CUG 2007
NEW FRONTIERS

Implementing the Operational Weather Forecast Model of Meteo Swiss on a Cray XT3

Tricia Balle and Neil Stringfellow
(CSCS)

CUG 2007
NEW FRONTIERS

Outline

- Current operational model
- Future operational model and its advantages and challenges
- COSMO 2K benchmark
- Porting details and timings
- Planned machine upgrades
- Conclusions

Production moves from NEC SX-5 to Cray XT3

“MANNO, Switzerland and SEATTLE, WA, March 12, 2007 -- In February 2007, MeteoSwiss began production weather forecasting using a Cray (Nasdaq GM: CRAY) supercomputer located at the Swiss National Supercomputing Centre (CSCS)…Plans to implement higher-resolution forecasts in January 2008 will make detailed forecasts of Switzerland's intricate Alpine topography possible for the first time.”

Current Operational Model

- Two aLMo/7 (lower resolution) runs per day
  - Cover whole of western Europe
  - Each run a 72 hour simulation
  - Grid size 385x325 points, 45 atmospheric levels
  - Time step 40 seconds

- NEC SX-5 (14 dedicated cpus) 78 mins
- Cray XT3 (50 dual core nodes) 73 mins

(66 mins with added opts)

Future Operational Model

- Supplement current forecasts with 18-hour forecast suites every 3 hours:
  - Cover the Alpine arc
  - Two assimilation runs PLUS one 18-hour low res forecast and one 18-hour high res forecast PLUS necessary interpolations
  - Together with time-critical postprocessing, must take less than 30 minutes

Advantages of Future Model

- More frequent forecasts over smaller domain
- Ability to provide warnings
- National security enhancement
- Collaboration with international partners
- Ability to capture extra features
  - Better modeling of local topography
  - Added modeling of extra physical effects
  - Possibility of some ensemble forecasting

Challenges of Future Model

- Chaining of programs in suite increases time criticality of results
- Failure of one part of chain can delay entire operational cycle
- Interruption level for other users is high: suite is run every 3 hours for 30 minutes
- PLUS unscheduled on-demand usage:
  - LPDM run 1-2 times a week
  - Two 24-hour emergency runs a year to be launched within 5 minutes of request

COSMO 2K Suite (1)

- aLMo/7
  - 385x325 (or 285x225) point mesh
  - 60 levels, 72s timestep
  - BCs applied at 3 simulation hour intervals
  - Rayleigh damping at top; Runge--Kutta numerical solver
  - Radiation routine called every 60 simulated mins
  - GRIB output files produced every 60 mins
- aLMo/2
  - 520x350 point mesh
  - 60 levels, 18s timestep (15s for assimilation)
  - Radiation routine called every 30 simulated mins
  - GRIB output files produced every 30 mins

COSMO 2K Suite (2)

1. Interpolate IFS to aLMo/7 (385x325)
2. aLMo/7 assimilation 6 hours (385x325, 72s)
3. Interpolate IFS to aLMo/7 (285x225)
4. aLMo/7 short forecast 18 hrs (285x225, 72s)
5. Interpolate aLMo/7 to aLMo/2 (520x350)
6. aLMo/2 assimilation 3 hours (520x350, 15s)
7. aLMo/2 forecast 18 hours (520x350, 18s)

LM model: Detail

- RAPS version of LokalModell (LM) of COSMO
- Numerical weather prediction code used by several European national weather centers
- Pole placed at 32.5^{o}N (longitude -170^{o}) so virtual equator passes through Bern and region of interest shows little distortion
- Code can produce a variety of ascii output files for timing, diagnostics and verification
- Execution driven by Fortran namelists

Porting COSMO 2K to Cray XT3

- XT3 at CSCS:
  - 2.6Ghz 1664 node ~~single~~ core (3.3 terabytes mem)

    Now dual core!
  - 2.6Ghz 74 node dual core
- PGI compiler version 6.2.5 (-fastsse)
- GRIB library: remove elements not available on Catamount. Allow IOBUF buffering.
- Remove repeated closing and flushing of ascii output files
- Remove unnecessary target attributes
- Reduce communication with mpi_alltoallv call

Running on the Cray XT3

- ltime_barrier=.false.
- Use IOBUF for larger input/output GRIB files
- -small_pages
- 1MB lustre stripes over 1 or 2 OSTs
- Run with 4 dedicated I/O processors
- Default MPI rank ordering

COSMO 2K Suite…timings

1.	Interp IFS to aLMo/7	0:11.70
2.	aLMo/7 assimilation	1:57.44
3.	Interpolate IFS to aLMo/7	0:14.42
4.	aLMo/7 short forecast	2:18.91
5.	Interp aLMo/7 to aLMo/2	1:25.96
6.	aLMo/2 assimilation	3:26.25
7.	aLMo/2 forecast	17:09.75

Total time to solution: 26.7minutes

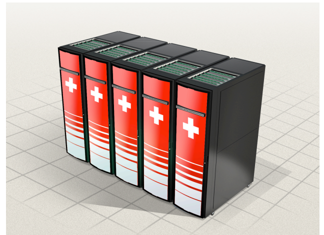
(26x25 + 4 I/O = 654 single core processors)

Single to dual core

- 862 dual core processors equivalent to 654 single core processors
- 20 to 25% penalty, mainly due to logic in fast waves Runge-Kutta section of code:
  - Introduced with high-resolution logic of LM model to support low gridpoint spacing
  - Heavily reliant on mem bandwidth and very cache unfriendly
- Needs a significant rewrite!

Moving from XT3 to XT4

- Problem: future operational suite demands 3 hourly 30 minute runs
  - Unacceptable to other users on system



- Solution: dedicated machine for Meteo Swiss runs
  - 5 cabinet Cray XT4 (2.6Ghz, 2GB compute nodes, 120 total blades)
  - To be installed summer 2007
  - Can meet time-to-solution requirements on 110 blades, leaving 5 blades for service and 5 spare
  - Could run postprocessing work on spare blades instead of offloading to separate machine

Conclusions

- Current MeteoSwiss operational suite successfully migrated from NEC SX-5 to Cray XT3
- Future high resolution forecasting suite to be run on dedicated Cray XT4 from January 2008

CUG 2007
NEW FRONTIERS

Thank you!

Roberto Ansaloni, Eckhard Tschirschnitz (Cray Inc.)

Neil Stringfellow, Mark Cheeseman, Angelo Mangili, Mauro Ballabio (CSCS)

Emanuele Zala, Guy de Morsier, Jean-Marie Bettems (MeteoSwiss)