Earth System Models: Challenges in a changing HPC environment

Patrick Jöckel

Institute of Atmospheric Physics, Oberpfaffenhofen

DLR

Outline

- Introduction

- “Coupling”: An attempt to classify different methods

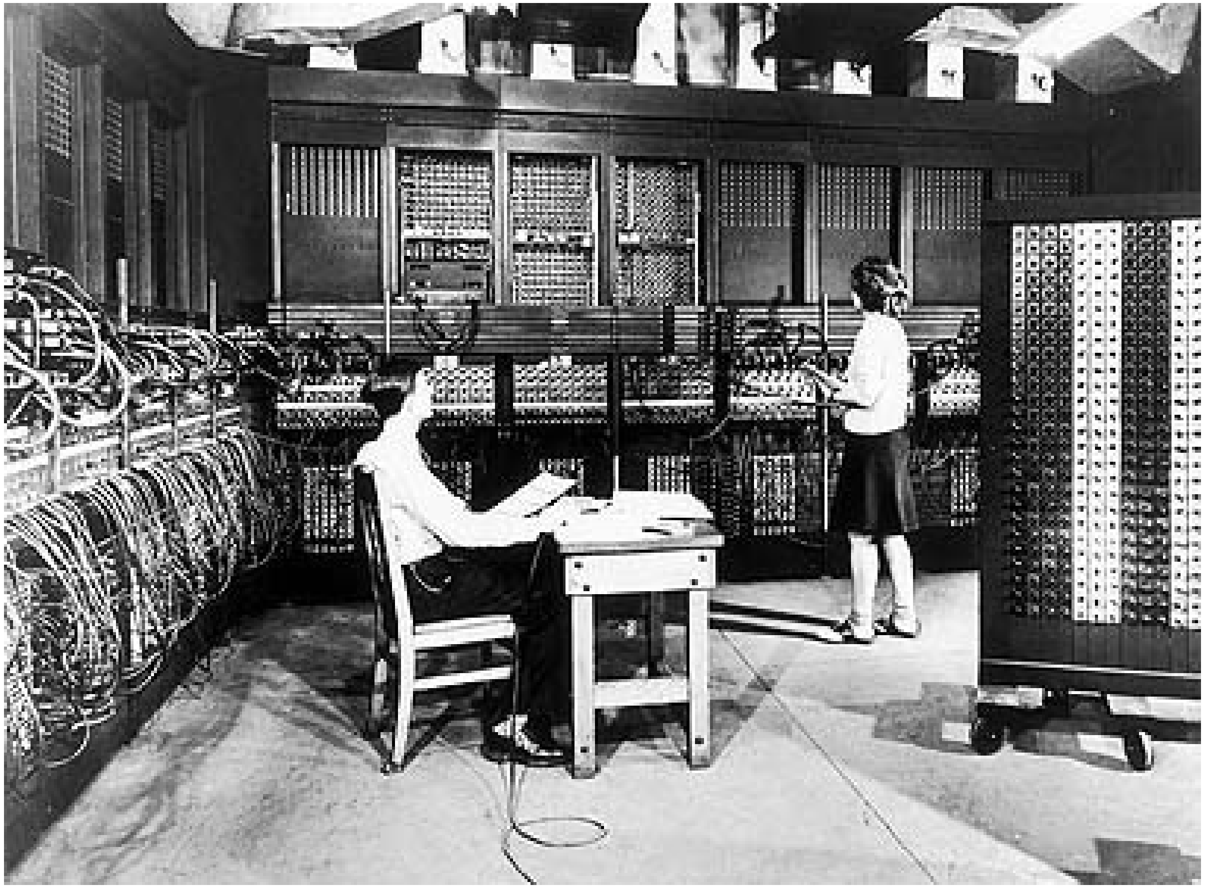
- Examples, challenges (and solutions)
    1. Atmospheric Chemistry in the Earth System
    2. Atmosphere – Ocean System
    3. On-line nesting: an alternative way to higher resolution

- Summary and Outlook

Computational Earth System Science
(numerical weather prediction and climate simulations)
was from the beginning on exploiting HPC up to the limits …

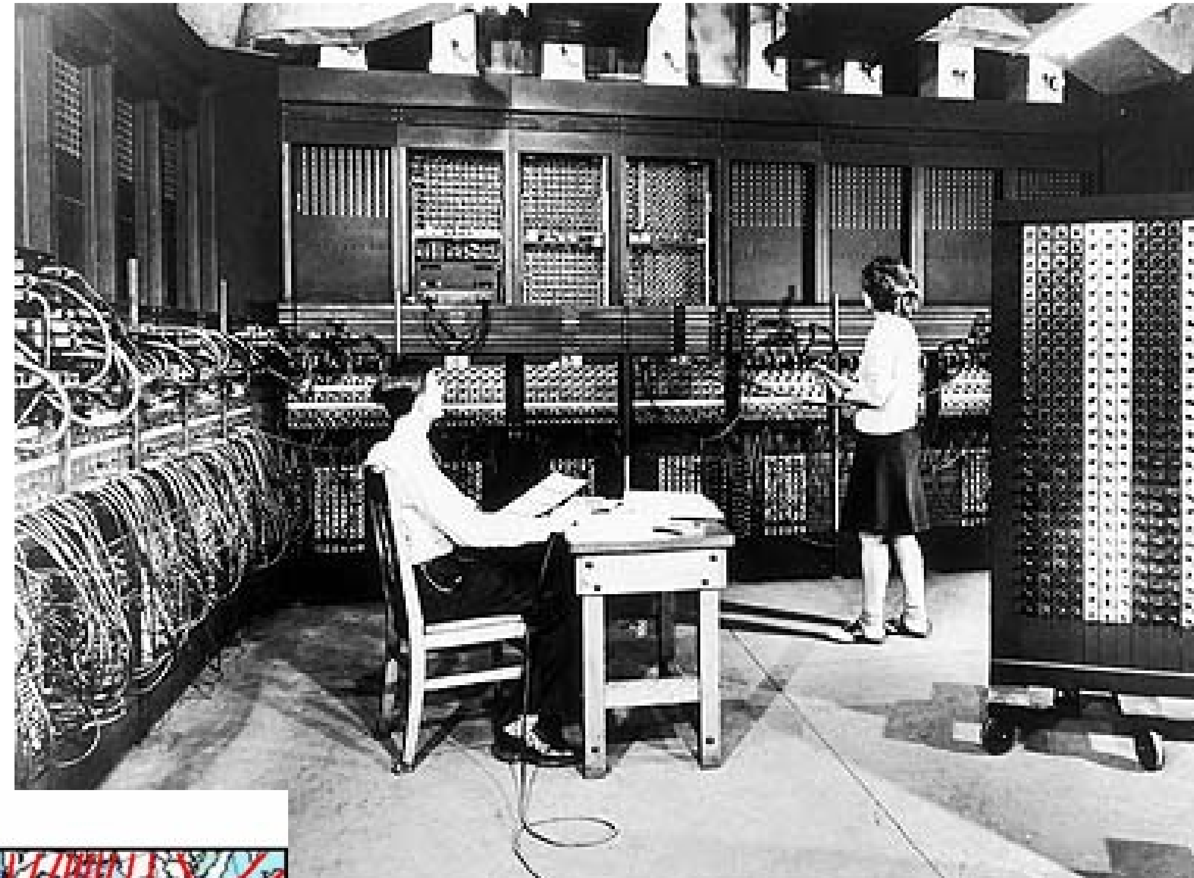
1950 (Charney, Fjørtoft, von Neumann):
first numerical weather forecast on ENIAC (Electronic Numerical Integrator and Computer)

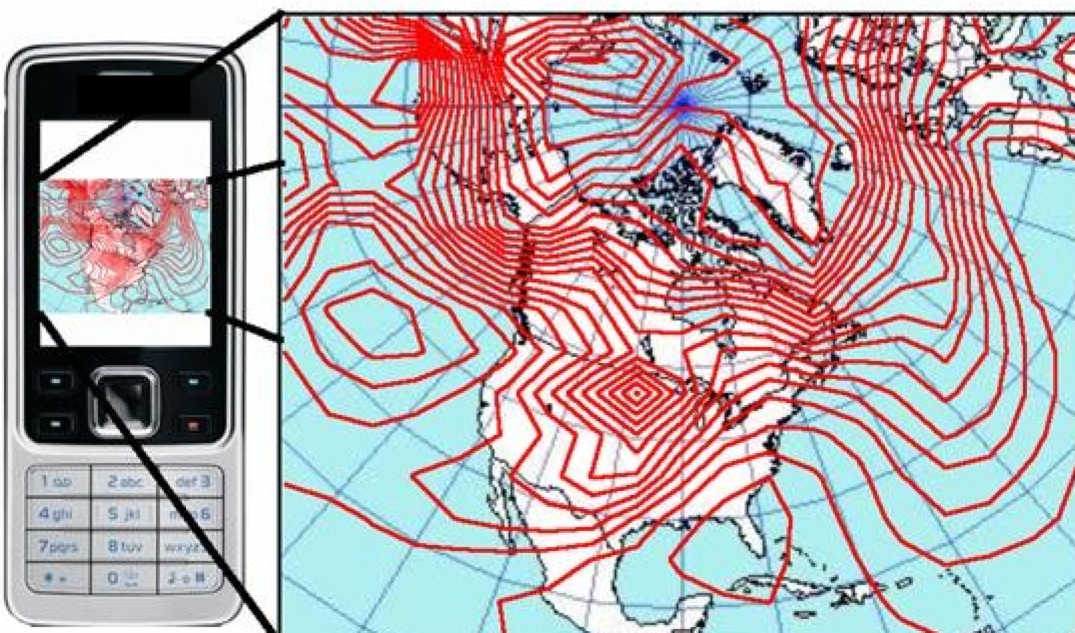
forecast time: 24 hours
computation: 24 hours



1950 (Charney, Fjørtoft, von Neumann):
first numerical weather forecast on ENIAC
(Electronic Numerical Integrator and Computer)

forecast time: 24 hours
computation: 24 hours





2008 (Lynch & Lynch):
reconstruction on mobile-phone (JAVA-application):

forecast time: 24 hours
computation: < 1 second (!!!)

Lynch & Lynch, Weather 63, 324-326, 2008.
http://mathsci.ucd.ie/~plynch/eniac/phoniac.html

Deutsches Klimarechenzentrum (DKRZ)

8448 Cores
158 TFlops/s

Top 500
June 2009: no. 27
Nov 2011: no. 98

http://www.dkrz.de

Computational Earth System Science
(numerical weather prediction and climate simulations)
was from the beginning on exploiting HPC up to the limits …

… mainly for two reasons:

increasing resolution of numerical discretisation(s)
(i.e., finer “grids”)

increasing complexity by incorporating more and more
processes

Computational Earth System Science
(numerical weather prediction and climate simulations)
was from the beginning on exploiting HPC up to the limits …

… mainly for two reasons:

