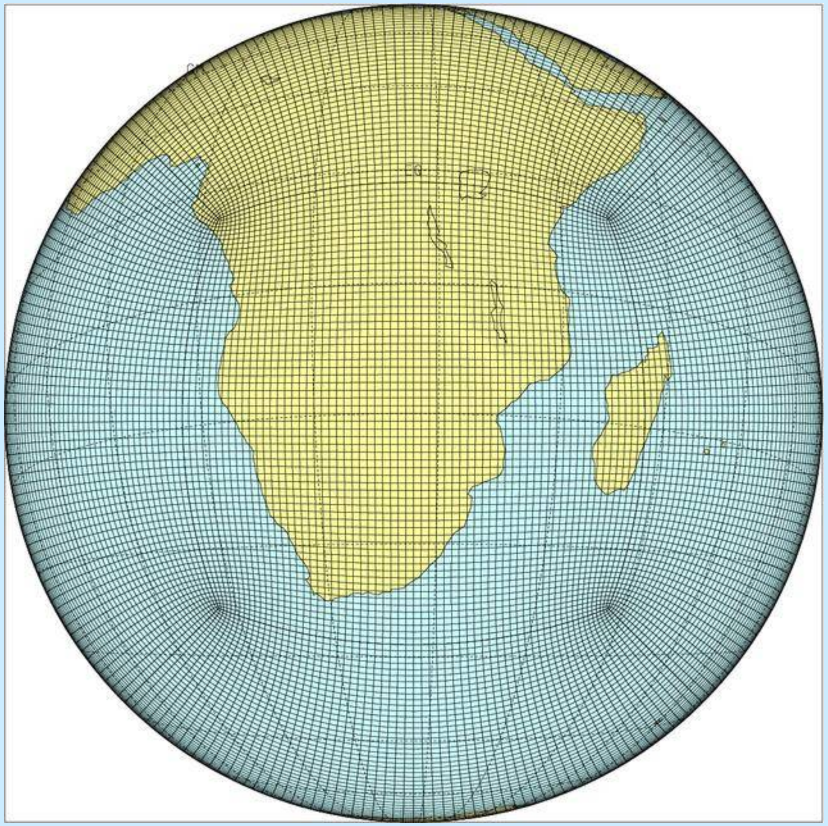
WGRC Expert meeting on Climate Information Distillation, Santander 2014

Some notes from the meeting and its relevance to WGNE

Francois Engelbrecht



CCAM C64 stretched-grid with resolution about 50 km over southern Africa

- Increased availability of projections of future climate change for climate change adaptation studies
- CMIP3&CMIP5 GCM projections
- Regional climate projections (CORDEX – Coordinated Regional Downscaling Experiment)
- Projections from statistical/empirical downscalings
- Multiple sources of evidence… end users experience this as “a sea of uncertainty”
- There needs to be a move from uncertainty to probability

RCM value-added contributions

Focus areas for regional climate modelling:

- Fine-scale processes and attributes, e.g. land heterogeneity, convective storms.
- Coupling of regional processes and feedbacks, e.g. ocean currents, upwelling, land-surface and aerosols
- Integrated suite of products/diagnostics within the context of the above.
- Coupled natural and human systems (e.g. a focus on mega cities)

RCM development aspects

Verification metrics

- Should be increasingly process-based

- Recognised as regionally dependent/ informed

- User-informed and user-informing

Testing and developing RCMs

- RCMs critisized as being tuned “to get today right”
- Representation of present-day climate (means and extremes)
- Representation of trends
- Representation of paleo-climate
- RCM CORDEX Phase I resolution correspond to that of operational global NWP
- RCM CORDEX Phase II resolutions to be beyond hydrostatic limit – corresponding to operational regional NWP
- RCM model development community may greatly benefit from NWP model development efforts (e.g. grey zone project, gravity wave drag) – potential contributions/interaction with WGNE were noted

ECMWF Scalability Workshop
ECMWF Reading, 14-15 April 2014

Some key thoughts from the working group summary report (Bauer et al., 2014)

- Of relevance: design of the next generation forecasting systems for future exa-scale HPC and data-management structures

- Typically this will be a partnership between an NWP centre, HPC centre and hardware providers

- Increasing alignment with the climate modelling community: The general development towards Earth System Modelling at high resolution for both NWP and climate modelling needs to deal with scalability and operability limits at NWP and climate centres - approach of maintaining a single model and data assimilation system (IFS/seamless system) for all applications vs promoting separate components tailored to forecast range and application.

- urgency of adaptation to highly parallel computing is different for each component of the forecasting system, namely data assimilation, forecasting and data post-processing/archiving.

Scalability: hard limits

Some key thoughts from the working group summary report (Bauer et al., 2014)

- Computing: the key figure is the electric power consumption per floating point operation per second (Watts/FLOP/s).
- I/O: the key figure is the absolute data volume to archive and the bandwidth available for transferring the data to the archive during production, and dissemination to multiple users.
- Both of the above are subject to hard limits, that is cost of power, network capacity and storage capacity.
- to what extent can fundamentally new scientific and technical methods and future HPC reduce current limitations?

Scalability: Status quo

Some key thoughts from the working group summary report (Bauer et al., 2014)

- Status quo: Despite ambitious targets being set for model resolution, complexity and ensemble size, today the bulk of the calculations are not performed with configurations that utilize the maximum possible number of processors. Data assimilation, extended range prediction and research experimentation mostly operate at relatively lower resolutions predominantly for technical and affordability reasons. However, the forecast suites always contain a cutting-edge component that fully exploits capabilities.

