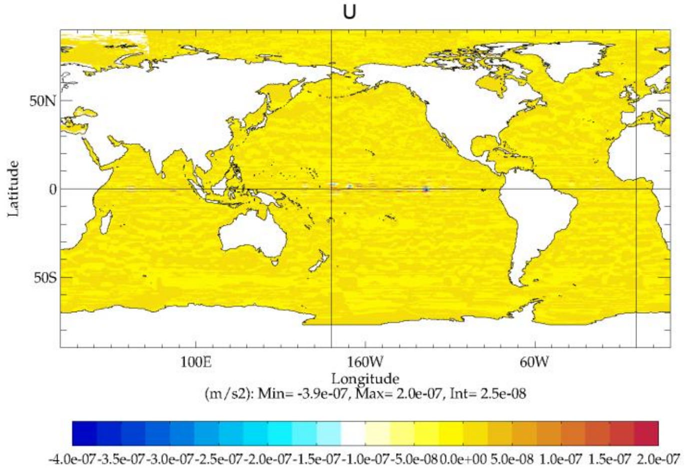
Up to 40% efficiency gains through enhanced memory utilisation; mostly relevant for ensemble forecasts

Need to protect sensitive code components (Adjoint, matrix inversions)

[F. Vana & P. Dueben]

Low hanging fruit: Single precision NEMOVAR

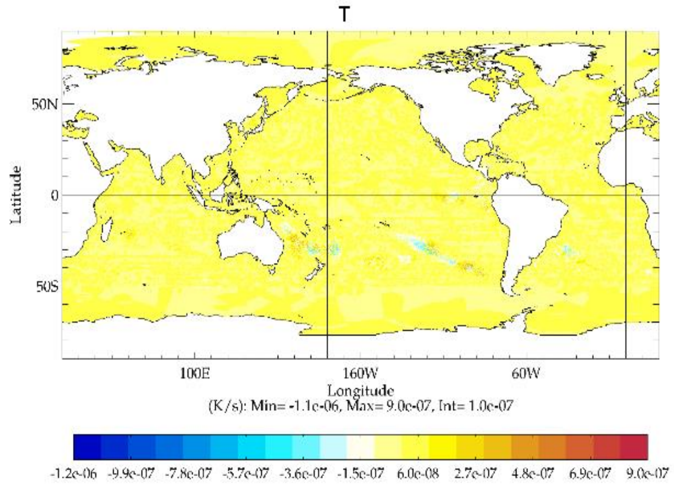


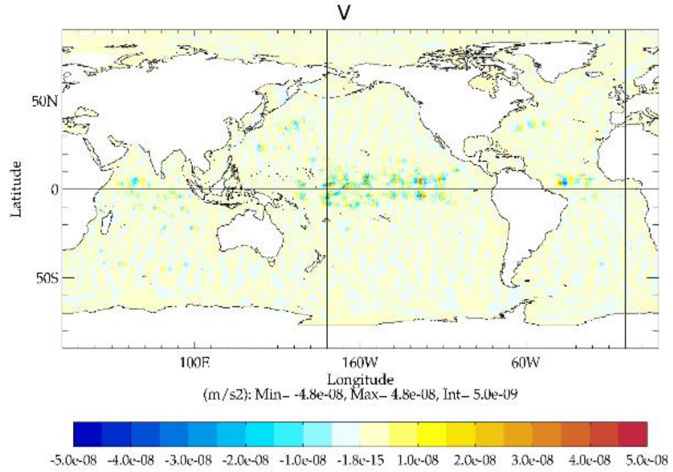


Difference from use of double – single precision in Chebyshev iteration solver

ORCA ¼ degree grid, 5-day assimilation window, at sea surface.

→ Reduction in run-time for the same case corresponds to a speed-up of 1.29x for 384 and 1.12x and for 786 MPI tasks.



