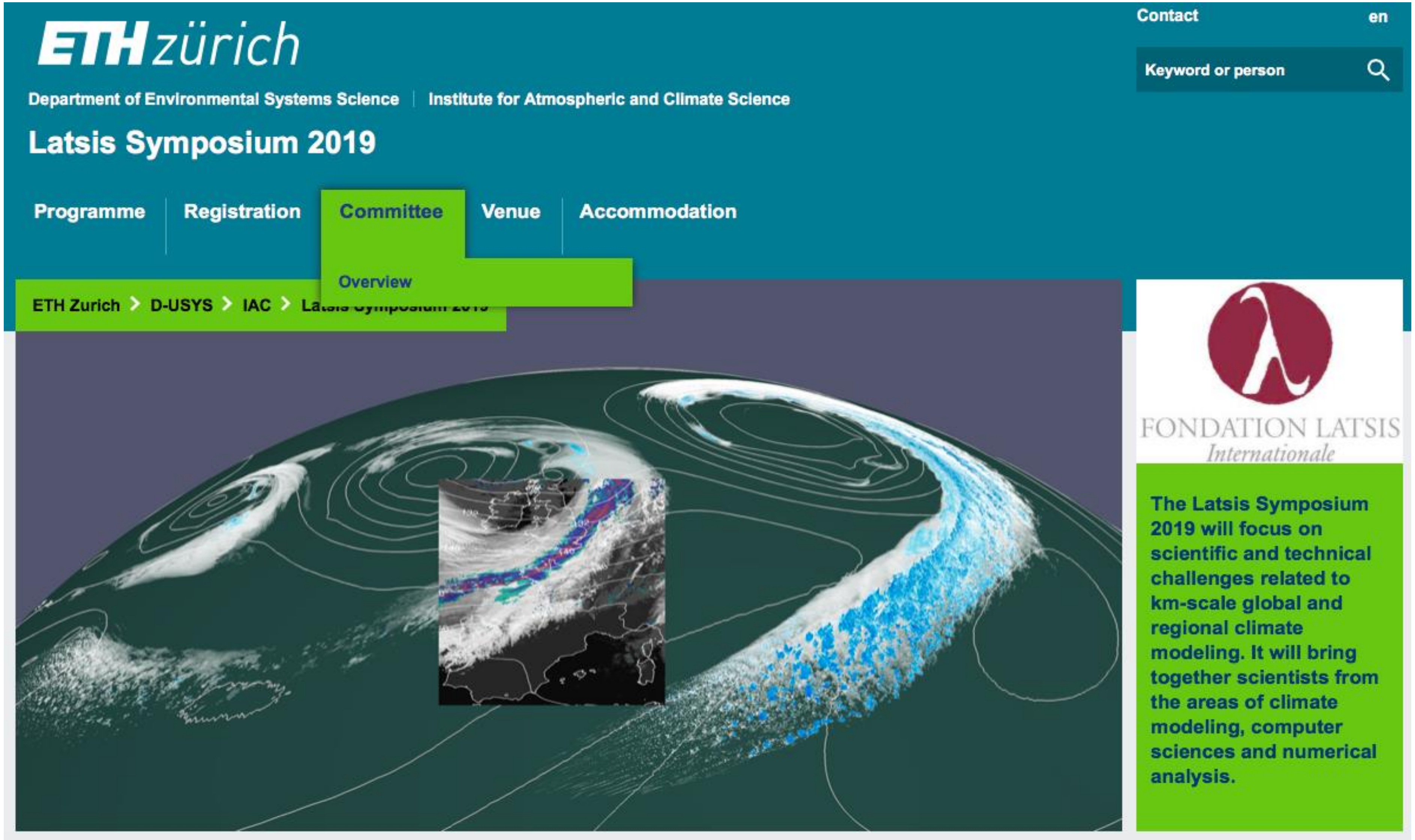
ECMWF's Scalability Programme

Peter Bauer

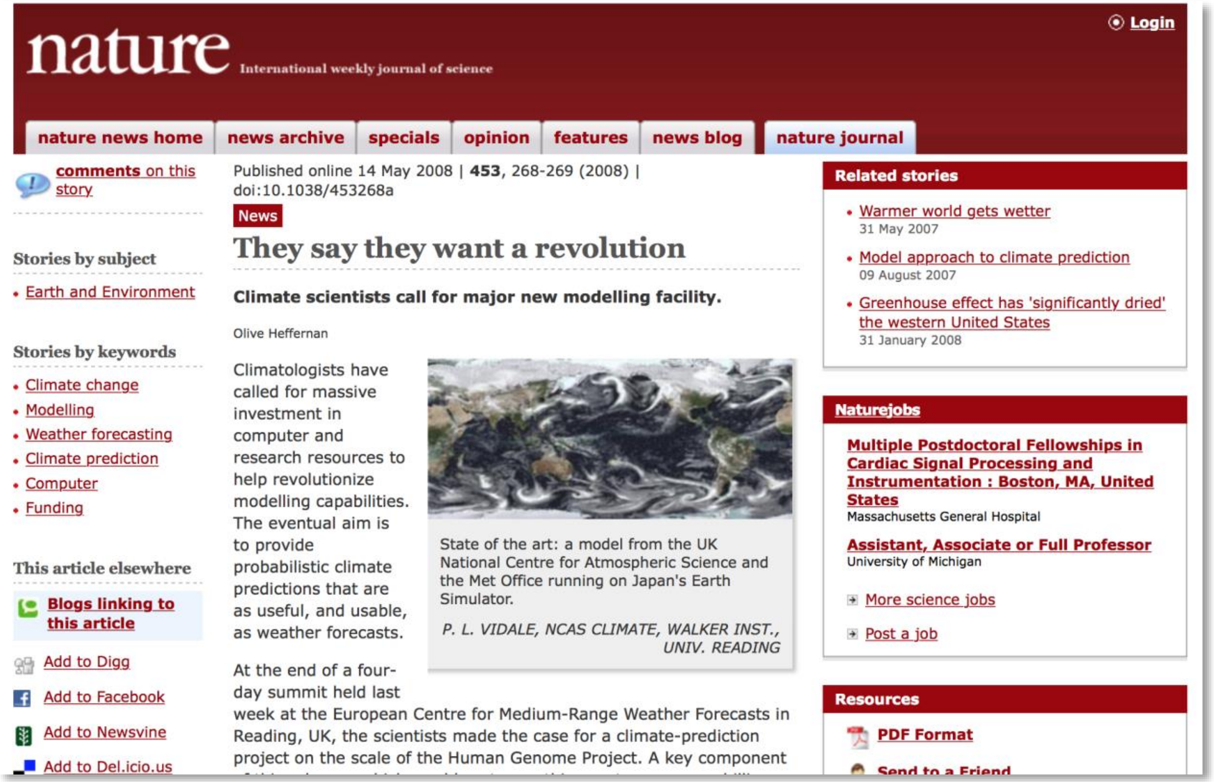
(This is a real team effort between many people at ECMWF and other international partners - and funding by the European Commission)

Overcoming key sources of model error



Targeting high resolution modelling: Athena

World Modeling Summit 2008



nature International weekly journal of science

Login

nature news home | news archive | specials | opinion | features | news blog | nature journal

comments on this story

Stories by subject

- Earth and Environment

Stories by keywords

- Climate change
- Modelling
- Weather forecasting
- Climate prediction
- Computer
- Funding

This article elsewhere

Blogs linking to this article

Add to Digg

Add to Facebook

Add to Newsvine

Add to Del.icio.us

Published online 14 May 2008 | **453**, 268-269 (2008) | doi:10.1038/453268a

News

They say they want a revolution

Climate scientists call for major new modelling facility.

Olive Heffernan

Climatologists have called for massive investment in computer and research resources to help revolutionize modelling capabilities. The eventual aim is to provide probabilistic climate predictions that are as useful, and usable, as weather forecasts.

State of the art: a model from the UK National Centre for Atmospheric Science and the Met Office running on Japan's Earth Simulator.

P. L. VIDALE, NCAS CLIMATE, WALKER INST., UNIV. READING

At the end of a four-day summit held last week at the European Centre for Medium-Range Weather Forecasts in Reading, UK, the scientists made the case for a climate-prediction project on the scale of the Human Genome Project. A key component

Related stories

- Warmer world gets wetter
  31 May 2007
- Model approach to climate prediction
  09 August 2007
- Greenhouse effect has 'significantly dried' the western United States
  31 January 2008

Naturejobs

Multiple Postdoctoral Fellowships in Cardiac Signal Processing and Instrumentation : Boston, MA, United States
Massachusetts General Hospital

Assistant, Associate or Full Professor
University of Michigan

More science jobs

Post a job

Resources

PDF Format

Send to a Friend

Cray XT4 called "Athena"

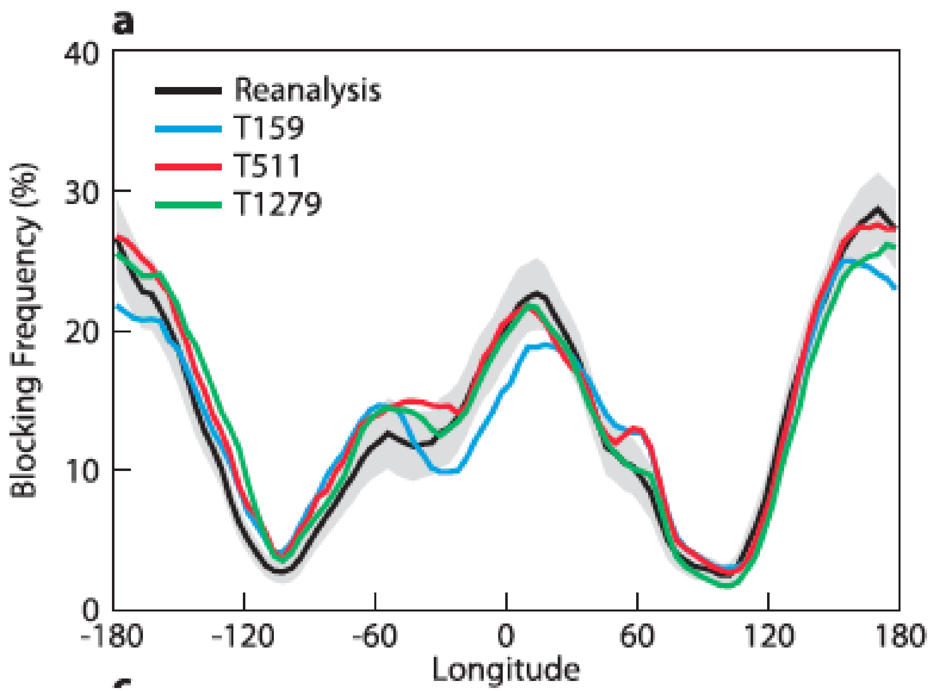
- National Institute for Computation Studies (NICS)
- ≈20.000 CPUs
- #30 on Top500 list (Nov 2009)

Key figures

- Dedicated access for 6 months from 10/2009–03/2010
- Technical support from NICS staff
- A total of 72 ▪ 10^6 CPUh
- Utilization above 95% of full capacity
- A total of ≈1.2 PB of data (≈ 1/3 of the entire CMIP5 archive)

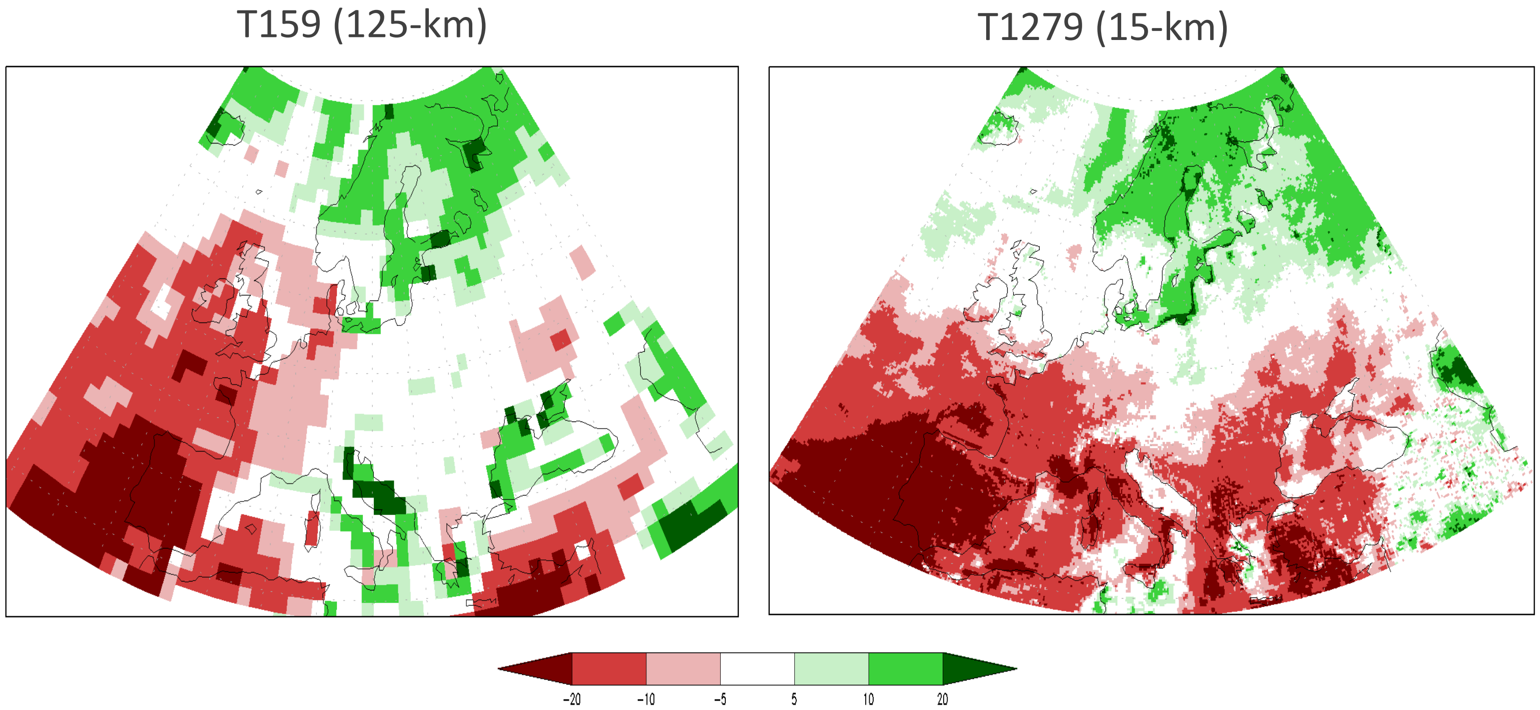
Targeting high resolution modelling: Athena

Blocking



Jung et al. (2012)

Mean temperature change



Kinter et al. (2013)

What is the ultimate target?