Scalability: assimilation and dissemination

Some key thoughts from the working group summary report (Bauer et al., 2014)

- Two main data assimilation development streams are being pursued at operational NWP centres, namely long-window 4D-Var and EnVar. Both have scientific and technical advantages and disadvantages but 4D-Var reaches efficiency limits very soon. OOPS will facilitate testing of scalability options at ECMWF.
- I/O limitations will become effective in the short term, for example linked to data bandwidth not growing at the same rate as computing, and data dissemination becoming impossible for large productions in NWP and climate.
- I/O optimization and the scientific component in data assimilation research should therefore assume very high priority in scalability research.
- For climate research and production the task of data assimilation for model initialization is only emerging now, but dissemination is of key importance

Scalability: numerical methods

Some key thoughts from the working group summary report (Bauer et al., 2014)

- More flexibility is needed and the development of shared libraries containing numerous kernels serving many different applications is becoming important.
- Advantage of the shared approach is that libraries can be co-developed with HPC-centres and vendors and thus optimization for specific architectures is greatly facilitated
- To what extent should libraries be interfaced with FORTRAN?
- Algorithms that are compute intensive but require little data communication are preferred. Examples are spectral element, finite element and discontinuous Galerkin methods – attractive for implementation on accelerators.

Scalability: numerical methods

Some key thoughts from the working group summary report (Bauer et al., 2014)

- More flexibility is needed and the development of shared libraries containing numerous kernels serving many different applications is becoming important.
- Advantage of the shared approach is that libraries can be co-developed with HPC-centres and vendors and thus optimization for specific architectures is greatly facilitated
- To what extent should libraries be interfaced with FORTRAN?
- Algorithms that are compute intensive but require little data communication are preferred. Examples are spectral element, finite element and discontinuous Galerkin methods – attractive for implemetation on accelerators.

Scalability: hardware/compilers

Some key thoughts from the working group summary report (Bauer et al., 2014)

- Given the expected stagnation of processor frequency and memory the main effort needs to be spent on distributed (load balanced) calculations with as much overlap between computing and communication as possible. This requires, most likely, hybrid architectures comprising CPUs and accelerators aiming at optimizing performance against energy consumption.

- It was demonstrated that energy efficiency gains are not easy to achieve with hybrid architectures since CPUs consume 80% of their peak energy when idle and fully exploiting accelerator capacity requires optimal compute-communication configurations.

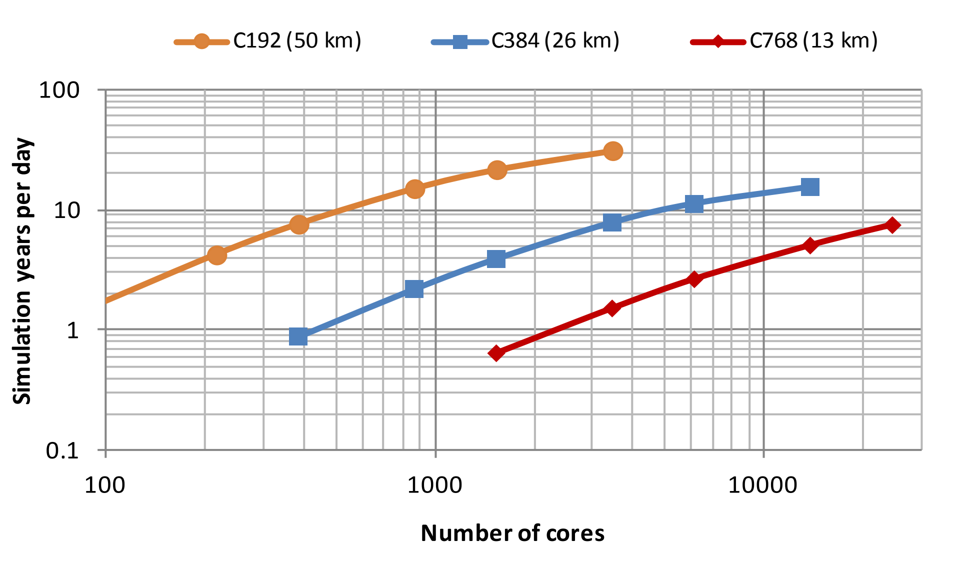
- The most important starting point for efficiency gains is a rigorous and continuous optimization, for example, with MPI/OpenMP, better vectorization, employing more flexible data structures, and cache optimization.

- It was recognized that bit-reproducibility for a fixed processor configuration is of crucial importance for code debugging, and the only means for distinguishing between hardware differences and code issues.

Scalability: Areas of collaboration (and WGNE's potential role)

Some key thoughts from the working group summary report (Bauer et al., 2014)

- Sharing of libraries, work-flows (e.g. data streaming with minimum I/O has received relatively little attention) and efficiency monitoring tools; Sharing common developments between NWP and climate prediction communities; Other areas identified for collaboration were benchmarking, bit-reproducibility versus fault tolerance and common strategies for I/O.



CCAM-ocean scaling at the CSIRO

- WGNE can run regular survey's that list the status quo and main developments in scalability at different centres
- Further afield it can coordinate projects aimed at sharing libraries, workflows, etc
- More discussion is needed on WGNE's role in scalability studies