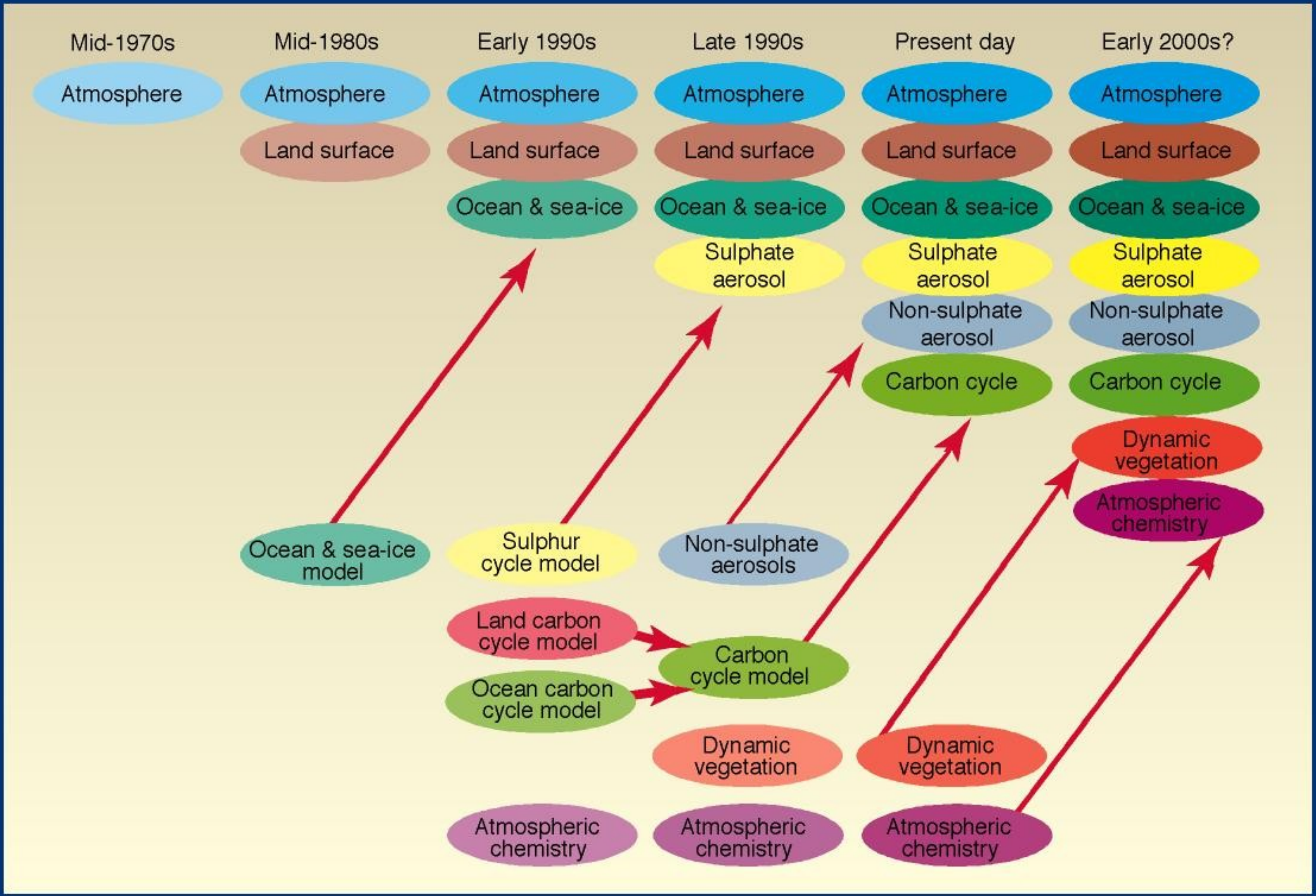
increasing resolution of numerical discretisation(s)
(i.e., finer "grids")

increasing complexity by incorporating more and more processes

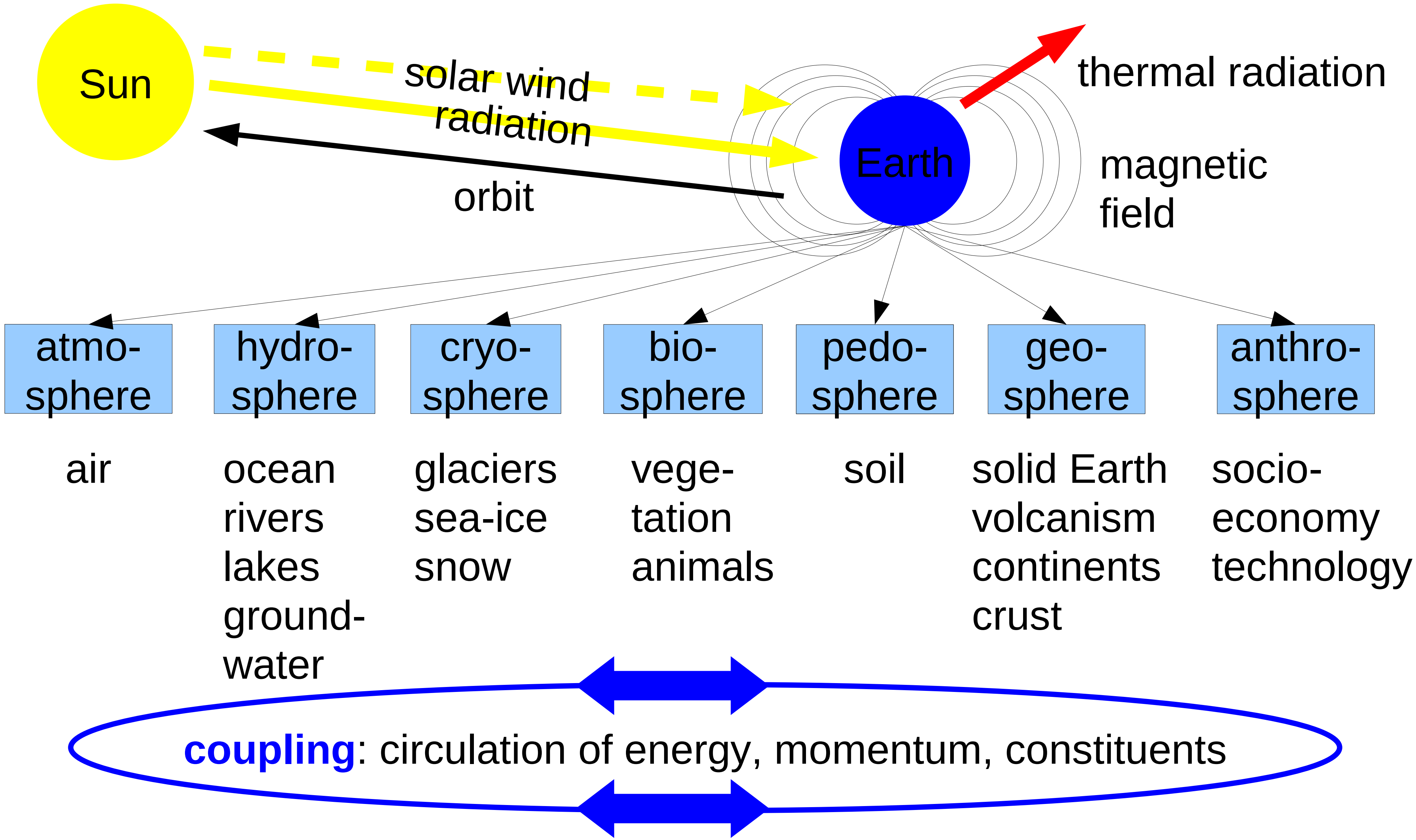
→ *"coupling"*

DLR

The development of climate models, past, present and future



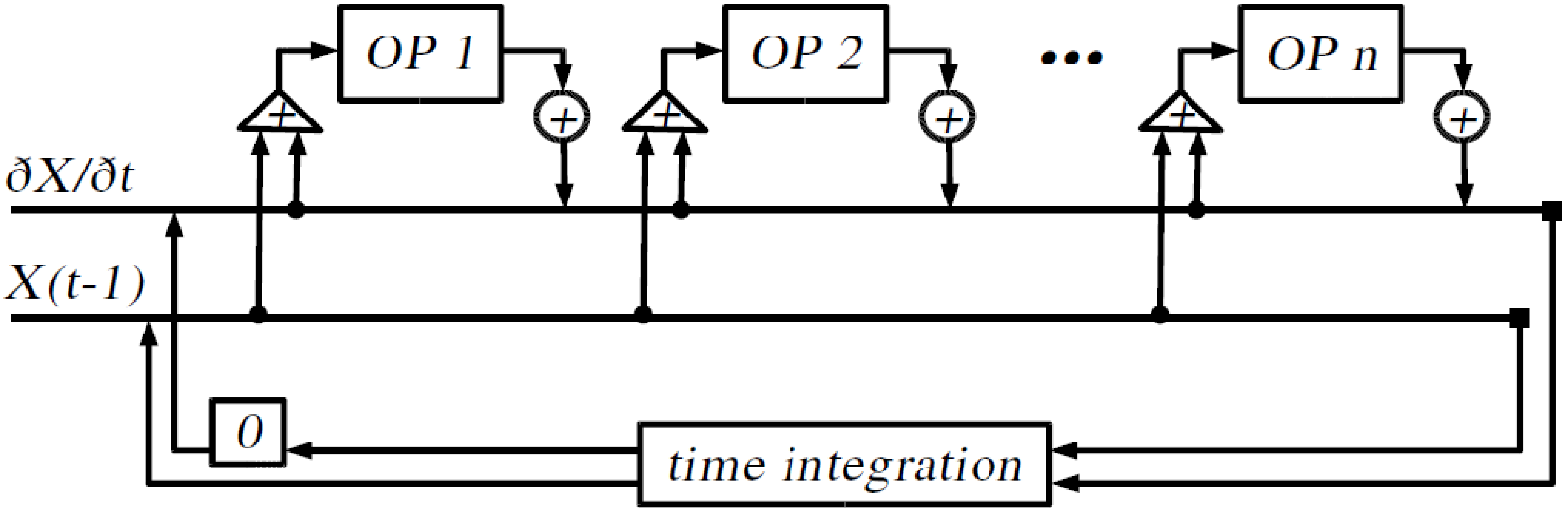
WG1 - TS BOX 3
FIGURE 1



change of state variables by physical, chemical, biological, socio-economic processes

Coupling ...

… the prerequisite is the *operator splitting* concept



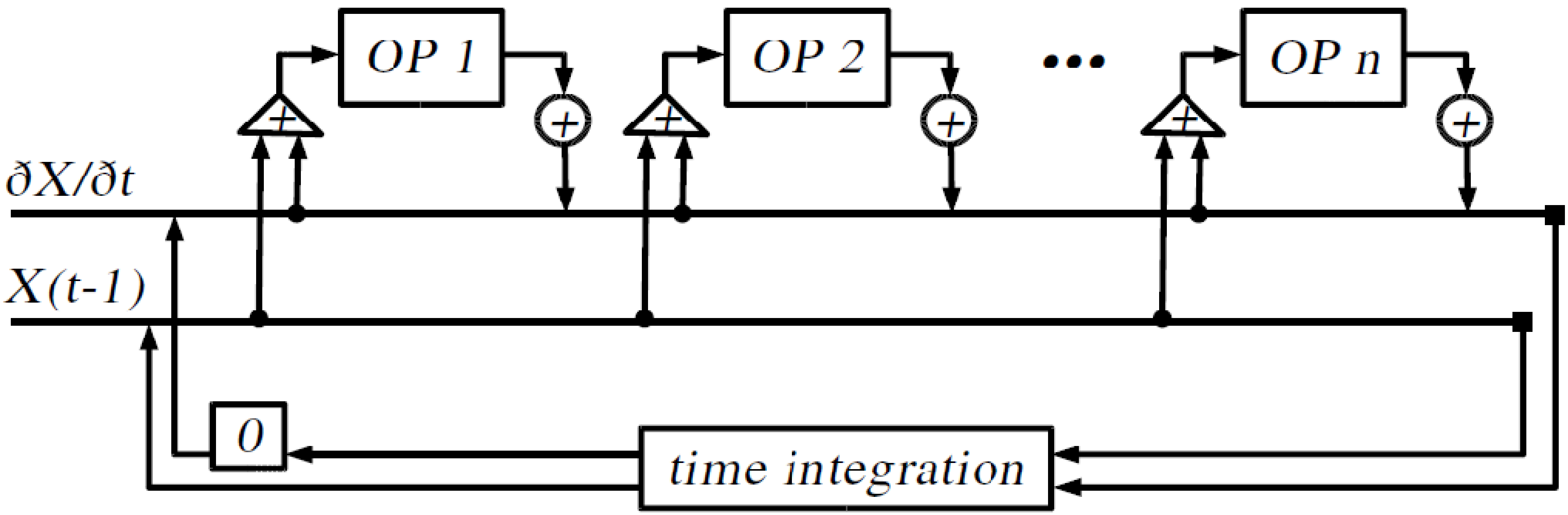
(Jöckel et al., ACP, 2005)

Coupling ...

