www.bsc.es

BSC
Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

INTRODUCTION TO SIMULATION ENVIRONMENT FOR EARTH SCIENCES

PRACE Advanced Training Centres (PATC) Course 2013-2014

Oriol Jorba, Georges Markomanolis and Kim Serradell

Barcelona, 12-13 December 2013

Course Agenda

Day 1

- Session 1
  - 9am – 9:15am: Reception and presentation of attendants
  - 9:15am – 10:30am: Introduction
    - BSC, PRACE and PATC Courses
    - Earth Sciences Modelling
  - 10:30am – 11am: Break
  - 11am – 12am: Introduction to the HPC environment applied to Earth Sciences
    - HPC environment
    - Models
    - Basic Visualization

Course Agenda

Day 1

- 12am – 1pm: Performance analysis of Earth Sciences Models
  - Performance of HPC applications
  - WRF model examples

- 1pm – 2pm: Lunch

- Session 2
  - 2pm – 3:30pm: HPC Environment Tutorial
    - Filesystem handle
    - Bashrc profile
    - Compilation
    - Job Submission
    - Job Monitoring
    - - Basic visualization

Course Agenda

Day 1

- 3:30pm – 4pm: Break

- 4pm – 6pm: Application cases tutorial
  - Weather Research and Forecasting Modeling System (WRF)
  - Nucleus for European Modelling of the Ocean model (NEMO)

- 6pm: End of first day

Course Agenda

Day 2

- Session 3
  - 9am -11am: Visualization and hands-on
    - GrADS, NCL, Python Map Generator
    - 3D visualization (Vapor, Visit)
  - 11am-11:30am: Break
  - 11:30am – 1 pm: Analysis packages and hands-on
    - CDO, NCO, R
  - 1pm – 2pm: Lunch
- Session 4
  - 2pm – 6pm: Practice on HPC environment

Objectives

- The objective of this course is to cover the main basic topics of HPC environment oriented to Earth Sciences applications.

- Attendants will learn how to access an HPC facility, install some Earth Sciences models and utilities, run specific test cases, monitoring an execution in batch mode, visualize and analyse the results.

- More specifically, the course will cover:
  - Basic usage of shell environment, compilers, and parallel programming paradigms (MPI, openMP)
  - Build a targeted Earth Science application
  - Execution and monitoring of submitted experiment
  - Introduction to some commonly used tools to visualize and analyse model outputs

Expected outcomes

The students who finish this course will be able to access, build, run, and visualize a collection of Earth Sciences numerical models. Furthermore, the students will gain a general knowledge on Earth Sciences applications within an HPC environment. The course will provide basic HPC skills for future Earth Sciences modellers.

Evaluation of the course

- At the end of the course, you will be requested to complete a brief questionnaire to evaluate the course:

  – http://events.prace-ri.eu/confDisplayEvaluation.py/display?confId=194

Attenders

- Presentation of participants, main background, previous HPC experience and interests

- Course expectations

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

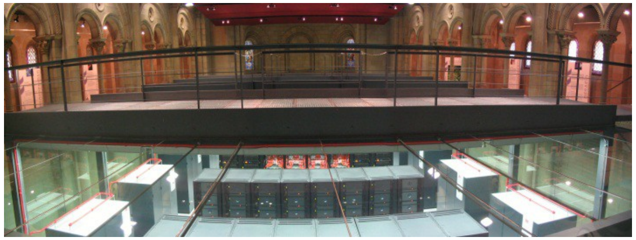
Session 1: Overview of BSC, PRACE and PATC course

Outline

- PATC Courses
  - The Barcelona Supercomputing Center (BSC)
  - The Earth Sciences Department of BSC
  - The PRACE Project and PATC Courses

BSC-CNS

- Barcelona Supercomputing Center – Centro Nacional de Supercomputación (BSC-CNS) is the Spanish National Laboratory in supercomputing.



- The BSC mission:
  - To investigate, develop and manage technology to facilitate the advancement of science.

- The BSC objectives:
  - To perform R&D in Computer Sciences and e-Sciences
  - To provide Supercomputing support to external research.

- BSC is a consortium that includes:
  - the Spanish Government – 51%
  - the Catalan Government – 37%
  - the Technical University of Catalonia – 12%

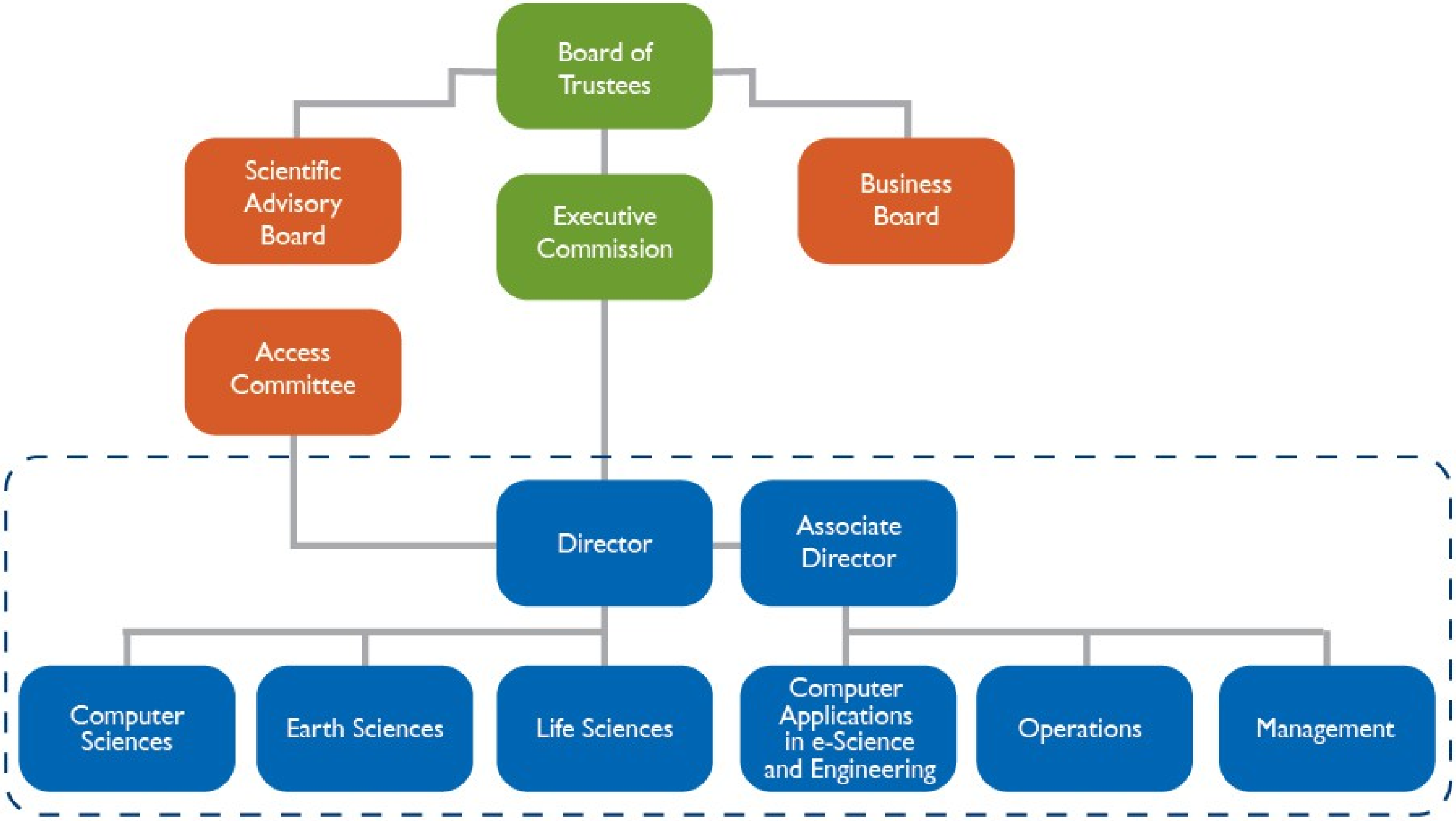
GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

Generalitat de Catalunya
Departament d'Economia i Coneixement

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

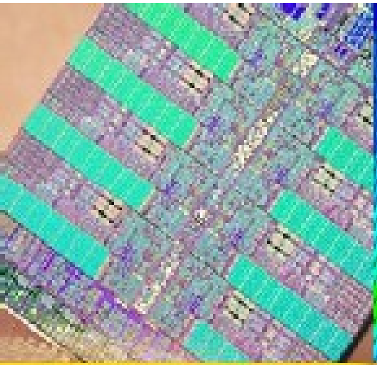
BSC Organization

The BSC is a fusion of a classic Scientific Support Structure and a classic Research Institute.

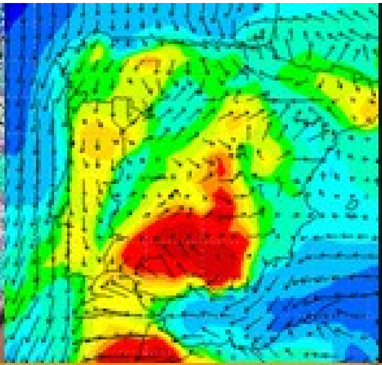


BSC Scientific & Technical Departments

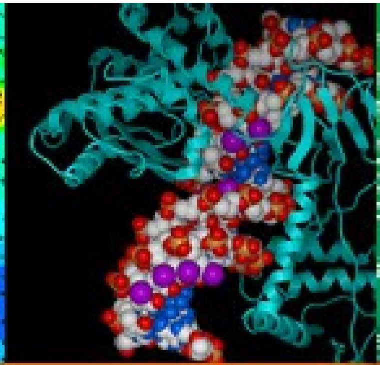
www.bsc.es



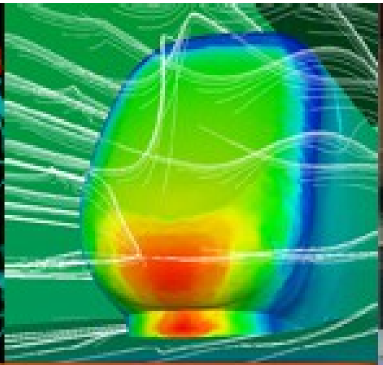
COMPUTER SCIENCES



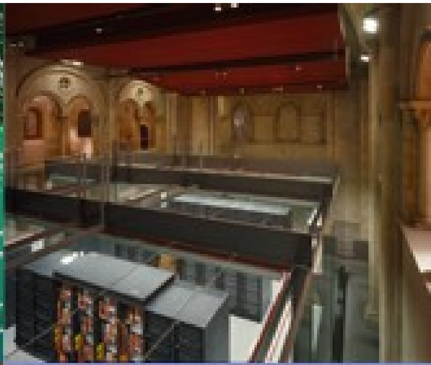
EARTH SCIENCES



LIFE SCIENCES



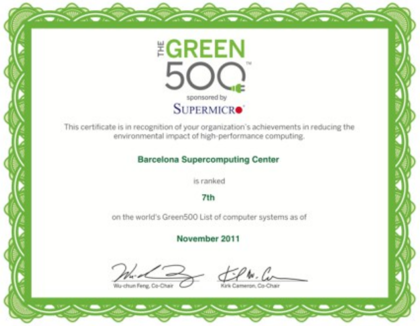
COMPUTER APPLICATIONS



MARENOSTRUM SUPPORT & SERVICES

BSC Current Resources

- MareNostrum 2013
  - 48448 Intel SandyBridge-EP cores
  - 1 PFlops


THE GREEN 500™

sponsored by SUPERMICRO

This certificate is in recognition of your organization's achievements in reducing the environmental impact of high-performance computing.

Barcelona Supercomputing Center

is ranked

7th

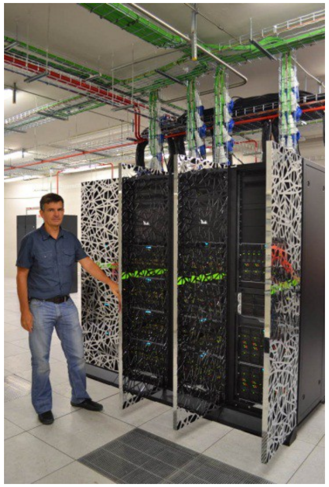
on the world's Green500 List of computer systems as of

November 2011

Wu-chun Feng, Co-Chair

Kirk Cameron, Co-Chair

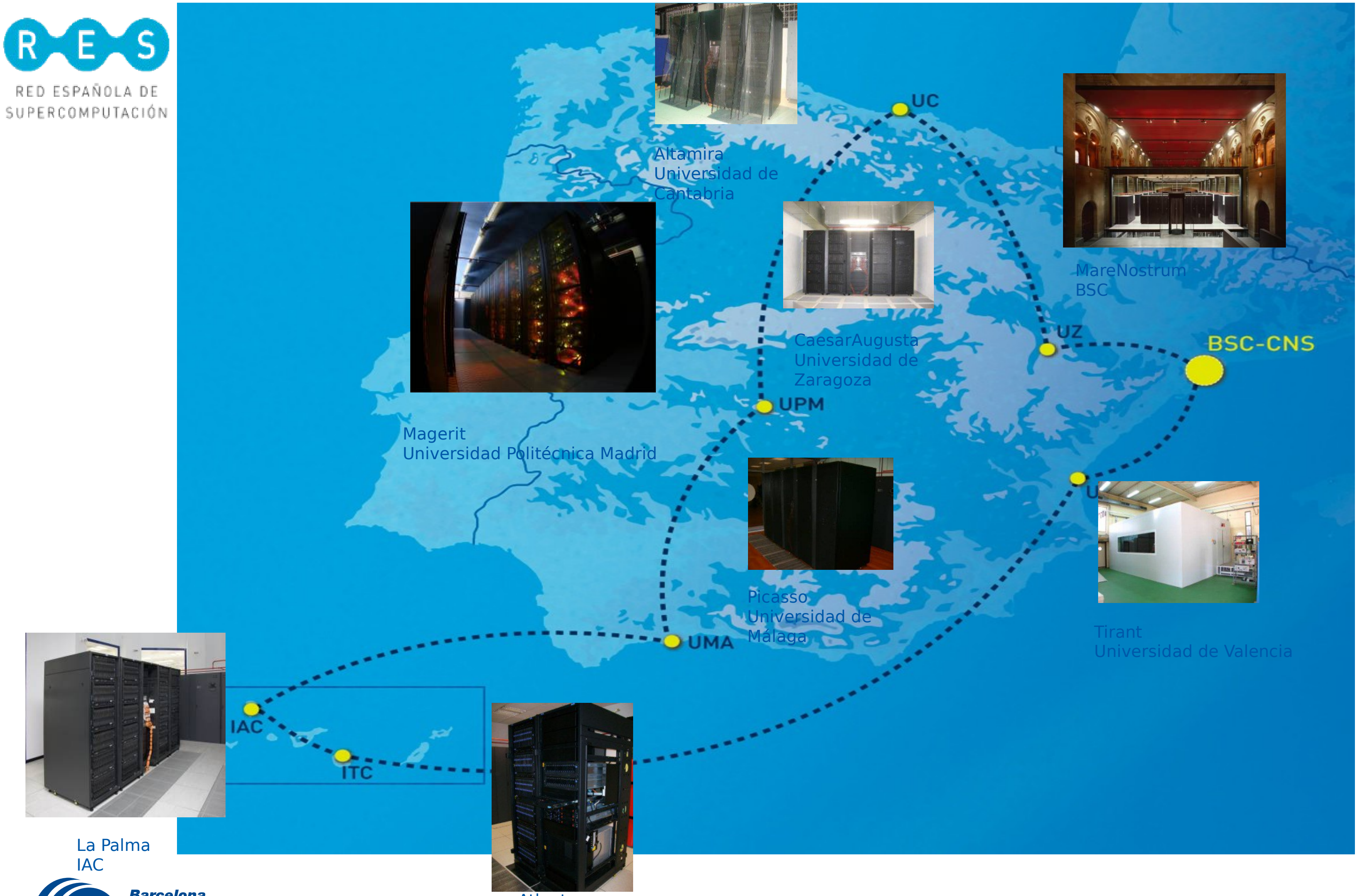
- MinoTauro 2011
  - 128 compute nodes
  - 182 TFlops



- HPC Storage and Backup:
  - 2.5 PB disk
  - 6.0 PB tapes Robot

BSC in Spain

RES
RED ESPAÑOLA DE SUPERCOMPUTACIÓN



Altamira
Universidad de Cantabria

MareNostrum
BSC

CaesarAugusta
Universidad de Zaragoza

Magerit
Universidad Politécnica Madrid

Picasso
Universidad de Málaga

Tirant
Universidad de Valencia

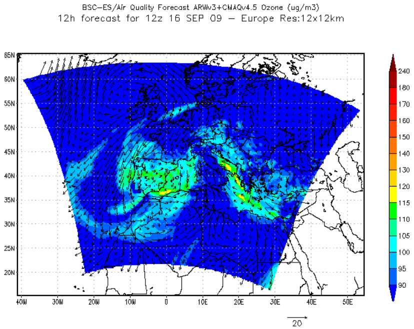
La Palma
IAC

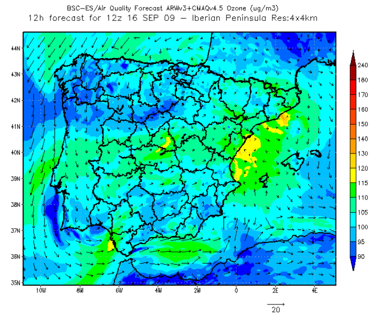
Atlante
ITC

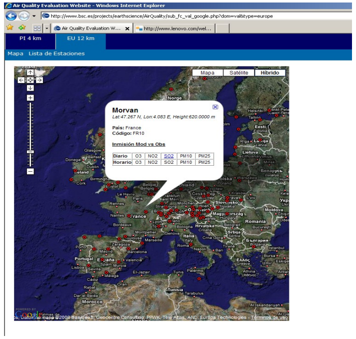
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Earth Sciences Department (www.bsc.es/earth-sciences)

- Research in the Earth Sciences area is devoted to the development and implementation of regional and global state-of-the-art models for short-term air quality forecast and long-term climate applications.


BSC–ES/Air Quality Forecast ARWv3+CMAQv4.5 Ozone (ug/m3)
12h forecast for 12z 16 SEP 09 – Europe Res:12x12km
65N
60N
55N
50N
45N
40N
35N
30N
25N
20N
40W
30W
20W
10W
0
10E
20E
30E
40E
50E
240
180
170
160
150
140
130
120
115
110
105
100
95
90
20


BSC–ES/Air Quality Forecast ARWv3+CMAQv4.5 Ozone (ug/m3)
12h forecast for 12z 16 SEP 09 – Iberian Peninsula Res:4x4km
44N
43N
42N
41N
40N
39N
38N
37N
36N
35N
10W
8W
6W
4W
2W
0
2E
4E
240
180
170
160
150
140
130
120
115
110
105
100
95
90
20


Air Quality Evaluation Website - Windows Internet Explorer

http://www.bsc.es/projects/earthscience/AirQuality/sub_fc_val_google.php?dom=val&type=europe

Air Quality Evaluation W... | http://www.lenovo.com/wel...

PI 4 km | EU 12 km

Mapa | Lista de Estaciones

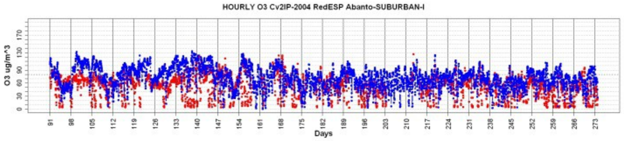
Morvan

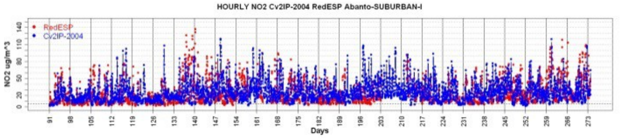
Lat:47.267 N, Lon:4.083 E, Height:620.0000 m

País: France
Código: FR10

Inmisión Mod vs Obs

Diario	O3	NO2	SO2	PM10	PM25
Horario	O3	NO2	SO2	PM10	PM25

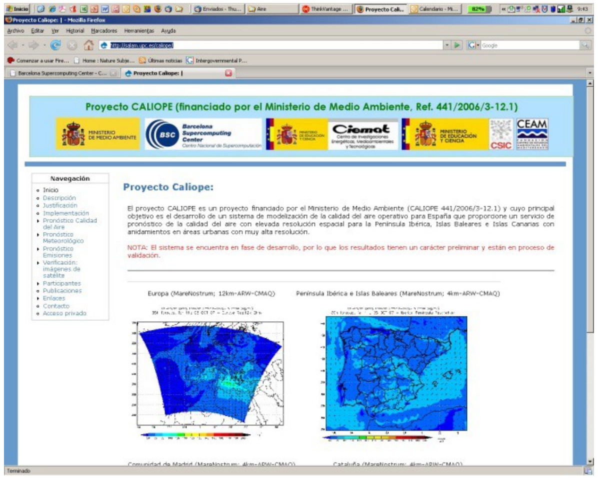




- ES maintains two daily operational systems: AQF CALIOPE and MD forecasts: BSC-DREAM8b and NMMB/BSC-CTM.

Earth Sciences research lines

Air Quality Forecast



Proyecto Caliope: | - Mozilla Firefox

Archivo Editar Ver Historial Marcadores Herramientas Ayuda
http://salam.upc.es/caliope/
Google
Comenzar a usar Fire... Home : Nature Subje... Últimas noticias Intergovernmental P...
Barcelona Supercomputing Center - C...
Proyecto Caliope: |
Terminado

Proyecto CALIOPE (financiado por el Ministerio de Medio Ambiente, Ref. 441/2006/3-12.1)

MINISTERIO DE MEDIO AMBIENTE
Barcelona Supercomputing Center Centro Nacional de Supercomputación
Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas
MINISTERIO DE EDUCACIÓN Y CIENCIA
CSIC
CEAM

Navegación

- Inicio
- Descripción
- Justificación
- Implementación
- Pronóstico Calidad del Aire
- Pronóstico Meteorológico
- Pronóstico Emisiones
- Verificación: imágenes de satélite
- Participantes
- Publicaciones
- Enlaces
- Contacto
- Acceso privado

Proyecto Caliope:

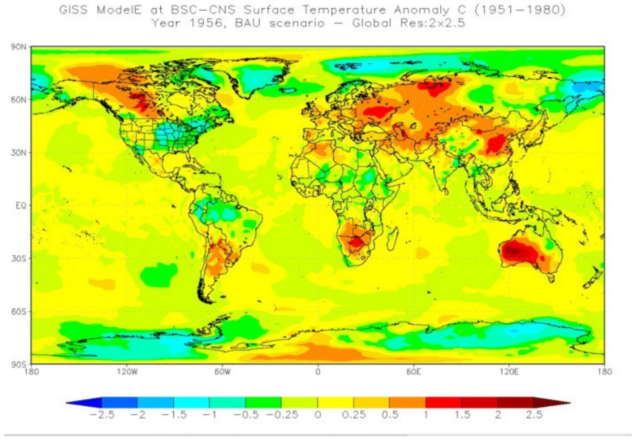
El proyecto CALIOPE es un proyecto financiado por el Ministerio de Medio Ambiente (CALIOPE 441/2006/3-12.1) y cuyo principal objetivo es el desarrollo de un sistema de modelización de la calidad del aire operativo para España que proporcione un servicio de pronóstico de la calidad del aire con elevada resolución espacial para la Península Ibérica, Islas Baleares e Islas Canarias con anidamientos en áreas urbanas con muy alta resolución.

NOTA: El sistema se encuentra en fase de desarrollo, por lo que los resultados tienen un carácter preliminar y están en proceso de validación.

Europa (MareNostrum; 12km-ARW-CMAQ) | Península Ibérica e Islas Baleares (MareNostrum; 4km-ARW-CMAQ)

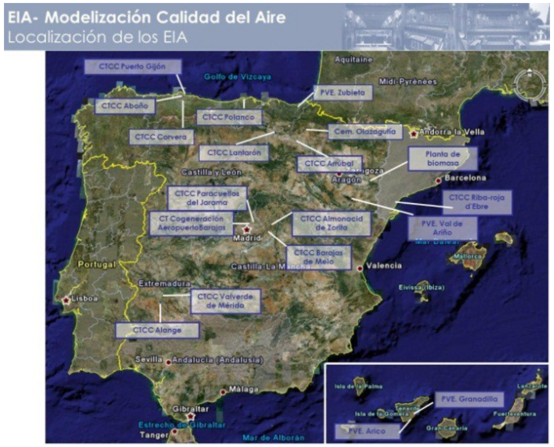
Comunidad de Madrid (MareNostrum; 4km-ARW-CMAQ) | Cataluña (MareNostrum; 4km-ARW-CMAQ)

Climate change modelling



GISS ModelE at BSC-CNS Surface Temperature Anomaly C (1951-1980)
Year 1956, BAU scenario – Global Res:2x2.5
90N
60N
30N
EQ
30S
60S
90S
180
120W
60W
0
60E
120E
180
-2.5
-2
-1.5
-1
-0.5
-0.25
0
0.25
0.5
1
1.5
2
2.5

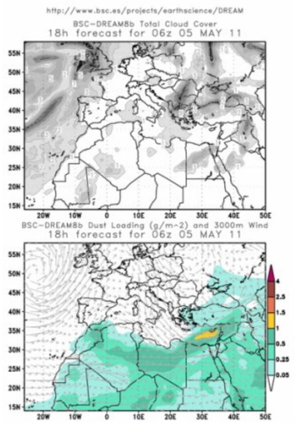
Transfer technology (EIA and AQ studies)



EIA- Modelización Calidad del Aire

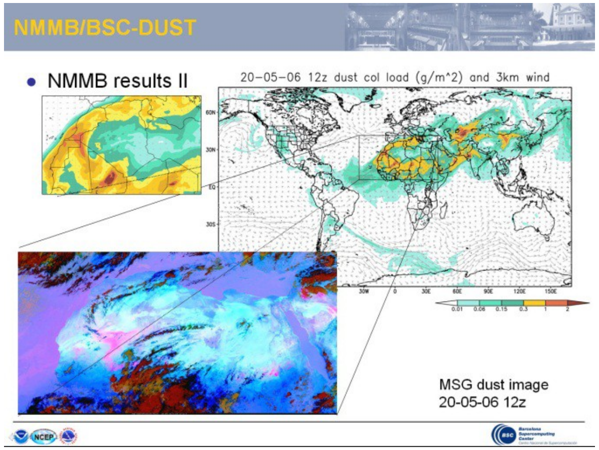
Localización de los EIA

Mineral dust transport: BSC-DREAM8b



http://www.bsc.es/projects/earthscience/DREAM
BSC-DREAM8b Total Cloud Cover
18h forecast for 06z 05 MAY 11
55N
50N
45N
40N
35N
30N
25N
20N
15N
20W
10W
0
10E
20E
30E
40E
50E
BSC-DREAM8b Dust Loading (g/m-2) and 3000m Wind
18h forecast for 06z 05 MAY 11
2.5
1.5
1
0.5
0.25
0.05

Atmospheric modelling: development of NMMB/BSC-CTM



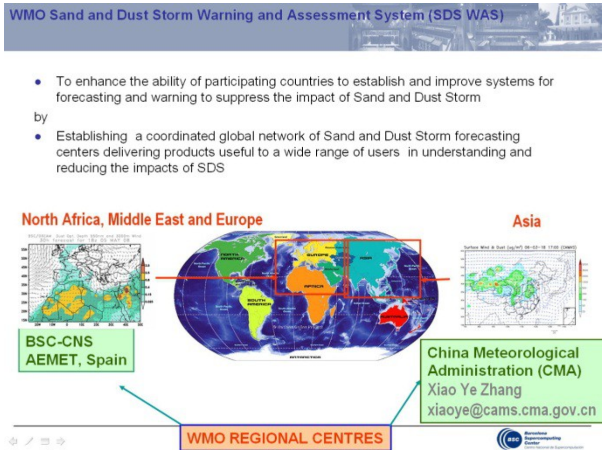
NMMB/BSC-DUST

- NMMB results II

20-05-06 12z dust col load (g/m^2) and 3km wind

MSG dust image 20-05-06 12z

WMO SDS WAS [AEMET-BSC]



WMO Sand and Dust Storm Warning and Assessment System (SDS WAS)

- To enhance the ability of participating countries to establish and improve systems for forecasting and warning to suppress the impact of Sand and Dust Storm

by

- Establishing a coordinated global network of Sand and Dust Storm forecasting centers delivering products useful to a wide range of users in understanding and reducing the impacts of SDS

North Africa, Middle East and Europe

Asia

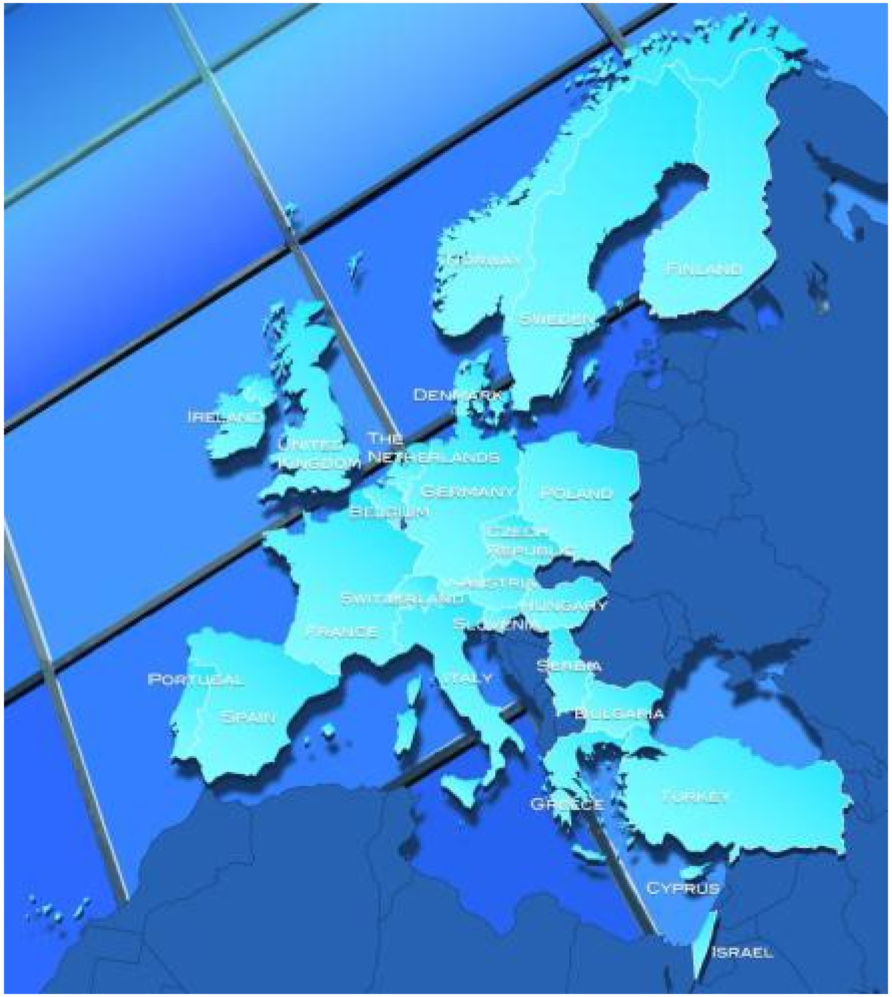
BSC-CNS
AEMET, Spain

China Meteorological Administration (CMA)
Xiao Ye Zhang
xiaoye@cams.cma.gov.cn

WMO REGIONAL CENTRES

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Partnership For Advanced Computing in Europe (PRACE) Project
www.prace-ri.eu



April, 23rd 2010 creation of the legal entity (AISBL) PRACE with seat location in Brussels, Belgium

25 PRACE Members

67+ Million € from EC FP7 for preparatory and implementation phases

Grants INFSO-RI-211528, 261557, and 283493

Complemented by ~ 50 Mio€ from PRACE members

And 400Mio€ by Hosting Members

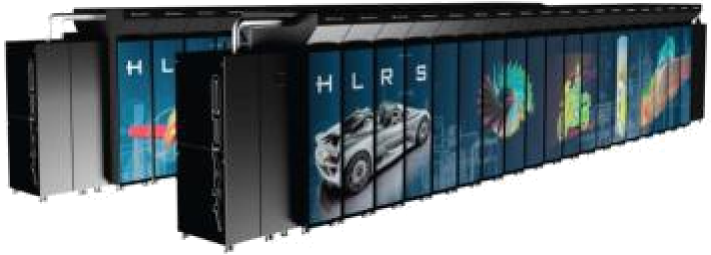
Tier-0 Petaflop-Capability in PRACE

- **2012: approx. 14,6 PF Peak Performance**

BlueGene/P
PRACE@Jülich

Bull Cluster **Curie**
PRACE@GENCI:

2012: BlueGene/Q
PRACE@(CINECA & Jülich)



CRAY HERMIT
PRACE@HLRS

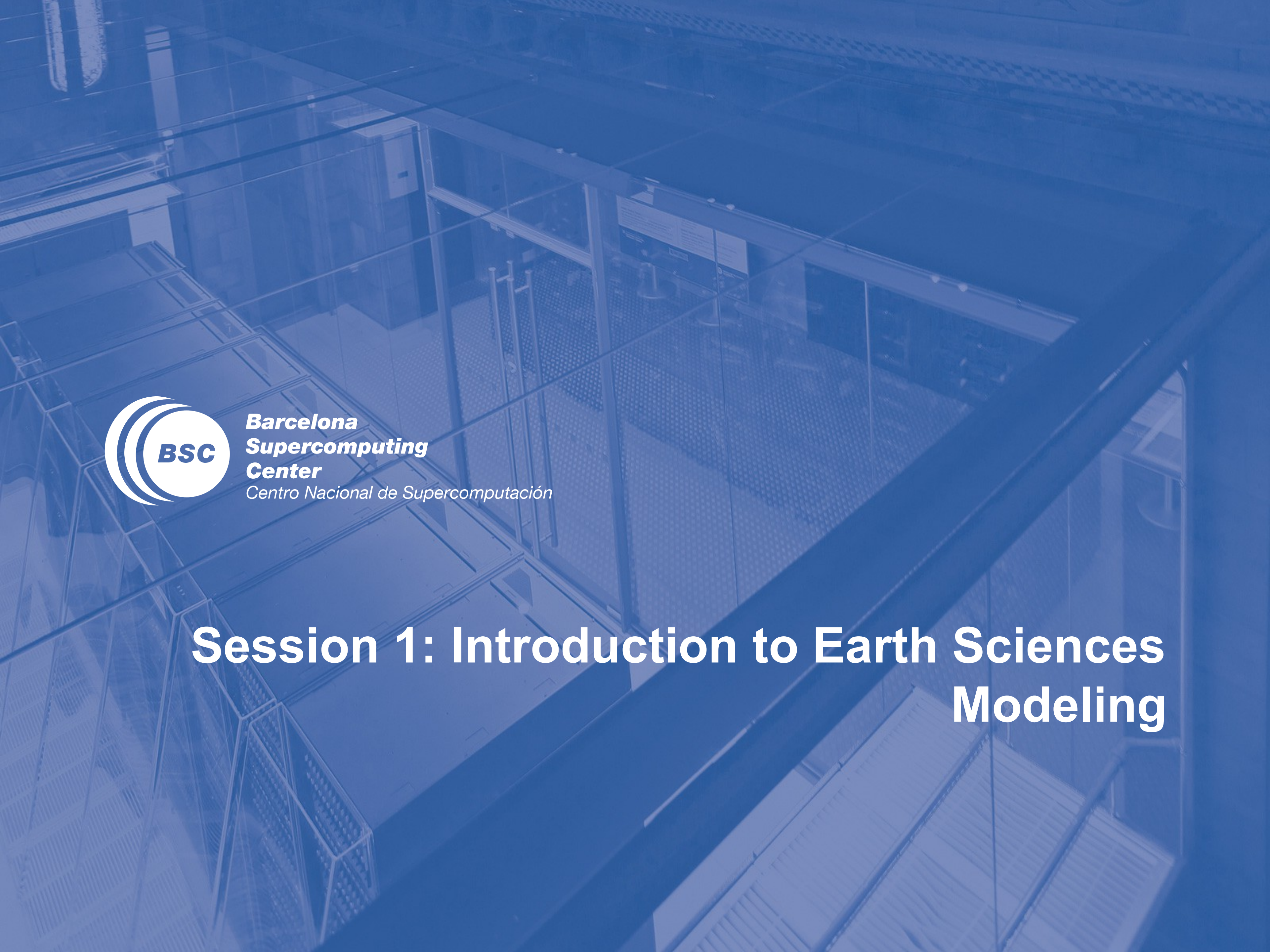
IBM SuperMUC
PRACE@LRZ

IBM in 2012
PRACE@BSC

PRACE Training Courses

- PRACE, the Partnership for Advanced Computing in Europe (www.prace-ri.eu), has selected six of its members' sites: Barcelona Supercomputing Center (Spain), CINECA - Consorzio Interuniversitario (Italy), CSC - IT Center for Science Ltd (Finland), EPCC at the University of Edinburgh (UK), Gauss Centre for Supercomputing (Germany) and Maison de la Simulation (France) as the **first PRACE Advanced Training Centres**.

- The **mission** of the PRACE Advanced Training Centres (PATCs) is to carry out and coordinate **training and education activities** that enable the European research community to **utilise the computational infrastructure available through PRACE**. The long-term vision is that such centres will become the hubs and key drivers of European high-performance computing education.



Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

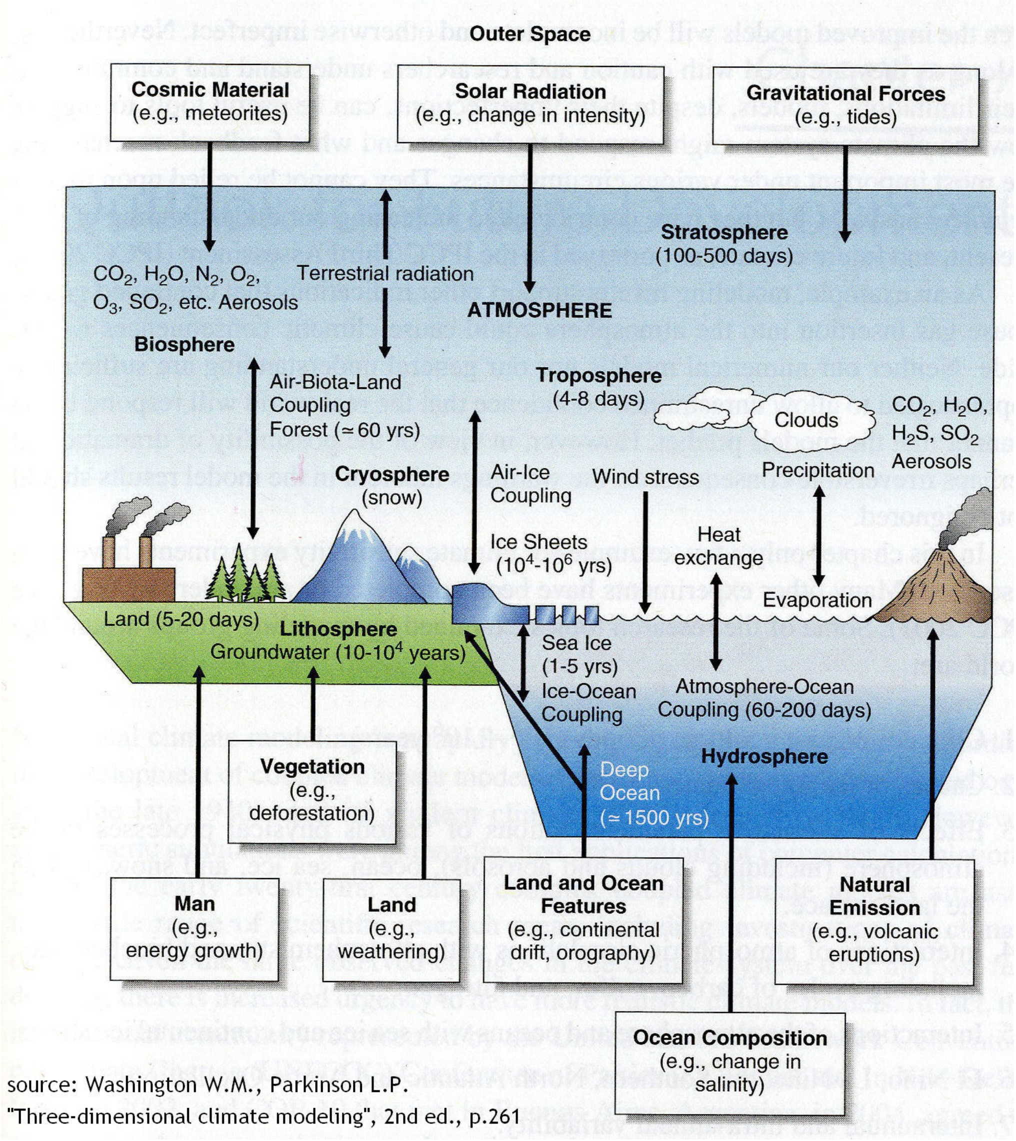
Session 1: Introduction to Earth Sciences Modeling

Outline

- The Earth system and impacts
- How do we model atmospheric processes?
- Examples of numerical modelling results
- Developments at the Earth Sciences Department of BSC

The Earth System

Readapted from: www.metoffice.gov.uk/research/hadleycentre/models/climate_system.html



source: Washington W.M., Parkinson L.P.
"Three-dimensional climate modeling", 2nd ed., p.261

- **Atmosphere:** circulation, the heat transfer to and from the sun, formation of clouds and atmospheric reactive flows that determine the concentrations of its chemicals.
- **Ocean:** interaction of ocean and atmosphere through exchange of momentum, heat and water. The ocean is a heat sink and it is a medium of transport of energy from continent to continent.
- **Land:** vegetation, man and soil play an important role in terms of air dynamics and chemicals transport.
- **Cryosphere:** snow, ice and sea-ice influence on the large-scale circulation.
- **Biosphere:** life on earth and in the water has an important impact on the CO2 cycle.

Impacts of atmospheric processes

Extreme weather events include droughts, floods and associated landslides, storms, cyclones and tornadoes, ocean and coastal surges, heat waves and cold snaps.

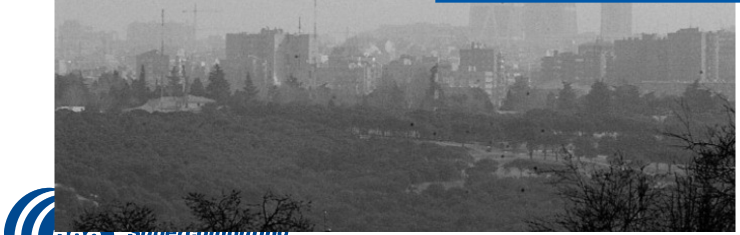
Impacts of anthropogenic activities

Los Angeles (USA)

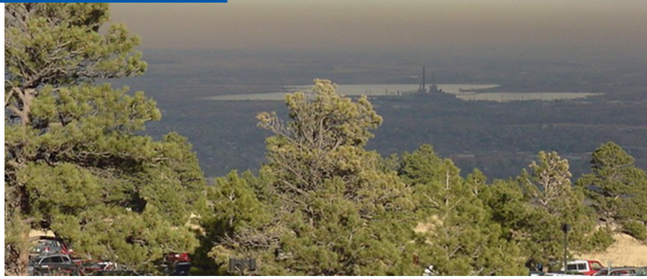
ico DF (México)

Not a new problem

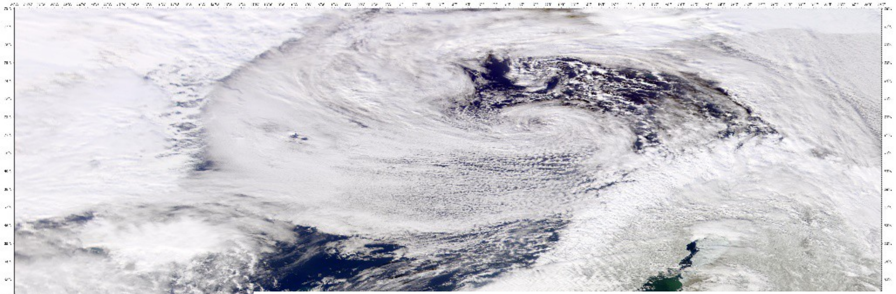
Madrid (Spain)

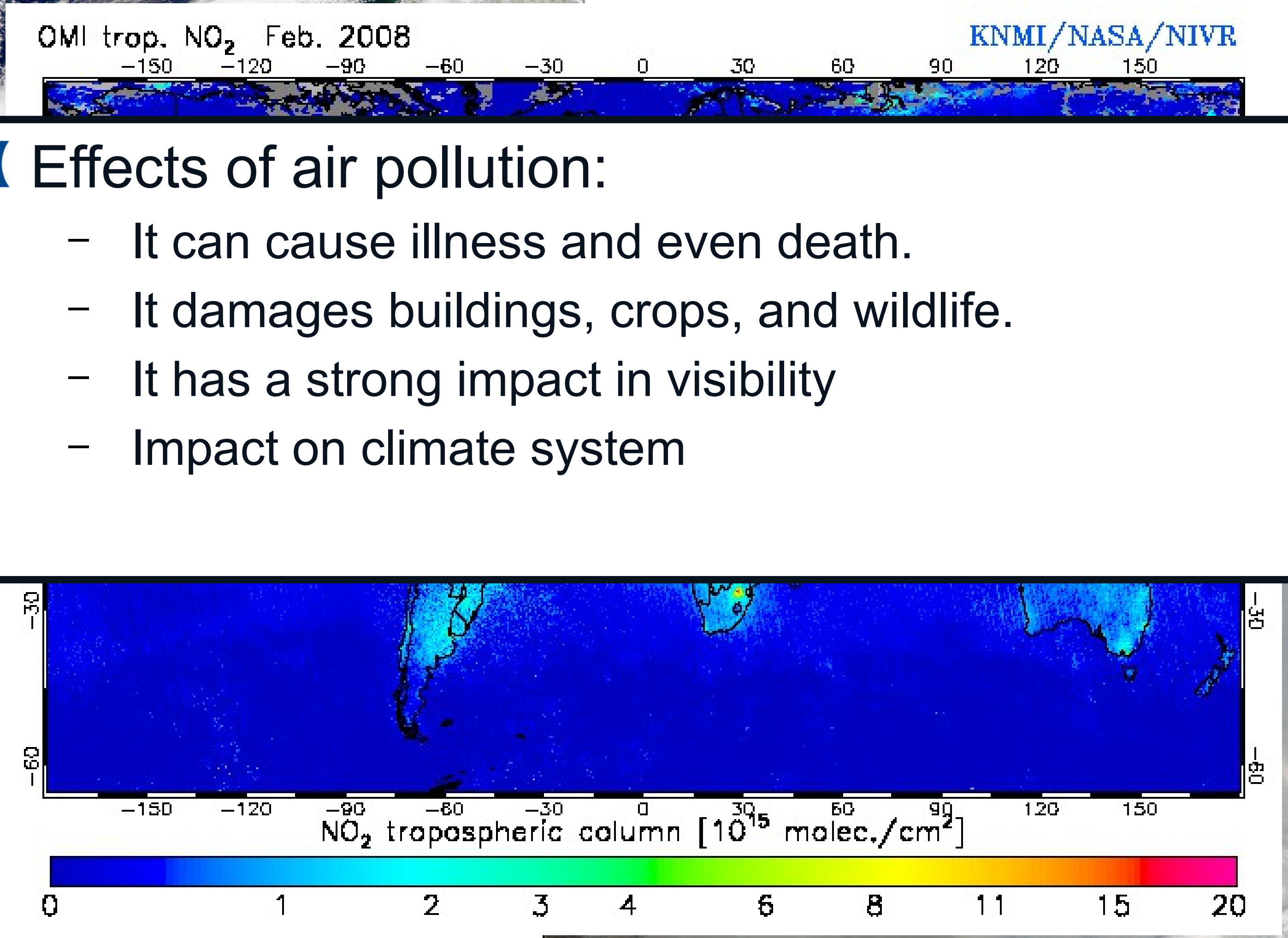


ulder (USA)

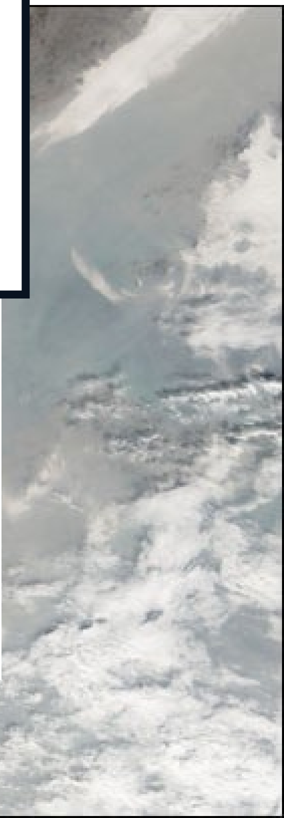


Air Pollution: Europe, South China, the Earth

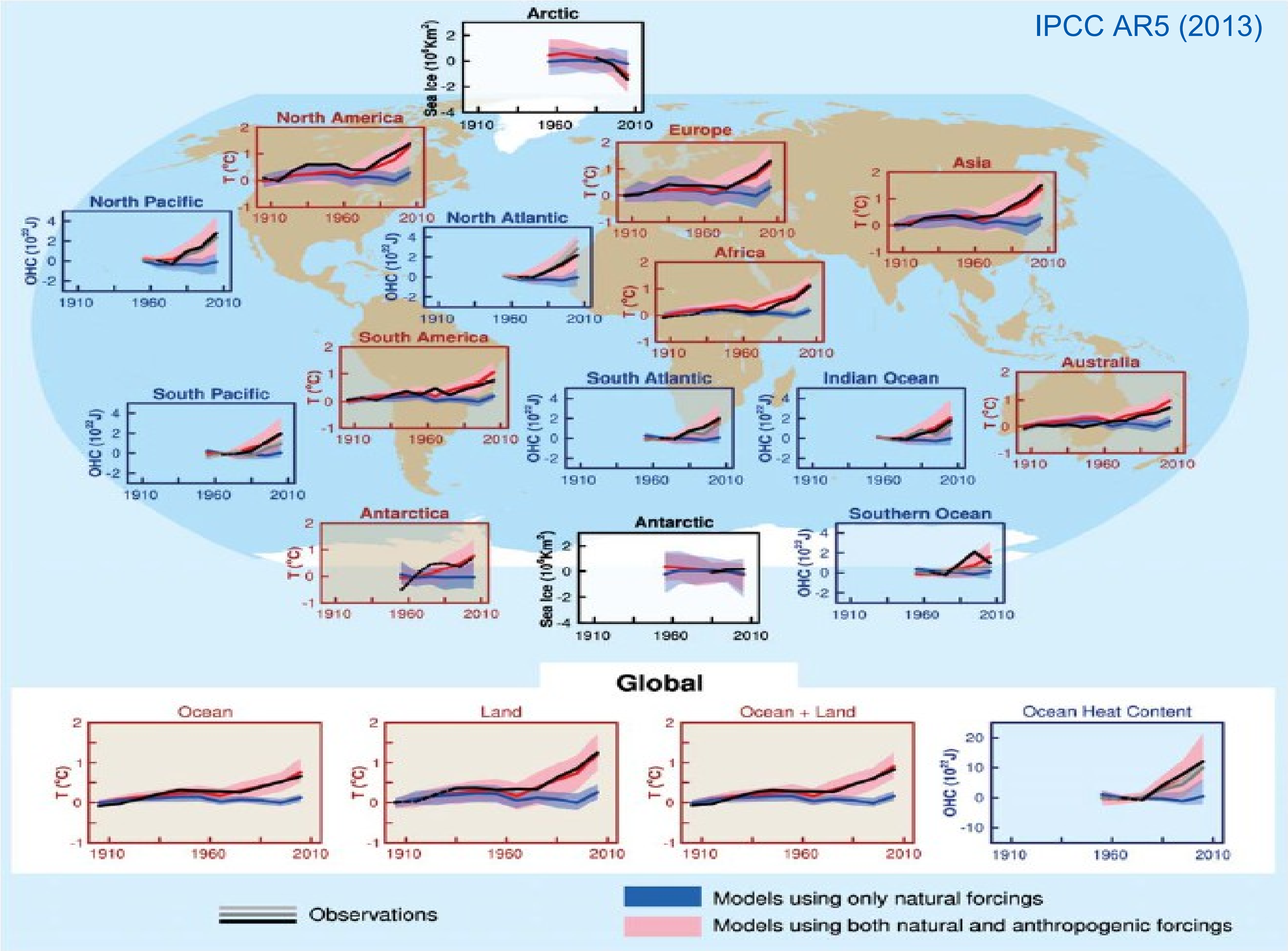



OMI trop. NO_2 Feb. 2008
KNMI/NASA/NIVR
−150
−120
−90
−60
−30
0
30
60
90
120
150
−30
−60
NO_2 tropospheric column [10^{15} molec./cm^2]
0
1
2
3
4
6
8
11
15
20

- Effects of air pollution:
  - It can cause illness and even death.
  - It damages buildings, crops, and wildlife.
  - It has a strong impact in visibility
  - Impact on climate system



Impact on Climate


IPCC AR5 (2013)

So, what can we do?

