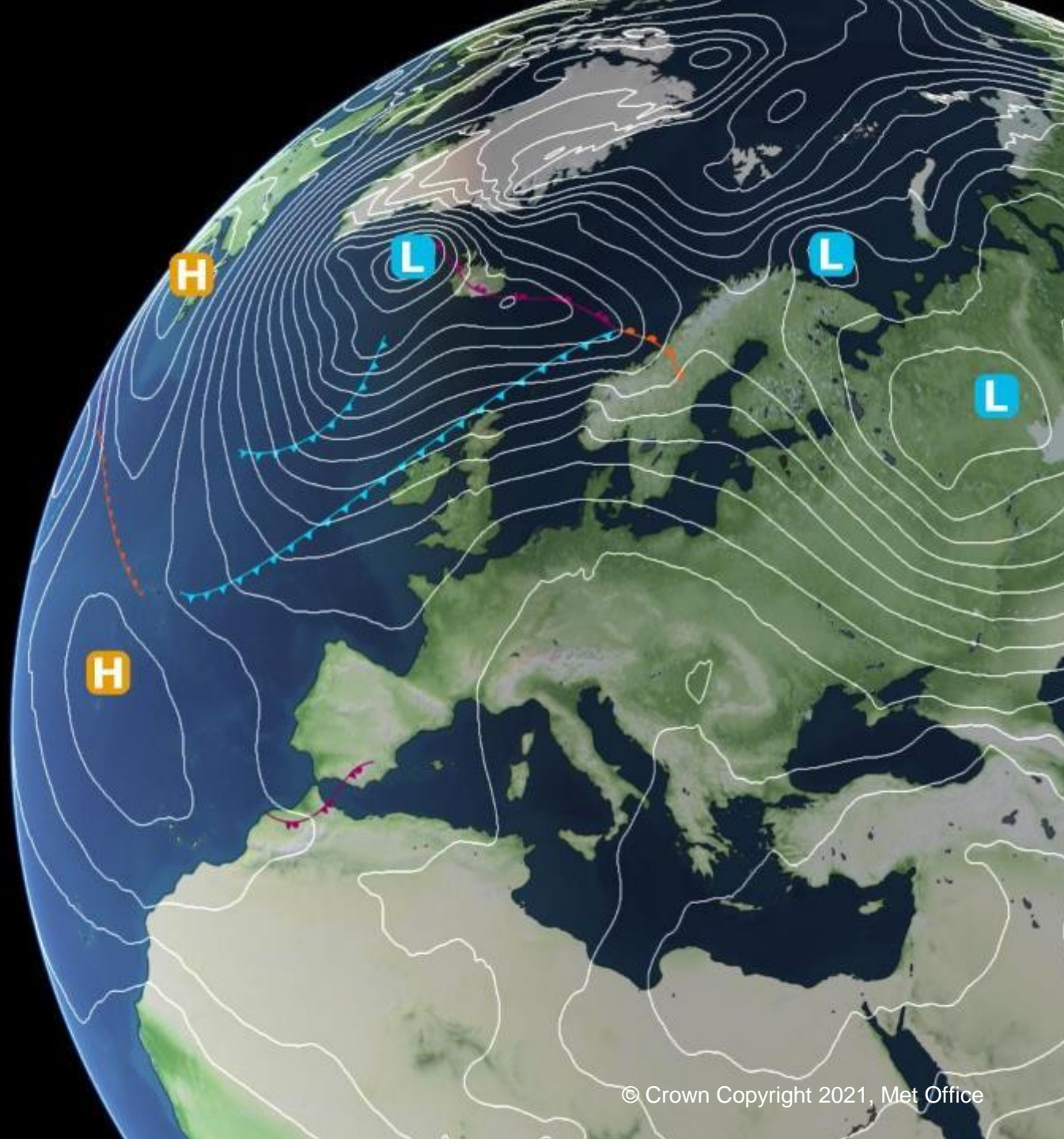


LFRic an update

Andrew Coughtrie, Met Office, UK &
*LFRic Science and Infrastructure
developers (Met Office, UK), PSyclons
(STFC, BOM, MO, ...), MO partners ...*



Introduction



- **LFRic** (*after **Lewis Fry Richardson***) is the new weather and climate modelling system being developed to replace the existing Unified Model in preparation for exascale computing in the 2020s

- Uses the **GungHo dynamical core**
- Runs on a **semi-structured, cubed-sphere mesh**



- **PSyclone** is a domain-specific compiler and source-to-source translator developed for use in finite element, finite volume and finite difference codes
 - Uses the **information** from a supported **API**
 - **Generates code** exploiting different **parallel programming models**