… the prerequisite is the *operator splitting* concept

Example:
X = air
temperature

radiation convection ... diffusion

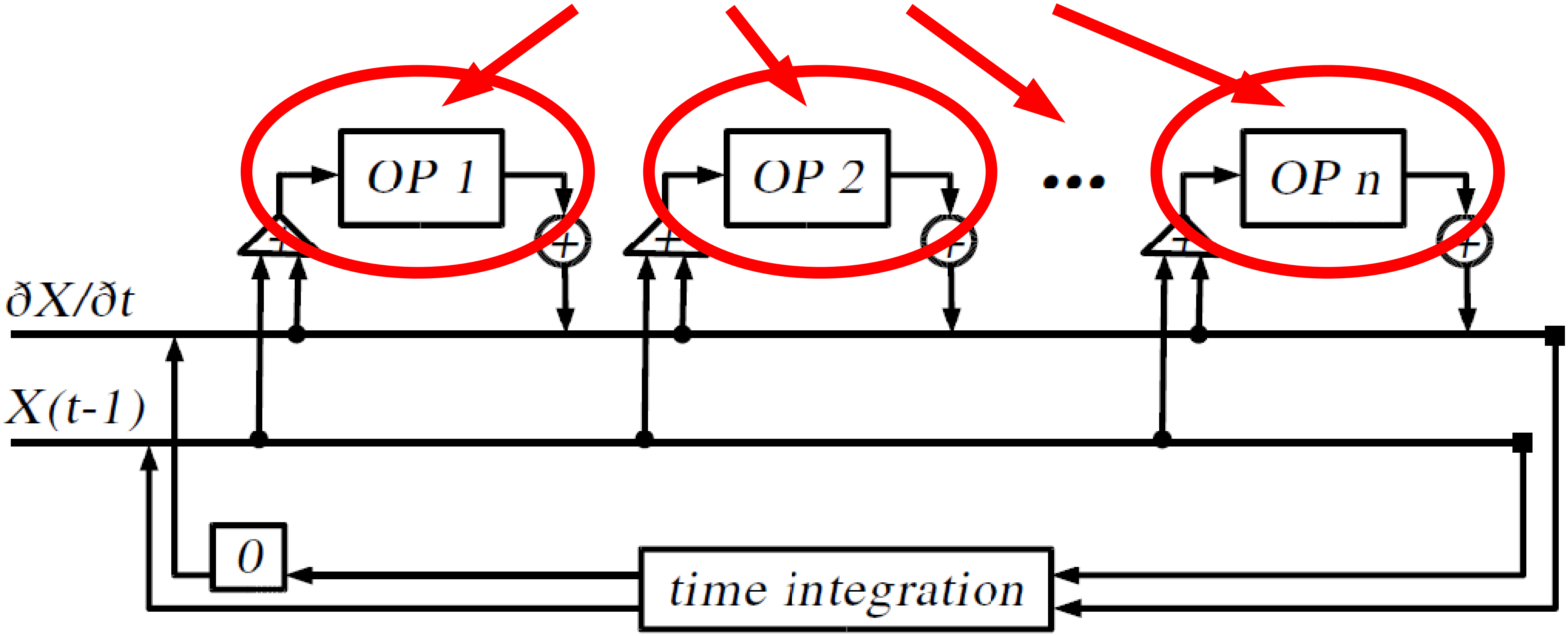


(Jöckel et al., ACP, 2005)

Coupling ...

… the prerequisite is the *operator splitting* concept

different numerical algorithms (discretisation, parallel decomposition, cache/vector blocking, ...)



(Jöckel et al., ACP, 2005)

DLR

Coupling ...

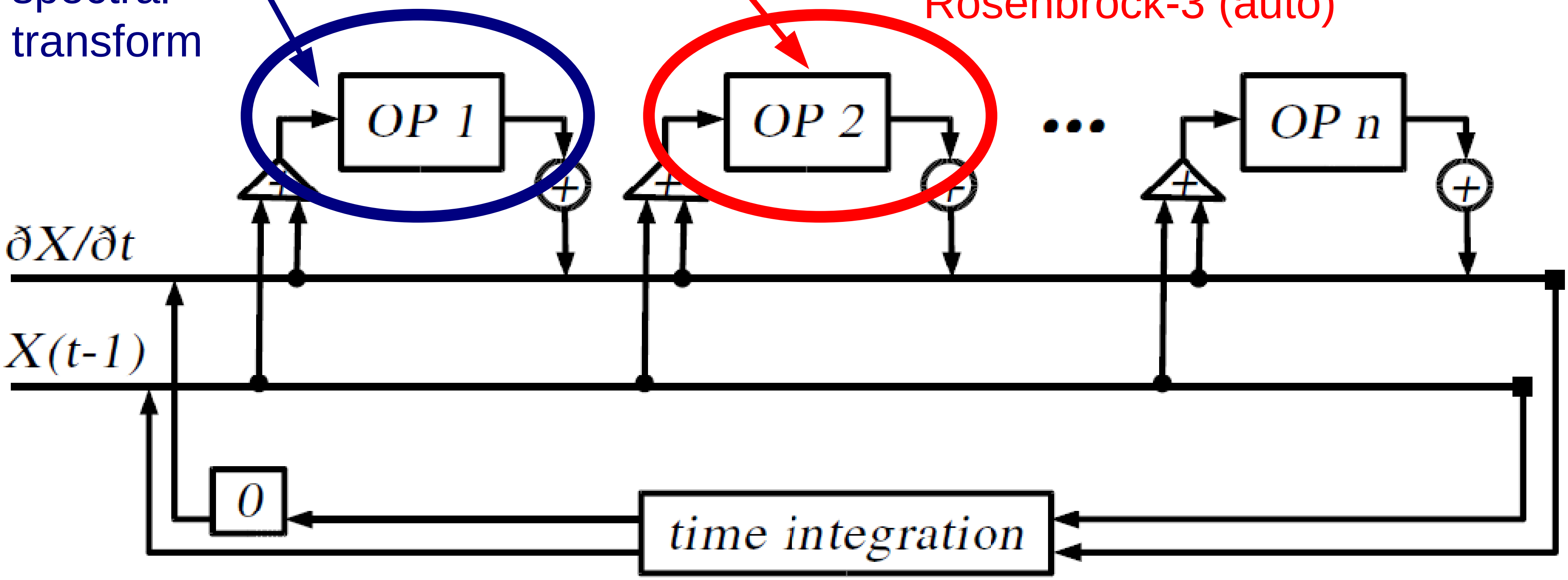
Example:

basic (dynamical) equations → coupled PDE system

reaction kinetics (chemistry) → coupled ODE system

spectral transform

Rosenbrock-3 (auto)



(Jöckel et al., ACP, 2005)

Coupling … a classification (of the “way” how operators “communicate”)

internal coupling ⟷ **external coupling**

internal coupling

- Os are part of the *same program unit* (“task”)
- data exchange via *working memory*

{ on-line, direct }

external coupling

- Os are split into *different program units* (“tasks”)

off-line

- data exchange via *external storage* (files)

on-line

- data exchange via *working memory*

indirect

- involves an *add. external* program (“*coupler*”)

direct

DLR

Coupling … a classification (of the "way" how operators "communicate")

internal coupling ⟷ **external coupling**

internal coupling

- Os are part of the *same program unit* ("task")
- data exchange via *working memory*
  - on-line
  - direct

external coupling

- Os are split into *different program units* ("tasks")
  - **off-line**
    - data exchange via *external storage* (files)
  - **on-line**
    - data exchange via *working memory*
      - **indirect**
        - involves an *add. external* program ("*coupler*")
      - **direct**

choice depends on:
- application
- implementation effort (legacy code!)
- desired sustainability, flexibility, re-usability

→ compromise in minimizing computational and communication overheads

Example 1: internal coupling

Language level: CALL subroutine_1(..., A, …) [INTENT(OUT)]

↓

CALL subroutine_2(..., A, …) [INTENT(IN)]

more formal: standard model infrastructure + coding standard

- Earth System Modeling Framework (ESMF) (*http://www.earthsystemmodeling.org*)
- Modular Earth Submodel System (MESSy) (*http://www.messy-interface.org*, *Jöckel et al., ACP, 2005*)
- … (many others)

key: strict separation of *model infrastructure (4 layer!)* (memory management, I/O, parallel decomp., time control, etc.) from *“process” (and “diagnostic”) formulations*