- Understanding the mechanisms that lead to extreme weather and air pollution episodes
  - Measurement campaigns
  - Chemical experiments: smog cambers
  - ***Modelling techniques***

- To provide information and guidance to the authorities
  - Information to the population, policy makers

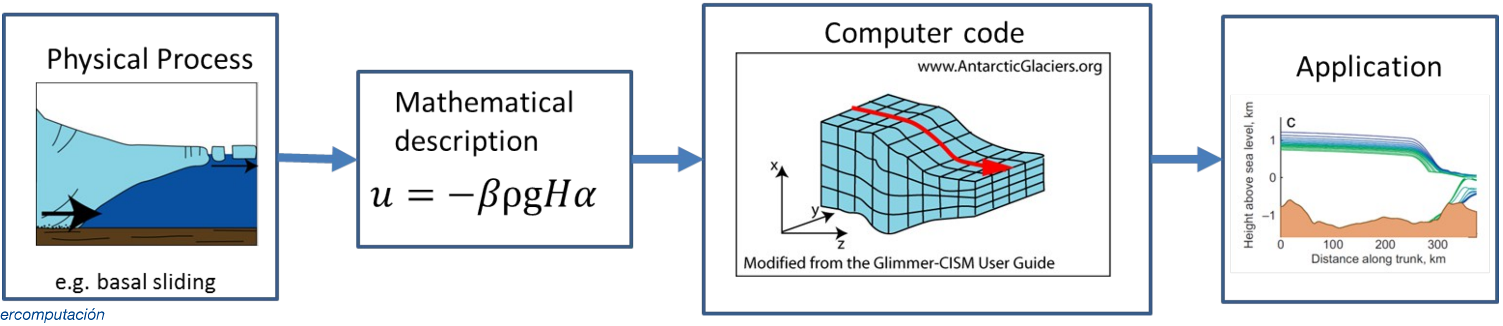
- Developing new numerical forecasting systems
  - ***Towards the chemical weather forecast and the Earth System Models***

Models: from theory to numerics

- Physical model
  - Reproduction at smaller scale of a process (e.g., wind tunnels and reproduction of buildings).

- Mathematical model
  - Set of mathematical equations with physical basis that describes a specific process.

- Numerical model
  - Computer program where the mathematical model is discretized and codified.



$$u = -\beta\rho g H\alpha$$

Types of numerical models

- Spatial scale
  - Street canyon
  - Local models
  - Mesoscale models
  - Synoptic or regional models
  - Global models
- Temporal scales
  - Short-term
  - Long-term or climate models
- Past or future scenarios
  - Hindcast
  - Nowcast
  - Forecast
  - Climate projections
- Reference framework:
  - Box model
  - Gaussian model
  - Lagrangian model
  - Eulerian model

How do we model the evolution of the atmosphere?
Meteorological models – Climate models

$$\underbrace{\frac{\partial u}{\partial t}}_{\text{Time changes in the west-to-east wind}} = \underbrace{-u\frac{\partial u}{\partial x} - v\frac{\partial u}{\partial y}}_{\text{Horizontal advection of the west-to-east wind (momentum)}} \underbrace{-\omega\frac{\partial u}{\partial p}}_{\text{Vertical advection of the west-to-east wind (momentum)}} + \overbrace{fv - g\frac{\partial z}{\partial x}}^{\text{Deviations from geostrophic balance for the south-to-north wind}} + \underbrace{F_x}_{\text{Surface friction and turbulent mixing acting on the west-to-east wind}}$$

change in vertical motion with respect to time = advection terms - nonhydrostatic vertical pressure gradient force in a grid box + buoyancy as it deviates from the large-scale average - precipitation drag

$$\frac{\partial w}{\partial t} = \text{advection terms} - \frac{1}{\rho_0}\frac{\partial p'}{\partial z} + gB - gq$$

P′ is the pressure departure from the large-scale hydrostatic balance
ρ_0 is the density of the environment
g is gravity
B is buoyancy
-gq is precipitation drag

$$\underbrace{\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}}_{\text{Horizontal Divergence}} + \underbrace{\frac{\partial \omega}{\partial p}}_{\text{Vertical Divergence}} = 0$$

$$\overbrace{\frac{\partial z}{\partial p}}^{\text{Difference in height between upper \& lower isobaric sufaces}} = -\frac{R\overbrace{T}^{}}{pg}\qquad \text{(}T\text{: Mean temperature within a layer)}$$

$$\underbrace{\frac{\partial T}{\partial t}}_{\text{Time changes in temperature}} = \underbrace{-u\frac{\partial T}{\partial x} - v\frac{\partial T}{\partial y}}_{\text{Horizontal advection of temperature}} \underbrace{-\omega\left(\frac{\partial T}{\partial p} - \frac{RT}{c_p p}\right)}_{\text{Difference between vertical temperature advection \& adiabatic processes}} + \overbrace{\frac{H}{c_p}}^{\text{Other processes (i.e., radiation, mixing, and condensation)}}$$

Source: COMET

$$\underbrace{\frac{\partial q}{\partial t}}_{\text{Time changes in moisture}} = \underbrace{-u\frac{\partial q}{\partial x} - v\frac{\partial q}{\partial y}}_{\text{Horizontal advection of moisture}} \underbrace{-\omega\frac{\partial q}{\partial p}}_{\text{Vertical advection of moisture}} + \overbrace{E}^{\text{Evaporation and sublimation}} - \underbrace{P}_{\text{Condensation (Precipitation)}}$$

Resolving the full compressible primitive equations:

Momentum conservation

Mass conservation

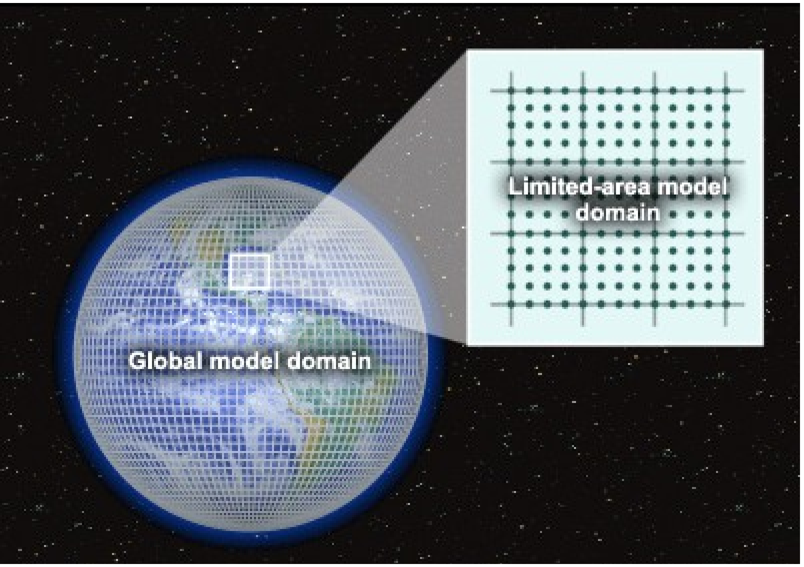
Energy conservation

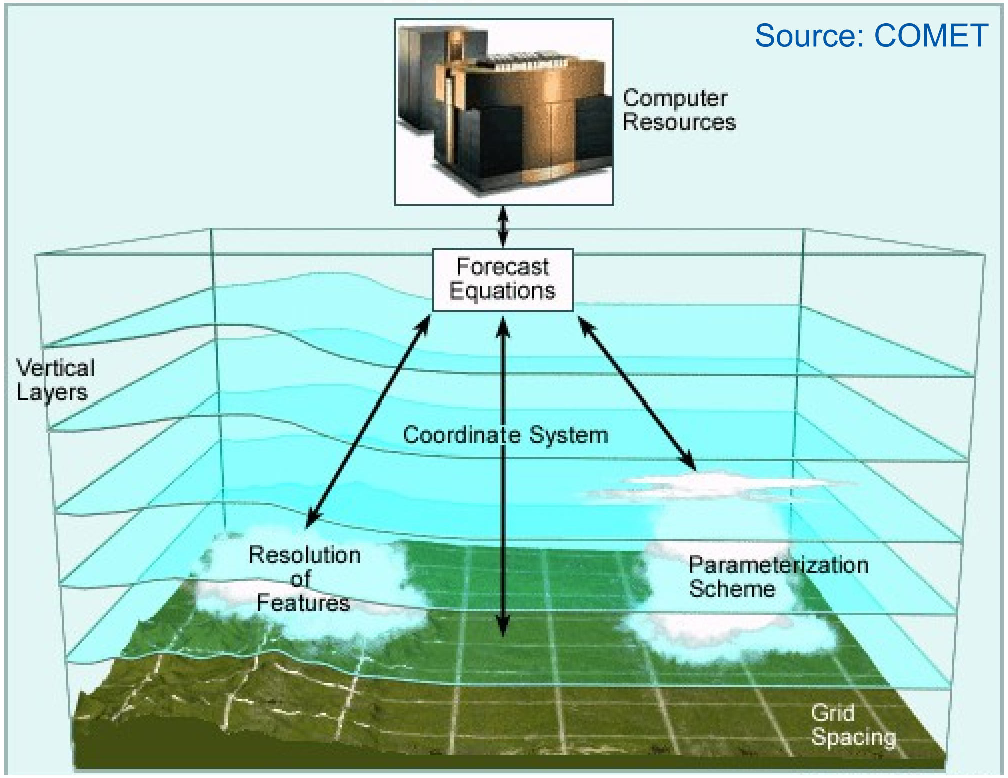
Moisture conservation

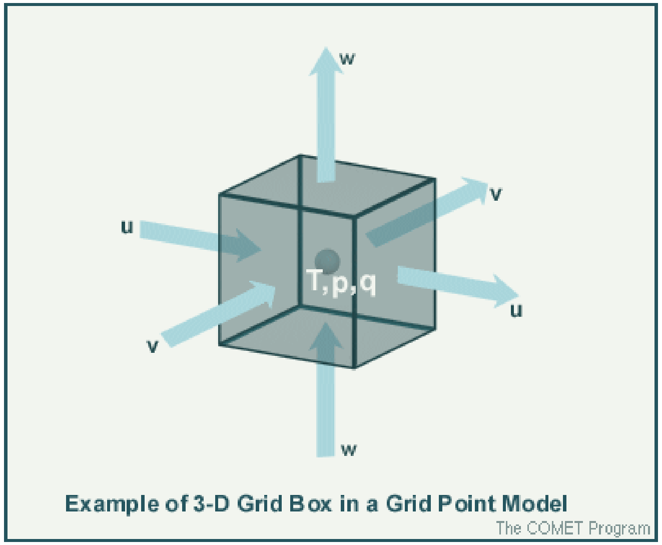
Partial Differential Equations

No analytic solution available: need of numerical approximations

Where do we solve the primitive equations? Grid discretization







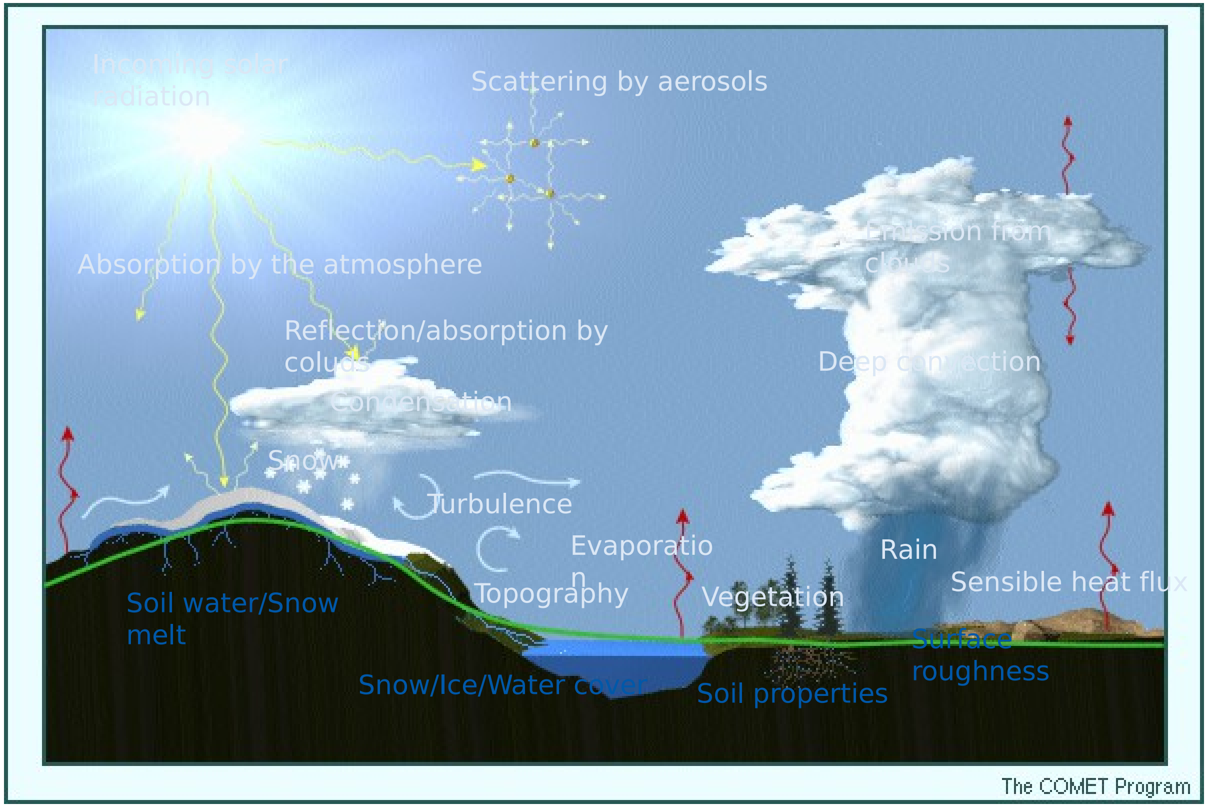
Example of 3-D Grid Box in a Grid Point Model

High performance computing resources:

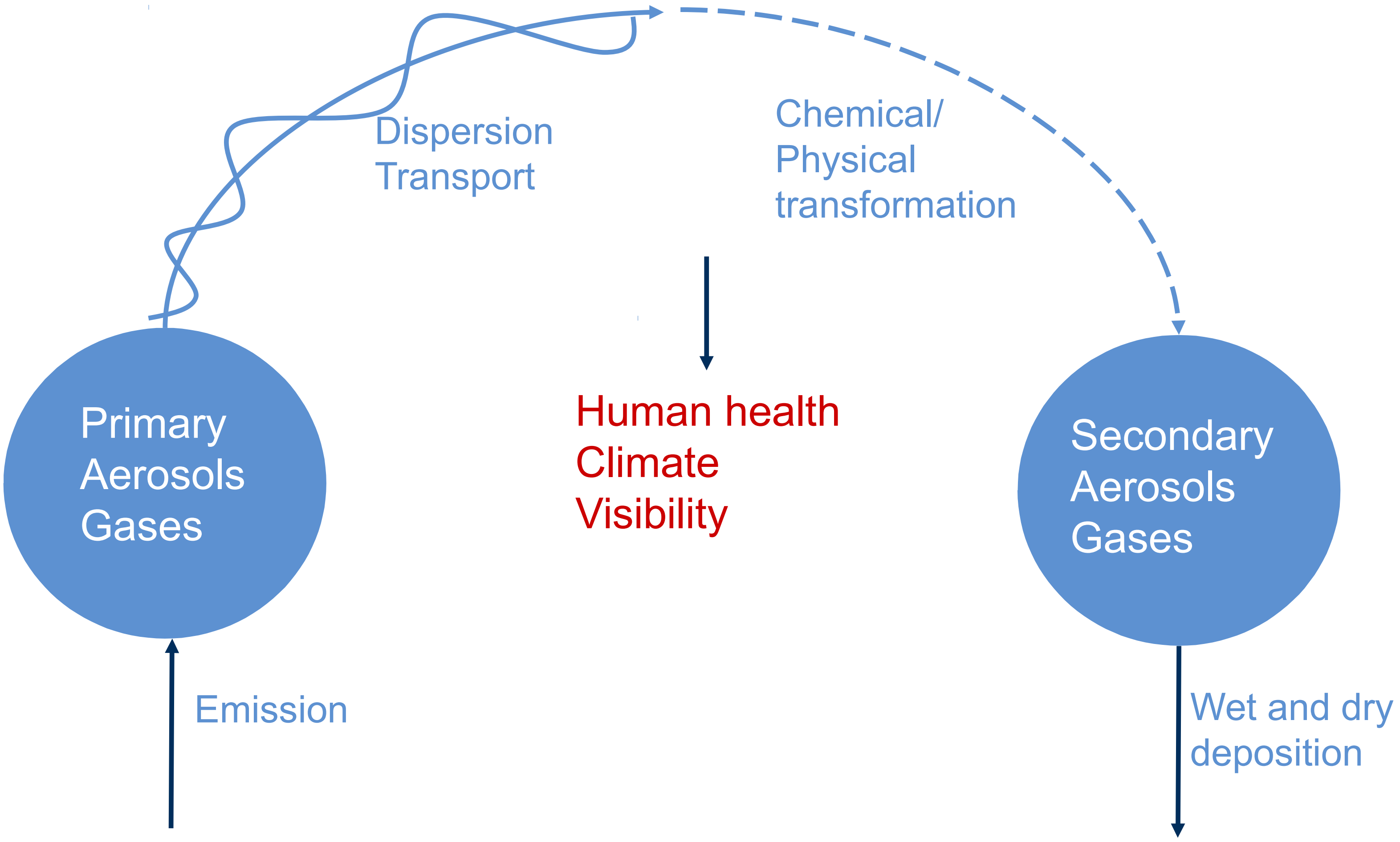
If we plan to solve small scale features we need higher resolution in the mesh and so more HPC resources are required.

Don’t forget important physical processes: physical parameterizations

- Parameterization: account for unresolved grid scale processes.



Atmospheric chemistry and Air pollution processes



Sources:
Anthropogenic
Biogenic
Geogenic

Aquatic ecosystems
Forests
Agricultural crops
Materials

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

How do we model air pollution and atmospheric chemistry?

The dispersion equation:

- Advection, diffusion, reaction-chemistry, emissions and deposition

$$\frac{\partial c}{\partial t} + U\frac{\partial c}{\partial x} + V\frac{\partial c}{\partial y} + W\frac{\partial c}{\partial z} =$$

$$\frac{\partial}{\partial x}\left(K_h\frac{\partial c}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_h\frac{\partial c}{\partial y}\right) + \frac{\partial}{\partial z}\left(K_z\frac{\partial c}{\partial z}\right)$$

$$+\ \text{chemistry} + \text{emissions} - \text{dry deposition} - \text{wet deposition}$$

Meteorology

Emissions

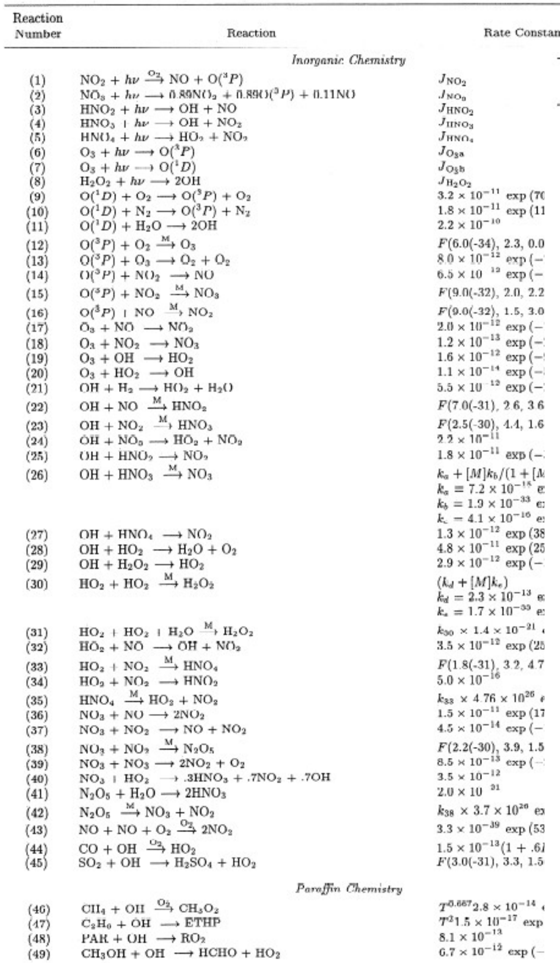
Chemistry

C – concentration of pollutant

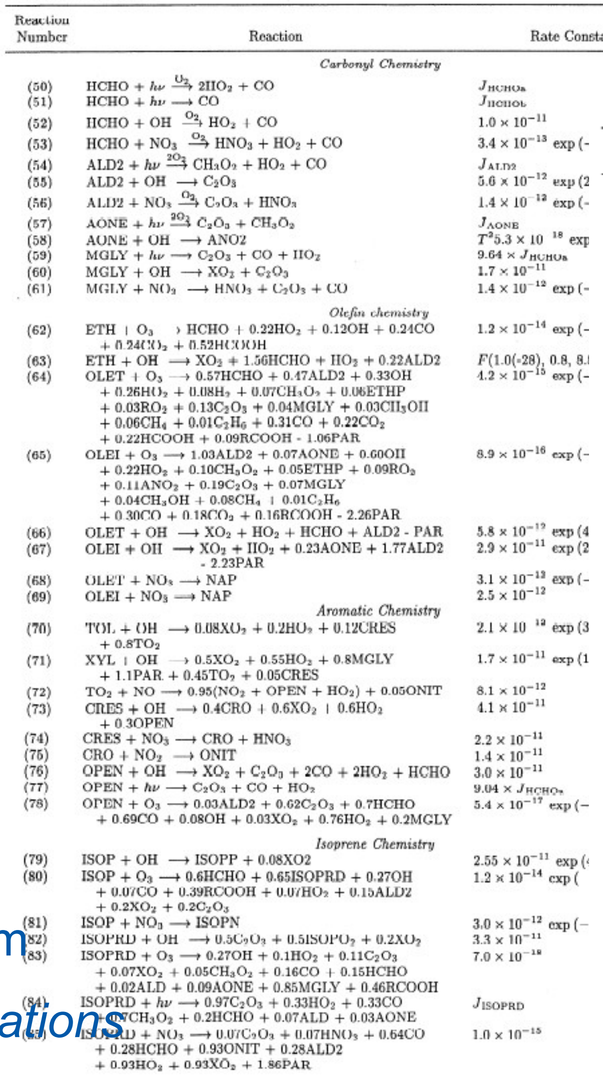
Kh – lateral diffusion coefficient

Kz – turbulence exchange coefficient

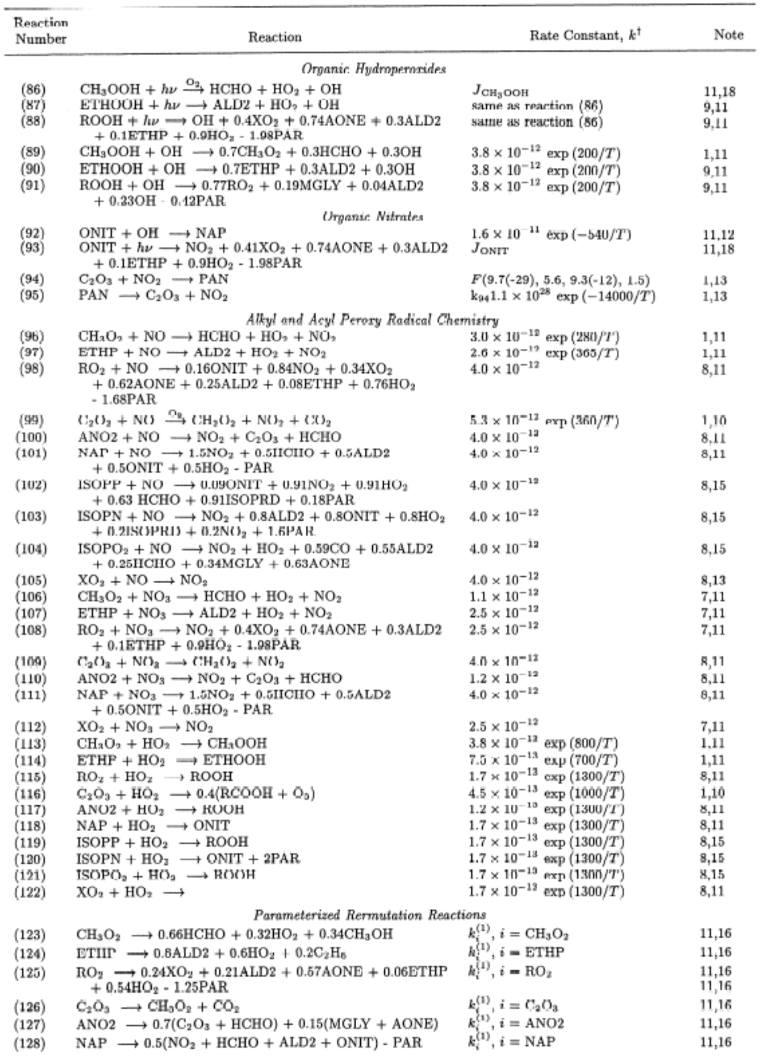
The chemical term



Reaction Number	Reaction	Rate Constant, $k^\dagger$	Note
	Inorganic Chemistry		
(1)	$NO_2 + h\nu \xrightarrow{O_2} NO + O(^3P)$	J_{NO_2}	
(2)	$NO_3 + h\nu \longrightarrow 0.89NO_2 + 0.89O(^3P) + 0.11NO$	J_{NO_3}	
(3)	$HNO_2 + h\nu \longrightarrow OH + NO$	J_{HNO_2}	
(4)	$HNO_3 + h\nu \longrightarrow OH + NO_2$	J_{HNO_3}	
(5)	$HNO_4 + h\nu \longrightarrow HO_2 + NO_2$	J_{HNO_4}	
(6)	$O_3 + h\nu \longrightarrow O(^3P)$	J_{O_3a}	
(7)	$O_3 + h\nu \longrightarrow O(^1D)$	J_{O_3b}	
(8)	$H_2O_2 + h\nu \longrightarrow 2OH$	$J_{H_2O_2}$	
(9)	$O(^1D) + O_2 \longrightarrow O(^3P) + O_2$	3.2×10^{-11} exp (7	
(10)	$O(^1D) + N_2 \longrightarrow O(^3P) + N_2$	1.8×10^{-11} exp (11	
(11)	$O(^1D) + H_2O \longrightarrow 2OH$	2.2×10^{-10}	
(12)	$O(^3P) + O_2 \xrightarrow{M} O_3$	F(6.0(-34), 2.3, 0.0	
(13)	$O(^3P) + O_3 \longrightarrow O_2 + O_2$	8.0×10^{-12} exp (−	
(14)	$O(^3P) + NO_2 \longrightarrow NO$	6.5×10^{-12} exp (−	
(15)	$O(^3P) + NO_2 \xrightarrow{M} NO_3$	F(9.0(-32), 2.0, 2.2	
(16)	$O(^3P) + NO \xrightarrow{M} NO_2$	F(9.0(-32), 1.5, 3.0	
(17)	$O_3 + NO \longrightarrow NO_2$	2.0×10^{-12} exp (−	
(18)	$O_3 + NO_2 \longrightarrow NO_3$	1.2×10^{-13} exp (−	
(19)	$O_3 + OH \longrightarrow HO_2$	1.6×10^{-12} exp (−	
(20)	$O_3 + HO_2 \longrightarrow OH$	1.1×10^{-14} exp (−	
(21)	$OH + H_2 \longrightarrow HO_2 + H_2O$	5.5×10^{-12} exp (−	
(22)	$OH + NO \xrightarrow{M} HNO_2$	F(7.0(-31), 2.6, 3.6	
(23)	$OH + NO_2 \xrightarrow{M} HNO_3$	F(2.5(-30), 4.4, 1.6	
(24)	$OH + NO_3 \longrightarrow HO_2 + NO_2$	2.2×10^{-11}	
(25)	$OH + HNO_2 \longrightarrow NO_2$	1.8×10^{-11} exp (−	
(26)	$OH + HNO_3 \xrightarrow{M} NO_3$	$k_a + [M]k_b/(1 + [M$ $k_a = 7.2 \times 10^{-15}$ e $k_b = 1.9 \times 10^{-33}$ e $k_c = 4.1 \times 10^{-16}$ e	
(27)	$OH + HNO_4 \longrightarrow NO_2$	1.3×10^{-12} exp (38	
(28)	$OH + HO_2 \longrightarrow H_2O + O_2$	4.8×10^{-11} exp (25	
(29)	$OH + H_2O_2 \longrightarrow HO_2$	2.9×10^{-12} exp (−	
(30)	$HO_2 + HO_2 \xrightarrow{M} H_2O_2$	$(k_d + [M]k_e)$ $k_d = 2.3 \times 10^{-13}$ e $k_e = 1.7 \times 10^{-33}$ e	
(31)	$HO_2 + HO_2 + H_2O \xrightarrow{M} H_2O_2$	$k_{30} \times 1.4 \times 10^{-21}$	
(32)	$HO_2 + NO \longrightarrow OH + NO_2$	3.5×10^{-12} exp (25	
(33)	$HO_2 + NO_2 \xrightarrow{M} HNO_4$	F(1.8(-31), 3.2, 4.7	
(34)	$HO_2 + NO_2 \longrightarrow HNO_2$	5.0×10^{-16}	
(35)	$HNO_4 \xrightarrow{M} HO_2 + NO_2$	$k_{33} \times 4.76 \times 10^{26}$	
(36)	$NO_3 + NO \longrightarrow 2NO_2$	1.5×10^{-11} exp (17	
(37)	$NO_3 + NO_2 \longrightarrow NO + NO_2$	4.5×10^{-14} exp (−	
(38)	$NO_3 + NO_2 \xrightarrow{M} N_2O_5$	F(2.2(-30), 3.9, 1.5	
(39)	$NO_3 + NO_3 \longrightarrow 2NO_2 + O_2$	8.5×10^{-13} exp (−	
(40)	$NO_3 + HO_2 \longrightarrow .3HNO_3 + .7NO_2 + .7OH$	3.5×10^{-12}	
(41)	$N_2O_5 + H_2O \longrightarrow 2HNO_3$	2.0×10^{-21}	
(42)	$N_2O_5 \xrightarrow{M} NO_3 + NO_2$	$k_{38} \times 3.7 \times 10^{26}$ e	
(43)	$NO + NO + O_2 \xrightarrow{O_2} 2NO_2$	3.3×10^{-39} exp (53	
(44)	$CO + OH \xrightarrow{O_2} HO_2$	$1.5 \times 10^{-13}(1 + .6$	
(45)	$SO_2 + OH \longrightarrow H_2SO_4 + HO_2$	F(3.0(-31), 3.3, 1.5	
	Paraffin Chemistry		
(46)	$CH_4 + OH \xrightarrow{O_2} CH_3O_2$	$T^{0.667}2.8 \times 10^{-14}$	
(47)	$C_2H_6 + OH \longrightarrow ETHP$	$T^2 1.5 \times 10^{-17}$ exp	
(48)	$PAR + OH \longrightarrow RO_2$	8.1×10^{-13}	
(49)	$CH_3OH + OH \longrightarrow HCHO + HO_2$	6.7×10^{-12} exp (−	



Reaction Number	Reaction	Rate Constant, $k^\dagger$	Note
	Carbonyl Chemistry		
(50)	$HCHO + h\nu \xrightarrow{O_2} 2HO_2 + CO$	J_{HCHOa}	13,18
(51)	$HCHO + h\nu \longrightarrow CO$	J_{HCHOb}	13,18
(52)	$HCHO + OH \xrightarrow{O_2} HO_2 + CO$	1.0×10^{-11}	
(53)	$HCHO + NO_3 \xrightarrow{O_2} HNO_3 + HO_2 + CO$	3.4×10^{-13} exp (-	
(54)	$ALD2 + h\nu \xrightarrow{2O_2} CH_3O_2 + HO_2 + CO$	J_{ALD2}	
(55)	$ALD2 + OH \longrightarrow C_2O_3$	5.6×10^{-12} exp (2	
(56)	$ALD2 + NO_3 \xrightarrow{O_2} C_2O_3 + HNO_3$	1.4×10^{-12} exp (-	
(57)	$AONE + h\nu \xrightarrow{2O_2} C_2O_3 + CH_3O_2$	J_{AONE}	
(58)	$AONE + OH \longrightarrow ANO2$	$T^2 5.3 \times 10^{-18}$ exp	
(59)	$MGLY + h\nu \longrightarrow C_2O_3 + CO + HO_2$	$9.64 \times J_{HCHOa}$	
(60)	$MGLY + OH \longrightarrow XO_2 + C_2O_3$	1.7×10^{-11}	
(61)	$MGLY + NO_3 \longrightarrow HNO_3 + C_2O_3 + CO$	1.4×10^{-12} exp (-	
	Olefin chemistry		
(62)	$ETH + O_3 \longrightarrow HCHO + 0.22HO_2 + 0.12OH + 0.24CO$ $+ 0.24CO_2 + 0.52HCOOH$	1.2×10^{-14} exp (-	
(63)	$ETH + OH \longrightarrow XO_2 + 1.56HCHO + HO_2 + 0.22ALD2$	F(1.0(-28), 0.8, 8.	
(64)	$OLET + O_3 \longrightarrow 0.57HCHO + 0.47ALD2 + 0.33OH$ $+ 0.26HO_2 + 0.08H_2 + 0.07CH_3O_2 + 0.06ETHP$ $+ 0.03RO_2 + 0.13C_2O_3 + 0.04MGLY + 0.03CH_3OH$ $+ 0.06CH_4 + 0.01C_2H_6 + 0.31CO + 0.22CO_2$ $+ 0.22HCOOH + 0.09RCOOH - 1.06PAR$	4.2×10^{-15} exp (-	
(65)	$OLEI + O_3 \longrightarrow 1.03ALD2 + 0.07AONE + 0.60OH$ $+ 0.22HO_2 + 0.10CH_3O_2 + 0.05ETHP + 0.09RO_2$ $+ 0.11ANO_2 + 0.19C_2O_3 + 0.07MGLY$ $+ 0.04CH_3OH + 0.08CH_4 + 0.01C_2H_6$ $+ 0.30CO + 0.18CO_2 + 0.16RCOOH - 2.26PAR$	8.9×10^{-16} exp (-	
(66)	$OLET + OH \longrightarrow XO_2 + HO_2 + HCHO + ALD2 - PAR$	5.8×10^{-12} exp (4	
(67)	$OLEI + OH \longrightarrow XO_2 + HO_2 + 0.23AONE + 1.77ALD2$ $- 2.23PAR$	2.9×10^{-11} exp (2	
(68)	$OLET + NO_3 \longrightarrow NAP$	3.1×10^{-13} exp (-	
(69)	$OLEI + NO_3 \longrightarrow NAP$	2.5×10^{-12}	
	Aromatic Chemistry		
(70)	$TOL + OH \longrightarrow 0.08XO_2 + 0.2HO_2 + 0.12CRES$ $+ 0.8TO_2$	2.1×10^{-12} exp (3	
(71)	$XYL + OH \longrightarrow 0.5XO_2 + 0.55HO_2 + 0.8MGLY$ $+ 1.1PAR + 0.45TO_2 + 0.05CRES$	1.7×10^{-11} exp (1	
(72)	$TO_2 + NO \longrightarrow 0.95(NO_2 + OPEN + HO_2) + 0.05ONIT$	8.1×10^{-12}	
(73)	$CRES + OH \longrightarrow 0.4CRO + 0.6XO_2 + 0.6HO_2$ $+ 0.3OPEN$	4.1×10^{-11}	
(74)	$CRES + NO_3 \longrightarrow CRO + HNO_3$	2.2×10^{-11}	
(75)	$CRO + NO_2 \longrightarrow ONIT$	1.4×10^{-11}	
(76)	$OPEN + OH \longrightarrow XO_2 + C_2O_3 + 2CO + 2HO_2 + HCHO$	3.0×10^{-11}	
(77)	$OPEN + h\nu \longrightarrow C_2O_3 + CO + HO_2$	$9.04 \times J_{HCHOa}$	
(78)	$OPEN + O_3 \longrightarrow 0.03ALD2 + 0.62C_2O_3 + 0.7HCHO$ $+ 0.69CO + 0.08OH + 0.03XO_2 + 0.76HO_2 + 0.2MGLY$	5.4×10^{-17} exp (-	
	Isoprene Chemistry		
(79)	$ISOP + OH \longrightarrow ISOPP + 0.08XO2$	2.55×10^{-11} exp (	
(80)	$ISOP + O_3 \longrightarrow 0.6HCHO + 0.65ISOPRD + 0.27OH$ $+ 0.07CO + 0.39RCOOH + 0.07HO_2 + 0.15ALD2$ $+ 0.2XO_2 + 0.2C_2O_3$	1.2×10^{-14} exp (	
(81)	$ISOP + NO_3 \longrightarrow ISOPN$	3.0×10^{-12} exp (-	
(82)	$ISOPRD + OH \longrightarrow 0.5C_2O_3 + 0.5ISOPO_2 + 0.2XO_2$	3.3×10^{-11}	
(83)	$ISOPRD + O_3 \longrightarrow 0.27OH + 0.1HO_2 + 0.11C_2O_3$ $+ 0.07XO_2 + 0.05CH_3O_2 + 0.16CO + 0.15HCHO$ $+ 0.02ALD + 0.09AONE + 0.85MGLY + 0.46RCOOH$	7.0×10^{-18}	
(84)	$ISOPRD + h\nu \longrightarrow 0.97C_2O_3 + 0.33HO_2 + 0.33CO$ $+ 0.07CH_3O_2 + 0.2HCHO + 0.07ALD + 0.03AONE$	J_{ISOPRD}	
(85)	$ISOPRD + NO_3 \longrightarrow 0.07C_2O_3 + 0.07HNO_3 + 0.64CO$ $+ 0.28HCHO + 0.93ONIT + 0.28ALD2$ $+ 0.93HO_2 + 0.93XO_2 + 1.86PAR$	1.0×10^{-15}	



Reaction Number	Reaction	Rate Constant, $k^\dagger$	Note
	Organic Hydroperoxides		
(86)	$CH_3OOH + h\nu \xrightarrow{O_2} HCHO + HO_2 + OH$	J_{CH_3OOH}	11,18
(87)	$ETHOOH + h\nu \longrightarrow ALD2 + HO_2 + OH$	same as reaction (86)	9,11
(88)	$ROOH + h\nu \longrightarrow OH + 0.4XO_2 + 0.74AONE + 0.3ALD2$ $+ 0.1ETHP + 0.9HO_2 - 1.98PAR$	same as reaction (86)	9,11
(89)	$CH_3OOH + OH \longrightarrow 0.7CH_3O_2 + 0.3HCHO + 0.3OH$	3.8×10^{-12} exp $(200/T)$	1,11
(90)	$ETHOOH + OH \longrightarrow 0.7ETHP + 0.3ALD2 + 0.3OH$	3.8×10^{-12} exp $(200/T)$	9,11
(91)	$ROOH + OH \longrightarrow 0.77RO_2 + 0.19MGLY + 0.04ALD2$ $+ 0.23OH - 0.42PAR$	3.8×10^{-12} exp $(200/T)$	9,11
	Organic Nitrates		
(92)	$ONIT + OH \longrightarrow NAP$	1.6×10^{-11} exp $(-540/T)$	11,12
(93)	$ONIT + h\nu \longrightarrow NO_2 + 0.41XO_2 + 0.74AONE + 0.3ALD2$ $+ 0.1ETHP + 0.9HO_2 - 1.98PAR$	J_{ONIT}	11,18
(94)	$C_2O_3 + NO_2 \longrightarrow PAN$	F(9.7(-29), 5.6, 9.3(-12), 1.5)	1,13
(95)	$PAN \longrightarrow C_2O_3 + NO_2$	$k_{94} 1.1 \times 10^{28}$ exp $(-14000/T)$	1,13
	Alkyl and Acyl Peroxy Radical Chemistry		
(96)	$CH_3O_2 + NO \longrightarrow HCHO + HO_2 + NO_2$	3.0×10^{-12} exp $(280/T)$	1,11
(97)	$ETHP + NO \longrightarrow ALD2 + HO_2 + NO_2$	2.6×10^{-12} exp $(365/T)$	1,11
(98)	$RO_2 + NO \longrightarrow 0.16ONIT + 0.84NO_2 + 0.34XO_2$ $+ 0.62AONE + 0.25ALD2 + 0.08ETHP + 0.76HO_2$ $- 1.68PAR$	4.0×10^{-12}	8,11
(99)	$C_2O_3 + NO \xrightarrow{O_2} CH_3O_2 + NO_2 + CO_2$	5.3×10^{-12} exp $(360/T)$	1,10
(100)	$ANO2 + NO \longrightarrow NO_2 + C_2O_3 + HCHO$	4.0×10^{-12}	8,11
(101)	$NAP + NO \longrightarrow 1.5NO_2 + 0.5HCHO + 0.5ALD2$ $+ 0.5ONIT + 0.5HO_2 - PAR$	4.0×10^{-12}	8,11
(102)	$ISOPP + NO \longrightarrow 0.09ONIT + 0.91NO_2 + 0.91HO_2$ $+ 0.63\ HCHO + 0.91ISOPRD + 0.18PAR$	4.0×10^{-12}	8,15
(103)	$ISOPN + NO \longrightarrow NO_2 + 0.8ALD2 + 0.8ONIT + 0.8HO_2$ $+ 0.2ISOPRD + 0.2NO_2 + 1.6PAR$	4.0×10^{-12}	8,15
(104)	$ISOPO_2 + NO \longrightarrow NO_2 + HO_2 + 0.59CO + 0.55ALD2$ $+ 0.25HCHO + 0.34MGLY + 0.63AONE$	4.0×10^{-12}	8,15
(105)	$XO_2 + NO \longrightarrow NO_2$	4.0×10^{-12}	8,13
(106)	$CH_3O_2 + NO_3 \longrightarrow HCHO + HO_2 + NO_2$	1.1×10^{-12}	7,11
(107)	$ETHP + NO_3 \longrightarrow ALD2 + HO_2 + NO_2$	2.5×10^{-12}	7,11
(108)	$RO_2 + NO_3 \longrightarrow NO_2 + 0.4XO_2 + 0.74AONE + 0.3ALD2$ $+ 0.1ETHP + 0.9HO_2 - 1.98PAR$	2.5×10^{-12}	7,11
(109)	$C_2O_3 + NO_3 \longrightarrow CH_3O_2 + NO_2$	4.0×10^{-12}	8,11
(110)	$ANO2 + NO_3 \longrightarrow NO_2 + C_2O_3 + HCHO$	1.2×10^{-12}	8,11
(111)	$NAP + NO_3 \longrightarrow 1.5NO_2 + 0.5HCHO + 0.5ALD2$ $+ 0.5ONIT + 0.5HO_2 - PAR$	4.0×10^{-12}	8,11
(112)	$XO_2 + NO_3 \longrightarrow NO_2$	2.5×10^{-12}	7,11
(113)	$CH_3O_2 + HO_2 \longrightarrow CH_3OOH$	3.8×10^{-13} exp $(800/T)$	1,11
(114)	$ETHP + HO_2 \longrightarrow ETHOOH$	7.5×10^{-13} exp $(700/T)$	1,11
(115)	$RO_2 + HO_2 \longrightarrow ROOH$	1.7×10^{-13} exp $(1300/T)$	8,11
(116)	$C_2O_3 + HO_2 \longrightarrow 0.4(RCOOH + O_3)$	4.5×10^{-13} exp $(1000/T)$	1,10
(117)	$ANO2 + HO_2 \longrightarrow ROOH$	1.2×10^{-13} exp $(1300/T)$	8,11
(118)	$NAP + HO_2 \longrightarrow ONIT$	1.7×10^{-13} exp $(1300/T)$	8,11
(119)	$ISOPP + HO_2 \longrightarrow ROOH$	1.7×10^{-13} exp $(1300/T)$	8,15
(120)	$ISOPN + HO_2 \longrightarrow ONIT + 2PAR$	1.7×10^{-13} exp $(1300/T)$	8,15
(121)	$ISOPO_2 + HO_2 \longrightarrow ROOH$	1.7×10^{-13} exp $(1300/T)$	8,15
(122)	$XO_2 + HO_2 \longrightarrow$	1.7×10^{-13} exp $(1300/T)$	8,11
	Parameterized Rermutation Reactions		
(123)	$CH_3O_2 \longrightarrow 0.66HCHO + 0.32HO_2 + 0.34CH_3OH$	$k_i^{(1)}$, $i = CH_3O_2$	11,16
(124)	$ETHP \longrightarrow 0.8ALD2 + 0.6HO_2 + 0.2C_2H_6$	$k_i^{(1)}$, $i = ETHP$	11,16
(125)	$RO_2 \longrightarrow 0.24XO_2 + 0.21ALD2 + 0.57AONE + 0.06ETHP$ $+ 0.54HO_2 - 1.25PAR$	$k_i^{(1)}$, $i = RO_2$	11,16 11,16
(126)	$C_2O_3 \longrightarrow CH_3O_2 + CO_2$	$k_i^{(1)}$, $i = C_2O_3$	11,16
(127)	$ANO2 \longrightarrow 0.7(C_2O_3 + HCHO) + 0.15(MGLY + AONE)$	$k_i^{(1)}$, $i = ANO2$	11,16
(128)	$NAP \longrightarrow 0.5(NO_2 + HCHO + ALD2 + ONIT) - PAR$	$k_i^{(1)}$, $i = NAP$	11,16

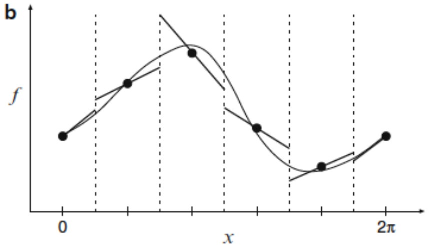
The chemical mechanism

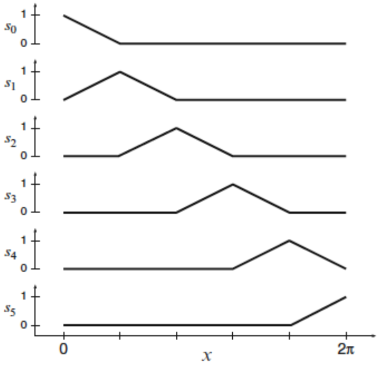
Ordinary differential equations

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Numerical approaches

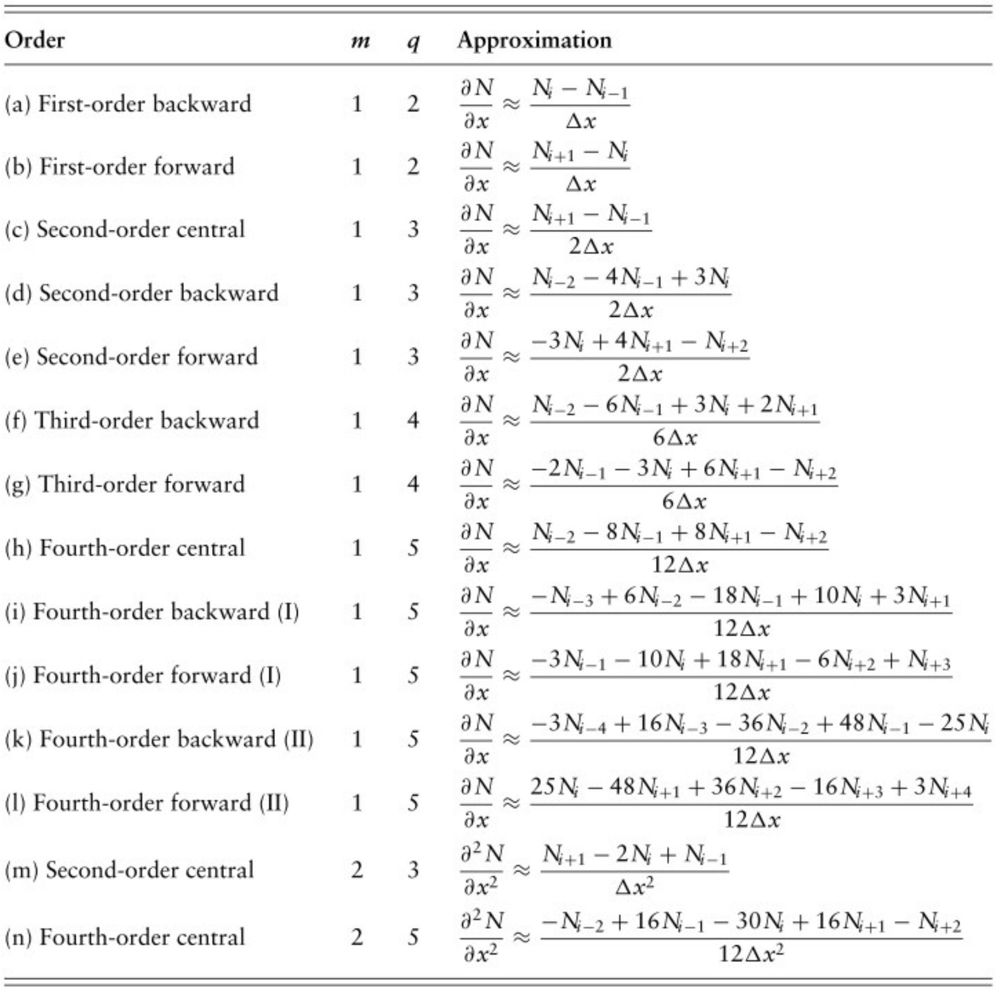
- Computers are not designed to solve differential equations directly
- Approximating Calculus with Algebra
  - Finite differences $\frac{df}{dx}(x_0) \approx \frac{f(x_0+\Delta x) - f(x_0-\Delta x)}{2\Delta x}$
  - Finite volumes $f(x) \approx f_j + \sigma_j(x - j\Delta x)$
  - Spectral transforms $a_1 + a_2 \cos x + a_3 \sin x + a_4 \cos 2x + a_5 \sin 2x$
  - Finite elements $b_0 s_0(x) + b_1 s_1(x) + \cdots + b_5 s_5(x)$



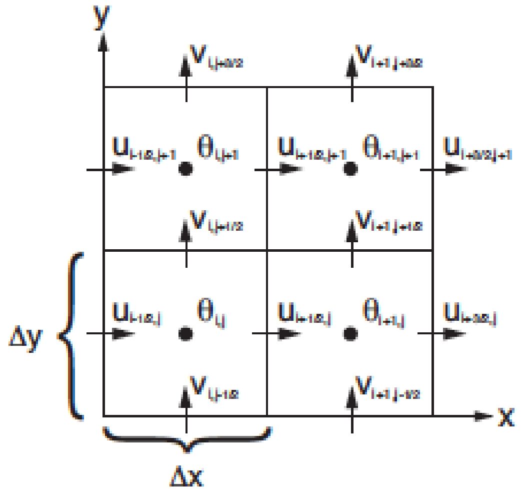


Barcelona Supercomputing Center
Centro Nacional de Supercomputación

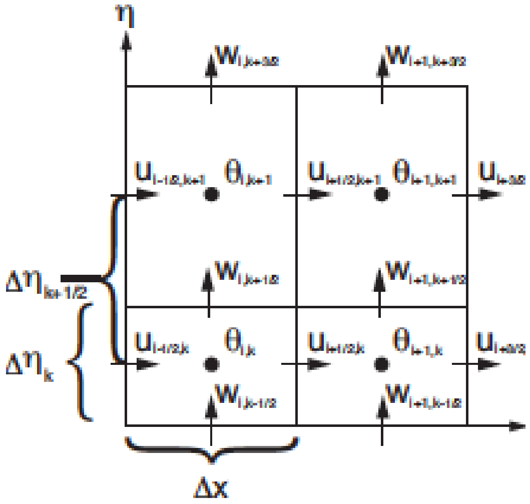
Finite differences: Discretization in space