(R)Evolution in preparation for exascale:
From UM/ENDGame to LFRic/GungHo



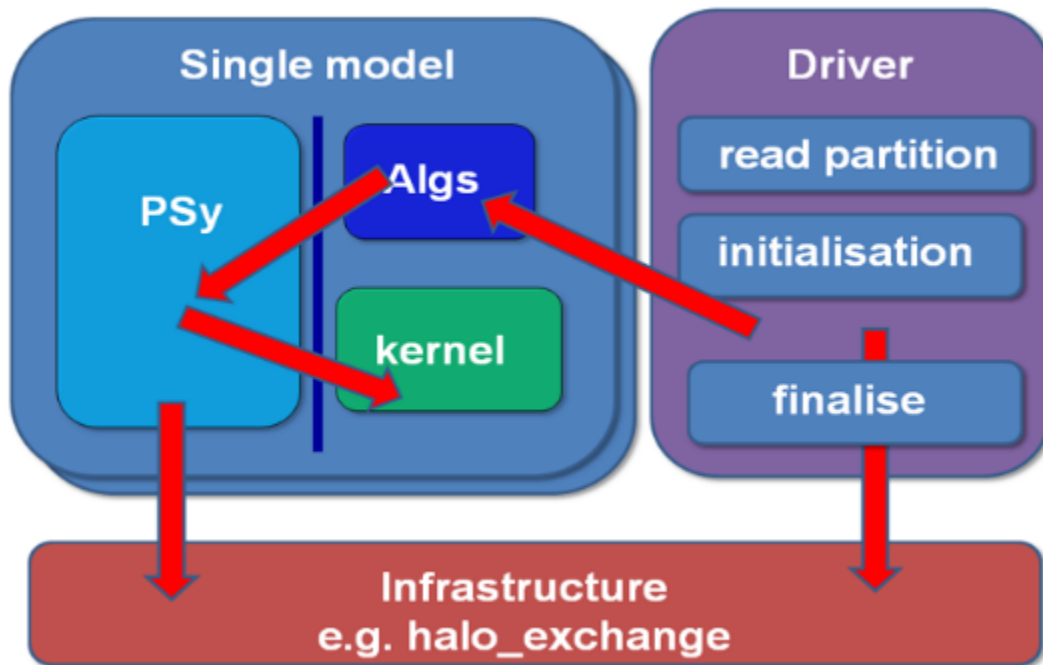
Lon-lat grid (Poles) *structured*
Finite-difference
Hard-coded optimisations

Maintain accuracy
→
Improve scalability
Exploit other programming
models



Cubed-sphere mesh – unstructured
Mixed finite element method
Generated optimisations

PSyKAI Separations of Concerns



- **Algorithms** (Natural Science: operations on *whole field objects*)
- **Parallel-Systems** (Computational Science: accesses field data and applies optimisations – *generated code*)
- **Kernels** (Natural Science: operations on *data points*)

Computational Science: Generated parallel code (STFC “PSy” + clone)

- **Optimisations:** generates optimised PSy-layer code
- *(work in progress on transformed kernels)*
- Optimisations encoded as a ‘recipe’ rather than baked into the scientific source code.
- Different recipes for different architectures.
- **Development**
- Generates **kernel stubs** (argument declarations and ordering).
- **fparser2** (F2003-2008): base for the LFRic code style checker (**stylist**).
- **Tools (profiling, DataAPI)**
- Insert calls to profiling tools (interface in PSyclone) – tested (and used) in LFRic.
- Extract data for running smaller code units as stand-alone applications (microbenchmarks) – work in progress.



