ARMing the IFS: Experiments and experiences from porting the ECMWF model to Fugaku

Sam Hatfield *with help* from Seiya Nishizawa, Hirofumi Tomita, Ioan Hadade, Balthasar Reuter, Peter Dueben, Olivier Marsden, Willem Deconinck and Michael Lange

samuel.hatfield@ecmwf.int

ESiWACE2 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823988.
This research used resources of the Oak Ridge Leadership Computing Facility (OLCF), which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.

Exascale NWP at ECMWF


ECMWF
European Centre for Medium-Range Weather Forecasts

earth system science

From the Scalability Programme to Destination Earth

Peter Bauer, Tiago Quintino, Nils Wedi

This article recalls the origins of ECMWF's Scalability Programme eight years ago and how it helped to prepare ECMWF for the EU-funded Destination Earth initiative, which emerged through a considerable effort from some of the leading scientists in the field of weather and climate prediction and computational science. The Scalability Programme was designed to adapt weather prediction codes to emerging computing paradigms, and Destination Earth is an initiative to develop a highly accurate digital twin of our planet. Both initiatives illustrate how fruitful international collaboration can be, but they also highlight the challenges for devising and implementing large and ambitious programmes supporting the digital revolution of Earth system science.

The scalability challenge

ECMWF launched its Scalability Programme in 2014 as a result of the realisation that the efficiency of our substantial codes and complex workflows would not be future-proof in view of new, emerging computing paradigms. In particular, CPU processor performance was being capped to reduce power consumption. At the same time, substantial investments in software were needed so that emerging, alternative processor technologies like GPUs could be explored. GPUs have been around in high-performance computing (HPC) since 2006, but they presented serious code adaptation challenges that limited their use as a general-purpose computing device and delayed their wider uptake in operational weather and climate prediction. A notable exception was MeteoSwiss, which invested in a substantial, manual GPU code rewrite early and operationalised its limited-area prediction system on CPUs and GPUs in 2016.

It also became clear that enhanced computing requires enhanced data handling capabilities, as big computing creates big data. The classical approach of storing all model output on tapes and performing frequent read and write operations was going to change radically in the future. The data handling challenge appeared to be more critical at the model output than at the observational data input stage as model output increases at least by an order of magnitude when resolution is doubled.

Advances in research and operational prediction lead to a doubling of horizontal resolution about every eight years. Upgrades in vertical resolution and the use of ensembles in assimilation and forecasts add to the multiplication of the computing and data burden. With no sign of such demands reaching saturation, there was increased pressure to make a dedicated investment in scalability.

In the past, continuous efforts were made to exploit shared and distributed processor memory in ECMWF's Integrated Forecasting System (IFS). However, a rethink of algorithmic choices and computational efficiency became necessary because of a number of developments: the complex setup of semi-implicit spectral and grid-point calculations; the semi-Lagrangian advective transport and its necessary data communication; and the coupling to increasingly complex Earth-system-component models for waves, ocean, sea ice and land, based on different grids, solvers and computational patterns. Any improvement also needed to be sustainable and thus portable to emerging and future computing architectures without repeatedly redesigning models, algorithms and work- and dataflows.

Scalability Programme – the research phase

The Scalability Programme was launched as a concerted effort across ECMWF and its Member States. Contrary to most projects run at the same time at most weather services and climate prediction centres, the ECMWF programme targeted both efficiency and sustainability of the entire work- and dataflow.

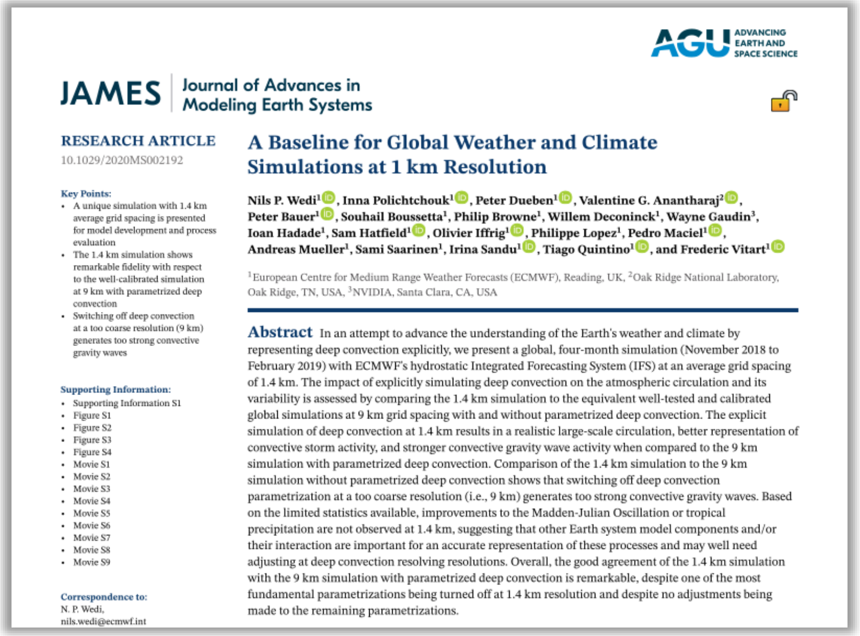
Initially, the programme was formulated around several projects focusing on observational data ingestion and pre-processing, data assimilation and the forecast model, model output data handling and product generation.

Finding resources

Once initiated, the programme needed resources. Our Member States were very supportive as performance and technology concerns were widely shared across the weather and climate community. The ECMWF Council agreed to add five staff members in 2014. The first generation of scalability staff came mostly from diverse, non-meteorological backgrounds.