Order	m	q	Approximation
(a) First-order backward	1	2	$\frac{\partial N}{\partial x} \approx \frac{N_i - N_{i-1}}{\Delta x}$
(b) First-order forward	1	2	$\frac{\partial N}{\partial x} \approx \frac{N_{i+1} - N_i}{\Delta x}$
(c) Second-order central	1	3	$\frac{\partial N}{\partial x} \approx \frac{N_{i+1} - N_{i-1}}{2\Delta x}$
(d) Second-order backward	1	3	$\frac{\partial N}{\partial x} \approx \frac{N_{i-2} - 4N_{i-1} + 3N_i}{2\Delta x}$
(e) Second-order forward	1	3	$\frac{\partial N}{\partial x} \approx \frac{-3N_i + 4N_{i+1} - N_{i+2}}{2\Delta x}$
(f) Third-order backward	1	4	$\frac{\partial N}{\partial x} \approx \frac{N_{i-2} - 6N_{i-1} + 3N_i + 2N_{i+1}}{6\Delta x}$
(g) Third-order forward	1	4	$\frac{\partial N}{\partial x} \approx \frac{-2N_{i-1} - 3N_i + 6N_{i+1} - N_{i+2}}{6\Delta x}$
(h) Fourth-order central	1	5	$\frac{\partial N}{\partial x} \approx \frac{N_{i-2} - 8N_{i-1} + 8N_{i+1} - N_{i+2}}{12\Delta x}$
(i) Fourth-order backward (I)	1	5	$\frac{\partial N}{\partial x} \approx \frac{-N_{i-3} + 6N_{i-2} - 18N_{i-1} + 10N_i + 3N_{i+1}}{12\Delta x}$
(j) Fourth-order forward (I)	1	5	$\frac{\partial N}{\partial x} \approx \frac{-3N_{i-1} - 10N_i + 18N_{i+1} - 6N_{i+2} + N_{i+3}}{12\Delta x}$
(k) Fourth-order backward (II)	1	5	$\frac{\partial N}{\partial x} \approx \frac{-3N_{i-4} + 16N_{i-3} - 36N_{i-2} + 48N_{i-1} - 25N_i}{12\Delta x}$
(l) Fourth-order forward (II)	1	5	$\frac{\partial N}{\partial x} \approx \frac{25N_i - 48N_{i+1} + 36N_{i+2} - 16N_{i+3} + 3N_{i+4}}{12\Delta x}$
(m) Second-order central	2	3	$\frac{\partial^2 N}{\partial x^2} \approx \frac{N_{i+1} - 2N_i + N_{i-1}}{\Delta x^2}$
(n) Fourth-order central	2	5	$\frac{\partial^2 N}{\partial x^2} \approx \frac{-N_{i-2} + 16N_{i-1} - 30N_i + 16N_{i+1} - N_{i+2}}{12\Delta x^2}$



horizontal grid



vertical grid

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Finite differences: Discretization in time

$$\frac{\partial u}{\partial t} = f(u,t)$$

Forward Euler (explicit)

$$\frac{U_{n+1}-U_n}{\Delta t} = f(U_n, t_n) \quad \text{is} \quad U_{n+1} = U_n + \Delta t\, f_n$$

Backward Euler (implicit)

$$\frac{U_{n+1}-U_n}{\Delta t} = f(U_{n+1}, t_{n+1}) \quad \text{is} \quad U_{n+1} - \Delta t\, f_{n+1} = U_n$$

Trapezoidal rule/Crank-Nicolson $\frac{U_{n+1}-U_n}{\Delta t} = \frac{1}{2}(f_{n+1}+f_n)$

Many discretization options with different stability, accuracy and cost

Backward differences $\frac{3U_{n+1}-4U_n+U_{n-1}}{2\Delta t} = f(U_{n+1}, t_{n+1})$

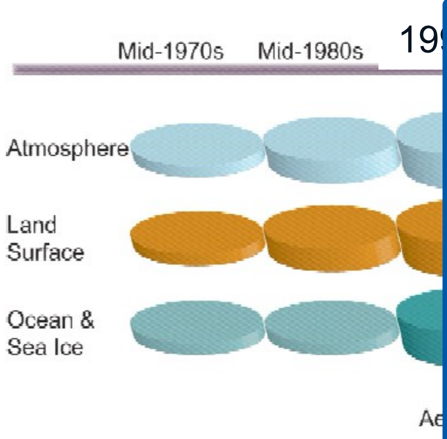
"Adams-Bashforth" $\frac{U_{n+1}-U_n}{\Delta t} = \frac{3}{2} f(U_n, t_n) - \frac{1}{2} f(U_{n-1}, t_{n-1})$

Runge-Kutta $\frac{U_{n+1}-U_n}{\Delta t} = \frac{1}{3}(k_1+2k_2+2k_3+k_4)$

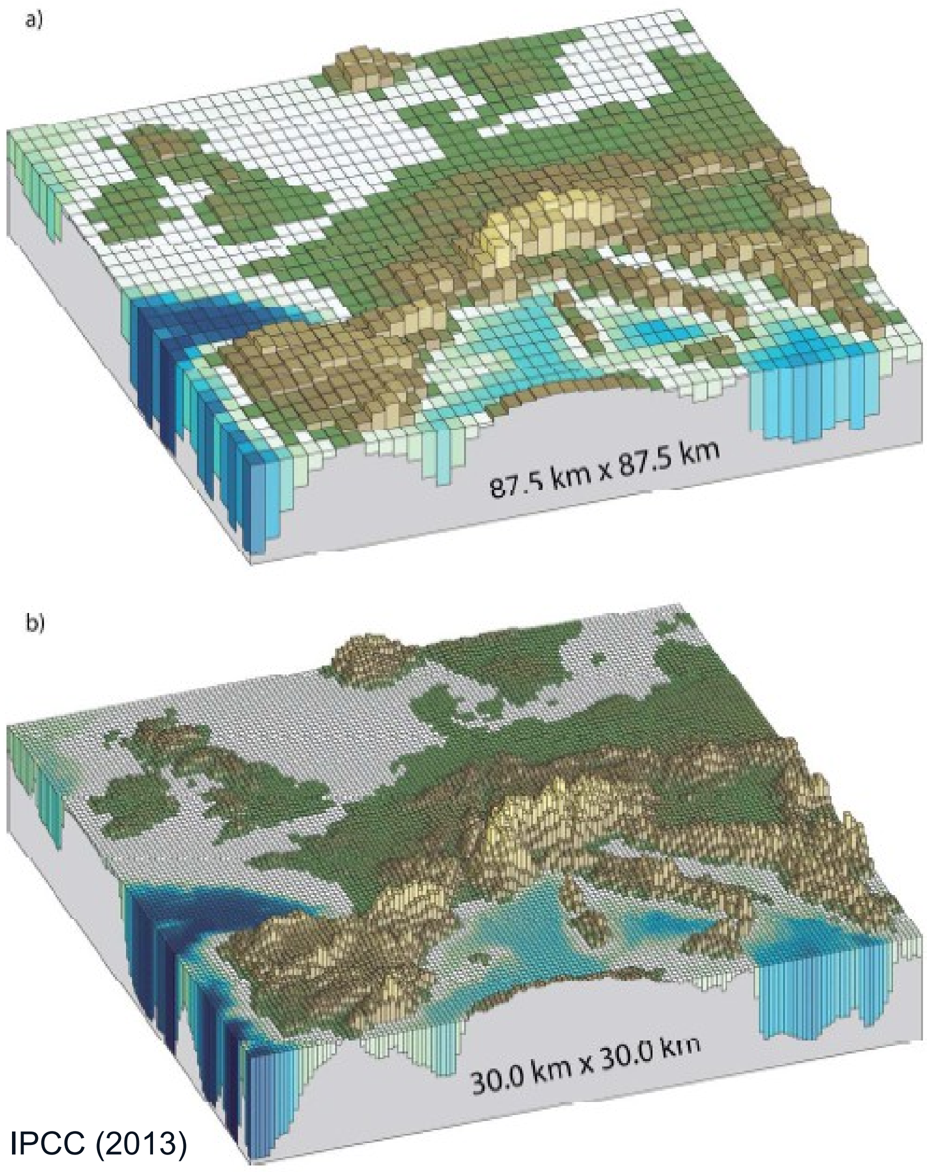
$$k_1 = \frac{1}{2} f(U_n, t_n) \qquad k_3 = \frac{1}{2} f(U_n + \Delta t\, \boldsymbol{k_2}, t_{n+1/2})$$

$$k_2 = \frac{1}{2} f(U_n + \Delta t\, \boldsymbol{k_1}, t_{n+1/2}) \qquad k_4 = \frac{1}{2} f(U_n + 2\Delta t\, \boldsymbol{k_3}, t_{n+1})$$

Historical evolution of climate models



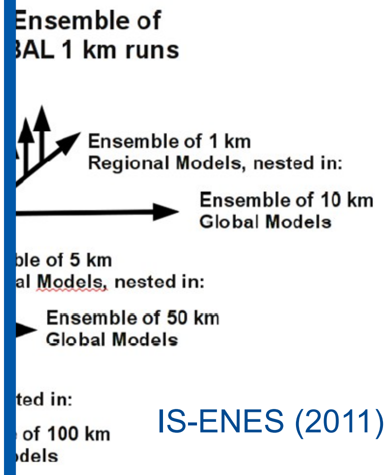
IPCC (2013)



IPCC (2013)

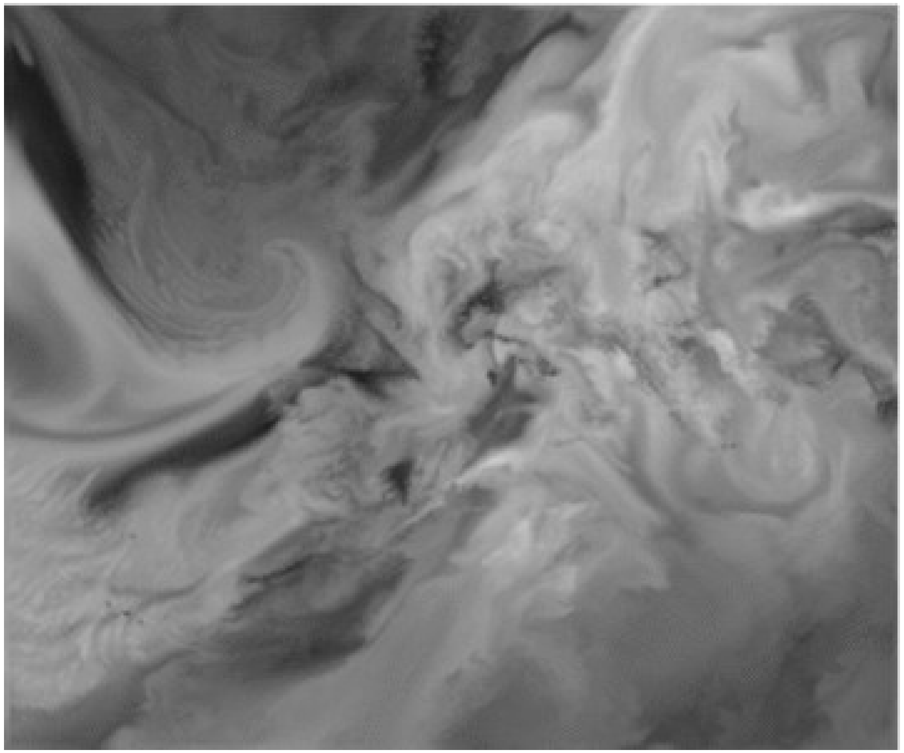
hor. res. 9 layers

or. res. 95 layers

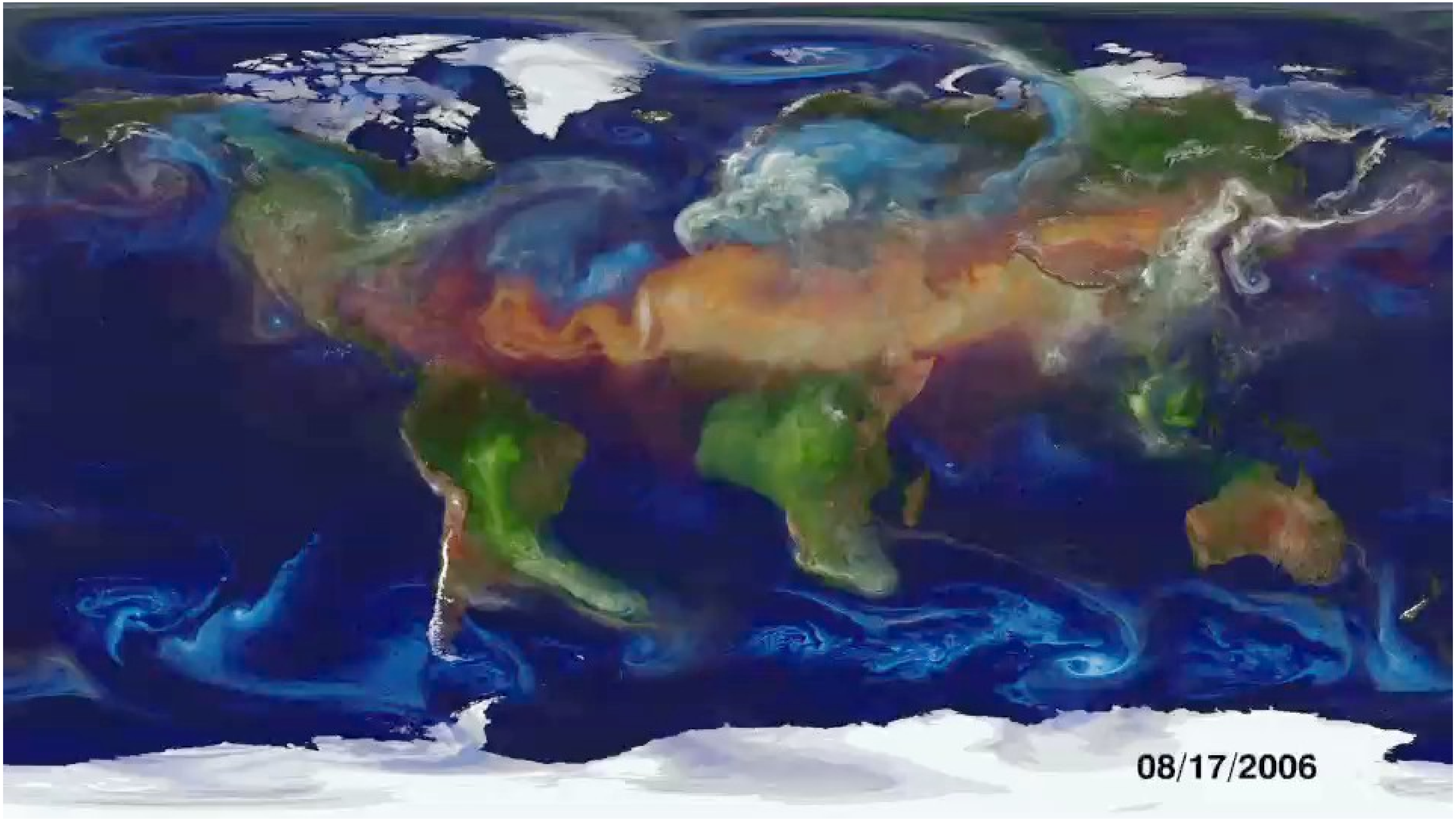


IS-ENES (2011)

Water vapor meteorological simulation



Global aerosol simulation



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Source: NASA GSFC

Climate change simulations – EC-Earth

- EC-Earth 2 m temperature projection
- EC-Earth Sea Ice coverage projection

BSC Model results:
NMMB/BSC-Chemical Transport Model

Fully on-line atmospheric-chemistry model

Multiscale: global to regional scales allowed

NMMB/ BSC-CTM

Nonhydrostatic Multiscale Model on the B-grid (NMMB)
meteo variables/parameters

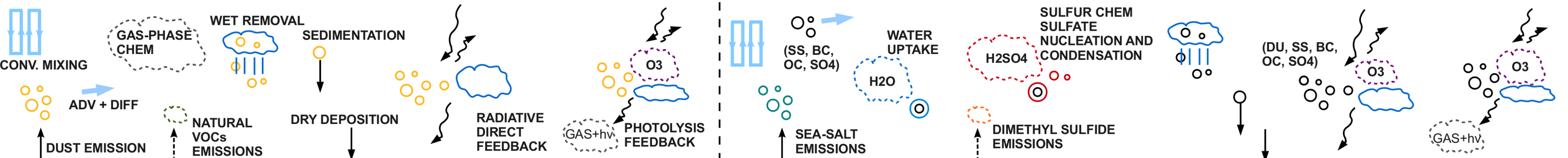
→ *Janjic and Gall (NCAR/TN 2012)*
→ *Janjic and Vasic (EGU2012)*
→ *Janjic et al. (MWR 2011)*
→ *(...)*

BSC Chemical Transport Model
(gas/aerosol variables: mass mixing ratios)

GAS-PHASE CHEM (52 species) → *Jorba et al. (JGR 2012)*

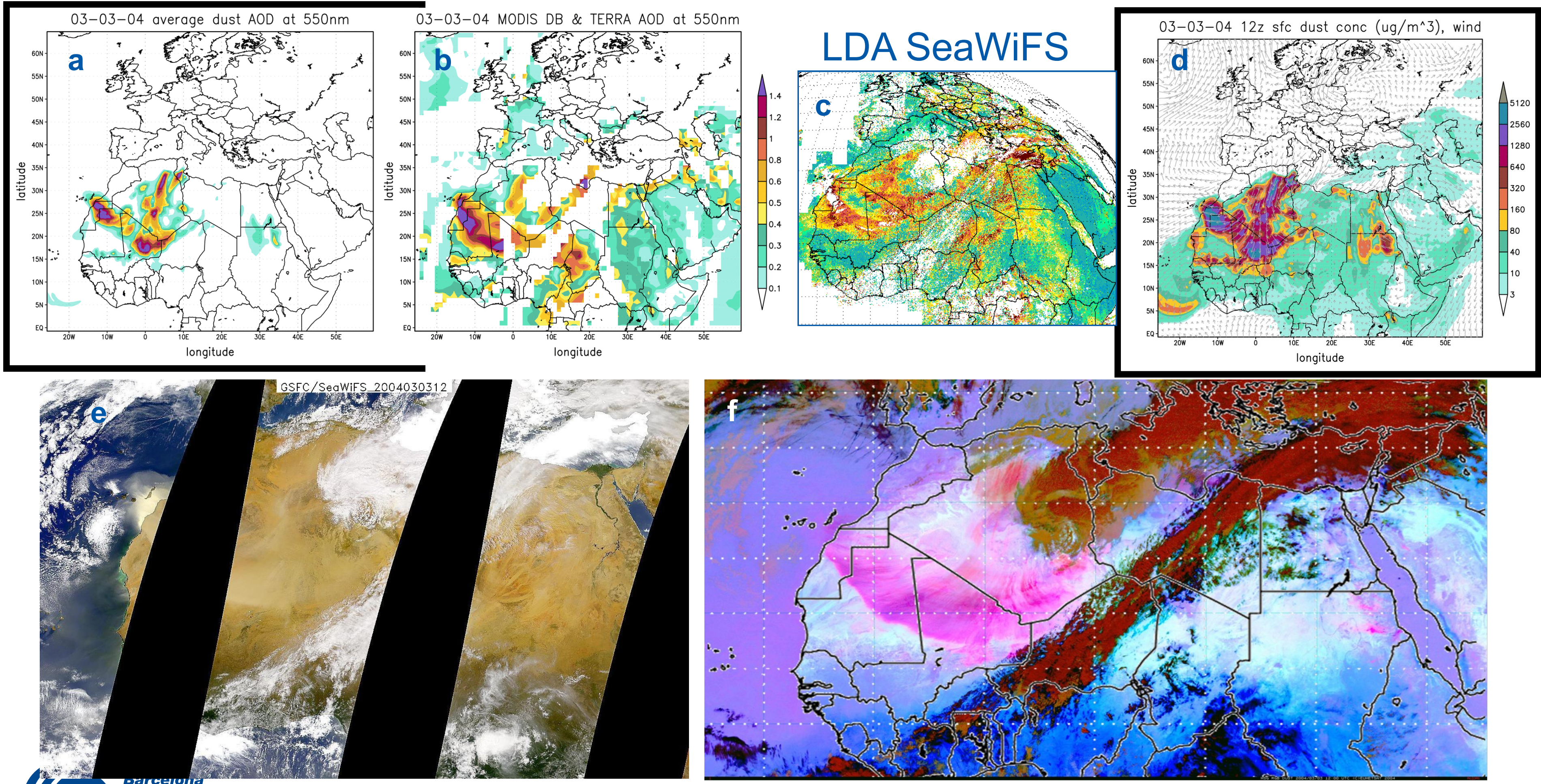
DUST (8 bins) → *Pérez et al. (ACP 2011)* → *Haustein et al. (ACP 2012)*

SEA-SALT → *Spada et al. (ACP 2013)*



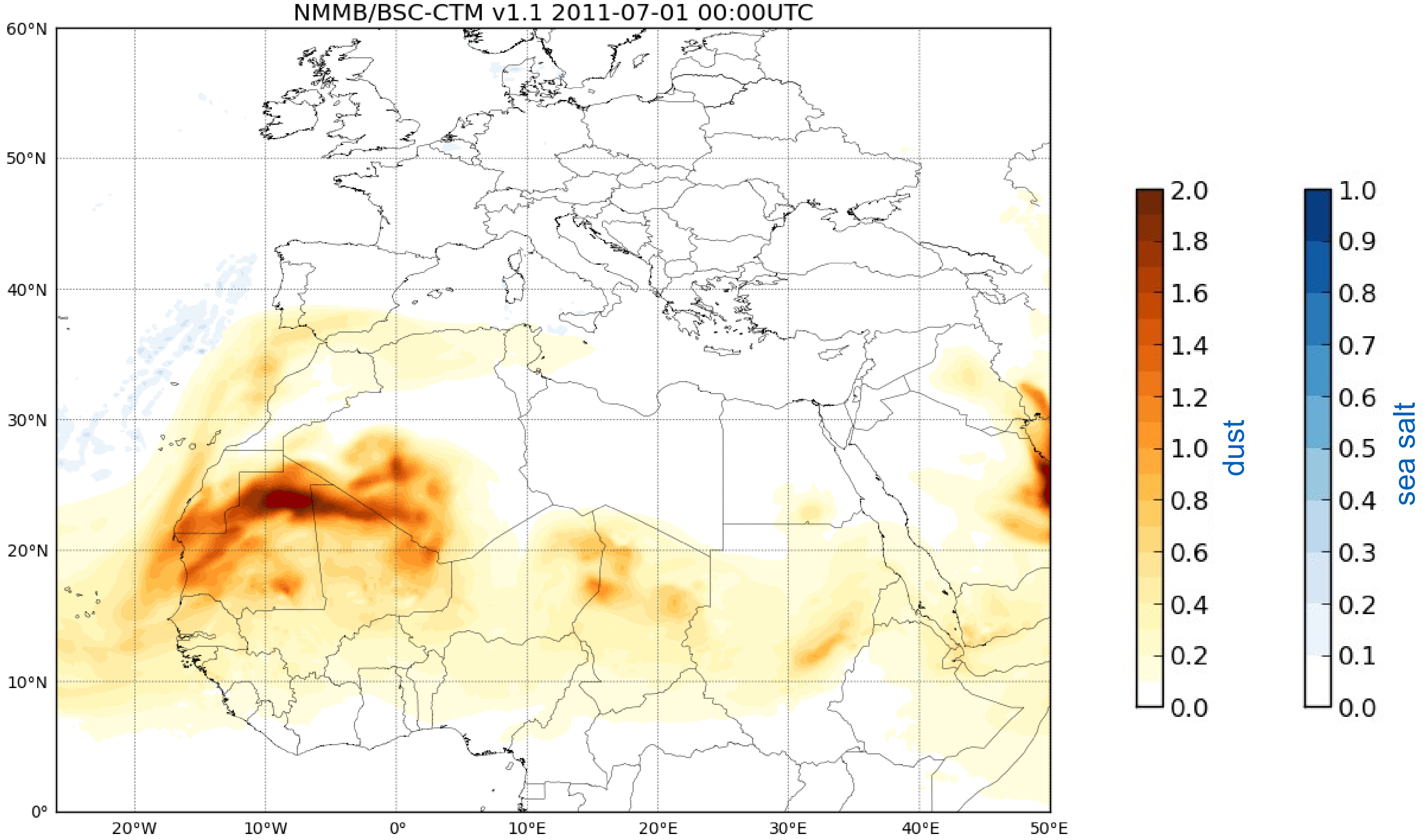
BSC model results: NMMB/BSC-CTM

- Global and regional annual simulations evaluated with:
  - Aeronet sun-photometer networks
  - LIDAR vertical profiles
  - Several satellite products
  - Surface concentrations
  - Emission and deposition fluxes


03-03-04 average dust AOD at 550nm
a
latitude
longitude
03-03-04 MODIS DB & TERRA AOD at 550nm
b
0.1
0.2
0.3
0.4
0.5
0.6
0.8
1
1.2
1.4
LDA SeaWiFS
c
03-03-04 12z sfc dust conc (ug/m^3), wind
d
3
10
40
80
160
320
640
1280
2560
5120
GSFC/SeaWiFS_2004030312
e
f

Pérez et al., 2011; Haustein et al., 2012

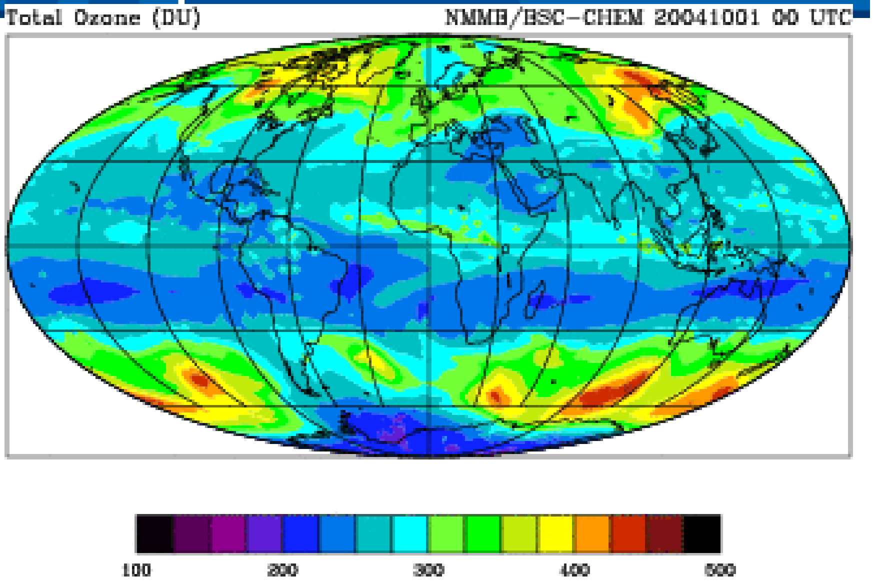
BSC model results: mineral dust and sea salt aerosols



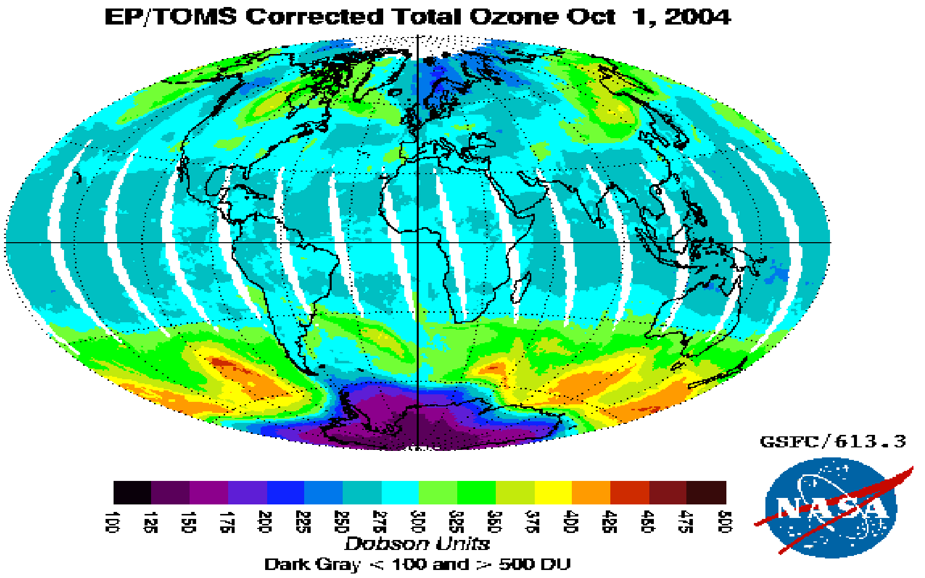
Aerosol optical depth

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

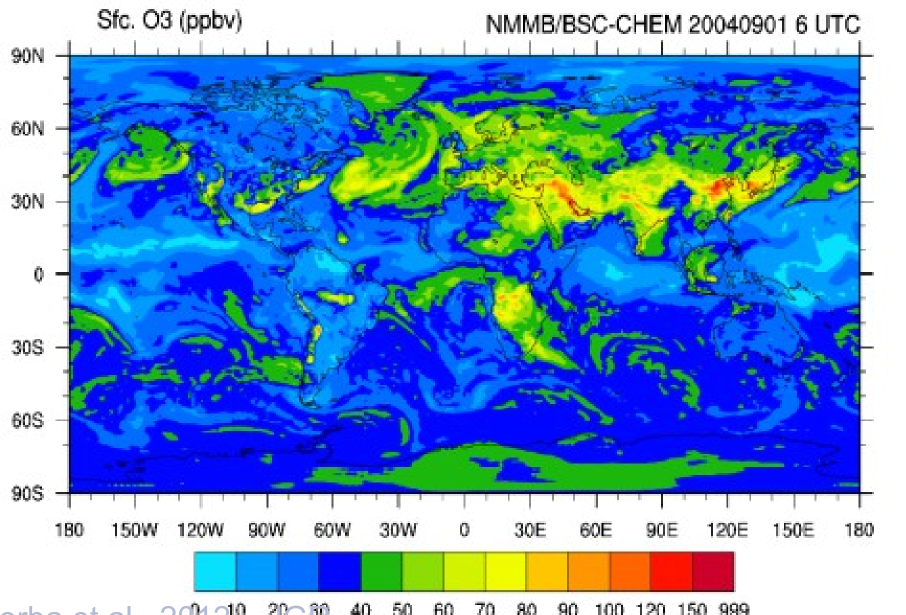
Gas phase results


Total Ozone (DU)

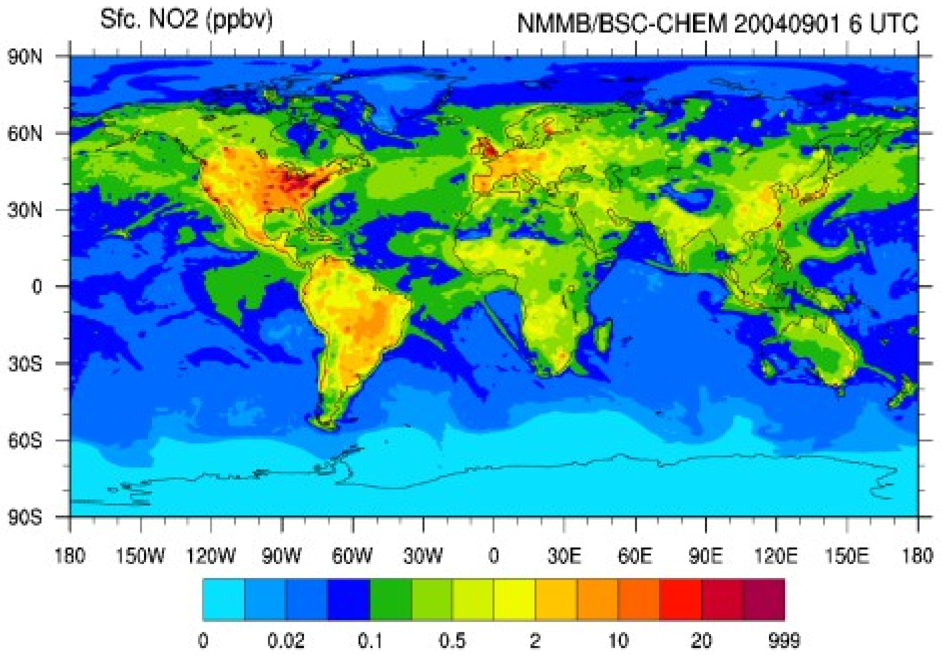
NMMB/BSC-CHEM 20041001 00 UTC


EP/TOMS Corrected Total Ozone Oct 1, 2004

GSFC/613.3


Sfc. O3 (ppbv)

NMMB/BSC-CHEM 20040901 6 UTC


Sfc. NO2 (ppbv)

NMMB/BSC-CHEM 20040901 6 UTC

Jorba et al., 2012 – JGR

Providing forecast products for

Log in

NORTHERN AFRICA-MIDDLE EAST-EUROPE (NA-ME-E) REGIONAL CENTER

WMO Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS)

WMO SDS WAS || Asia Regional Center

HOME | ABOUT US | FORECAST & PRODUCTS | PROJECTS & RESEARCH | MATERIALS | NEWS | EVENTS | CONTACT US

DUST FORECASTS
DUST OBSERVATIONS
GUIDANCE FOR FORECASTERS
TIME-AVERAGED VALUES
FORECAST EVALUATION
REANALYSIS
DATA POLICY

Forecast & Products

Home
- About us
- Forecast & Products
- Projects & Research
- Materials
- News
- Events
- Public Newsletter

Search

Search Site | Search

Latest News

UN Envoy Supports Greenbelts in Iraq to Combat Sandstorms

Feb 2

UNEP Alert S 'Forec warnir

Feb 1

Schola the Ur

rica-Middle East-Europe (NA-ME-E) Regional

— last modified May 29, 2012 03:33 PM

vailable for dust forecast

Multimodel Products

WMO SDS-WAS NA-ME-E Regional Center will be a Regional Specialized Meteorological Center

Forecast evaluation

Compared dust forecasts

Dust forecasts

Subscribe to the Public Newsletter!

To be informed about our activities, news and events related to dust. Frequency is almost monthly.

Full Name

Your email

Subscribe

· Mineral dust forecasts for SDS-WAS North Africa, Middle East and Europe portal

-forecast

NMMB/BSC-CTM selected to provide operational mineral dust forecast for the First Specialized Center for Mineral Dust Prediction of the World Meteorological Organization

try-

· Participate in the AQMEII on-line Air Quality model intercomparison project

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Session 1: Introduction to THE HPC ENVIRONMENT APPLIED TO Earth Sciences Applications

Objectives

- Introduce the technical point of view of Earth Sciences Applications
- Present some High Performance Computing topics
- Discover some models and how it works
- Discover basic visualization tools

- *Feel free to ask whenever you want…*

Outline

- Introduction
- HPC Environment
- Models
- Basic Visualization

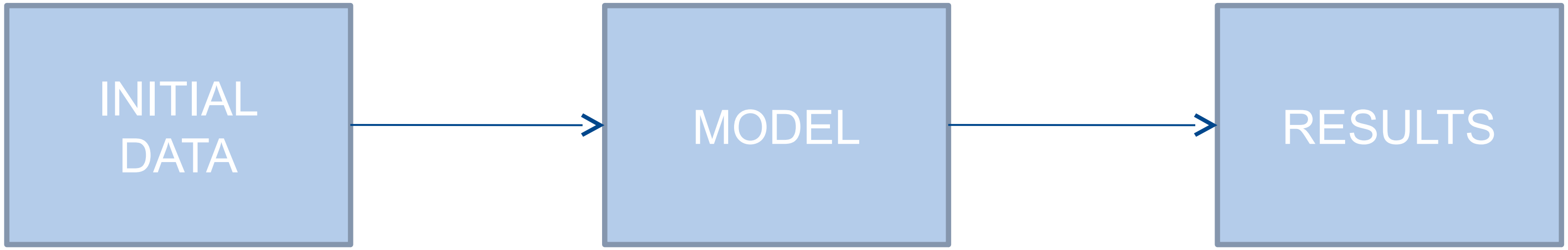
BSC

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Introduction

HPC environment
models
Basic visualization

What does “simulate” means in IT context



- Observations
- Data from other models
- Empty

- A collection of codes

- Binary data
- Maps
- Plots
- Text files

Why HPC?

- HPC: High Performance Computing

- Definition: *High-performance computing (HPC) is the use of parallel processing for running advanced application programs efficiently, reliably and quickly.**

- We need HPC to calculate the operations inside the models

** http://searchenterpriselinux.techtarget.com/definition/high-performance-computing*

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Introduction

HPC environment

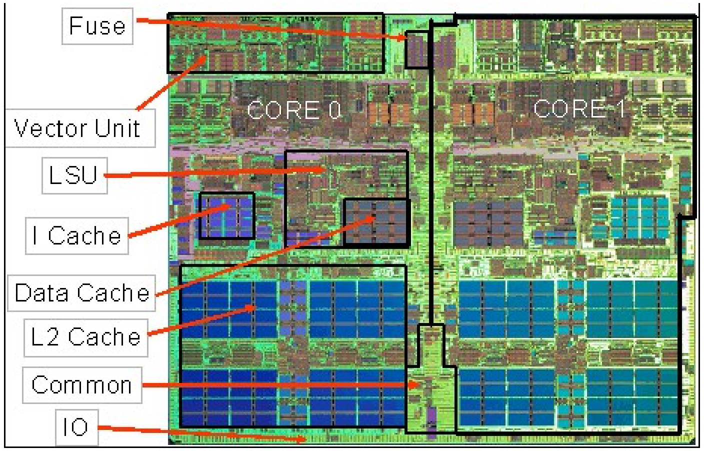
models

Basic visualization

What is it a Supercomputer

- Processors, Blades, BladeCenters and Racks & network

CORE



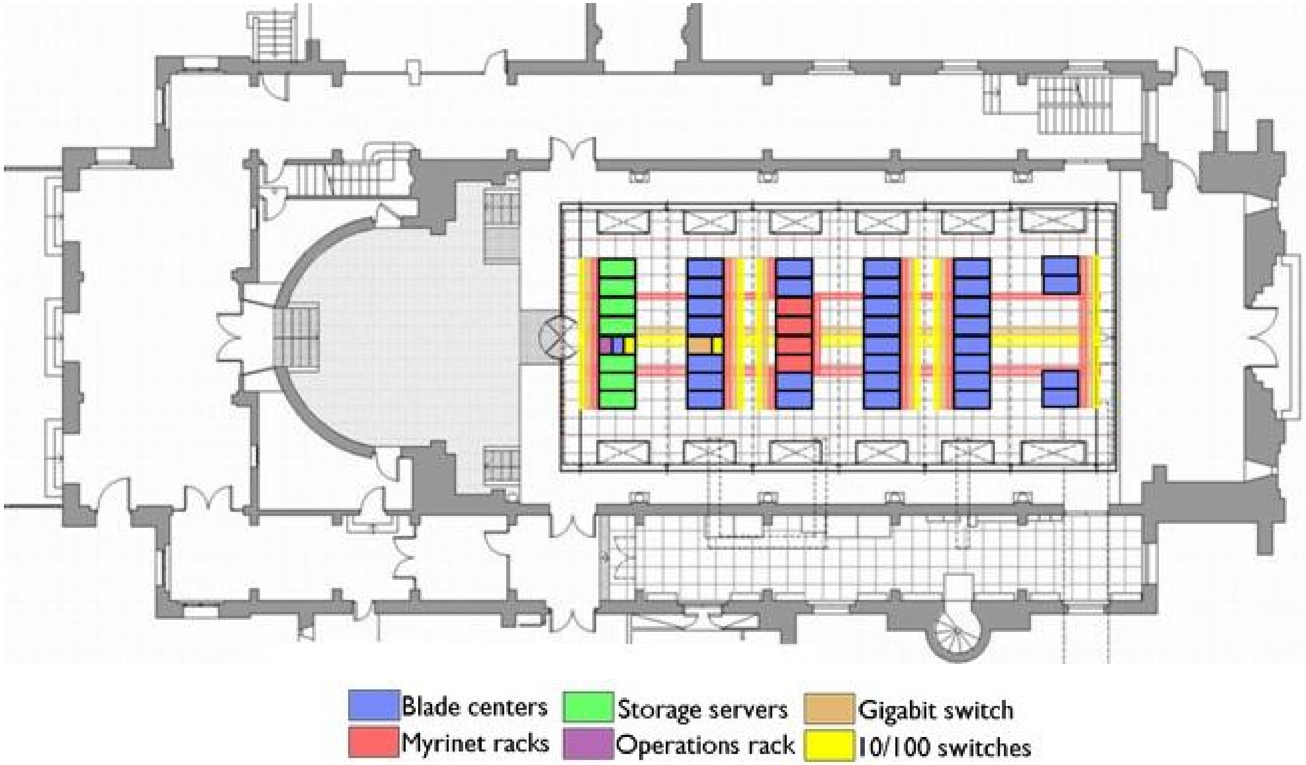
PROCCESSOR

BLADE

BLADE CENTER

RACKS

What is a Supercomputer?

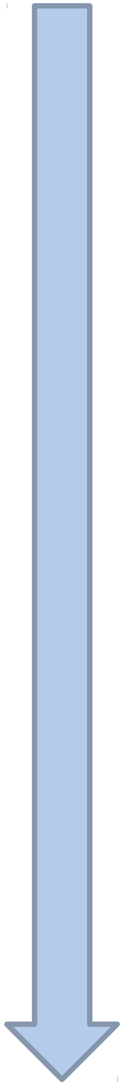


Mare Nostrum

- Peak Performance of 1,1 Petaflops
- 100.8 TB of main memory
- Homogeneous Nodes
  - 3,056 compute nodes
  - 2x Intel SandyBridge-EP E5-2670/1600 20M 8-core at 2.6 GHz
  - 8x4GB DDR3-1600 DIMMS (2GB/core)
- Heterogeneous Nodes
  - 42 heterogeneous compute nodes
  - 2x Intel SandyBridge-EP E5-2670/1600 20M 8-core at 2.6 GHz
  - 2x Xeon Phi 5110 P
  - 8x8GB DDR3-1600 DIMMS (4GB/core)
- 2 PB of disk storage
- Interconnection networks:
  - Infiniband FDR10
  - Gigabit Ethernet
- Operating System: Linux - SuSe Distribution

HPC systems evolution

- Vector Processors
  - Cray-1
- SIMD, Array Processors
  - Goodyear MPP, MasPar 1 & 2, TMC CM-2
- Parallel Vector Processors (PVP)
  - Cray XMP, YMP, C90 NEC Earth Simulator, SX-6
- Massively Parallel Processors (MPP)
  - Cray T3D, T3E, TMC CM-5, Blue Gene/L
- Commodity Clusters
  - Beowulf-class PC/Linux clusters
  - Constellations
- Distributed Shared Memory (DSM)
  - SGI Origin
  - HP Superdome
- Hybrid HPC Systems
  - Roadrunner
  - Chinese Tianhe-1A system
  - GPGPU systems
  - Accelerators



MORE CALCULATION POWER

ECMWF Supercomputer History

1979: Cray 1-A

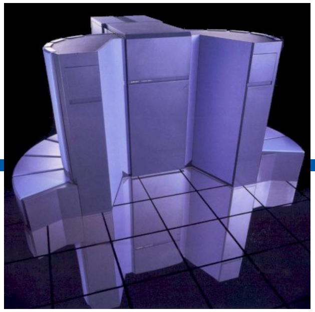
1 cpu (80 MHz)

Peak 160Mf

1986: Cray XMP-48

4 cpu (112 MHz)

Peak 880Mf



1990: Cray Y-MP 8/8-64

8 cpu (166 MHz)

1992: Cray C90

16 cpu

Peak 16 Gf

1994: Cray T3D

128 cpu

1996: Fujitsu VPP700

116 cpu

Peak 255 Gf

1999: Fujitsu VPP5000

100 cpu (80 MHz)

Peak 960Gf

2002: 2 IBM Cluster 1600

30 p690 SMP

Upgrade: 70 p690+

2006: 2 IBM Cluster

310 p5-575

Peak 38 Tf

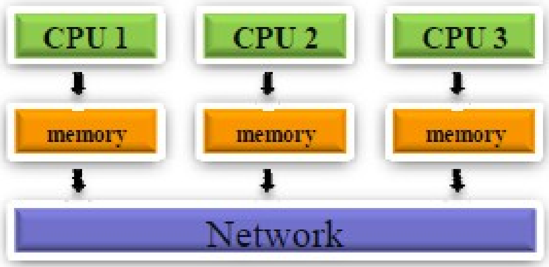
2009: IBM Cluster 286 POWER6 p6-575 servers

2012-2013: IBM Cluster 768 POWER7-775 servers

Peak 1.5 Pf

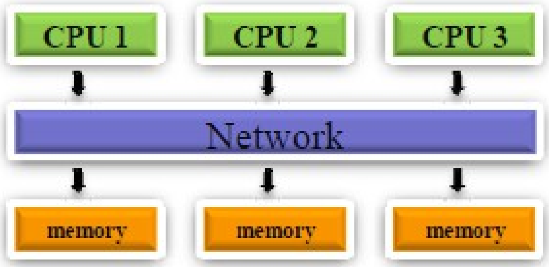
Different Architectures

- Distributed Memory
  - MPI, OpenMP



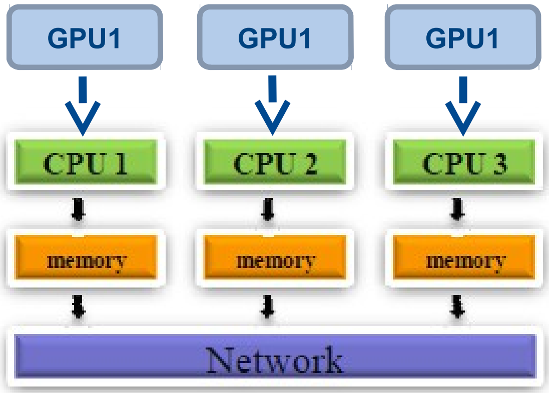
Distributed Shared Memory

- Shared Memory
  - OpenMP



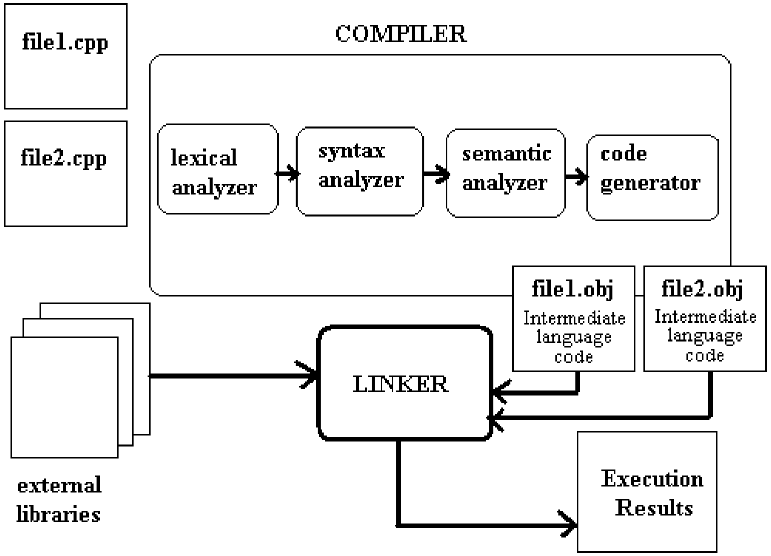
Shared Memory Systems

- Accelerators
  - CUDA, OpenGL, Cell, Xeon Phy…



Compiler

- We need to pass from codes written in languages that we understand (Fortran, C, C++,…)
- In a code that the machine can understand.



- Each final executable, will be different in each machine.
  - Sometimes compatible, sometimes not.

Programming Models

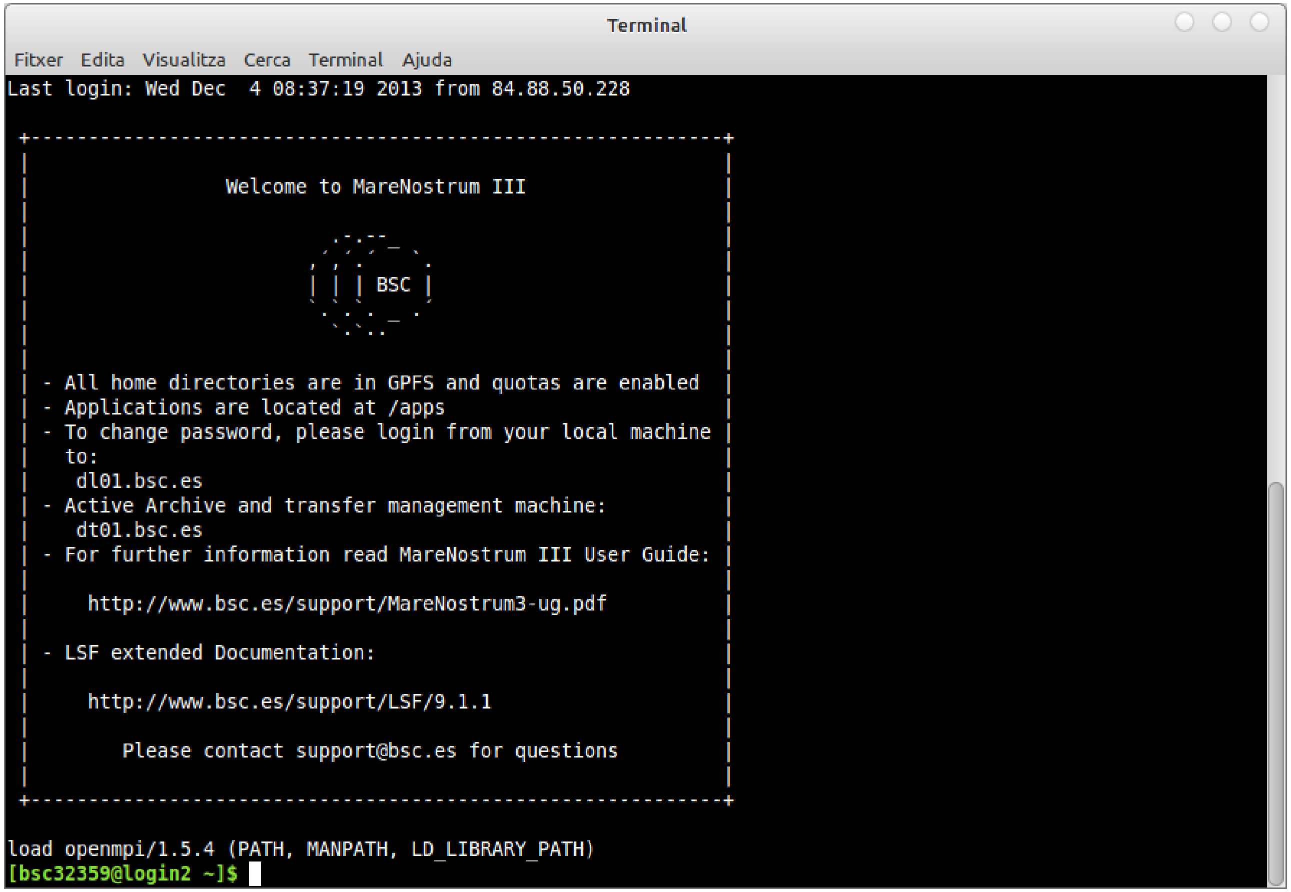
- Message Passing (MPI)
- Shared Memory (OpenMP)
- Partitioned Global Address Space Programming (PGAS) Languages
  - UPC, Coarray Fortran, Titanium
- Next Generation Programming Languages and Models
  - Chapel, X10, Fortress, ompSs
- Languages and Paradigm for Hardware Accelerators
  - CUDA, OpenCL
- Hybrid: MPI + OpenMP + CUDA/OpenCL
- New Xeon Phy Co-processors

Optimizing your code

- Optimizing the code
  - Getting better performances to run faster the models.
- Many optimizations can be done
  - Take time to write your code. Ask engineers.
  - In compilation time, ask for the usual compilations flags in your machine.
  - But be careful. Optimizations sometimes are no for free:
    - Compilation time is increased
    - It’s a tradeoff with precision.