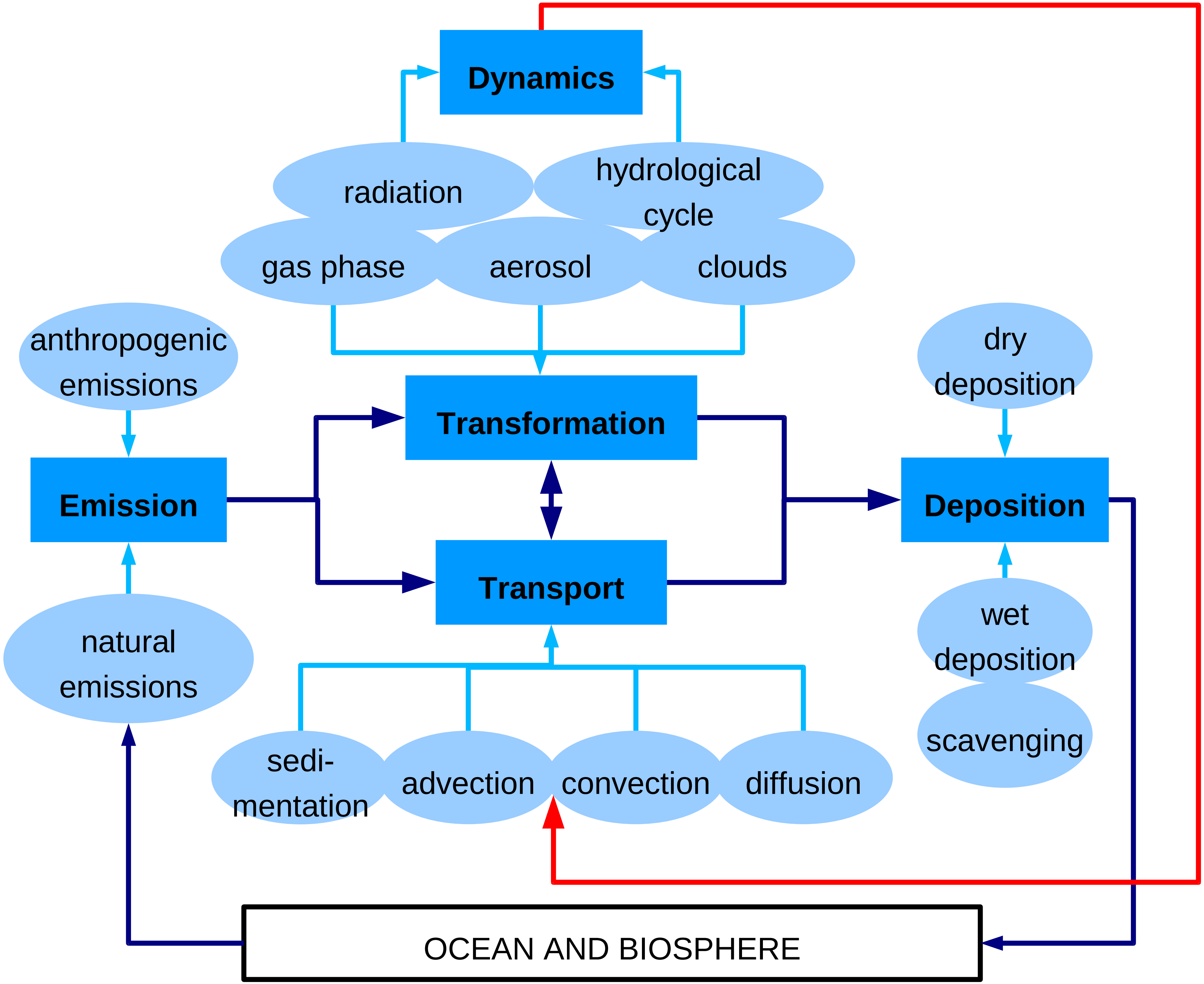
DLR

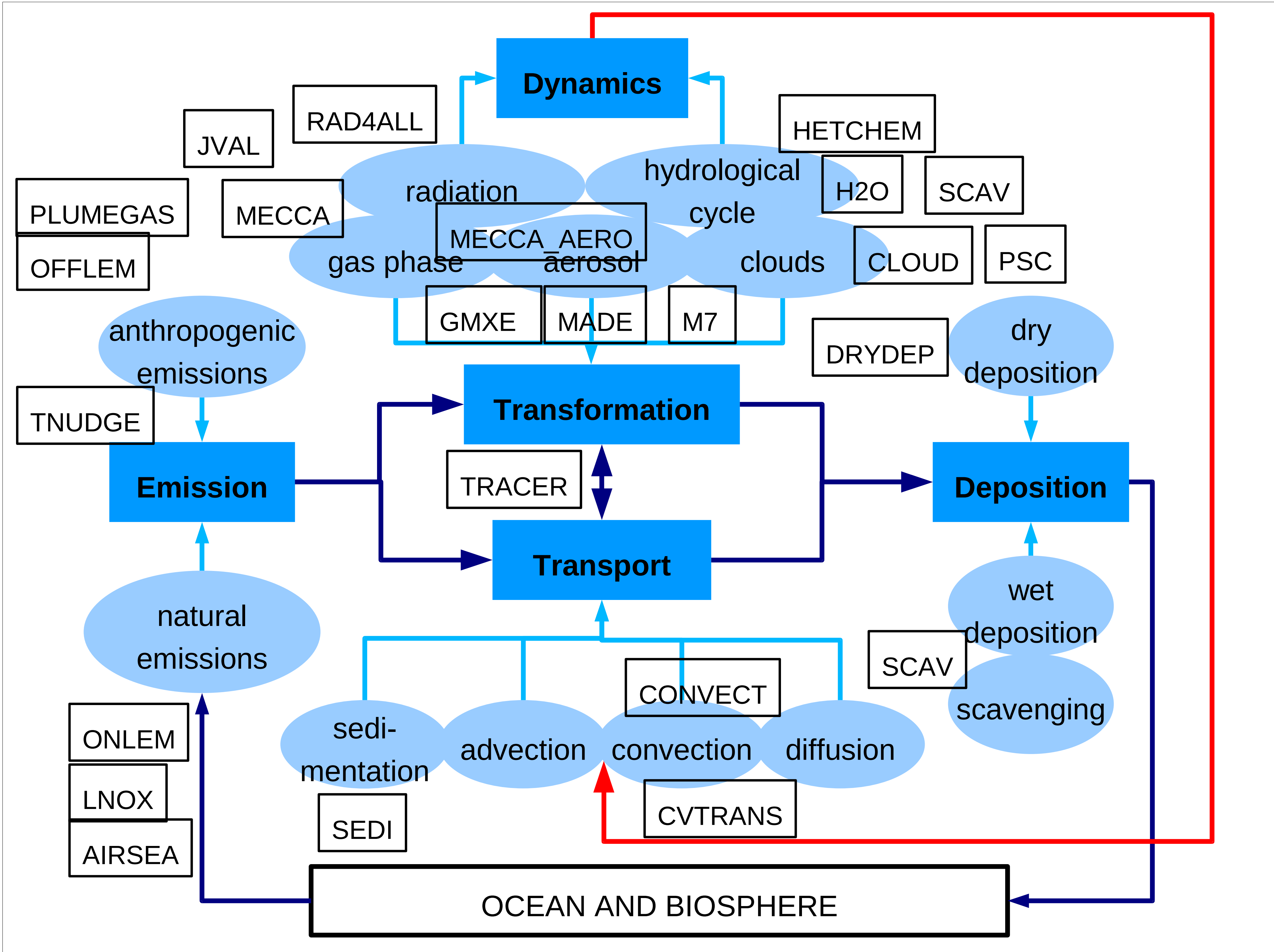
Example 1a: internal coupling of Atmospheric Chemistry in MESSy

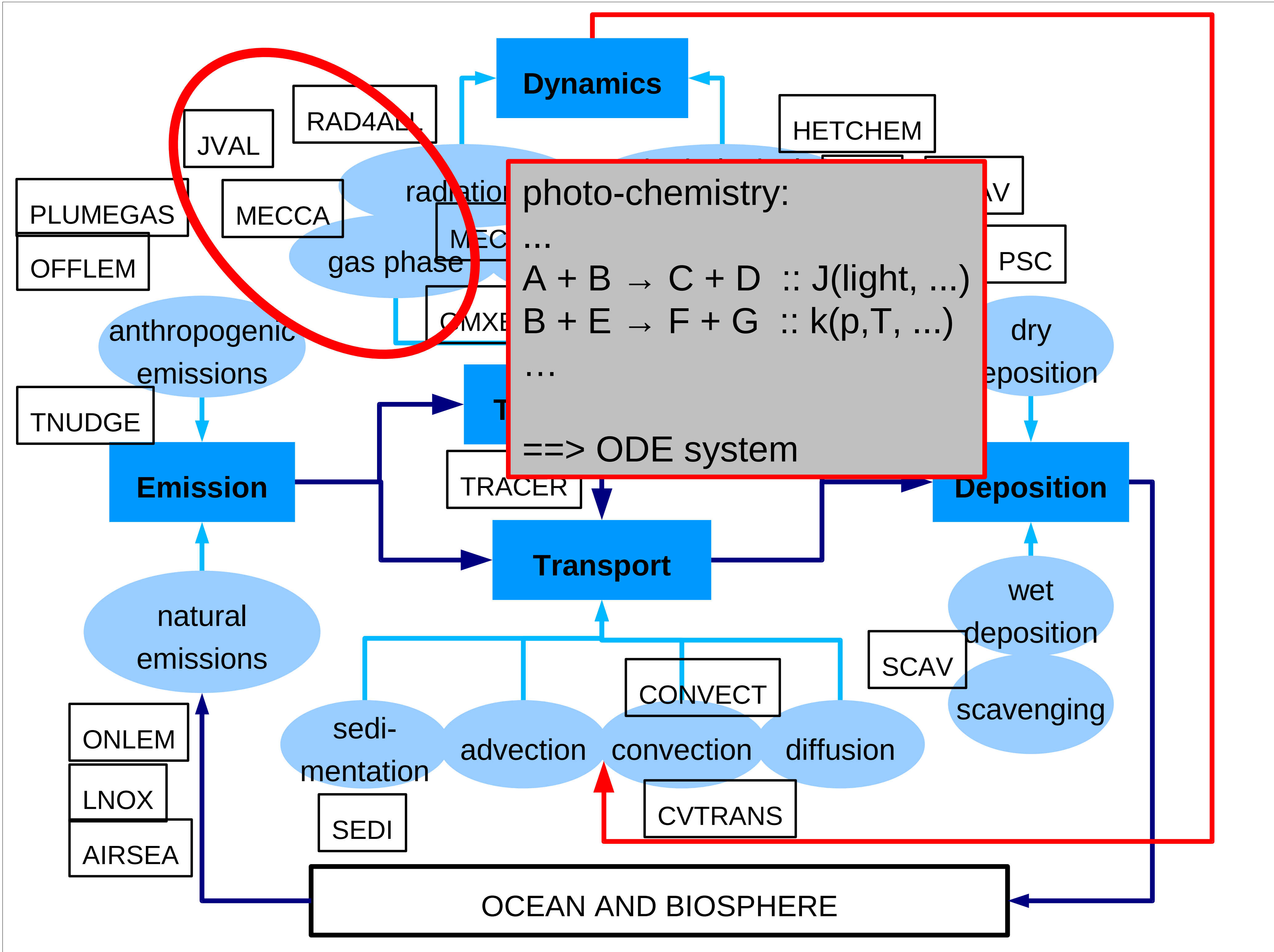
coupling via model infrastructure (nearly object oriented)

- TIMER
- CHANNEL (pointer based memory management and I/O)
- TRACER (special for chemical constituents)
- …

“operators” = “processes” = “submodels”


Dynamics
radiation
hydrological cycle
gas phase
aerosol
clouds
anthropogenic emissions
dry deposition
Transformation
Emission
Deposition
Transport
natural emissions
wet deposition
scavenging
sedi-mentation
advection
convection
diffusion
OCEAN AND BIOSPHERE

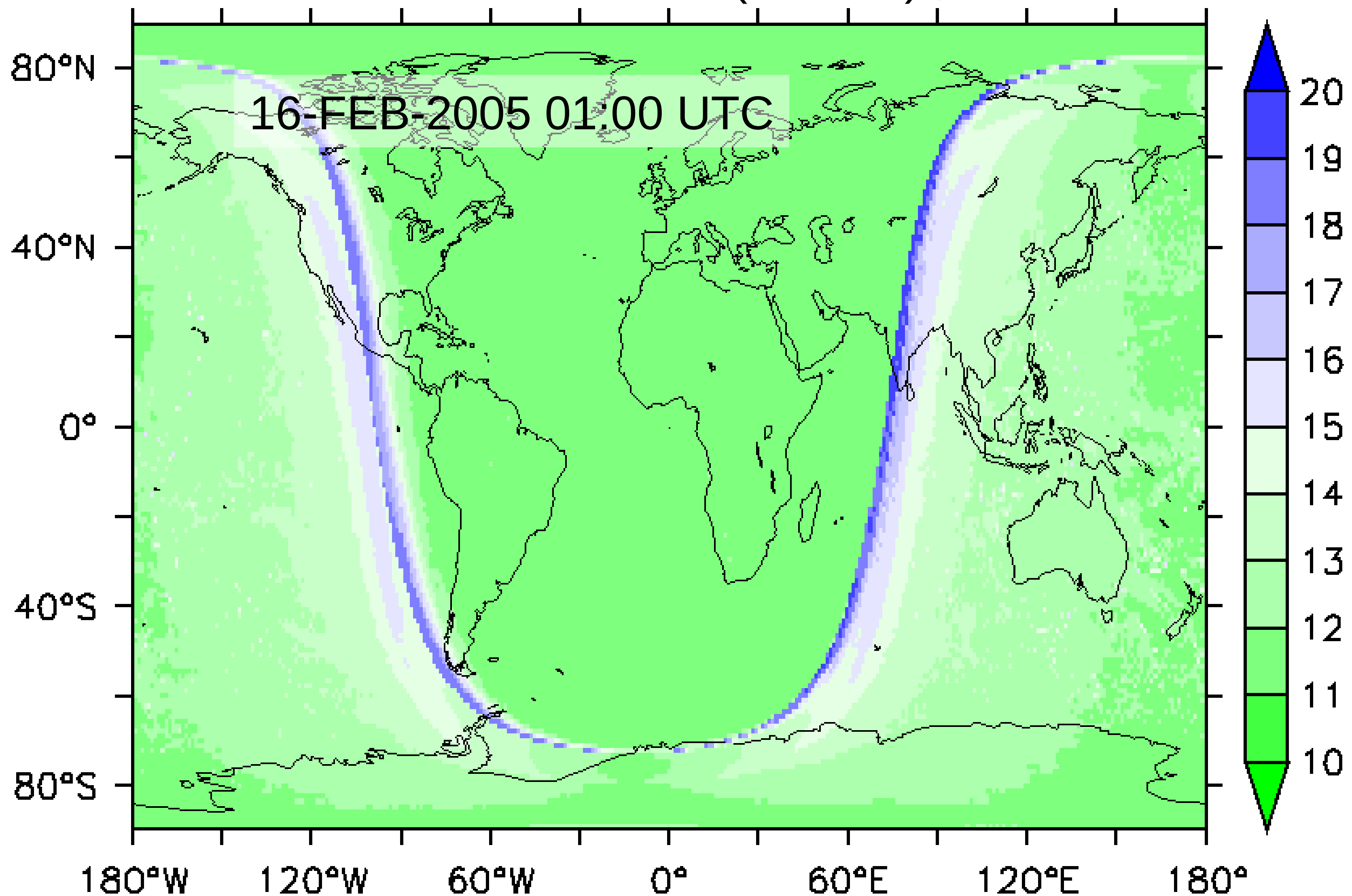

Dynamics
RAD4ALL
JVAL
HETCHEM
PLUMEGAS
MECCA
radiation
hydrological cycle
H2O
SCAV
OFFLEM
MECCA_AERO
gas phase
aerosol
clouds
CLOUD
PSC
GMXE
MADE
M7
DRYDEP
anthropogenic emissions
dry deposition
TNUDGE
Transformation
Emission
TRACER
Deposition
Transport
natural emissions
wet deposition
SCAV
scavenging
CONVECT
ONLEM
sedi-mentation
advection
convection
diffusion
LNOX
SEDI
CVTRANS
AIRSEA
OCEAN AND BIOSPHERE