ECMWF Newsletter 171 • Spring 2022 doi: 10.21957/pb2vnp59ks 15

ECMWF Newsletter, April 2022


AGU ADVANCING EARTH AND SPACE SCIENCE

JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2020MS002192

A Baseline for Global Weather and Climate Simulations at 1 km Resolution

Nils P. Wedi[1], Inna Polichtchouk[1], Peter Dueben[1], Valentine G. Anantharaj[2], Peter Bauer[1], Souhail Boussetta[1], Philip Browne[1], Willem Deconinck[1], Wayne Gaudin[3], Ioan Hadade[1], Sam Hatfield[1], Olivier Iffrig[1], Philippe Lopez[1], Pedro Maciel[1], Andreas Mueller[1], Sami Saarinen[1], Irina Sandu[1], Tiago Quintino[1], and Frederic Vitart[1]

[1]European Centre for Medium Range Weather Forecasts (ECMWF), Reading, UK, [2]Oak Ridge National Laboratory, Oak Ridge, TN, USA, [3]NVIDIA, Santa Clara, CA, USA

Key Points:
- A unique simulation with 1.4 km average grid spacing is presented for model development and process evaluation
- The 1.4 km simulation shows remarkable fidelity with respect to the well-calibrated simulation at 9 km with parametrized deep convection
- Switching off deep convection at a too coarse resolution (9 km) generates too strong convective gravity waves

Supporting Information:
- Supporting Information S1
- Figure S1
- Figure S2
- Figure S3
- Figure S4
- Movie S1
- Movie S2
- Movie S3
- Movie S4
- Movie S5
- Movie S6
- Movie S7
- Movie S8
- Movie S9

Correspondence to:
N. P. Wedi,
nils.wedi@ecmwf.int

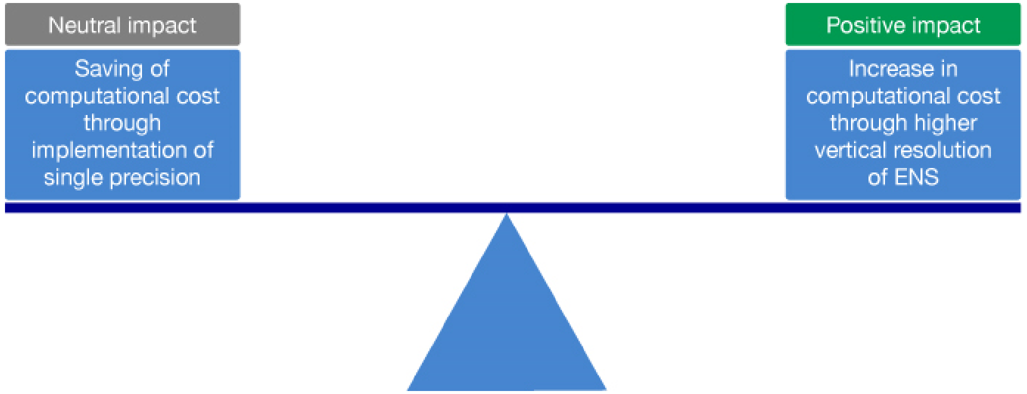
Abstract In an attempt to advance the understanding of the Earth's weather and climate by representing deep convection explicitly, we present a global, four-month simulation (November 2018 to February 2019) with ECMWF's hydrostatic Integrated Forecasting System (IFS) at an average grid spacing of 1.4 km. The impact of explicitly simulating deep convection on the atmospheric circulation and its variability is assessed by comparing the 1.4 km simulation to the equivalent well-tested and calibrated global simulations at 9 km grid spacing with and without parametrized deep convection. The explicit simulation of deep convection at 1.4 km results in a realistic large-scale circulation, better representation of convective storm activity, and stronger convective gravity wave activity when compared to the 9 km simulation with parametrized deep convection. Comparison of the 1.4 km simulation to the 9 km simulation without parametrized deep convection shows that switching off deep convection parametrization at a too coarse resolution (i.e., 9 km) generates too strong convective gravity waves. Based on the limited statistics available, improvements to the Madden-Julian Oscillation or tropical precipitation are not observed at 1.4 km, suggesting that other Earth system model components and/or their interaction are important for an accurate representation of these processes and may well need adjusting at deep convection resolving resolutions. Overall, the good agreement of the 1.4 km simulation with the 9 km simulation with parametrized deep convection is remarkable, despite one of the most fundamental parametrizations being turned off at 1.4 km resolution and despite no adjustments being made to the remaining parametrizations.

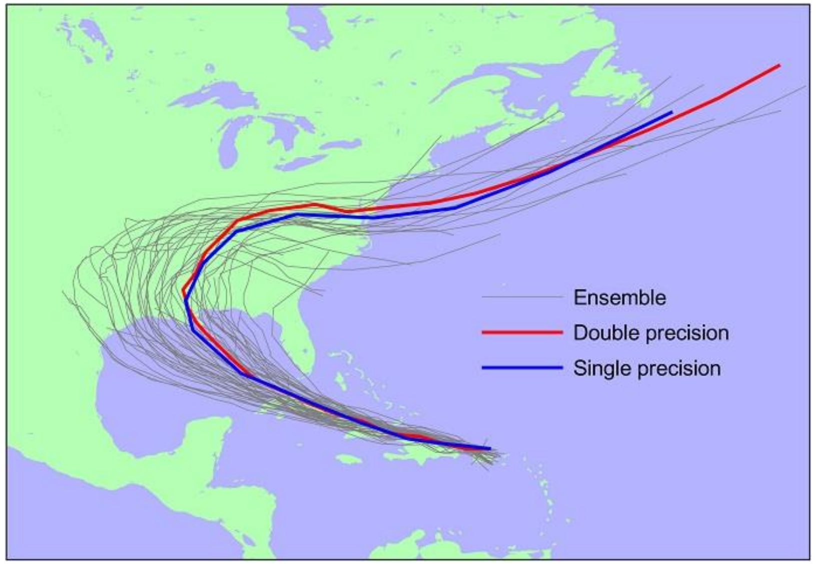
Wedi et al., 2020

Current HPC activities at ECMWF

Machine	Use	Peak perf. (PFLOP/s)	Hardware	Toolchain
CCA/CCB (Cray XC40)	Operations (old)	8	CPU (Intel)	Cray
AA/AB/AC/AD (Atos BullSequana XH2000)	Operations (new)	30	CPU (AMD) + GPU (Nvidia)	Intel + NVHPC
JUWELS Booster	nextGEMS	70	GPU (Nvidia)	NVHPC
Summit	INCITE	200	GPU (Nvidia)	NVHPC
LUMI-G (test nodes)	Benchmarking	550	GPU (AMD)	CCE/ROCm
Frontier	INCITE (provisional)	1500	GPU (AMD)	TBC
Fugaku	Benchmarking	500	CPU (ARM)	Fujitsu

