

Porting Climate Models to NEC SX-Aurora TSUBASA

<Panagiotis Adamidis>
Deutsches Klimarechenzentrum (DKRZ)

NEC SX-Aurora TSUBASA @ DKRZ

- People who contributed
 - Panagiotis Adamidis, Jan Frederik Engels, Hendryk Bockelmann (DKRZ)
 - Günther Zängl (DWD)
 - Rene Redler (MPI-Met)

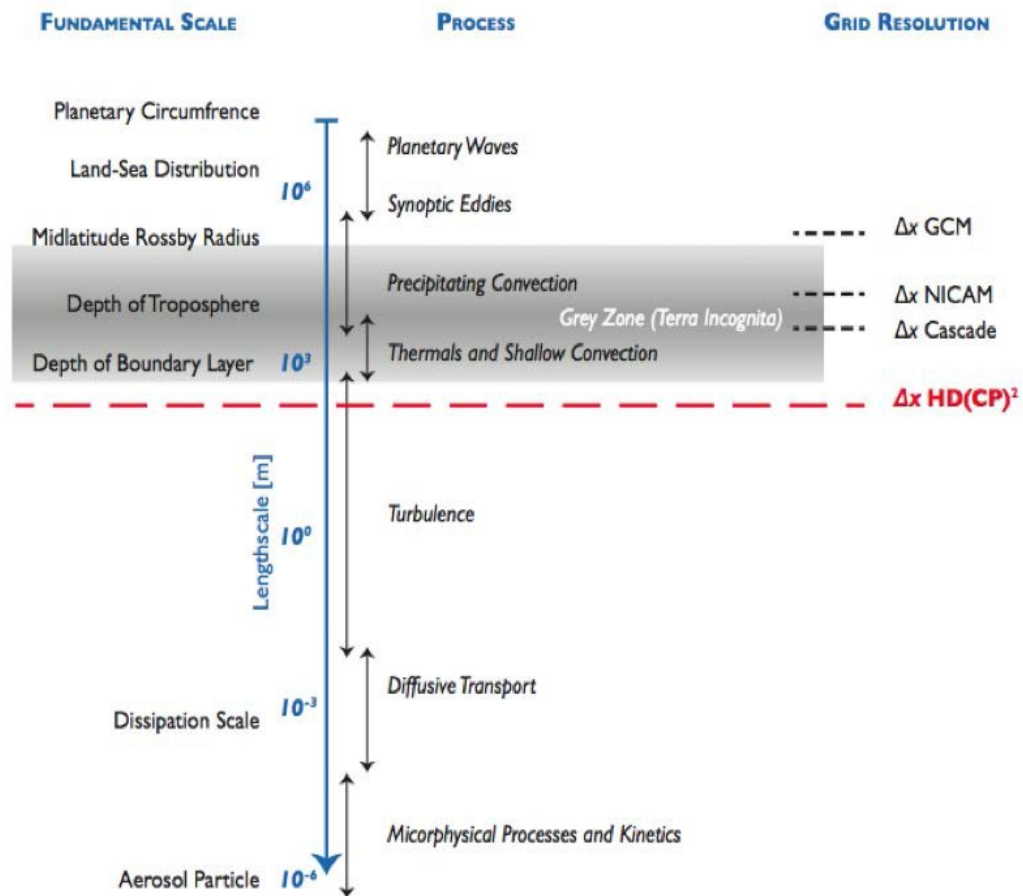
Climate Models @ DKRZ

- High Resolution
 - Resolving small scale physical processes
- Coarse Resolution
 - Simulating longer periods (80000 – 100000 years)
 - Complete glacial cycles