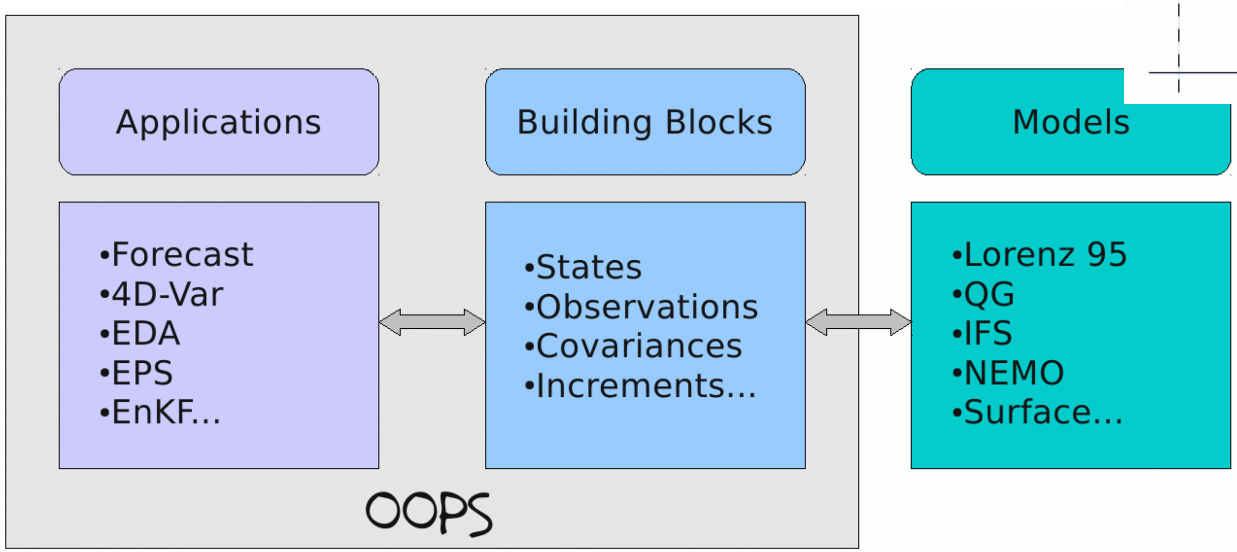


[M. Chrust]

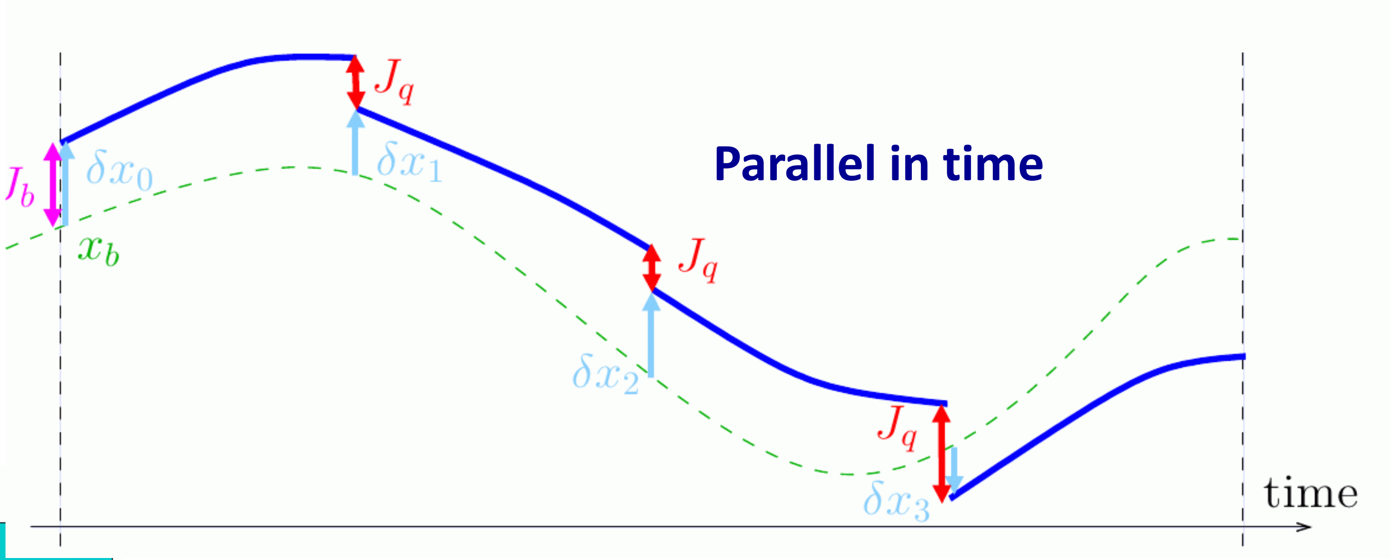
Flexibility and efficiency: Data assimilation

Object Oriented Prediction System (OOPS)