Towards a fully single-precision Earth-system model





https://www.ecmwf.int/en/about/media-centre/news/2021/
forecast-upgrade-innovates-single-precision-and-ensemble-resolution

A zoo of number formats

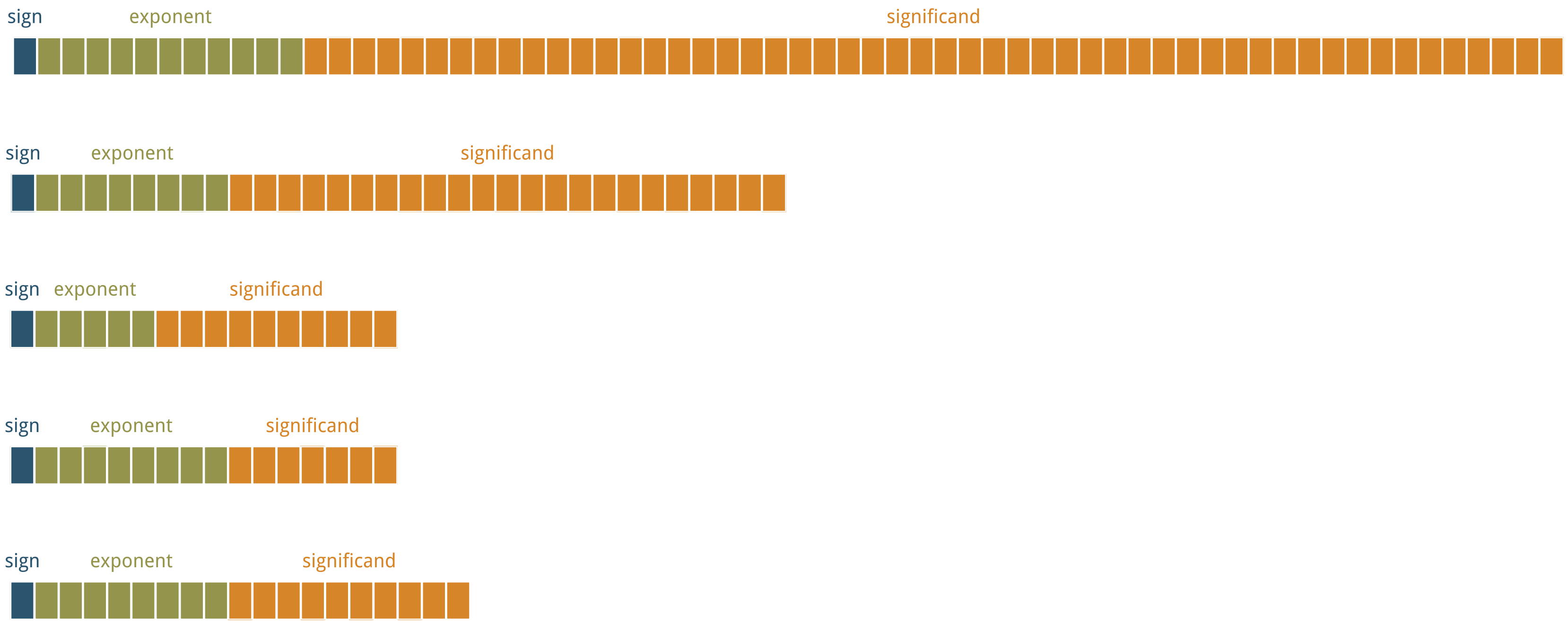
FP64 “double precision”

FP32 “single precision”

FP16 “half precision”

bfloat16

“TF32” (effectively 19-bit)

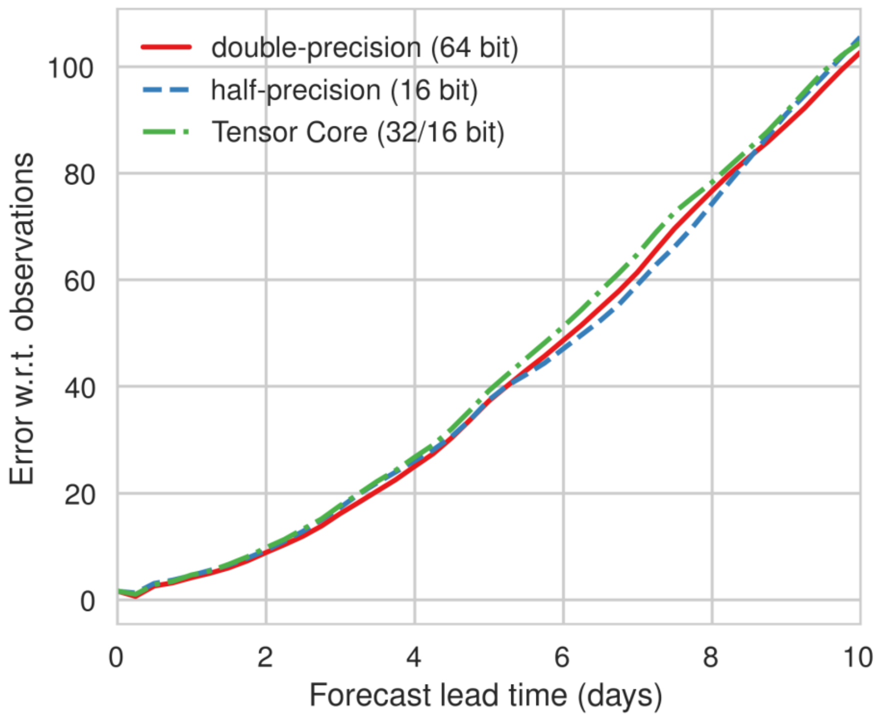


A zoo of number formats

FP16 “half precision”

Half-precision spectral transforms

NH Z500, TCO1279

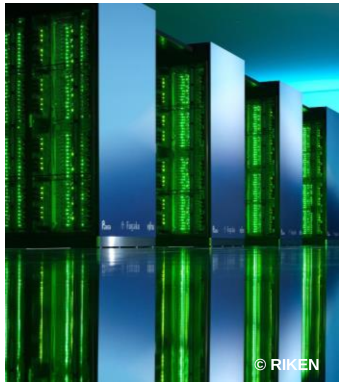


Skill of forecasts using **half-precision Legendre transforms** compared with **double precision**