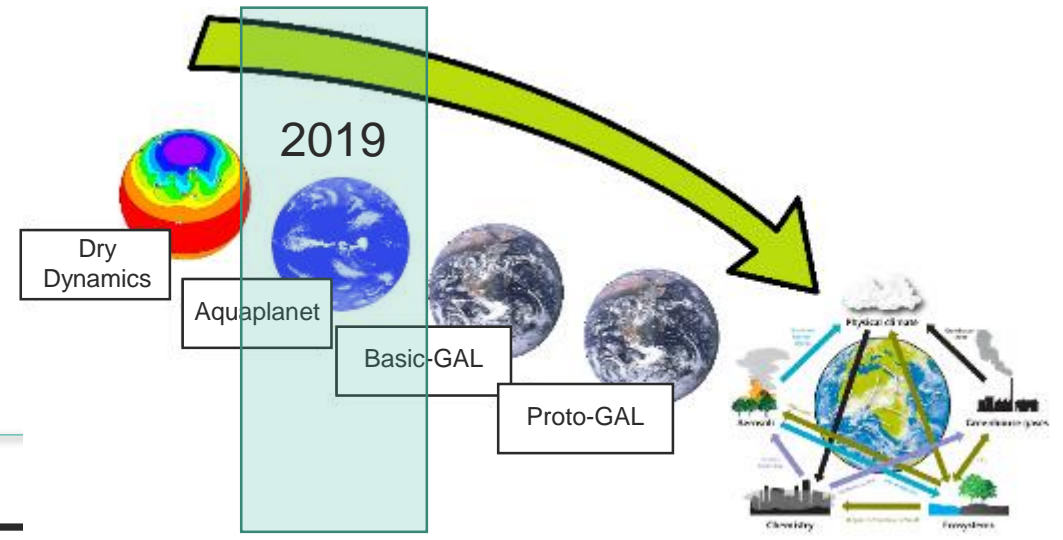
GungHo: capabilities

Simple physics & land surface (GAL

Model top 40km (38 levels - stretched)

Restricted timestep (Courant < 1)

Fixed SSTs



GungHo: capabilities

Full physics & land surface (GAL7)

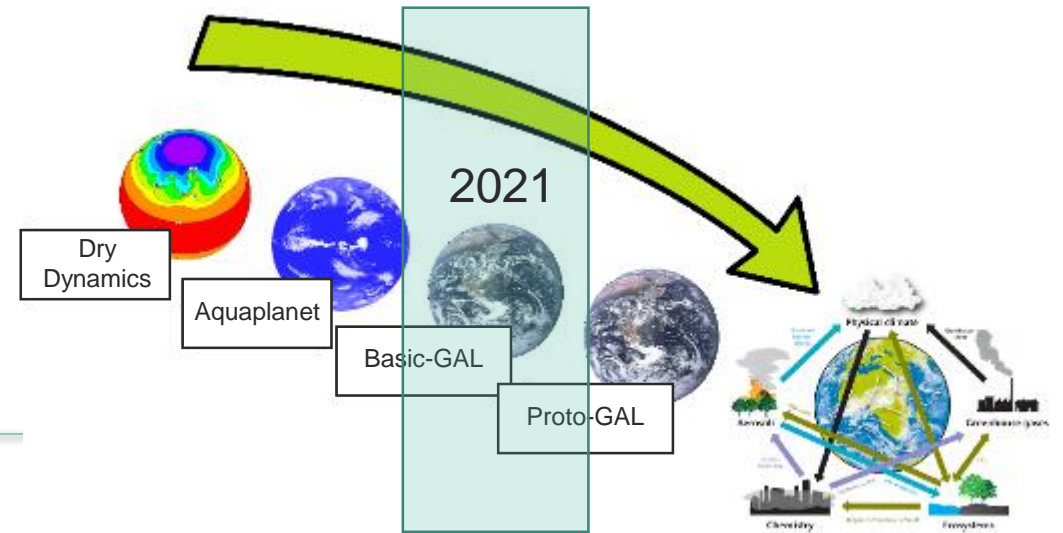
Model top 80km (70 levels - stretched)