Dynamics
RAD4ALL
JVAL
HETCHEM
PLUMEGAS
MECCA
OFFLEM
radiation
gas phase
PSC
anthropogenic emissions
dry deposition
TNUDGE
Emission
TRACER
Deposition
Transport
natural emissions
wet deposition
SCAV
scavenging
CONVECT
sedi-mentation
advection
convection
diffusion
ONLEM
LNOX
AIRSEA
SEDI
CVTRANS
OCEAN AND BIOSPHERE
photo-chemistry:
...
A + B → C + D :: J(light, ...)
B + E → F + G :: k(p,T, ...)
...
==> ODE system

Number* of sub-time steps of the ODE solver for the kinetic system

T106L90MA → ~ 1.125° x 1.125° x 90 (~80 km) ; Δt = 6 min

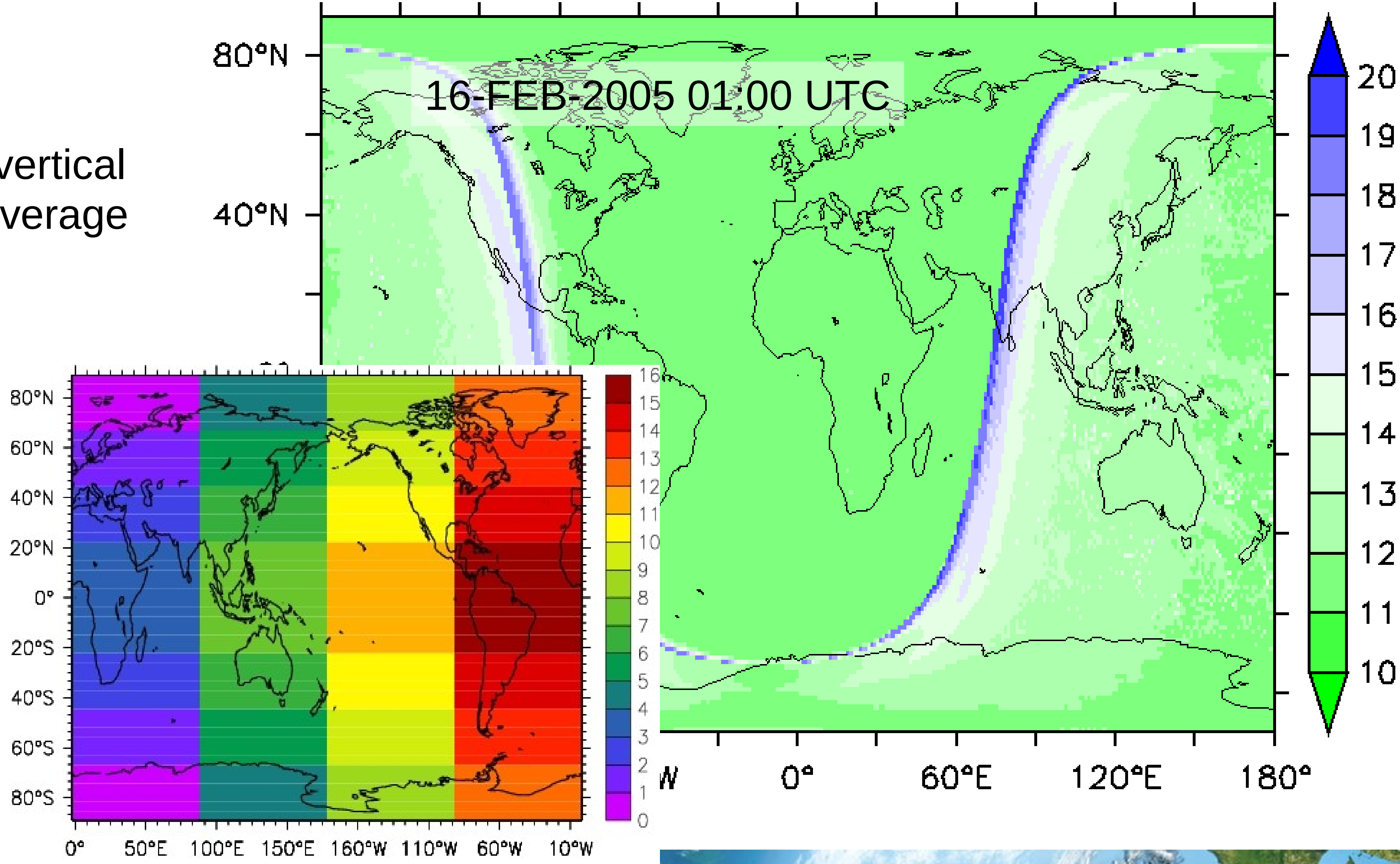
*vertical average



Number* of sub-time steps of the ODE solver for the kinetic system

T106L90MA → ~ 1.125° x 1.125° x 90 (~80 km) ; Δt = 6 min

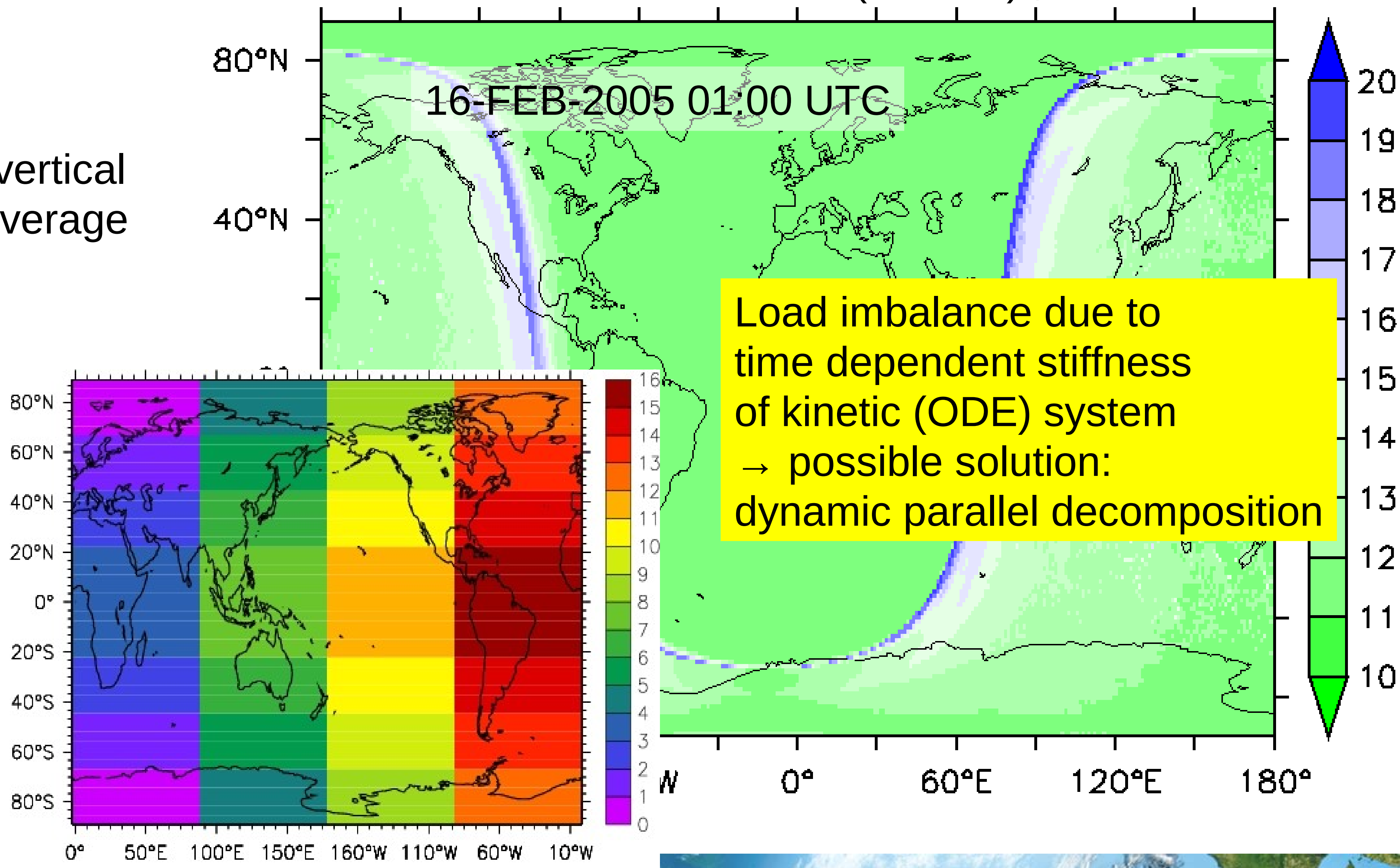
*vertical average



Number* of sub-time steps of the ODE solver for the kinetic system

T106L90MA → ~ 1.125° x 1.125° x 90 (~80 km) ; Δt = 6 min

*vertical
average



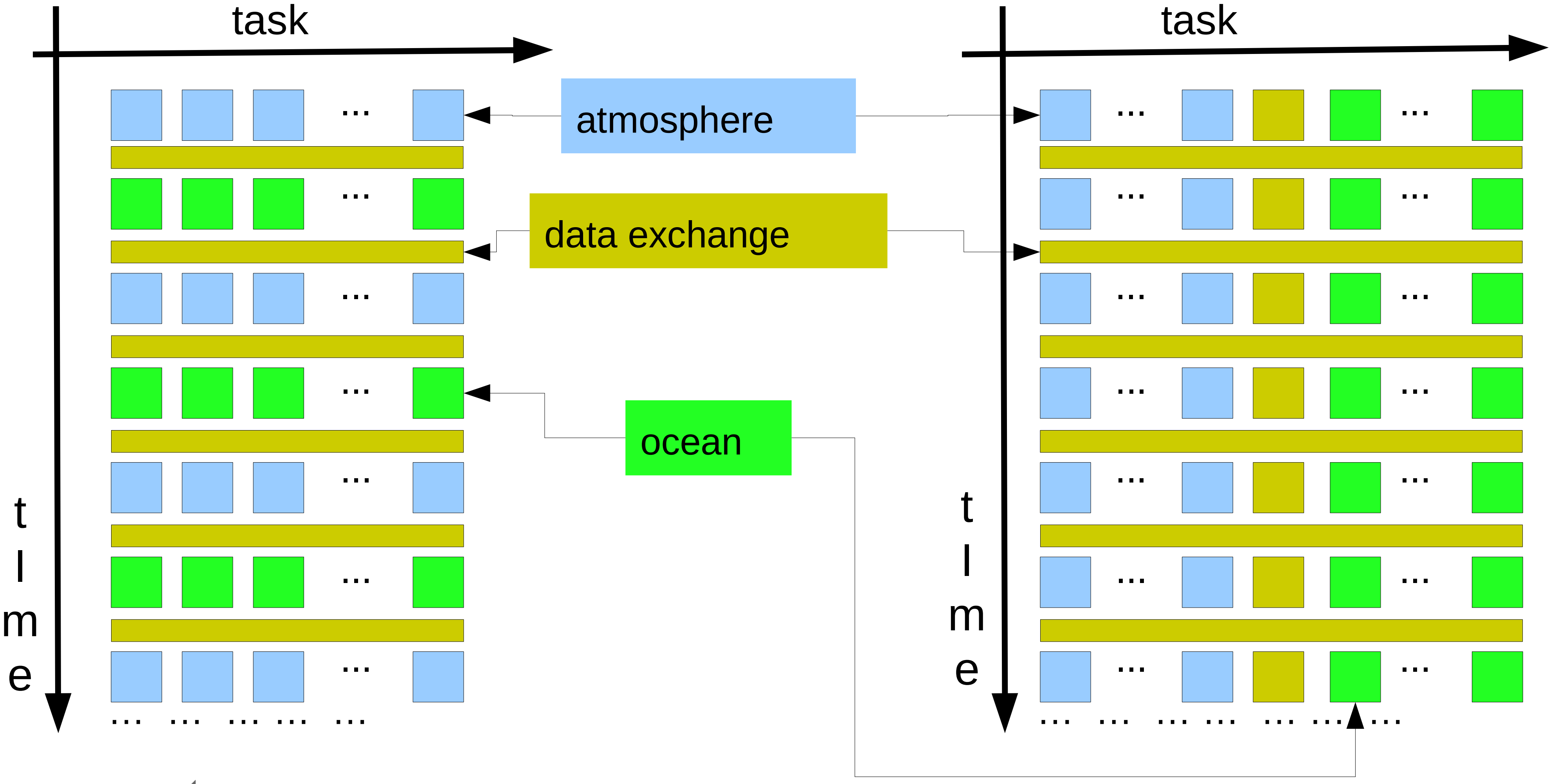
Load imbalance due to
time dependent stiffness
of kinetic (ODE) system
→ possible solution:
dynamic parallel decomposition