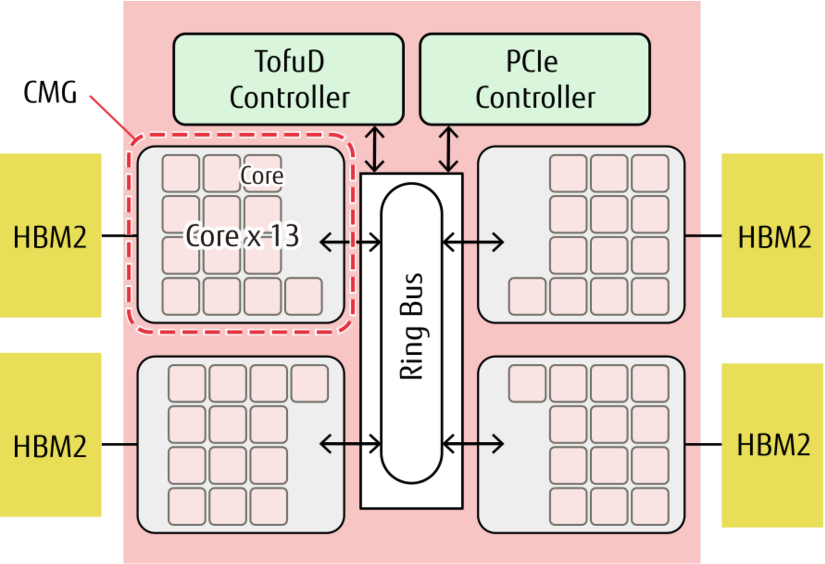
Hatfield et al. 2019, https://doi.org/10.1145/3324989.3325711

- **Legendre transforms** of the IFS a good target for half precision
  - Bottleneck at high resolution
  - Compact code
  - Algorithmically simple → series of GEMMs
- Preliminary software emulation studies (Hatfield et al. 2019):
  - Half precision can be used in Legendre transforms even up to TCO1279 (9 km globally) resolution
  - Necessary to **rescale** inputs/outputs, as before

The first half-precision CPU: Fujitsu A64FX



Fugaku = 1 A64FX/node
× 158,976 nodes
540 PFLOP/s peak



- AArch64 (ARM) instruction set
- 48 cores split among 4 “CMGs” (core memory groups)
- 32 GB High Bandwidth Memory
- No DDR RAM/L3 cache
- **Native support for FP16**

ECMWF/R-CCS collaboration

ECMWF

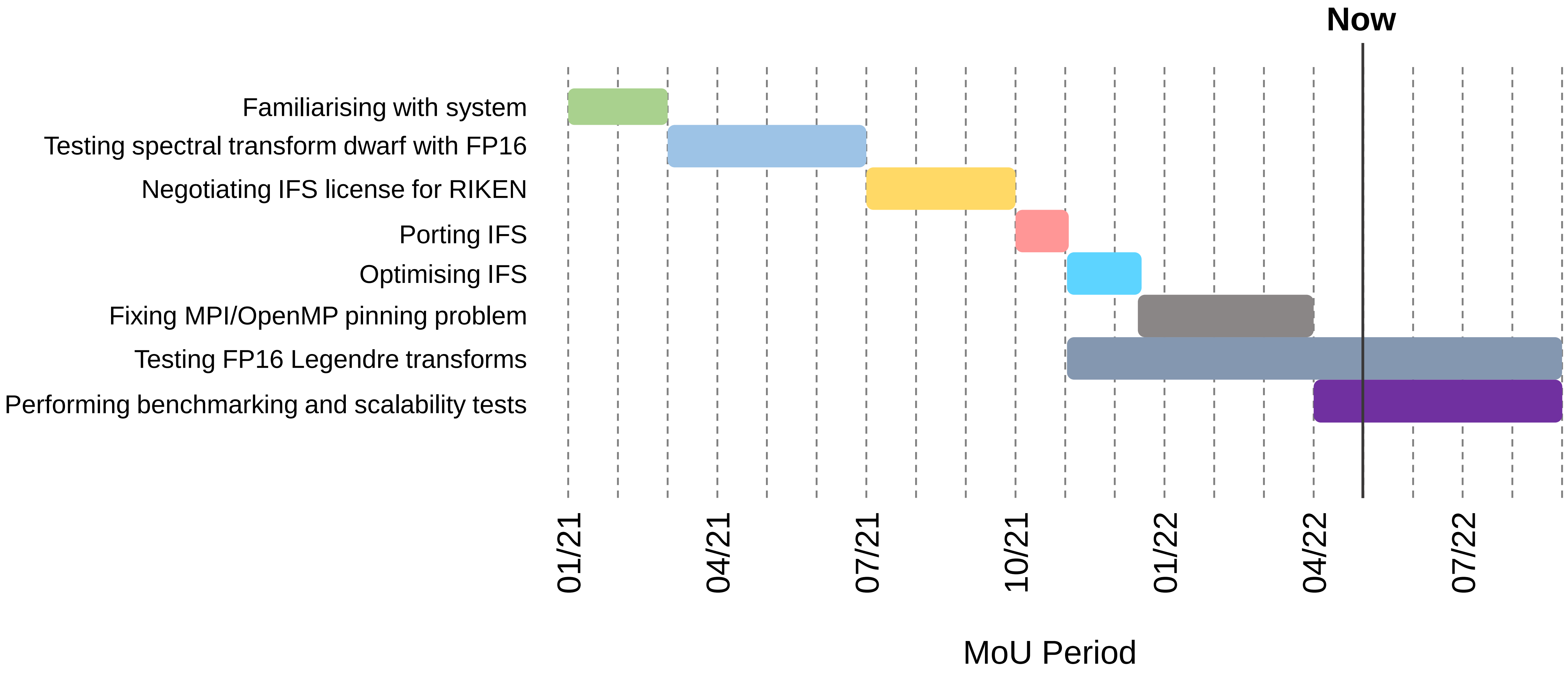
RIKEN R-CCS

- Initiated between R-CCS and ECMWF in January 2021
- **R-CCS:** Hirofumi Tomita, Seiya Nishizawa, Tsuyoshi Yamaura
- **ECMWF:** Sam Hatfield, Peter Dueben
- Modest budget: **~20,000 node-hours/year**