Conservation of momentum, energy, mass and moisture:

$$\frac{\partial \vec{V}}{\partial t} = -(\vec{V}\cdot\nabla)\vec{V} - \frac{1}{\rho}\nabla p - \vec{g} - 2\vec{\Omega}\times\vec{V} + \nabla\cdot(k_\omega \nabla\vec{V}) + \vec{F}_d$$

$$\rho c_p \frac{\partial T}{\partial t} = -\rho c_p(\vec{V}\cdot\nabla)T - \nabla\cdot\vec{R} + \nabla\cdot(k_\tau \nabla T) + C + S$$

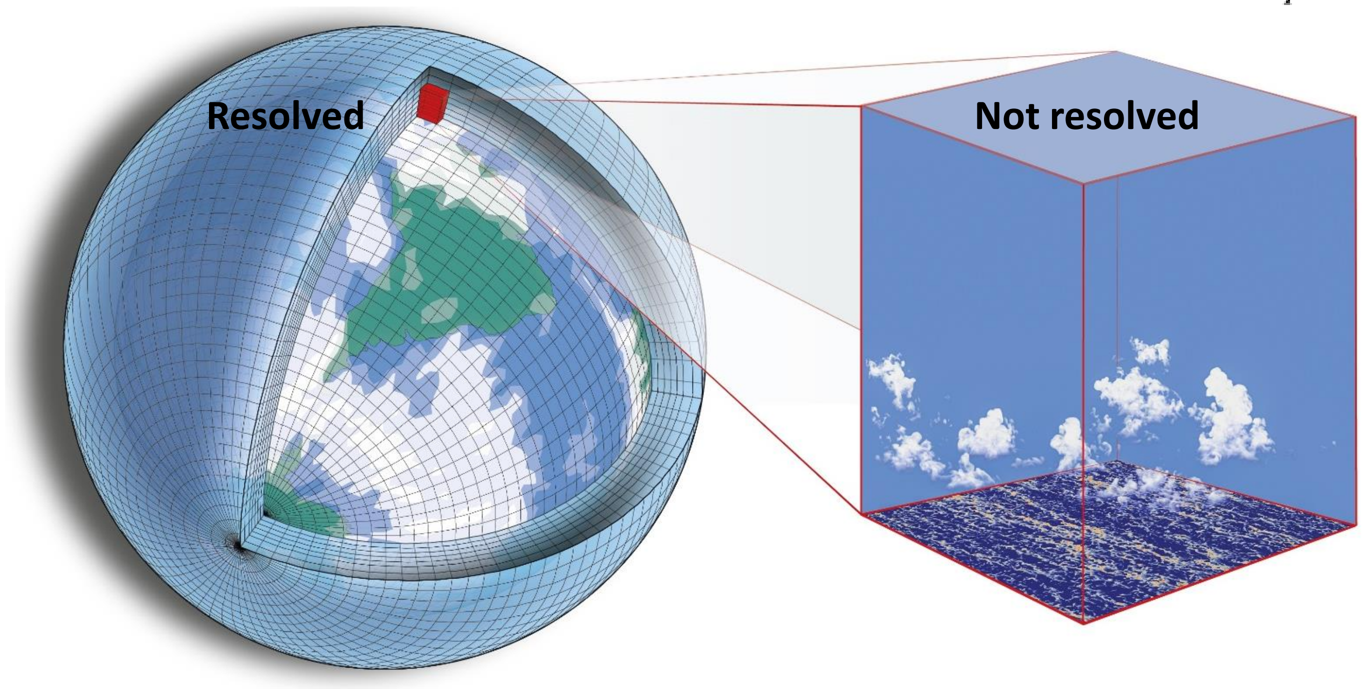
$$\frac{\partial \rho}{\partial t} = -(\vec{V}\cdot\nabla)\rho - \rho(\nabla\cdot\vec{V})$$

$$\frac{\partial q}{\partial t} = -(\vec{V}\cdot\nabla)q + \nabla\cdot(k_q\nabla q) + S_q + E$$

Equation of state:

$$p = \rho R_d T$$

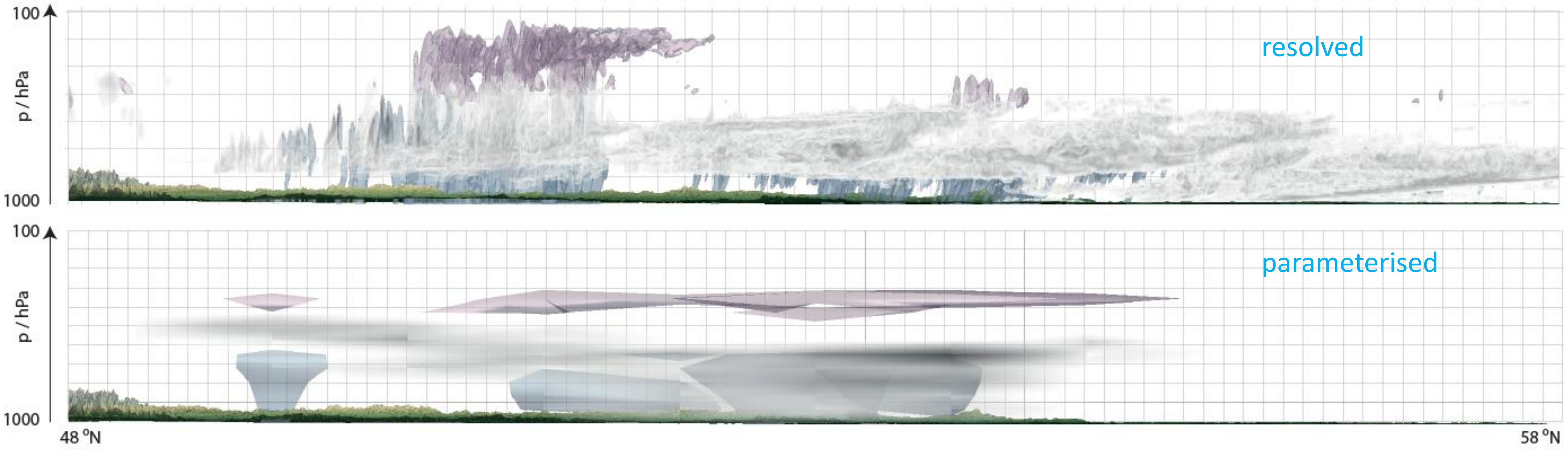
V = velocity
T = temperature
p = pressure
ρ = density
q = specific humidity
g = gravity
Ω = rotation of Earth
F_d = drag force of Earth
R = radiation vector
C = conductive heating
c_p = heat capacity, constant p
E = evaporation
S = latent heating
S_q = phase change source
k = diffusion coefficients
R_d = dry air gas constant



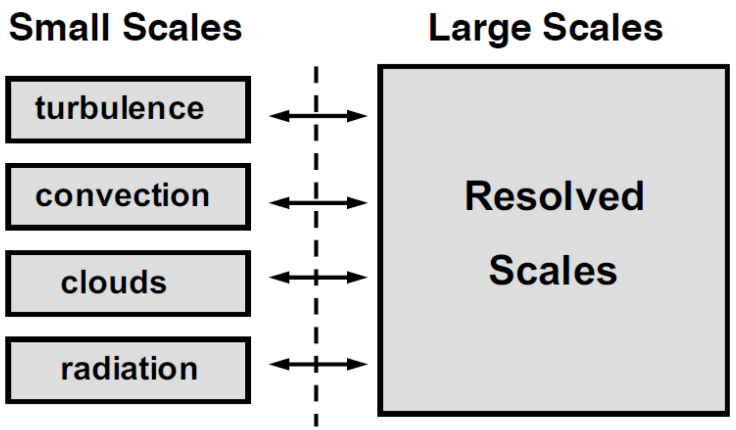
… from **parameterizations** for radiation, cloud, convection, turbulence, waves…

PETER BAUER 2019

What is the ultimate target?

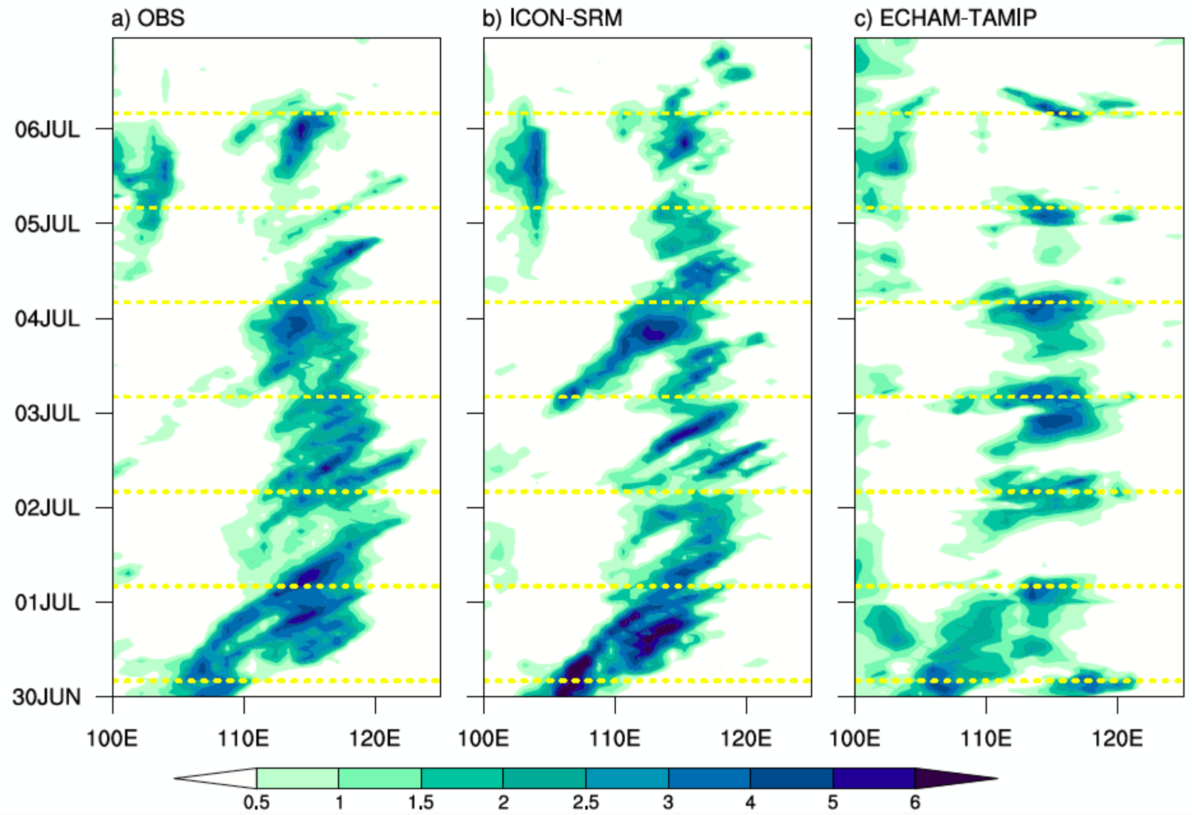


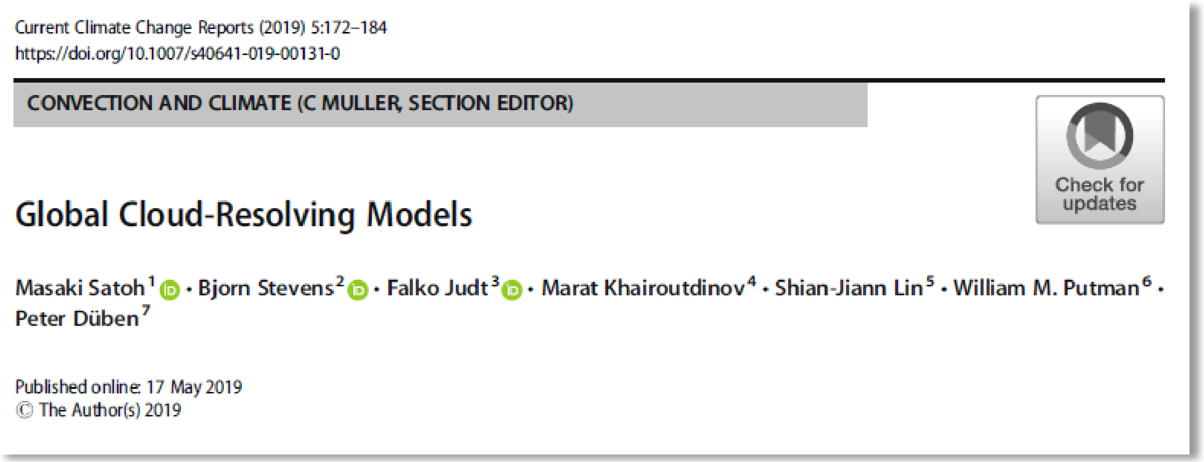
They are not the same:



[Courtesy Bjorn Stevens]

What is the ultimate target?




Current Climate Change Reports (2019) 5:172–184
https://doi.org/10.1007/s40641-019-00131-0

CONVECTION AND CLIMATE (C MULLER, SECTION EDITOR)

Check for updates

Global Cloud-Resolving Models

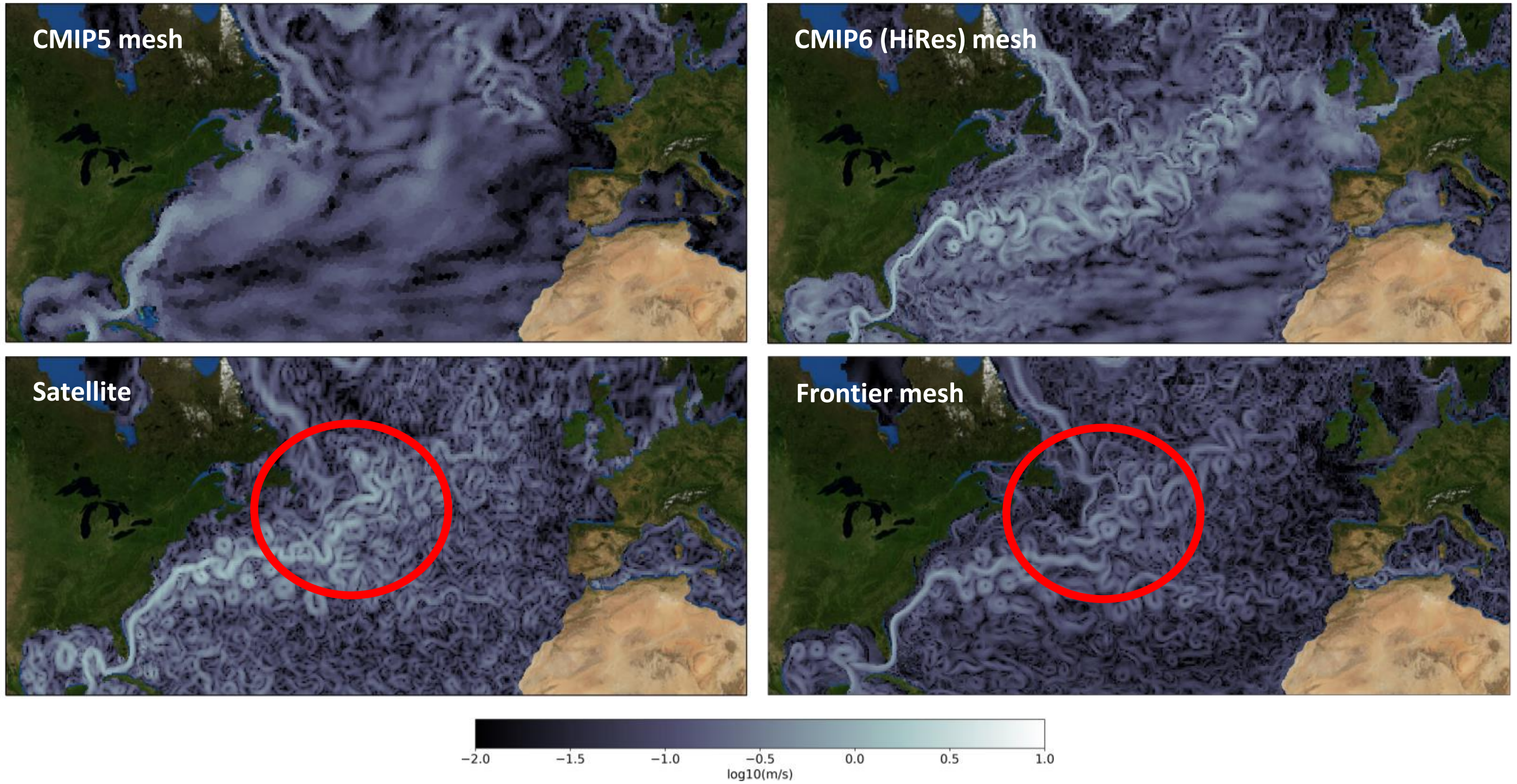
Masaki Satoh[1] · Bjorn Stevens[2] · Falko Judt[3] · Marat Khairoutdinov[4] · Shian-Jiann Lin[5] · William M. Putman[6] · Peter Düben[7]

Published online: 17 May 2019

- Representation of the global mesoscale
- Multi-scale scale interactions of convection
- Circulation-driven microphysical processes
- Turbulence and gravity waves
- Synergy with satellite observations
- Downscaling for impact studies
- Etc.

[Courtesy Bjorn Stevens]

What is the ultimate target?