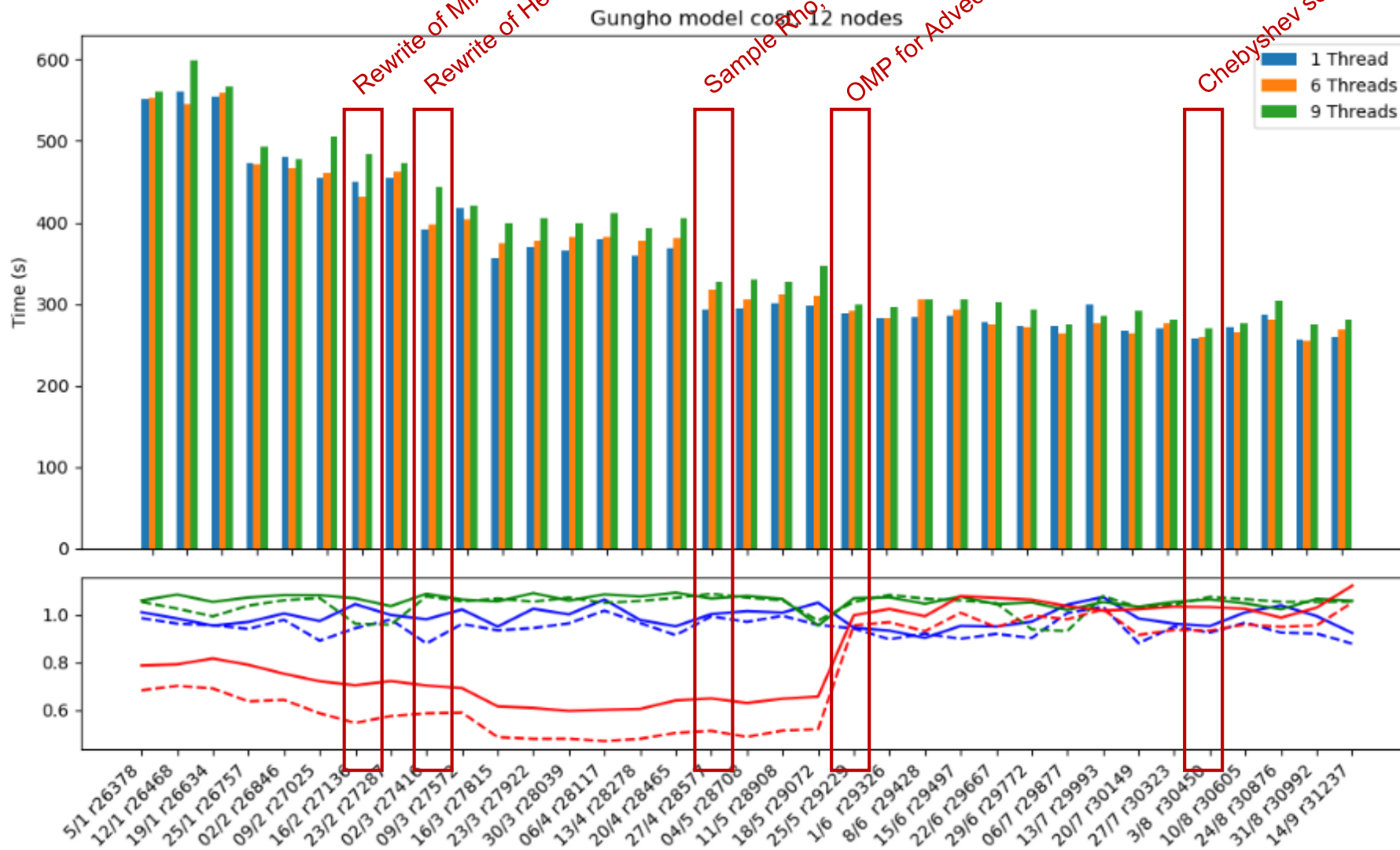
Long timestep capability (Courant $\gg 1$)

Surface orography

5 day C192 and 1 year C48 demo runs

'Credible' simulations





Gungho dynamical core, Held-Suarez test case:

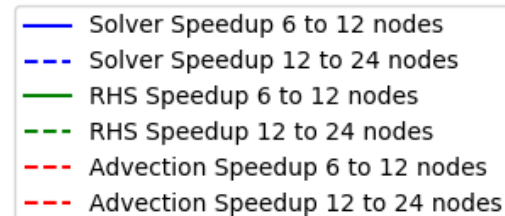
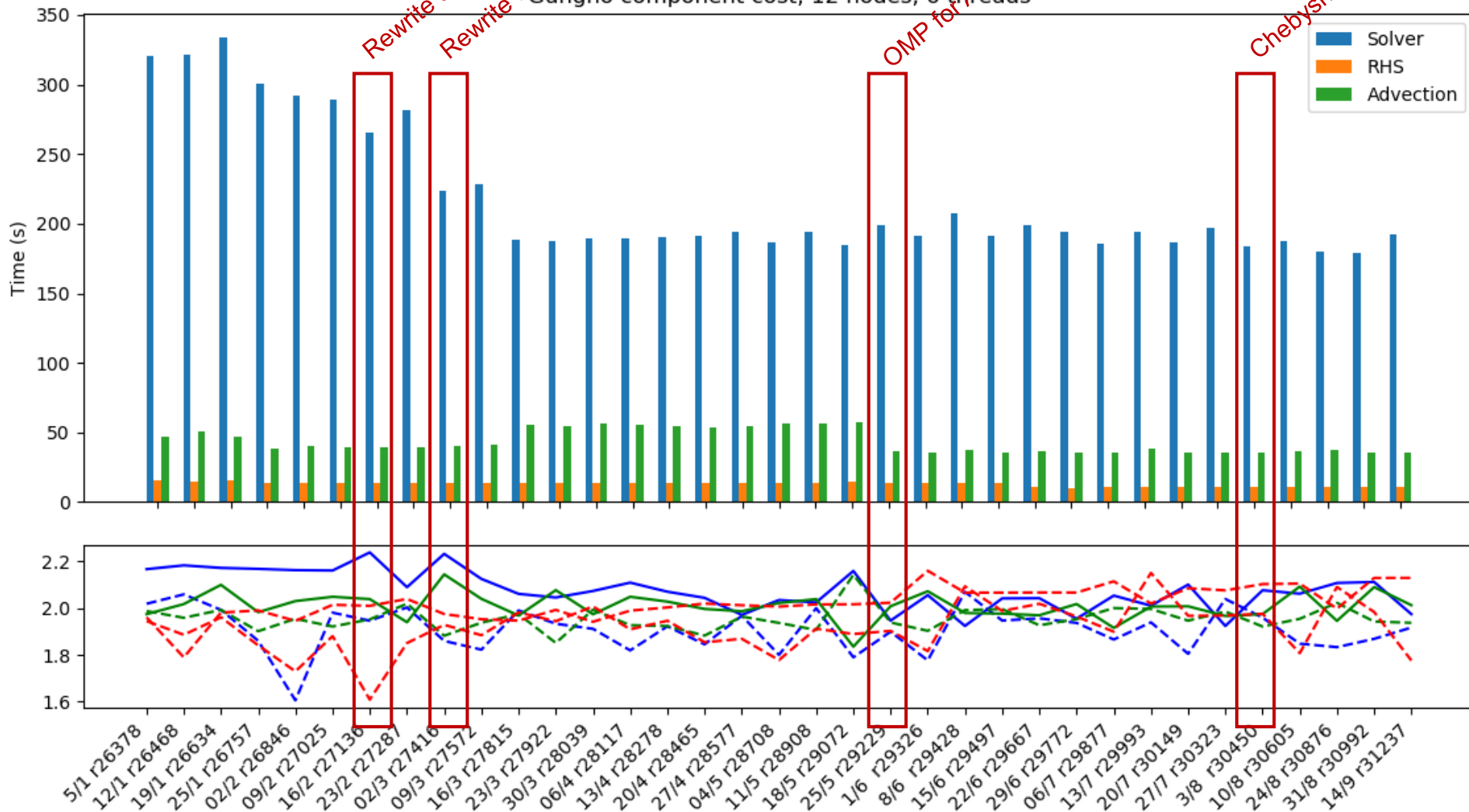
- C192 MG (6×192^2 columns ≈ 50 km horizontal)
- L30 DCMIP vert
- $dt = 1200$ s (SI)
- Local volume: 32×16 (1 OMP), 64×48 (6 OMP), 96×48 (9 OMP)
- Cray XC40, 2x18-core Broadwell