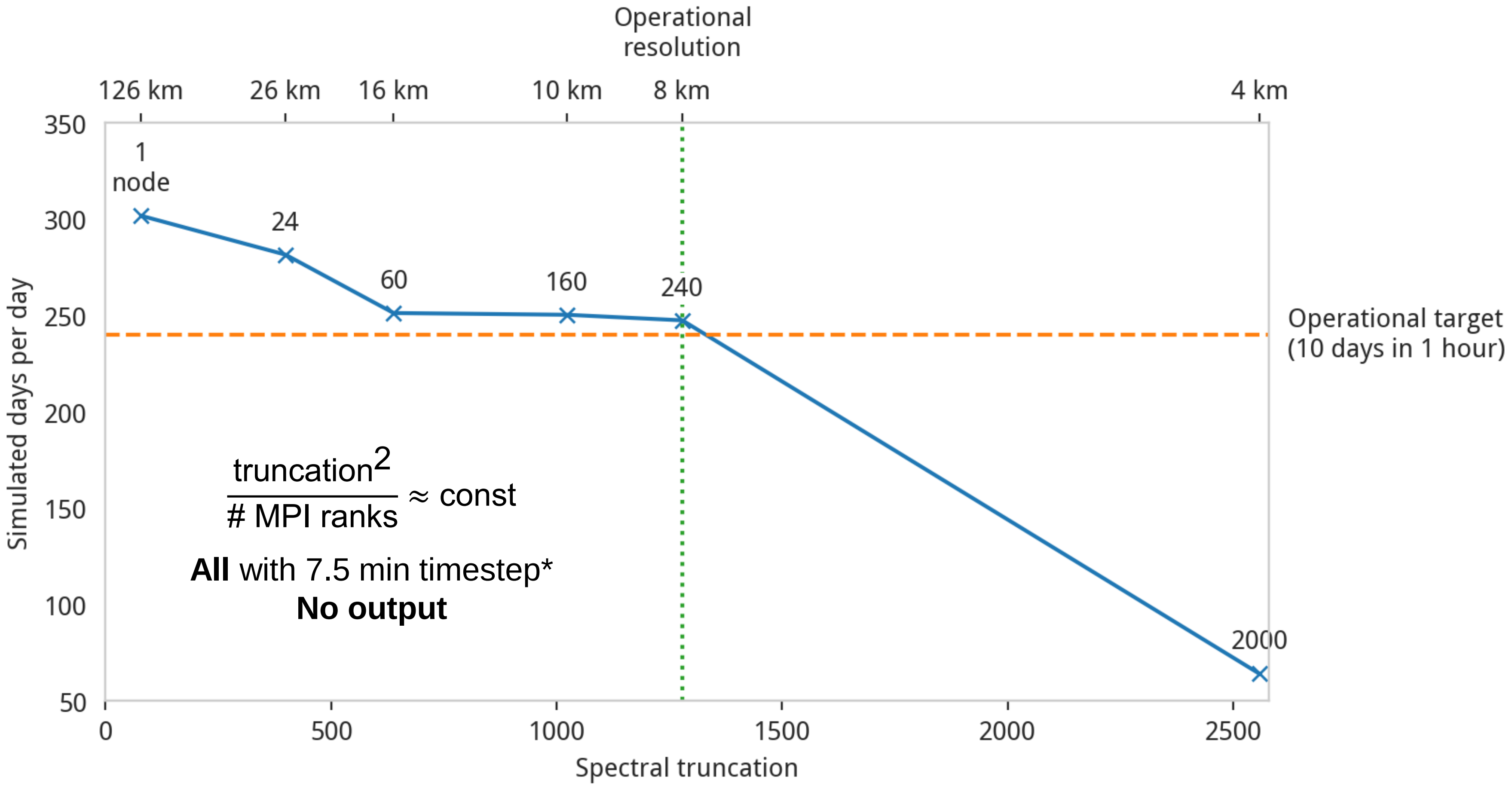
Key questions:

- How easy is it to port existing weather and climate codes to ARM? (focusing on **Fugaku**)
- How can FP16 limitations (low range, large rounding errors) be accommodated by algorithmic changes?
- What FP16 speed-up can be realised in real world applications?