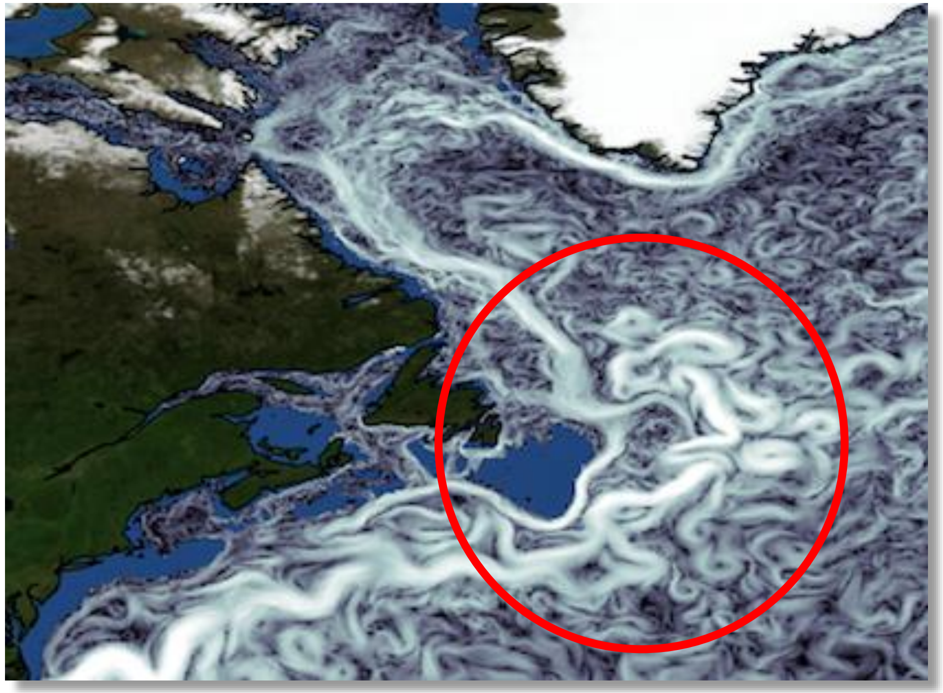


Displayed on a common 1/4º mesh

Surface current simulation with FESOM-2 ocean/sea-ice model on adaptive mesh refining resolution in coastal areas and towards the poles using the Rossby radius of deformation

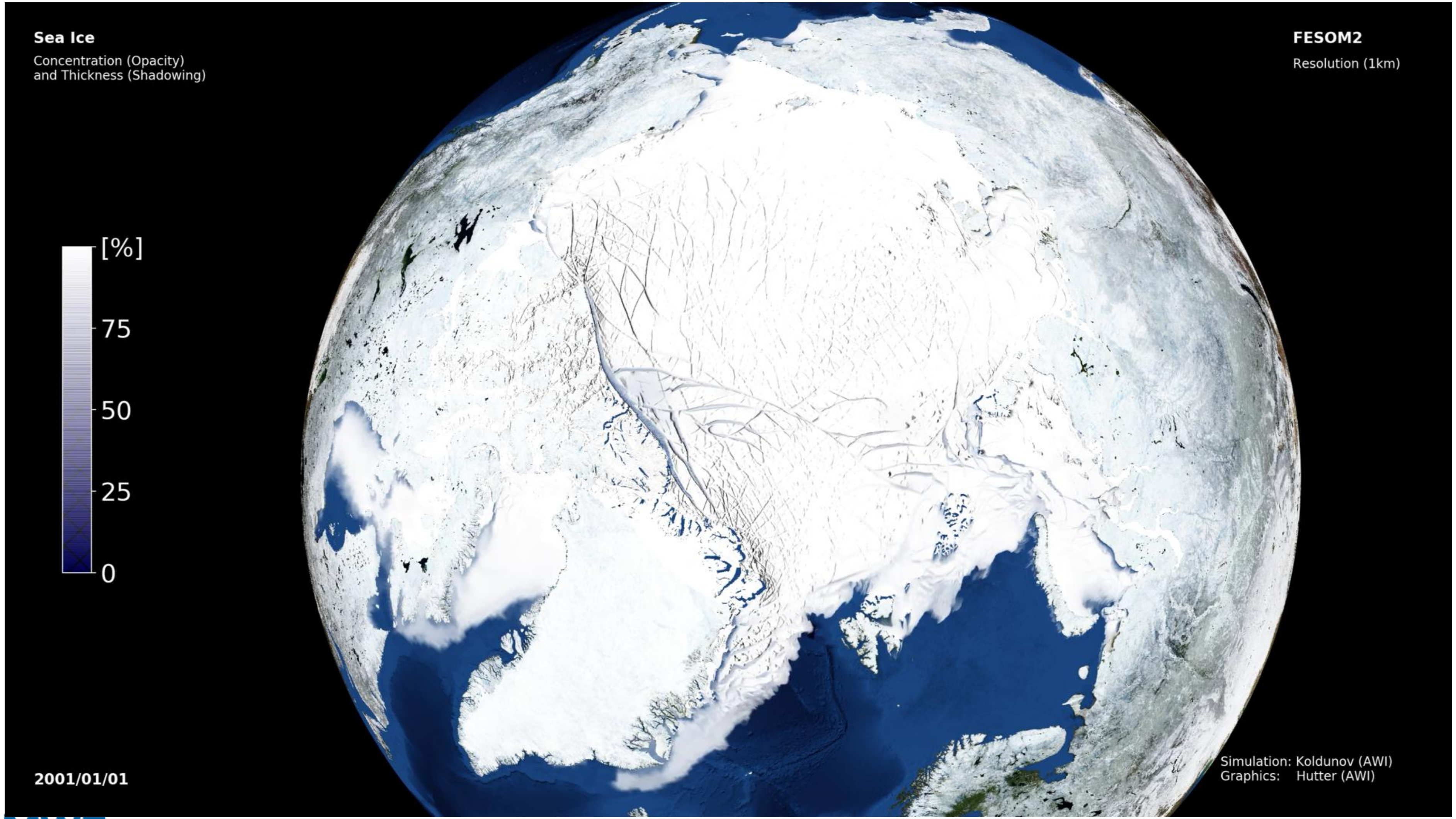
(Courtesy T Jung and S Danilov, AWI)

¼ Rossby radius of deformation

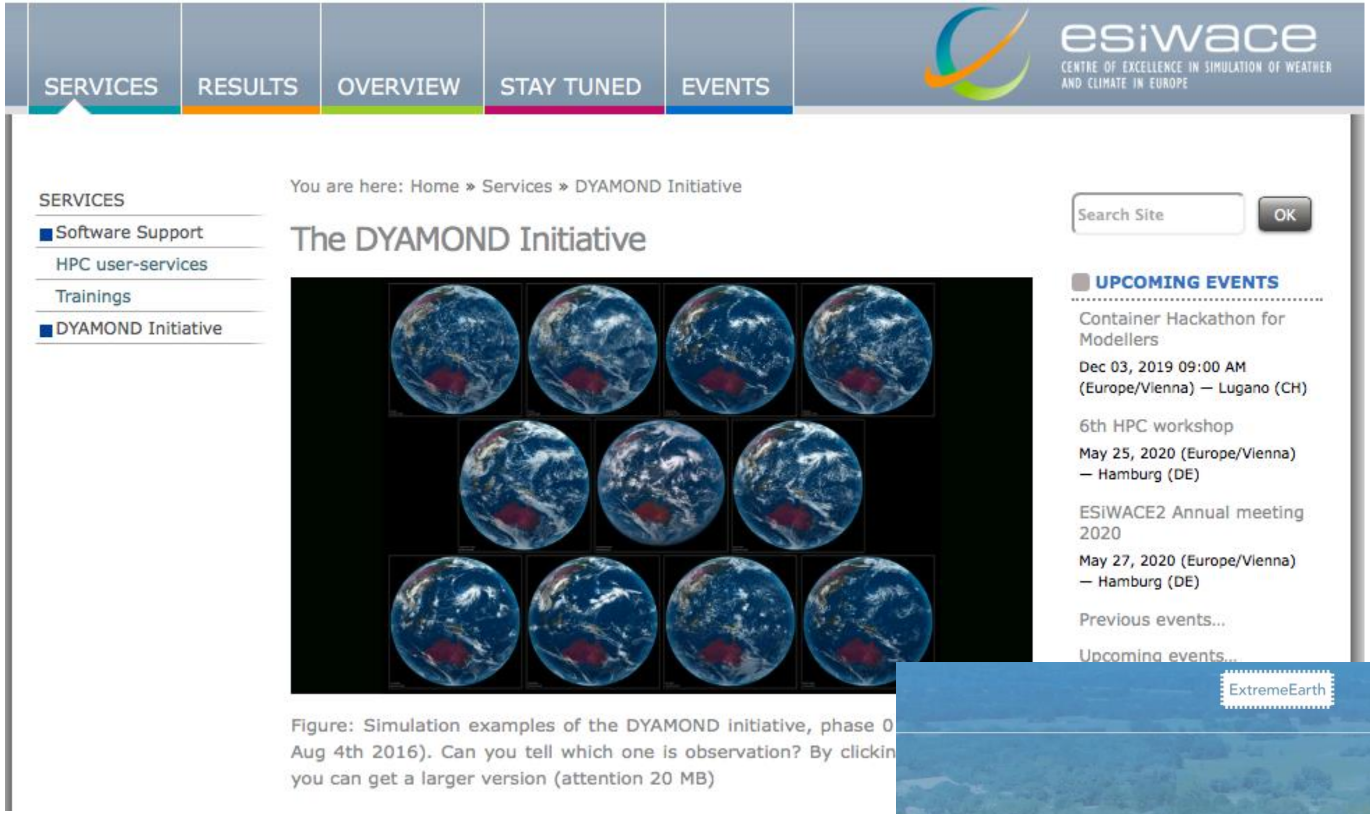


What is the ultimate target?

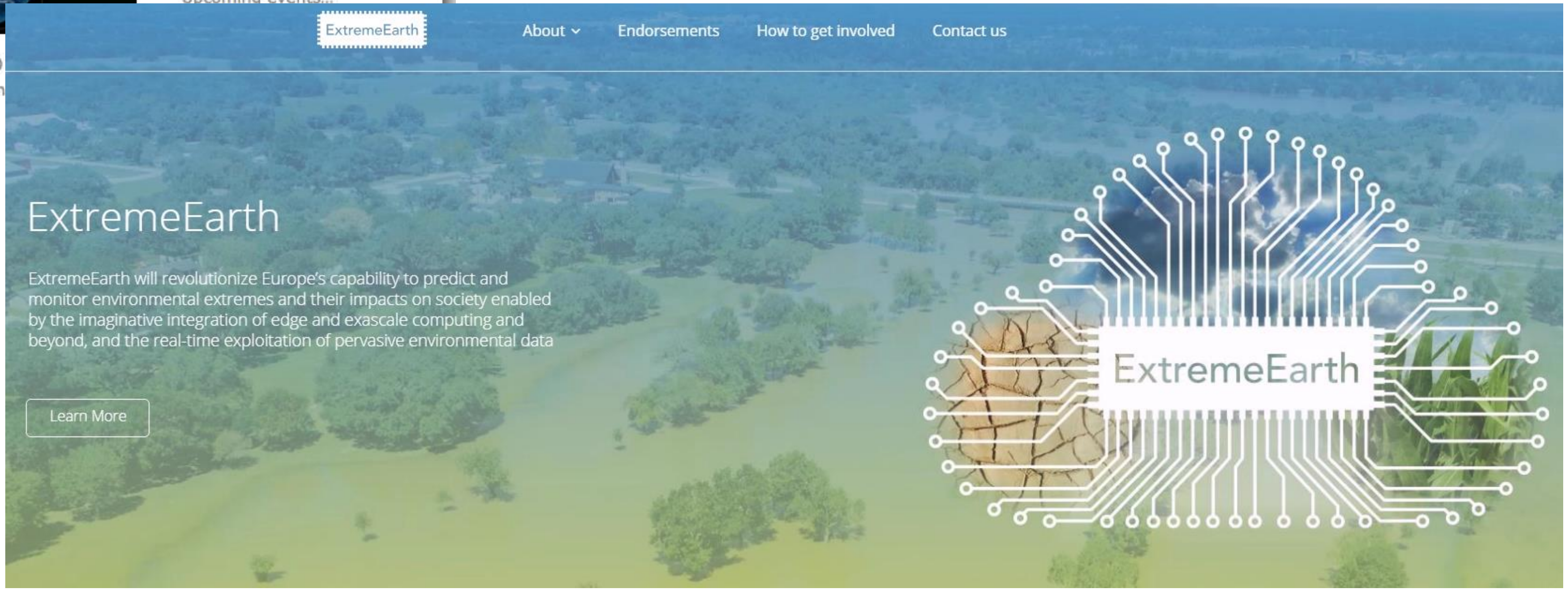
Sea-ice simulation with FESOM-2 ocean/sea-ice model (Courtesy T Jung and S Danilov, AWI)



1-km as a proxy for qualitatively different models



https://www.esiwace.eu/



https://www.extremeearth.eu/

ECMWF Scalability Programme – Present capability @ 1.45km

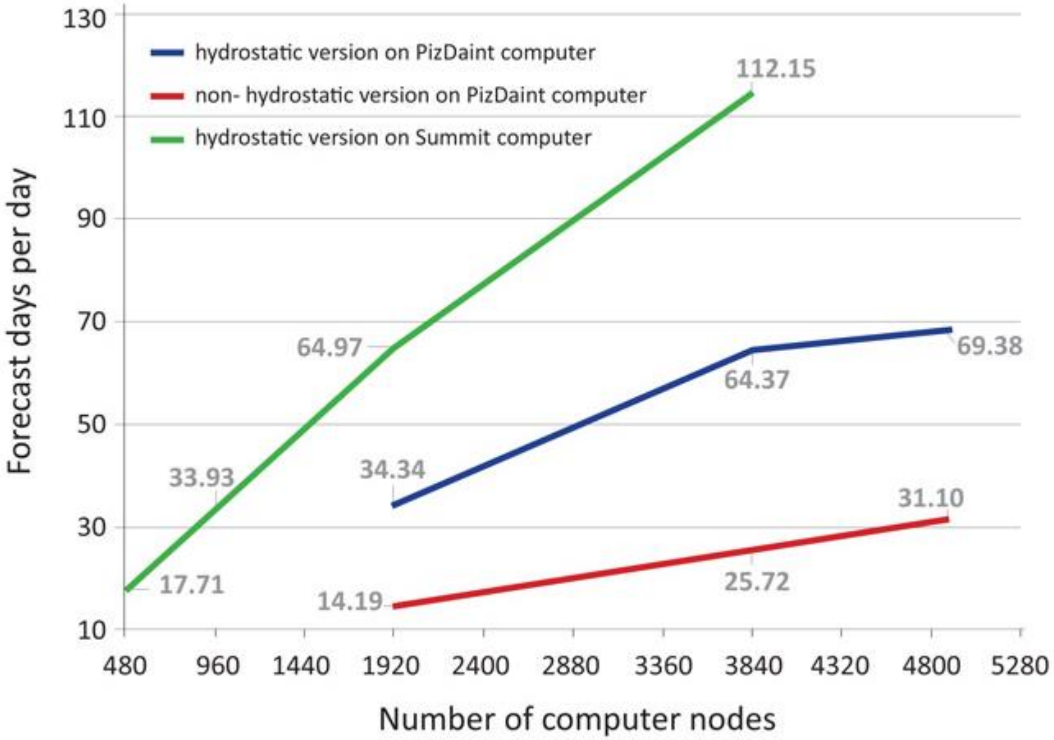
U.S. DEPARTMENT OF ENERGY
INCITE
LEADERSHIP COMPUTING

Type: New
Title: "Unprecedented scales with ECMWF's medium-range weather prediction model"

Principal Investigator: Nils Wedi, European Center for Medium-Range Weather Forecasts
Co-Investigators: Peter Bauer, European Center for Medium-Range Weather Forecasts

Scientific Discipline: Earth Science: Climate Research

INCITE Allocation:
Site: Oak Ridge National Laboratory
Machine (Allocation): IBM AC922 (102,000 node-hours)



→ O(3-10) too slow (atmosphere only, no I/O)