Rewrite of Mixed Operator
Rewrite of Helmholtz Operator

OMP for Advection

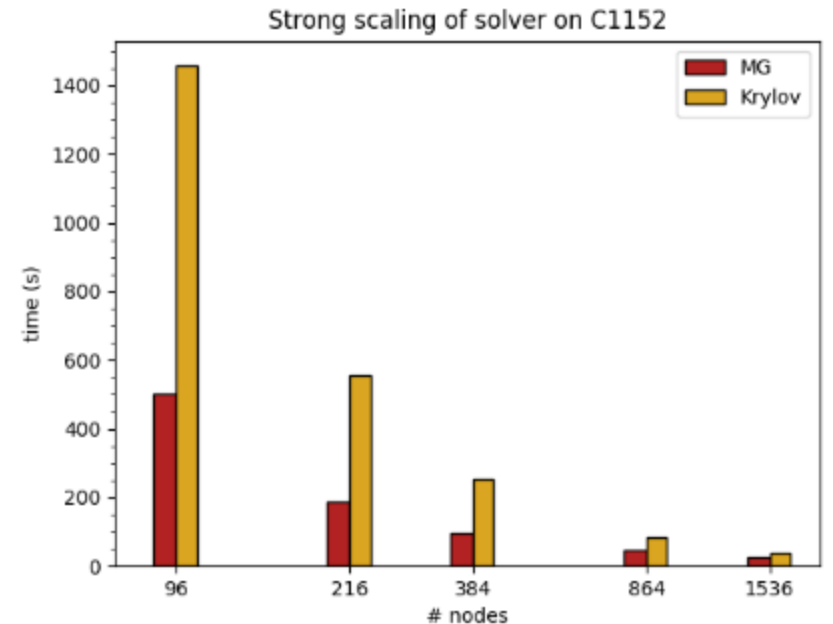
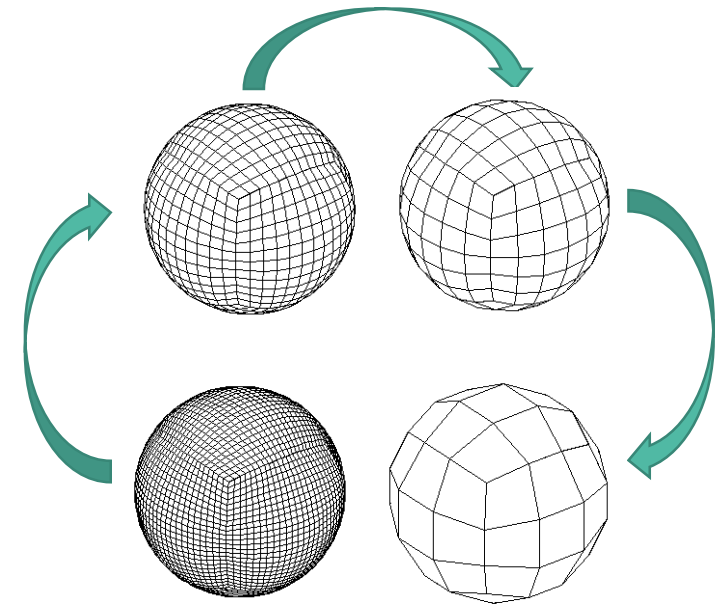
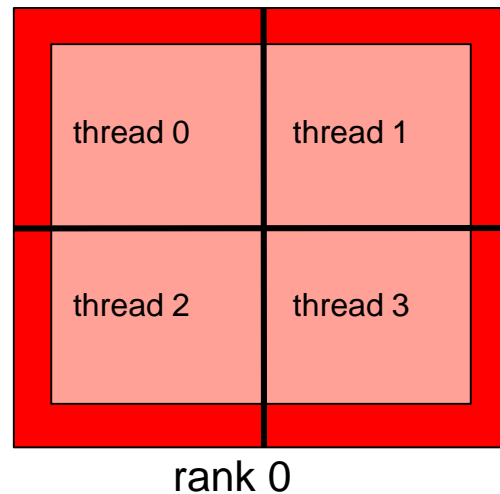
Chebyshev solver

Gungho component cost, 12 nodes, 6 threads



Reducing communication

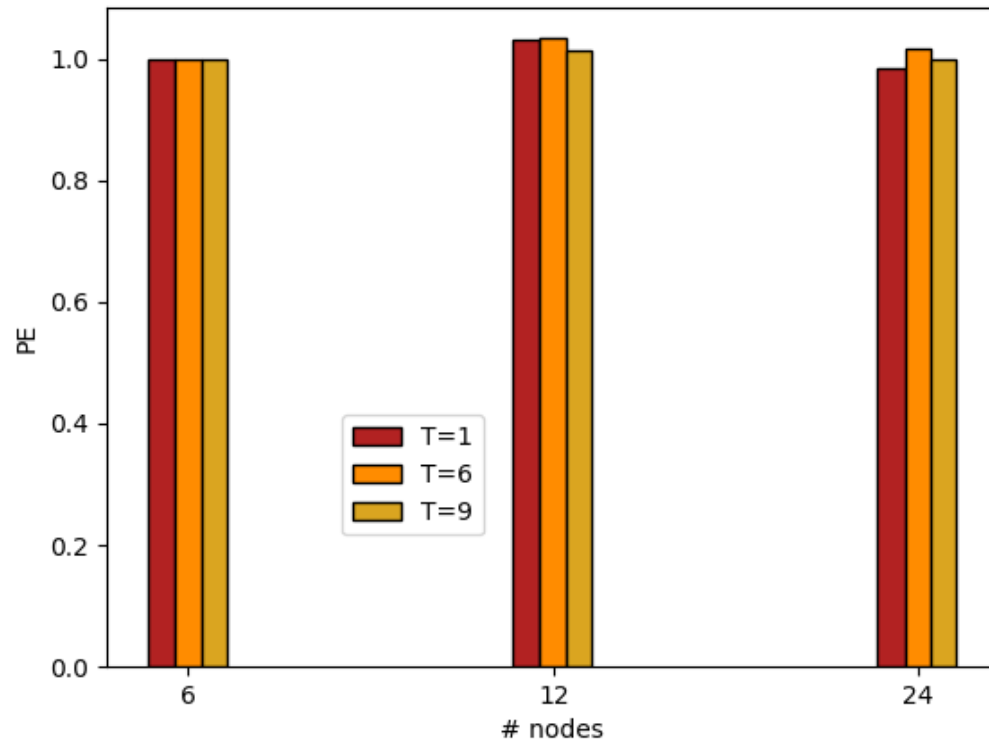
- Replacing Krylov subspace solver (BiCGstab) with Multigrid (reducing the number of global sums required)
- OpenMP - increasing parallelisation while decreasing communication