Working in HPC

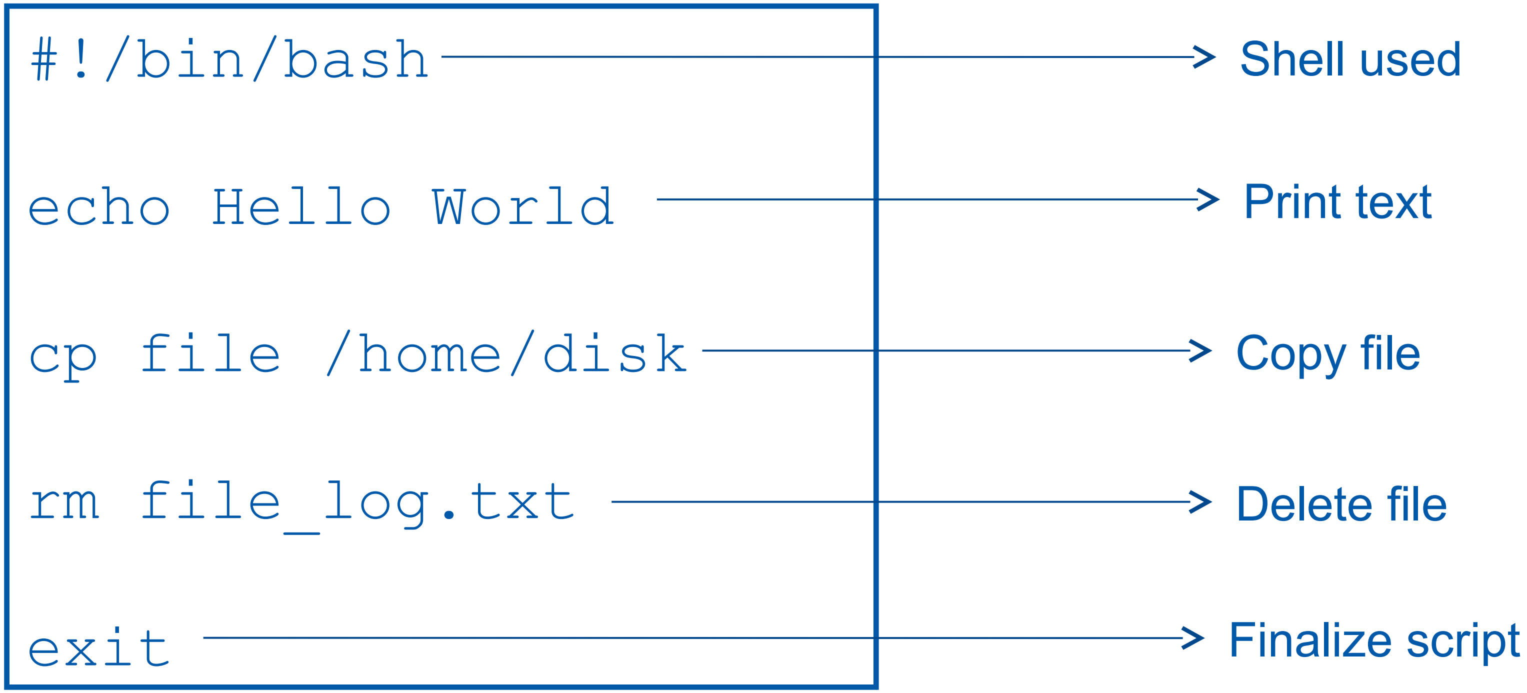
- Interface is done through SSH terminal
- Enable X11 forwarding (ssh -X)
- Unix terminal with a shell (can be shell, bourne shell, cshell…)



```
Last login: Wed Dec  4 08:37:19 2013 from 84.88.50.228

 +------------------------------------------------------------+
 |                                                            |
 |                 Welcome to MareNostrum III                 |
 |                                                            |
 |                          .-.--                             |
 |                        ,',',_   `.                         |
 |                       ; ; :       :                        |
 |                       | | | BSC |                          |
 |                       ` ` `       '                        |
 |                        . . . _   .                         |
 |                          `.`..                             |
 |                                                            |
 | - All home directories are in GPFS and quotas are enabled  |
 | - Applications are located at /apps                        |
 | - To change password, please login from your local machine |
 |   to:                                                      |
 |    dl01.bsc.es                                             |
 | - Active Archive and transfer management machine:          |
 |    dt01.bsc.es                                             |
 | - For further information read MareNostrum III User Guide: |
 |                                                            |
 |     http://www.bsc.es/support/MareNostrum3-ug.pdf          |
 |                                                            |
 | - LSF extended Documentation:                              |
 |                                                            |
 |     http://www.bsc.es/support/LSF/9.1.1                    |
 |                                                            |
 |        Please contact support@bsc.es for questions         |
 |                                                            |
 +------------------------------------------------------------+

load openmpi/1.5.4 (PATH, MANPATH, LD_LIBRARY_PATH)
[bsc32359@login2 ~]$ 
```

Working in HPC

- To run models, we use scripts.
- Script: text file containing a set of commands.



```
#!/bin/bash

echo Hello World

cp file /home/disk

rm file_log.txt

exit
```

- To edit scripts: many editors (vi, emacs, joe…)

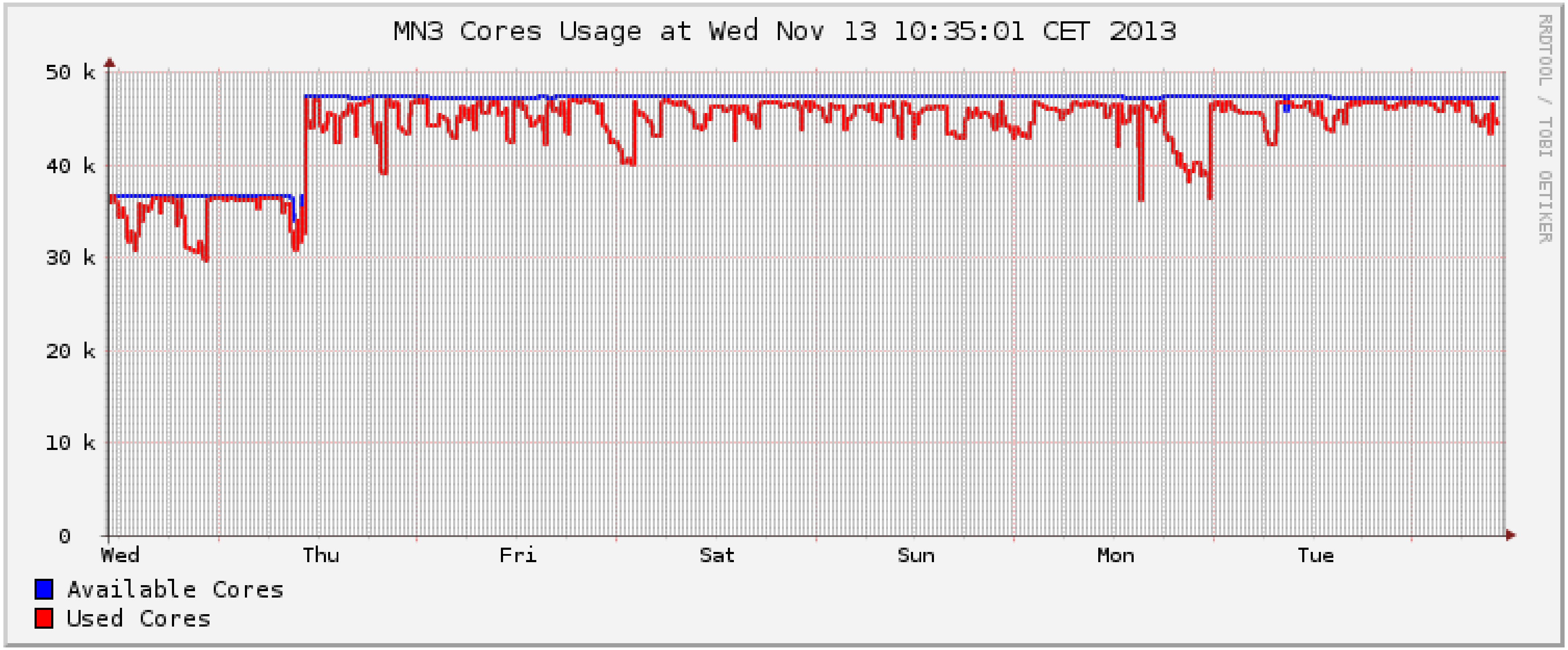
Working in HPC

- Defining variables
  - Some models, need variables to run or build
  - Also, we need access to usual programs

- To make variables persistent (not only in the current session), we need to include them in /gpfs/home/…/.bashrc

```
#NETCDF PROCESSING
export PATH=$PATH:/gpfs/apps/MN3/NETCDF/3.6.3/bin
export PATH=$PATH:/gpfs/apps/MN3/CDO/1.5.9/bin
export PATH=$PATH:/gpfs/apps/MN3/NCO/4.2.3/bin
export PATH=$PATH:/gpfs/apps/MN3/NCVIEW/2.1.2/bin
export
LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/gpfs/apps/MN3/NETCDF/3.6.3/lib
#IMAGES
export PATH=$PATH:/gpfs/apps/MN3/IMAGEMAGICK/6.8.1-9/bin
#GRADS
export PATH=$PATH:/gpfs/apps/MN3/GRADS/2.0.2/bin
export GADDIR=/gpfs/apps/MN3/GRADS/2.0.2/data/
#NCL
export NCARG_ROOT=/gpfs/apps/MN3/NCL/6.1.2
export PATH=$PATH:/gpfs/apps/MN3/NCL/6.1.2/bin
#R
export PATH=$PATH:/gpfs/apps/MN3/R/2.15.2/bin
#PANOPLY
export PATH=$PATH:/gpfs/apps/MN3/PANOPLY/3.1.7/
```

Sharing resources



Working in HPC

- Constraints
  - Many users
  - Many jobs
  - Limited ressources

- We need a job sumitter and scheduler
  - Distributes jobs through machine
  - Gives priority
  - Each job has an id
  - Makes a waiting queue
- The user submits the jobs and waits for the result.
- Commands in our machine
  - bsub
  - bkill
  - bjobs

Job submit

```
#!/bin/bash
#BSUB -n 8
#BSUB -o %J.out
#BSUB -e %J.err
#BSUB -cwd .
#BSUB -J helloworld_parallel
#BSUB -W 00:01
#BSUB -U patc

mpirun ./my_job

exit
```

Queue example

```
[bsc32359@login2 ~]$ bjobs
JOBID   USER     STAT  QUEUE      FROM_HOST    EXEC_HOST    JOB_NAME   SUBMIT_TIME
649736  bsc3235 RUN   bsc_es     s16r2b45     s11          *AG_AND1km Dec  4 06:17
649782  bsc3235 RUN   bsc_es     s16r2b45     s16          IMAG_CAT   Dec  4 07:30
649830  bsc3235 RUN   xlarge     login2       s10          WRF-TEST   Dec  4 08:38
649831  bsc3235 RUN   xlarge     login2       s03          *-WRF-2013 Dec  4 08:39
649841  bsc3235 RUN   bsc_es     login2       s15          *GE-NETCDF Dec  4 08:56
649852  bsc3235 RUN   bsc_es     login1       s15          KRIGING_AM Dec  4 09:14
649850  bsc3235 RUN   xlarge     s10r1b29     s03          *OPE-CAN_t Dec  4 09:12
649854  bsc3235 RUN   xlarge     s03r1b15     s08          *LIOPE-CAN Dec  4 09:15
649858  bsc3235 RUN   bsc_es     s03r1b15     16*s16       NDOWN_IP   Dec  4 09:17
                                              16*s02
```

Manage data

- In Earth Sciences community, HUGE size of data can be generated.

- We need different filesystems to work and store this data.
  - Immediate disk: disk in the supercomputer where the model is run. Very fast disk.
  - Medium term disk: one the data is generated, we need to work with the data to analyse it.
  - Long term storage: store the data, in case to reuse later. Usually tapes, speed is not a constraint.

- Data storage costs money !!!

Moving data

- Usually, we have to move data from our local computer to the supercomputer
  - Result files
  - Init files

- We use a secure copy
  - scp file.local user@machine.bsc.es:/PATH/.

File types

- Binary data is simple, but hard to read it.
  - I need to know how this data was created.
- We need standards to build files in order to exchange between groups.
  - NETCDF
  - HDF
  - GRIB1
  - GRIB2

```
netcdf grid_modWRF {
dimensions:
                x = 266 ;
                y = 169 ;
                time = UNLIMITED ; // (7 currently)
variables:
                double time(time) ;
                     time:standard_name = "time" ;
                     time:units = "day as %Y%m%d.%f" ;
                     time:calendar = "proleptic_gregorian" ;
                float imask(time, y, x) ;
                     imask:FieldType = 104 ;
                     imask:MemoryOrder = "XY " ;
                     imask:description = "LAND MASK (1 FOR LAND, 0 FOR WATER)" ;
                     imask:stagger = "" ;
                float lat(time, y, x) ;
                     lat:units = "degree_north" ;
                     lat:FieldType = 104 ;
                     lat:MemoryOrder = "XY " ;
                     lat:description = "LATITUDE, SOUTH IS NEGATIVE" ;
                     lat:stagger = "" ;
                float lon(time, y, x) ;
                     lon:units = "degree_east" ;
                     lon:FieldType = 104 ;
                     lon:MemoryOrder = "XY " ;
                     lon:description = "LONGITUDE, WEST IS NEGATIVE" ;
                     lon:stagger = "" ;
}
```

To get this information:

```
ncdump -h
```

File types

- Once we have filetypes, we need software to work with:
  - NCO: http://nco.sourceforge.net/
    - NetCDF operators
  - CDO: https://code.zmaw.de/projects/cdo
    - Climate Data Operators
  - LIBGRIB:
    - Library to work with gridded binaries.

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Introduction

HPC environment

models

basic visualization

Types of simulations

- Climate Simulations
  - Global scale
  - Large periods
  - Huge amount of data created
  - Execution time is not a critical constraint
  - Example: EC-EARTH model for 1900 to 2100, year simulation

- Operational Simulations
  - Global/Regional Scale
  - Small periods
  - Data created is smaller but postprocess products are more important
  - Execution time and reliabilty are very critical
  - Example: Daily weather forecast

Setting up a model

- A model is a collection of source codes
- We need to compile to build an executable
- The executable will run and produce results

- Usually, models have a building producedure
  - Configure
  - Makefiles
  - Scripting…

Computational demands

- Which domains are we simulating?
  - Barcelona
  - Catalunya
  - Spain
  - World

- Which resolution?
  - 1 km2
  - 4 km2
  - 12 km2
  - 50 km2

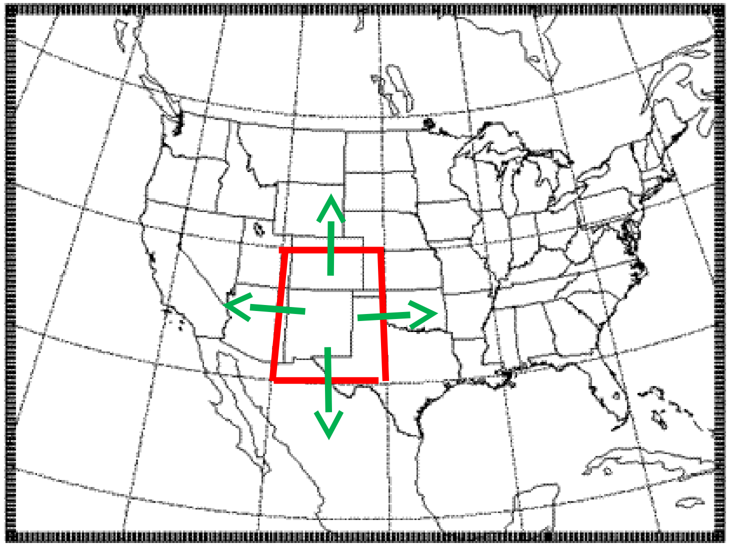
- How many variables we want to compute?
  - T2
  - U10, V10
  - QRAIN, QVAPOR

- Increasing this parameters, increases the system constraints
  - Computation Needs (CPU’s, Memory Bandwith…)
  - Data Storage

- Define this parameters in function of your hardware and time to serve forecast.

Parallelizing Atmospheric Models

- We need to be able to run this models in Multi-core architectures.
- Model domain is decomposed in patches
- Patch: portion of the model domain allocated to a distributed/shared memory node.

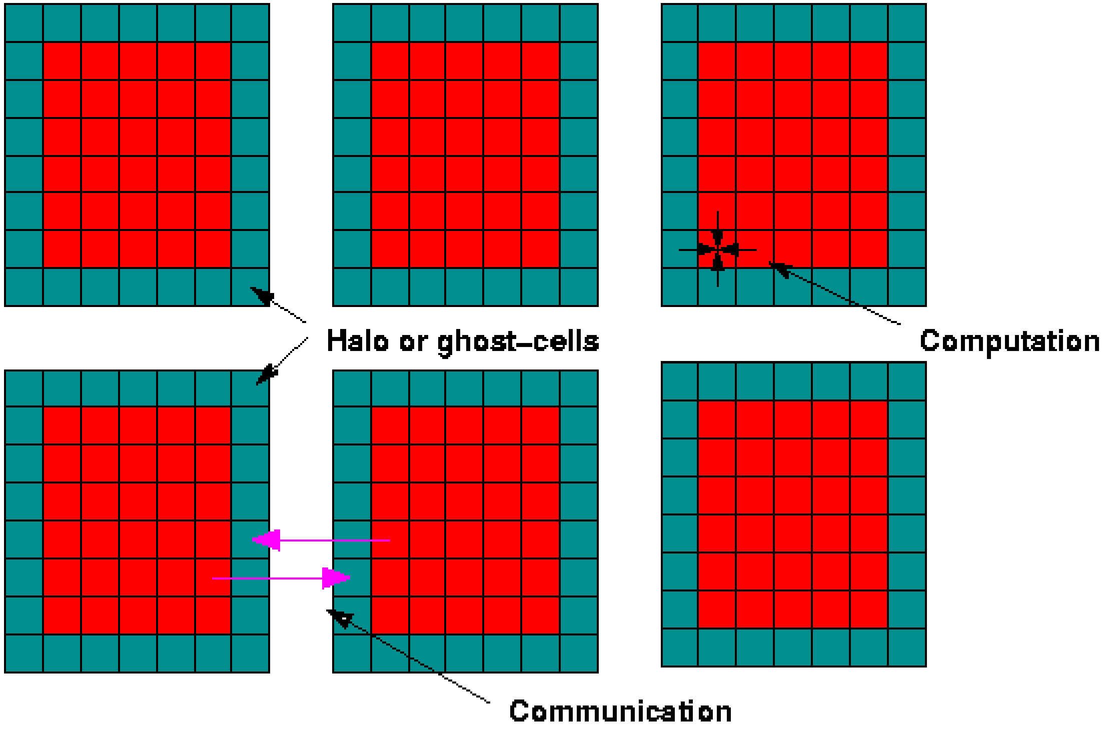


□ Patch

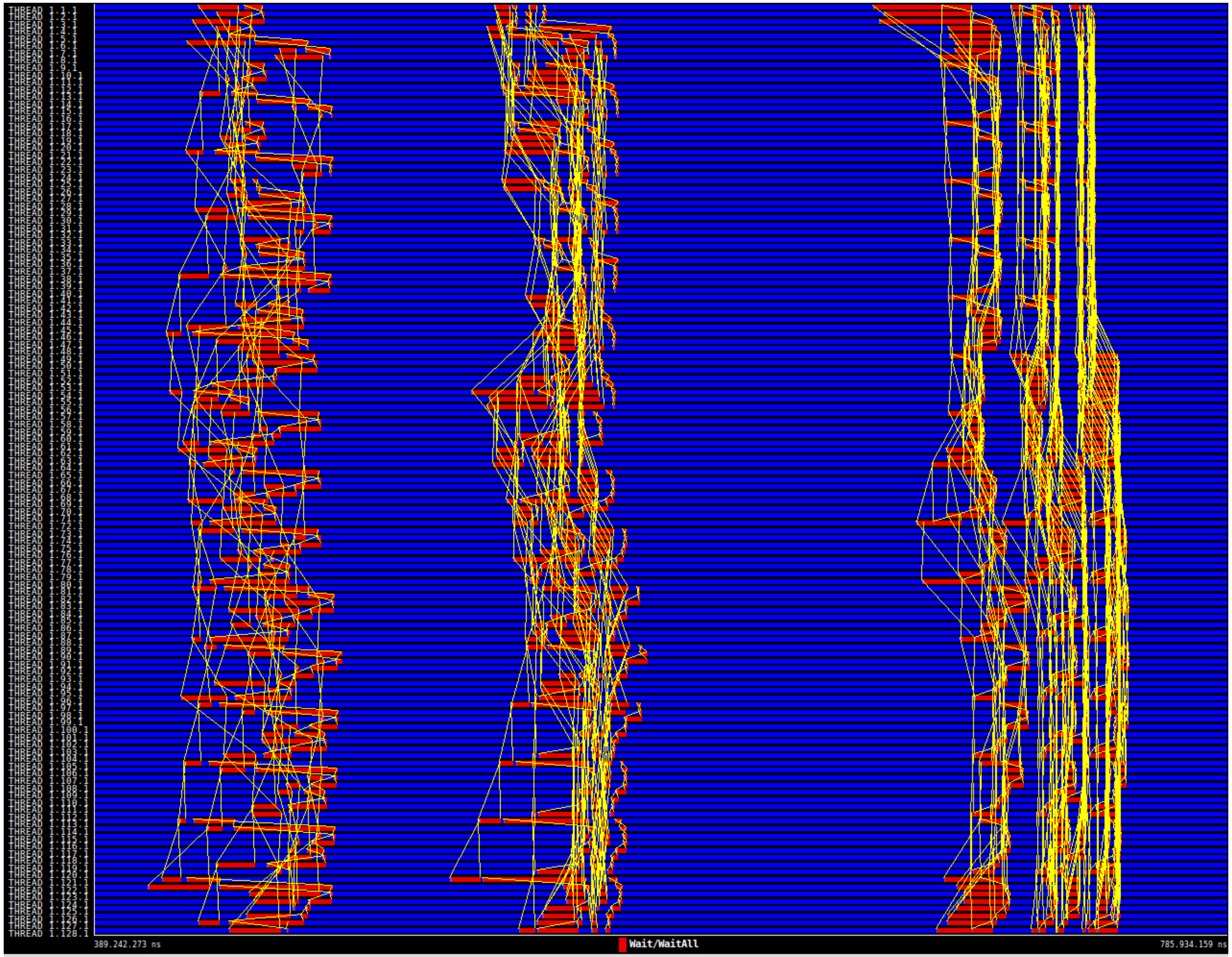
→ MPI/OpenMP Communication with neighbours

Parallelizing Atmospheric Models

- Halo exchange



Parallelizing Atmospheric Models

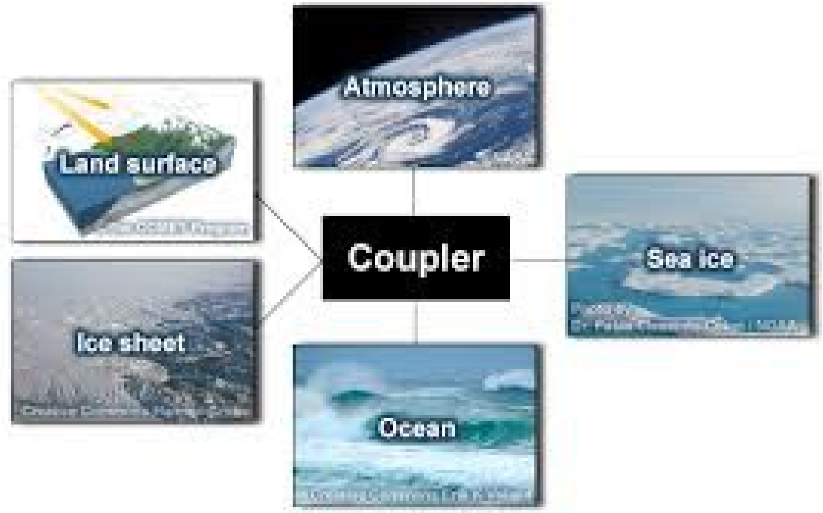


⬛	Idle
🟦	Running
⬜	Not created
🟥	Waiting a message
🟪	Bloking Send
🟫	Thd. Synchr.
🟩	Test/Probe
🟨	Sched. and Fork/Join
🟥	Wait/WaitAll
🟩	Blocked
🟪	Inmediate Send
🟪	Inmediate Receive
🟨	I/O
🟧	Group Communication
🟩	Tracing Disabled
🟨	Others
🟪	Send Receive

128 CPU’s
Iberian
Peninsula
Mare Nostrum

Couplers

- What is the role of a coupler ?
  - Exchange and transform information through two or more diferent models.
  - Manage the execution and synchronization of the codes.
  - Example: couple an ocean model and atmosphere.



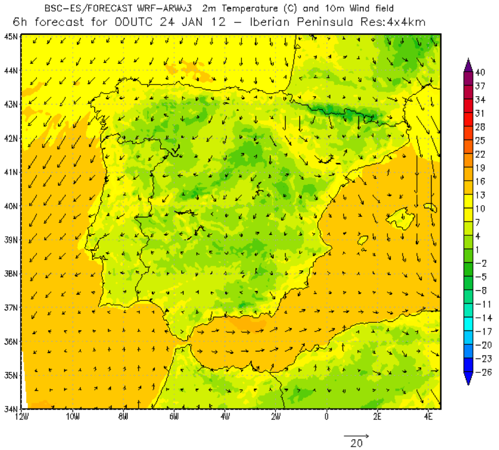
Couplers

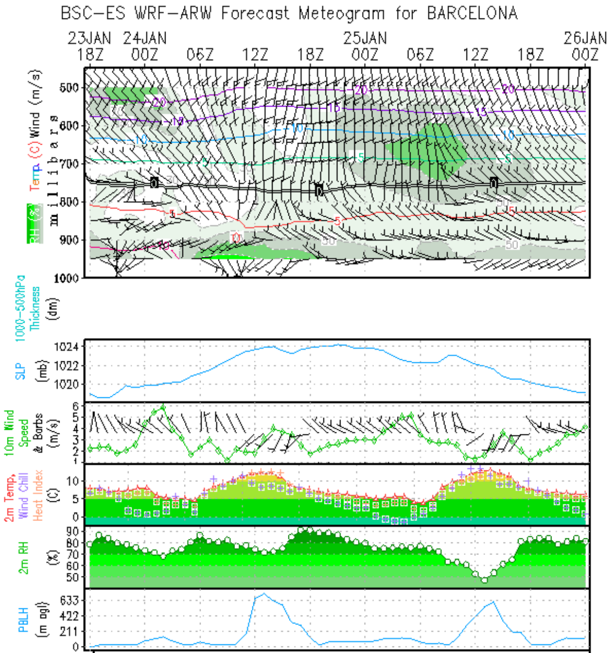
- Existing couplers
  - ESMF
  - OASIS
  - CESM

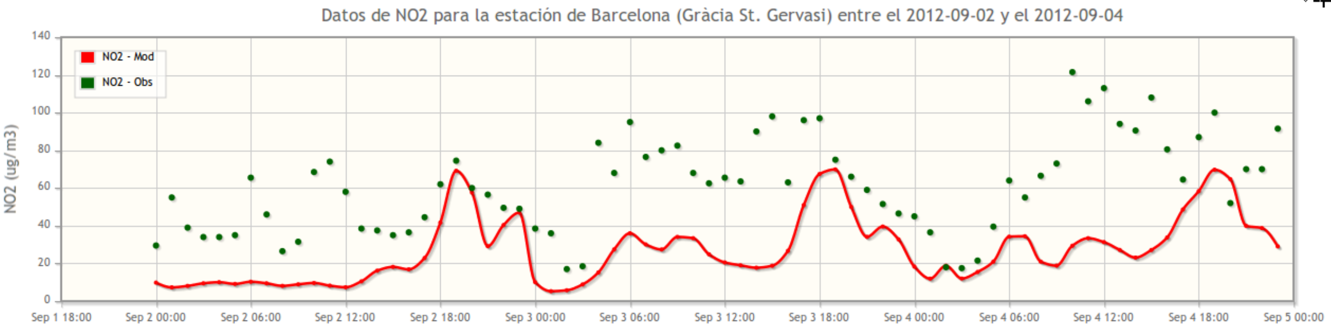
- Kinds of coupling
  - Offline: a model is run with the output of another one.
  - Online: models are run simultaneously.
    - Feedback between models
      - Example: chemistry and solar radiation.

Post-processing

- Once the model is run successfully, we need to post-process results to visualise data
  - Maps
  - Plots
  - Text files
  - 3D Animations


BSC-ES/FORECAST WRF-ARWv3 2m Temperature (C) and 10m Wind field
6h forecast for 00UTC 24 JAN 12 – Iberian Peninsula Res:4x4km

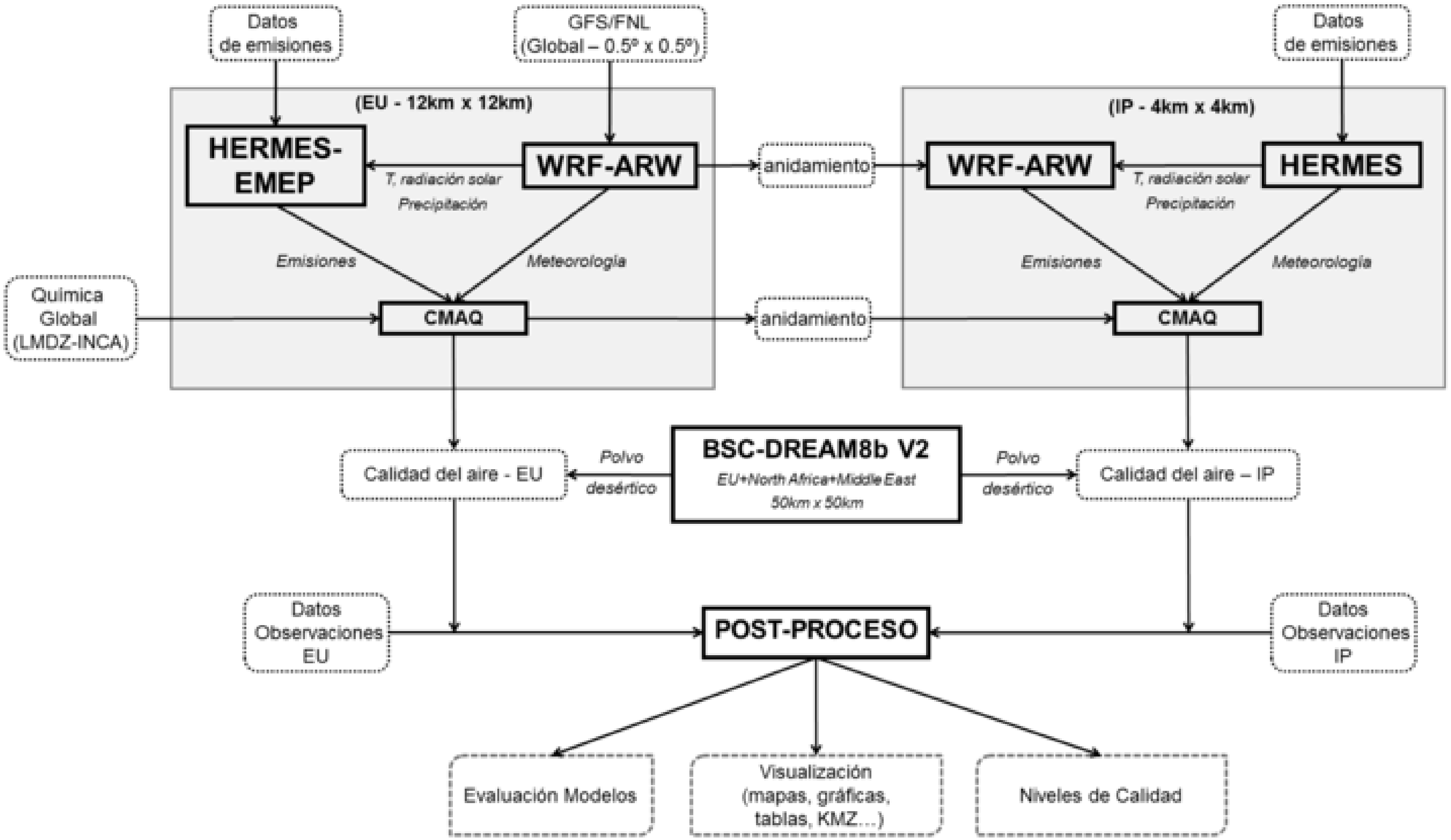

BSC-ES WRF-ARW Forecast Meteogram for BARCELONA


Datos de NO2 para la estación de Barcelona (Gràcia St. Gervasi) entre el 2012-09-02 y el 2012-09-04

Post-processing

- This stage can be done in the same supercomputer or in independent workstations
- Software to generate these products has to be installed and run
  - Examples
    - GRADS for pictures
    - Extract for text files to insert in a Database
    - GNU-plot to plot time series

Workflow



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Introduction

HPC environment

models

BASIC VISUALIZATION

Objective

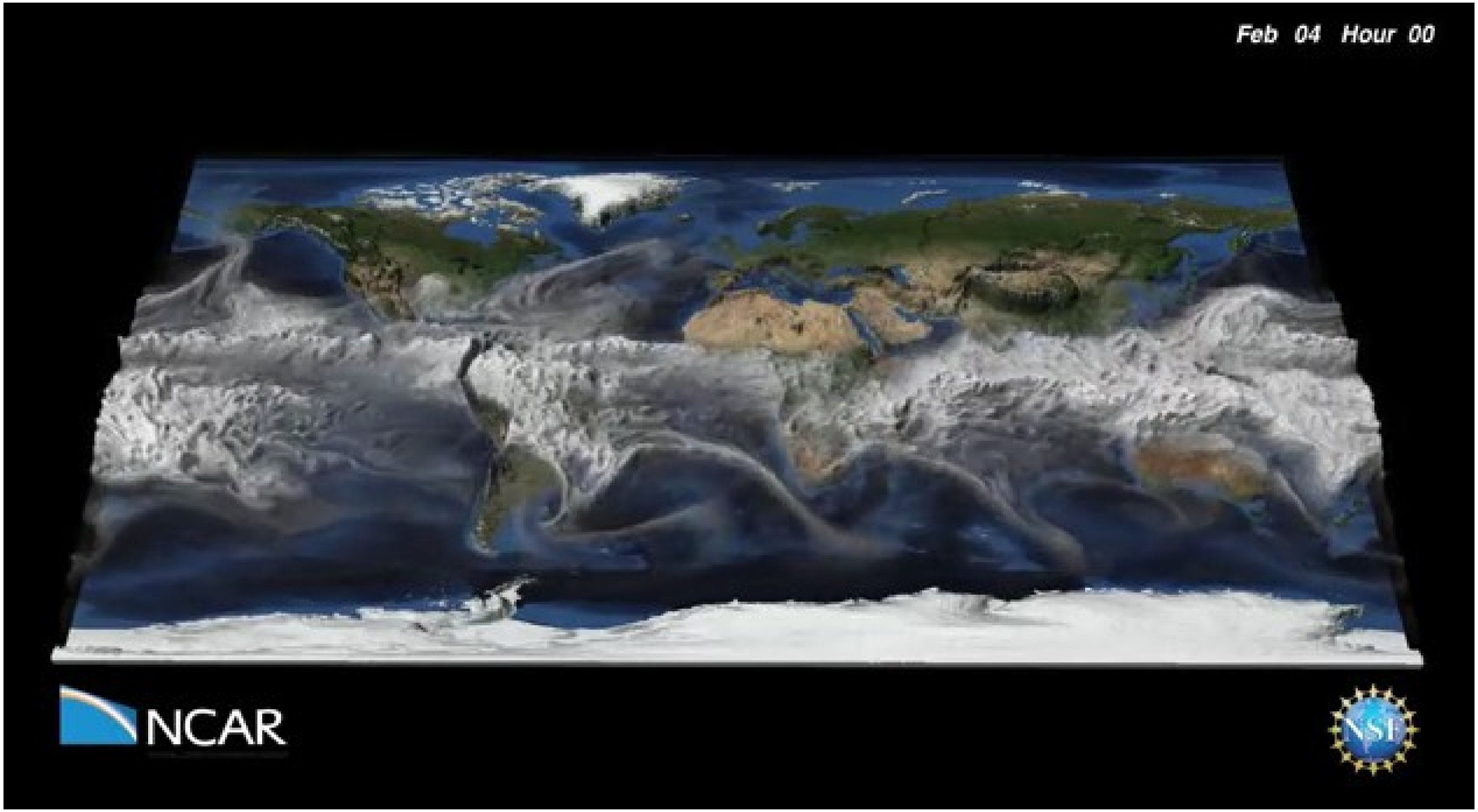
- Numerical models produce a huge amount of data on a variety of formats
  - Binary
  - NetCDF
  - ASCII
  - HDF5
  - GRIB
  - ...

- We need tools to analyse and visualise them

- In this section we will introduce some utilities freely available and widely used within the Earth Sciences community

- Many more are available…

Example



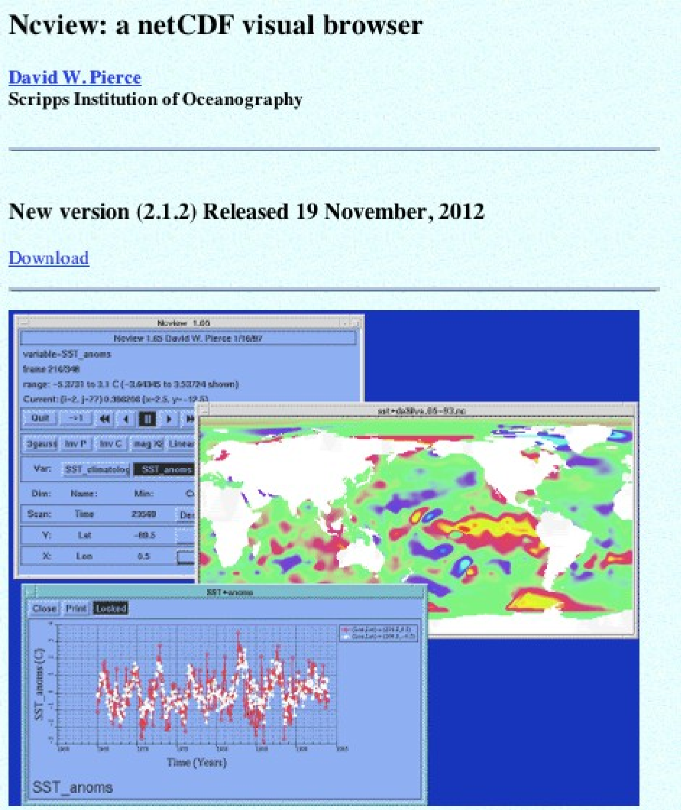
http://www.youtube.com/watch?v=am90gOaM16Q

Many Packages exists out there

- Visualization platforms
  - NCVIEW
  - PANOPLY
  - GRADS
  - NCL
  - MapGenerator

Visualization
Ncview: a netCDF visual browser

- Ncview is a visual browser for netCDF format files.
- Very useful to get a quick and easy look at your netCDF files.
- You can view:
  - simple movies of the data
  - view along various dimensions
  - take a look at the actual data values
  - change color maps
  - invert the data
  - etc.
- It runs on UNIX/Linux platforms.



Ncview: a netCDF visual browser

David W. Pierce
Scripps Institution of Oceanography

New version (2.1.2) Released 19 November, 2012

Download

- For more information:

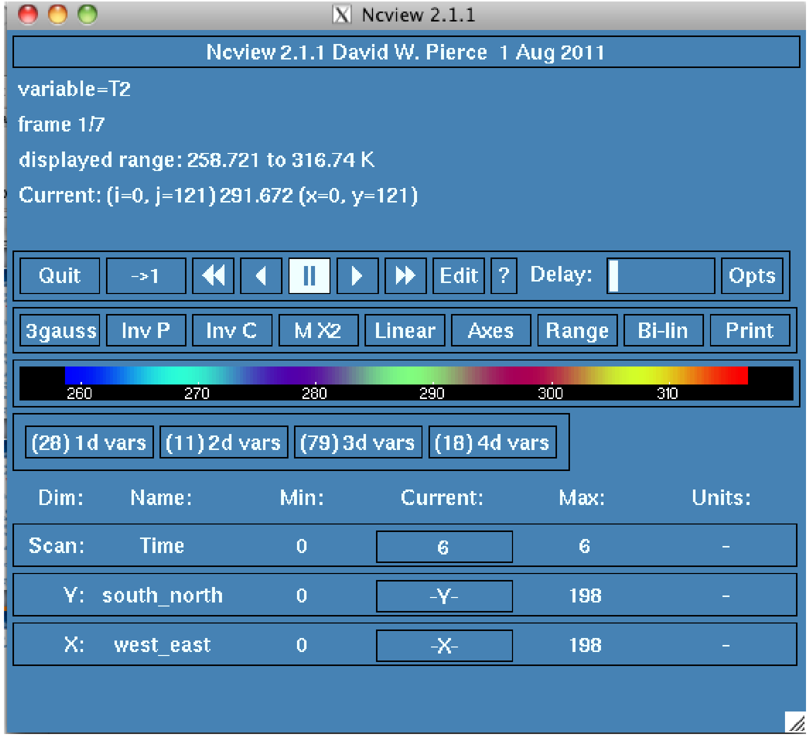
http://meteora.ucsd.edu/~pierce/ncview_home_page.html

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Ncview: examples

- Simple console
  - Meta data display
  - Time controls
  - Different color scales
  - Enlarge images
  - Define axes
  - Define range
  - Plot interpolated 2D fields
  - Plot time series

- Usage:
  - ncview wrfout_d01



Ncview 2.1.1 David W. Pierce 1 Aug 2011

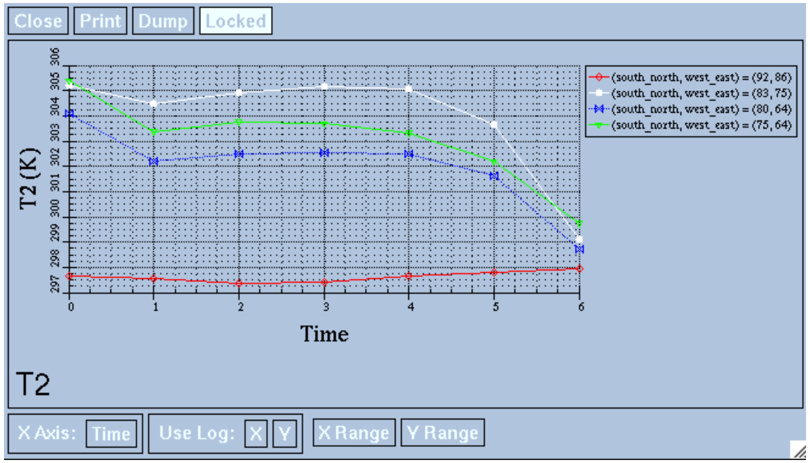
variable=T2

frame 1/7

displayed range: 258.721 to 316.74 K

Current: (i=0, j=121) 291.672 (x=0, y=121)

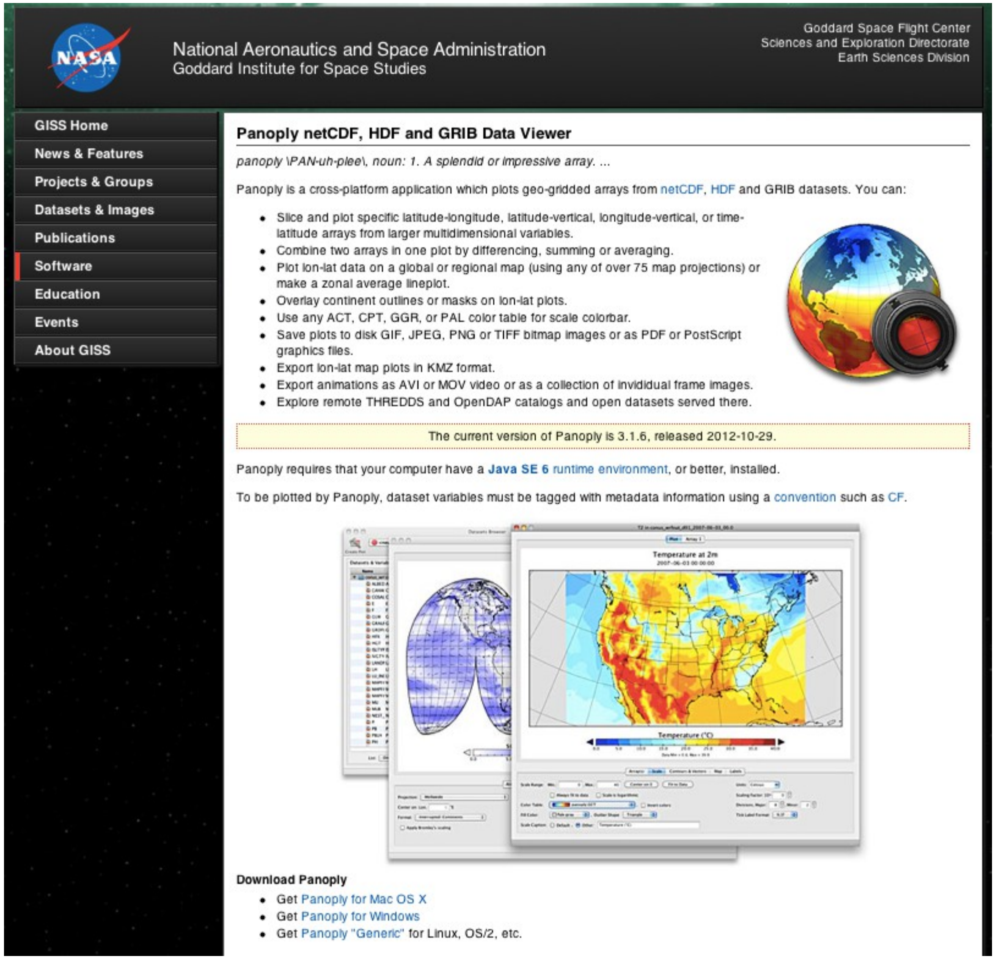
Dim:	Name:	Min:	Current:	Max:	Units:
Scan:	Time	0	6	6	-
Y:	south_north	0	-Y-	198	-
X:	west_east	0	-X-	198	-


Close
Print
Dump
Locked
T2 (K)
297
298
299
300
301
302
303
304
305
306
0
1
2
3
4
5
6
Time
(south_north, west_east) = (92, 86)
(south_north, west_east) = (83, 75)
(south_north, west_east) = (80, 64)
(south_north, west_east) = (75, 64)
T2
X Axis:
Time
Use Log:
X
Y
X Range
Y Range

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Visualization
Panoply: netCDF, HDF and GRIB Data Viewer

- Cross-platform application to plot geo-gridded arrays from netCDF, HDF and GRIB datasets
- Features:
  - Slice and plot specific latitude-longitude, latitude-vertical, longitude-vertical, or time-latitude arrays from larger multidimensional variables.
  - Combine two arrays in one plot by differencing, summing or averaging.
  - Plot lon-lat data on a global or regional map (using any of over 75 map projections) or make a zonal average lineplot.
  - Use any ACT, CPT, GGR, or PAL color table for scale colorbar.
  - Save plots to disk GIF, JPEG, PNG or TIFF bitmap images or as PDF or PostScript graphics files.
  - Export lon-lat map plots in KMZ format.
  - Export animations as AVI or MOV video or as a collection of individual frame images.
  - Explore remote THREDDS and OpenDAP catalogs and open datasets served there.
- It runs on Linux, UNIX, Mac OS X, Windows
- http://www.giss.nasa.gov/tools/panoply/



National Aeronautics and Space Administration
Goddard Institute for Space Studies

Goddard Space Flight Center
Sciences and Exploration Directorate
Earth Sciences Division

Panoply netCDF, HDF and GRIB Data Viewer

panoply \PAN-uh-plee\, noun: 1. A splendid or impressive array. ...

Panoply is a cross-platform application which plots geo-gridded arrays from netCDF, HDF and GRIB datasets. You can:

- Slice and plot specific latitude-longitude, latitude-vertical, longitude-vertical, or time-latitude arrays from larger multidimensional variables.
- Combine two arrays in one plot by differencing, summing or averaging.
- Plot lon-lat data on a global or regional map (using any of over 75 map projections) or make a zonal average lineplot.
- Overlay continent outlines or masks on lon-lat plots.
- Use any ACT, CPT, GGR, or PAL color table for scale colorbar.
- Save plots to disk GIF, JPEG, PNG or TIFF bitmap images or as PDF or PostScript graphics files.
- Export lon-lat map plots in KMZ format.
- Export animations as AVI or MOV video or as a collection of invididual frame images.
- Explore remote THREDDS and OpenDAP catalogs and open datasets served there.

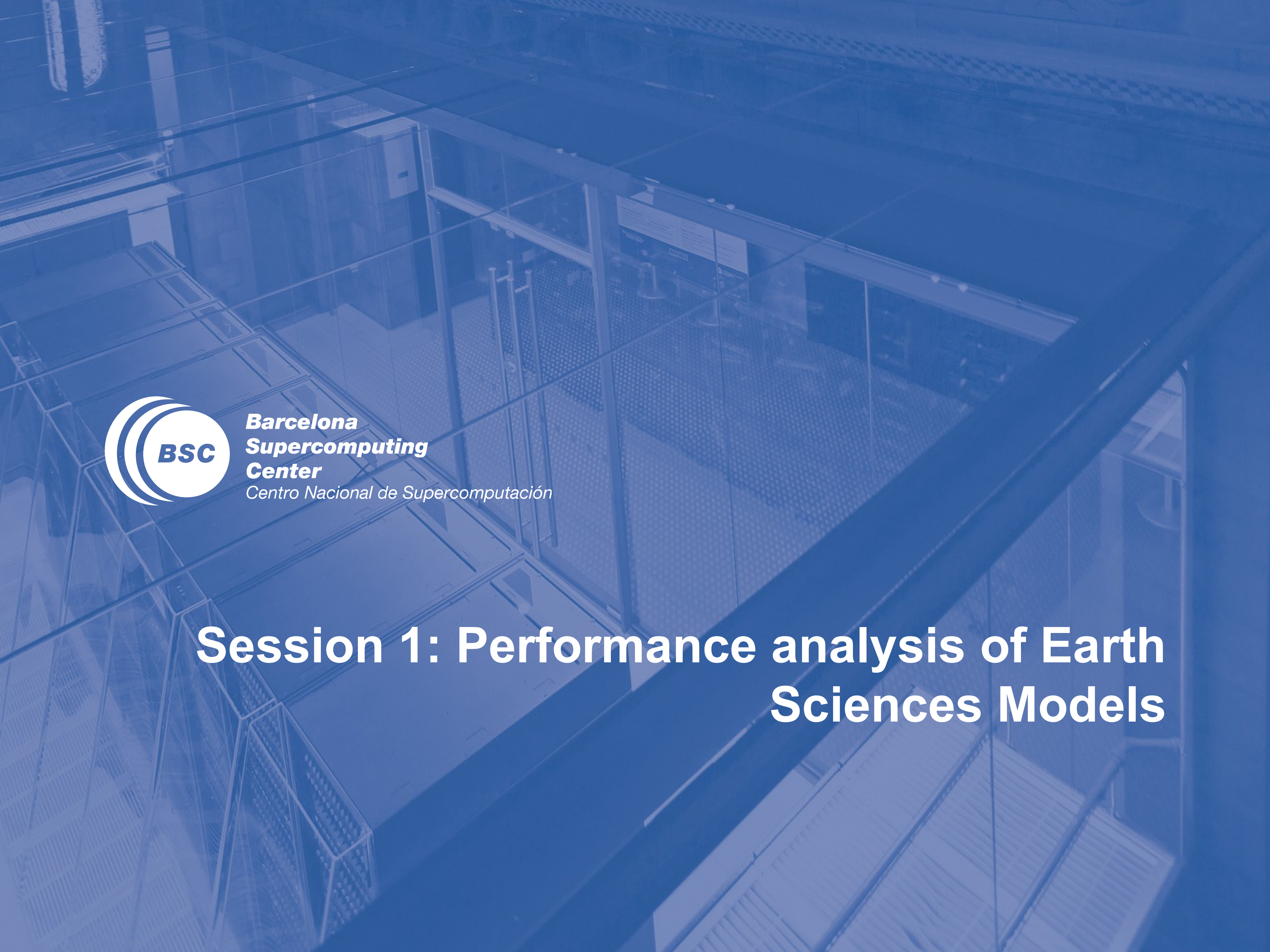
The current version of Panoply is 3.1.6, released 2012-10-29.

Panoply requires that your computer have a **Java SE 6** runtime environment, or better, installed.

To be plotted by Panoply, dataset variables must be tagged with metadata information using a convention such as CF.

Download Panoply

- Get Panoply for Mac OS X
- Get Panoply for Windows
- Get Panoply "Generic" for Linux, OS/2, etc.



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Session 1: Performance analysis of Earth Sciences Models

Why Performance Analysis?

Is it needed?

- Understand the behaviour of an application
  - Can we execute an application faster?
  - Is a simulation finished on time?

- Optimize an application
- Do we use the resources optimal?
- Could we decrease the electricity bills of a supercomputer?
- Predict physical catastrophes on time?
- Is needed for the next-generation of supercomputers
- How do you know that your application can not be improved?

What is WRF?

Weather Research & Forecasting (WRF) Model

- Mesocale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs.
- Suitable for a broad spectrum of meteorological applications across scales ranging from meters to thousands of kilometers.
- It is a supported “community model” - www.wrf-model.org
- Development led by NCAR, NOAA/GSD and NOAA/NCEP/EMS with partnerships at AFWA, FAA, NRL, and collaborations with universities and other government agencies in the US and worldwide.
- WRF is going to be explained in another session.