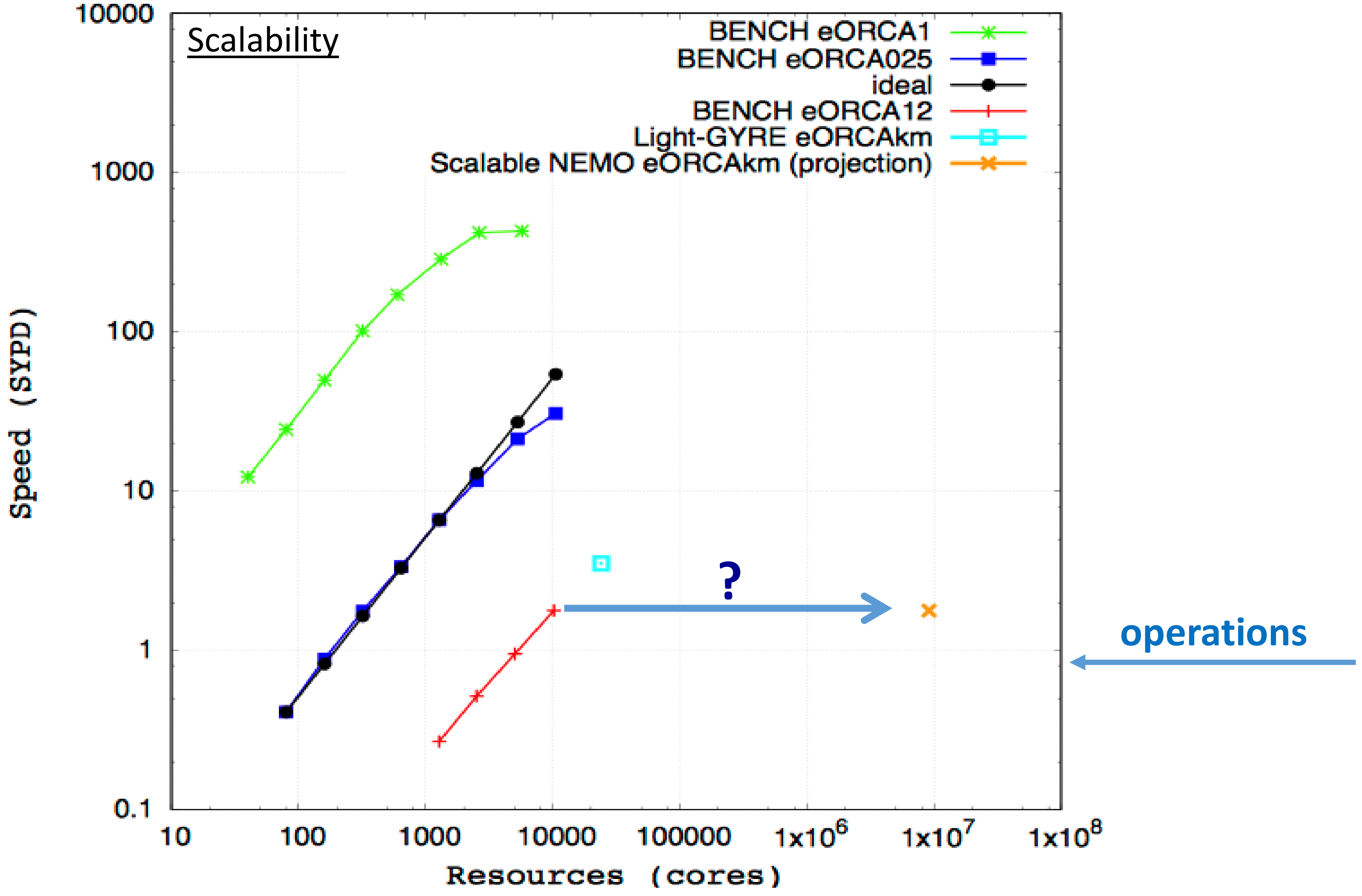
	Near-global COSMO[15]		Global IFS[16]	
	Value	Shortfall	Value	Shortfall
Horizontal resolution	0.93 km (non-uniform)	0.81x	1.25 km	1.56x
Vertical resolution	60 levels (surface to 25 km)	3x	62 levels (surface to 40 km)	3x
Time resolution	6 s (split-explicit with substepping)*	-	120 s (semi-implicit)	4x
Coupled	No	1.2x	No	1.2x
Atmosphere	Non-hydrostatic	-	Non-hydrostatic	-
Precision	Single	-	Single	-
Compute rate	0.043 SYPD	23x	0.088 SYPD	11x
Other (e.g. physics, ...)	microphysics	1.5x	Full physics	-
Total shortfall		101x		247x

→ O(100-250) too slow (still no I/O)

→ O(1000) incl. everything (ensembles, Earth system, etc.)

[Schulthess et al. 2019, Computing in Science & Engineering]

Present capability @ 1km: NEMO (ocean)



Météo-France Bullx Intel Broadwell processors

[Courtesy CERFACS, IPSL, BSC @ESiWACE]

But we don’t have to move to 1km to be worried

Computing:

https://www.ecmwf.int/sites/default/files/ECMWF_Strategy_2016-2025.pdf: “[…] An ambitious target that depends on scientific, computing and scalability advances is for this ensemble to have a horizontal resolution of about 5 km by 2025. […]”

$$\text{HPC cost growth} = \frac{Cost\ (50\ members,\ 5\ km,\ 200\ levels)}{Cost\ (50\ members,\ 18\ km,\ 137\ levels)} = O(100)$$

(more if we count significant ocean model upgrades and atmospheric composition)

Data:

***Public** access **per year**:*

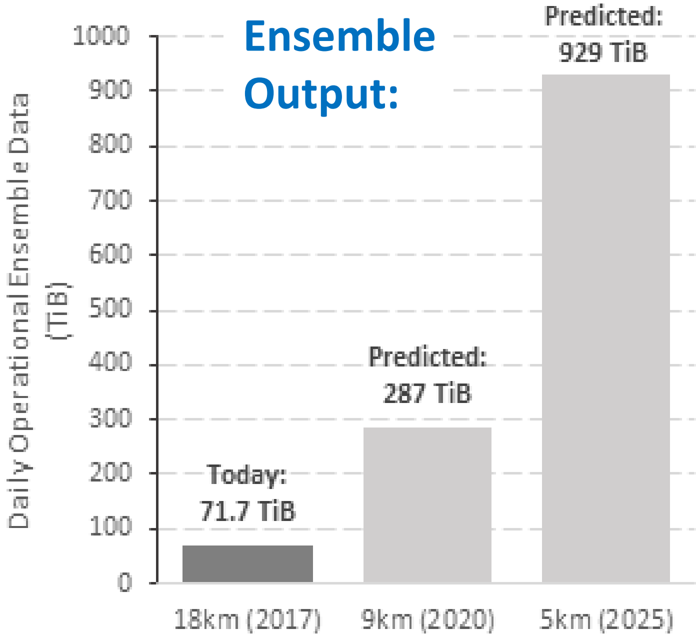
- 40 billions fields
- 20 PB retrieved
- 25,000 users

Total **activity (Member States and commercial customers)** **per day**:

- 450 TBytes retrieved
- 200 TBytes archived
- 1.5 million requests

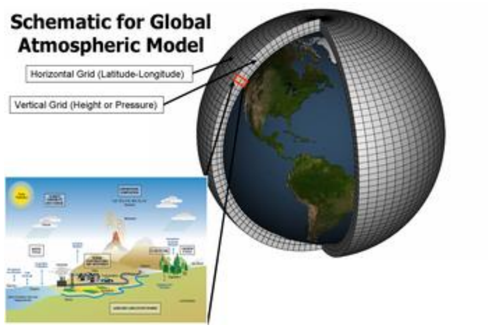
Total volume in MARS: 220 PiB

Ensemble Output:



[Courtesy T Quintino]

ECMWF Scalability Programme – Holistic approach



Data acquisition

- Lean workflow in critical path
- Object based data store
- Load balancing obs-mod
- Quality control and bias correction with ML

Data assimilation

- OOPS control layer
- Algorithms: 4DV, En4DV, 4DEnVar
- Models: IFS, NEMO, QG
- Coupling
- Surrogate models with ML

Forecast

- IFS-ST & IFS-FVM on same grid and with same physics
- Coupling
- Separation of concerns
- Surrogate models with ML

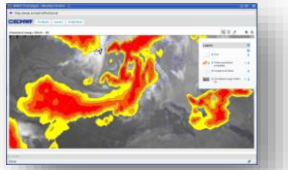
Product generation

- Lean workflow in critical path
- Object based data store
- Use deep memory hierarchy
- Broker-worker separation
- Integration in Cloud (EWC)
- Data analytics with ML

Dissemination → RMDCN, Internet

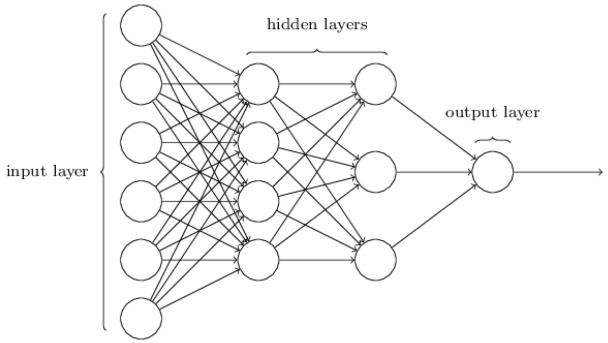
Web services → Internet

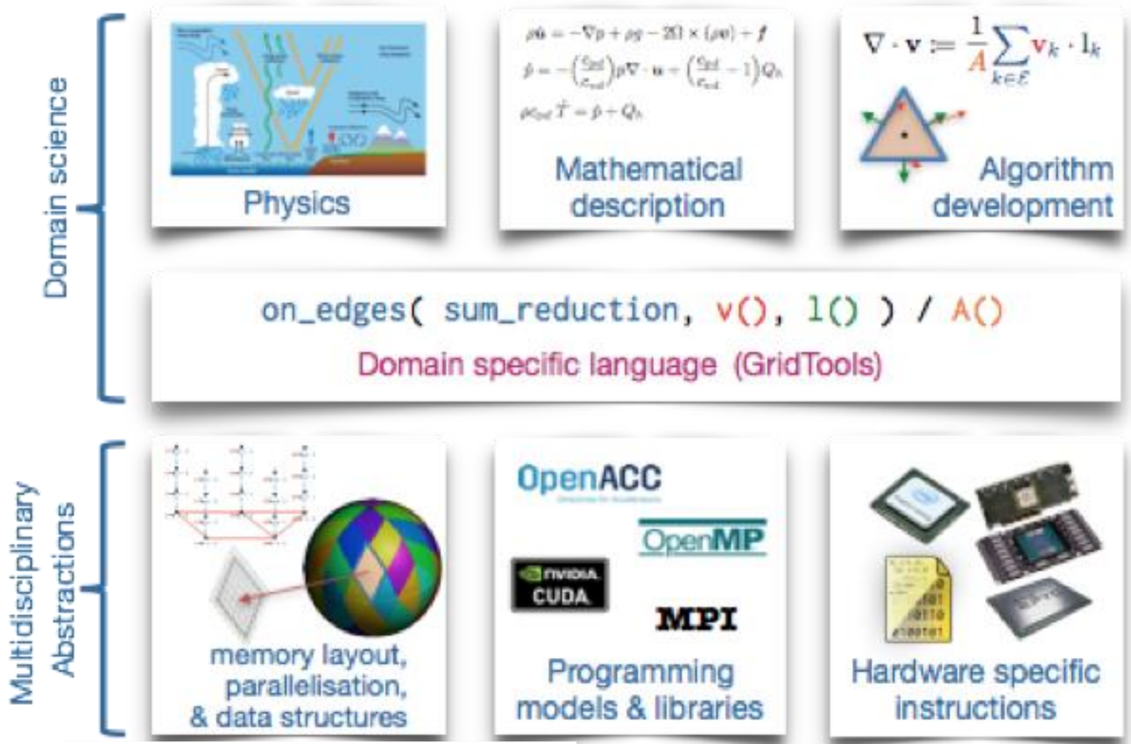
Archive → Data Handling System



ECMWF Scalability Programme – Ultimately, touch everything

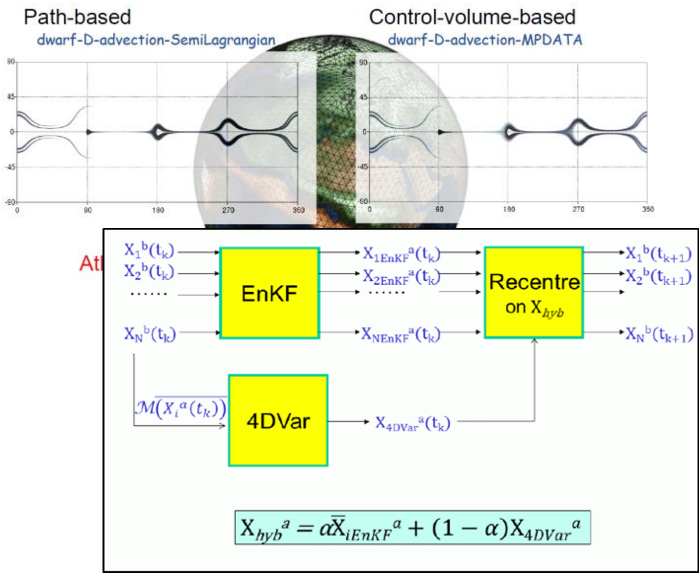
Neural networks



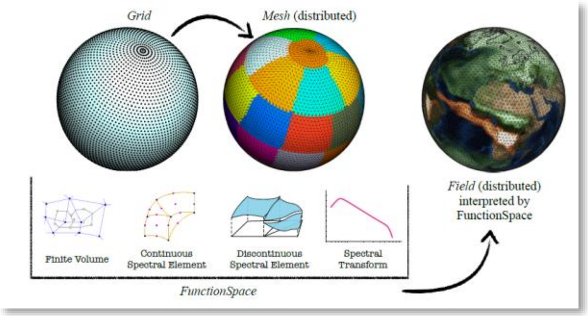


Mathematics&algorithms

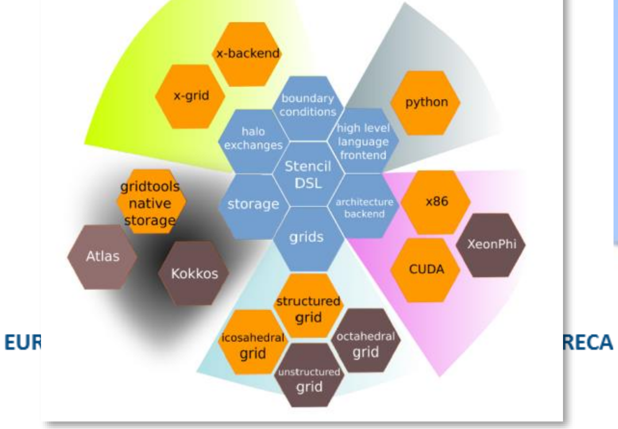
Rossby-Haurwitz test case after 7 days



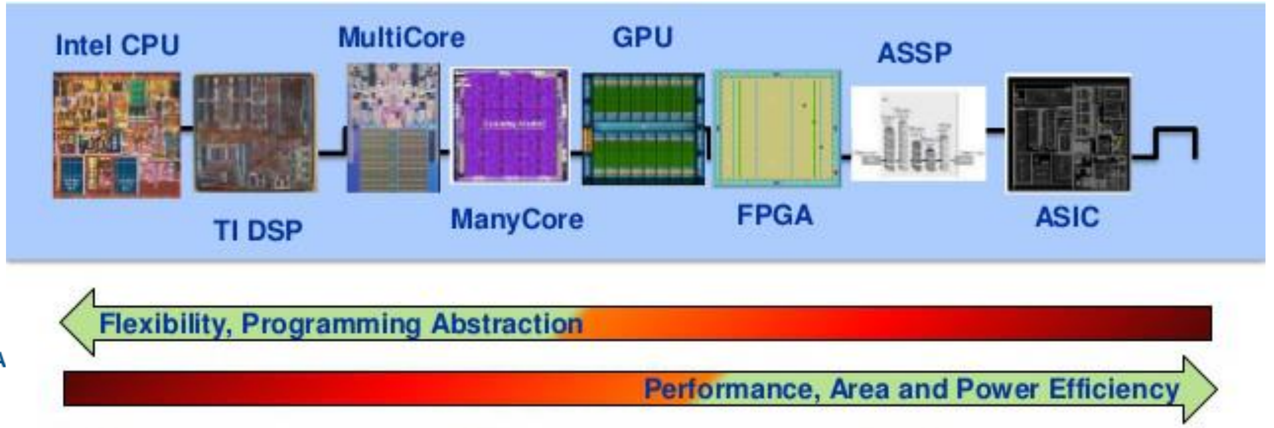
Data structures: Atlas

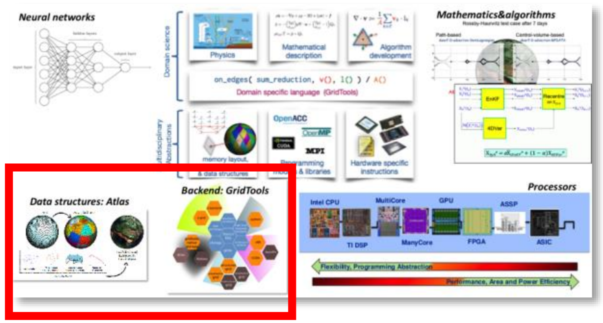


Back-end: GridTools



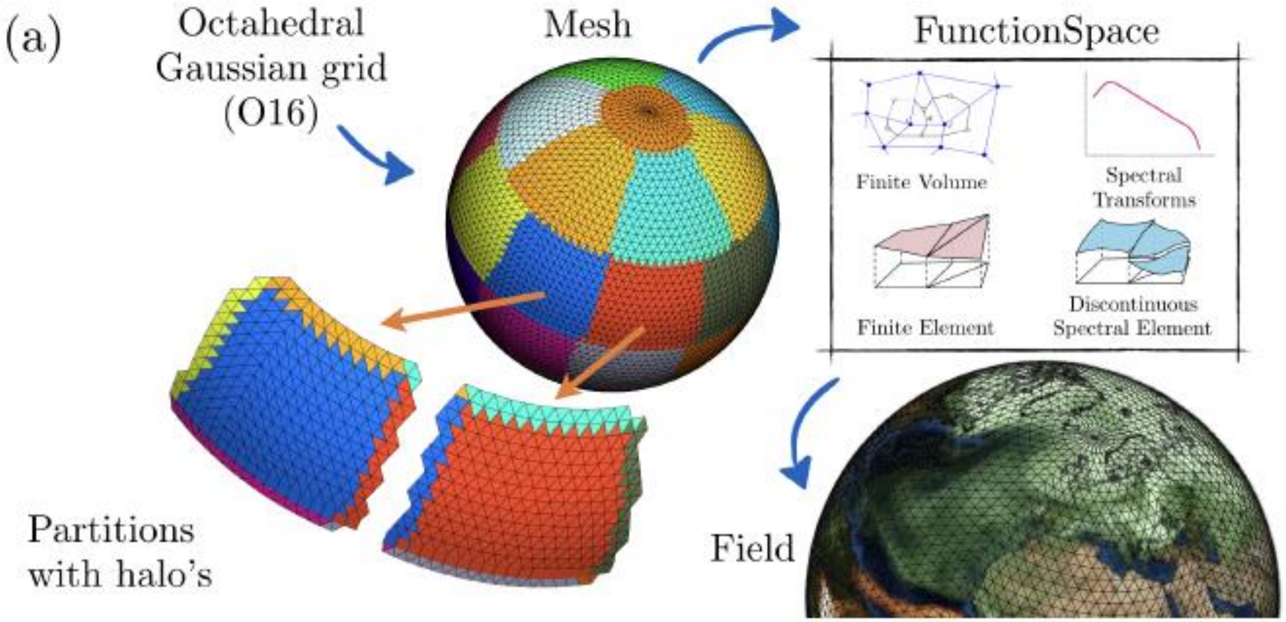
Processors





Generic data structure library *Atlas*

Funded by the European Union



Computer Physics Communications 220 (2017) 188–204

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

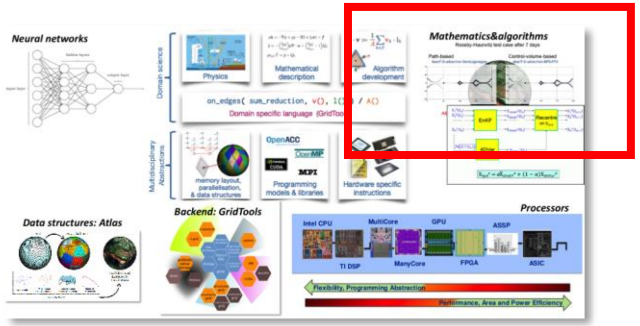
Atlas: A library for numerical weather prediction and climate modelling