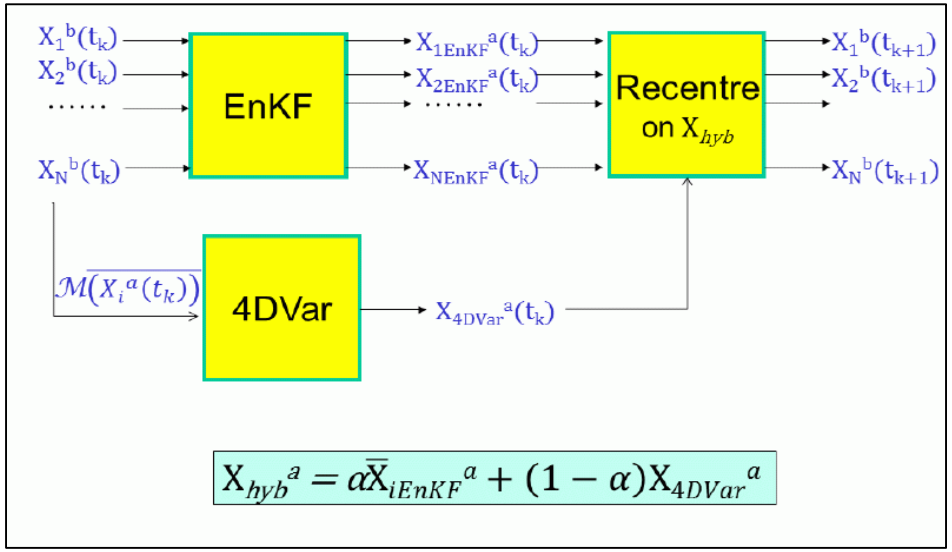
Object oriented



Parallel in time

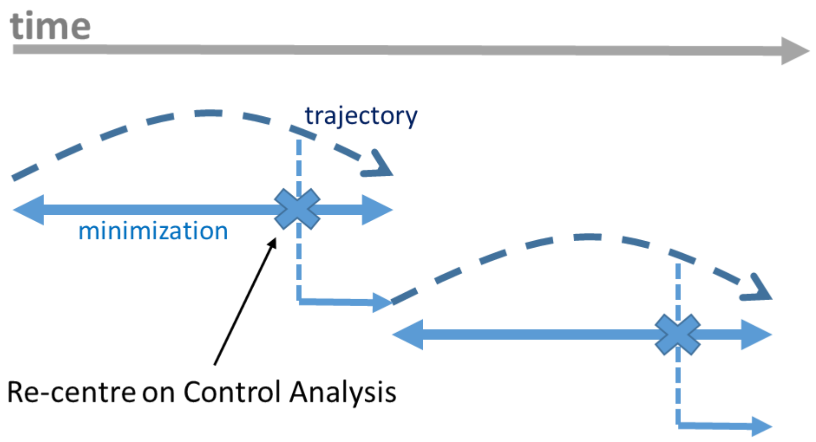


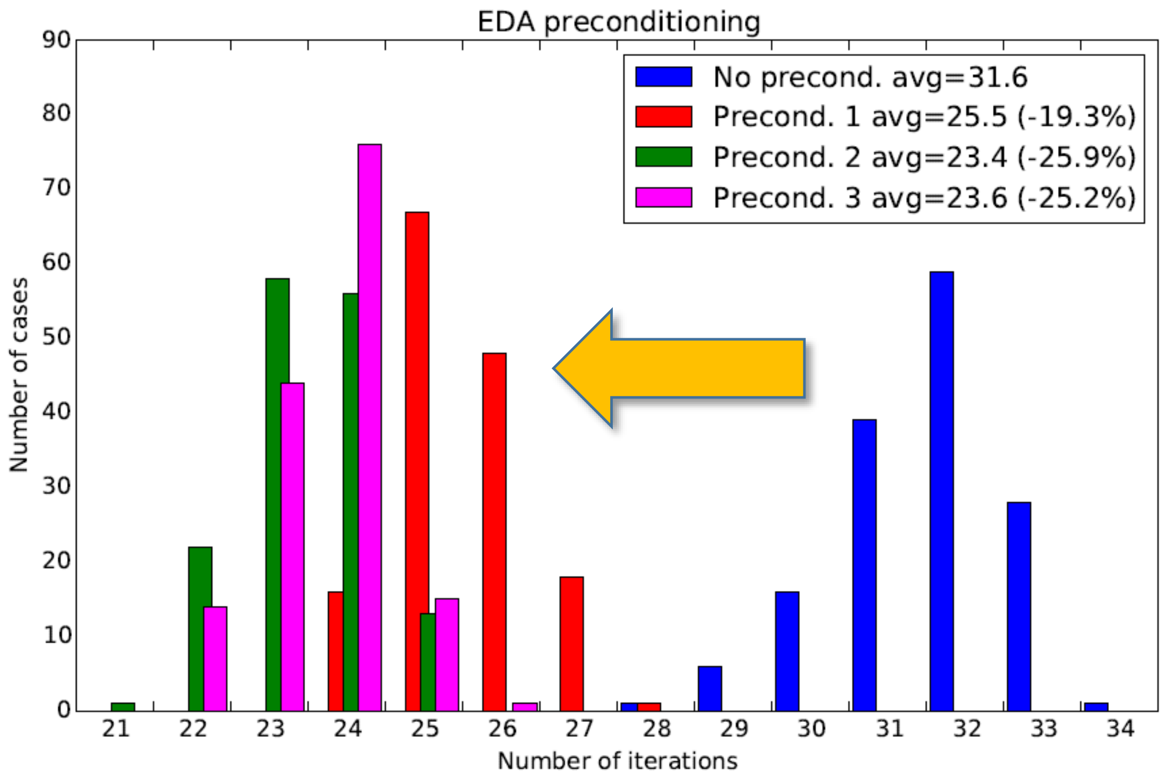
Algorithms



[Y. Tremolet]

Where efficiency defines science: EDA Design





Old: all ensemble members with same configuration
→ **New**: high-resolution control, low-resolution perturbed members

= 40% efficiency gains (and significant skill/reliability improvements)