Our case

Used data

- Real data from Caliope system
  - Simulation of 36 hours
  - Grid size: 400 x 480

Marenostrum Supercomputer

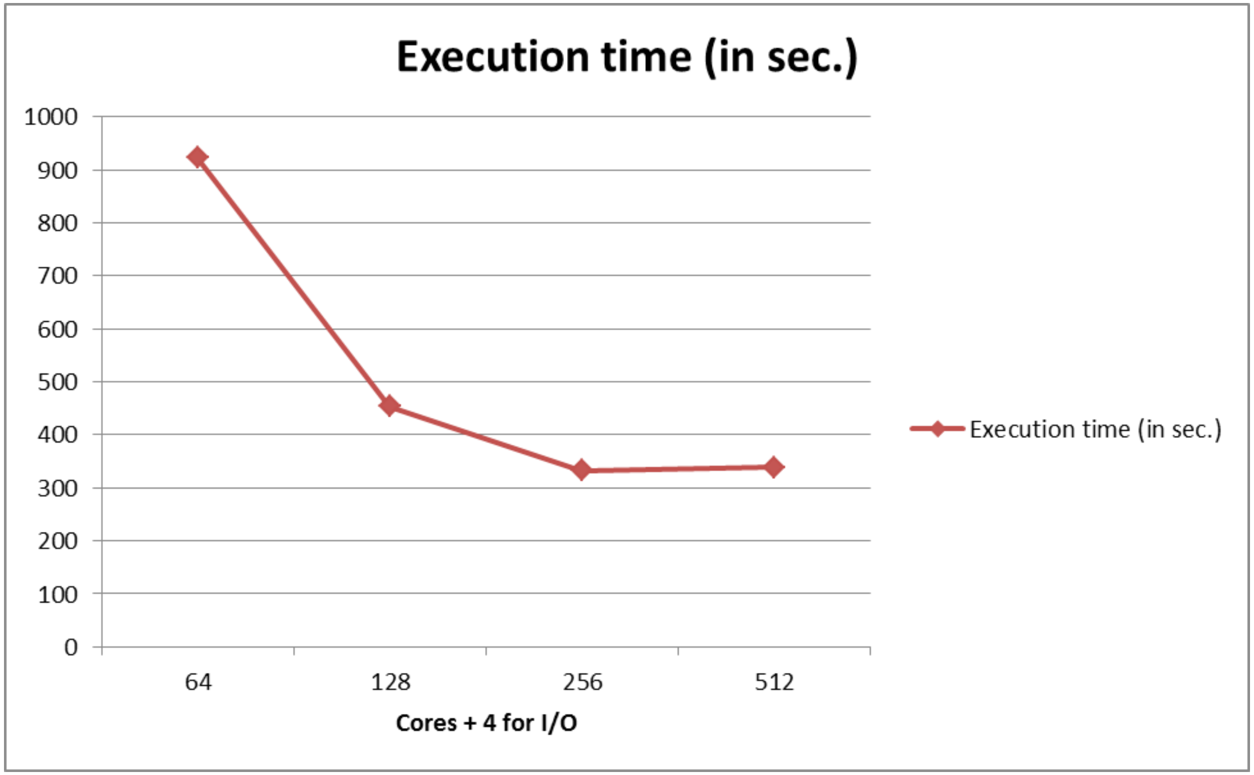
- We used up to 516 cores (33 nodes)
- Intel SandyBridge-EP E5-2670/1600 20M 8-core at 2.6 GHz
- Interconnection: Infiniband

Objectives

- Observe the performance of the WRF model
- Indicate any optimization technique through the namelist file
- Compare both MPI and MPI/OpenMP versions
- Present the various I/O modes
- Visualize Paraver traces

Execution Time

- Execution time in seconds



- Bad scaling for 256+ cores
- Is it because of the workload or the setup of I/O servers?

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Remarks

- According to WRF FAQ in order to use the model as expected, there should be at least 15 x 15 points per process

- The previous behavior means that we should increase the grid resolution or setup a different I/O configuration (explained later)

- Before decide about some results we should be familiar with the internals of an application

Speedup

- Speedup refers to how much a parallel algorithm is faster than a corresponding sequential algorithm

- Is defined by the following formula:

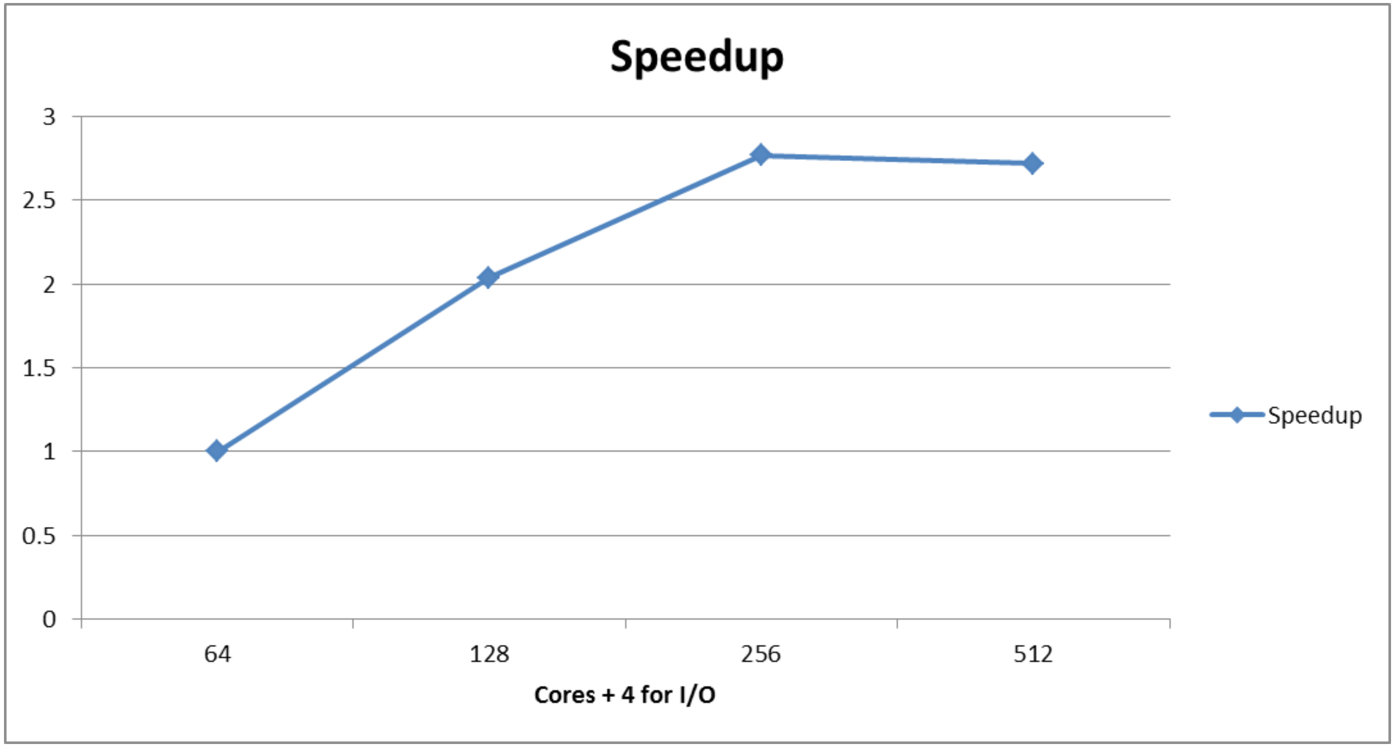
$$Sp = \frac{T_1}{T_p}$$

  - p is the number of processors
  - T_1 is the execution time of the sequential algorithm
  - Tp is the execution time of the parallel algorithm with p processors

- We do not have the duration of the serial version so we'll compute the speedup with the 64 processors as base

Speedup

- We can see the speedup for our execution
- From 256 to 512 cores there is no improvement



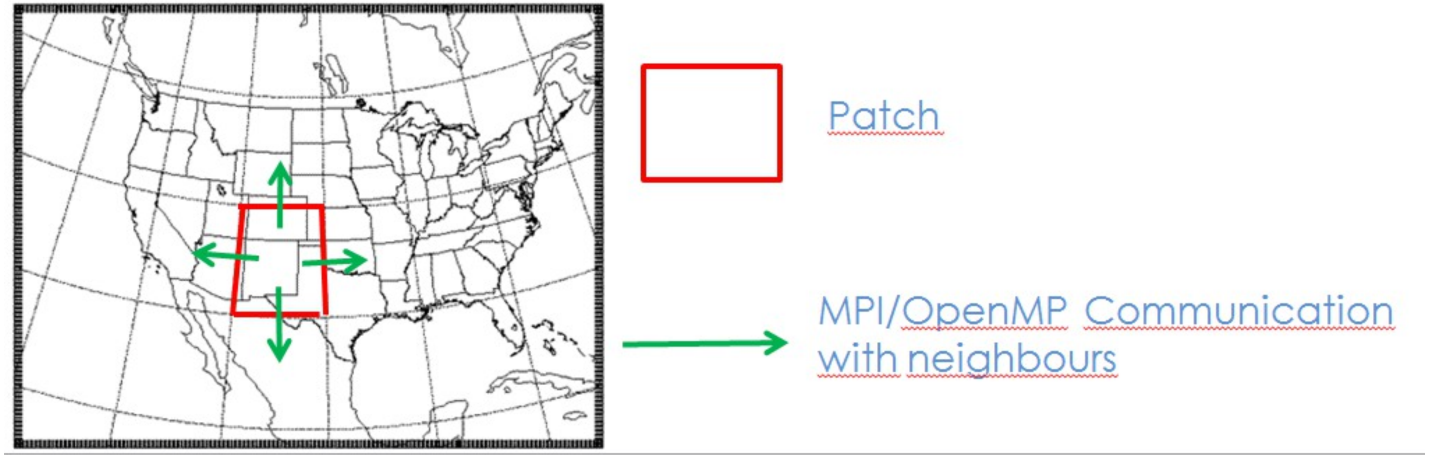
WRF Namelist

- The file namelist.input contains a lot of necessary information for the execution of the model such as start/end dates, information about resolution etc.

- We study the variables that can influence the performance of WRF

Decomposition (nproc_x, nproc_y)

- By default WRF will use the square root of the processors for values in nproc_x and nproc_y. If it is not possible, it will use some values that are close to each other.

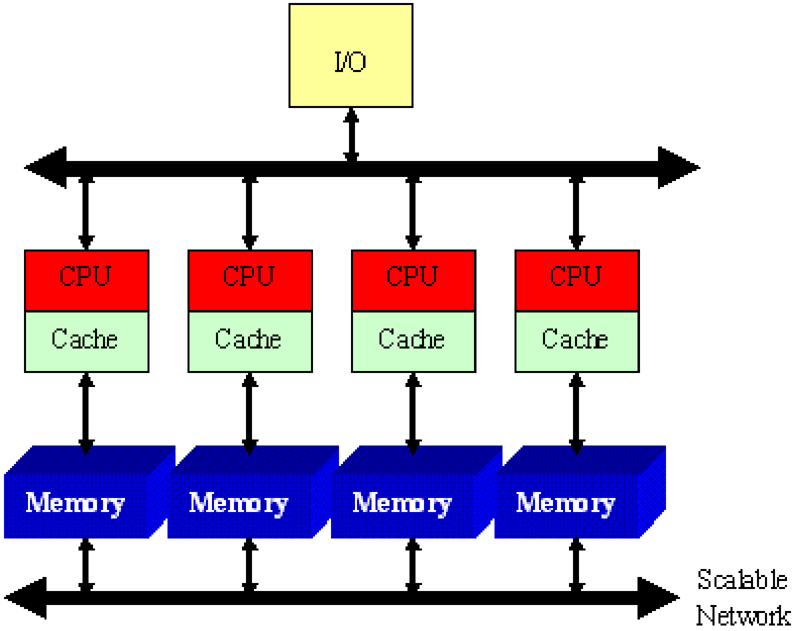


Patch

MPI/OpenMP Communication with neighbours

Decomposition (nproc_x, nproc_y)

- This is not correct as WRF responds better to a more rectangular decomposition (i.e. X<<Y). This leads to longer inner loops for better vector and register reuse, better cache blocking, and more efficient halo exchange communication pattern.

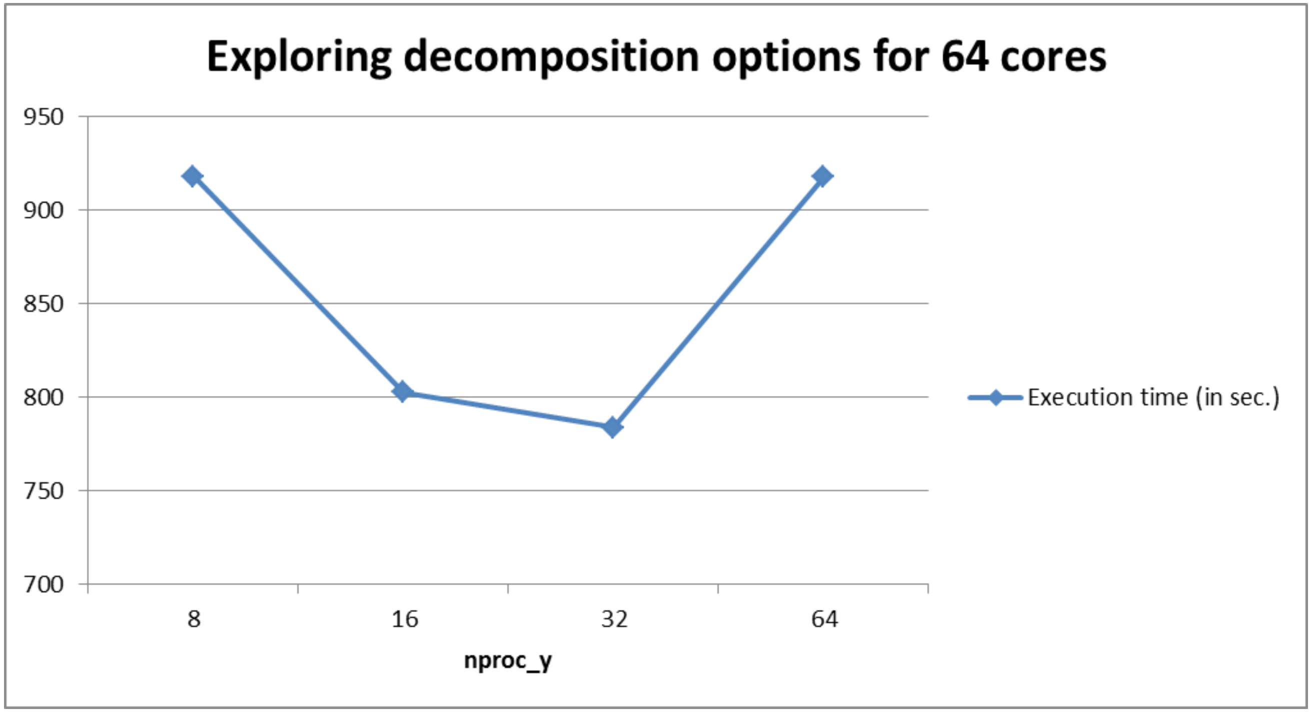


- The values depend on the software and hardware.

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Decomposition (nproc_x, nproc_y)

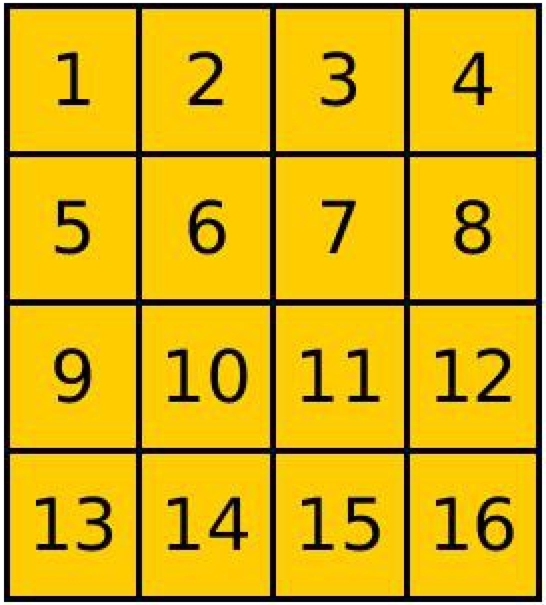
- The values can be computed only by trial and error



- The default value of nproc_y when we use 64 cores (+4 cores for I/O) is equal to 8. By changing this value to 32 (nproc_x = 2), the execution time was decreased by 14.58%.

Patch and tile

- In this example there is one patch on a processor which is split to 16 tiles.

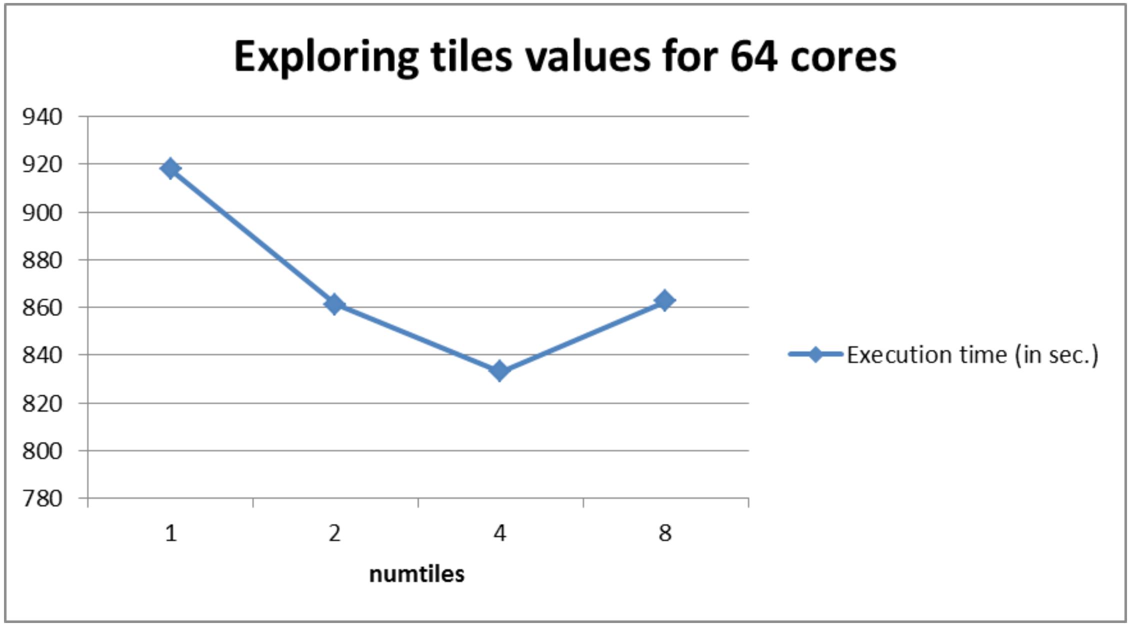


1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

- The use of tiling has greatest effect on lower processors when the patches do not fit into cache.

Patch and tile (numtile)

- Again we have to follow the trial and error approach.



- When we declare 4 tiles the execution time is decreased by 9.26%.

Processor Affinity

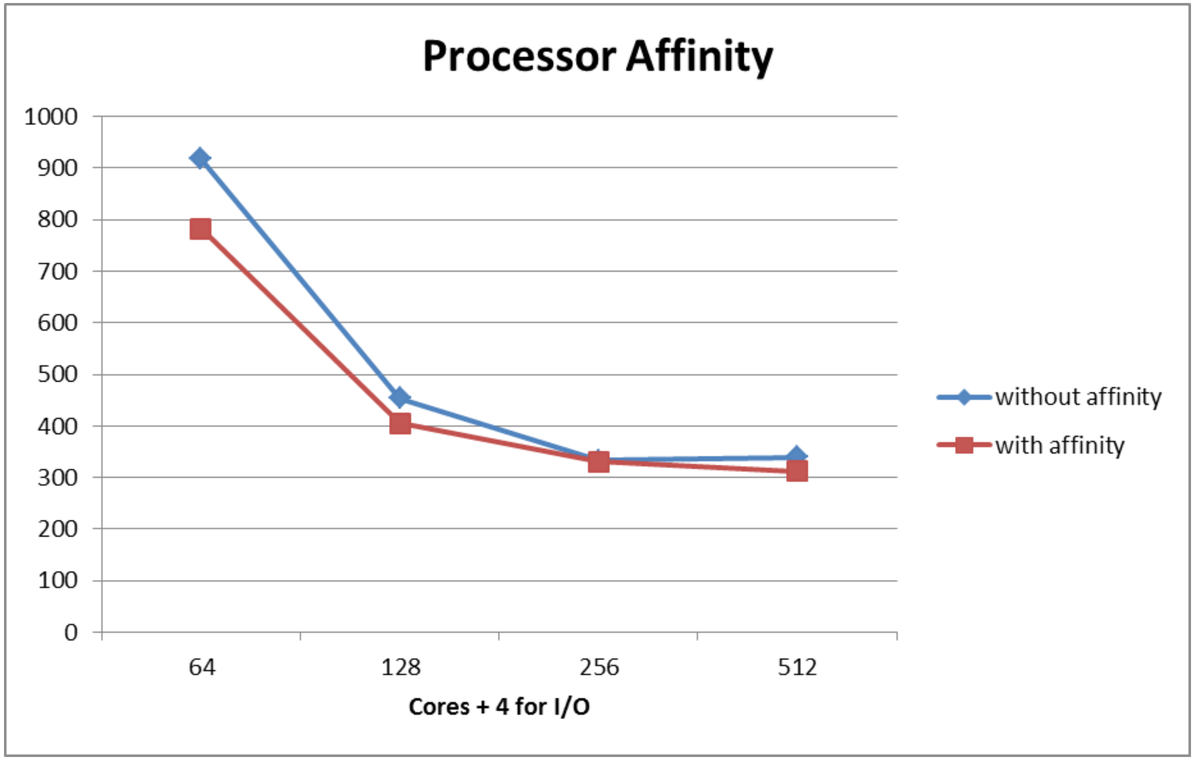
- Processor affinity is the procedure where we declare a mapping between a process and hardware (processor, core etc.).

- In order to apply processor affinity (per core) add the option “--bind-to-core”. This procedure maps each process on each core and does not allow any process migration.

- This is important because if a process migrates then it can not find the previous saved data in the new cache memory.

Processor Affinity

- The execution time is decreased till 14.8% for 64 cores.



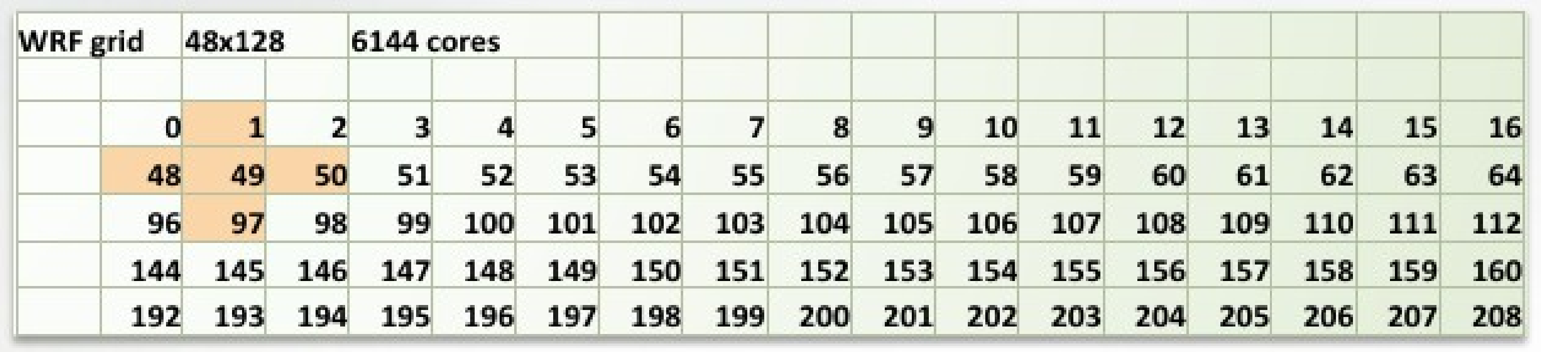
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Rank Ordering

- By rank reordering we can achieve less communication
- Some times it can be a difficult task, it depends on the hardware configuration
- Some people achieved +18% better performance by rank reordering
- Cray machines provide their own tool
- What is the currently mapping of the processes?

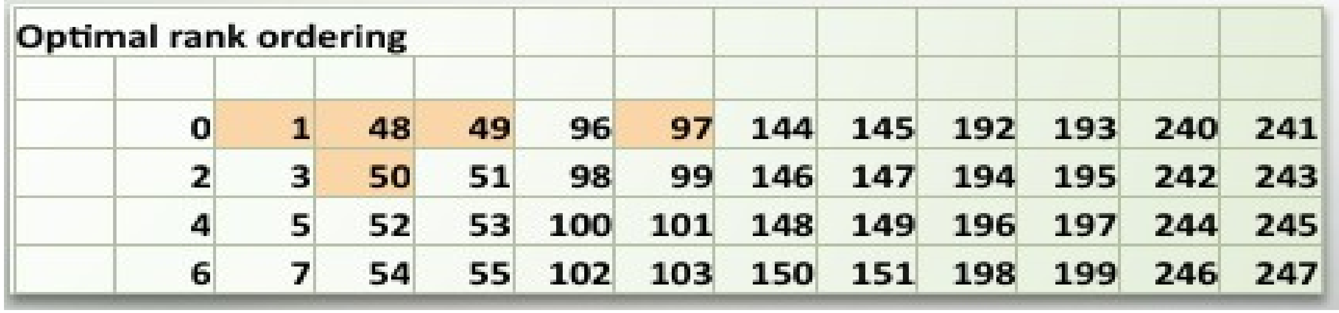
Rank Ordering

- Images from “Tuning WRF Forecasts on the Cray XT”, Peter Johnsen
- Default rank ordering



WRF grid		48x128		6144 cores													
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112
	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160
	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208

- Optimal rank ordering



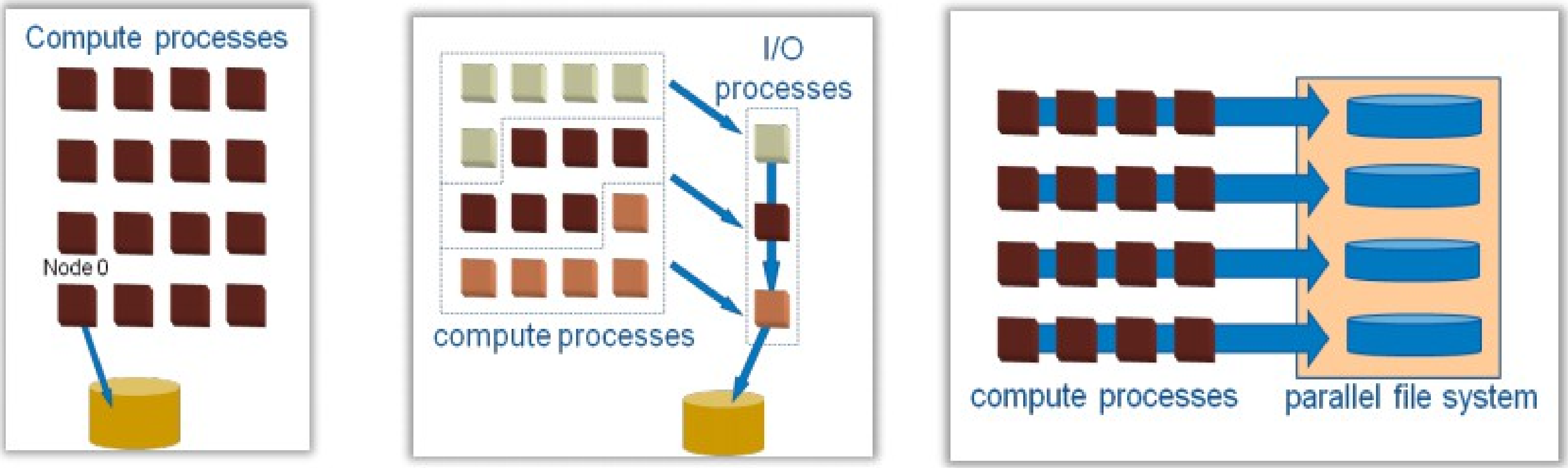
Optimal rank ordering												
	0	1	48	49	96	97	144	145	192	193	240	241
	2	3	50	51	98	99	146	147	194	195	242	243
	4	5	52	53	100	101	148	149	196	197	244	245
	6	7	54	55	102	103	150	151	198	199	246	247

Rank Ordering Results

- The variable reorder_mesh in the namelist.input did not improve the performance.
- We implemented the exact same mapping for our platform. The results were not improved by the new mapping.
- We use at most 32 nodes, thus mapping can not be really improve the execution as more nodes are needed.

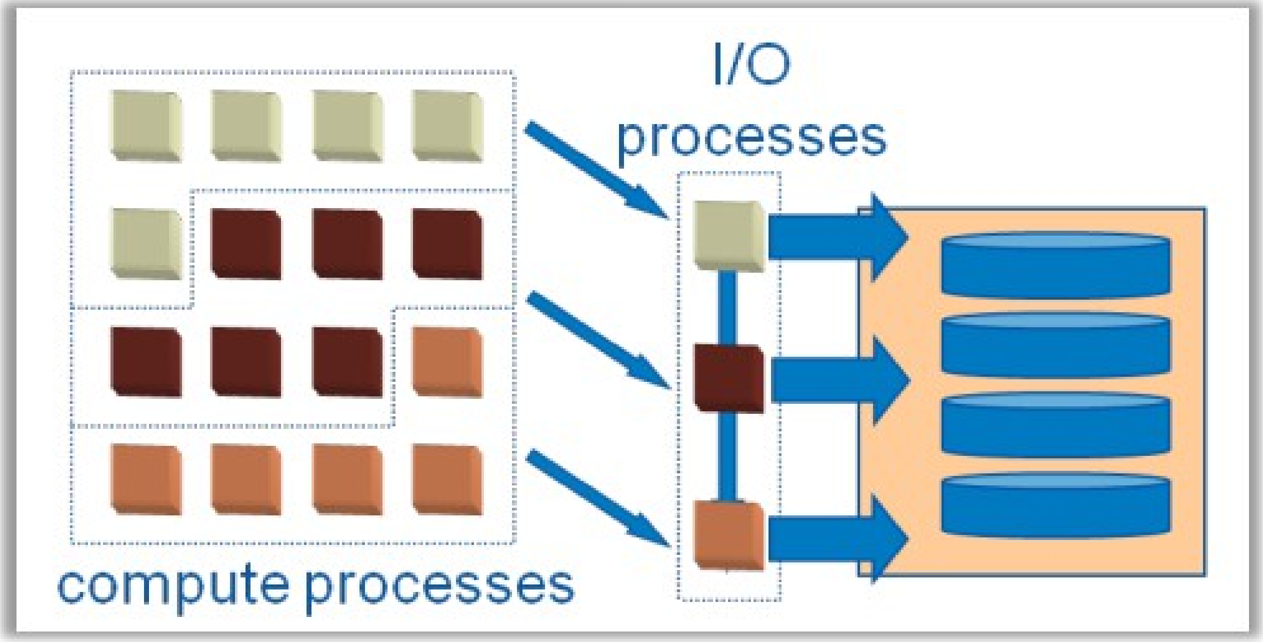
WRF I/O

- Images from “Opportunities for WRF Model Acceleration”, John Michalakes, Andrew Porter



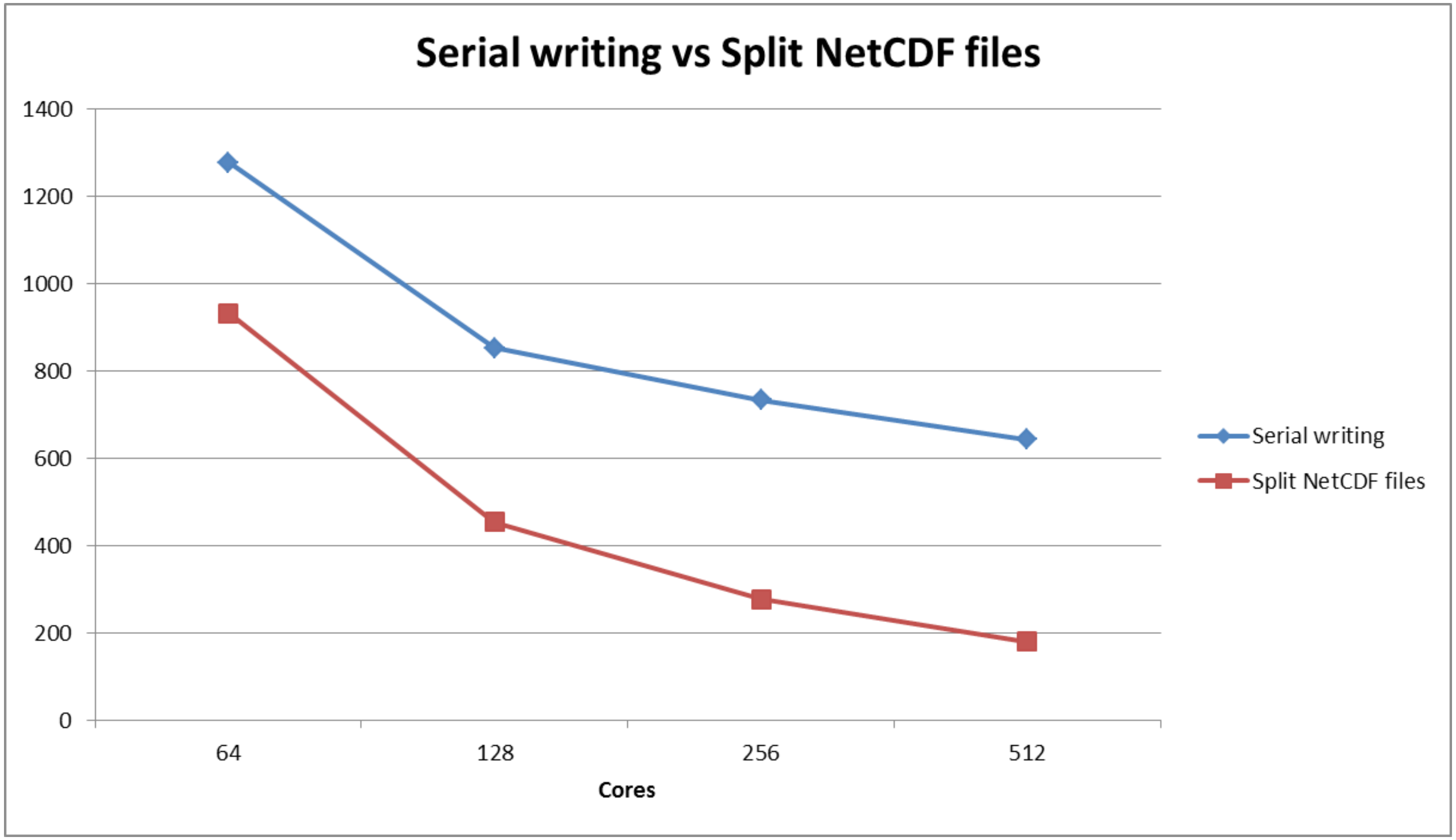
WRF I/O

- Parallel NetCDF written to single files by all MPI tasks.



Serial writing vs Split NetCDF files

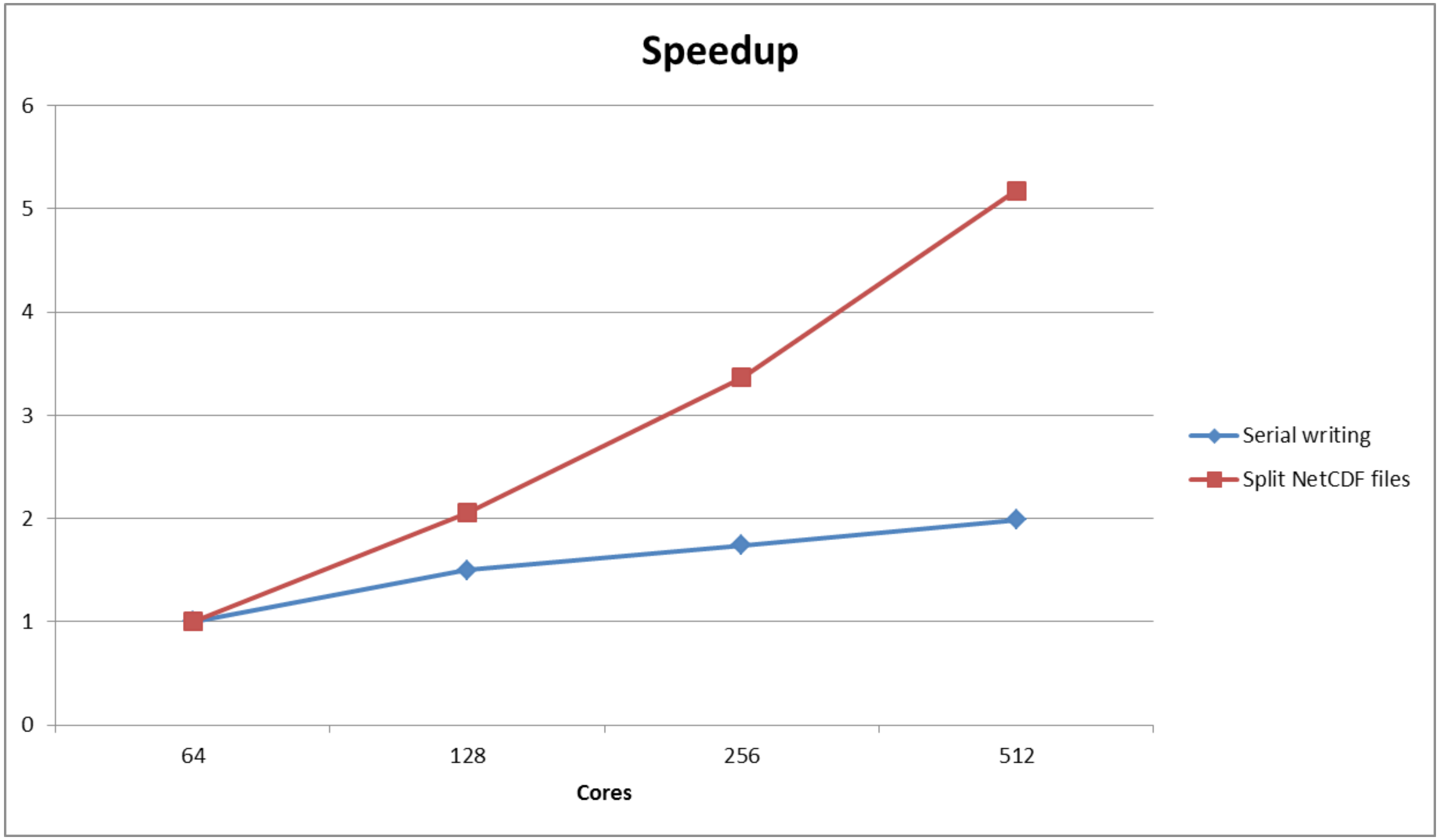
- When we split the NetCDF files, every MPI process saves its own file.
- The improvement depends on the size of the data per process. The performance improved by 2.5x times for 512 cores.
- Disadvantage: We have to merge all the files, WRF does not support it officially.



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

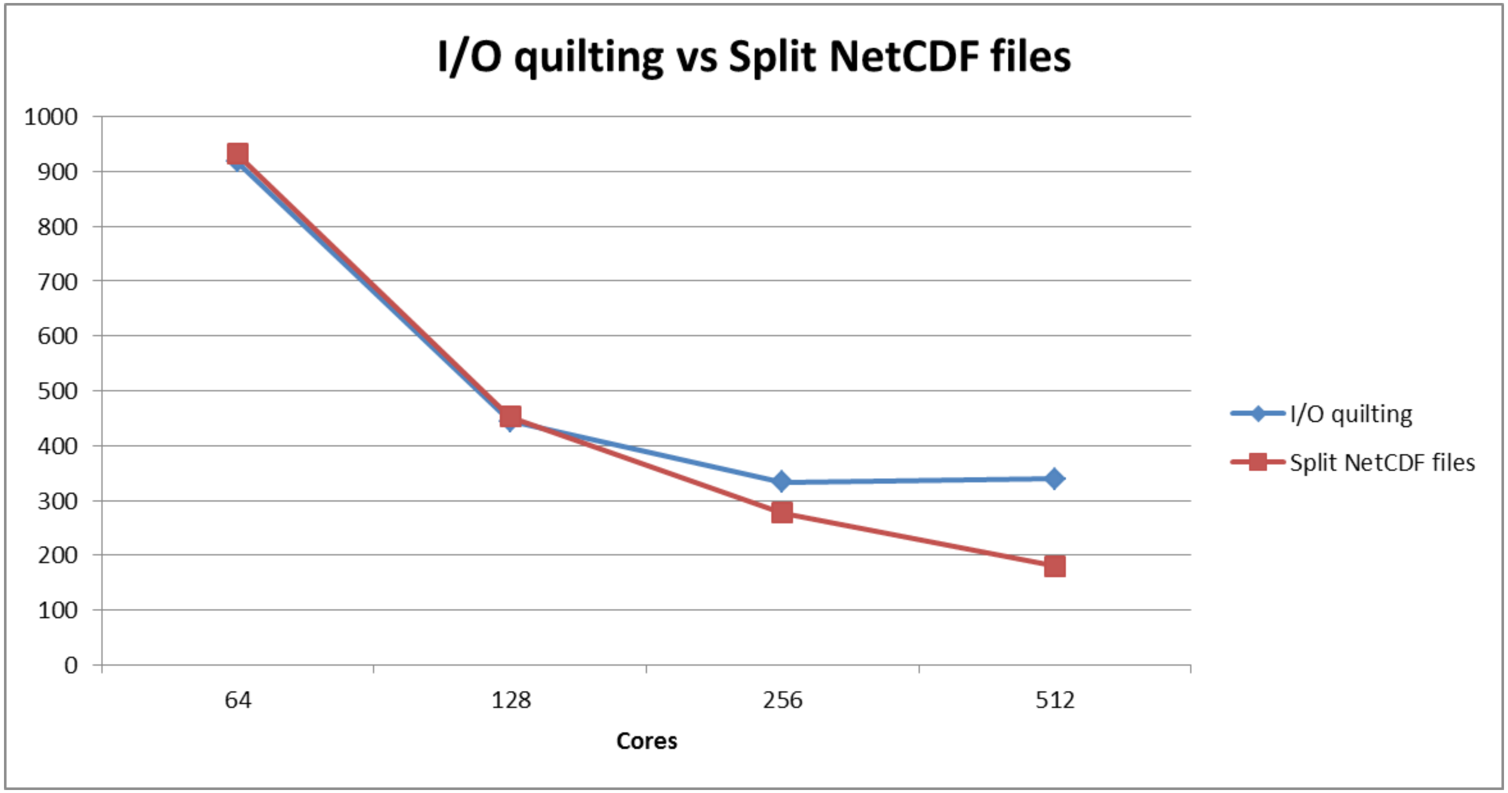
Speedup Serial writing vs Split NetCDF files

- Now the speedup is improved in comparison to the initial results, I/O is a bottleneck.
- We have a speedup of 5.1 with 512 cores (64 cores is the base) while before we had speedup of 2.71.



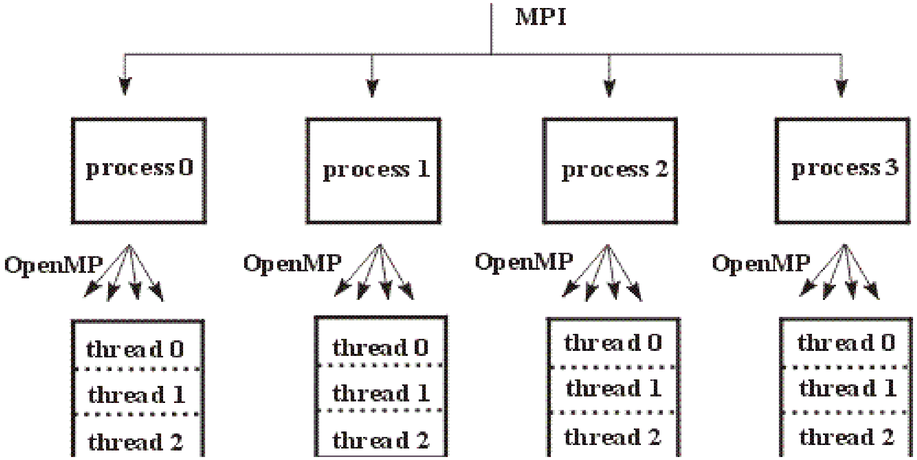
I/O Quilting vs Split NetCDF files

- One problem with splitting the NetCDF files is that the requested time to handle them, could be significant.
- Does the number of the servers play an important role for I/O quilting? It depends on the case.



MPI/OpenMP

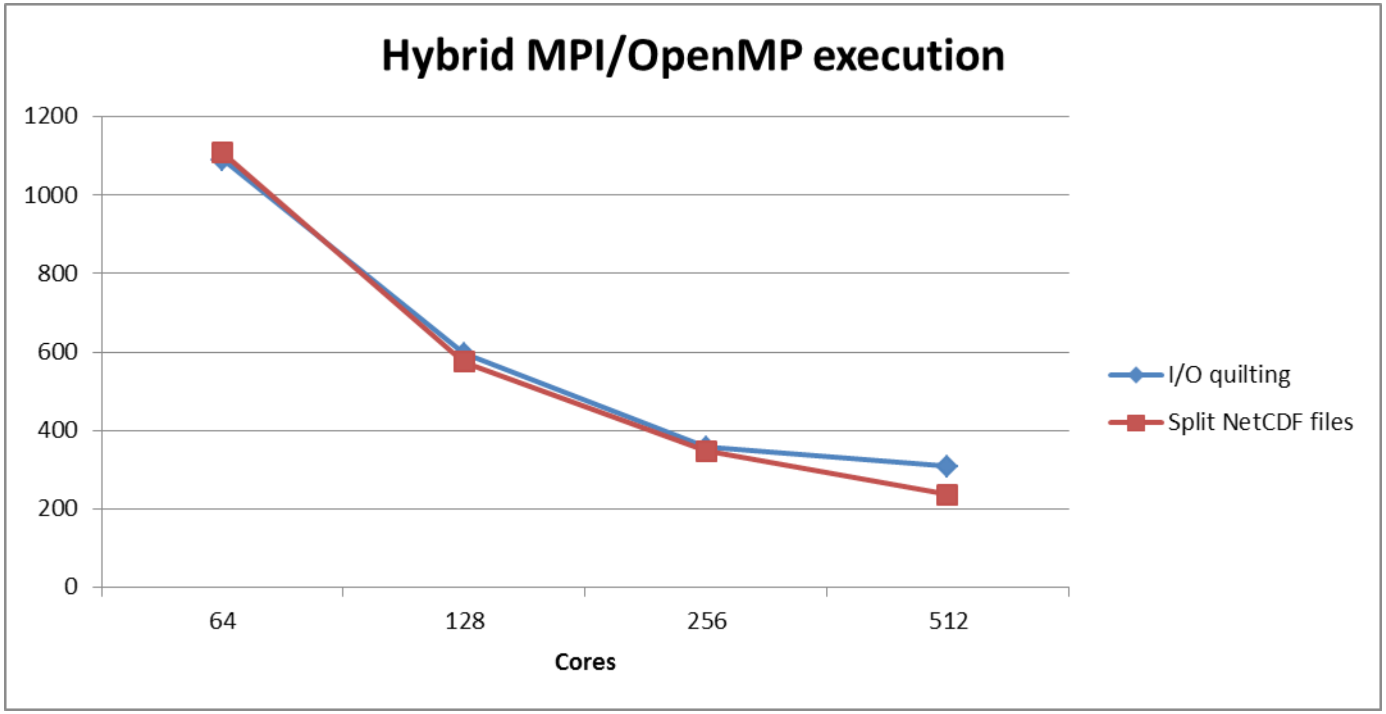
- OpenMP is an API that supports multi-platform shared memory multiprocessing programming (http://www.openmp.org)
- OpenMP is used for problems that can be solved efficient on shared memory machines
- Message Passing Interface (MPI) is used for distributed memory problems (http://www.open-mpi.org)
- MPI is used for large scale experiments across many nodes.



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

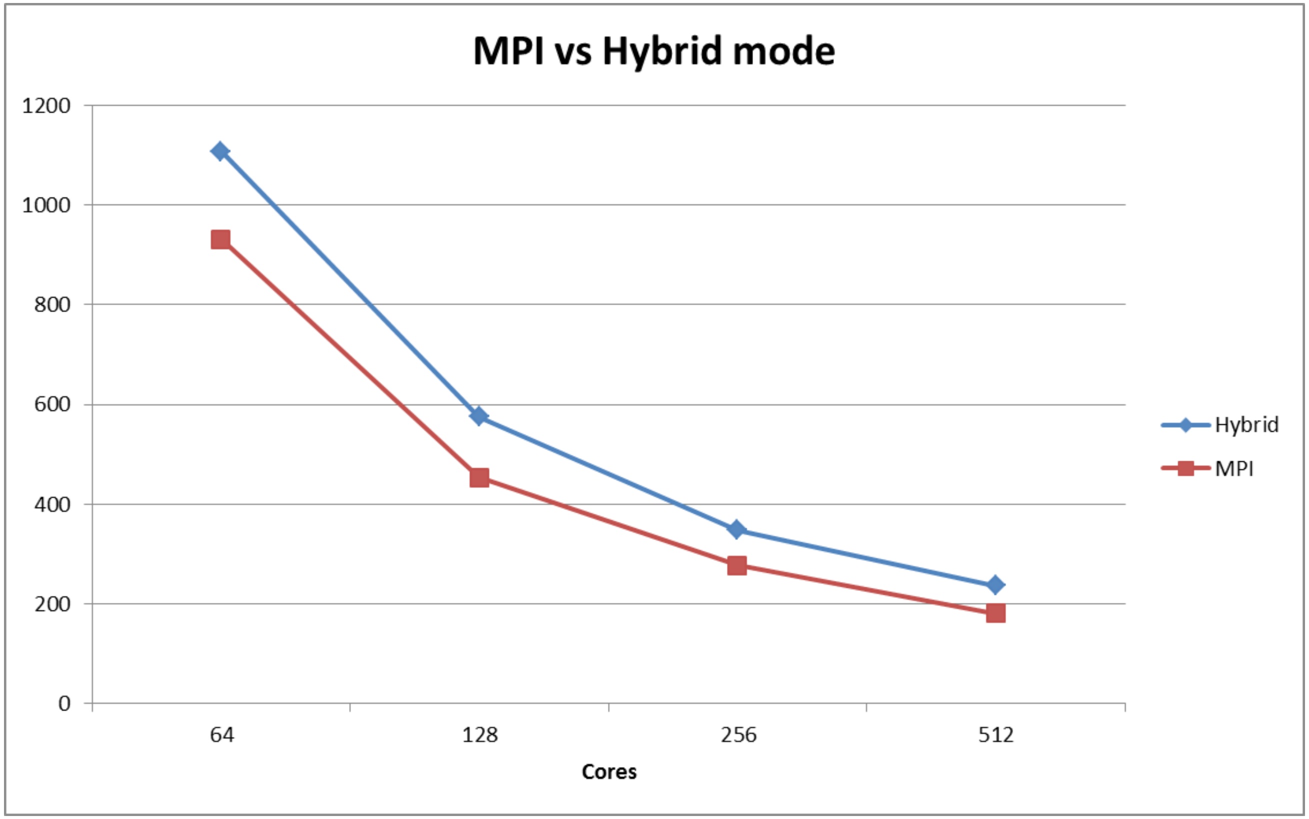
WRF Hybrid MPI/OpenMP

- We execute the hybrid version of WRF with various I/O modes.
- The I/O performance does not seem to be improved significant by splitting the NetCDF files.



WRF MPI vs MPI/OpenMP

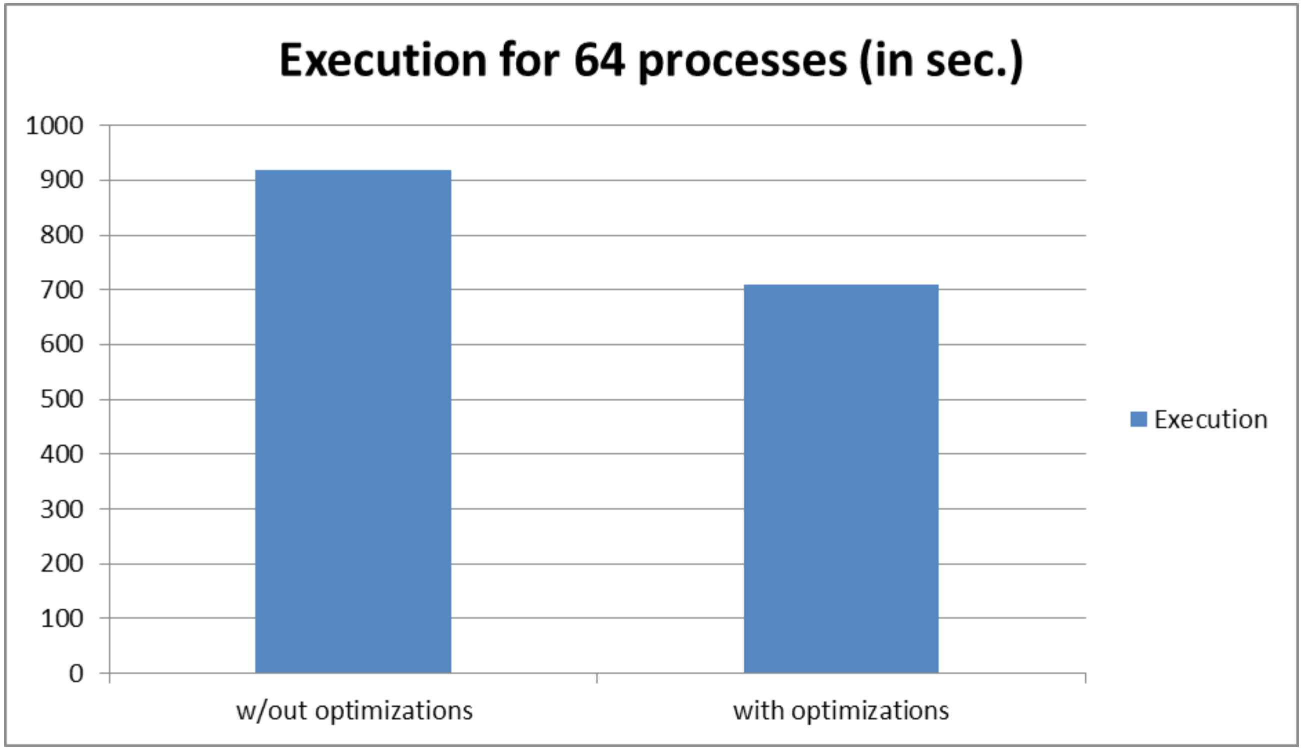
- Executing both versions with splitting NetCDF files.
- For MPI/OpenMP we use one MPI process and 16 OpenMP threads per node.
- MPI is between 15.8% and 23.75% faster than hybrid mode.