Willem Deconinck *, Peter Bauer, Michail Diamantakis, Mats Hamrud, Christian Kühnlein, Pedro Maciel, Gianmarco Mengaldo, Tiago Quintino, Baudouin Raoult, Piotr K. Smolarkiewicz, Nils P. Wedi

European Centre for Medium-Range Weather Forecasts (ECMWF), Shinfield Park, Reading RG2 9AX, United Kingdom

ECMWF

[Courtesy W Deconinck]



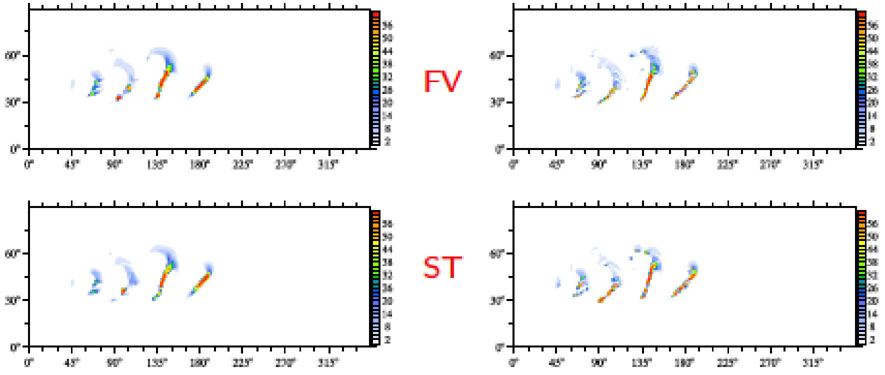
New IFS-FVM dynamical core

- finite-volume discretisation operating on a compact stencil
- deep-atmosphere non-hydrostatic fully compressible equations in generalised height-based vertical coordinate
- fully conservative and monotone advective transport
- flexible horizontal and vertical meshes
- robustness wrt steep slopes of orography
- Atlas built in

Moist baroclinic instability using IFS-FVM and IFS-ST with parametrization for large-scale condensation and diagnostic precipitation following Reed and Jablonowski 2011:

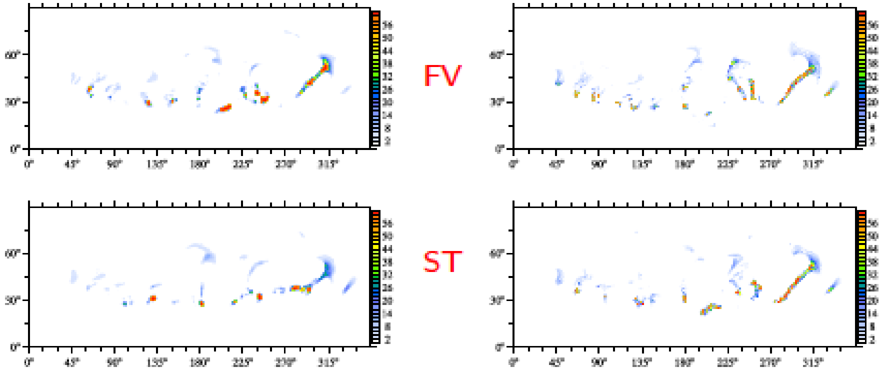
Precipitation (mm/day) at day 10

O160/TCo159, $\Delta_h \approx 62$ km   O640/TCo639, $\Delta_h \approx 18$ km

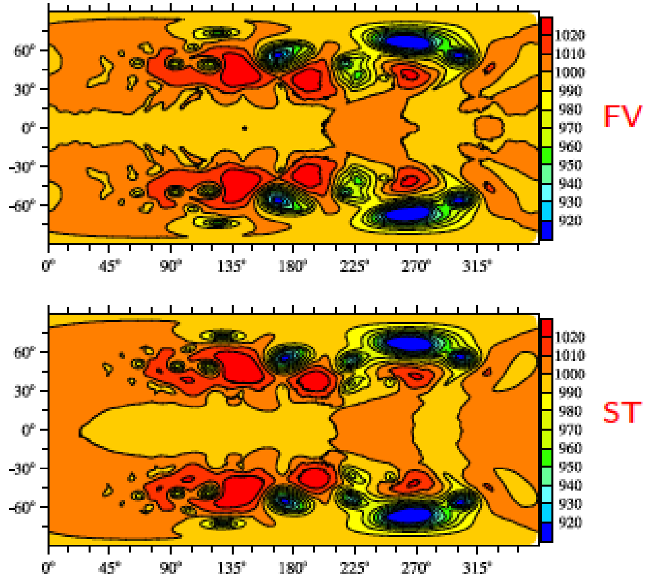


Precipitation (mm/day) at day 15

O160/TCo159, $\Delta_h \approx 62$ km   O640/TCo639, $\Delta_h \approx 18$ km



Surface pressure O640/TCo639, $\Delta_h \approx 18$ km, day 15

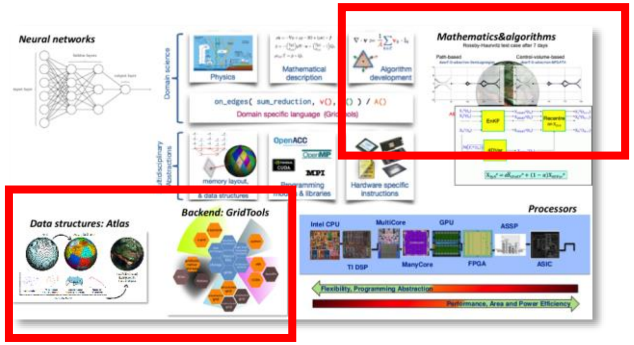


- Finite-volume solutions can achieve accuracy of established spectral-transform IFS for moist flows

[Kühnlein et al. 2019, Geoscientific Model Development]

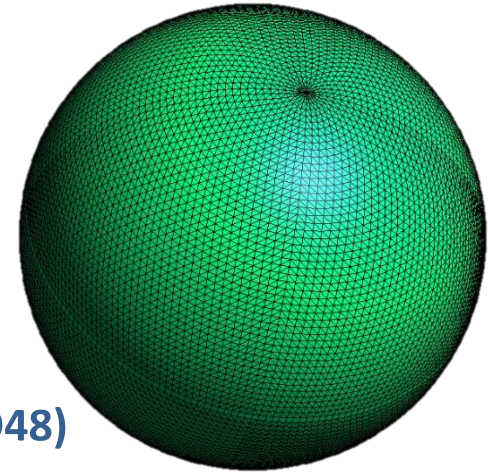
ECMWF

[Courtesy C Kühnlein, P Smolarkiewicz, N Wedi]

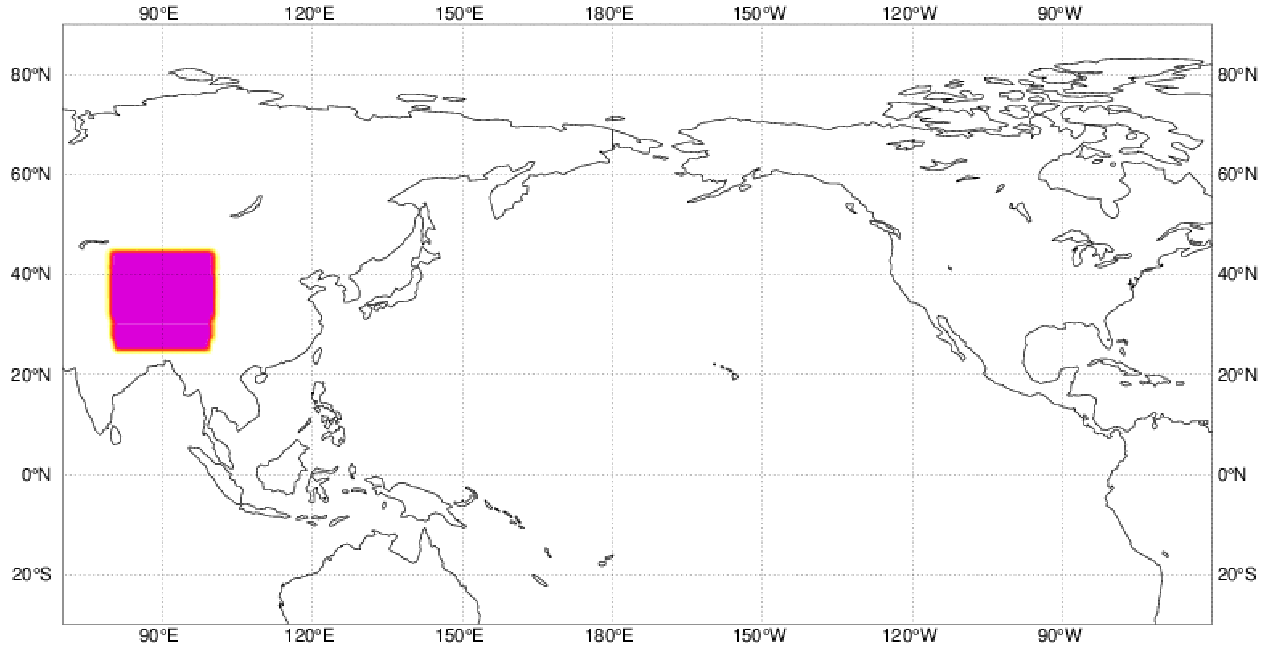


IFS-ST vs IFS-FVM advection using *Atlas*

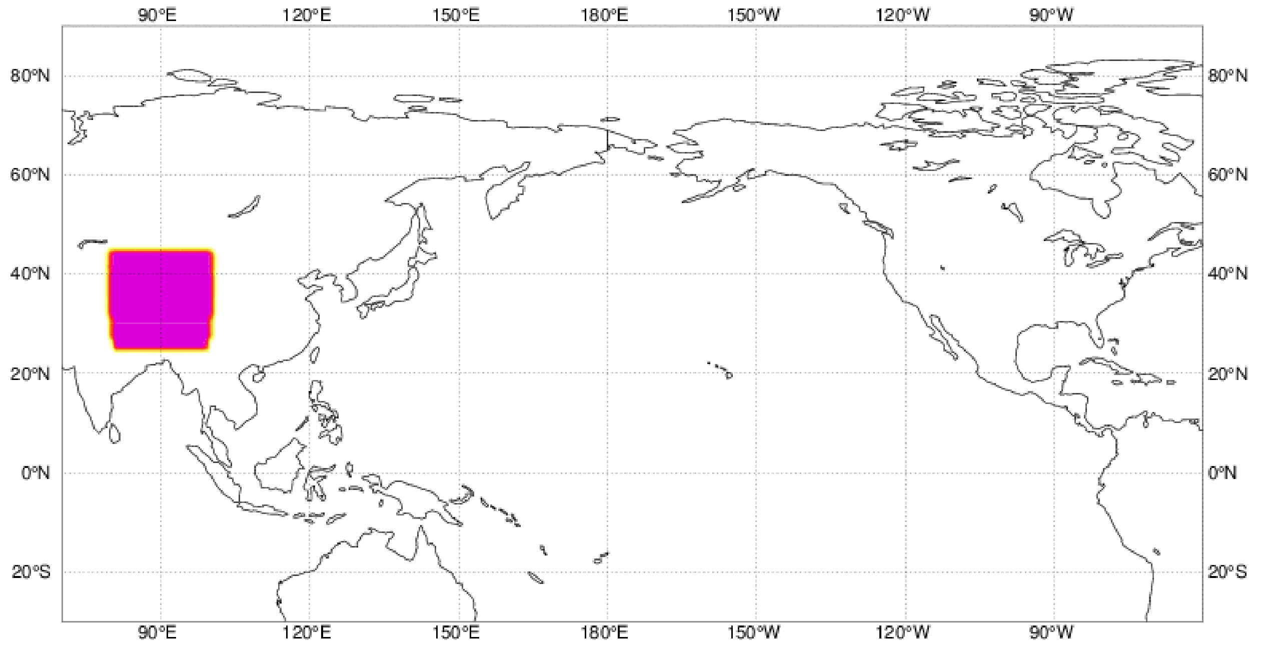
- Native winds on fine grid (~125km)
- Parallel remapping with *Atlas*
- Tracer advection on coarse grid O48 (~200 km)