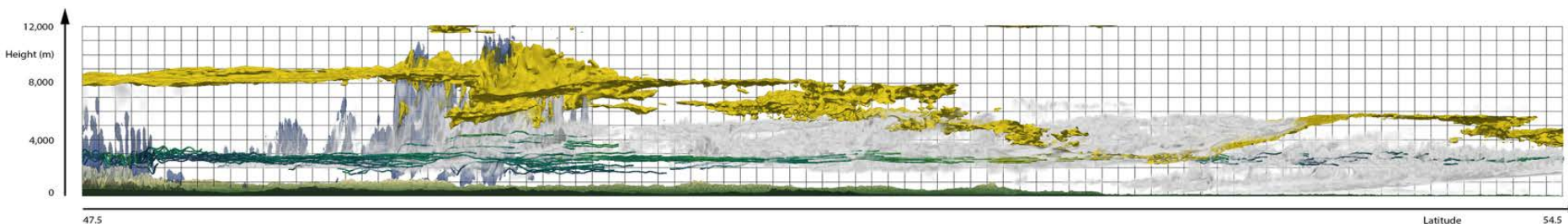
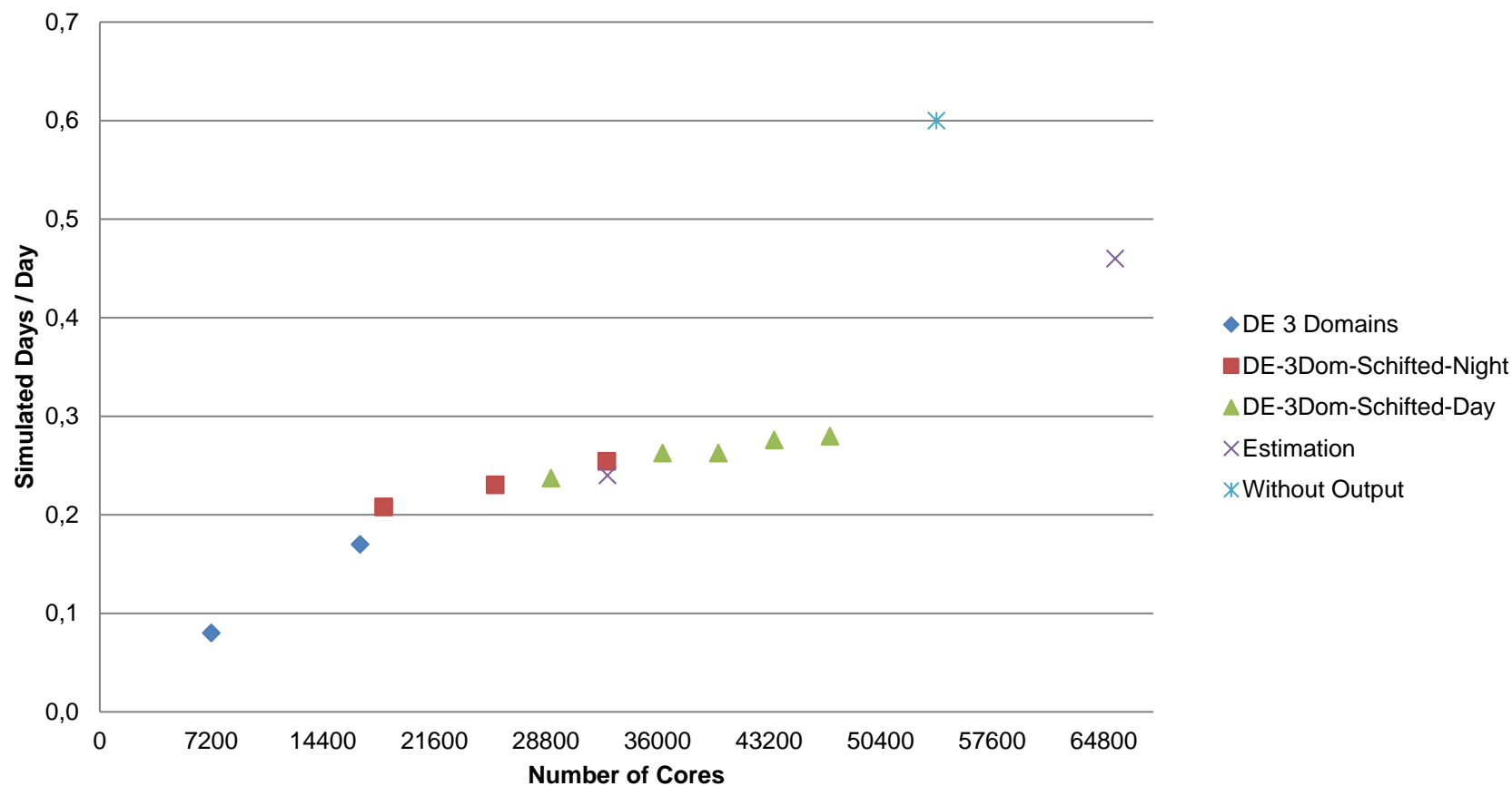
ICON Grid Resolutions