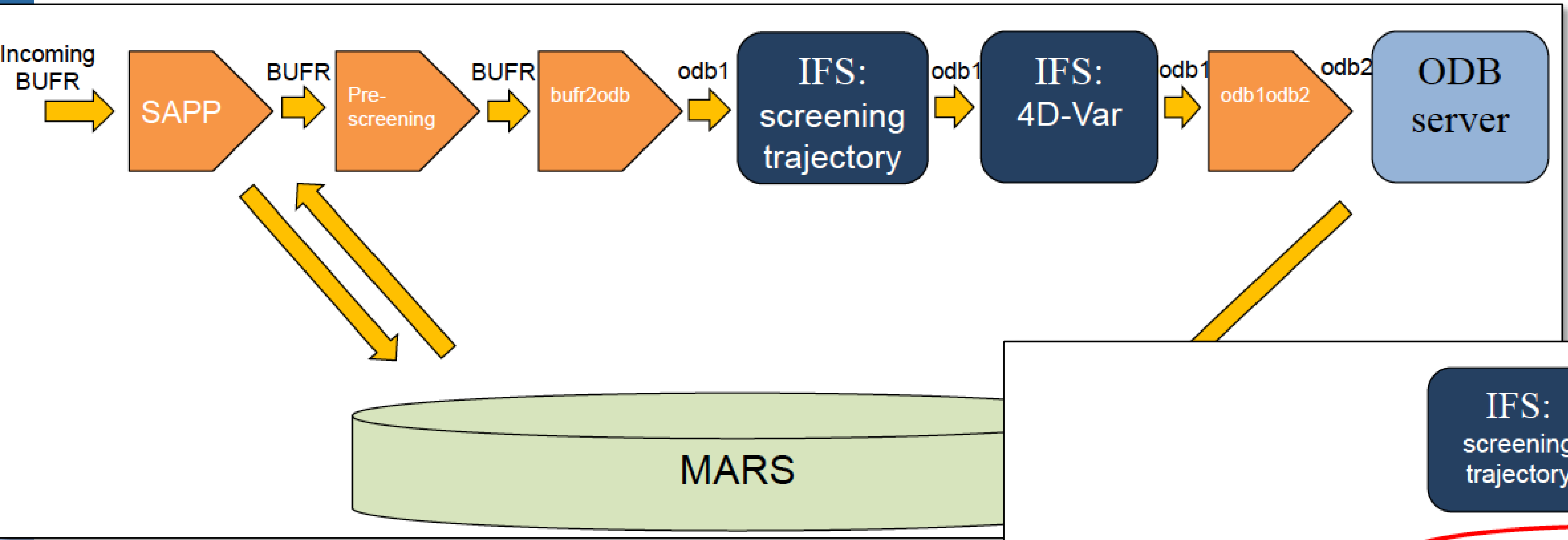
Preconditioning of perturbed members with co-variance statistics from control member

= 25% efficiency gains

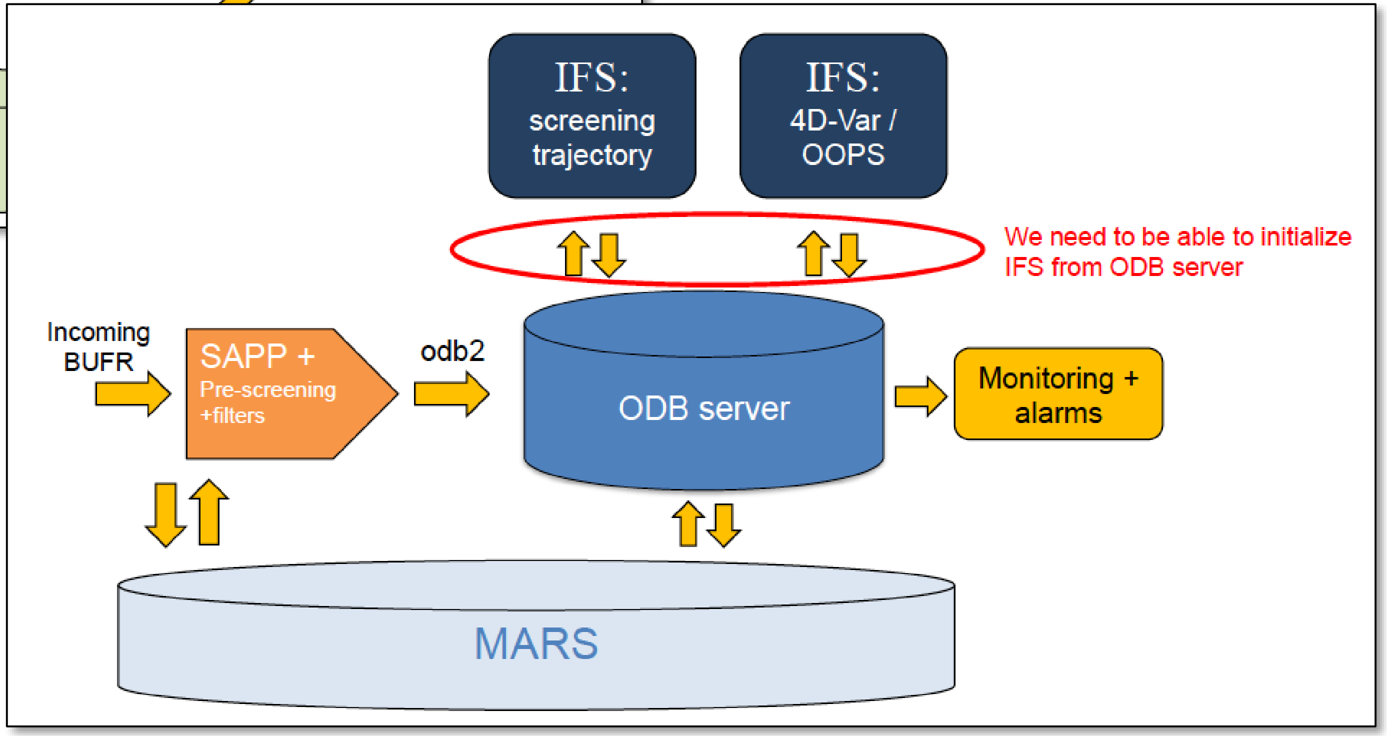
[S. Lang & Y. Tremolet]