semi-Lagrangian **on coarse grid (O48)**



flux-form Eulerian **on coarse grid (O48)**

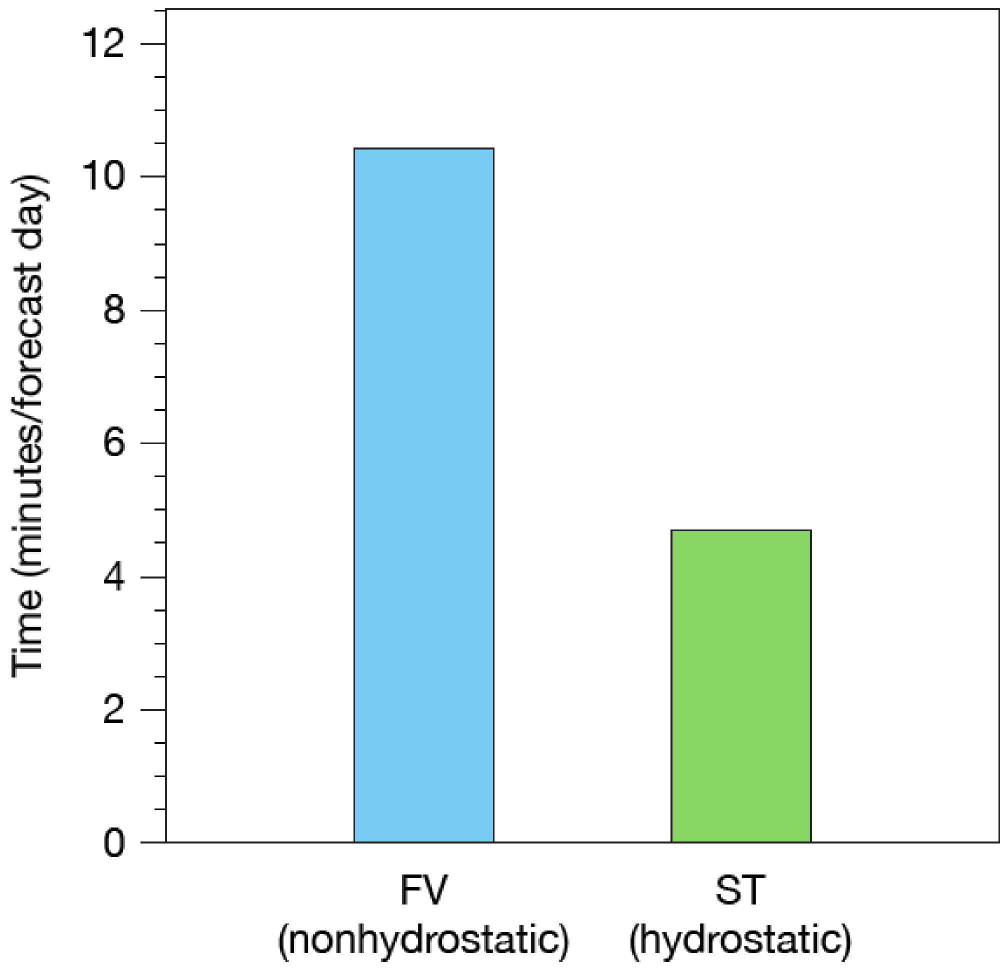


ECMWF

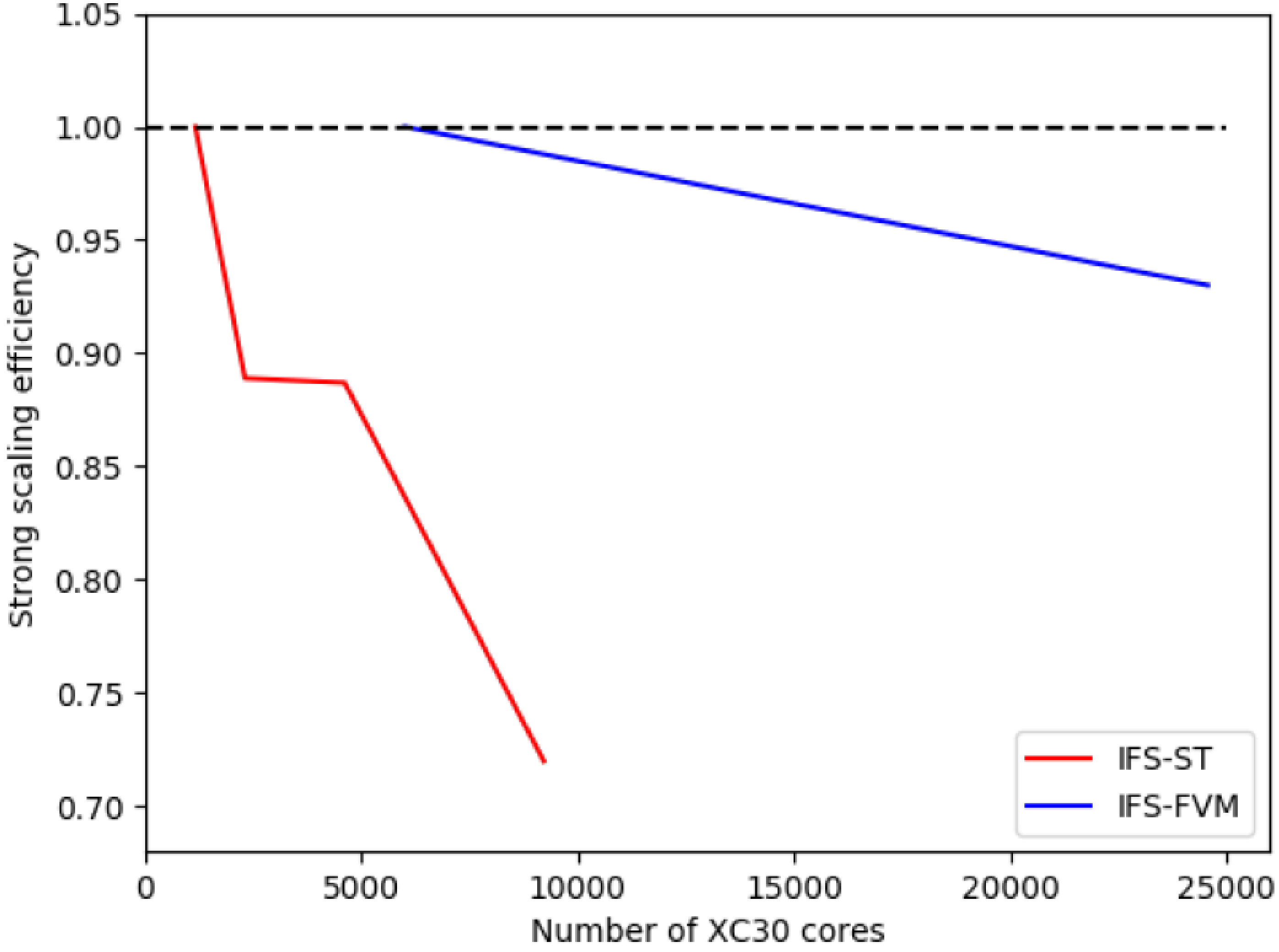
[Courtesy C Kühnlein, P Smolarkiewicz, N Wedi]

IFS-ST vs IFS-FVM advection using *Atlas*

Dry baroclinic instability at 10 km and 137 levels on 350 Cray XC40 nodes



Strong scaling of dynamical core at 13 km resolution



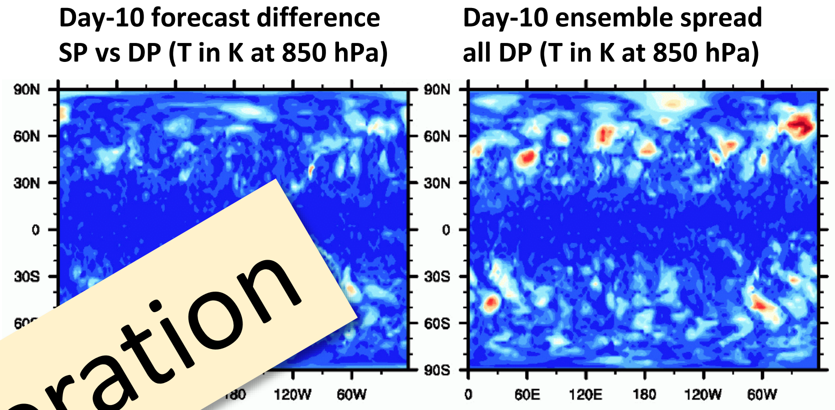
[Courtesy C Kühnlein, N Wedi]

ECMWF Scalability Programme – Do less and do it cheaper

Single precision (Vana et al. 2017, MWR; Dueben et al. 2018, MWR):

- running IFS with single precision arithmetics saves 40% of runtime, IFS-ST offers options like precision by wavenumber;
- storing ensemble model output at even more reduced precision can save 67% of data volume;

→ to be implemented in **operations** asap (capability + capacity)

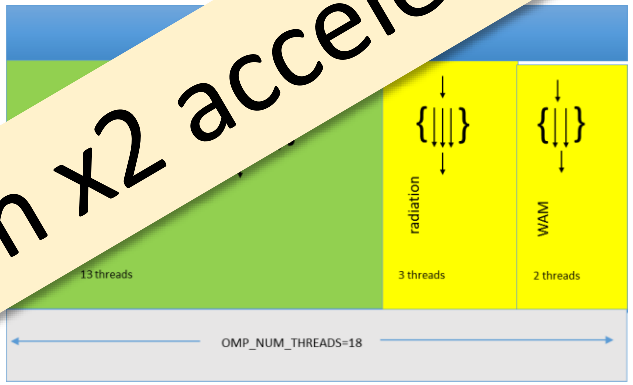


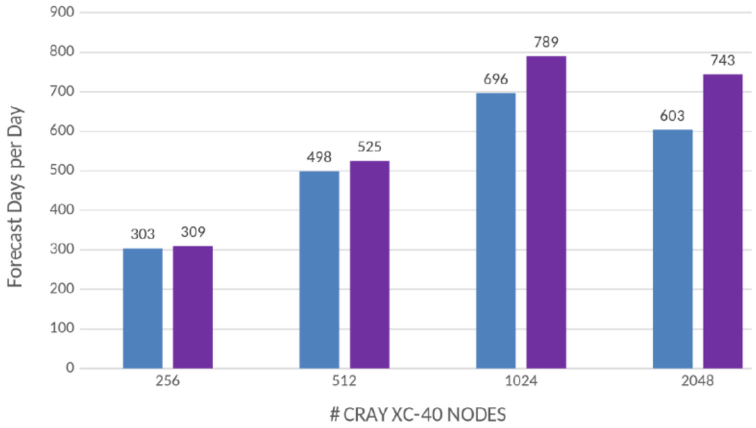
Maximum x2 acceleration

Concurrency:

- allocating threads/task (/across tasks) to model components like radiation or waves can save 20% (gain increases with resolution);

→ to be implemented in **operations** asap (capability + capacity)

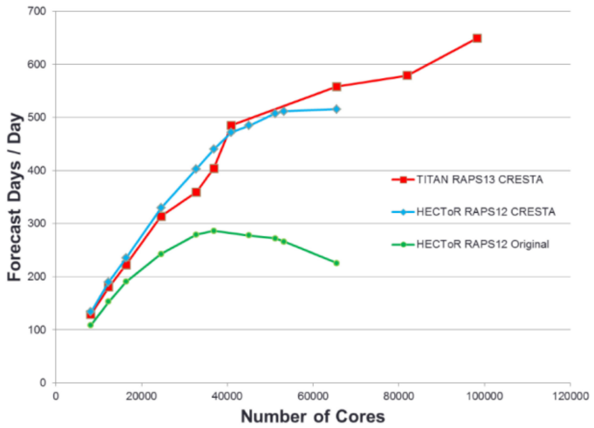




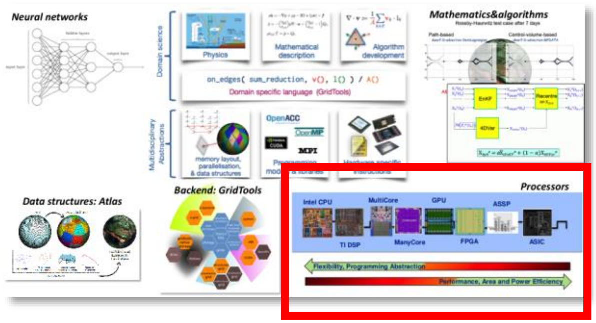
Overlapping communication & computation:

- through programming models (Fortran co-array vs GPI2 vs MPI), gave substantial gains on Titan w/Gemini,
- on XC-30/40 w/ Aries there is no overall performance benefit over default MPI implementation;

→ to be explored further



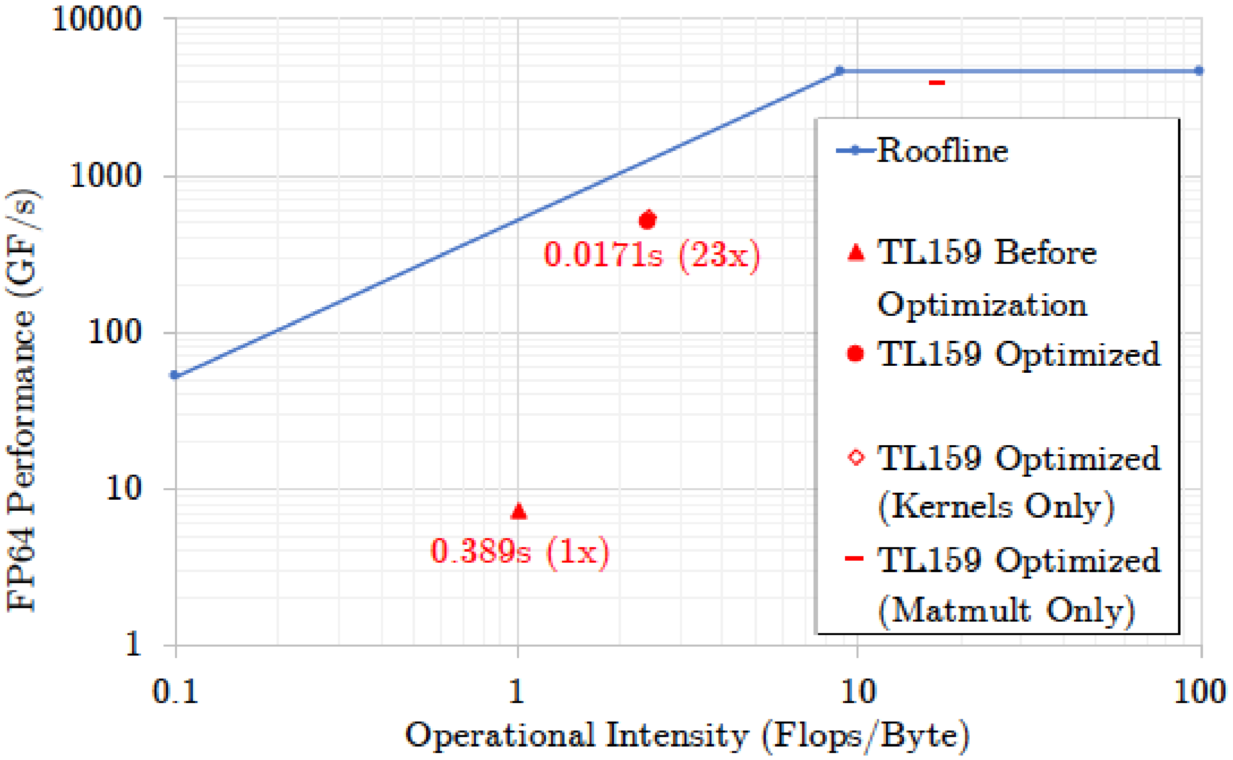
Tasks x threads	Nodes	Experiments	Forecast days/day
2160Tx12t	360	control	**1104.9**
2160Tx12t	360	control	1116.1
2160Tx12t	360	coarray2	815.6
2160Tx12t	360	coarray2	846.0
2160Tx12t	360	gpi2	788.9



Funded by the European Union

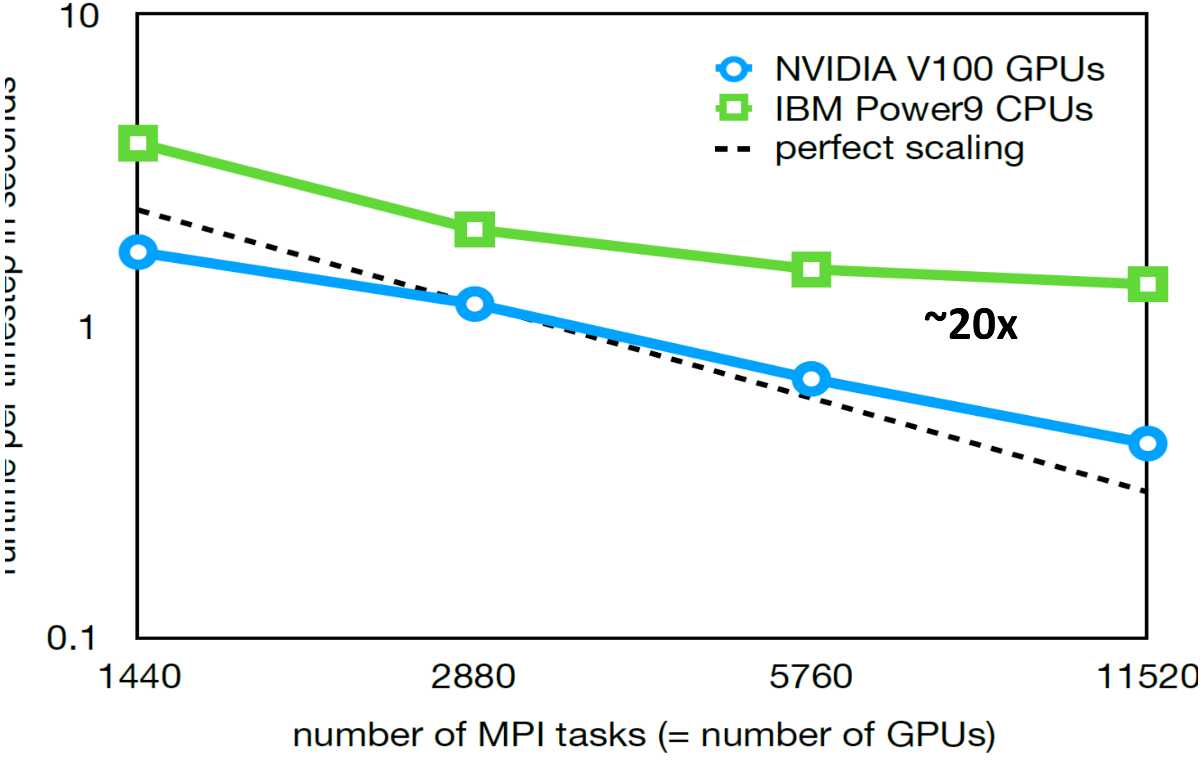
ESCAPE dwarfs on GPU

Spectral transforms on GPU - single core

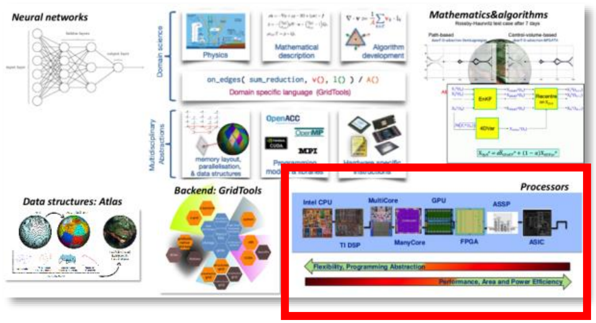


[Müller et al. 2019, Geoscientific Model Development] FORECASTS

Spectral transform dwarf @ 2.5 km, 240 fields on Summit GPU (2 CPU vs 6 GPU):