OpenMP scaling in Advection

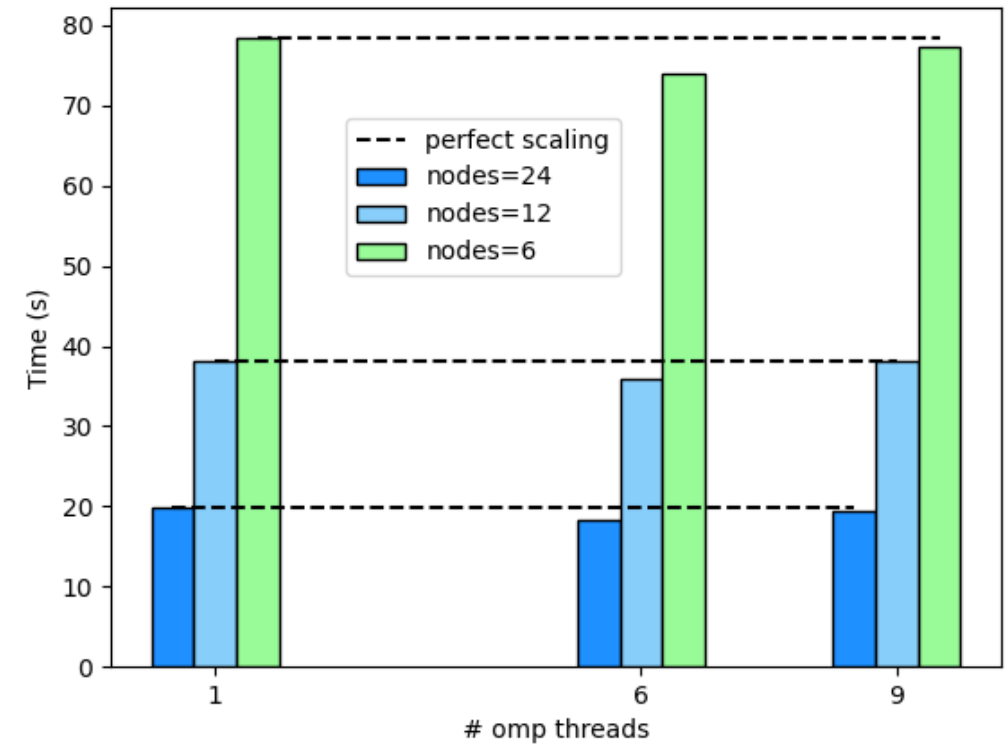
Equal is best

Node scaling



Lower is better

Thread scaling



Redundant Computation

- **Dynamical core run, baroclinic test case** with four multigrid resolutions and 30 vertical levels.
- Cray XC40/XCS, Aries interconnect, dual socket 18-core Broadwell Intel Xeon node, i.e. 36 cores per node.
- PSystem configurations: **ANY_SPACE** (general function space, assumes continuity); **ANY_DISCONTINUOUS_SPACE** (no shared dofs); **AD** and **Redundant Computation**.
- Columns per task: 1536 for the base MPI run and 768 for 2 x MPI tasks.
- Reduction in number of halo exchanges: **≈ 10% for AD** (no significant change in runtimes) and **≈ 57% for AD+RC**.

MG Resolution	MPI tasks in base run	AD+RC (MPI)	AD+RC (1/2 x MPI, 2 OMP)	AD+RC (2 x MPI)	AD+RC (MPI, 2 OMP)
C48 (≈ 200 km)	18	2.31	NotMul6	2.91	3.71
C96 (≈ 100 km)	72	1.50	2.89	4.75	3.26
C192 (≈ 50 km)	288	2.34	3.58	5.24	6.96
C384 (≈ 25 km)	1152	7.60	3.11	SubFail	5.78

Speed-up (%) compared to AC runtimes for each configuration.