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Combine optimizations

- Could we really combine some of the previous optimizations? The answer is not clear and it depends on many factors.
- By combining processor affinity, tiles declaration, and decomposition of 2 x 32 for 64 processors, the execution time was improved by 22.8%.



Used optimizations

- Try any of the following declarations in the namelist.input file under the domains section. This work was done with WRF v3.5.1. The following values apply for a specific example with 64 processes, change them as you wish.

Declarations for the namelist.input (domains section)

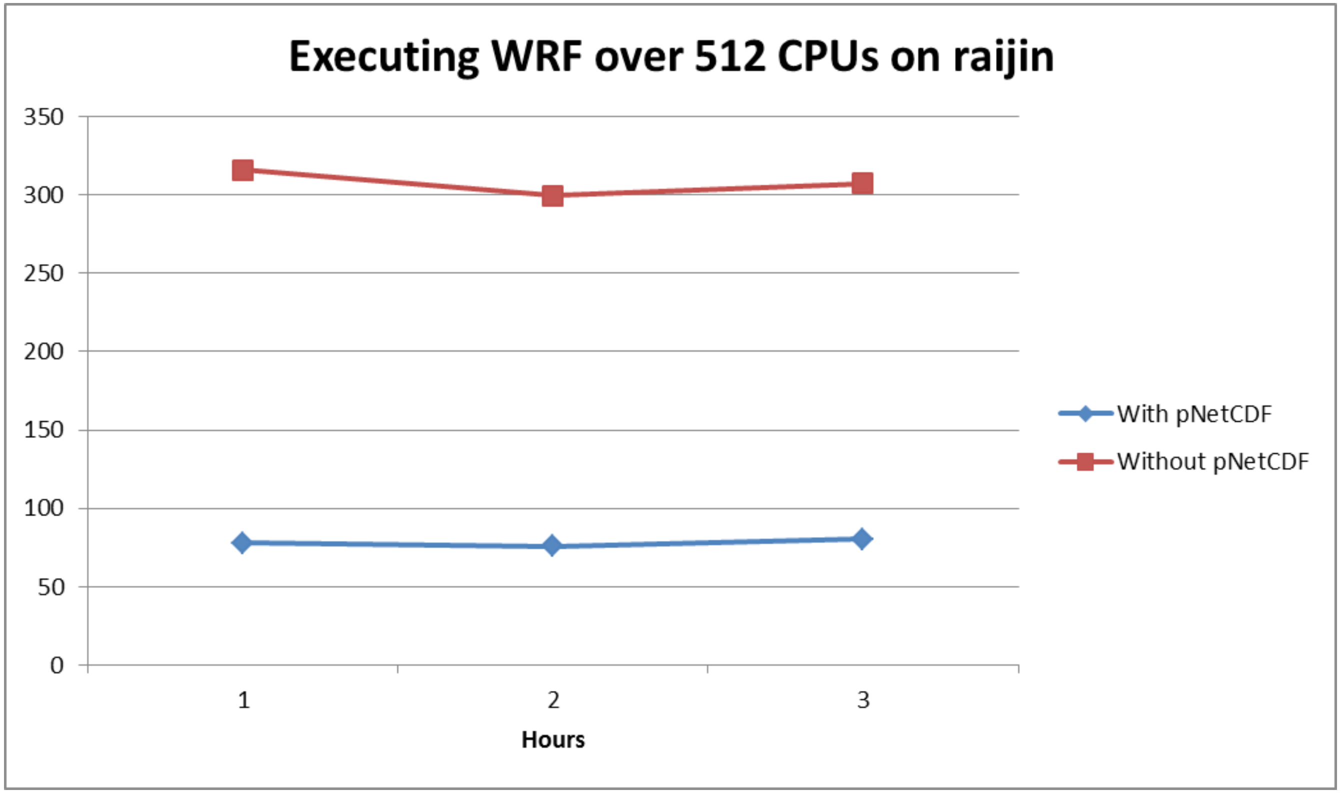
```
nproc_x = 2
nproc_y = 32

numtiles = 4

reorder_mesh = .true.
```

Used optimizations II

- Processor affinity: add the flag --bind-to-core after the mpirun call.
- Split NetCDF files: declare io_form_history = 102 (time_control section)
- Parallel NetCDF: io_form_history = 11 (not working currently). Do not forget to declare nocolons = .true. in the namelist files of WPS and WRF.
- Disable I/O quilting: declare nio_tasks_per_group = 0 (namelist_quilt section)
- Add more I/O servers: increase nio_groups (namelist_quilt section)

Parallel NetCDF

- Results from ARC centre of excellence for Climate System Science.
- Parallel NetCDF for this case improves the performance by 3.8 to 4 times.



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

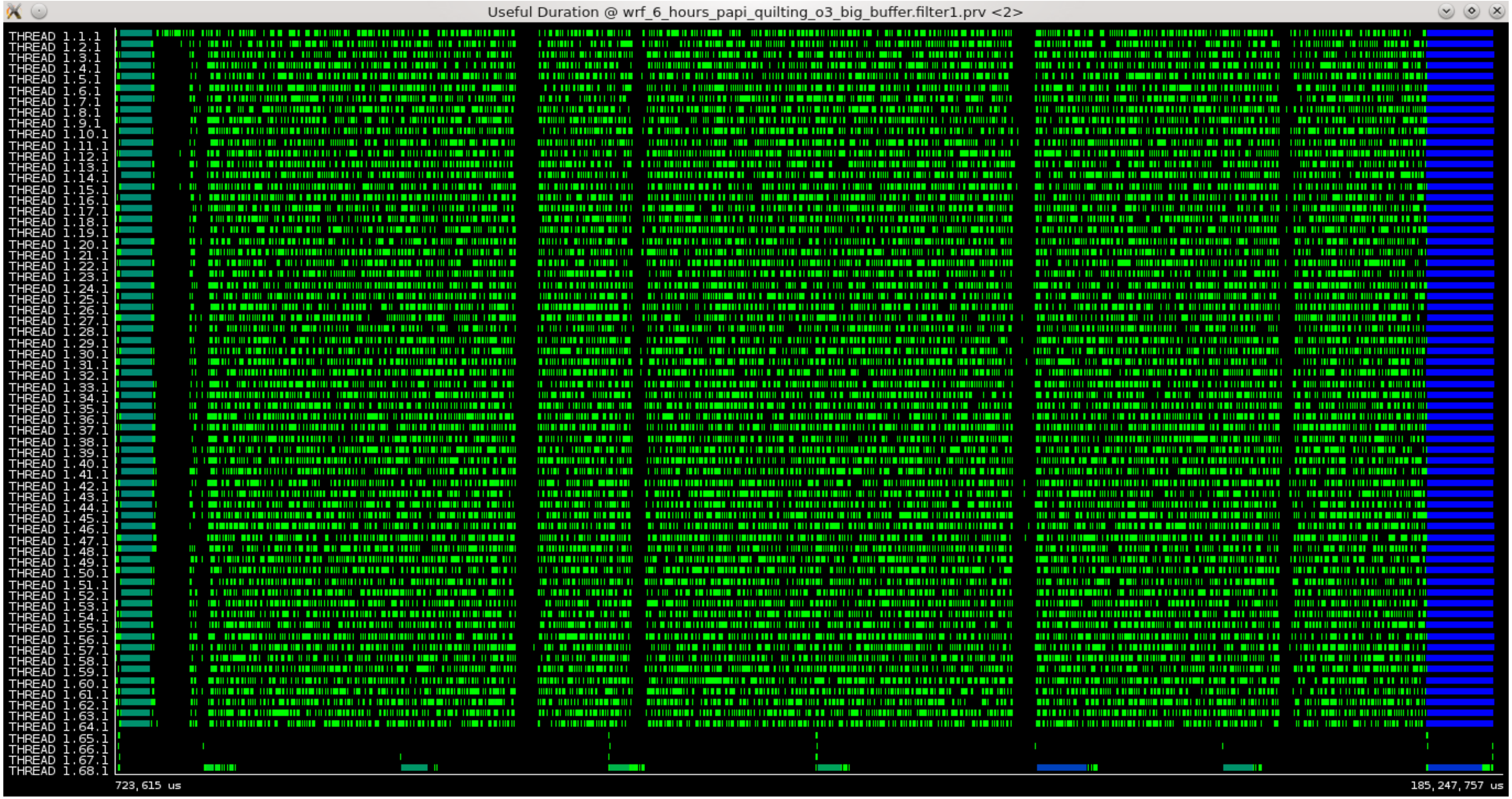
Performance Analysis with Paraver

Paraver

- An application to analyze traces
- Discover bottlenecks
- Possible to do visual and statistical analysis of the various events
- Customizable semantics of the visualized information
- Provides views
- Information: http://www.bsc.es/paraver/

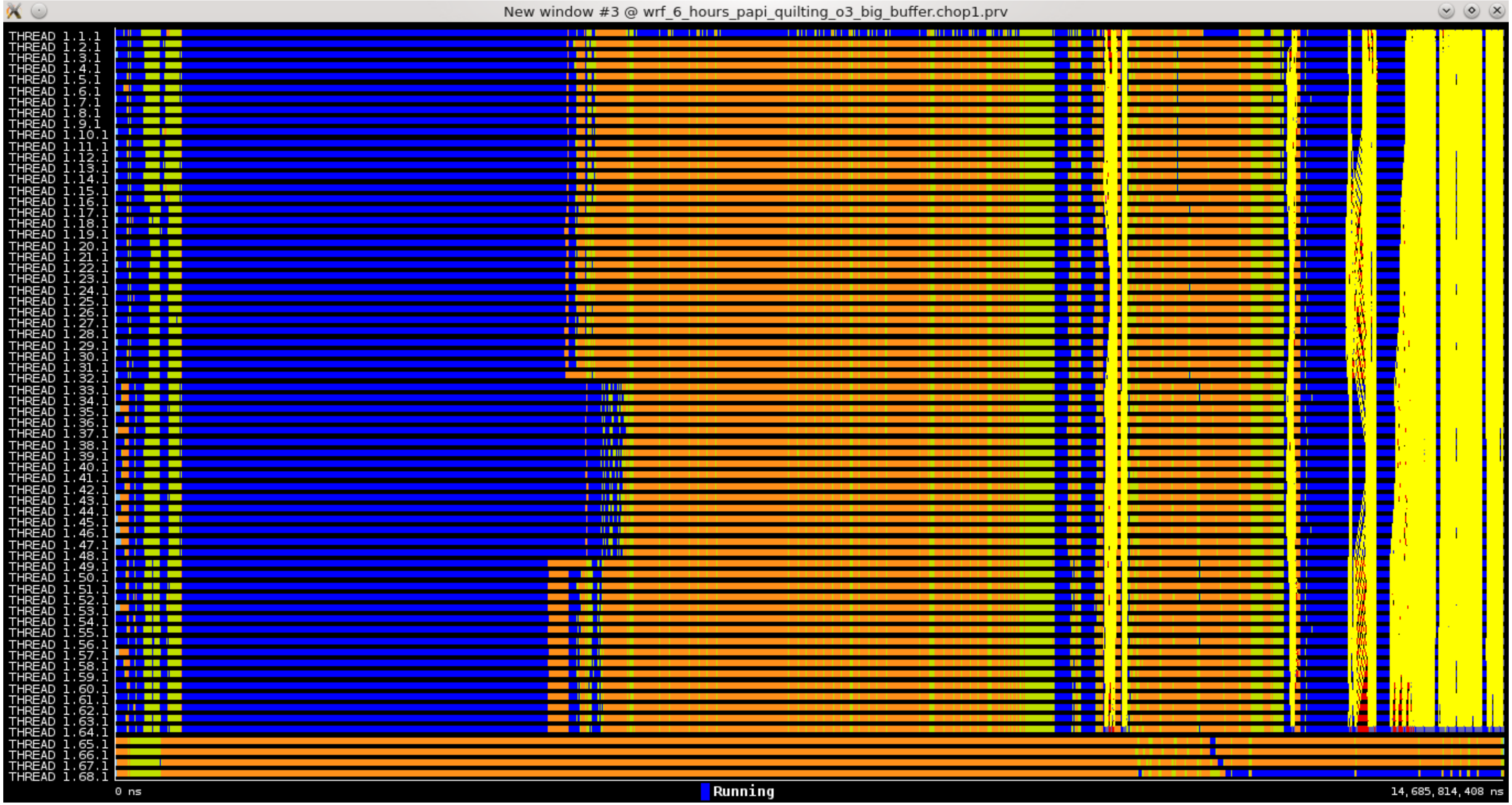
Visualizing the computation of a whole trace

- Visualizing computation duration of 6 hours simulation (5.2GB initial trace, 68 cores)



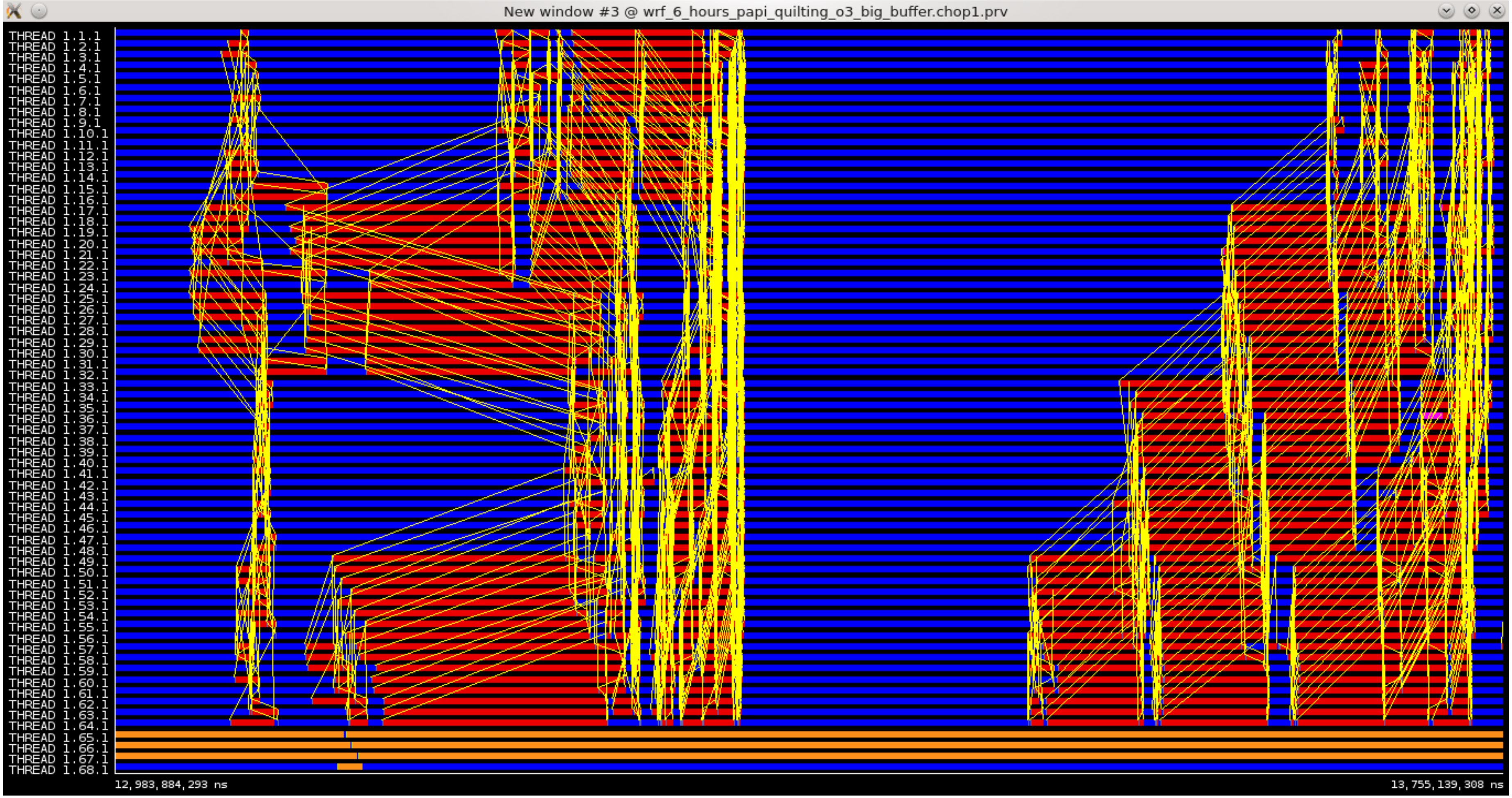
Trace Analysis – Beginning of the trace

- Beginning of the trace, 4 cores for I/O quilting
- Blue colour is running part, no communication, yellow colour is message transfer (send/recv etc.)


New window #3 @ wrf_6_hours_papi_quilting_o3_big_buffer.chop1.prv
THREAD 1.1.1 … THREAD 1.68.1
0 ns
Running
14,685,814,408 ns

Trace Analysis – Beginning of the trace

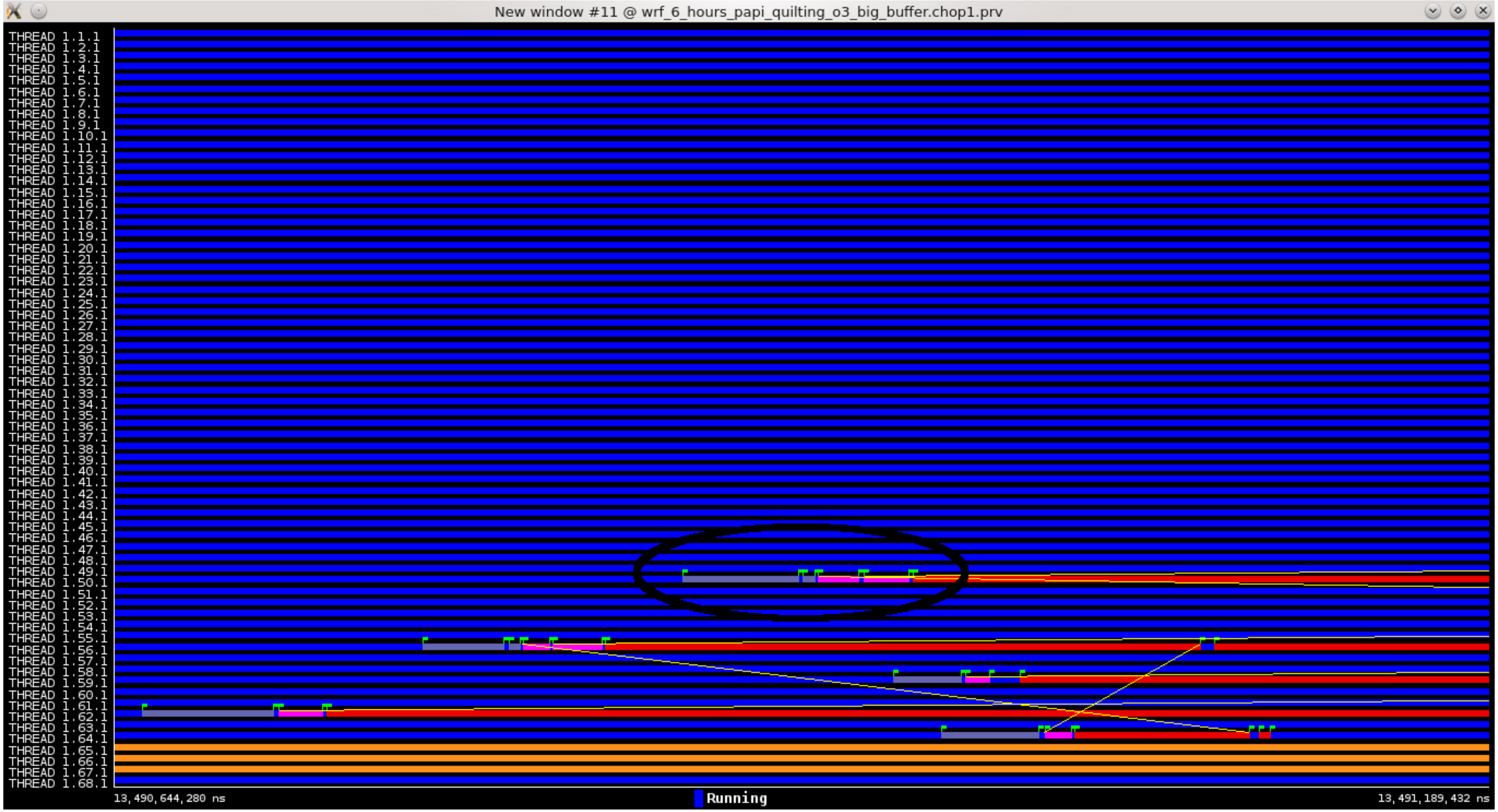
- Let's zoom a bit before the end of the previous visualization
- There are some long MPI_Wait calls (red colour)



New window #3 @ wrf_6_hours_papi_quilting_o3_big_buffer.chop1.prv

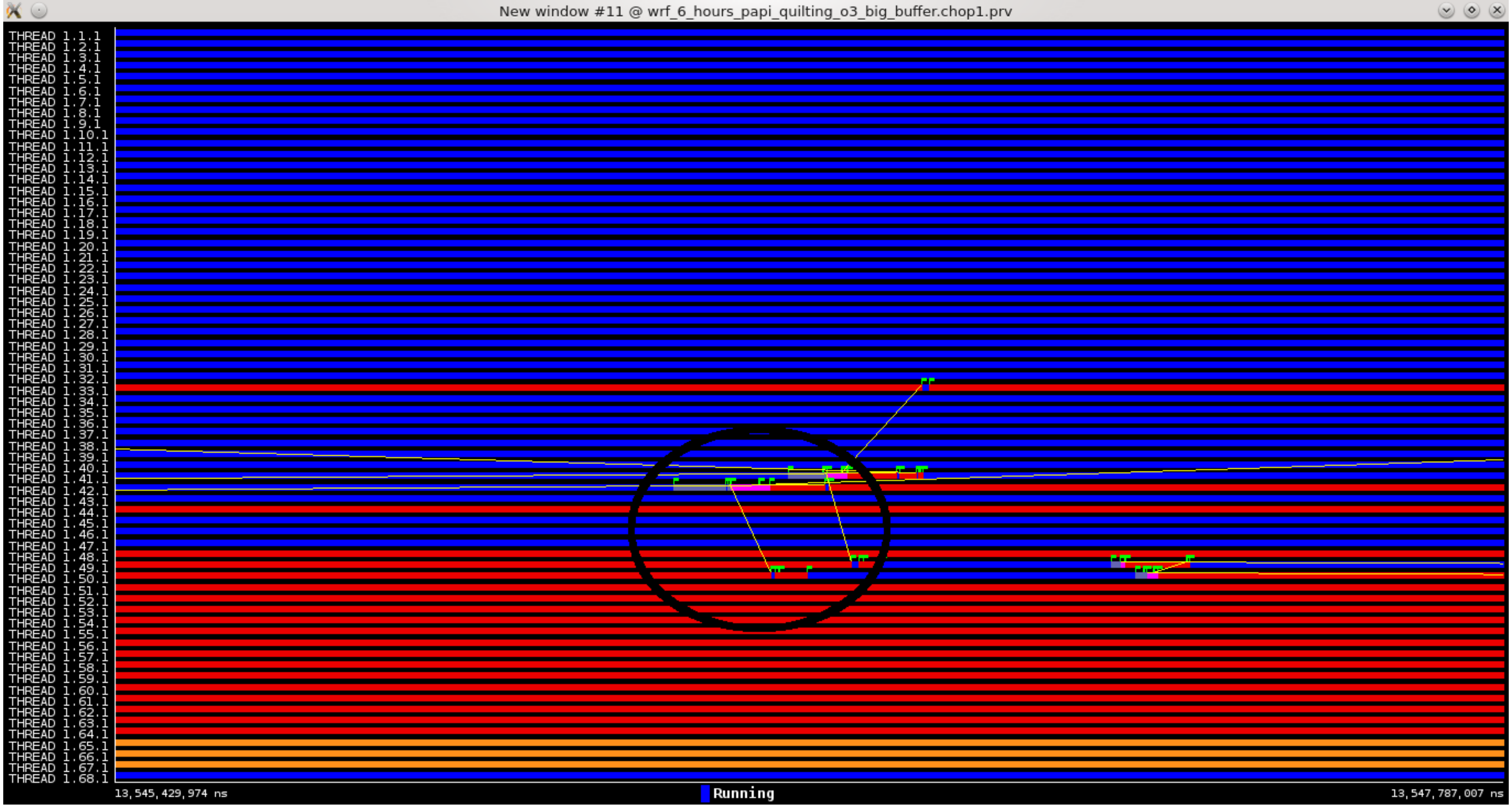
Trace Analysis – Beginning of the trace

- We zoom at the beginning of the second half of the previous plot and we focus on rank 50
- There are two MPI_Irecv and MPI_Isend calls before the MPI_Wait call



Trace Analysis – Beginning of the trace

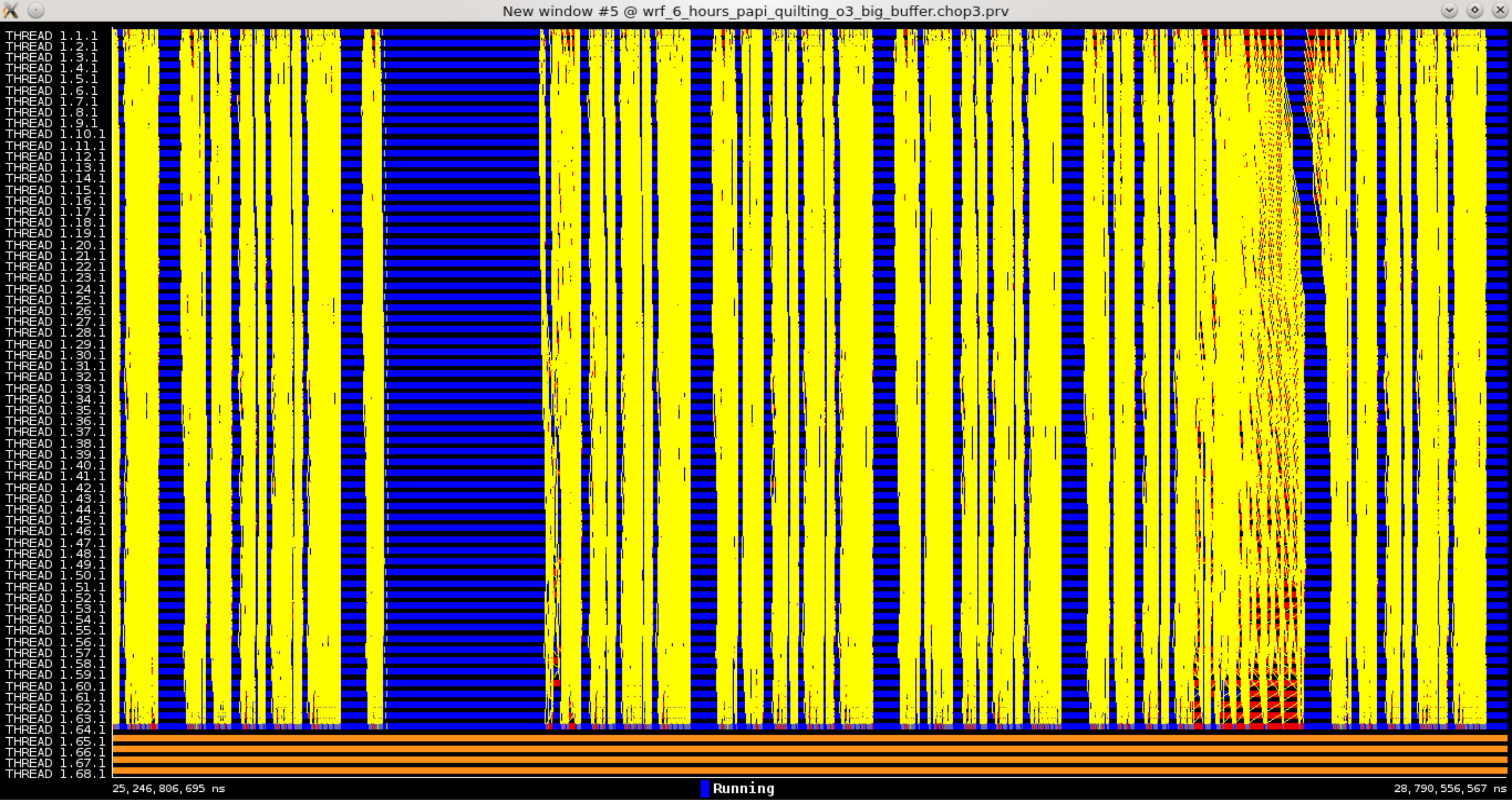
- The corresponding MPI_Isend for the previous MPI_Irecv is called too late
- Possible solution move MPI_Wait of rank 50 after some computation phases



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

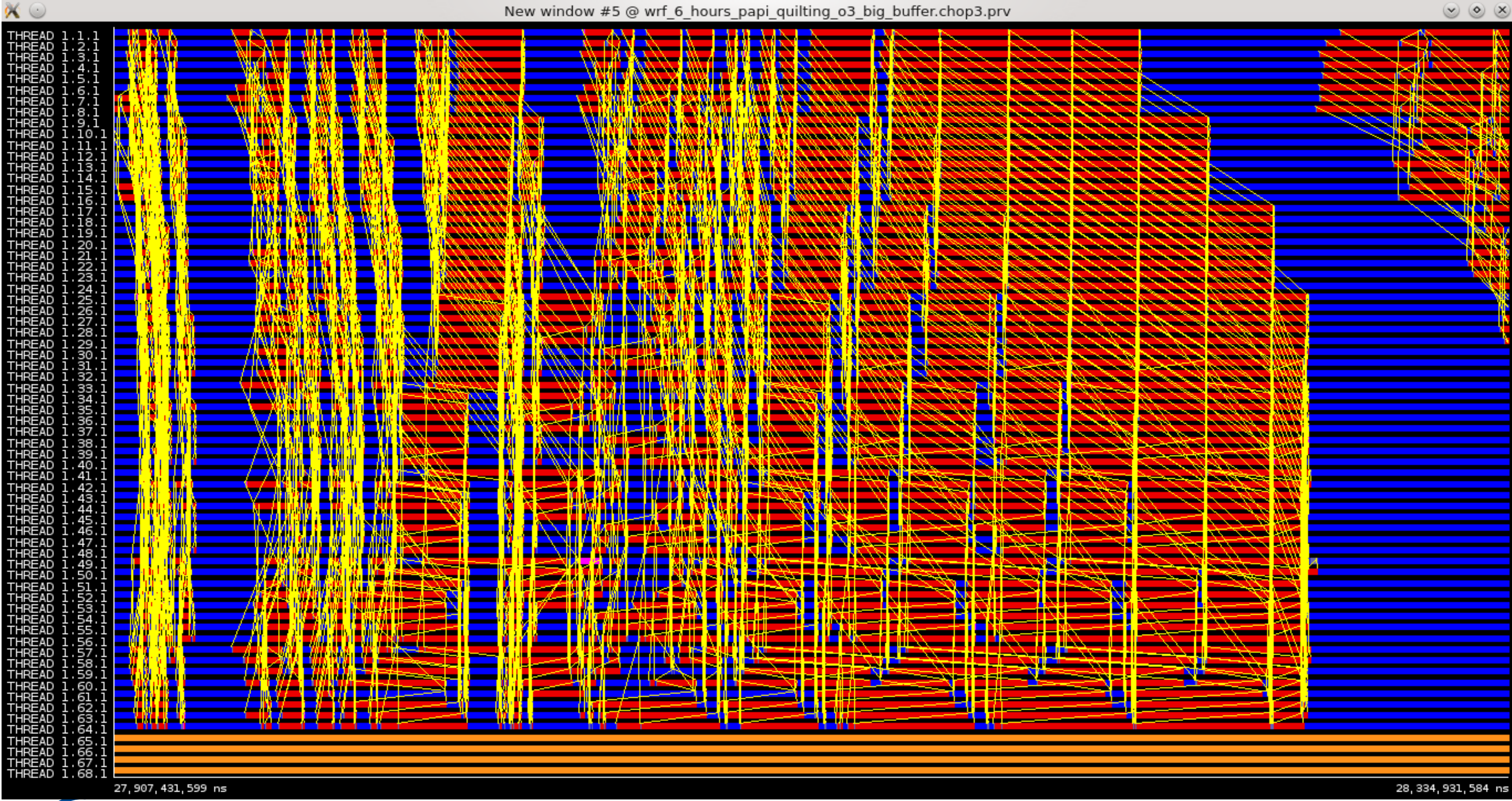
Trace Analysis

- We can observe some communications at the right that behave different than the rest ones



Trace Analysis

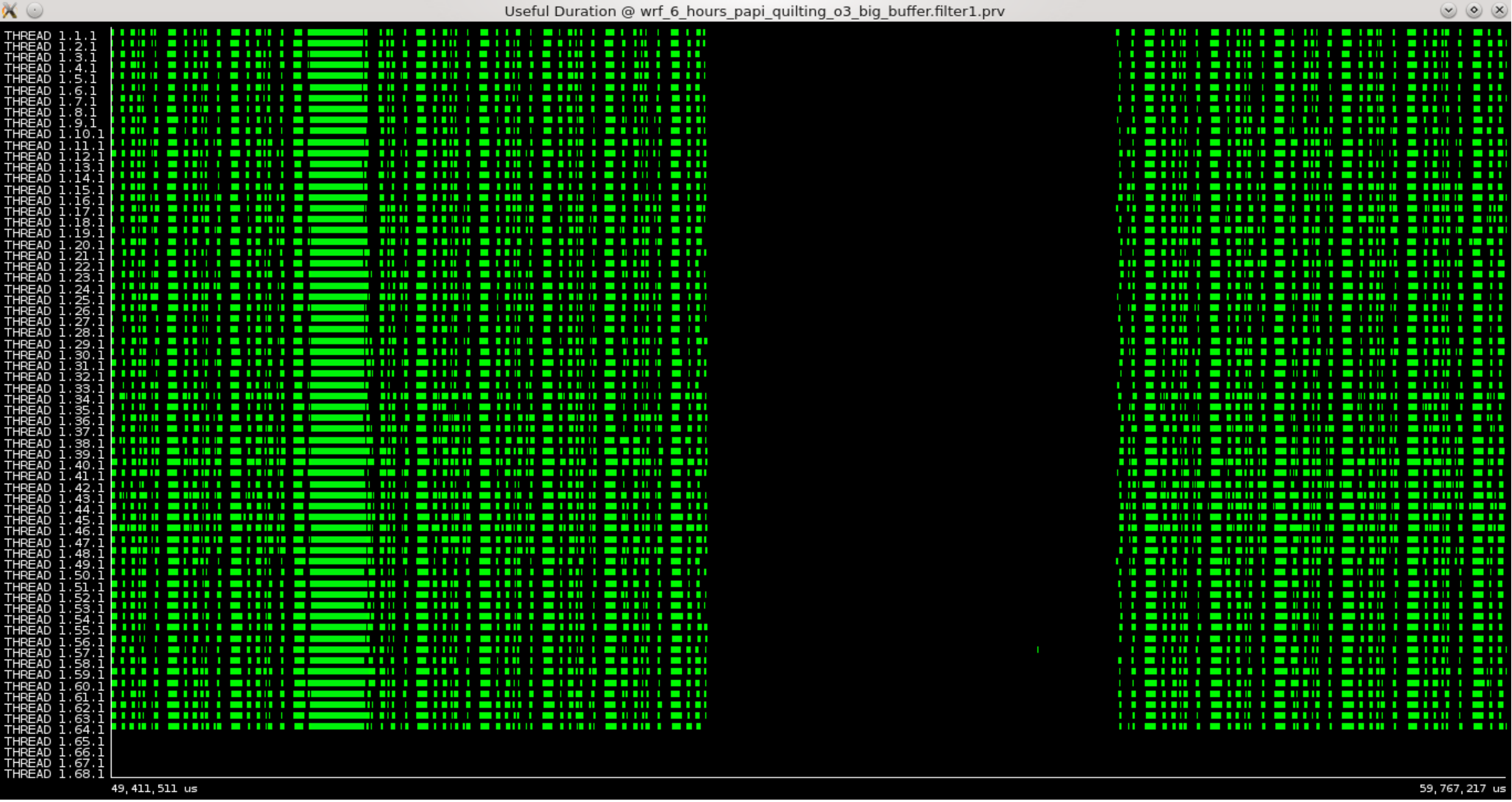
- If we zoom, we have the following
- Similar problems with some MPI Wait calls



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Trace Analysis

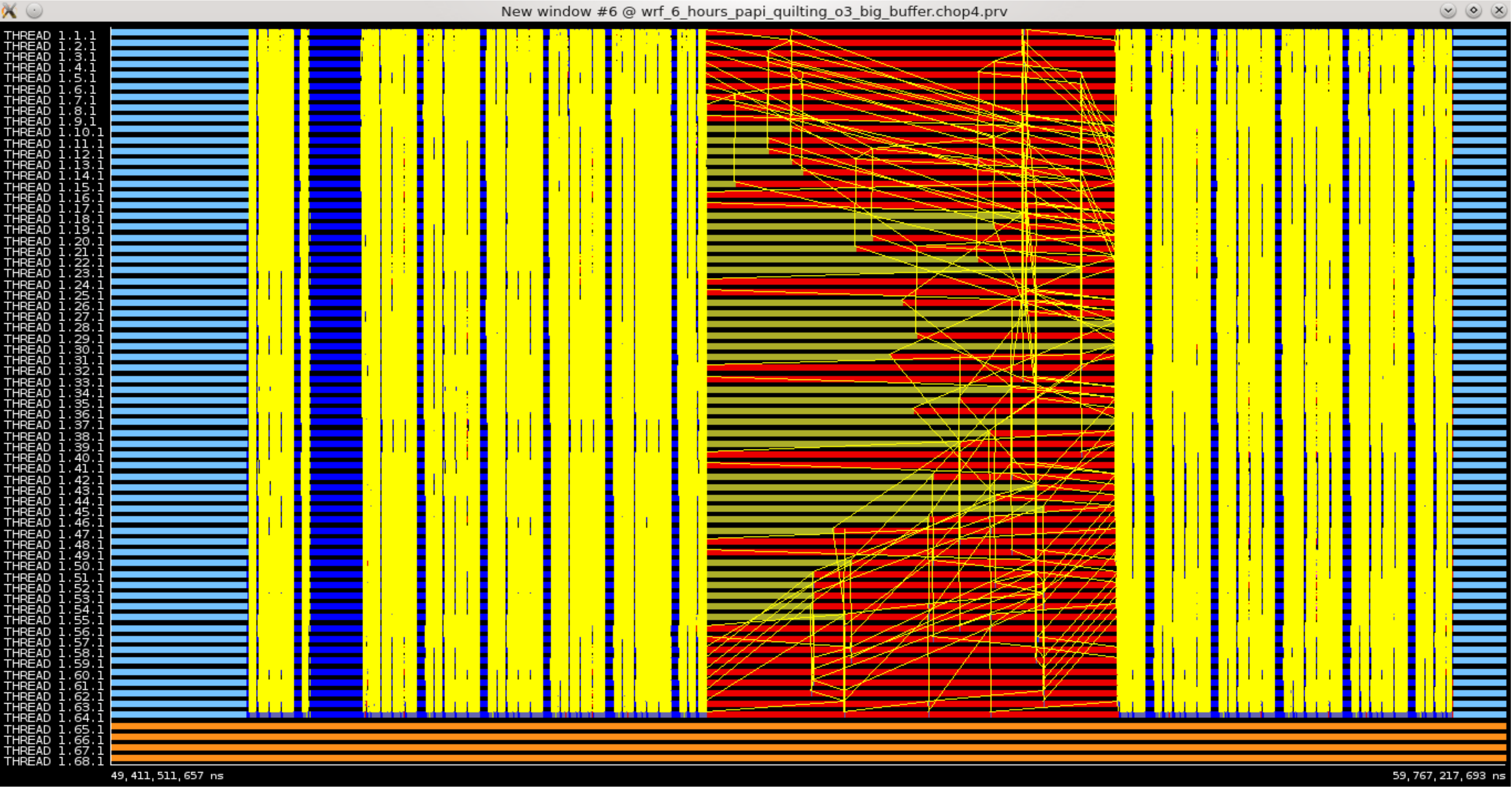
- During the visualization of the computation areas we can see a large black area



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Trace Analysis

- The previous black area is caused by communication perturbation
- The brown area is the I/O caused from the flushing of the traces on the hard disk

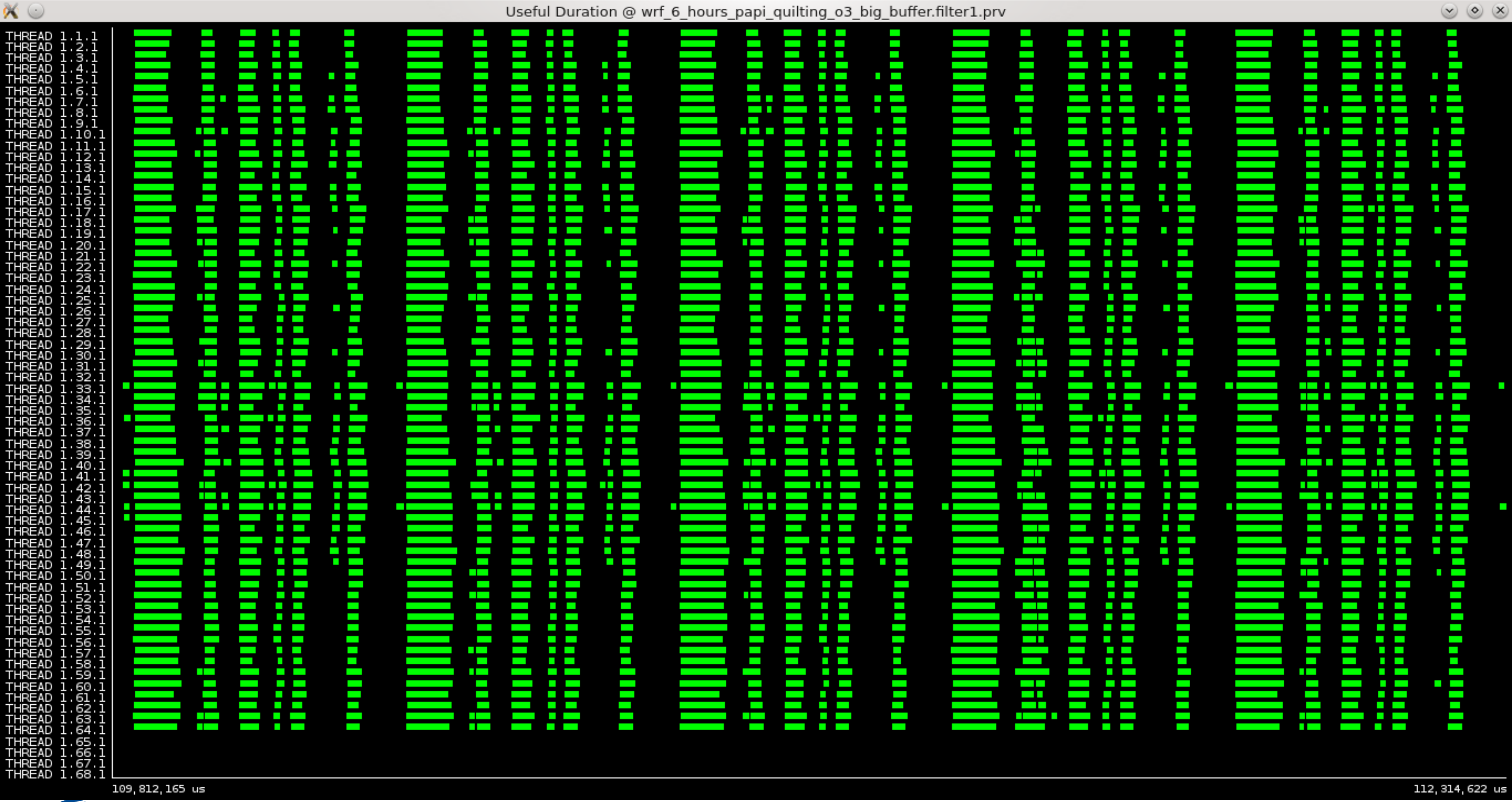


New window #6 @ wrf_6_hours_papi_quilting_o3_big_buffer.chop4.prv

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Trace Analysis

- Observing the patterns from the computation phases is a good approach to know where we should focus (we have 5 similar phases)

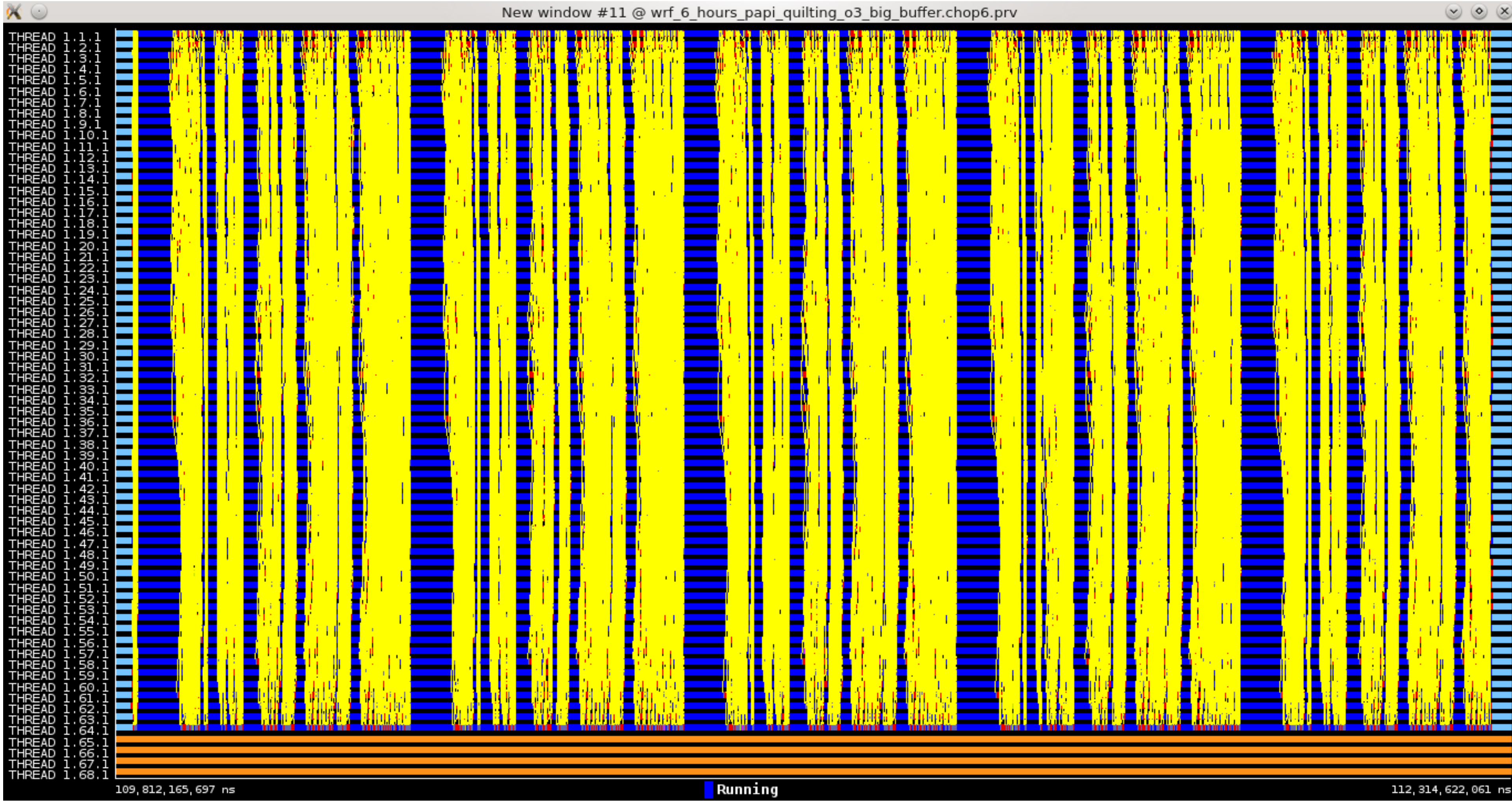


Useful Duration @ wrf_6_hours_papi_quilting_o3_big_buffer.filter1.prv

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

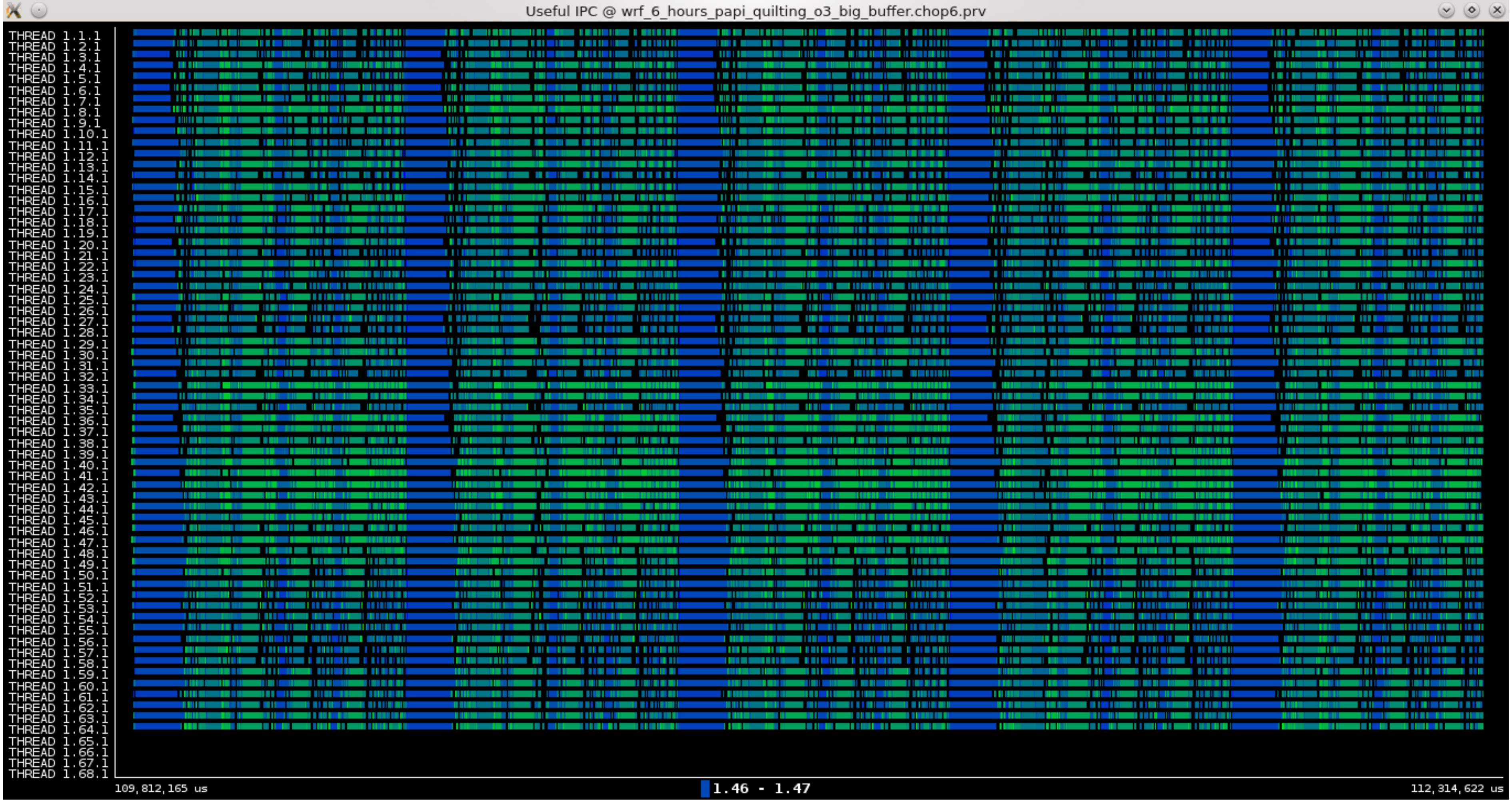
Trace Analysis

- The previous visualization with the communications and any extra metric



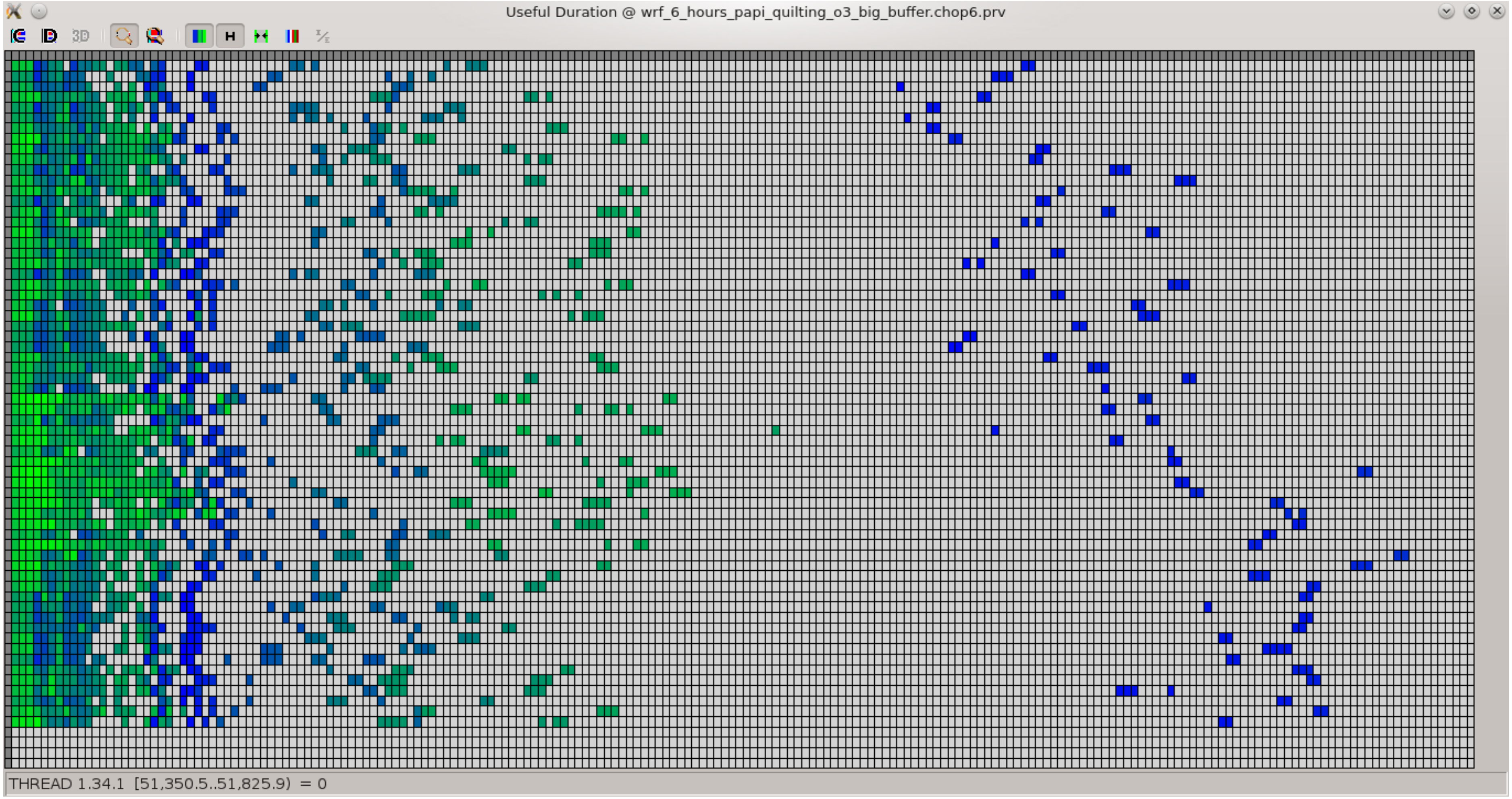
Trace Analysis

- Useful instructions per cycle. A value close to 2 is good. Much lower value means that the code should be improved