Example 2: internal coupling **versus** indirect external coupling

of an Atmosphere – Ocean System (domain coupling)

MPIOM as MESSy submodel
“coupled to” ECHAM5

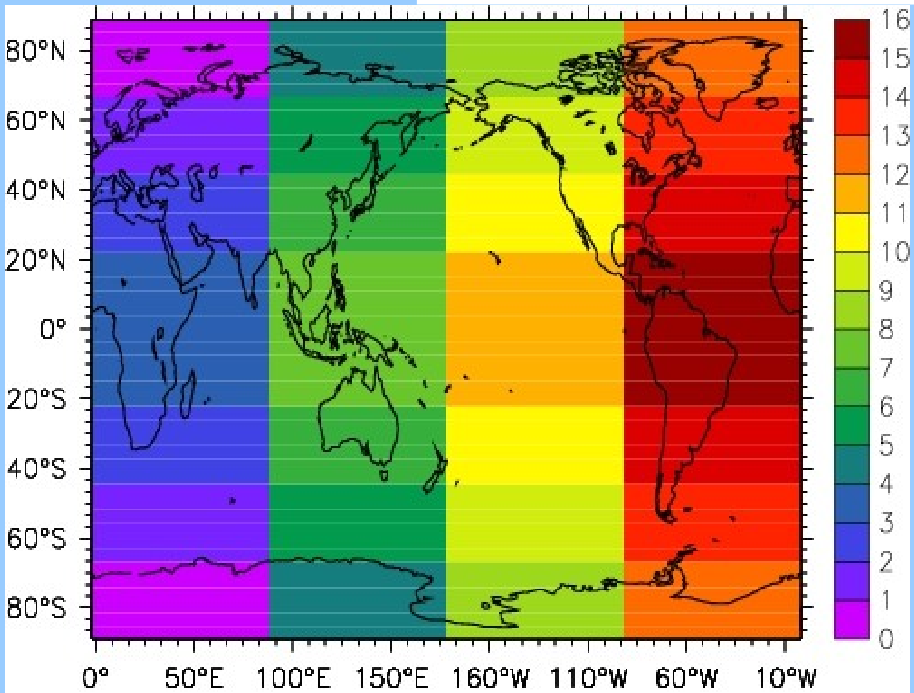
MPIOM – OASIS3 – ECHAM5



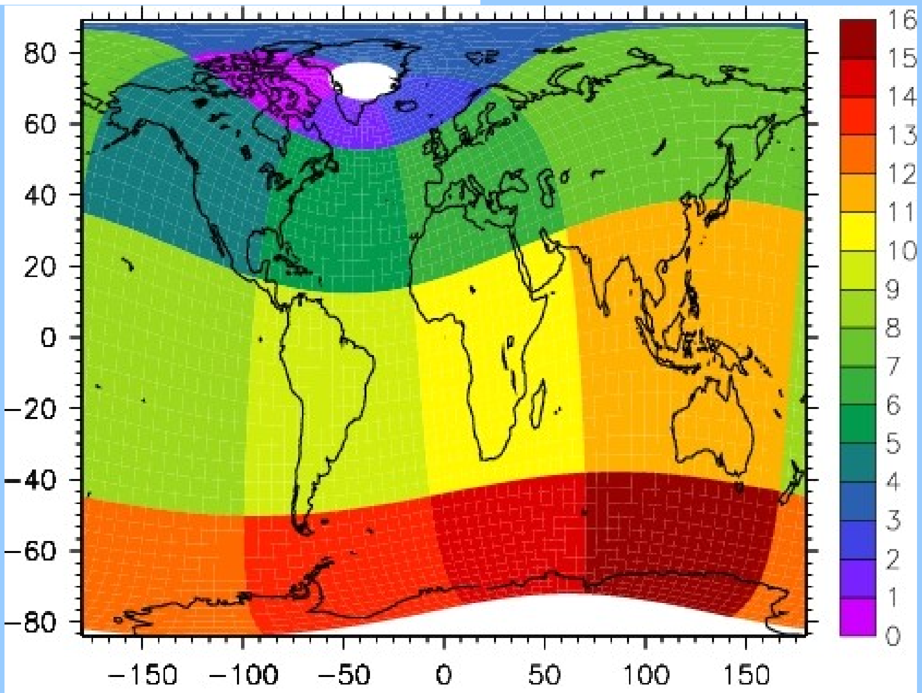
Example 2: internal coupling **versus** indirect external coupling

of an Atmosphere – Ocean System (domain coupling)

Processor ID # for 4 x 4 decomposition



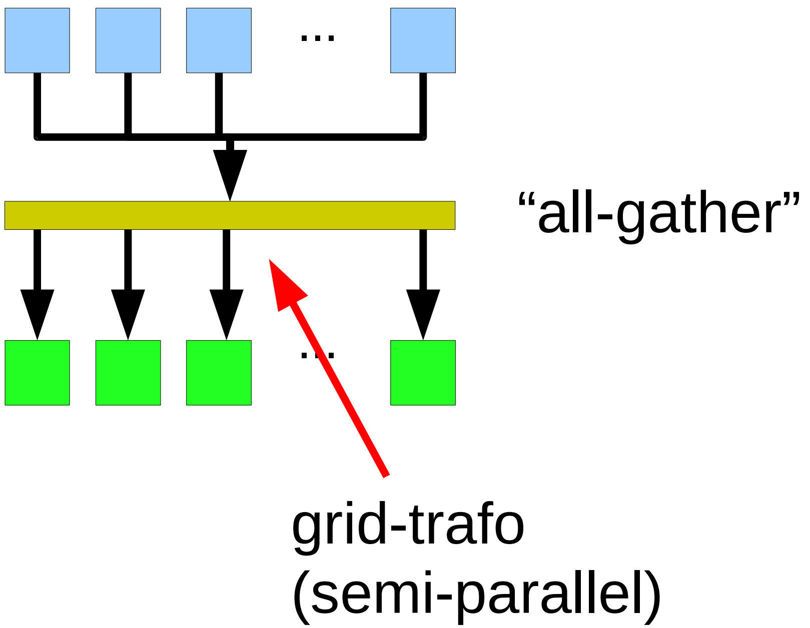
Atmosphere

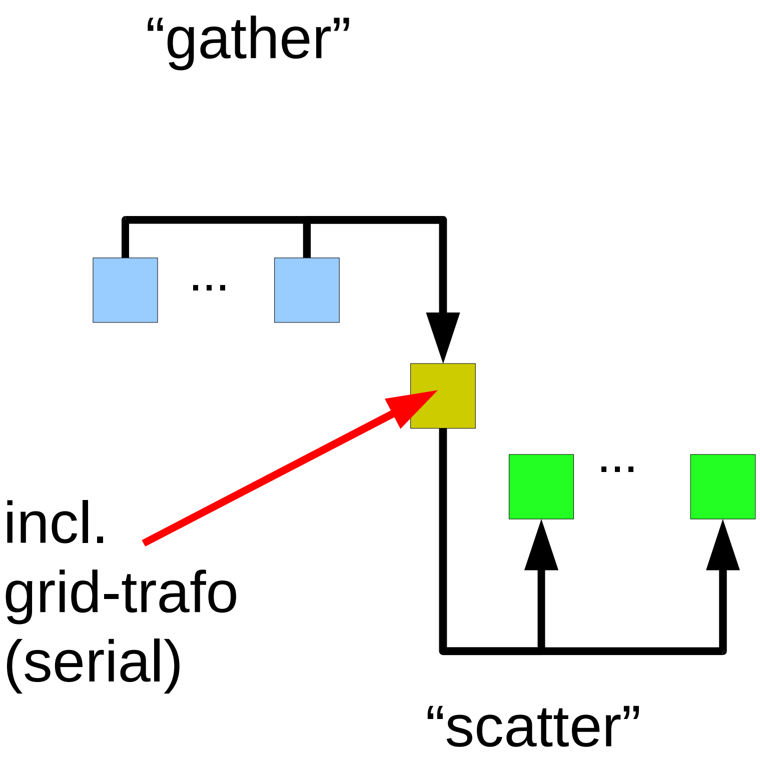


Ocean (rotated grid)

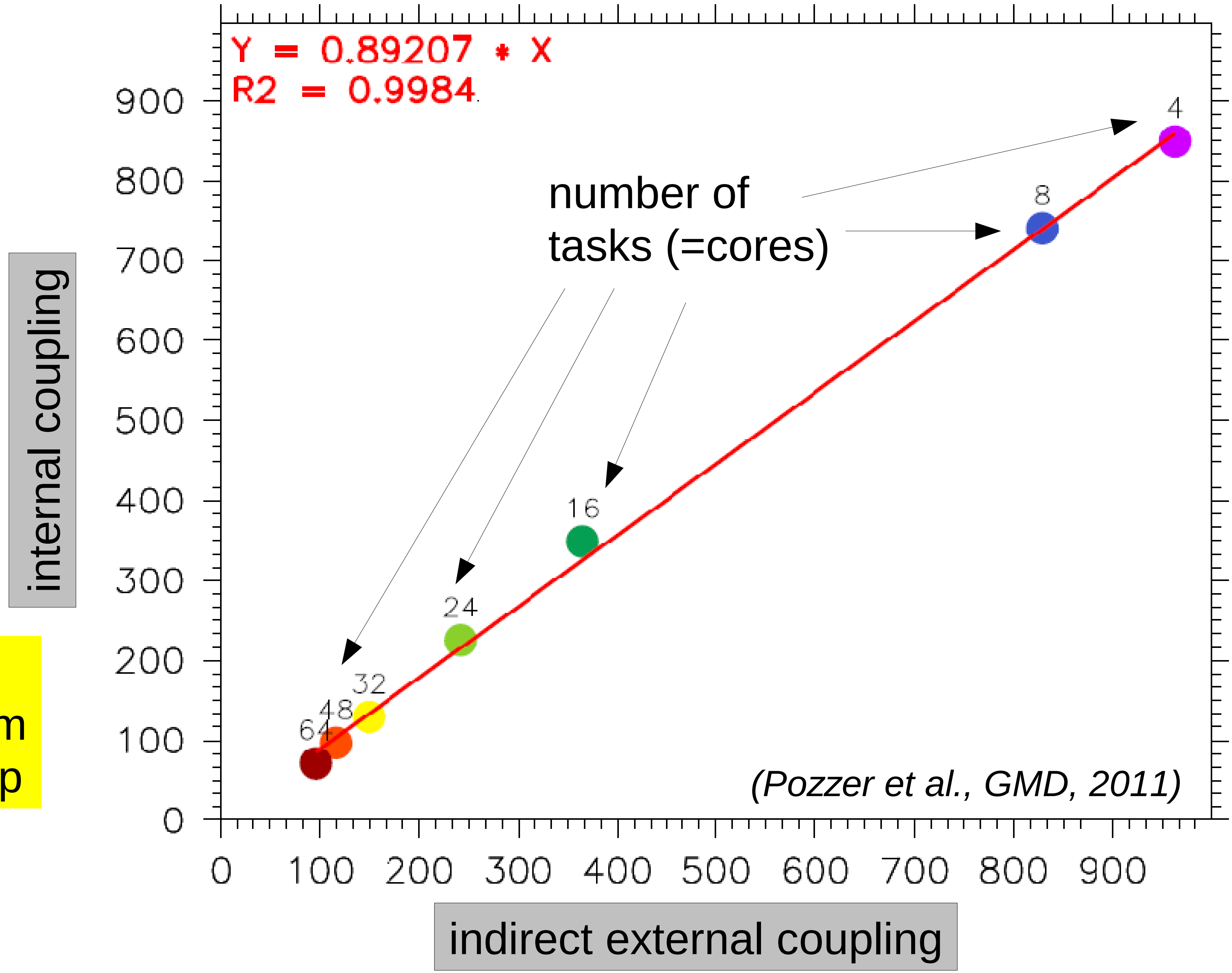
Example 2: internal coupling **versus** indirect external coupling

of an Atmosphere – Ocean System (domain coupling)