[Courtesy A Müller]



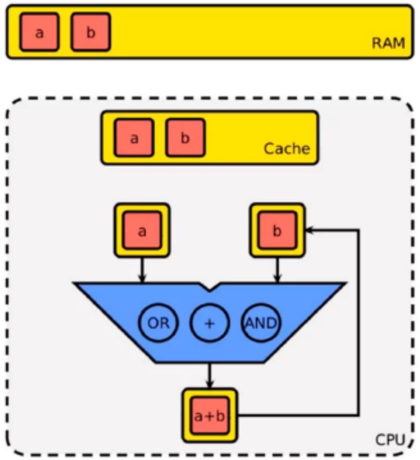
ESCAPE dwarfs on FPGA

Funded by the European Union

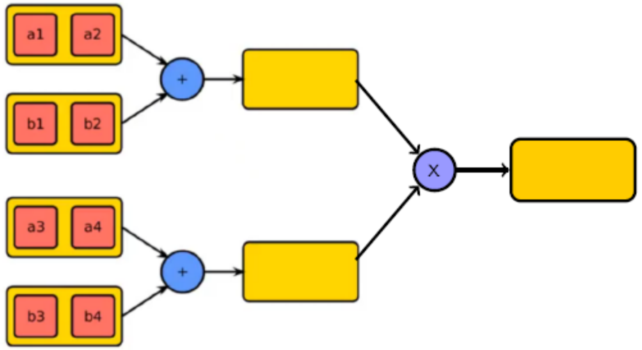
EuroEXA

- Converted complex Fortran code and data structures to C via source-to-source translation
- Hand-ported to MaxJ via Maxeler IDE and emulator

CPU (Von Neumann)



FPGA (dataflow)

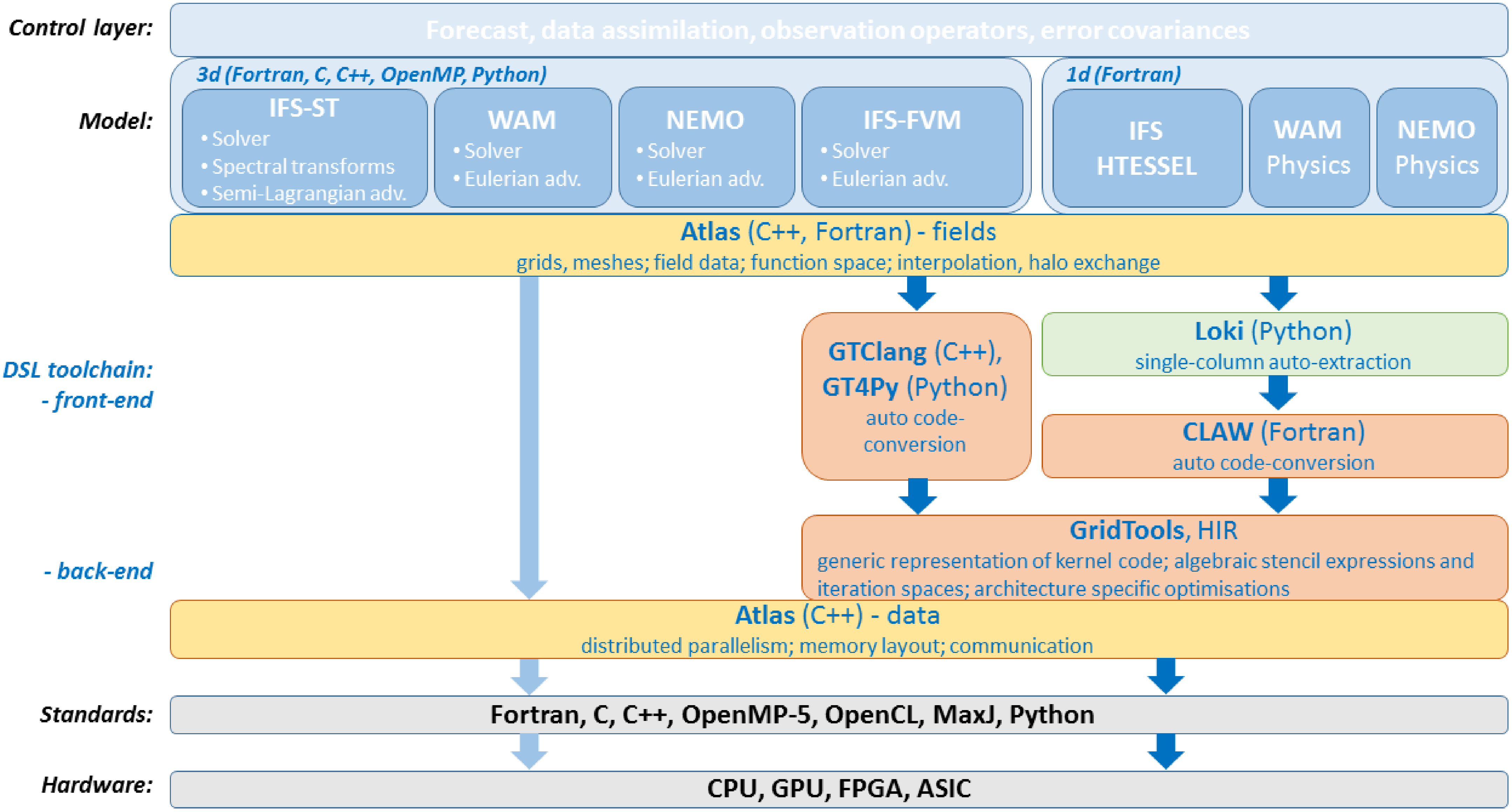


- On-board memory bandwidth limit (no PCIe): 1.13 million columns/s
- Dataflow kernel compiled at 156MHz
- 156 million cells/s, equivalent to 1.07 million columns/s
- Average flops / column estimated on CPU; **Extrapolated equivalent FPGA performance of 133.6 Gflops/s**
- Reference run on 12-core 2.6 GHz Intel Haswell, single socket **CPU is about 21 Gflops/s**, but with double precision!
- Dynamic power usage is < 30W compared to 95W single socket CPU (Haswell)

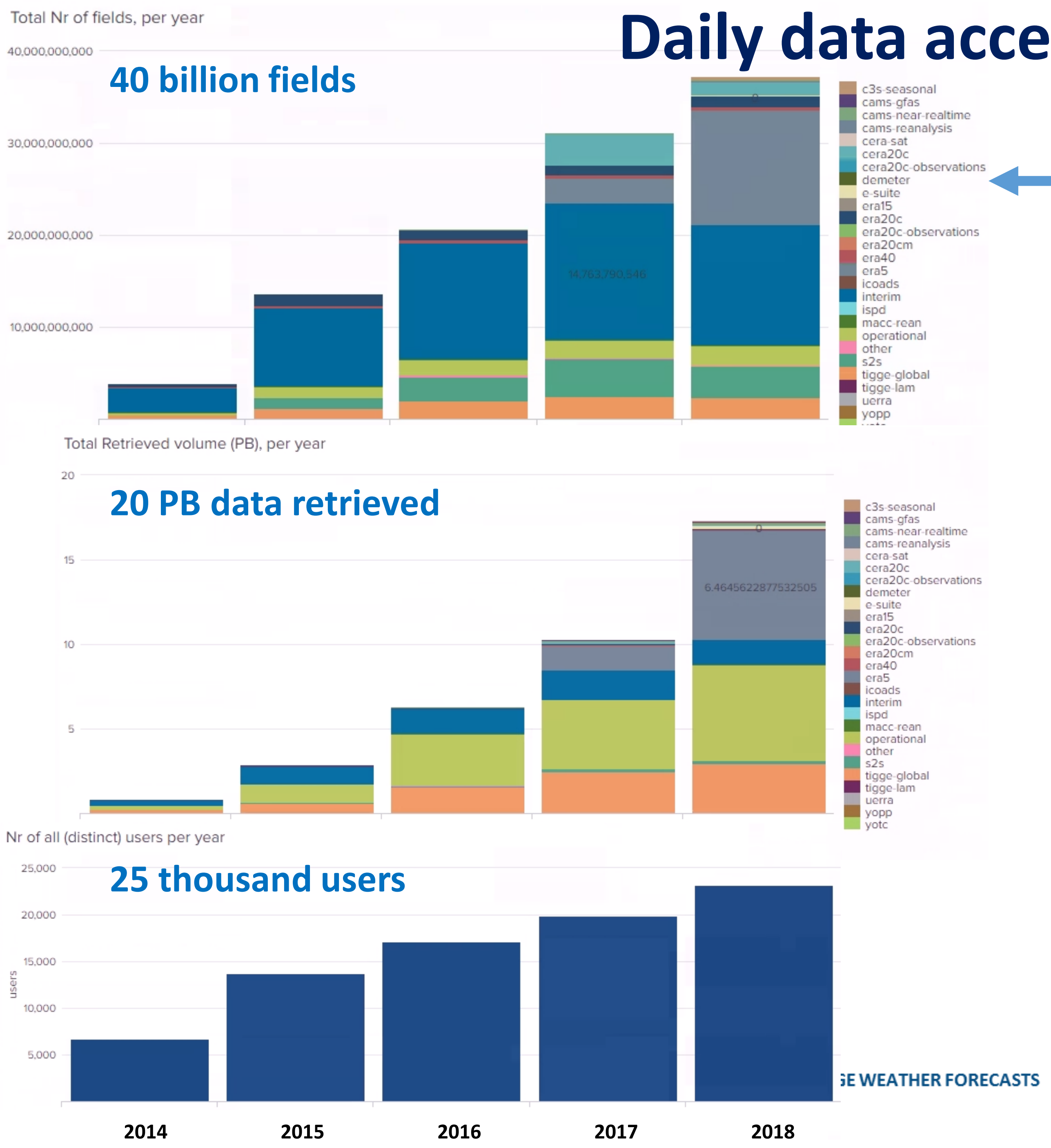
→ x3 time to solution times x3 energy to solution

[Courtesy M Lange, O Marsden, J Taggert]

Separation of Concerns with IFS (in stages)



Daily data access at ECMWF



← *Public*

Total activity (Member States and commercial customers) per day:

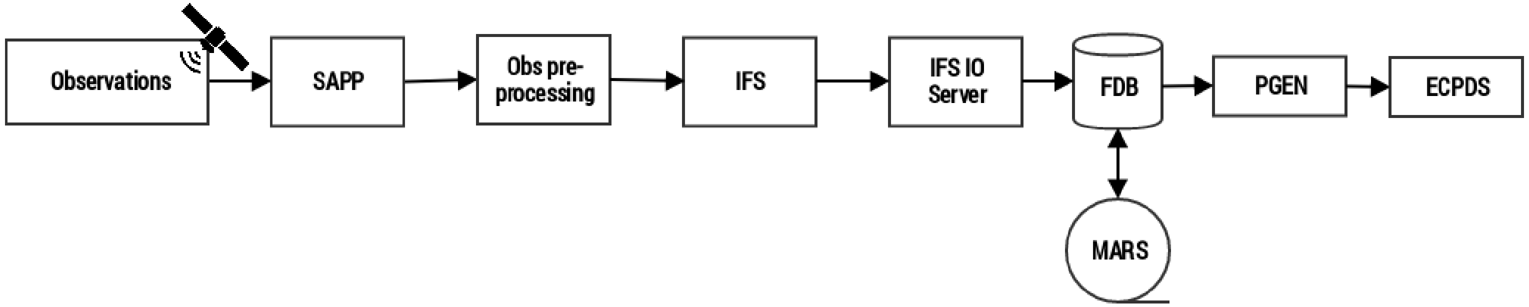
- 450 TBytes retrieved
- 200 TBytes archived
- 1.5 million requests

Total volume in MARS: 220 PiB

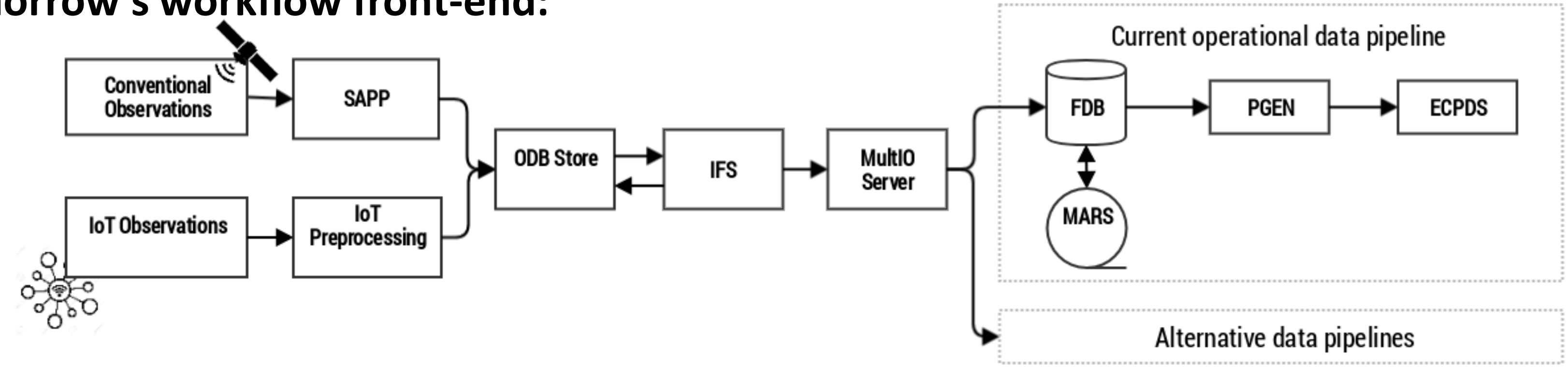
[Courtesy M Manoussakis]

Numerical Weather Prediction Data Flow

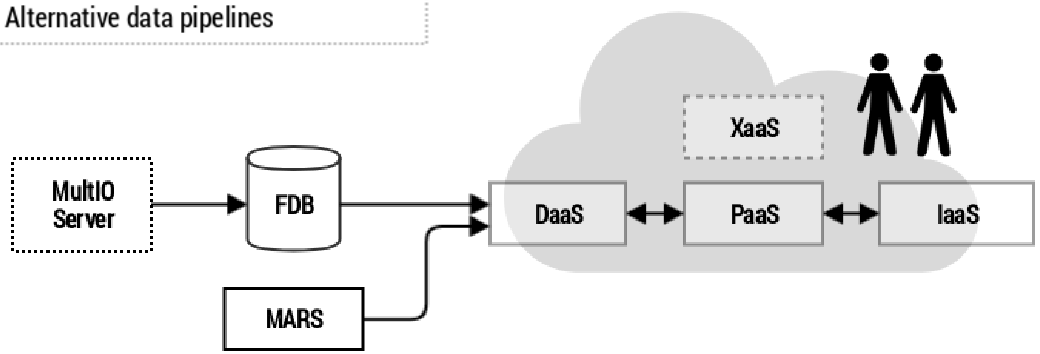
Today's workflow:



Tomorrow's workflow front-end:

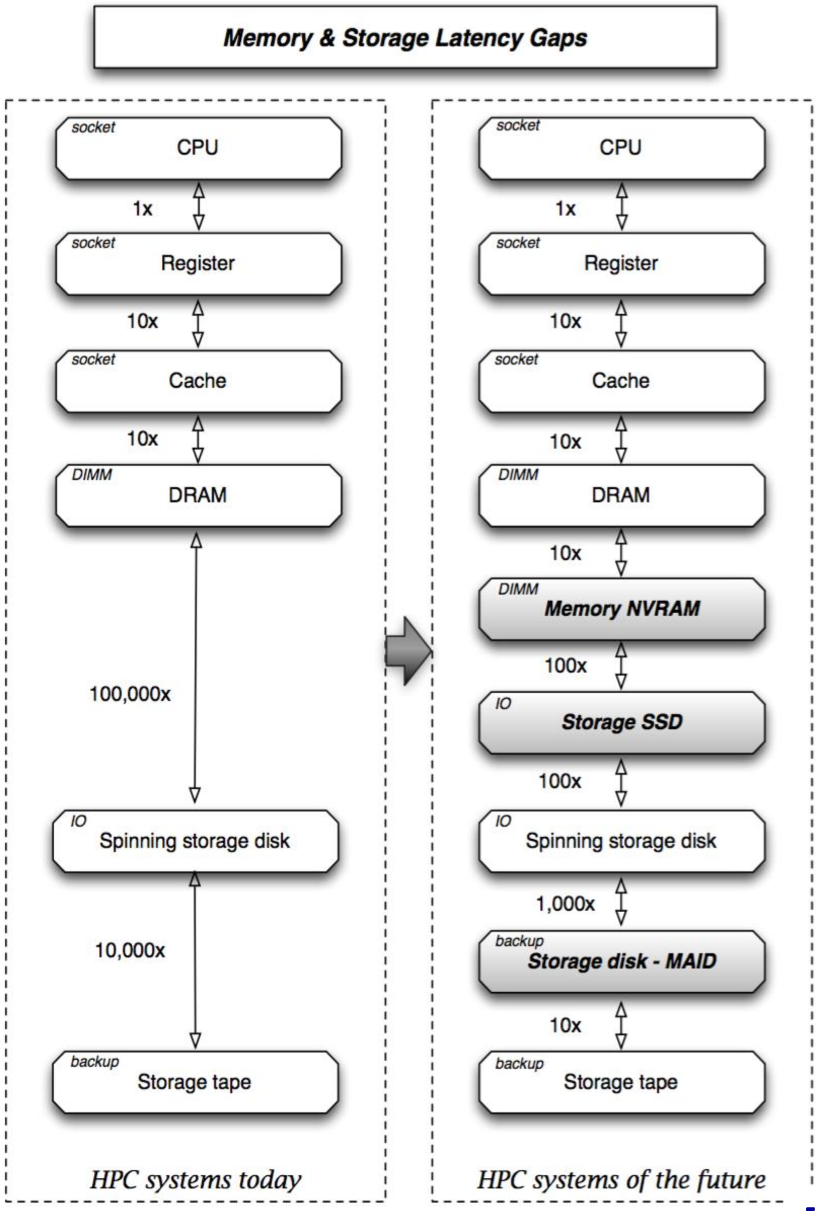


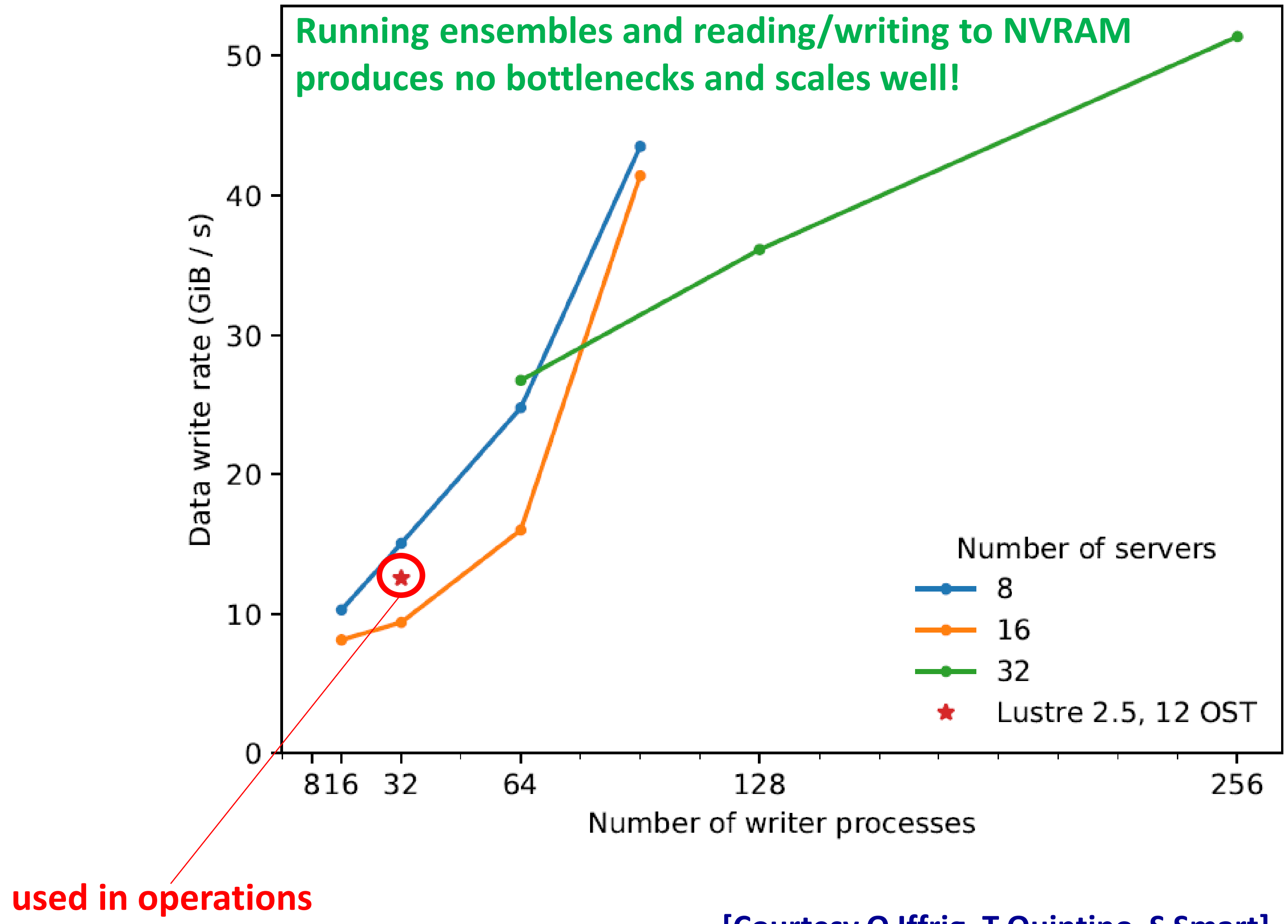
Tomorrow's workflow back-end:



[Courtesy J Hawkes, T Quintino]

ECMWF Scalability Programme – Use new memory technology





[Courtesy O Iffrig, T Quintino, S Smart]

ANGE WEATHER FOREC

[Courtesy S Smart, T Quintino]

Funded by the European Union

nextgenio

Machine learning application areas in workflow

Observational data processing *(edge & cloud &HPC)*:

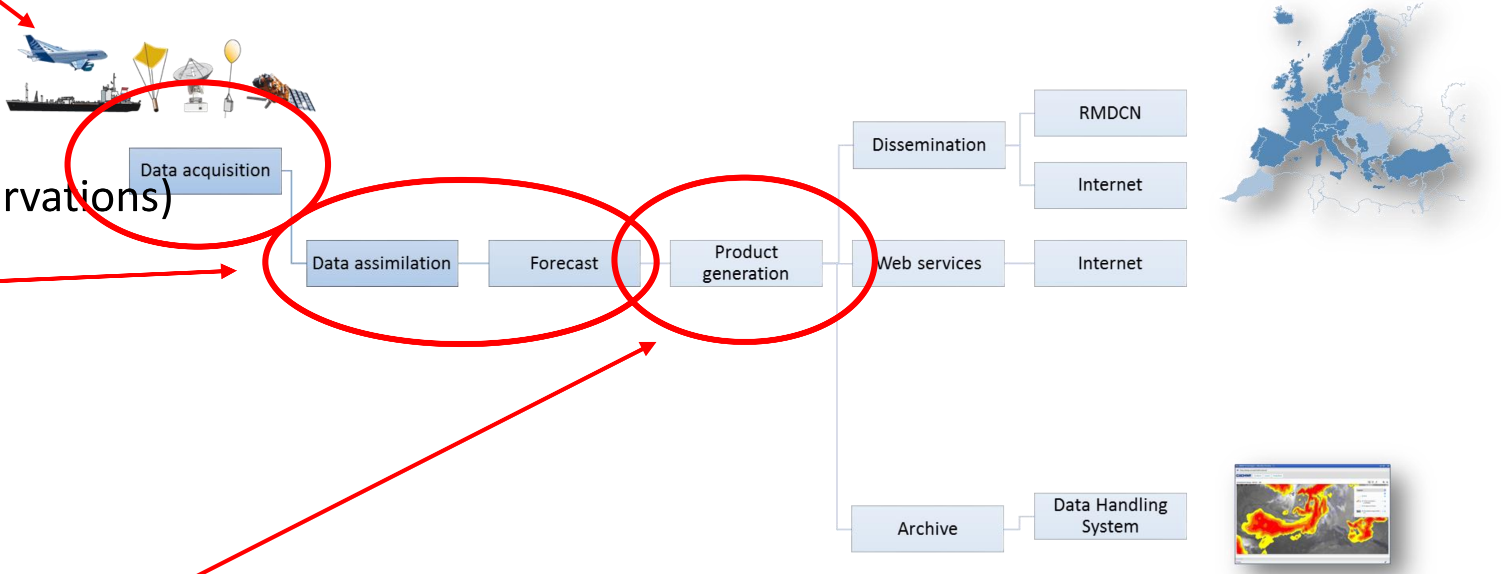
- Quality control and bias correction
- Data selection
- Inversion (=retrieval)
- Data fusion (combining observations)
- …

Prediction models *(cloud & HPC)*:

- Data assimilation (combining models w/ observations)
- Surrogate model components
- Prediction itself
- Model error statistics
- …

Service output data processing *(cloud &HPC)*:

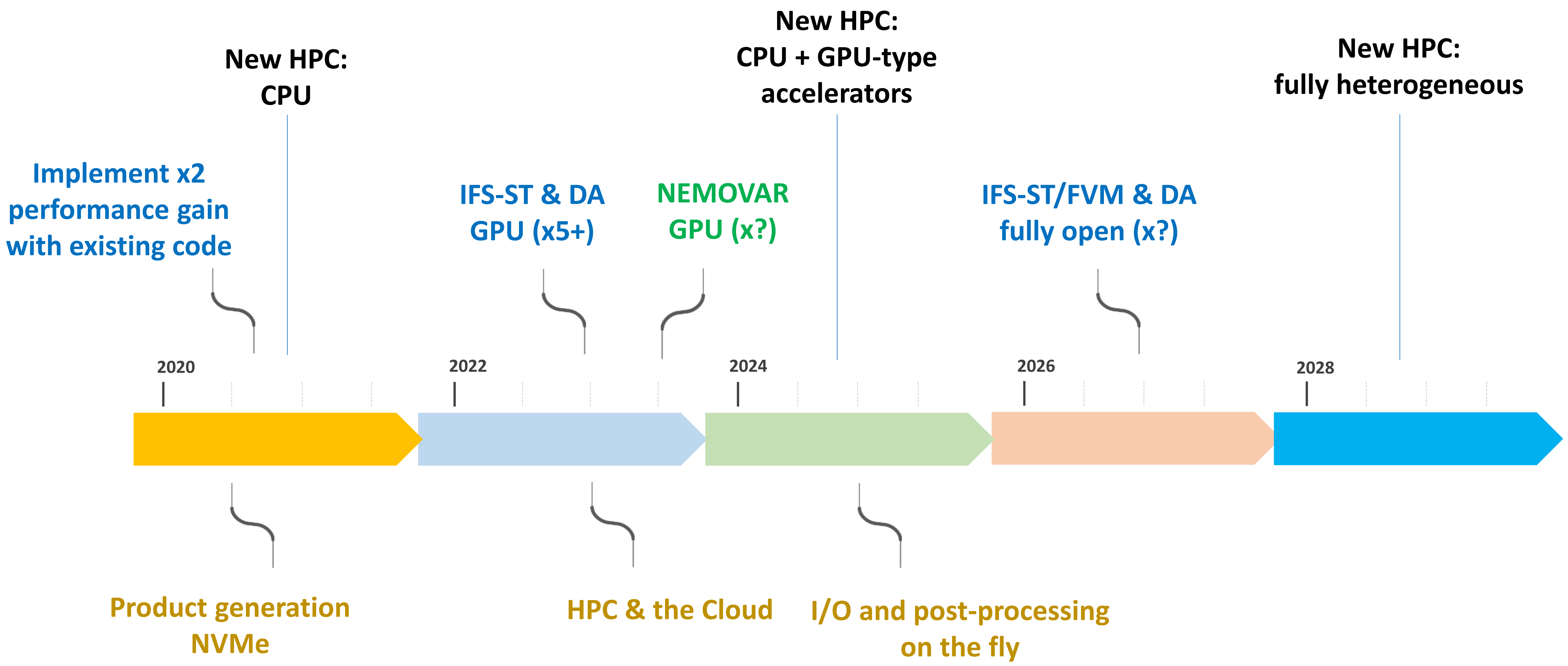
- Product generation and dissemination
- Product feature extraction (data mining)
- Product error statistics
- Interactive visualisation and selection
- Data handling (access prediction)
- …



Existing projects (Peter Dueben):

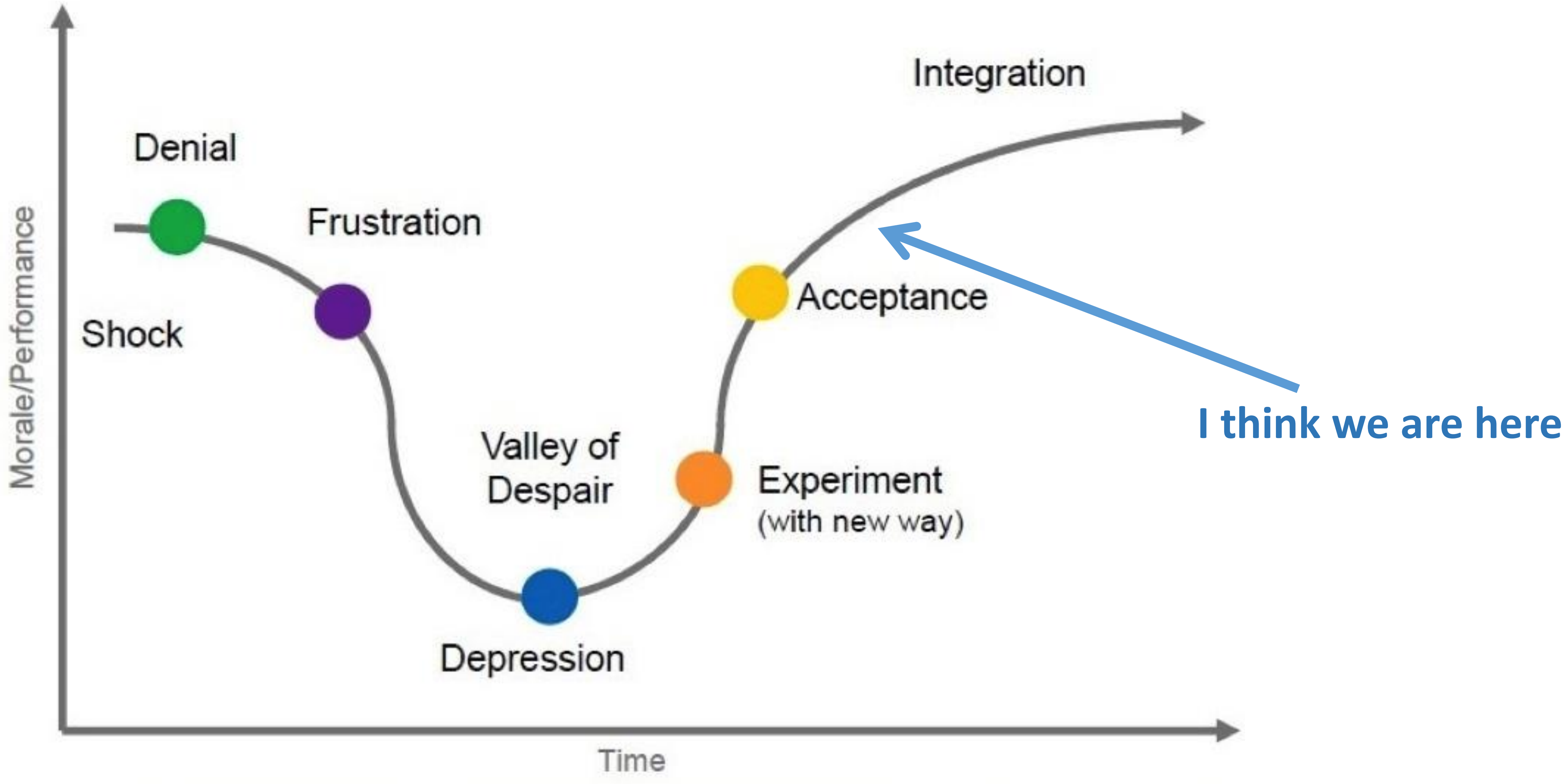
- Radiation code emulation (NVIDIA)
- Predicting uncertainty from poor ensembles (U Oxford)
- Refining variational bias correction in data assimilation
- Refining uncertain parameter settings
- and more

So, where are we with all this?



Open questions:

- What about code that is not in our control, e.g. NEMO?
- Do we have sufficient expertise – collaboration?
- Do we have sufficient funding?


Integration
Denial
Frustration
Acceptance
Shock
Morale/Performance
I think we are here
Valley of
Despair
Experiment
(with new way)
Depression
Time