Trace Analysis

Useful duration per process

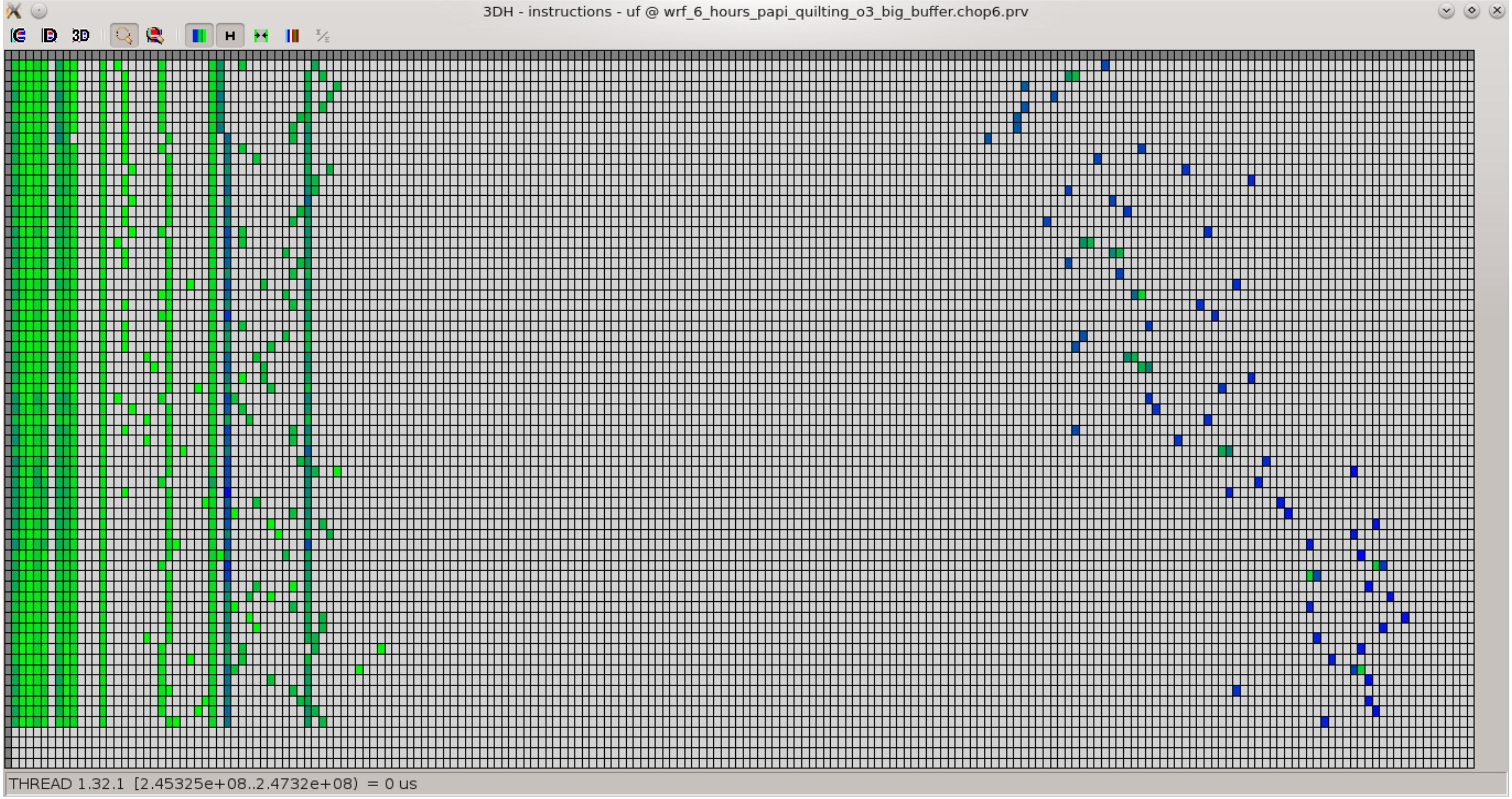


Useful Duration @ wrf_6_hours_papi_quilting_o3_big_buffer.chop6.prv

THREAD 1.34.1 [51,350.5..51,825.9) = 0

Trace Analysis

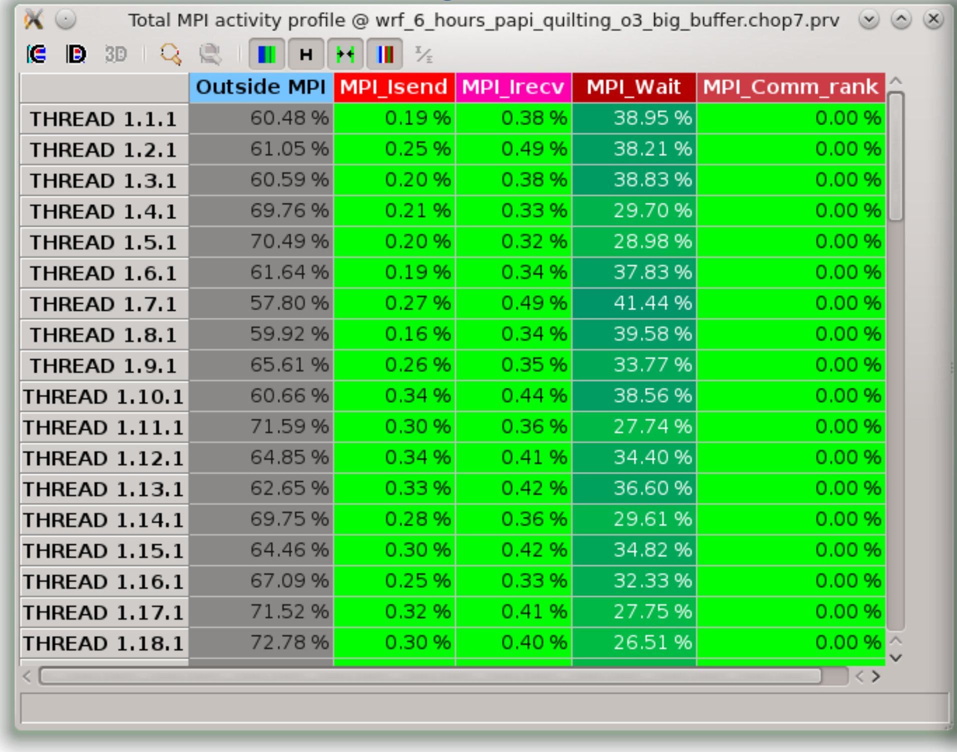
- Instructions per process
- In general we should have vertical lines


3DH - instructions - uf @ wrf_6_hours_papi_quilting_o3_big_buffer.chop6.prv

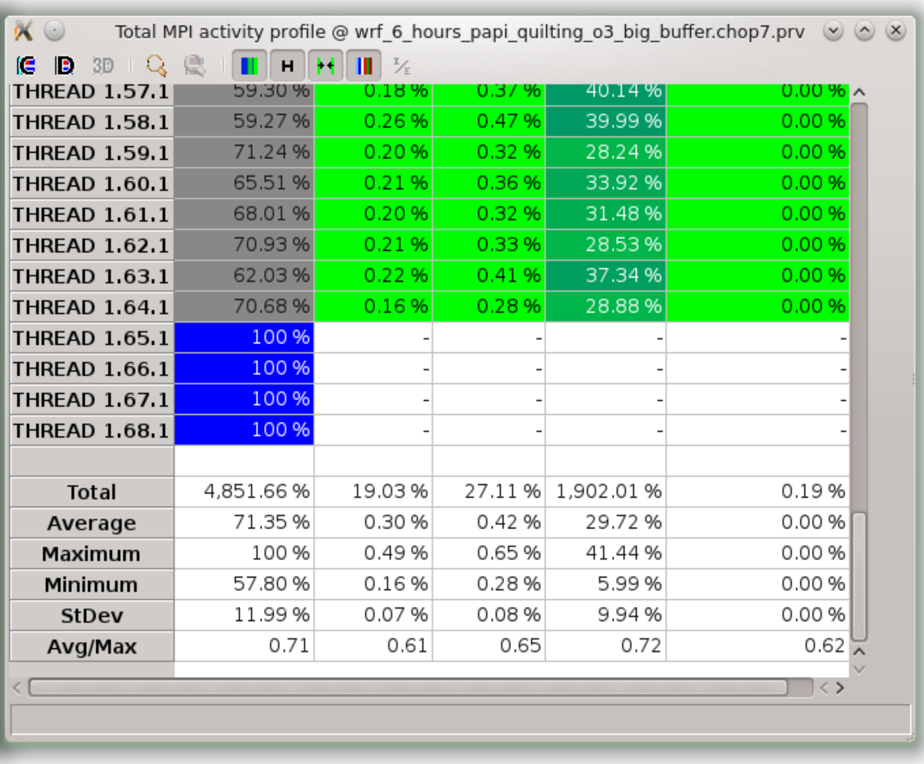
THREAD 1.32.1 [2.45325e+08..2.4732e+08) = 0 us

Trace Analysis

MPI calls profiling


Total MPI activity profile @ wrf_6_hours_papi_quilting_o3_big_buffer.chop7.prv
3D
H

	Outside MPI	MPI_Isend	MPI_Irecv	MPI_Wait	MPI_Comm_rank
THREAD 1.1.1	60.48 %	0.19 %	0.38 %	38.95 %	0.00 %
THREAD 1.2.1	61.05 %	0.25 %	0.49 %	38.21 %	0.00 %
THREAD 1.3.1	60.59 %	0.20 %	0.38 %	38.83 %	0.00 %
THREAD 1.4.1	69.76 %	0.21 %	0.33 %	29.70 %	0.00 %
THREAD 1.5.1	70.49 %	0.20 %	0.32 %	28.98 %	0.00 %
THREAD 1.6.1	61.64 %	0.19 %	0.34 %	37.83 %	0.00 %
THREAD 1.7.1	57.80 %	0.27 %	0.49 %	41.44 %	0.00 %
THREAD 1.8.1	59.92 %	0.16 %	0.34 %	39.58 %	0.00 %
THREAD 1.9.1	65.61 %	0.26 %	0.35 %	33.77 %	0.00 %
THREAD 1.10.1	60.66 %	0.34 %	0.44 %	38.56 %	0.00 %
THREAD 1.11.1	71.59 %	0.30 %	0.36 %	27.74 %	0.00 %
THREAD 1.12.1	64.85 %	0.34 %	0.41 %	34.40 %	0.00 %
THREAD 1.13.1	62.65 %	0.33 %	0.42 %	36.60 %	0.00 %
THREAD 1.14.1	69.75 %	0.28 %	0.36 %	29.61 %	0.00 %
THREAD 1.15.1	64.46 %	0.30 %	0.42 %	34.82 %	0.00 %
THREAD 1.16.1	67.09 %	0.25 %	0.33 %	32.33 %	0.00 %
THREAD 1.17.1	71.52 %	0.32 %	0.41 %	27.75 %	0.00 %
THREAD 1.18.1	72.78 %	0.30 %	0.40 %	26.51 %	0.00 %

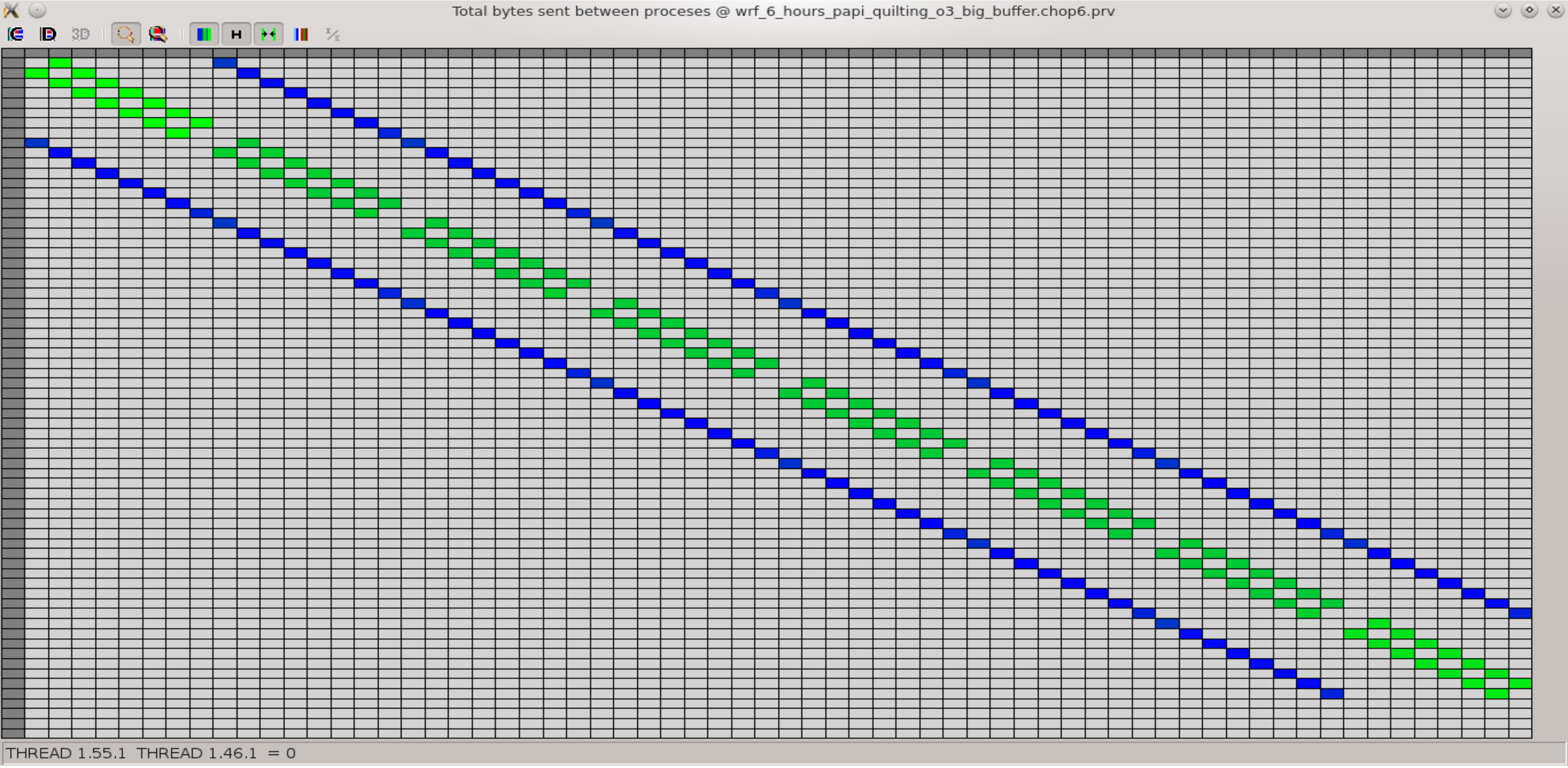

Total MPI activity profile @ wrf_6_hours_papi_quilting_o3_big_buffer.chop7.prv
3D
H

THREAD 1.57.1	59.30 %	0.18 %	0.37 %	40.14 %	0.00 %
THREAD 1.58.1	59.27 %	0.26 %	0.47 %	39.99 %	0.00 %
THREAD 1.59.1	71.24 %	0.20 %	0.32 %	28.24 %	0.00 %
THREAD 1.60.1	65.51 %	0.21 %	0.36 %	33.92 %	0.00 %
THREAD 1.61.1	68.01 %	0.20 %	0.32 %	31.48 %	0.00 %
THREAD 1.62.1	70.93 %	0.21 %	0.33 %	28.53 %	0.00 %
THREAD 1.63.1	62.03 %	0.22 %	0.41 %	37.34 %	0.00 %
THREAD 1.64.1	70.68 %	0.16 %	0.28 %	28.88 %	0.00 %
THREAD 1.65.1	100 %	-	-	-	-
THREAD 1.66.1	100 %	-	-	-	-
THREAD 1.67.1	100 %	-	-	-	-
THREAD 1.68.1	100 %	-	-	-	-
Total	4,851.66 %	19.03 %	27.11 %	1,902.01 %	0.19 %
Average	71.35 %	0.30 %	0.42 %	29.72 %	0.00 %
Maximum	100 %	0.49 %	0.65 %	41.44 %	0.00 %
Minimum	57.80 %	0.16 %	0.28 %	5.99 %	0.00 %
StDev	11.99 %	0.07 %	0.08 %	9.94 %	0.00 %
Avg/Max	0.71	0.61	0.65	0.72	0.62

- For the study of the statistics we exclude the I/O processes (scripting)
- Maximum value: 93.21% (communication efficiency)
- Average value: 69.55% (parallel efficiency)
- Avg/max value: 74.6% (global load balance)
- Note: we study just a small part of the whole execution

Trace Analysis

- Communication matrix
- The previous mentioned mapping from Peter Johnsen is validated

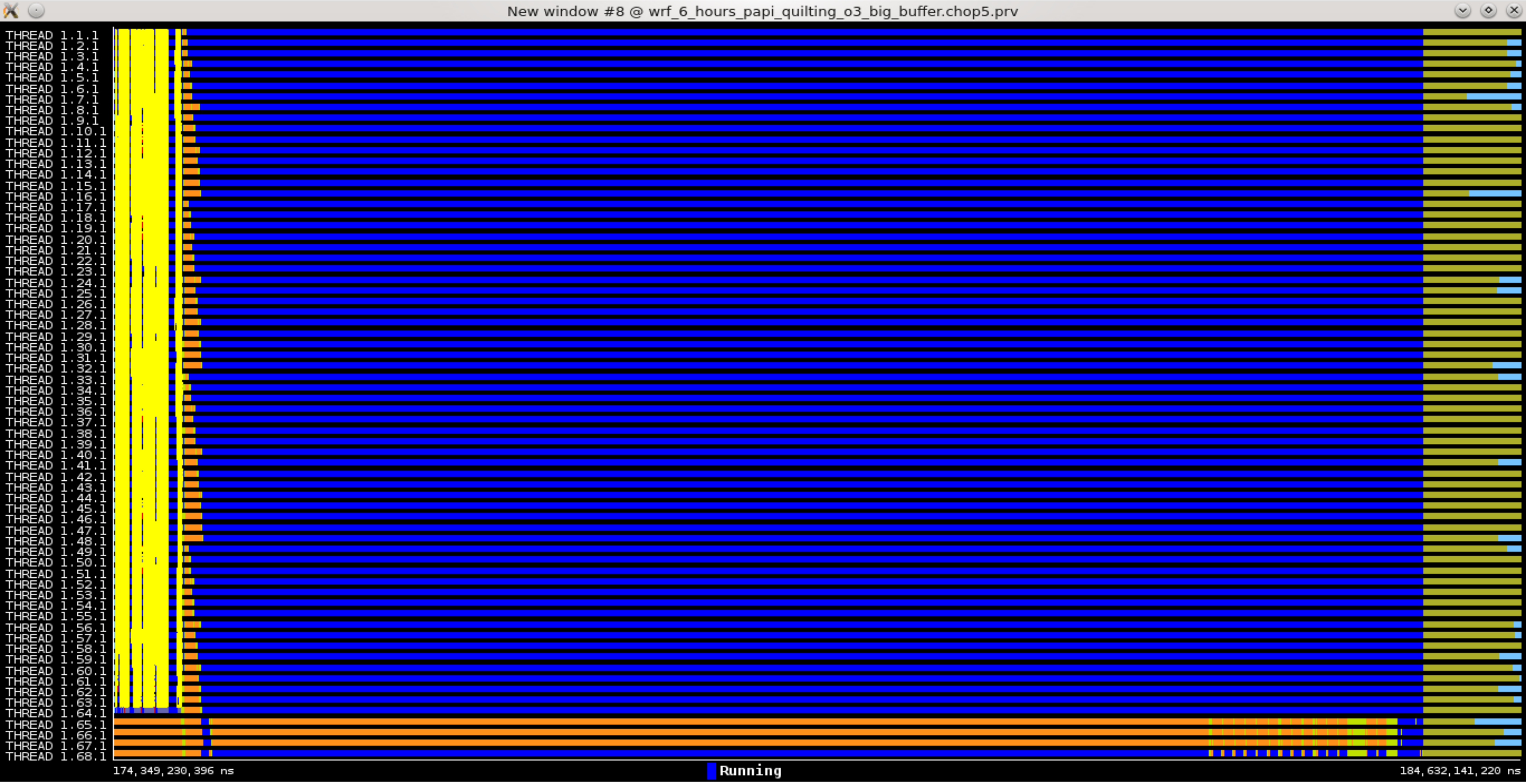


Total bytes sent between proceses @ wrf_6_hours_papi_quilting_o3_big_buffer.chop6.prv

THREAD 1.55.1 THREAD 1.46.1 = 0

Trace Analysis – End of the trace

- There is communication between the write tasks (last four).
- All the processes wait till the write tasks finish for the case of I/O quilting.



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Conclusions

- Optimize first your application through the provided options, you can be surprised
- Be careful about the combination of the optimization options
- Different number of processors and workload does not mean that they can be optimized with the same approach
- Paraver can provide a lot of insight information about the behavior of an earth science model
- Integrate new technologies

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Session 2: HPC ENVIRONMENT TUTORIAL

Outline

- Connect with SSH to Mare Nostrum 3
- Interact inside Mare Nostrum 3
- Compile and run serial program
- Compile and run parallel program
- Cancel a job
- Copy results to your local machine

- Basic visualization
  - Ncview
  - Panoply

Users in Mare Nostrum

- One user for each student
  - Username: nct010[01-15]
  - Password: NCT.2013.[01-15]

- Two folders
  - Home: /gpfs/home/nct00/nct010XX/
  - Projects: /gpfs/projects/nct00/nct010XX/ (working folder)

- In the hands on, when we talk of nct010XX, replace by your number !!!

Windows or Linux

- With Linux, no extra work is needed
- With Windows, need to install
  - SSH terminal: putty
    - http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html
  - X Windows Manager: Xming
    - http://sourceforge.net/projects/xming/files/Xming/6.9.0.31/Xming-6-9-0-31-setup.exe/download

Connect / Interact with SSH to MN3

- Open a terminal
- Type
  - ssh -X nct010XX@mn1.bsc.es
- List your home folder
  - ls
- Create folder in projects
  - mkdir /gpfs/projects/nct00/nct010XX/
- Copy exercises in your project folder
  - cp -r /gpfs/projects/nct00/nct00001/PRACE-Course /gpfs/projects/nct00/nct010XX/
- Verify the X windows is working properly
  - xeyes, xclock…
- Open text a editor (vi, emacs…)

Modifiying .bashrc

- In .bashrc we define variables and locations programs
- Edit your .bashrc (with emacs or vi)
  - emacs ~/.bashrc
- Paste these definitions

```
#NETCDF PROCESSING
export PATH=$PATH:/gpfs/apps/MN3/NETCDF/3.6.3/bin
export PATH=$PATH:/gpfs/apps/MN3/CDO/1.5.9/bin
export PATH=$PATH:/gpfs/apps/MN3/NCO/4.2.3/bin
export PATH=$PATH:/gpfs/apps/MN3/NCVIEW/2.1.2/bin
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/gpfs/apps/MN3/NETCDF/3.6.3/lib:/gpfs/apps/MN3/NETCDF/4.1.3/lib
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/gpfs/apps/MN3/UDUNITS/2.1.24/lib:/gpfs/apps/MN3/UDUNITS/1.12.11/lib
#IMAGES
export PATH=$PATH:/gpfs/apps/MN3/IMAGEMAGICK/6.8.1-9/bin
#GRADS
export PATH=$PATH:/gpfs/apps/MN3/GRADS/2.0.2/bin
export GADDIR=/gpfs/apps/MN3/GRADS/2.0.2/data/
# NCL
export NCARG_ROOT=/gpfs/apps/MN3/NCL/6.1.2
export PATH=$PATH:/gpfs/apps/MN3/NCL/6.1.2/bin
export
LD_LIBRARY_PATH=/gpfs/apps/MN3/PROJ/4.8.0/lib/:/gpfs/apps/MN3/GDAL/1.9.2/lib/:/gpfs/apps/MN3/HDF5/1.8.10/lib:/gpfs/apps/MN3/SZ
IP/2.1/lib/:$LD_LIBRARY_PATH
#R
export PATH=$PATH:/gpfs/apps/MN3/R/2.15.2/bin
#PANOPLY
export PATH=$PATH:/gpfs/apps/MN3/PANOPLY/3.1.7/
#MAPGENERATOR
export PATH=/gpfs/apps/MN3/PYTHON/2.7.3/bin/:$PATH
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/gpfs/apps/MN3/INTEL/mkl/lib/intel64
```

Compile and run serial program

- Open a terminal
- Log MN3
- Change dir to example_serial
- Compile
  - make
- Submit job
  - bsub < submit.cmd
- Analyze outputs (with vi or emacs…)
  - vi “job_id”.err
  - vi “job_id”.out

Compile and run parallel program

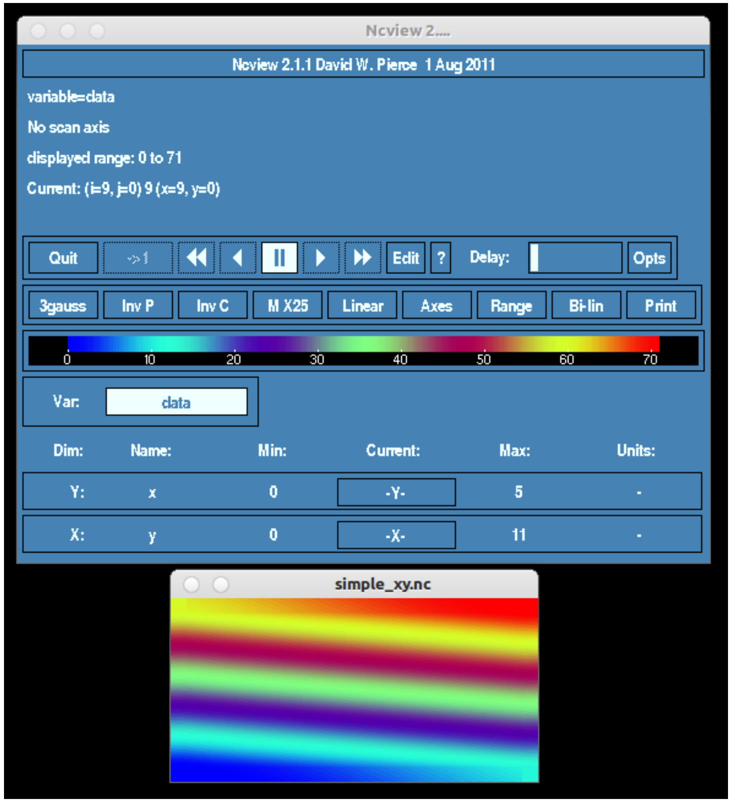
- Open a terminal
- Log MN3
- Change dir to example_parallel
- Compile
- Submit job
- Analyze outputs
- Run the example with 12 cores
  - Change total_tasks in submit.cmd
- Analyze outputs

Cancel a job

- Open a terminal
- Log MN3
- Change dir to example_parallel_cancel
- Compile
- Submit job
- Analyze outputs
- Queue jobs
  - bjobs
- Identify your job
  - bjobs ang get the job_id
- Kill job
  - bkill “job_id”

Job with Netcdf Library

- Open a terminal
- Log MN3
- Change dir to example_create_netcdf
- Open Makefile and analyze it
- Compile
- Submit
- Open netcdf file with ncview
  - ncview simple_xy.nc
- Open netcdf with ncdump
  - ncdump -h simple_xy.nc
  - Analyze header
- Copy netcdf file in your local Machine
  - From local: scp your nct010XX@mn1.bsc.es:/path/simple_xy.nc .



Ncview 2.1.1 David W. Pierce 1 Aug 2011

variable=data

No scan axis

displayed range: 0 to 71

Current: (i=9, j=0) 9 (x=9, y=0)

Var: data

Dim:	Name:	Min:	Current:	Max:	Units:
Y:	x	0	-Y-	5	-
X:	y	0	-X-	11	-

Basic Visualization - NCVIEW

- Open WRF file with NCVIEW
  - cd /gpfs/projects/nct00/nct010XX/PRACE-Course/visualization-hands_on
  - ncview wrfout_d01_2012-09-16_12:00:00
  - display 2D variables
    - T2
    - U10
    - V10
  - Display 3D variables
    - T
    - W
  - Use another color palette
  - Change range values
  - Plot a time series of two points

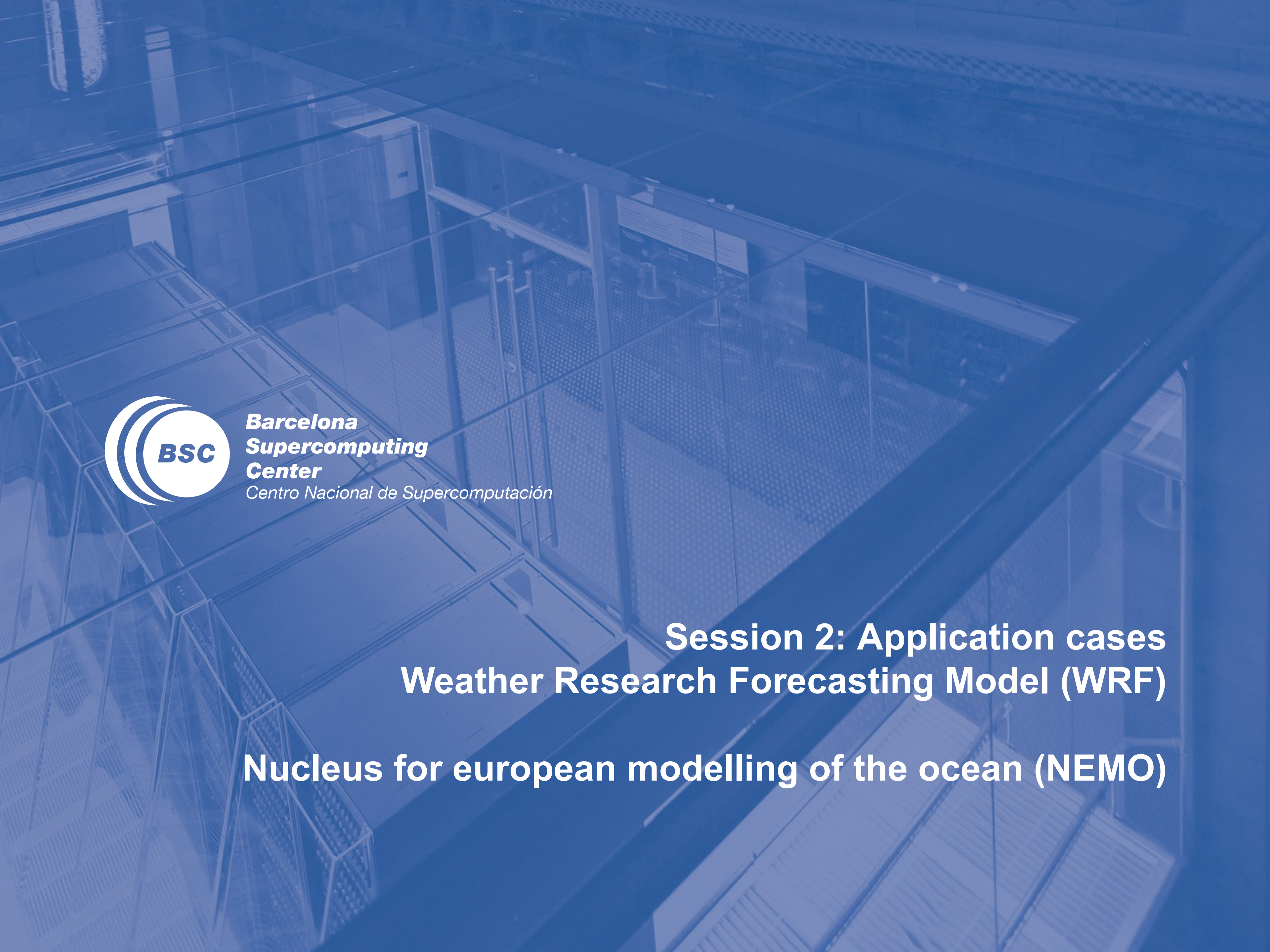
Basic Visualization - PANOPLY

- Example on Panoply
  - Open wrfout_d01 netcdf file
  - Open a GRIB file

- To open panoply.
  - Run ./panoply.sh

- WRF File
  - Choose another view (regional is better).
  - Generate plot of T2
  - Generate KMZ with U10

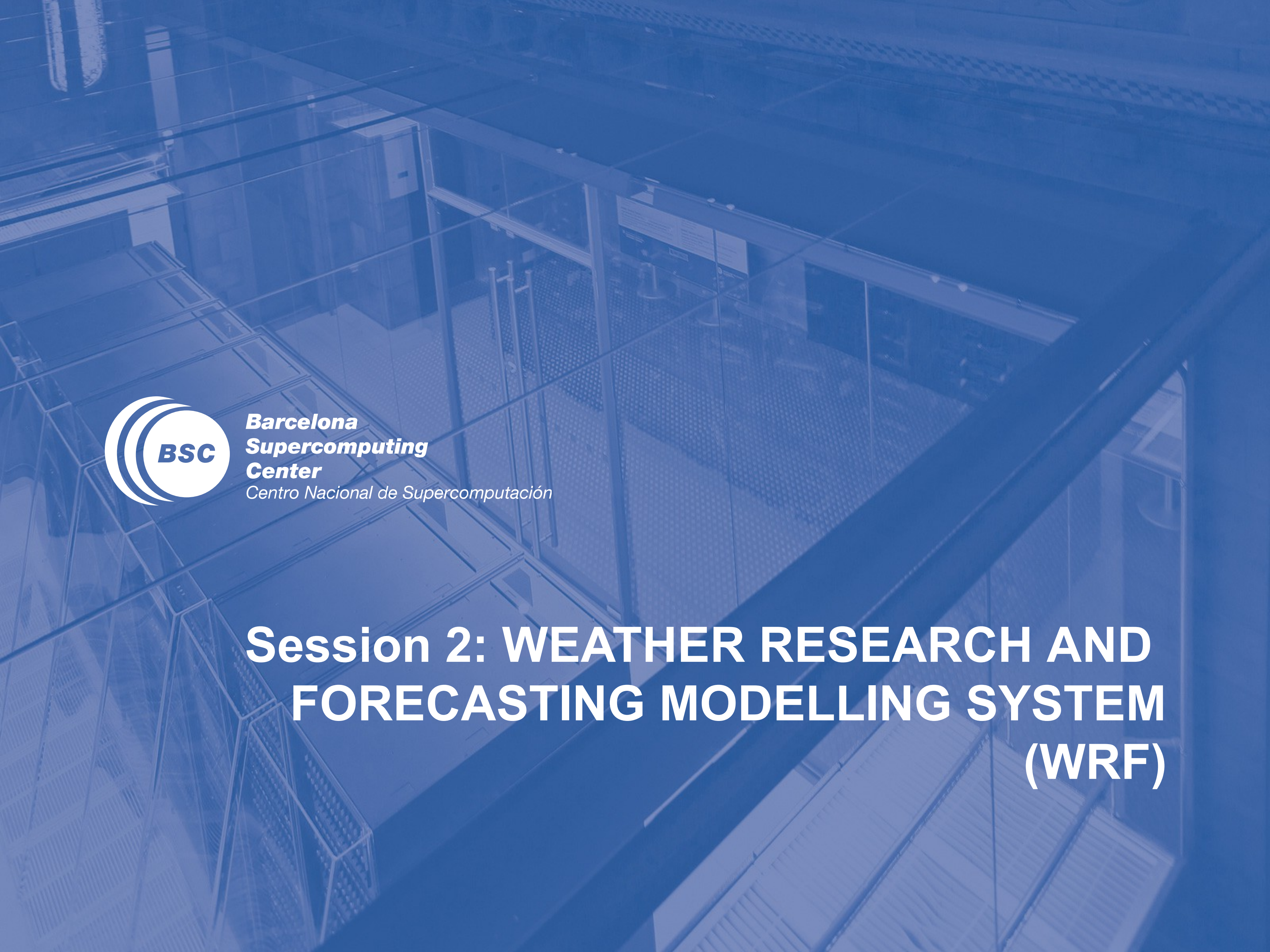
- GRIB File
  - Display “Temperature @ Ground or water surface”



Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Session 2: Application cases
Weather Research Forecasting Model (WRF)

Nucleus for european modelling of the ocean (NEMO)



Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Session 2: WEATHER RESEARCH AND FORECASTING MODELLING SYSTEM (WRF)

Outline

- Application case: the meteorological model *Weather and Research Forecasting System* (WRF)
  - Overview of WRF model (based on the online tutorial of the model)
  - Model Hands-on
    - Build and compile
    - Configuration
    - Execution

What is WRF?

Weather Reasearch & Forecasting (WRF) Model

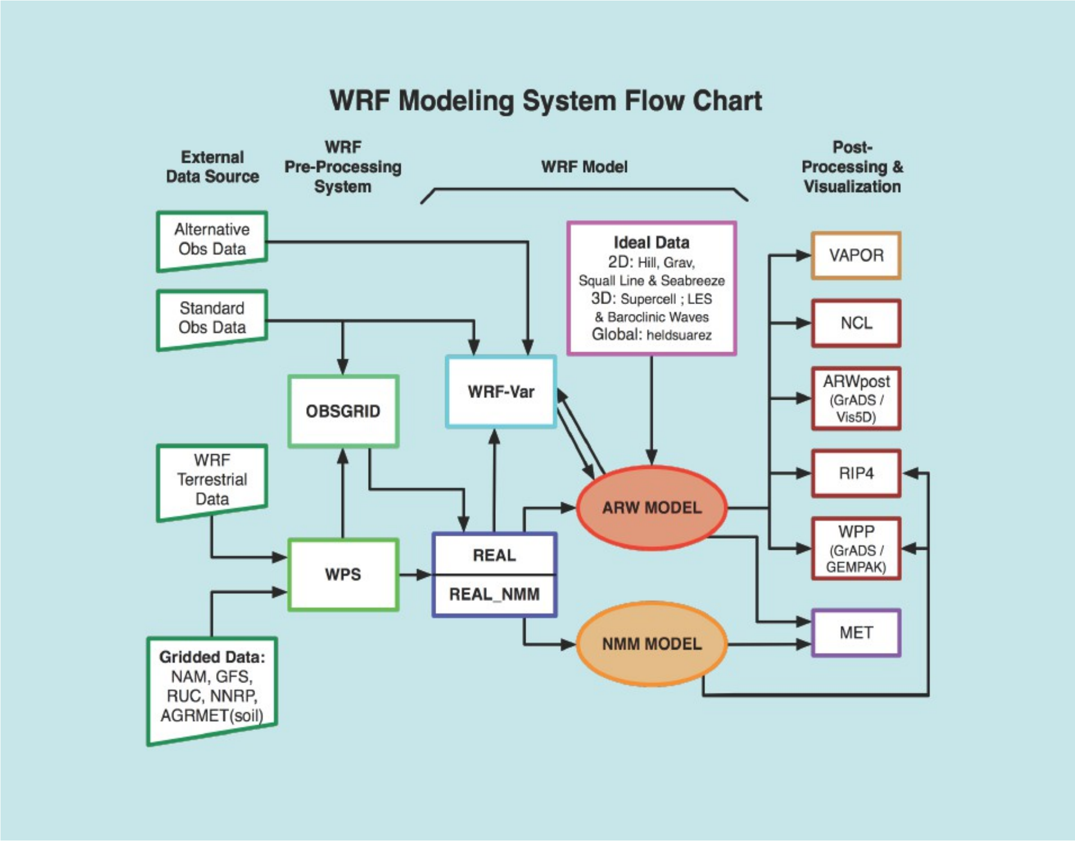
- Mesocale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs.

- Suitable for a broad spectrum of meteorological applications across scales ranging from meters to thousands of kilometers.

- It is a supported "community model" - www.wrf-model.org

- Development led by NCAR, NOAA/GSD and NOAA/NCEP/EMS with partnerships at AFWA, FAA, NRL, and collaborations with universities and other government agencies in the US and worlwide.

Modelling System Components


WRF Modeling System Flow Chart
External Data Source
WRF Pre-Processing System
WRF Model
Post-Processing & Visualization
Alternative Obs Data
Standard Obs Data
WRF Terrestrial Data
Gridded Data: NAM, GFS, RUC, NNRP, AGRMET(soil)
OBSGRID
WPS
WRF-Var
REAL
REAL_NMM
Ideal Data 2D: Hill, Grav, Squall Line & Seabreeze 3D: Supercell ; LES & Baroclinic Waves Global: heldsuarez
ARW MODEL
NMM MODEL
VAPOR
NCL
ARWpost (GrADS / Vis5D)
RIP4
WPP (GrADS / GEMPAK)
MET

- WRF Pre-processing System (WPS)
- WRF-VAR
- ***WRF Model***
- Post-processing and visualization tools

Wide information in WRF webpages:

www.mmm.ucar.edu/

www.dtcenter.org/wrf-nmm/users/

Documentation, online tutorials, source code…

- Large set of physical parameterizations
- Capabilities for data assimilation

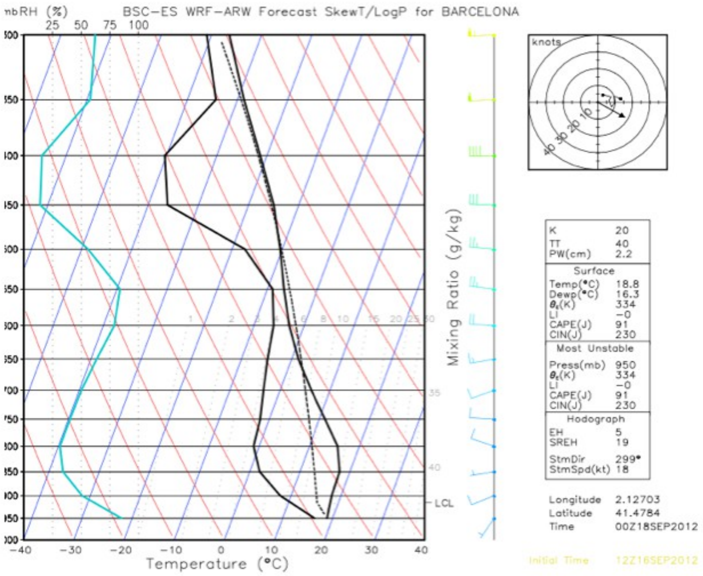
WRF Model Characteristics

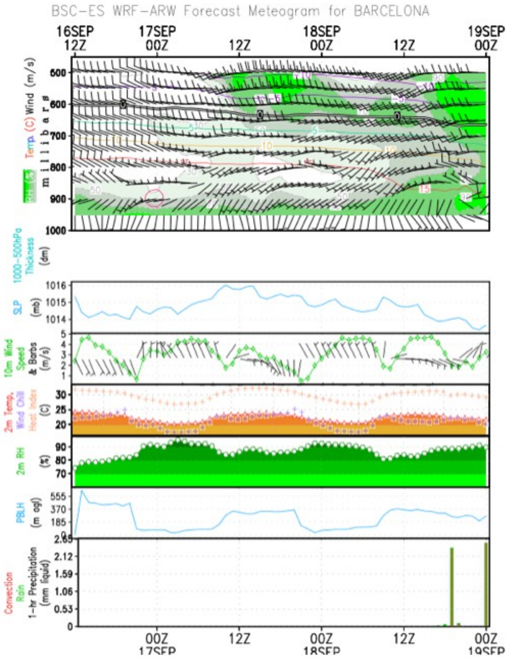
- We focus on the *WRF-ARW* model
  - Solves the Fully compressible, nonhydrostatic, Euler equations
  - Terrain following hydrostatic pressure coordinate
  - Arakawa C-grid
  - The model uses higher-order numerics
    - Runge-Kutta 2nd- and 3rd-order time integration schemes
    - 2nd- to 6th-order advection schemes in both horizontal and vertical directions
  - It uses a time-split small step for acoustic and gravity-wave modes
  - The dynamics conserves scalar variables
  - Modelling System contains: initialization programs, WRF model, nesting capabilities.

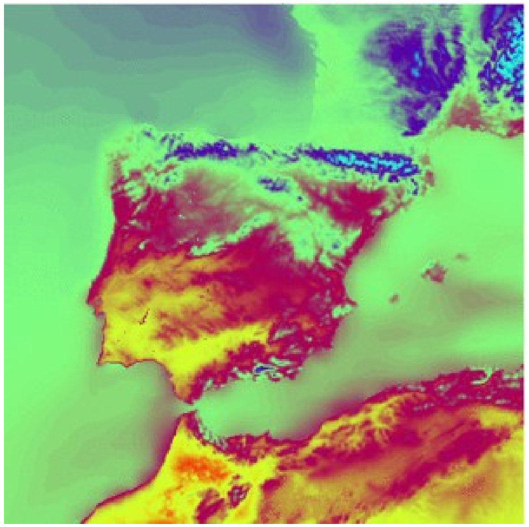
Software requirements

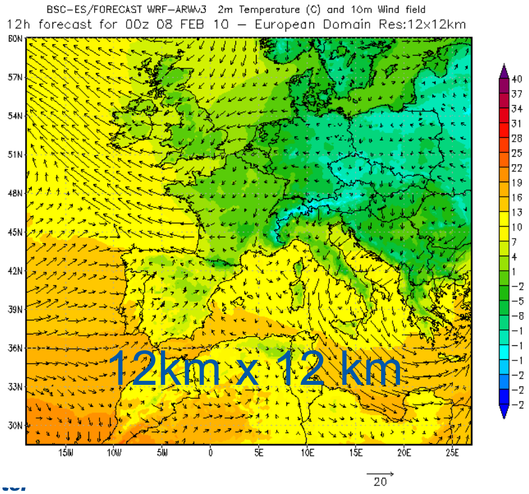
- Fortran 90 or 95 and C compiler
- perl 5.04 or later
- If MPI and OpenMP compilation is desired, MPI or OpenMP libraries are required
- WRF I/O API supports netCDF, pnetCDF, PHD5, GriB 1 and GriB 2 formats
- UNIX utilities: csh and Bourne shell, make, M4, sed, awk, and the uname command

Products



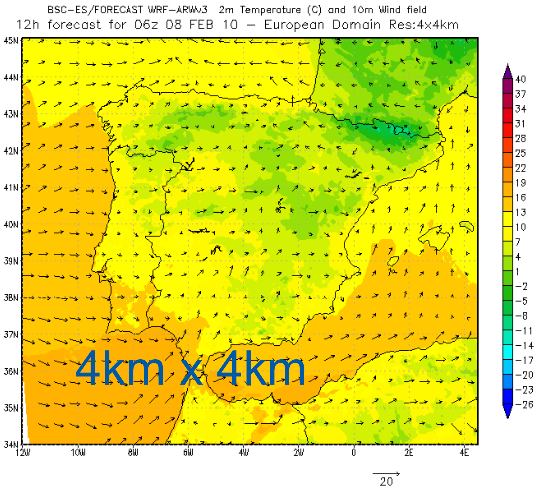






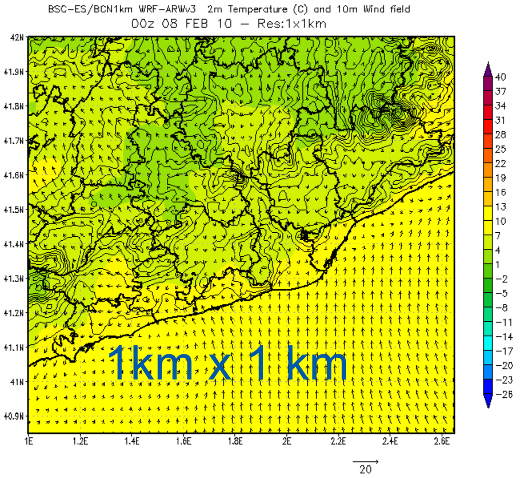
BSC-ES/FORECAST WRF-ARWv3 2m Temperature (C) and 10m Wind field
12h forecast for 00z 08 FEB 10 – European Domain Res:12x12km
60N 57N 54N 51N 48N 45N 42N 39N 36N 33N 30N
15W 10W 5W 0 5E 10E 15E 20E 25E
40 37 34 31 28 25 22 19 16 13 10 7 4 1 −2 −5 −8 −1 −1 −1 −2 −2 −2
20

12km x 12 km



BSC-ES/FORECAST WRF-ARWv3 2m Temperature (C) and 10m Wind field
12h forecast for 06z 08 FEB 10 – European Domain Res:4x4km
45N 44N 43N 42N 41N 40N 39N 38N 37N 36N 35N 34N
12W 10W 8W 6W 4W 2W 0 2E 4E
40 37 34 31 28 25 22 19 16 13 10 7 4 1 −2 −5 −8 −11 −14 −17 −20 −23 −26
20

4km x 4km



BSC-ES/BCN1km WRF-ARWv3 2m Temperature (C) and 10m Wind field
00z 08 FEB 10 – Res:1x1km
42N 41.9N 41.8N 41.7N 41.6N 41.5N 41.4N 41.3N 41.2N 41.1N 41N 40.9N
1E 1.2E 1.4E 1.6E 1.8E 2E 2.2E 2.4E 2.6E
40 37 34 31 28 25 22 19 16 13 10 7 4 1 −2 −5 −8 −11 −14 −17 −20 −23 −26
20

1km x 1 km

Model Hands-on

- This is not a WRF tutorial – for a full description visit online tutorials

- Objective:
  - Acces to MN3 supercomputer
  - Build and compile WRF modelling system
  - Edit and configure a model run
  - Submit a batch job
  - Monitor job execution
  - Visualize model input and output

Model Hands-on

- Get the source code:
  - Already copied in MN3:
  - Go to the wrf folder
    - cd /gpfs/projects/nct00/nct010XX/PRACE-Course/wrf/source/

- Unpack the code:
  - gunzip WRFV3.5.1.TAR.gz
  - tar -xf WRFV3.5.1.TAR

- This will create the WRFV3/ directory

Build and compile WRF

- Move into WRFV3 directory
  - cd WRFV3

- Scripts to compile and configure the model
  - clean: script to clean created files, executables
  - compile: script for compiling WRF code
  - configure: script for configure the configure.wrf file for compile

- Define Shell environment
  - export NETCDF=/gpfs/apps/MN3/NETCDF/3.6.3

- Configure WRF environment
  - ./configure
  - Select option 19. Linux x86_64 i486 i586 i686, ifort compiler with icc (dmpar)
  - Compile for nesting? Select basic (1)
  - File configure.wrf will be created
    - **Copy configure.wrf from** /gpfs/projects/nct00/nct00001/PRACE-Course/wrf/source/WRFV3 if you have problems

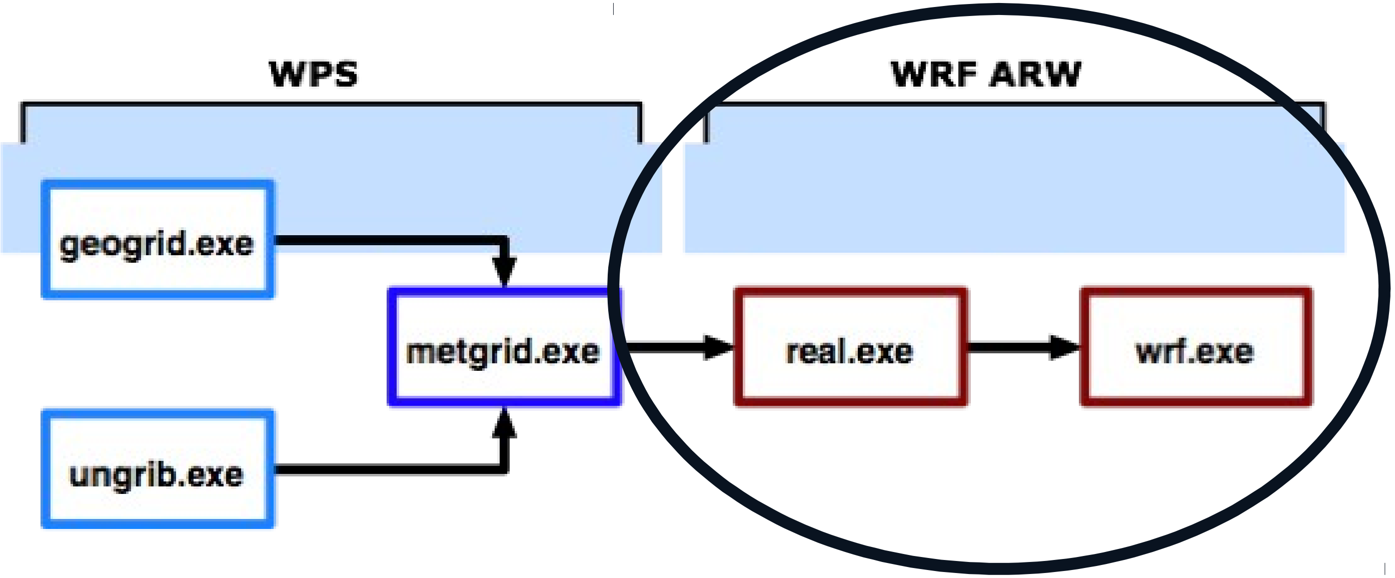
Build and compile WRF

- Edit configure.wrf to speed-up compilation
  - Set FCOPTIM= -O0

- Compile WRF for real data cases
  - Set the following environment variable before compile:
    - Export WRF_EM_CORE=1
  - ./compile em_real >& compile.log
  - Check the compile.log file for any errors.

