The ECMWF forecast model, quo vadis?
by Nils Wedi
European Centre for Medium-Range Weather Forecasts
wedi@ecmwf.int

contributors: Piotr Smolarkiewicz, Mats Hamrud, George Mozdzynski,
Sylvie Malardel, Christian Kuehnlein, Willem Deconinck, Jean Bidlot, Linus Magnusson
Orography – T1279 (16km)

Alps
Orography - T7999 (2.5km)

Alps
IFS Evolution

The energy cost of computing and affordability!

accelerators
4 foci:
1. Enabling technology for computations on massively parallel computers for science today
2. A cost-effective, low-energy consuming but highly accurate model and data assimilation system
3. Code resilience and understanding, what does each part of the model contribute, where is accuracy required, where not?
4. Quantitative measures on how good the forecast is

GPUs, Accelerators?!

Energy consumption?!
10 km IFS model scaling on TITAN (CRESTA project)

NO USE OF GPU

dissemination requirement
5 km IFS model scaling on TITAN (CRESTA project)

Successful overlapping of communication and computation!

NO USE OF GPU
The “cost” of communication:

- Assume the following:
  - model time step of 30 seconds
  - 10 day forecast
  - model on 4M cores
  - max 1 hour wall clock
- 1 step needs to run in under 0.125 seconds
- Using 32 threads per task with 128,000 MPI tasks
- A simple MPI_SEND from 1 task to all other 128K tasks will take an estimated 128,000 x 1 microsec = 0.128 seconds

- Implies global communications cannot be used, likely
- each task needs to run with 100’s or 1000’s of threads or equivalent
  => aim for max O(10,000) MPI tasks
PantaRhei + CRESTA + Loughborough University

- Flexible unstructured mesh data structure with options to retain reduced Gaussian grid nodes
- Change of equations
- Compact stencil, minimizing communication and data movement
- Fast route to competitive real weather simulations
10km/L137 dynamical core comparison

- Ideal
- PantaRhei
- IFS T₀ 2047
- IFS T₀ 1023

X3C0

# cores vs. speedup (log-log scale)

# cores vs. efficiency (log-log scale)
10km/L137 dynamical core comparison

- IFS T$_l$ 2047
- IFS T$_c$ 1023

Forecast Days / Day vs. # cores

XC30
Underlying physics

- *Multiscale nature* of the atmosphere may be exploited to extract further parallelism.
- Further resolution increases coincide with two major changes in how we model the physical system, with far reaching consequences for numerical algorithms and equations used:
  - *Resolved convection*
  - *Non-hydrostatic effects*
- This furthers the need to assess and *quantify forecast uncertainty* but also provides opportunities for new algorithms and/or hardware design.
  - E.g. single precision or low-accuracy computations
Berkeley Dwarfs and ECMWF’s IFS
Recognizing patterns of computation and communication

♦ Dense Linear Algebra (matrix/matrix multiply) 😊
♦ Sparse Linear Algebra (Newton – Krylov solver) 😊
♦ Spectral methods 😊
♦ N-body methods (FMM) 😊
♦ Structured grids (data layout) 😊
♦ Unstructured grids (data layout) 😊
♦ MapReduce (I/O, data processing)

♦ New: combinational logic, graph traversal (indirect addressing), dynamic programming (parallel-in-time algorithms), branch-and-bound, graphical models, finite-state machines (event driven) …..
Titan (Cray), Oakridge, US, No 2 in the world (Top 500, Nov 2013)

~8.2MW

299,008 cores AND 18,688 NVIDIA GPUs

♦ Need to explore hybrid architectures
♦ Accelerator adaptation of IFS key components
Warning the public and saving lives
Wave height 72h forecast, T3999 (~5km)
Efficient coupling to surface, boundary layer, ocean, waves, …
Tianhe-2 (MilkyWay-2)
No 1 in the world (Top 500, Nov 2013)

This computer has 3,120,000 cores, with Intel Xeon Phi co-processor technology

A 50 member ensemble at ~5km may need this number of cores in order to run in 1 hour.
Energy-aware computing

♦ ECMWF uses the equivalent energy comparable to the annual consumption of ~8000 4-bedroom houses!

♦ 51 ENS members consume about 330KWh, approximately the same as a single (~5km) global 10-day forecast

♦ Today the energy consumption of one ENS member is equivalent to leaving the Kettle on for 3 hours!

http://ukbusinessblog.co.uk
Preparing for the future means for us ...

- Flexibility on the equations solved
- Flexibility on the numerical algorithms used
- Flexibility on the horizontal and vertical discretization used
- Options for the data layout to adapt to massively-parallel, heterogeneous computing architectures
- Reduce communication requirements
- Develop strategies for resilience
- Limit the Mega Watts used per forecast produced!