Workflow: Observational data processing

From this …



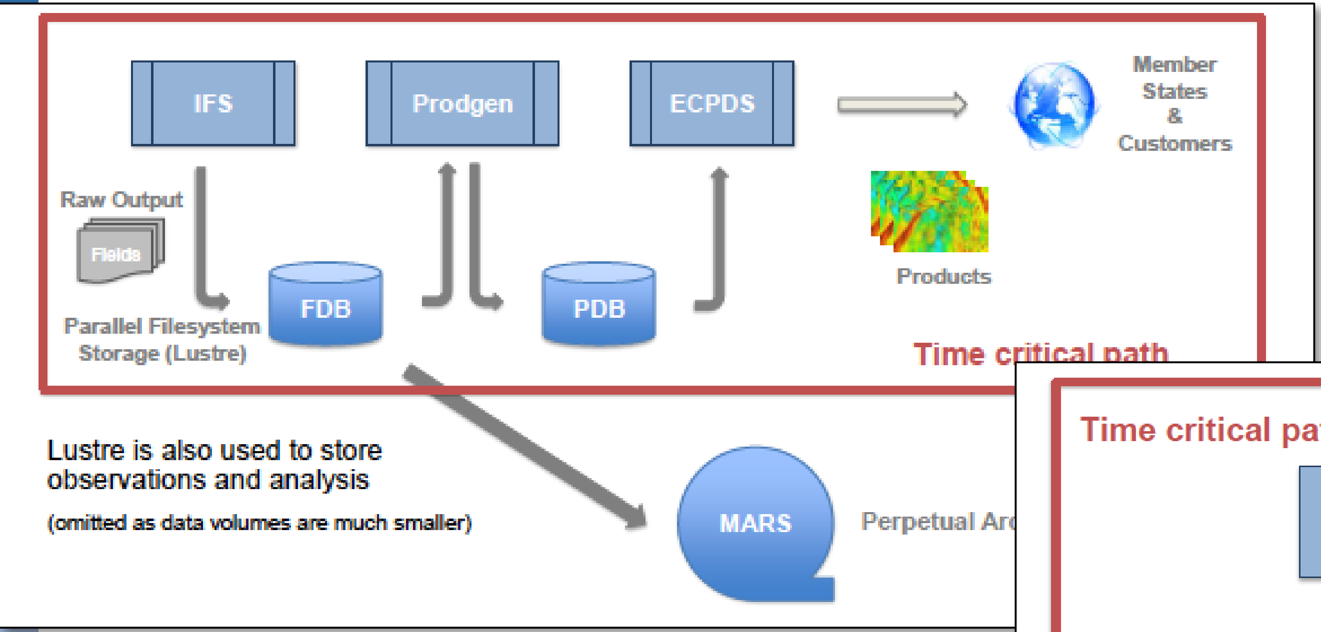
… to this



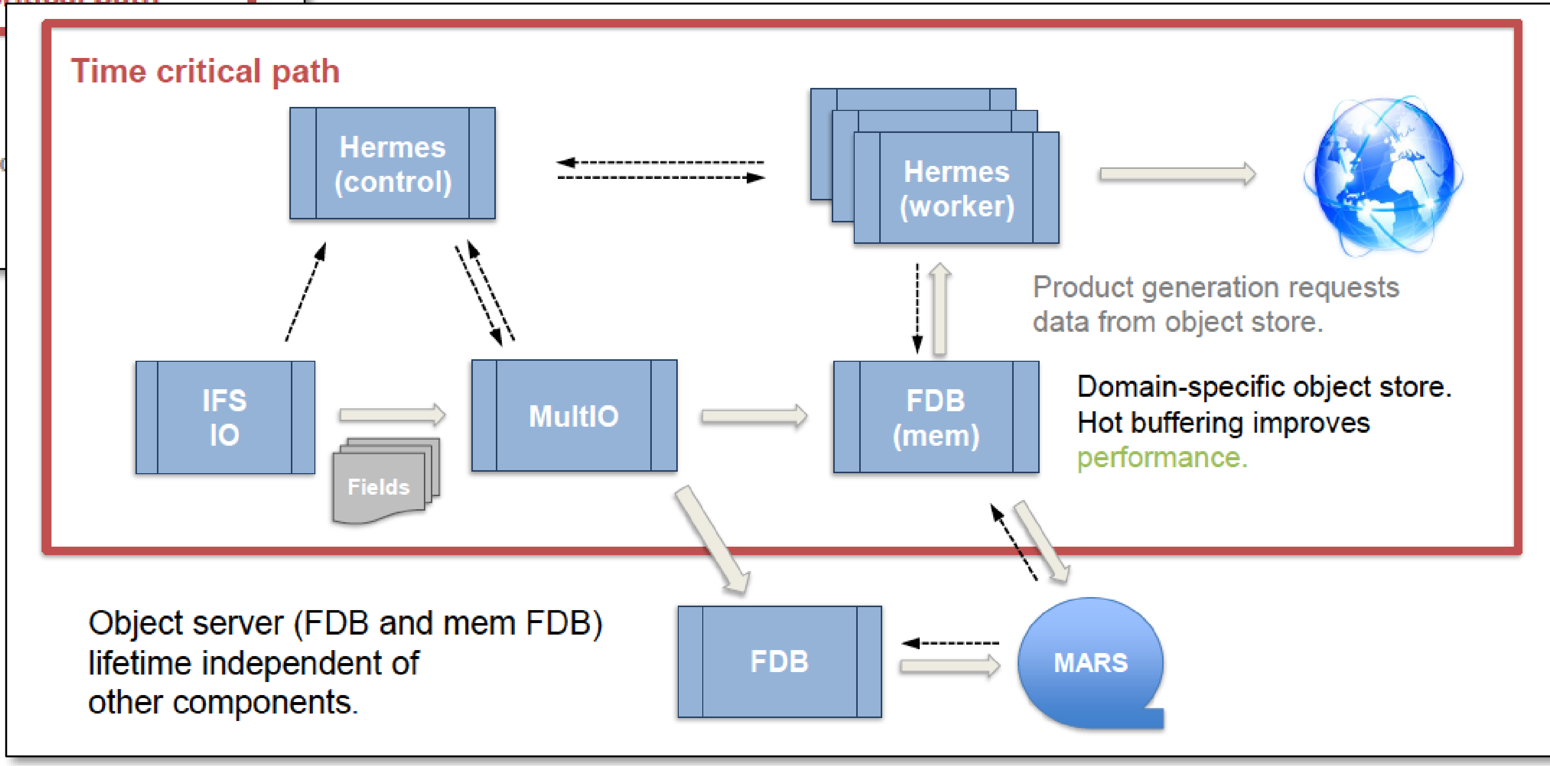
[E. Fucile et al.]

Workflow: Model output data processing

