MAX-PLANCK-INSTITUT
FÜR METEOROLOGIE

50th anniversary of ECMWF

Good ideas are persistent - pick them up!

Luis Kornblueh

Acknowlegements

David Dent, Nils Wedi, Clive Temperton, Tony Hollingworth, Clive Temperton, Norbert Kreitz, George Mozdzynski, Deborah Salmond, Tiago Quintino, Matthew Griffith, Paul Dando, Mats Hamrud, Florence Rabier, Klaus Arpe, Lennart Bengtsson, Ioan Hadade and an unknown group of CRAY engineers of the 80's.

Not to forget many more computing center and vendor staff!

Reminders

Response time on bug/performance compiler problem reports

CRAY,PVP	no access any more
Fujitsu,VPP	unknown - no bug found
Hitachi	2-4 weeks
NEC	1 day-4 week
DEC, Compaq	2-4 weeks
Fujitsu,WS	2 days
HP	$<$ 4 weeks
IBM	$<$ 4 weeks — 1.5 years
SGI	unknown - no bug found
SUN	unknown - too many bugs found
NAG (SUN)	2 days
Lahey/Fujitsu	2 days — 4 weeks
NAG (Linux)	2 days
INTEL (Linux)	unknown - too many bugs found

Technology

Hardware adaptive loop nesting

NPROMA

- (nlat, nlon) to (nproma, nblks)
- From IFS to ECHAM and later to ICON (using unstructured grids)
- supports CPUs, Vector and GPUs

Memory management

Dynamic memoray handling with meta data (provenance)

- Organized in traversable data structures
- GRIB2 tables and CF-convention (NetCDF/HDF5)

Model time handling

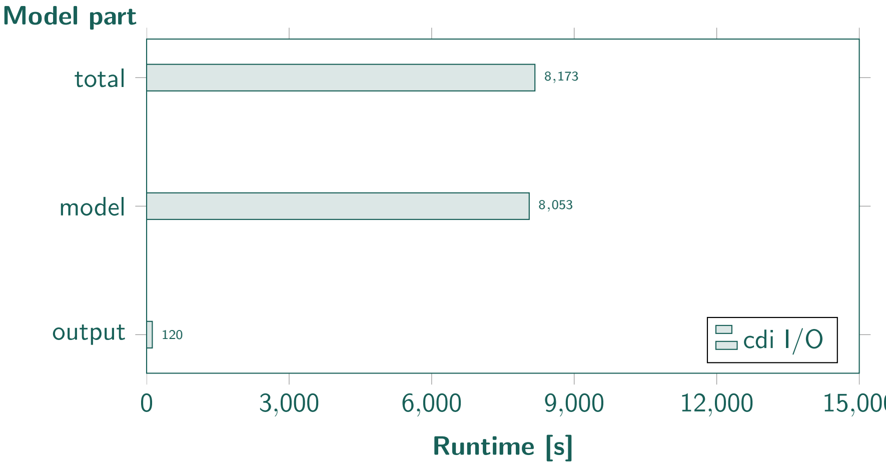
Long term time handling (enables million years range)

- old climate data dating (years of 360 days and 365 days)
- Astronomical clean dates (year 0 included, Gregorian date handling extended into the past, not historical)
- High precision Earth's orbit and Kepler available
- Event handling, pure integer arithmetic based

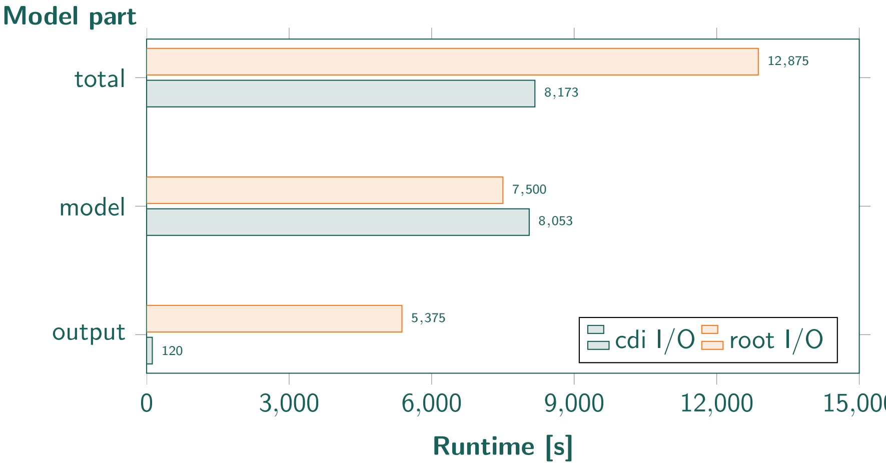
Unifying utility functionality

- I/O across model and processing tools: libcdi
- Collection on post-processing knowledge: cdo
- libaec provides fast compression for GRIB2