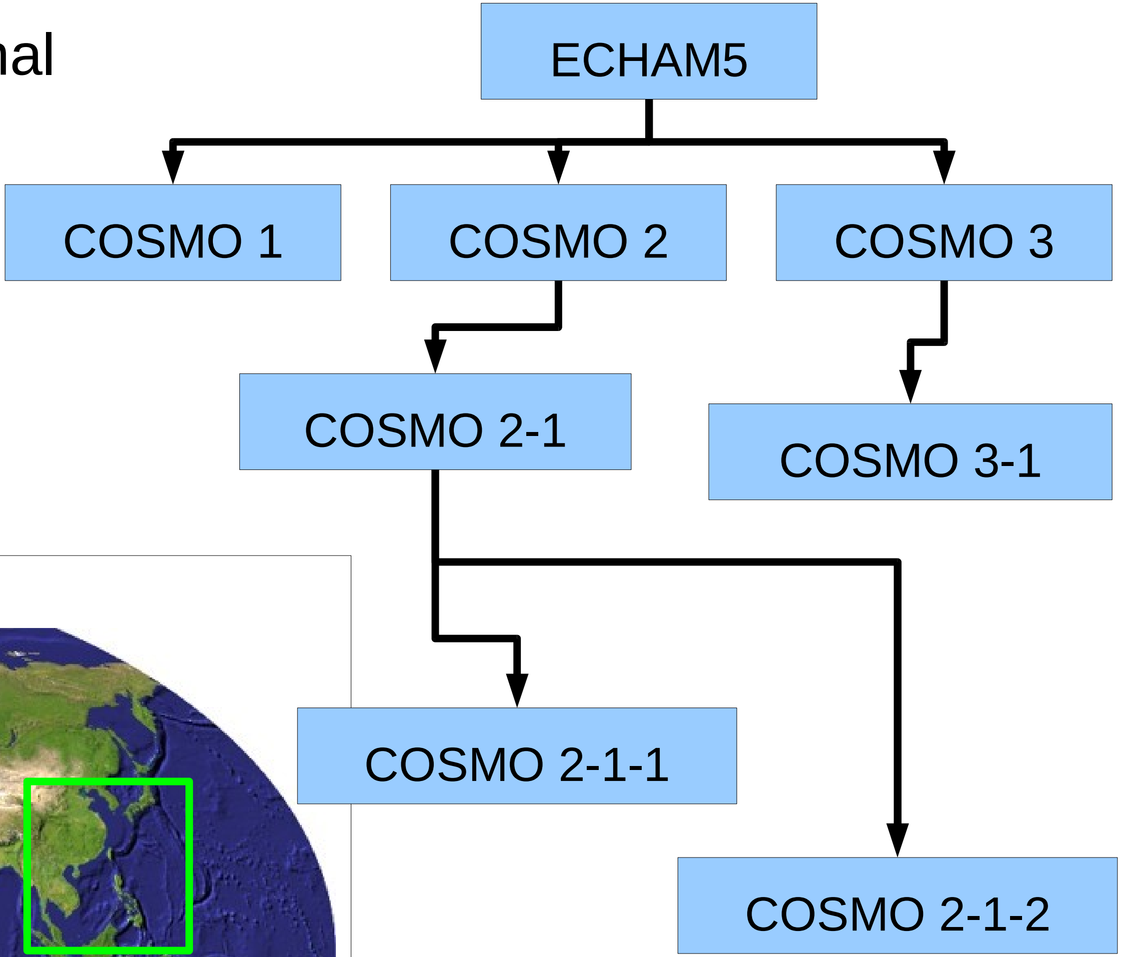
Example 2: Performance (seconds per simulated month)

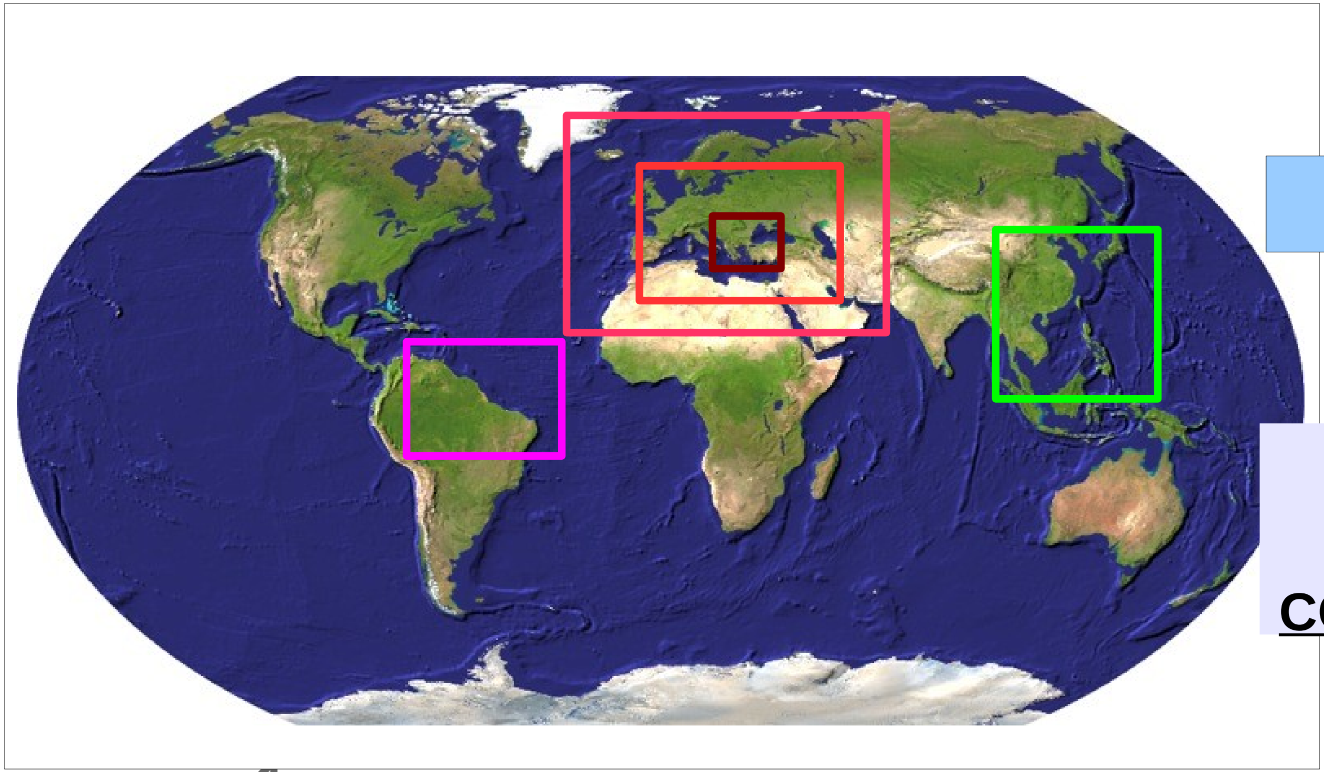


depends on:
- HPC system
- model setup

Example 3: On-line nesting: an alternative way to higher resolution

- 1-way on-line nested global-regional atmospheric model system (zoom)
- multiple instances possible due to client – server architecture of MMD ...





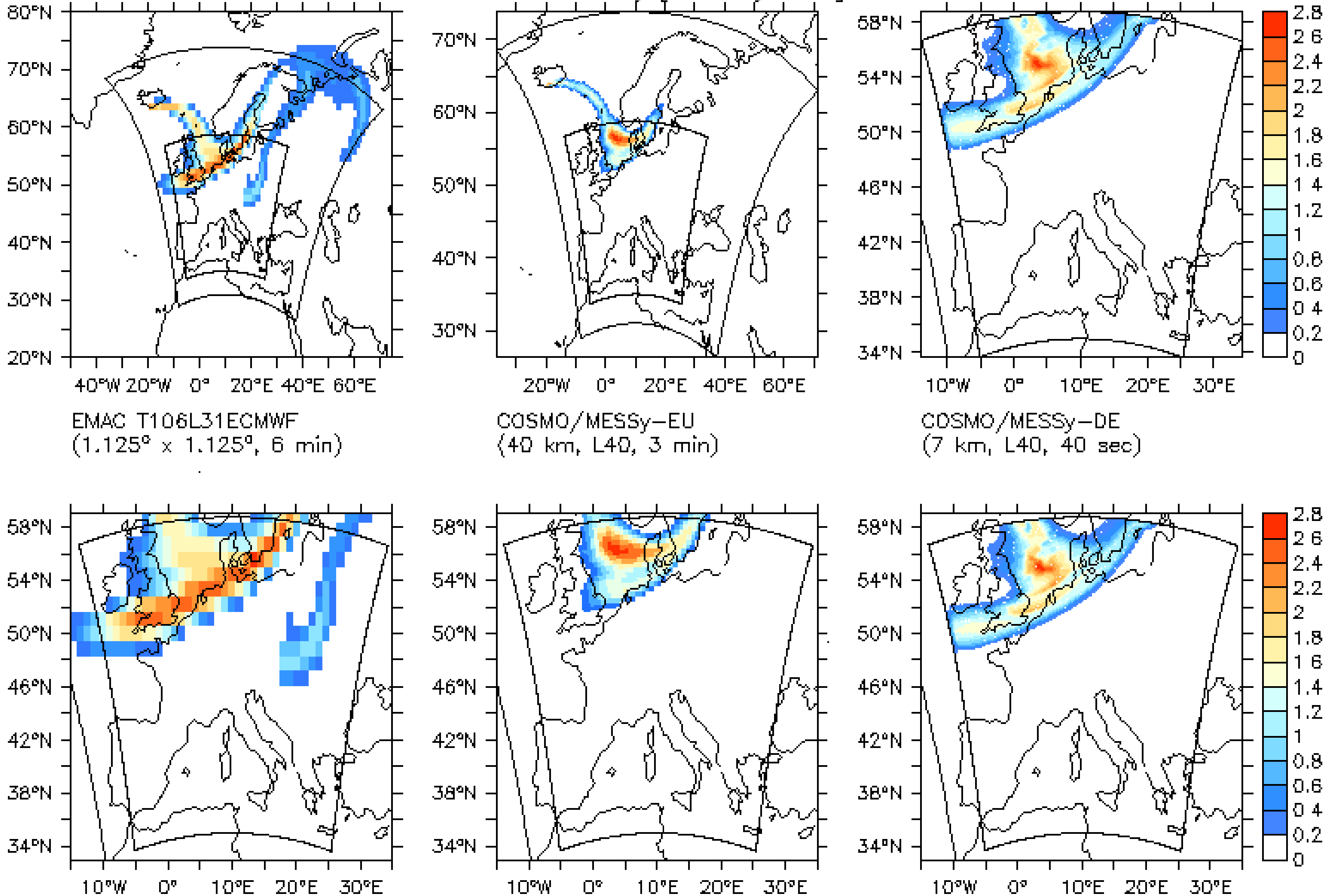
MECO(n):
<u>M</u>ESSy-fied <u>E</u>CHAM and
<u>CO</u>SMO models <u>n</u>-times nested

(Kerkweg & Jöckel, GMD, 2012a,b; Hofmann et al., GMD, 2012)