Porting the IFS to Fugaku



Weak scalability on Fugaku



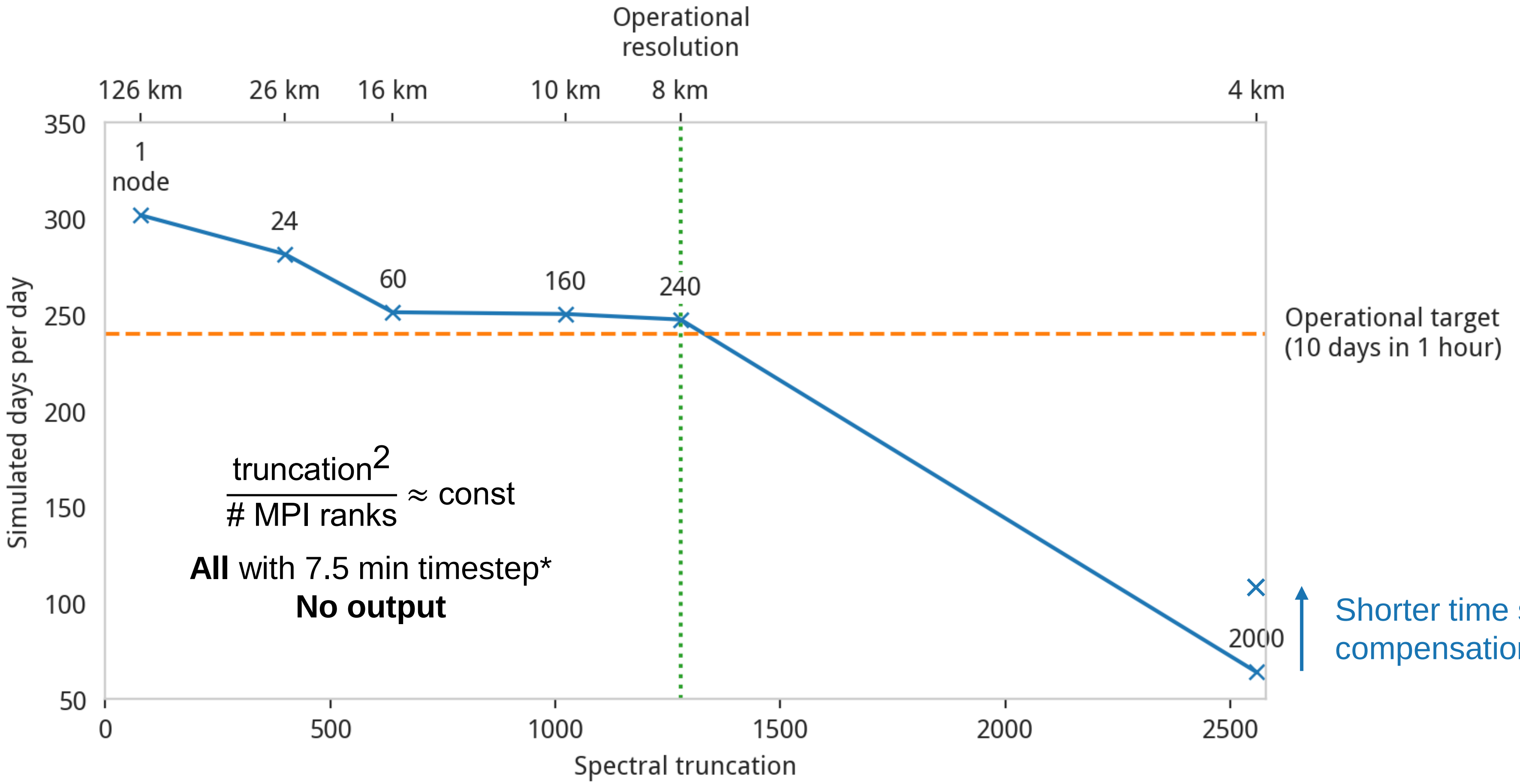
$$\frac{\text{truncation}^2}{\text{\# MPI ranks}} \approx \text{const}$$

All with 7.5 min timestep*

No output

**except TCO2559/4 km → timestep 4 min*

Weak scalability on Fugaku



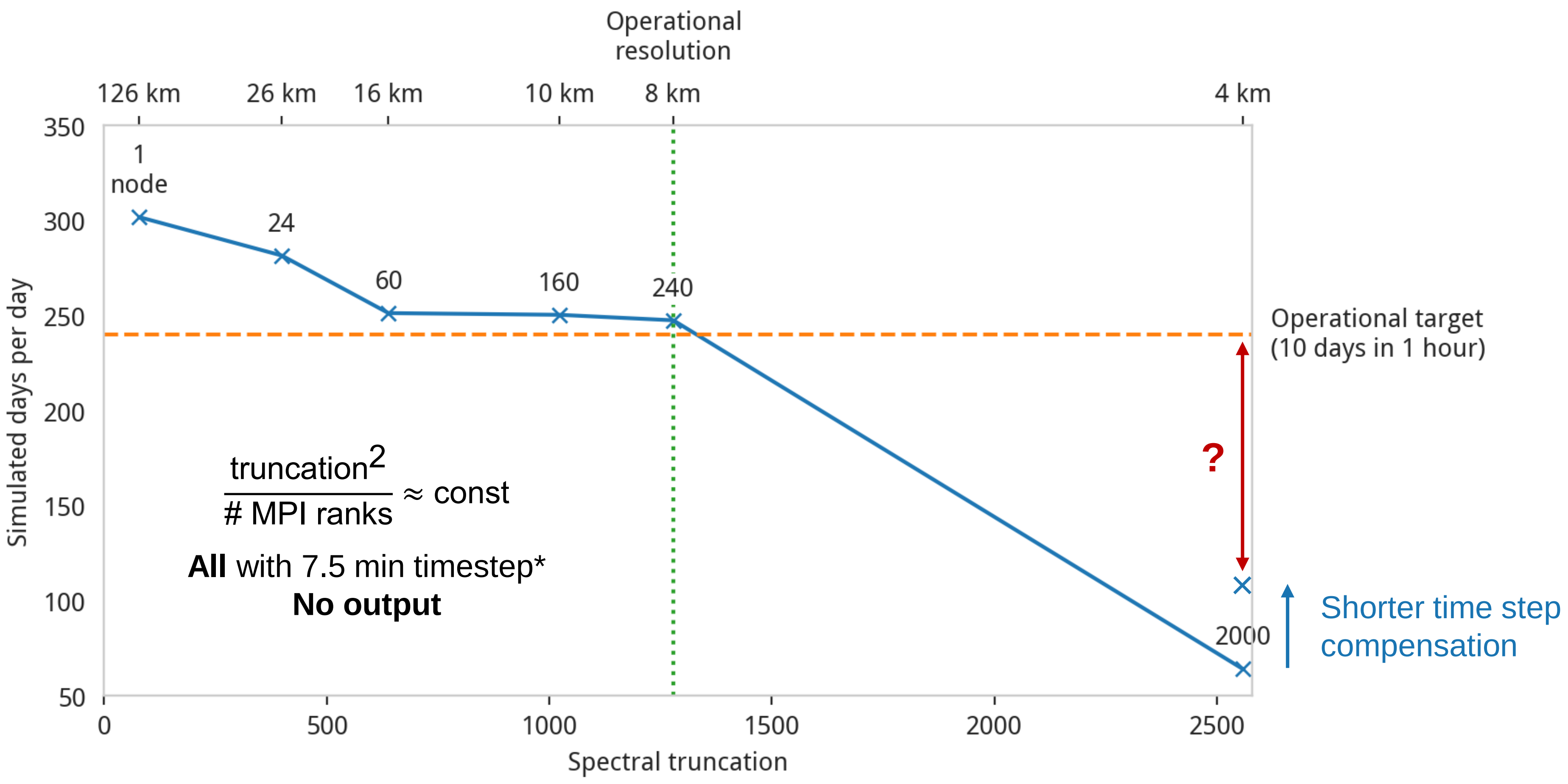
$$\frac{\text{truncation}^2}{\text{\# MPI ranks}} \approx \text{const}$$