Async I/O

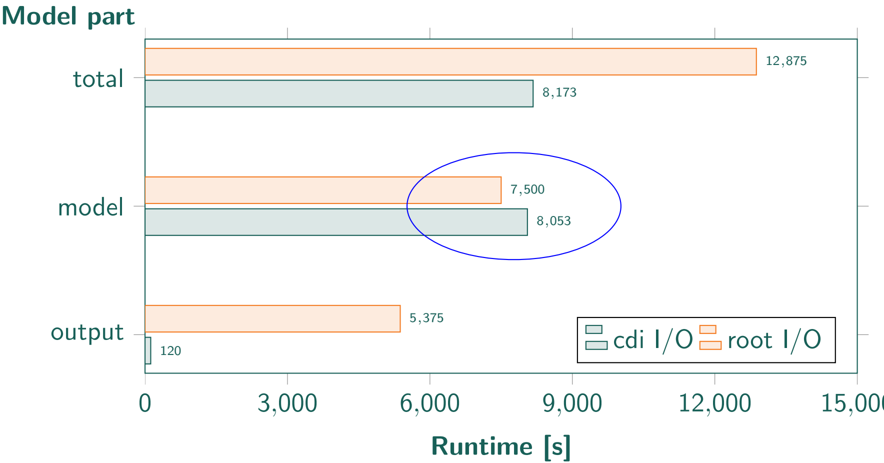

Model part
total
model
output
8,173
8,053
120
cdi I/O
0
3,000
6,000
9,000
12,000
15,00
Runtime [s]

Async I/O



Model part
total
model
output
12,875
8,173
7,500
8,053
5,375
120
cdi I/O
root I/O
0
3,000
6,000
9,000
12,000
15,000
Runtime [s]

Async I/O


Model part
total
model
output
12,875
8,173
7,500
8,053
5,375
120
cdi I/O
root I/O
0
3,000
6,000
9,000
12,000
15,000
Runtime [s]

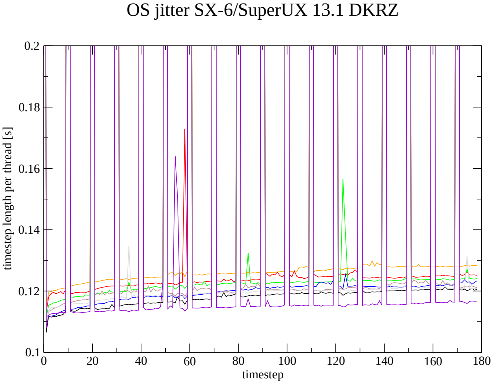
Working with WMO

- Adding new features to GRIB2
- Extending GRIB2 tables
- Complain about GRIB2 reference implementation

More ideas already picked-up

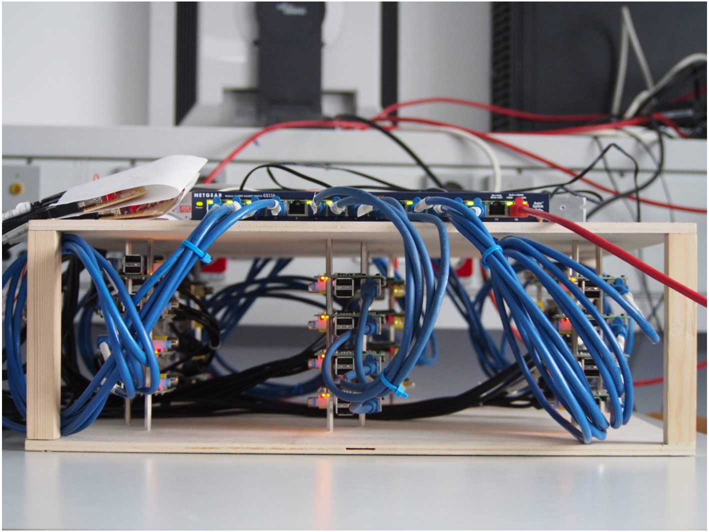
- ~~testing: unit- and application-tests~~
- ~~version control~~