grid	number of cells	avg. resolution
R2B04	20480	158 km
R2B05	81920	79 km
R2B06	327680	40 km
R2B07	1310720	20 km
R2B09	20971520	5 km
R2B10	83886080	2.5 km
R2B11	335544320	1.25 km

Very High Resolution Climate Modelling HD(CP)²



HDCP2 Modell - Sustained Performance

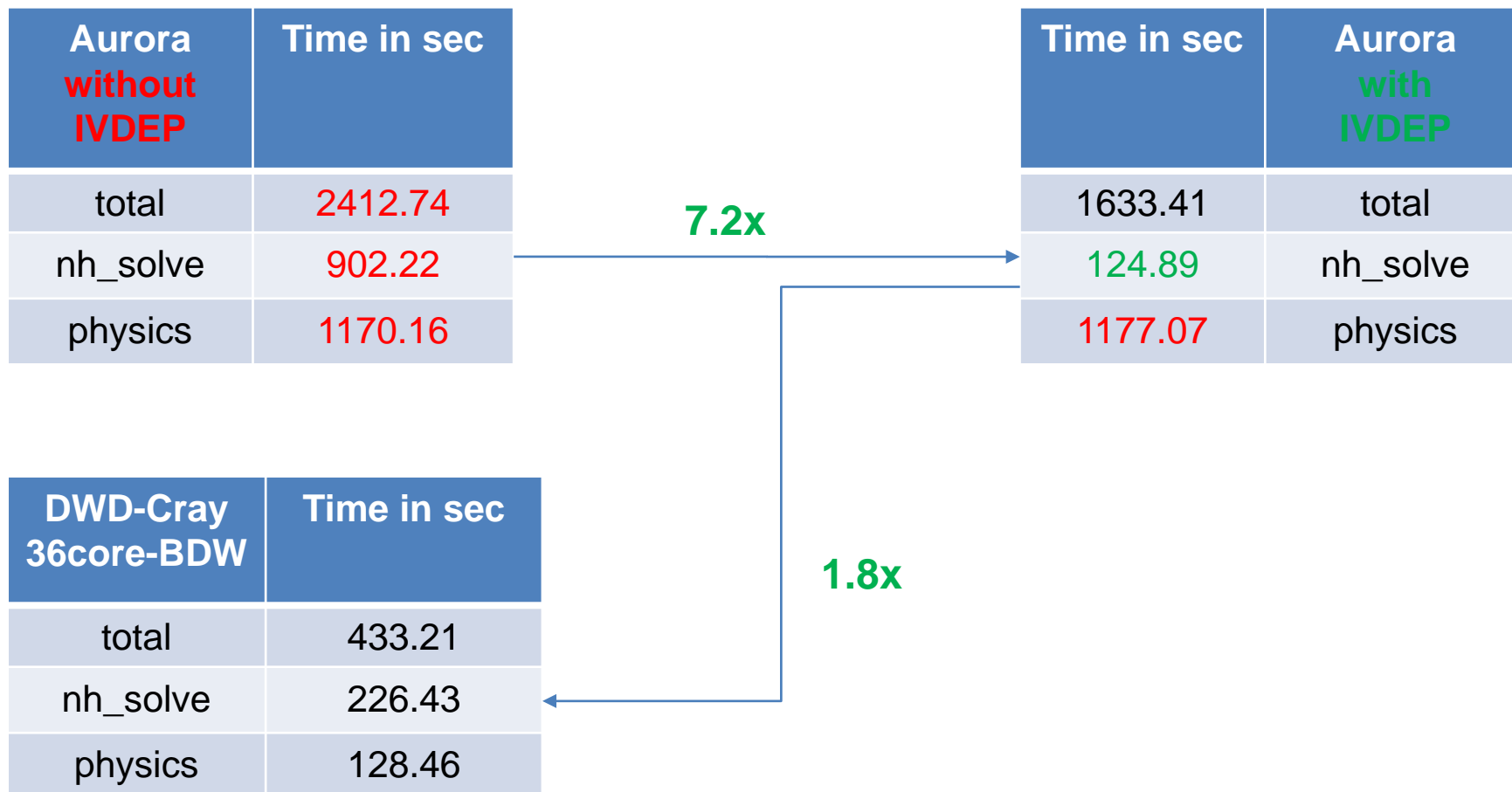


Some Statistics (01.2016-02.2018)

Total number of / amount of	
simulated days	26
variables written out	169
meteorological stations	36
data	output : 1,5 PB + meteogram : 1 TB