From this …



… to this



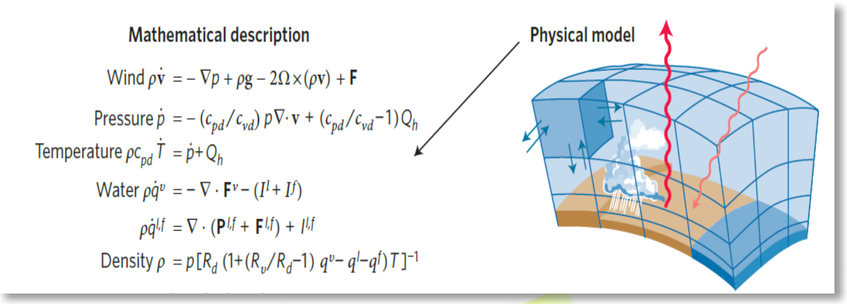
[T. Quintino et al.]

Model development: ESCAPE

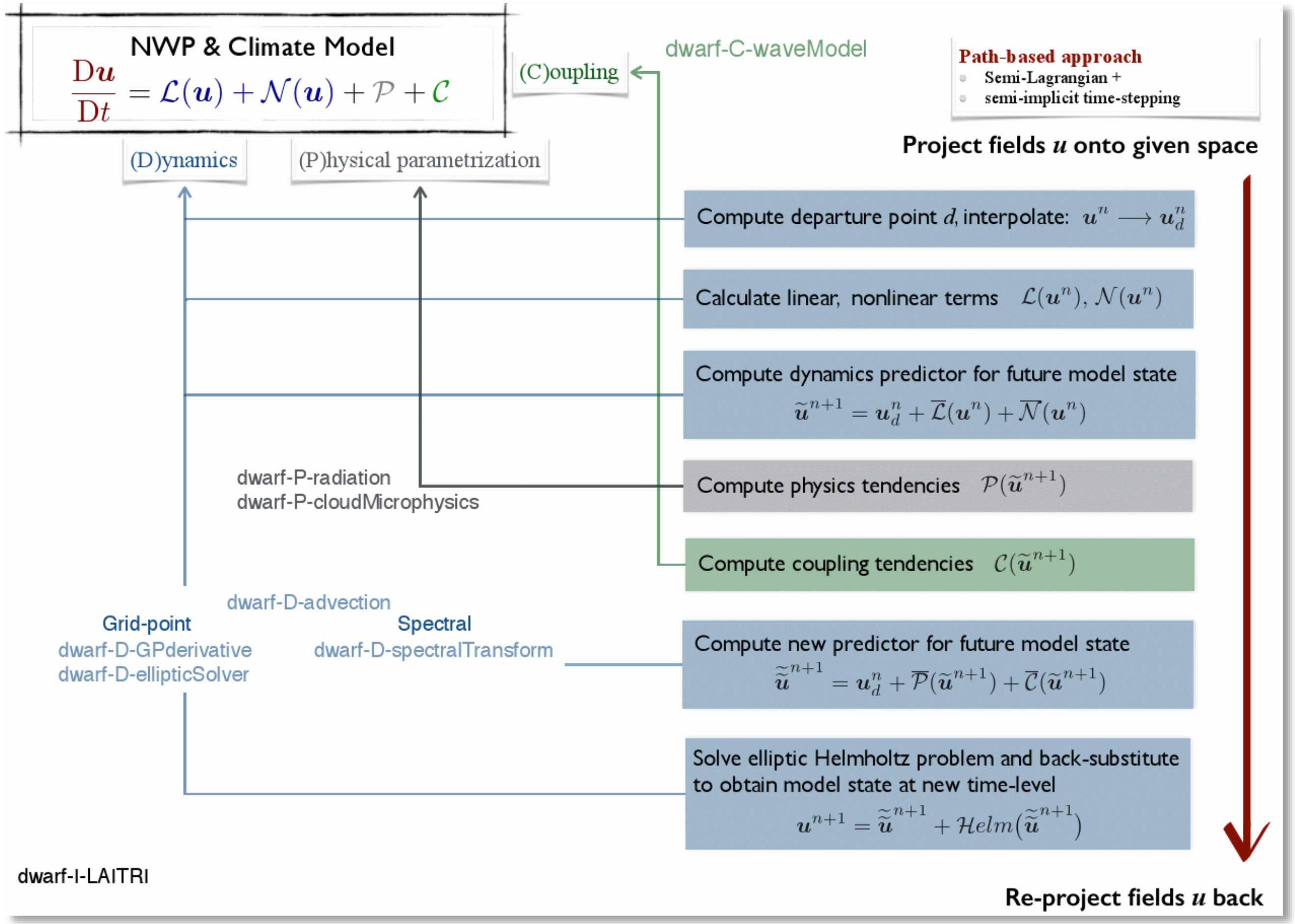
This project is funded by the European Union