Other On-going Development

- **GPU Acceleration**

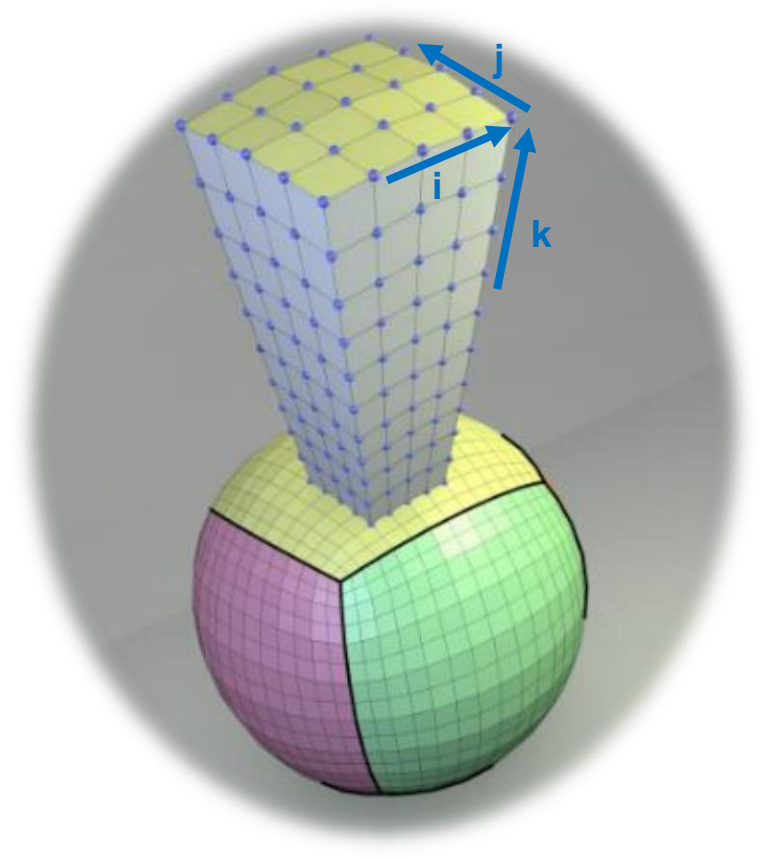
- Currently requires hand editing of kernels to add OpenACC directives
- Partially generated PSy-layer code.
- Need for copy on CPU before sending to GPU

- **I-First**

- Most UM Physics schemes written to be performed layer by layer currently
- LFRic Column based

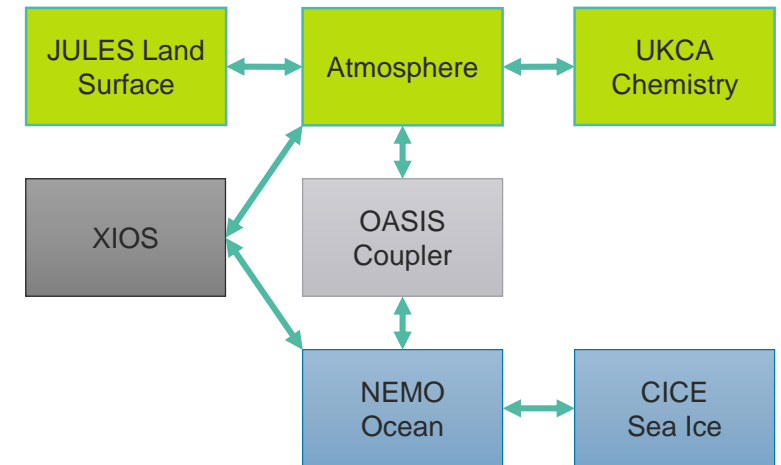
- **Offline Partitioning**

- For large domains start-up time expensive and very memory intensive currently



Other On-going Development

- **Mixed Precision**
 - Use of some form of templating required
- **Task Parallelism**
 - PSyclone investigation into task parallelism within core model (Rupert Ford, STFC)
- **Hybridisation solver**
- **Data Assimilation coupling**
 - Coupling LFRic models with JCSDA JEDI data assimilation software
- **Adjoint linear model**
 - PSyclone optimisations
- **OASIS coupling**
 - Coupling to external applications



Thanks & Contributions

LFRic Infrastructure

Steve Mullerworth (Project Manager)

Mike Hobson

Matthew Hambley

Iva Kavcic

Chris Maynard

Ed Hone

Ricky Wong

Andrew Coughtrie

GHASP

Ben Shipway

Tom Melvin

Tom Bendall

Christine Johnson

James Kent

Alex Brown

Mohamed Zerroukat

STFC Hartree Centre (PSyclone)

Rupert Ford

Andy Porter

Sergi Siso

Aidan Chalk

Australian Bureau of Meteorology

Joerg Henrichs (PSyclone)

And Many Many More....

Questions?