node hours used	3,25 Mio node hours

NEC SX-Aurora TSUBASA: ICON-NWP (R2B06L90)

- Successful compilation with **-O2 -finline-functions**



(G. Zängl, DWD)

NEC SX-Aurora TSUBASA

- Compiler shows reluctant behaviour towards vectorization
- Many unvectorizable dependencies detected
- Using Pointerarrays with the `CONTIGUOUS` attribute leads to unvectorizable dependencies

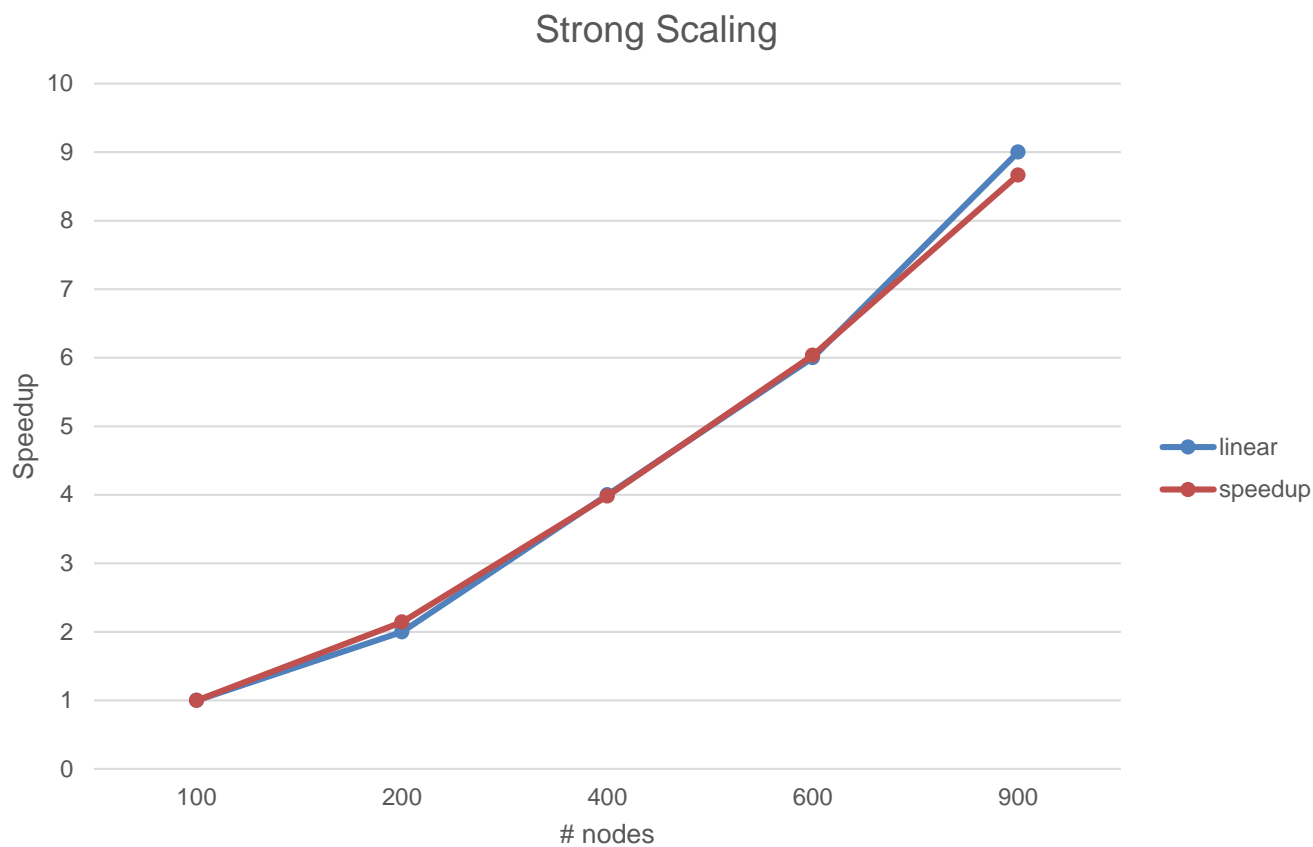
NEC SX-Aurora TSUBASA: ICON-NWP (R2B06L90)