Energy efficient **SC**alable **A**lgorithms for weather **P**rediction at **E**xascale

Separation of concerns:



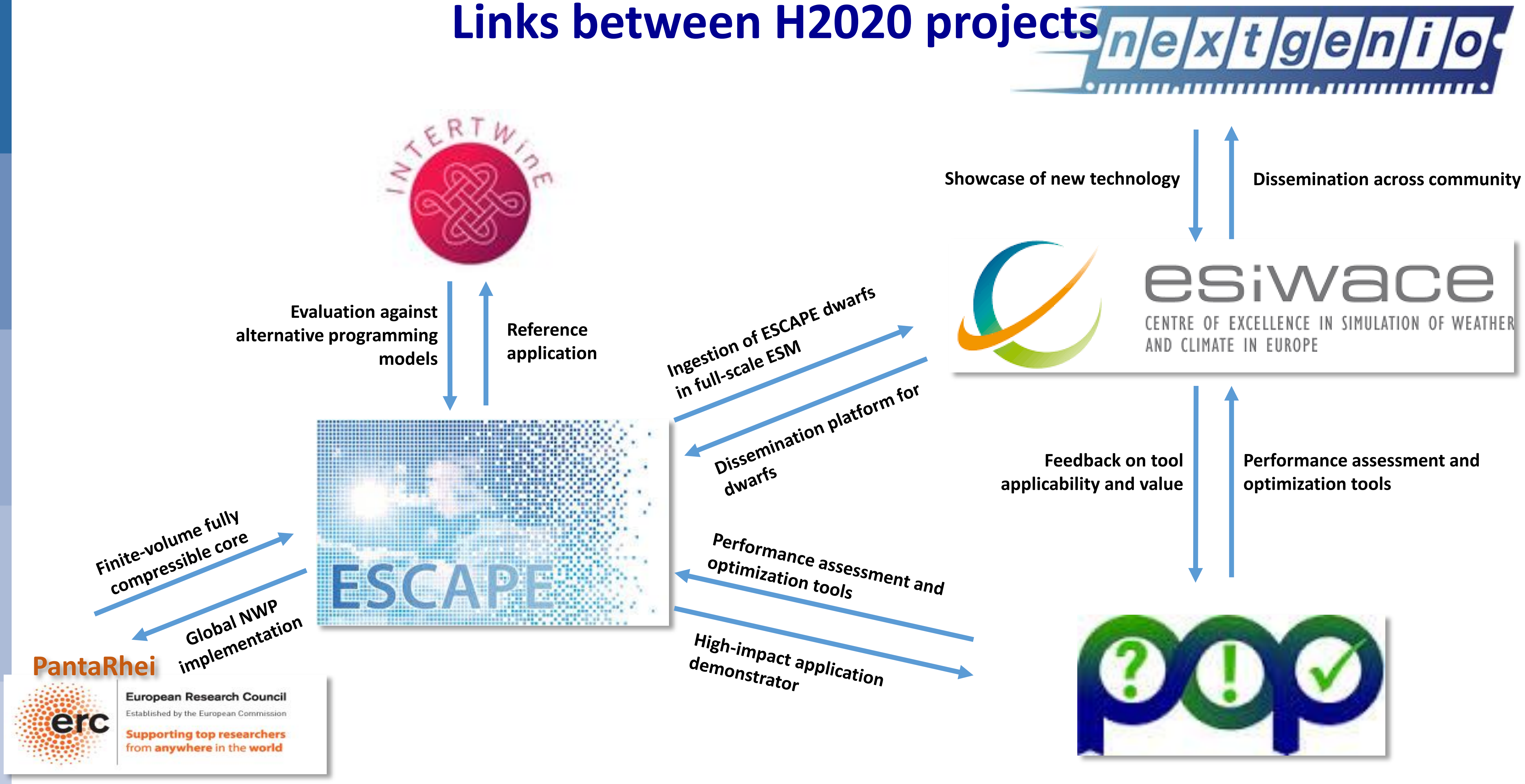
Workflow of model:



[G. Mengaldo]

ECMWF · MeteoSwiss · DMI · RMI · NVIDIA · METEO FRANCE Toujours un temps d'avance · ICHEC · Deutscher Wetterdienst Wetter und Klima aus einer Hand · Optalysys · Bull · PSNC · Loughborough University

Links between H2020 projects



Climate & weather prediction together

… European Flagship Programme On Extreme Computing and Climate. Drawing on existing climate modelling expertise in Europe and working closely with existing supercomputing centres, EPECC, would oversee the development of cloud-and-eddy-resolved global climate system models, and integration of these models into an extreme-scale computing technology platform …

Target:

1 km global coupled simulations,

1 year/day processing rate

= new ESiWACE demonstrator case!

Login | About | Legal notice | Cookies | Contact on Europa | Search on Europa

European Commission

FUTURIUM

Digital4Science

European Commission > Futurium > Digital4Science > Discussions > A Flagship European Programme on Extreme Computing and Climate

About | Discussions | Library | Blogs | Polls | Events

A Flagship European Programme on Extreme Computing and Climate

Contribution received to the FET Flagships consultation: A Flagship European Programme on Extreme Computing and Climate

Author(s): Tim Palmer + 20 Leading European Climate Scientists

You can add your comments on this topic at: ***https://ec.europa.eu/futurium/en/content/flagship-european-programme-ext…***

At the 2015 Paris Climate Conference, leaders from 194 countries of the world unanimously acknowledged the serious threat posed by anthropogenic emissions of greenhouse gases. Society must now become resilient to changes in climate over coming decades. Most importantly, this will require quantitative estimates of the changing character of climate extremes. Such extremes include not only exceptional weather events such as violent wind storms and flash floods, but also persistent anomalies in planetary- scale circulation patterns, which lead to pervasive flooding in some regions and seasons, and long--lived drought and extremes of heat in others. However, providing such information will require a step change in the quality of global climate models, which at present are simply not adequate for this purpose (1). In particular, the resolution of these models must increase to a level which allows both ocean eddies and individual cloud systems to become represented explicitly; this is the best hope for obviating long standing climate model biases.

Such capability can begin to be realised with the advent of exascale supercomputing (1018 floating point operations per second) anticipated around 2023. However, being able to realise exaFLOPS for practical applications will require considerable investment in climate modelling software. Indeed, such is the complexity of current climate system models, such an endeavour will require a pooling of expertise and resources from existing climate institutes. Here we argue for a European Flagship Programme On Extreme Computing and Climate. Drawing on existing climate modelling expertise in Europe and working closely with existing supercomputing centres, EPECC (pronounced "Epic") would oversee the development of cloud-- and eddy--resolved global climate system models, and integration of these models into an extreme--scale computing technology platform. Over the term of the programme, EPECC deliverables would be tailor--made for simulating climate extremes accurately on exaFLOP compute infrastructure. The key scientific focus for EPECC will be the drivers of changing extremes in Europe over the 21st century. EPECC will work in cooperation with national meteorological and oceanographic agencies, to provide information tailored to the needs of their clients in the public and private sector, with the aim of increasing European resilience to a

Join here

to take part

Search Digital4Science

Related tags

FET Flagships · FET Flagships consultation · FET · consultation · Future and Emerging Technologies · emerging technologies

What have we achieved?

Data pre-processing: Redesign of workflows for operations and research

Data assimilation: 3D-Var FGAT and simplified 4D-Var with OOPS

Model development: ESCAPE dwarfs concept established; separation of concern with Atlas (data structure framework) & GridTools; trials with GPU and Xeon Phi

Data post-processing: Broker-worker logic for product generation demonstrated; MultiIO I/O layer using NVRAM

Programming models: Single precision for ensembles; testing of GPI vs MPI vs Fortran co-array; DSL

Computer architectures: GPU cluster; Intel KNL rack; partner in H2020 co-design projects

Collaboration: ECMWF members states; vendors; weather – climate

Some final thoughts

- Do we need to reverse engineer our applications, say a single ensemble member?
  - *Example*: given that you want to avoid communication across nodes, what is the optimal model configuration on a (fat) node in terms of grid points, levels, model complexity?
    .... is this co-design?

- How do we replace work flow components in an operational setting that changes all the time?

- How will we do benchmarking in the next 5-10 years?
  - Now, we need to build-in flexibility and efficiency, while in the future we may be able to focus mostly on efficiency;
  - Therefore, we may need to produce a range of benchmarks with options (for example dynamical cores, advection schemes, physics schemes, DSL options etc.).