All with 7.5 min timestep*
No output

Shorter time step compensation

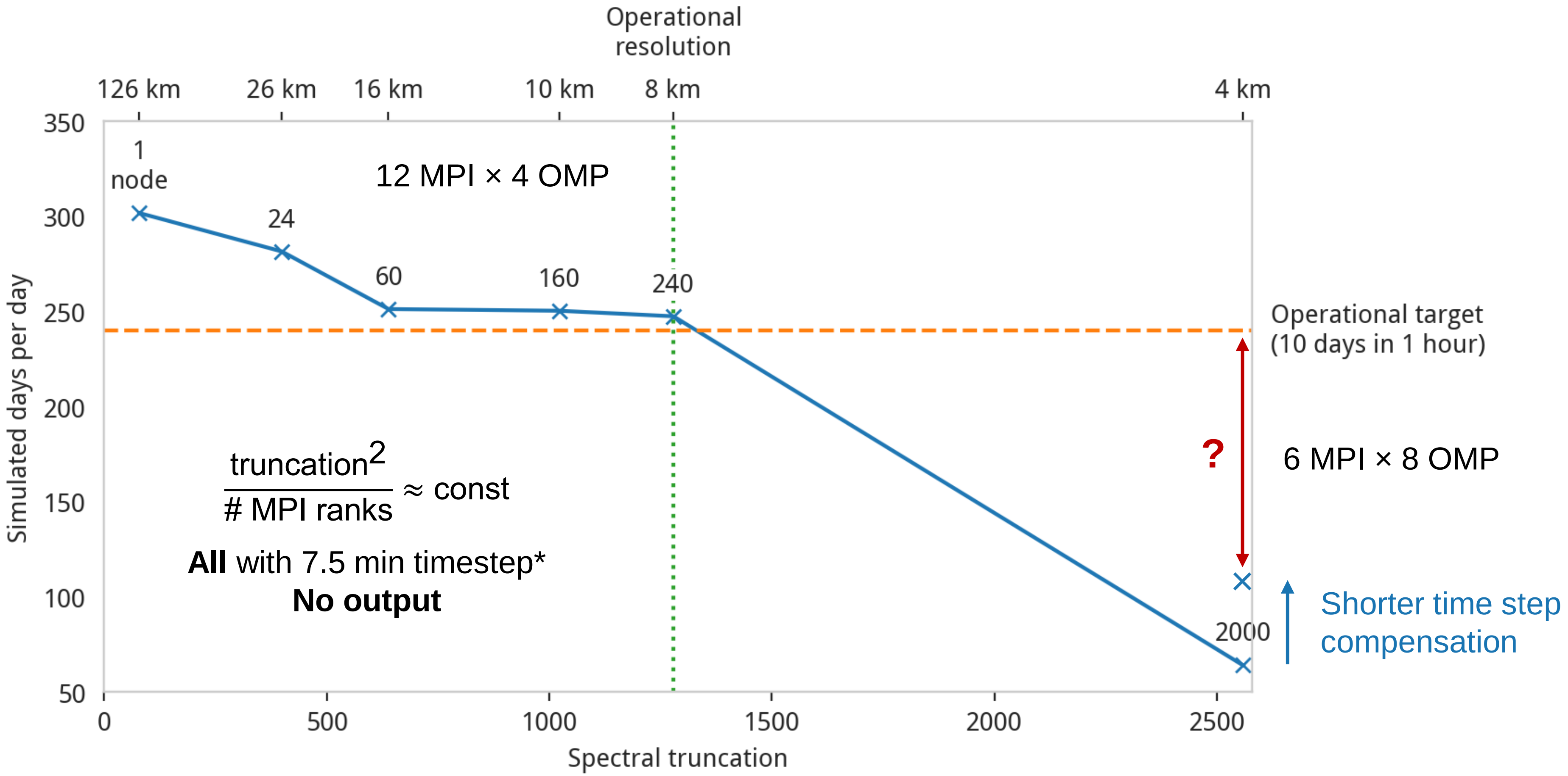
**except TCO2559/4 km → timestep 4 min*

Weak scalability on Fugaku



**except TCO2559/4 km → timestep 4 min*

Weak scalability on Fugaku



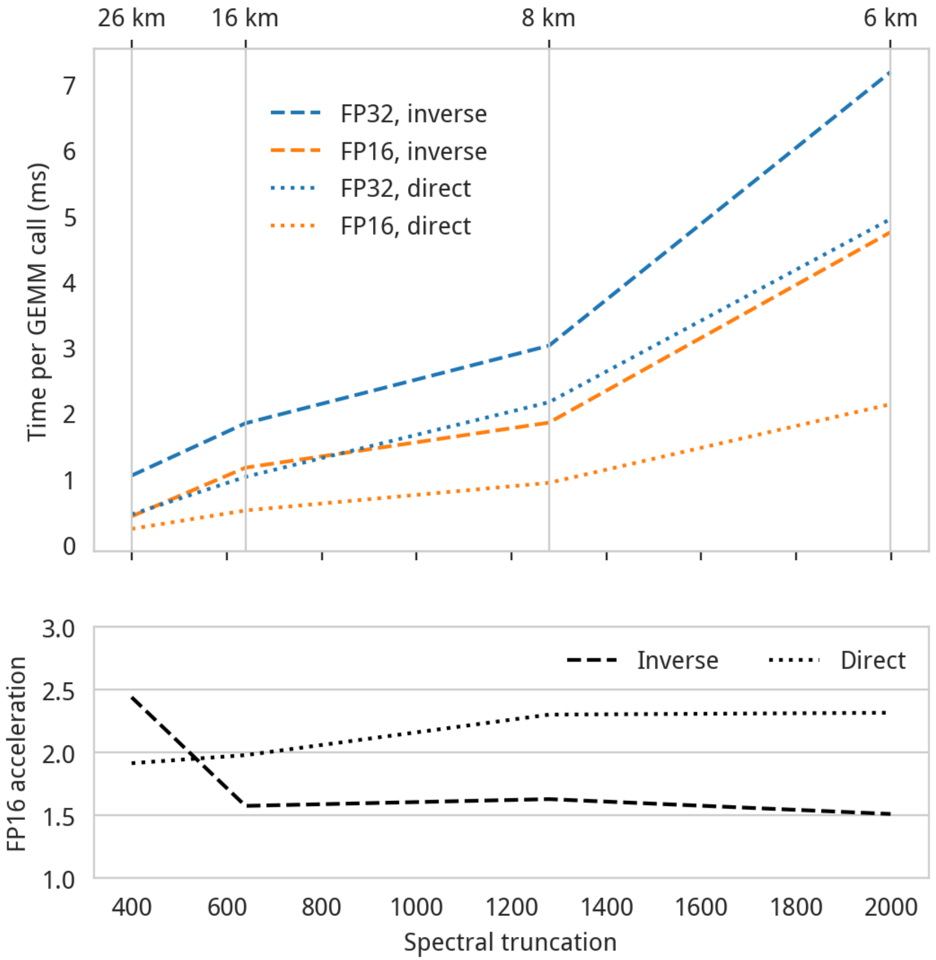
**except TCO2559/4 km → timestep 4 min*

Half-precision Legendre transforms

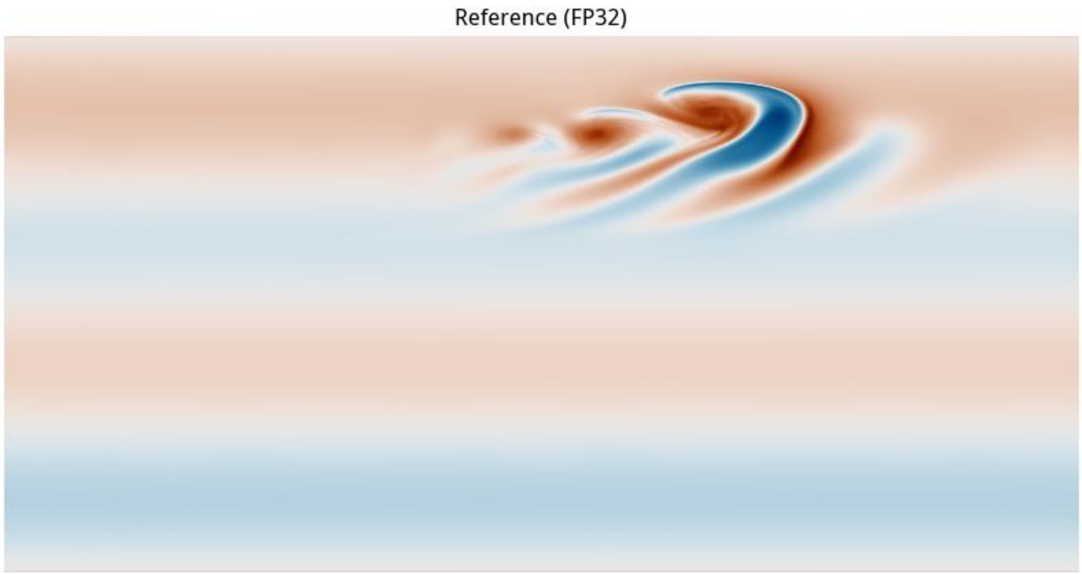
```
CALL SGEMM('N','N’, ...)
```

↓

```
REAL(KIND=2) :: ZBA(:,:)
...
CALL FJBLAS_GEMM_R16('N','N’, ...)
```