nh_solve	1 node NEC SX-Aurora TSUBASA 16 cores (time in sec)	1 node CRAY Broadwell 36 cores (time in sec)
nh_solve.cellcomp (direct addressing, memory-bound)	3.16	36.05
nh_solve.veltend (indirect addressing)	14	31.7
nh_solve.vimpl (good cache utilization)	11.9	53.5

(G. Zängl, DWD)

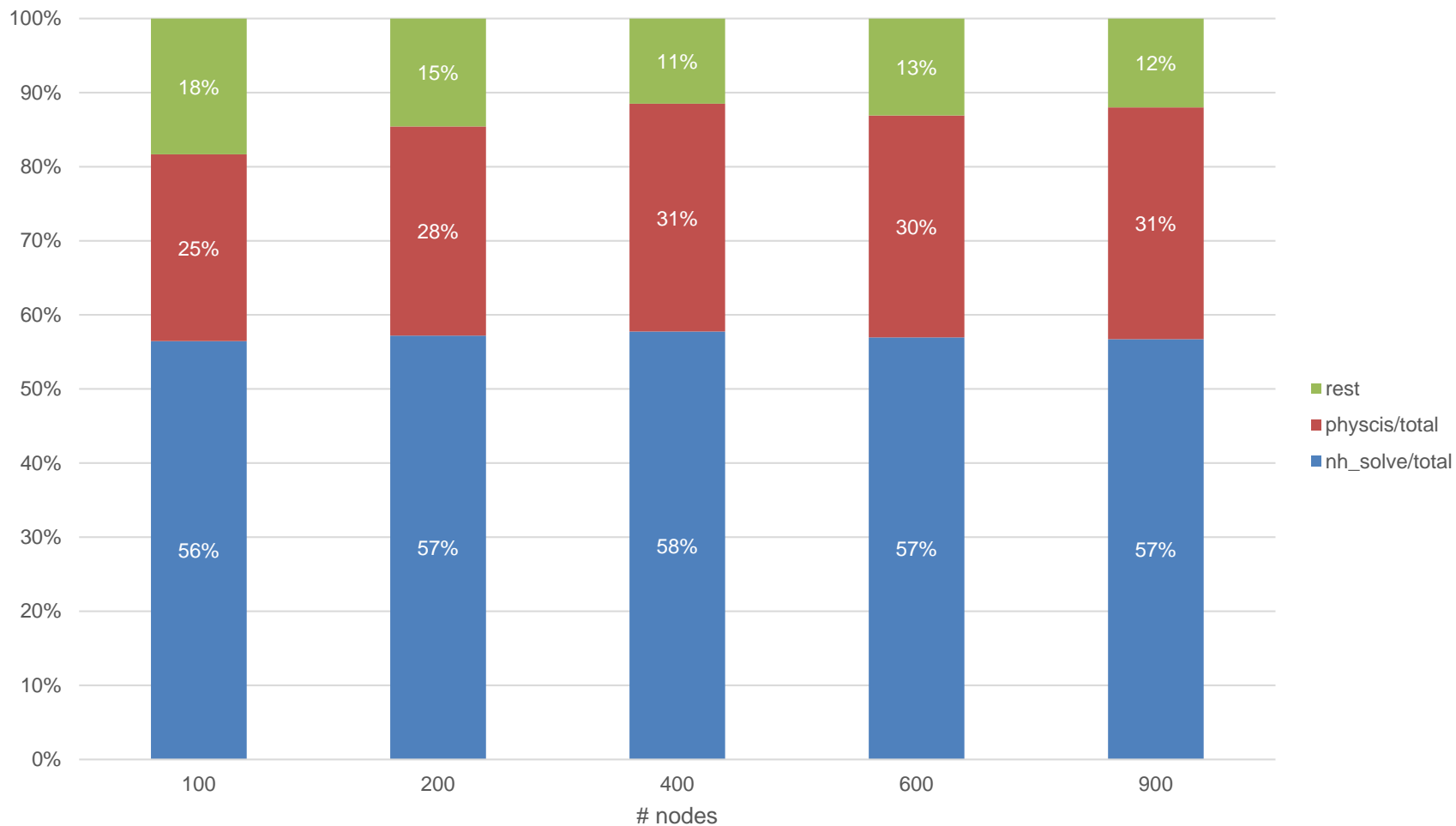
DYAMOND : ICON-AES / R2B09L90 / 1 Simulated Day

mistral BDW	100 nodes	200 nodes	400 nodes	600 nodes	900 nodes
total	8731.34	4072.34	2192.81	1446.71	1007.58