OS jitter effect



Most commented project

Hardware deployment of experimental cluster



Courtesy by Miriam, 13 years

System characteristics

- 24 nodes with Broadcom BCM2835 SoC (700 MHz ARM 1176JZF-S, VideoCore IV GPU, not used)
- Non-blocking fat tree high speed network IEEE 802.3u (100BASE-TX) via USB-2 Bus (aggregated 273.6 MB/s)
- NFSv4 network filesystem, SLURM, GCC, mpich
- Linux Debian jessie (Kernel 4.4)

System characteristics

- 24 nodes with Broadcom BCM2835 SoC (700 MHz ARM 1176JZF-S, VideoCore IV GPU, not used)
- Non-blocking fat tree high speed network IEEE 802.3u (100BASE-TX) via USB-2 Bus (aggregated 273.6 MB/s)
- NFSv4 network filesystem, SLURM, GCC, mpich
- Linux Debian jessie (Kernel 4.4)

System characteristics

- 24 nodes with Broadcom BCM2835 SoC (700 MHz ARM 1176JZF-S, VideoCore IV GPU, not used)
- Non-blocking fat tree high speed network IEEE 802.3u (100BASE-TX) via USB-2 Bus (aggregated 273.6 MB/s)
- NFSv4 network filesystem, SLURM, GCC, mpich
- Linux Debian jessie (Kernel 4.4)

System characteristics

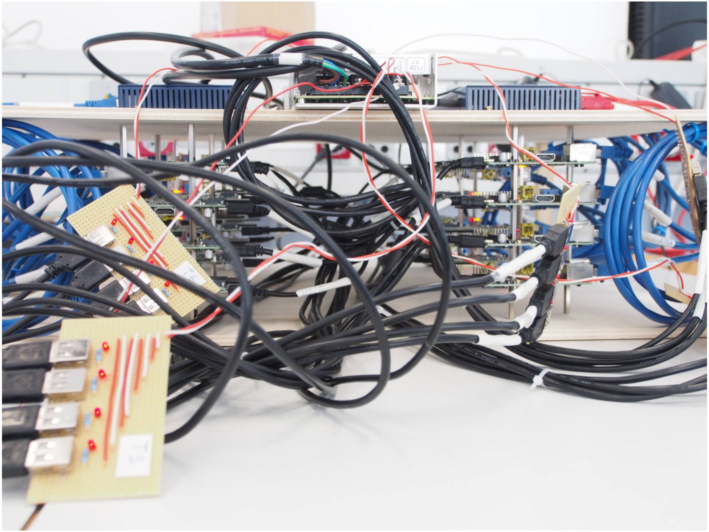
- 24 nodes with Broadcom BCM2835 SoC (700 MHz ARM 1176JZF-S, VideoCore IV GPU, not used)
- Non-blocking fat tree high speed network IEEE 802.3u (100BASE-TX) via USB-2 Bus (aggregated 273.6 MB/s)
- NFSv4 network filesystem, SLURM, GCC, mpich
- Linux Debian jessie (Kernel 4.4)

System characteristics

- 24 nodes with Broadcom BCM2835 SoC (700 MHz ARM 1176JZF-S, VideoCore IV GPU, not used)
- Non-blocking fat tree high speed network IEEE 802.3u (100BASE-TX) via USB-2 Bus (aggregated 273.6 MB/s)
- NFSv4 network filesystem, SLURM, GCC, mpich
- Linux Debian jessie (Kernel 4.4)

Successfully run echam 4.6 T31L19 (CVS version 6.00, 2000-09-19 08:26:58 (Git: da9d477) , no code changes) using the full system.

Energy consumption 100 W



Courtesy by Miriam, 13 years

Projects to come

What next?

- Next generation of self-designed and build cluster to train for next generation of complexity wizards, a dream of mine.
- C++ version/replacement of namelists.
- In code DAG based task scheduling (C++), might be crazy.

What next?

- Next generation of self-designed and build cluster to train for next generation of complexity wizards, a dream of mine.
- C++ version/replacement of namelists.
- In code DAG based task scheduling (C++), might be crazy.

What next?

- Next generation of self-designed and build cluster to train for next generation of complexity wizards, a dream of mine.
- C++ version/replacement of namelists.
- In code DAG based task scheduling (C++), might be crazy.