DLR

MECO(2) simulation of **Eyjafjallajokull eruption plume 2010**

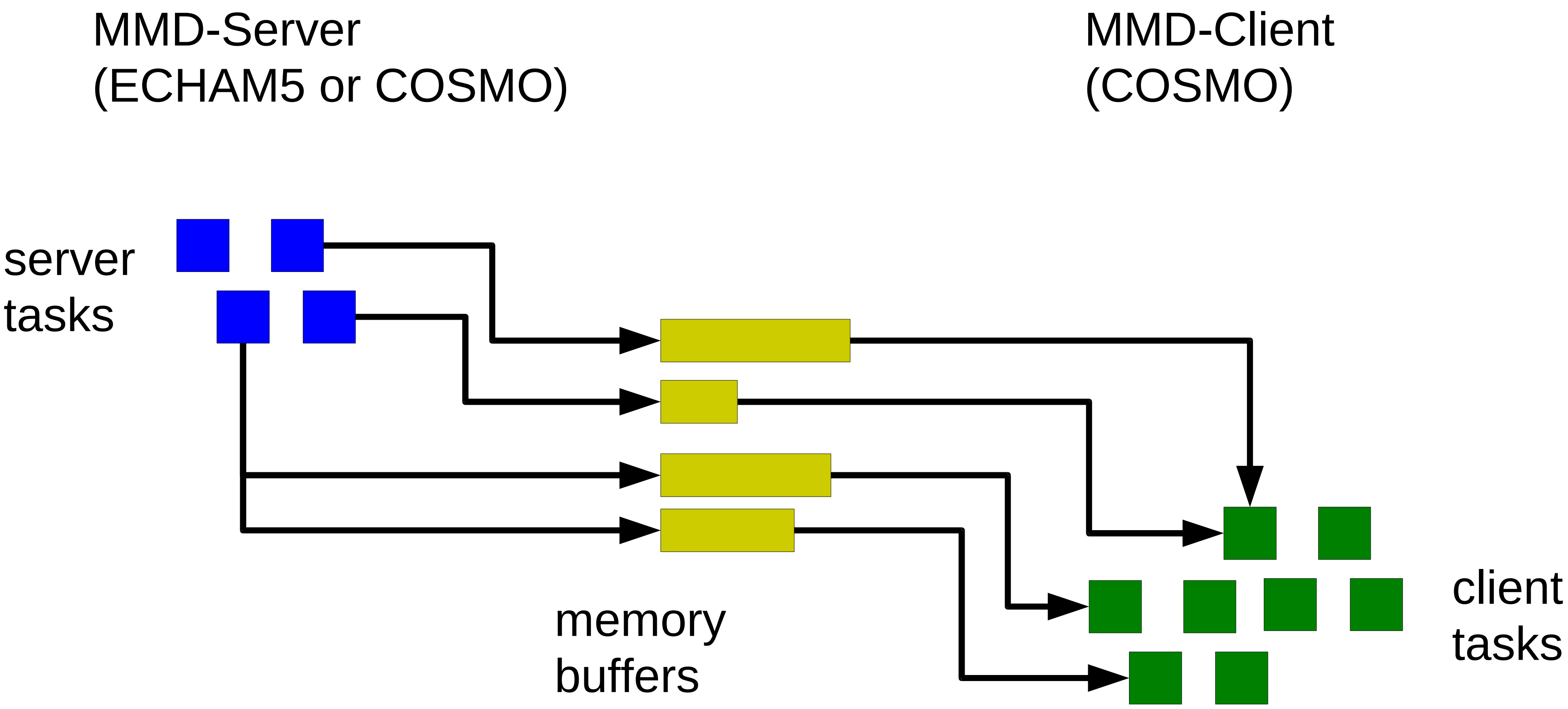
TIME : 17-APR-2010 00:00 column density [arbitrary units]



(c) Patrick Joeckel, DLR, Aug 2011

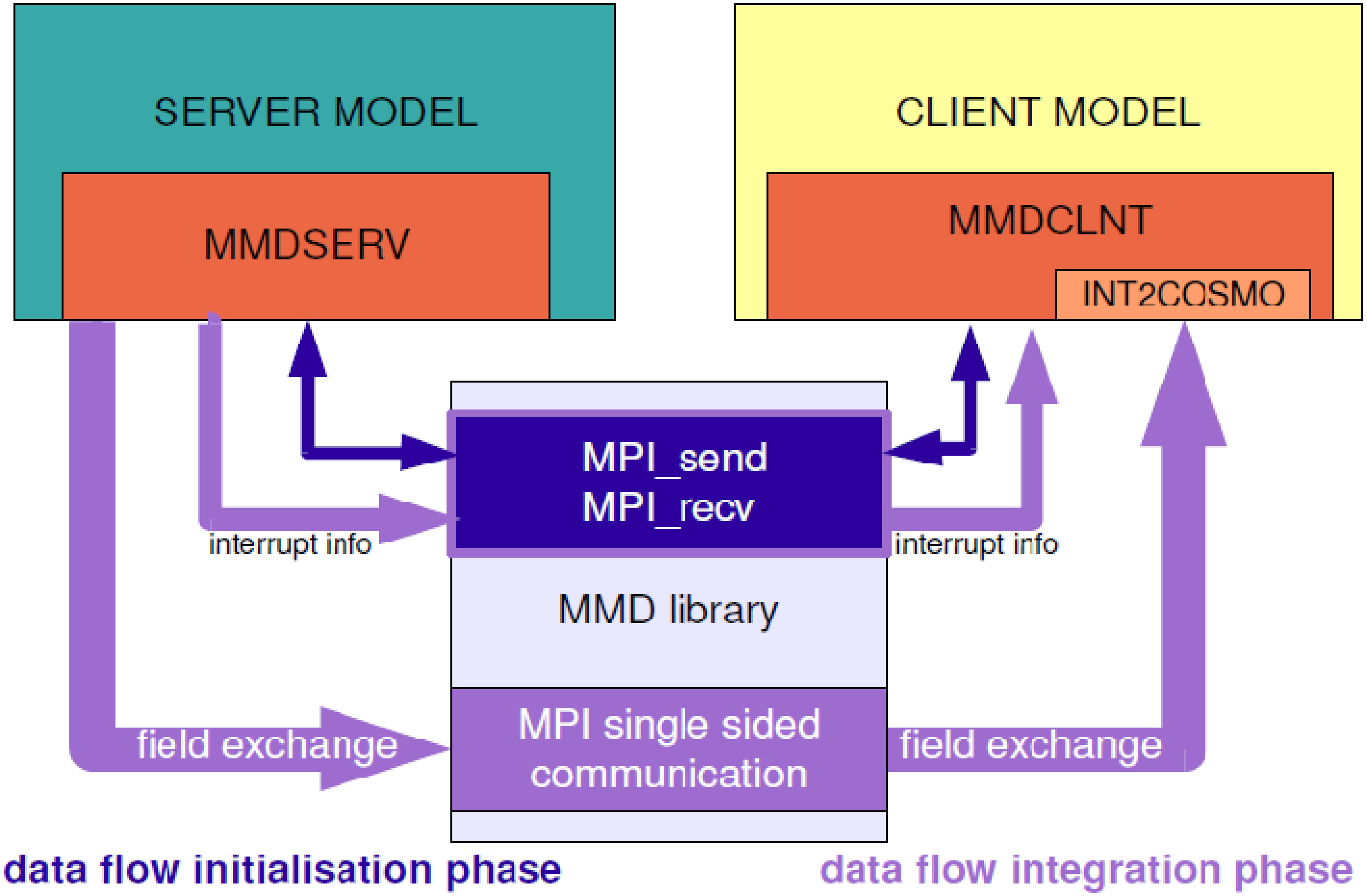
DLR

Example 3: On-line nesting: an alternative way to higher resolution



MPI based, single sided “point-to-point” communication between c&s tasks with overlapping grids

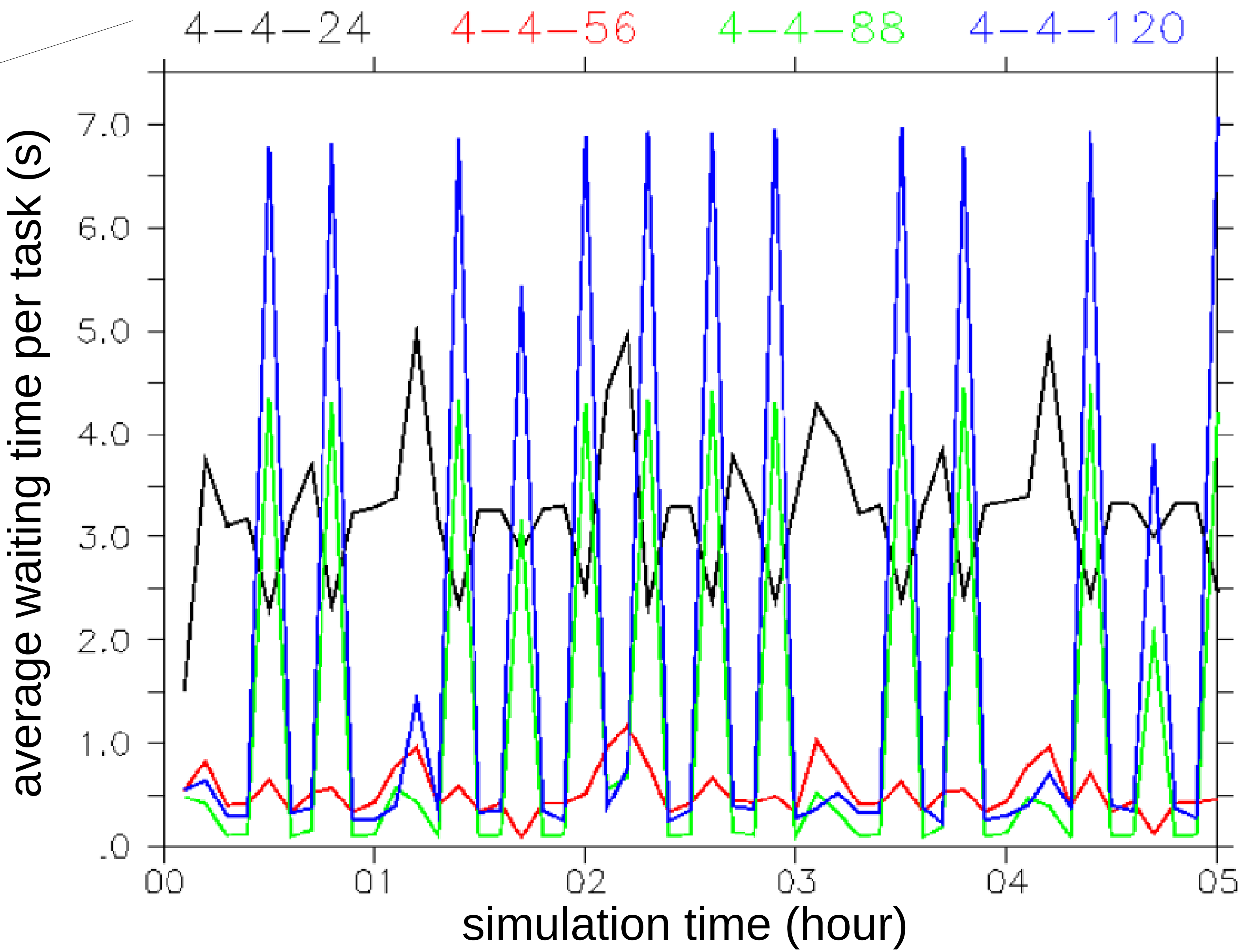
Example 3: On-line nesting: an alternative way to higher resolution



(Kerkweg & Jöckel, GMD, 2012b)

Example 3: On-line nesting: an alternative way to higher resolution

no. of tasks
ECHAM-
COSMO-40 km
COSMO-7 km



4-4-56
is most
efficient !

additional
effort to
optimize
efficiency !

Summary

- ESMs are computationally demanding due to *continuously increasing complexity*
- *Operator splitting* is basis for coupling of model components
- Different coupling methods exist;
  challenge: efficiency – computation versus communication
- Exemplary challenges:
  - Atmospheric Chemistry: *internal coupling*
    → Load Imbalance (parallel decomp.)
  - Atmosphere – Ocean System: *internal* vs. *indirect external* coupling
    → both feasible, best choice depends on model (legacy code!), model setup, HPC-system
  - Global – Regional Nesting: *direct on-line coupling* (client – server approach)
    → complex timing, add. effort to achieve efficiency
- (exascale parallelisation, parallel I/O, memory/core reduction)