DYAMOND : ICON-AES / R2B09L90 / 1 Simulated Day

Time Distribution



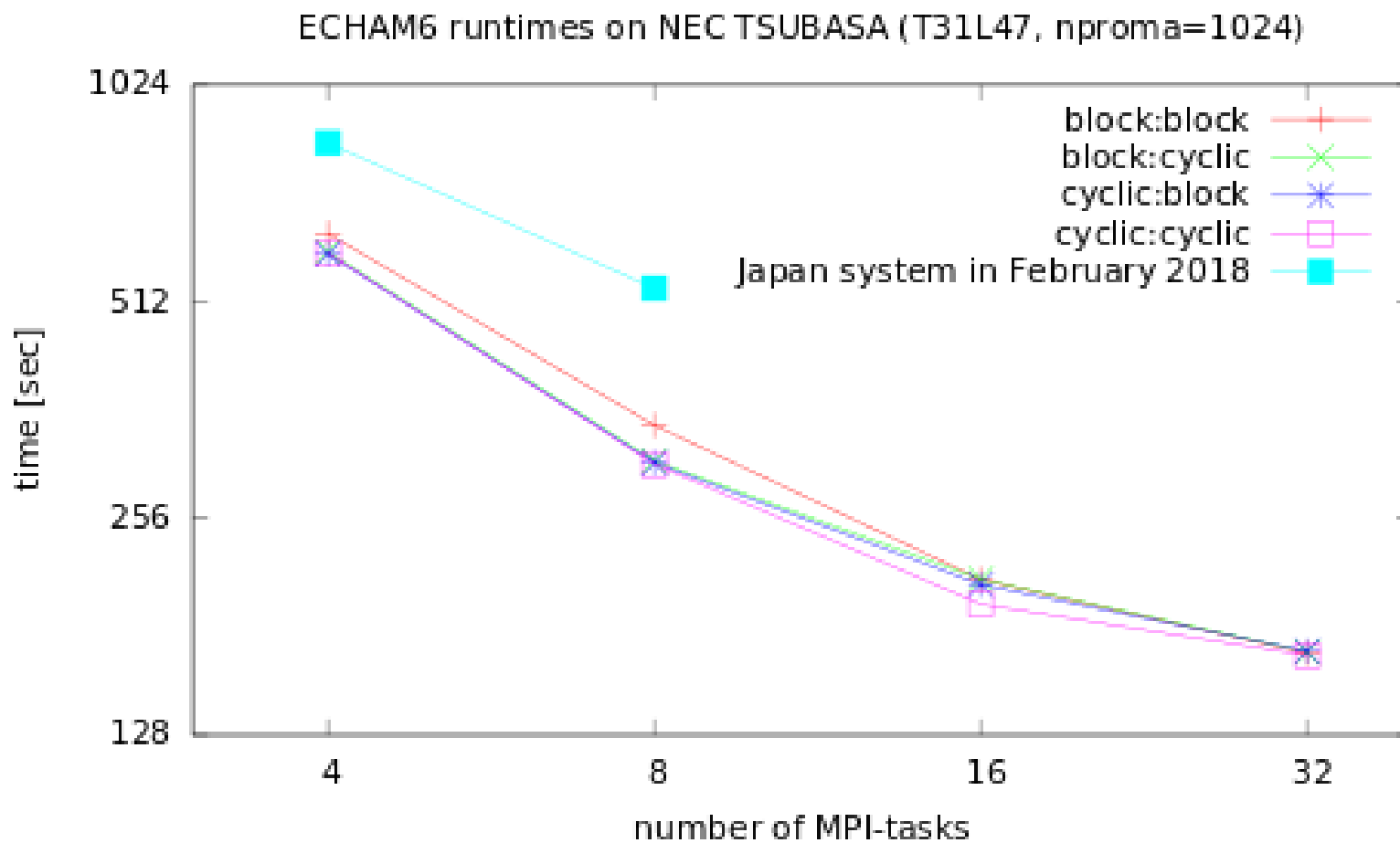
NEC SX-Aurora TSUBASA: ICON-coupled ATM/OCE

- Successful compilation **with** optimization level **O3**
- 1 Simulated Day
- Input seems to be a bottleneck

Hardware	MPI-Tasks	Total (time in sec)	echam_bcs (time in sec)	Difference of Total time in sec	Slowdown factor
1 mistral node : 24 HSW cores	22 Atmosphere + 2 Ocean	65,8	4,12		
1 Aurora node : 16 Vector cores	14 Atmosphere + 2 Ocean	296,7	16,7	230,9	4,5x

(R. Redler, MPI-Met)

NEC SX-Aurora TSUBASA : ECHAM6



NEC SX-Aurora TSUBASA: ECHAM6

- Successful compilation **only** when using low optimization level **O1**

Number MPI Tasks	HSW O3 + hiopt (time in sec)	NEC SX-Aurora O1 (time in sec)	Difference (time in sec)	Slowdown
16	48,657	210,49	161,83	13,15x
32	26,122	166,43	140,30	6,37x

NEC SX-Aurora TSUBASA : ftrace

	with ftrace (time in sec)	no ftrace (time in sec)
ICON-AES	478	28
ICON-Coupled	1600	297

- Overhead of ftrace is huge

Conclusion & Outlook

- NEC SX-Aurora architecture has good potential to deliver sustained performance
- Compiler with efficient vectorization is essential
- Good scaling over hundreds/thousands of nodes is necessary
- Efficient parallel I/O is vital for high resolution simulations
- The performance of the file system is very important