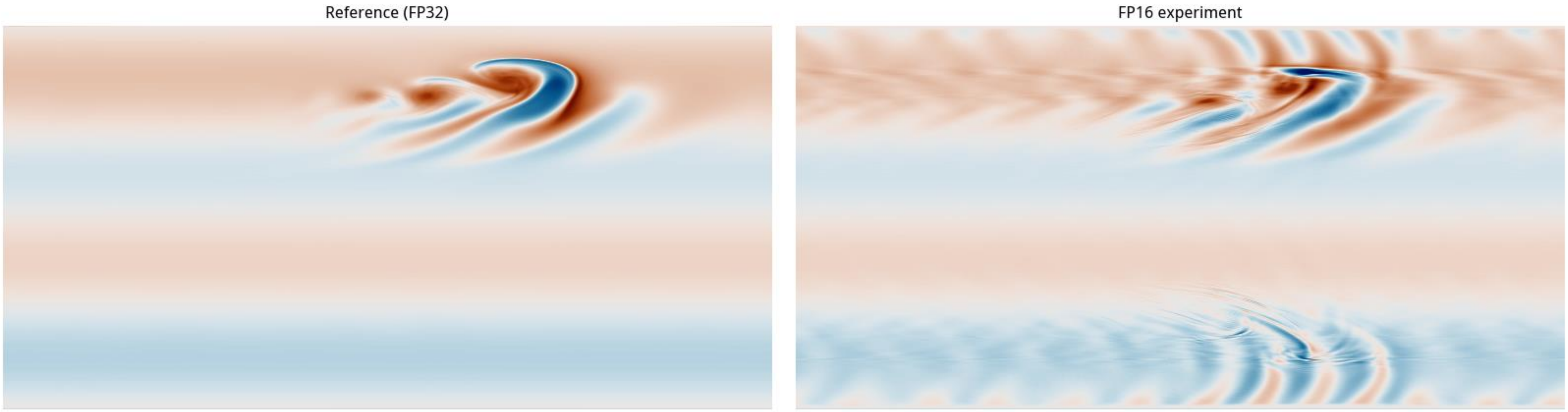


Half-precision Legendre transforms in the IFS



- Baroclinic wave test case
- 500 hPa vorticity after 10 days, TCO399L137 resolution (~25 km)

Half-precision Legendre transforms in the IFS



- Baroclinic wave test case
- 500 hPa vorticity after 10 days, TCO399L137 resolution (~25 km)

Future work

Scaling up

- Continue scaling: TCO3999 (2.5 km), TCO7999 (1.25 km)
- Direct comparison with Summit
- (Budget permitting) High-resolution coupled forecast

Half precision

- Keep debugging 😄