- If compilation is successful, it will create the following executables in the WRFV3/main/ directory
  - ndown.exe
  - nup.exe
  - ***real.exe***
  - ***wrf.exe***

Model Flow



real.exe

- This program vertically interpolates the met_em* files (generated by metgrid.exe), creates boundary and initial condition files, and does some consistency checks.

wrf.exe

- Generates the model forecast.

Preparing a model run – input files

- Get input files:
  - Already copied in MN3:
    - /gpfs/projects/nct00/nct010XX/PRACE-Course/wrf/input_data
      - namelist.input
      - met_em.d01.2012-09-16_12:00:00.nc
      - met_em.d01.2012-09-16_18:00:00.nc

- Copy input files in WRFV3/test/em_real/

- Edit namelist.input
  - Set Hours = 06, end_day = 16, end_hour = 18

- Visualize input files met_em.d01…
  - ncview met_em.d01.2012-09-16_12:00:00.nc

Preparing a model run – First step REAL.EXE

- Get submit script file:
  - Already copied in MN3:
    - /gpfs/projects/nct00/nct010XX/PRACE-Course/wrf/input_data
      - submit_real.cmd
  - Copy submit_real.cmd to WRFV3/test/em_real
- Edit submit_real.cmd:
  - Set total_tasks = 4
- Submit real pre-process:
  - bsub submit_real.cmd
- Monitor job status: bjobs

```
JOBID    NAME    USER     STATE      TIME TIMELIMIT CPUS  NODES NODELIST(REASON)
480031  0.REAL bsc32771  PENDING      0:00     12:00    4      1 (Priority)
```

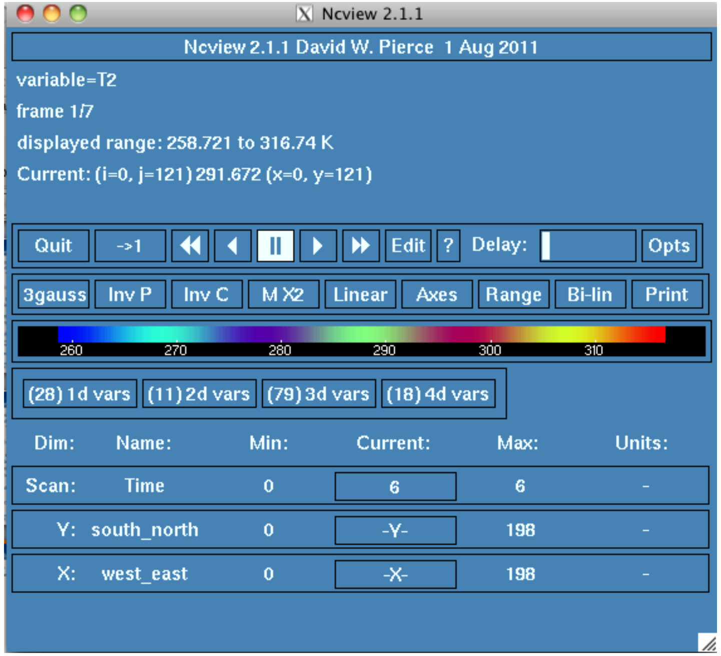
Check with ncview wrfinput_d01 and wrfbdy_d01 are created

Preparing a model run – Second step WRF.EXE

- Get submit script file:
  - Already copied in MN3:
    - /gpfs/projects/nct00/nct010XX/PRACE-Course/wrf/input_data
      - submit_wrf.cmd
  - Copy submit_wrf.cmd to WRFV3/test/em_real
- Edit submit_wrf.cmd:
  - Set total_tasks = 24
- Submit wrf process:
  - bsub submit_wrf.cmd
- Monitor job status: bjobs
- Check rsl.out.0000 and wrfout_d01_2012-09-16_12:00:00
  - At the end or rsl.out… “SUCCESS COMPLETE WRF”

Visualize WRF output

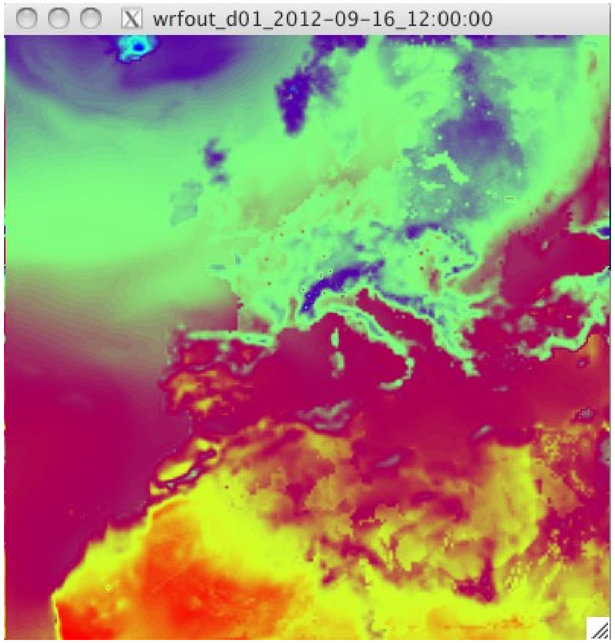
- Use ncview to visualize the wrf output netcdf file:
  - wrfout_d01_2012-09-16_12:00:00

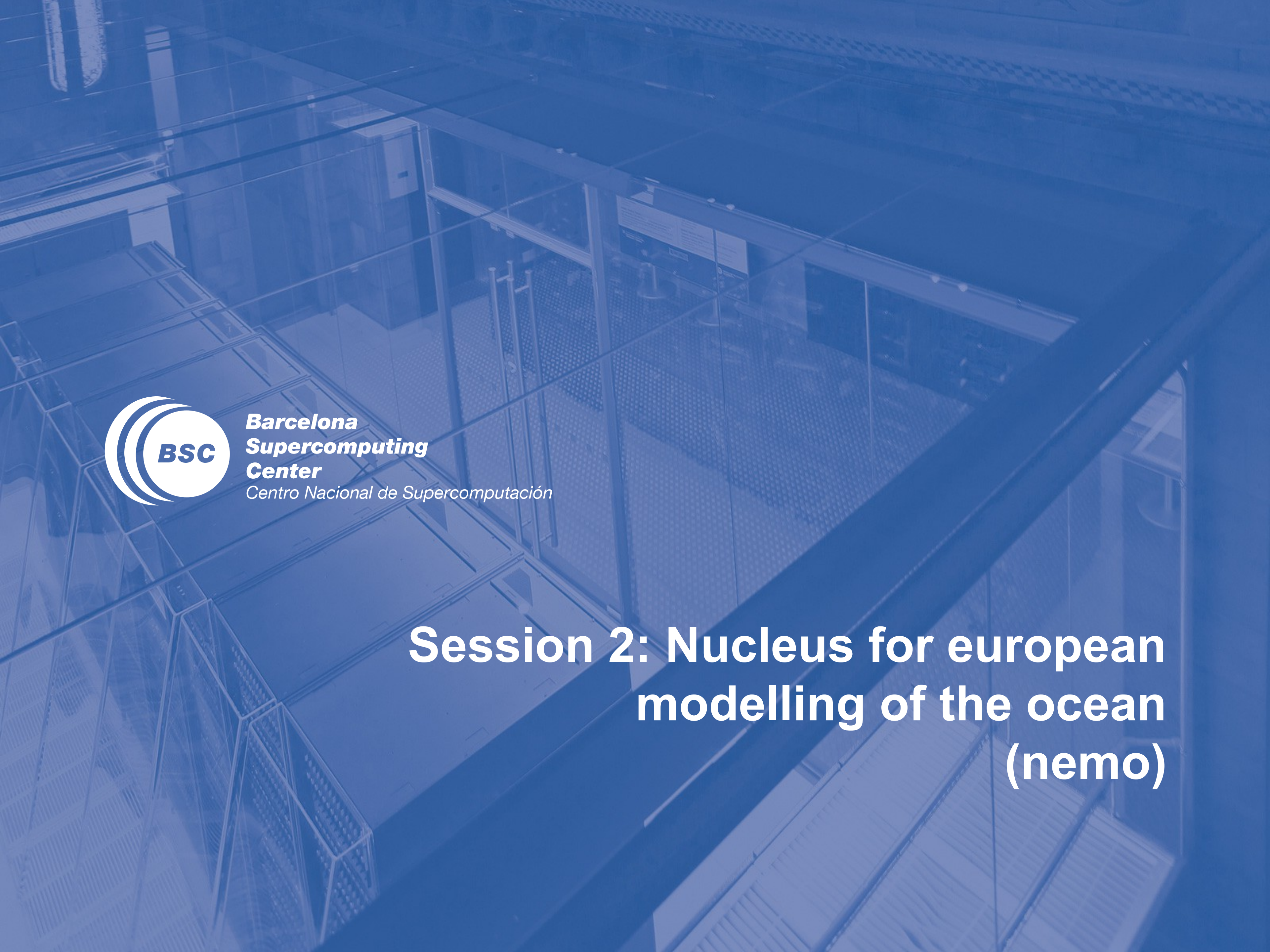


Ncview 2.1.1 David W. Pierce 1 Aug 2011

variable=T2
frame 1/7
displayed range: 258.721 to 316.74 K
Current: (i=0, j=121) 291.672 (x=0, y=121)

Dim:	Name:	Min:	Current:	Max:	Units:
Scan:	Time	0	6	6	-
Y:	south_north	0	-Y-	198	-
X:	west_east	0	-X-	198	-





Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

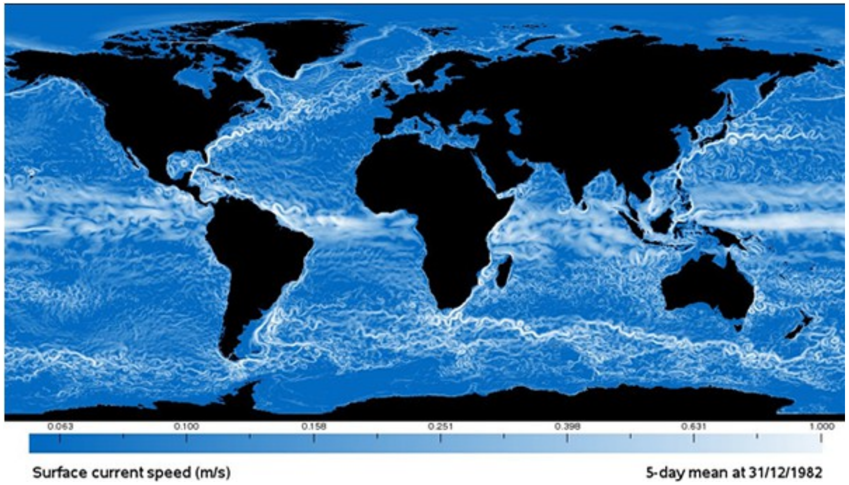
Session 2: Nucleus for european modelling of the ocean (nemo)

Outline

- Application case: NEMO (Nucleus for European Modelling of the Ocean) is a state-of-the-art modeling framework for oceanographic research, operational oceanography seasonal forecast and climate studies.

- More information in http://www.nemo-ocean.eu/

- Overview of NEMO model and Hands-on
  - Build and compile
  - Execution
  - Visualize results



What is NEMO?

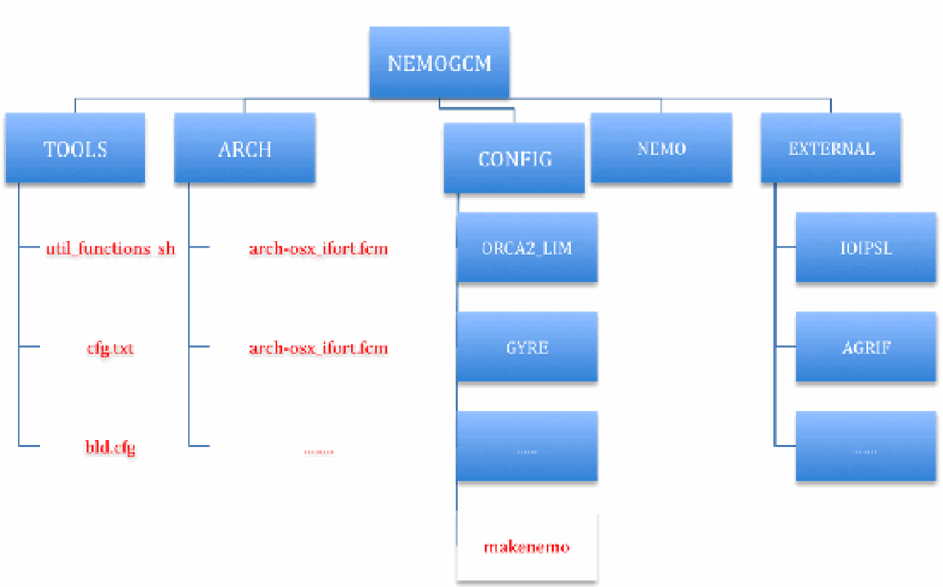
- NEMO is an ocean model with several modules associated with ocean processes.

- NEMO includes 5 major components
  - the blue ocean (ocean dynamics, NEMO-OPA)
  - the white ocean (sea-ice, NEMO-LIM)
  - the green ocean (biogeochemistry, NEMO-TOP) ;
  - the adaptative mesh refinement software (AGRIF) ;
  - the assimilation component NEMO_TAM

- In the hands-on, we will work with NEMO-OPA and NEMO-LIM.

Software requirements

- Fortran 90 or 95
- MPI if parallel execution is desired
- Perl
- Input/Ouput netCDF
- UNIX utilities: csh and Bourne shell, make, M4, sed, awk, and the uname command



NEMOGCM

TOOLS

util_functions.sh

cfg.txt

bld.cfg

ARCH

arch-osx_ifort.fcm

arch-osx_ifort.fcm

......

CONFIG

ORCA2_LIM

GYRE

......

makenemo

NEMO

EXTERNAL

IOIPSL

AGRIF

......

Output paralel behaviour

- In paralel execution, each process writes a piece of the netcdf.

```
-rw-r--r-- 1 bsc32353 bsc32  11368352 Dec  3 17:52 ORCA2_5d_00010101_00011231_grid_T_0000.nc
-rw-r--r-- 1 bsc32353 bsc32  10895152 Dec  3 17:52 ORCA2_5d_00010101_00011231_grid_T_0001.nc
-rw-r--r-- 1 bsc32353 bsc32  10895152 Dec  3 17:52 ORCA2_5d_00010101_00011231_grid_T_0002.nc
-rw-r--r-- 1 bsc32353 bsc32  10895152 Dec  3 17:52 ORCA2_5d_00010101_00011231_grid_T_0003.nc
-rw-r--r-- 1 bsc32353 bsc32  10421952 Dec  3 17:52 ORCA2_5d_00010101_00011231_grid_T_0004.nc
-rw-r--r-- 1 bsc32353 bsc32  10421952 Dec  3 17:52 ORCA2_5d_00010101_00011231_grid_T_0005.nc
-rw-r--r-- 1 bsc32353 bsc32  10421952 Dec  3 17:52 ORCA2_5d_00010101_00011231_grid_T_0006.nc
-rw-r--r-- 1 bsc32353 bsc32  10895152 Dec  3 17:52 ORCA2_5d_00010101_00011231_grid_T_0007.nc
```

- At the end of the execution,we need a tool to gather them all
- ../../../TOOLS/REBUILD_NEMO/rebuild_nemo outputs/ORCA2_5d_00010101_00011231_grid_T 8

```
-rw-r--r-- 1 bsc32353 bsc32 641622864 Dec  3 18:04 ORCA2_5d_00010101_00011231_grid_T.nc
```

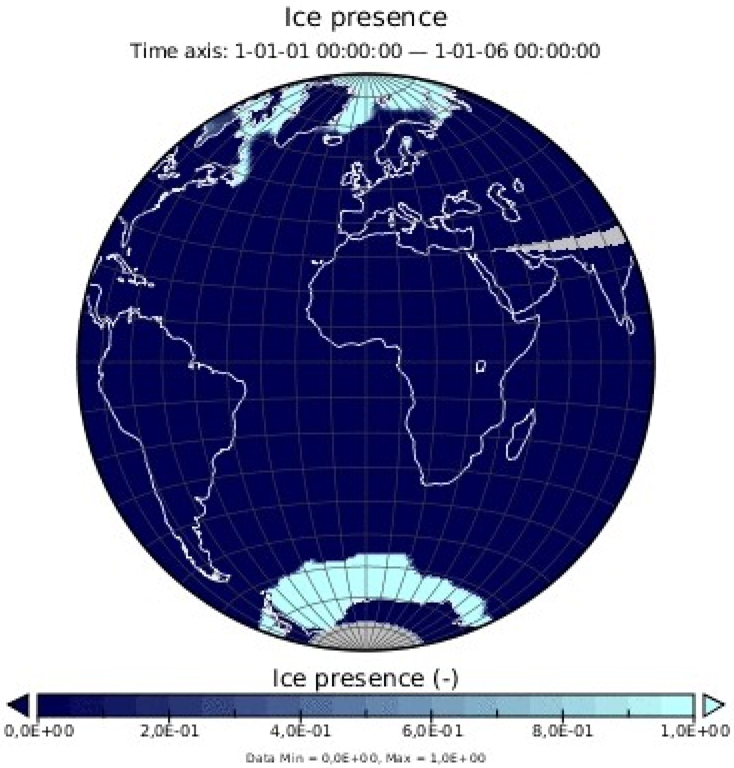
Compilation

- Compilations are done with MAKENEMO utility

```
./makenemo -h
Usage   : makenemo  [-h] [-n name] [-m arch] [-d dir1 dir2] [-r conf] [-s Path] [-e Path] [-j No] [-v No]
 -h               : help
 -h institute : specific help for consortium members
 -n name      : config name, [-n help] to list existing configurations
 -m arch      : choose compiler, [-m help] to list existing compilers
 -d dir       : choose NEMO sub-directories
 -r conf      : choose reference configuration
 -s Path      : choose alternative location for NEMO main directory
 -e Path      : choose alternative location for MY_SRC directory
 -j No        : number of processes used to compile (0=nocompilation)
 -v No        : set verbosity level for compilation [0-3]
 -t dir       : temporary directory for compilation
```

- Compiling GYRE, with ifort on linux to create a MY_GYRE configuration
  - makenemo –m ifort_linux –r GYRE -n MY_GYRE
- Recompile it
  - makenemo
- Now, create and compile ORCA_LIM3
  - makenemo –n ORCA_LIM3 (and answer)
- Now, create and compile ORCA2_LIM_2_2, add and delete keys, based on ORCA2_LIM
  - makenemo –n ORCA2_LIM_2_2 -r ORCA2_LIM add_key "key_mpp_mpi key_nproci=2 key_nprocj=2" del_key "key_agrif" (and answer)
- To remove a bad configuration
  - makenemo –n ORCA2_LIM_2_2 clean_config (and answer)

Products


Ice presence
Time axis: 1-01-01 00:00:00 — 1-01-06 00:00:00
Ice presence (-)
0,0E+00
2,0E-01
4,0E-01
6,0E-01
8,0E-01
1,0E+00
Data Min = 0,0E+00, Max = 1,0E+00

At the end of this session, you should be able to reproduce this visualization

Hands On - Compilation

- Enter in Nemo Folder
  - cd /gpfs/projects/nct00/nct010XX/
- cd nemo/dev_v3_4_STABLE_2012/NEMOGCM/CONFIG/

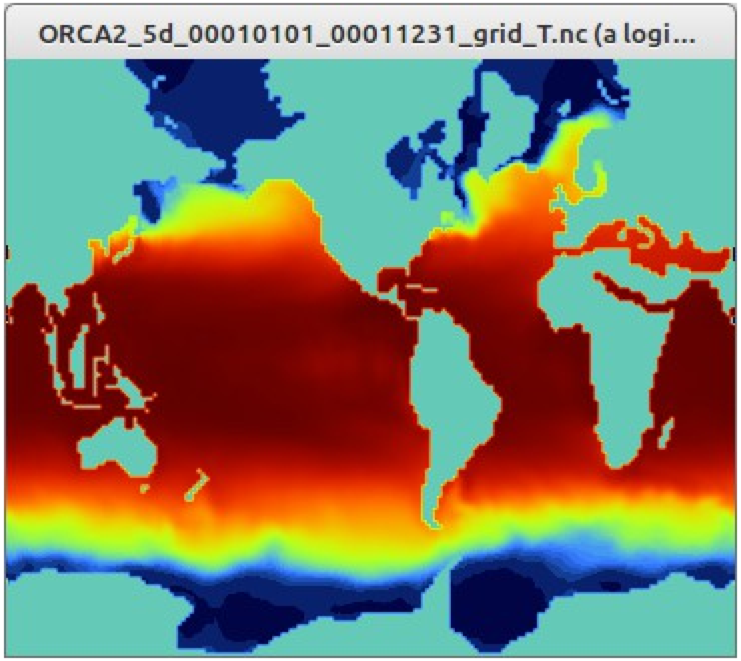
- Compile the model
- makenemo -m ifort_MN3 -n ORCA2_LIM_PRACE -r ORCA2_LIM

- Check compilation
  - nemo/dev_v3_4_STABLE_2012/NEMOGCM/CONFIG/ORCA2_LIM_PRACE/BLD/bin/nemo.exe
  - nemo/dev_v3_4_STABLE_2012/NEMOGCM/CONFIG/ORCA2_LIM_PRACE/BLD/bin/server.exe

Hands On – Running the model

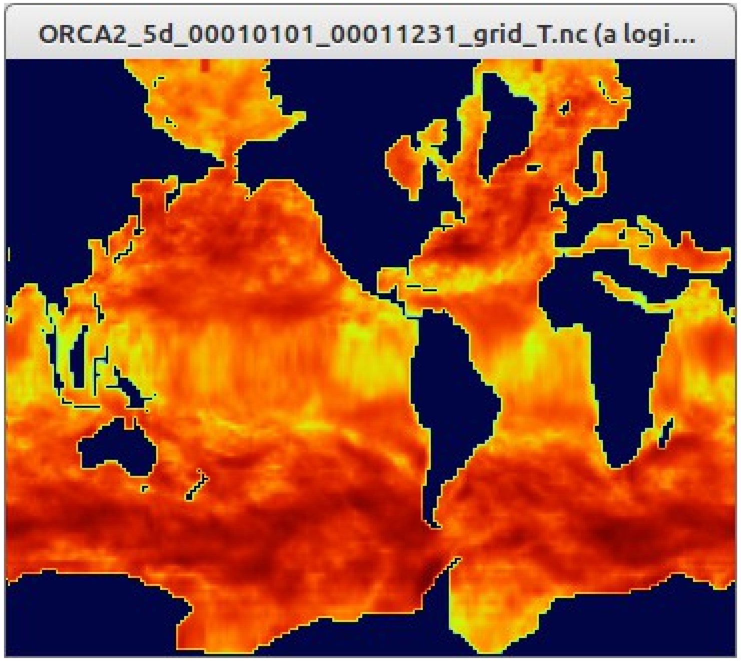
- Enter the experiment folder
  - cd nemo/dev_v3_4_STABLE_2012/NEMOGCM/CONFIG/ORCA2_LIM_PRACE/MY_WORK
- Untar and uncompress the test case files
  - tar -xvf ../../../../../ORCA2_LIM_nemo_v3.4.tar
  - gunzip *
- Go to the running folder
  - cd ../EXP00
- Link with the test case files
  - ln -sf ../MY_SRC/* .
- Copy a MN3 submit file
  - cp ../../ORCA2_LIM_BSC/EXP00/bsub_nemo.cmd .
- Submit job
  - bsub < bsub_nemo.cmd
- Wait for the results and check netcdf files are created
- Gather results
  - ../../../TOOLS/REBUILD_NEMO/rebuild_nemo outputs/ORCA2_1d_00010101_00011231_grid_T 32
  - ../../../TOOLS/REBUILD_NEMO/rebuild_nemo outputs/ORCA2_1m_00010101_00011231_grid_T 32
  - ../../../TOOLS/REBUILD_NEMO/rebuild_nemo outputs/ORCA2_1y_00010101_00011231_grid_T 32
  - ../../../TOOLS/REBUILD_NEMO/rebuild_nemo outputs/ORCA2_5d_00010101_00011231_grid_T 32
  - ../../../TOOLS/REBUILD_NEMO/rebuild_nemo outputs/ORCA2_5d_00010101_00011231_grid_U 32
  - ../../../TOOLS/REBUILD_NEMO/rebuild_nemo outputs/ORCA2_5d_00010101_00011231_grid_V 32
  - ../../../TOOLS/REBUILD_NEMO/rebuild_nemo outputs/ORCA2_5d_00010101_00011231_icemod 32
- View results
  - view outputs/ORCA2_5d_00010101_00010514_grid_T

Hands On – Results



ORCA2_5d_00010101_00011231_grid_T.nc (a logi...

SST



ORCA2_5d_00010101_00011231_grid_T.nc (a logi...

Wind Speed Module 10m

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Session 3: Visualization and hands on

Objective

- Numerical models produce a huge amount of data on a variety of formats
  - Binary
  - NetCDF
  - ASCII
  - HDF5
  - GRIB
  - ...
- We need tools to analyse and visualise them
- In this section we will introduce some utilities freely available and widely used within the Earth Sciences community
- Many more are available…

Many Packages exists out there

- Visualization platforms
  - NCVIEW
  - PANOPLY
  - GRADS
  - NCL
  - Python Visualization

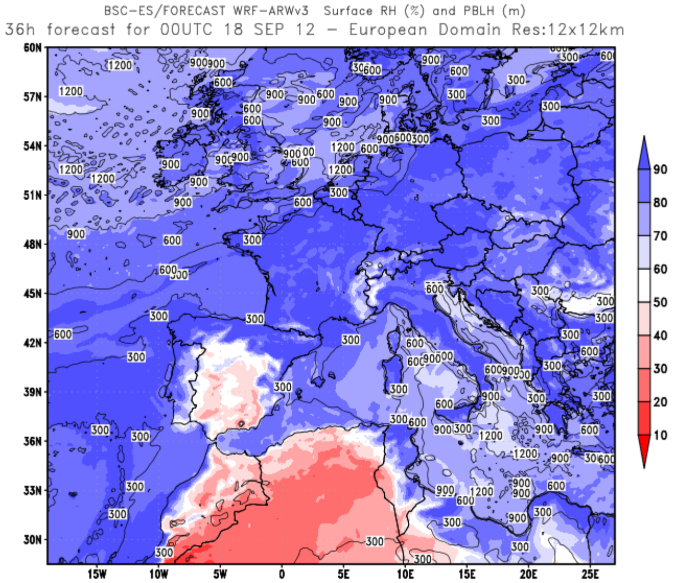
- Analytics packages
  - CDO
  - NCO
  - R
  - NCL

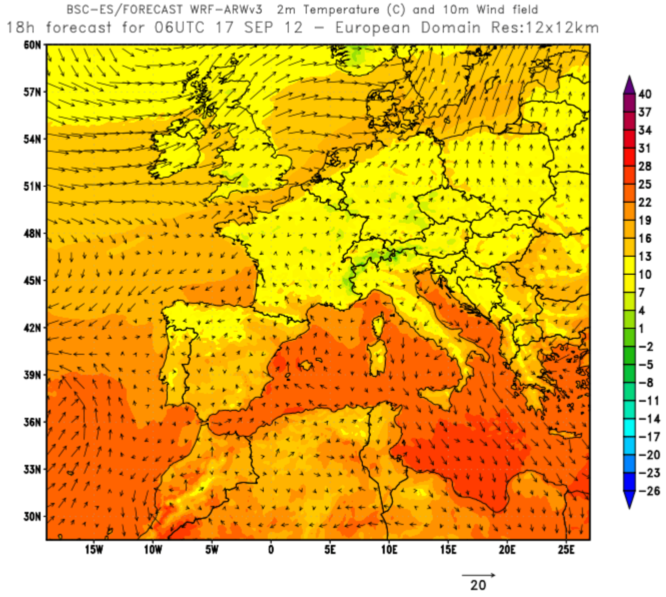
Visualization
GrADS: Grid Analysis and Display System

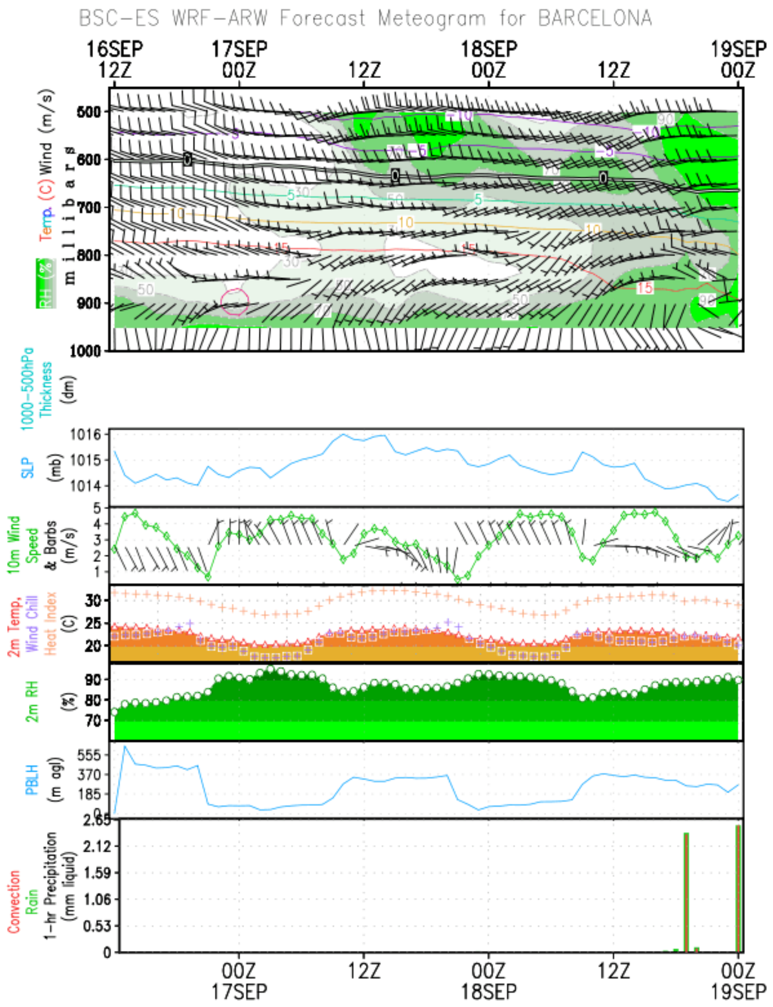
- Interactive desktop tool that is used for easy access, manipulation, and visualization of earth science data.
- Supports many data file formats, including binary (stream or sequential), GRIB (version 1 and 2), NetCDF, HDF (version 4 and 5), and BUFR (for station data).
- Freely distributed over the Internet.
- GrADS handles grids that are regular, non-linearly spaced, gaussian, or of variable resolution.
- Data from different data sets may be graphically overlaid, with correct spatial and time registration.
- Operations are executed interactively by entering FORTRAN-like expressions at the command line or batch commands to generate pictures.
- Data may be displayed using a variety of graphical techniques: line and bar graphs, scatter plots, smoothed contours, shaded contours, streamlines, wind vectors, grid boxes, shaded grid boxes, and station model plots.
- GrADS has a programmable interface (scripting language) that allows for sophisticated analysis and display applications.

- http://www.iges.org/grads/
- Grads tutorial: http://www.iges.org/grads/gadoc/tutorial.html

GRADS Examples


BSC-ES/FORECAST WRF-ARWv3 Surface RH (%) and PBLH (m)
36h forecast for 00UTC 18 SEP 12 - European Domain Res:12x12km
60N
57N
54N
51N
48N
45N
42N
39N
36N
33N
30N
15W
10W
5W
0
5E
10E
15E
20E
25E
90
80
70
60
50
40
30
20
10
1200
900
600
300


BSC-ES/FORECAST WRF-ARWv3 2m Temperature (C) and 10m Wind field
18h forecast for 06UTC 17 SEP 12 - European Domain Res:12x12km
60N
57N
54N
51N
48N
45N
42N
39N
36N
33N
30N
15W
10W
5W
0
5E
10E
15E
20E
25E
40
37
34
31
28
25
22
19
16
13
10
7
4
1
-2
-5
-8
-11
-14
-17
-20
-23
-26
20

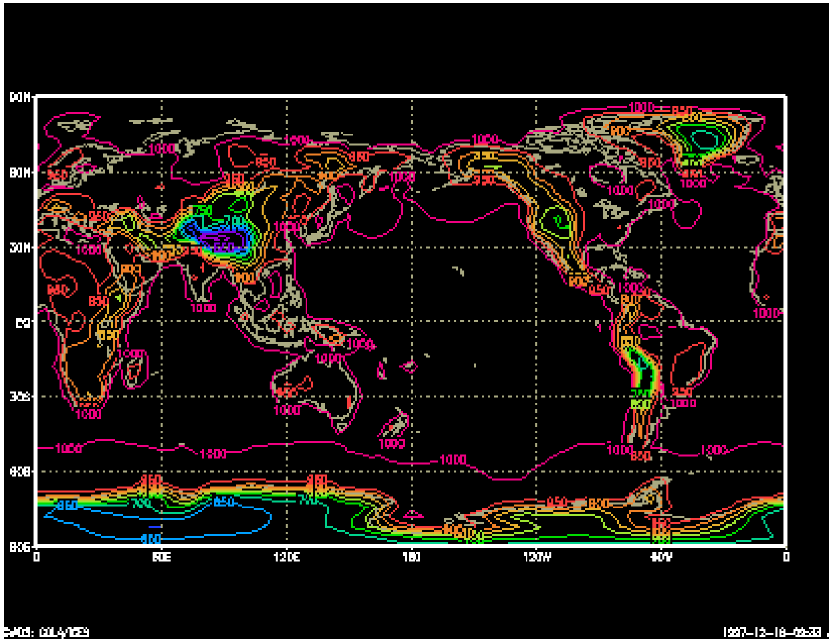


Visualization
GrADS: examples

- We need three files:
  - Binary data: .dat
  - Descriptor file: .ctl
  - Script: .gs

- From the control file a default display can be created using simply :

```
> open model.ctl
> d ps
```



GrADS: examples

- Copy visualization handson/grads
- Analyze *.ctl file

- Execute: grads
- open wrfout….ctl
- View variables, times, domain.
  - q file – extracts file content
  - q dims – provide information on projection
  - d t2
  - Set gxout shaded
  - d t2
  - d skip(u10,14,14);v10

- View the example with T2.
  - Generate picture (grads -blc t2.gs).
  - Open picture (display T2.gif)

- Make a zoom over Iberian Peninsula
  - Use 'set lat 34 45.05'
  - Use 'set lon -12 4.5'

- Modify the example to show T2 in Celsius instead of Kelvin
  - Hint: TºK=TºC + 273.15

- Create a plot with PSFC

- Create a plot with wind vectors u10,v10
  - Hint: use function ⎵ set gxout vector and d skip(u10m,10,10);v10m

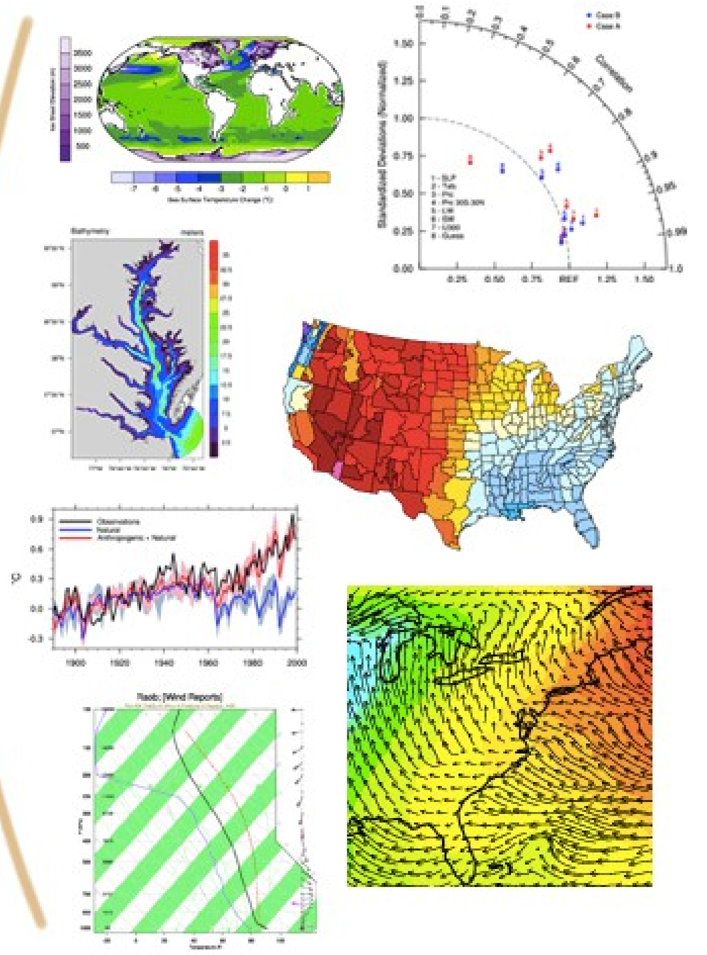
Visualization
NCL: NCAR Command Language

NCL NCAR Command Language NCAR

- Interpreted language designed for scientific data analysis and visualization
- Portable. Available as binaries or open source
- Supports netCDF3/4, GRIB1/2, HDF-SDS, HDF4-EOS, binary, shapefiles, and ascii files
- Numerous analysis functions built-in
- High quality graphics
- Many example scripts

- http://www.ncl.ucar.edu/

ble



©2012 UCAR | Privacy Policy | Terms of Use | Contact the Webmaster | Sponsored by NSF

NCL: scripts

- Large number of examples at the website

- http://www.ncl.ucar.edu/Applications/
  - Ready for several datasets
  - Different Map projections
  - Examples for specific models
  - Allow data analysis

- http://www.ncl.ucar.edu/Applications/wrf.shtml
  - Examples for WRF outputs

NCL: examples

- Try some examples:

```
/home/nct/nct00002/PATC-Course/visualization-hands_on/NCL

# NCL
export NCARG_ROOT=/gpfs/apps/MN3/NCL/6.1.2
export PATH=$PATH:/gpfs/apps/MN3/NCL/6.1.2/bin
```

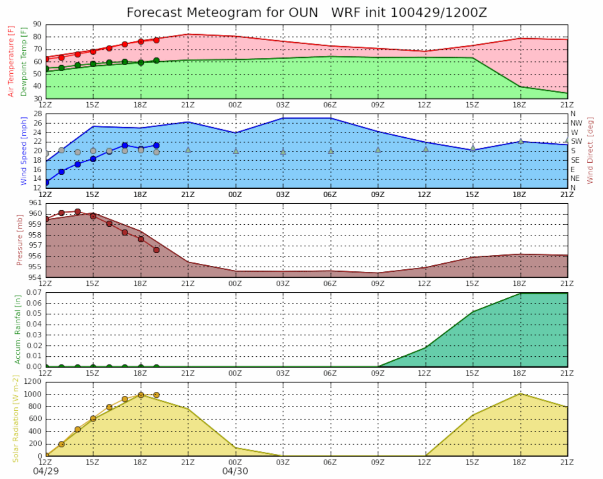
Execute:

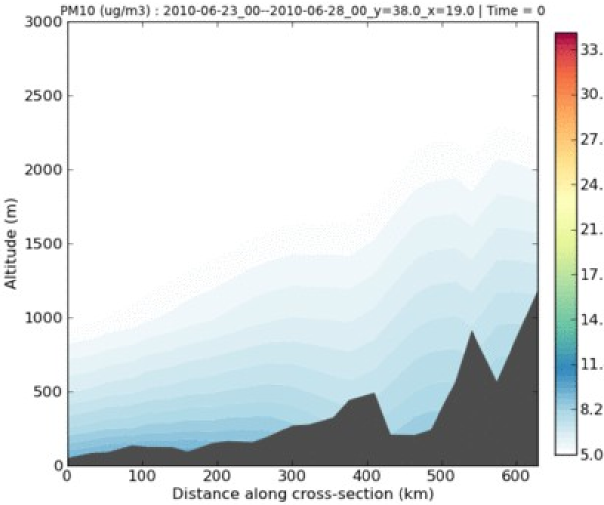
```
ncl wrf_script.ncl
```

Python Map Generator

- Using Python language to plot maps and time series

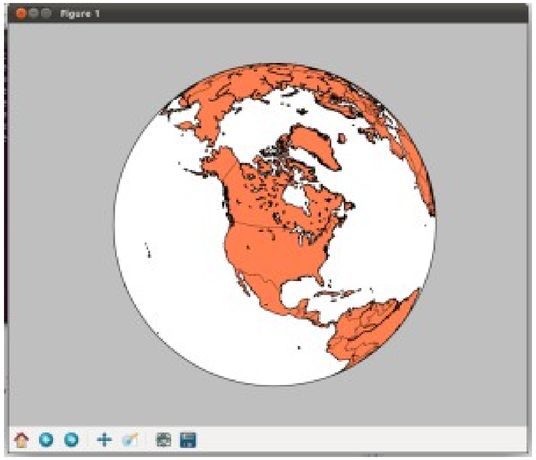
- Requirements
  - Python
  - Matplotlib 1.0 or greater
  - Basemap 0.9 or greater
  - Nio or scipy/netcdf (some Python library for reading in netCDF files is necessary)

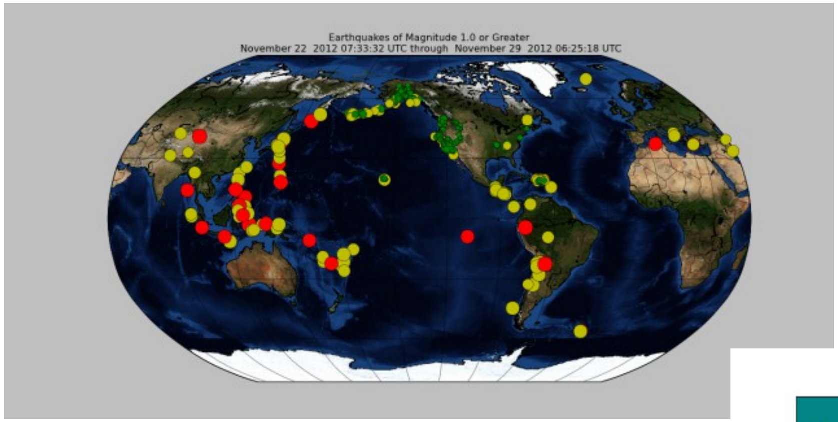


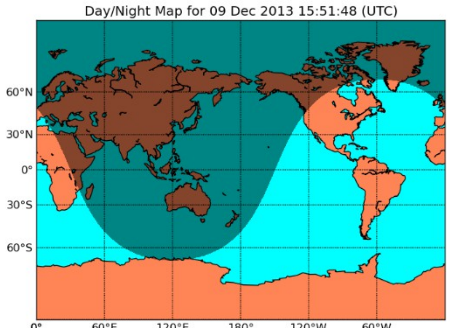


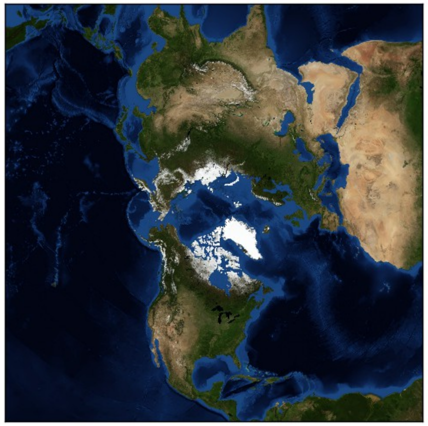
Python Map Generator

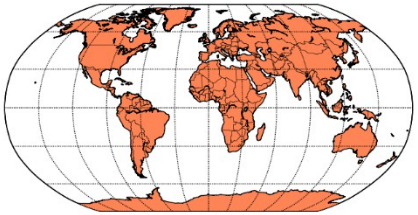
- Many options in map backgrounds and projections




Earthquakes of Magnitude 1.0 or Greater
November 22 2012 07:33:32 UTC through November 29 2012 06:25:18 UTC



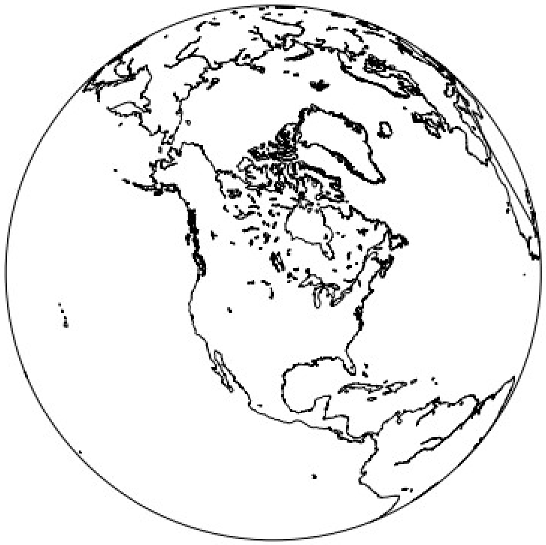




Python Map Generator

- To run a script
  - python script.py

```
from mpl_toolkits.basemap import Basemap
import matplotlib.pyplot as plt
import numpy as np

map = Basemap(projection='ortho', lat_0=50, lon_0=-100,
              resolution='l', area_thresh=1000.0)

map.drawcoastlines()

plt.show()
```



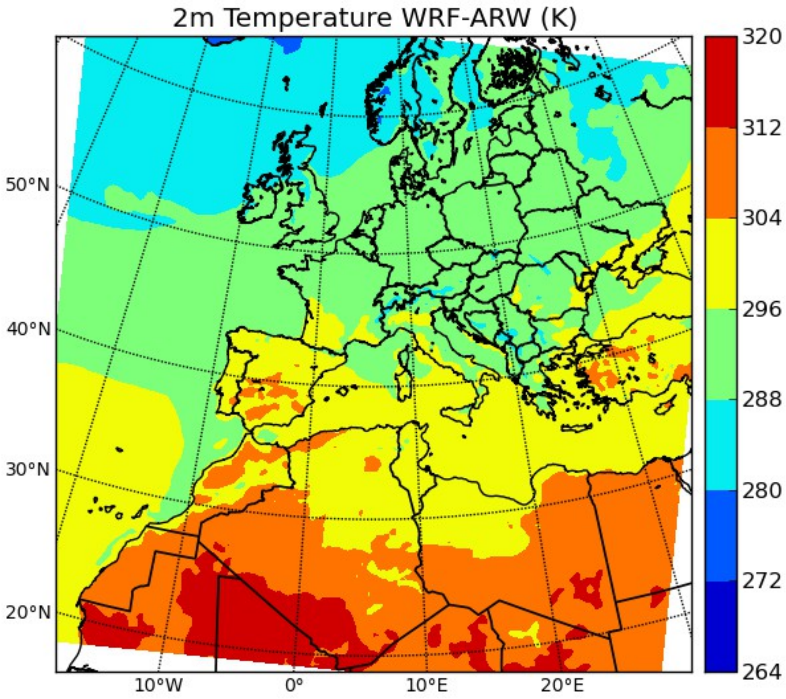
Python Map Generator Example

- Try some examples:
  - /home/nct/nct00002/PATC-Course/visualization-hands_on/MapGenerator

- To run the WRF example
  - python display_wrf_T2.py
  - display wrf_T2.png

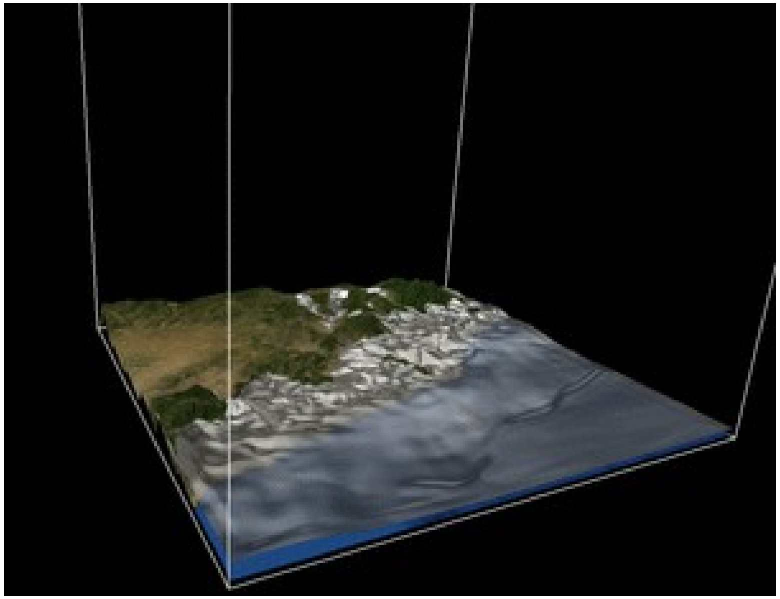
- Modify the example to show Pressure (PSFC)

Plot the third hour of the file.



3D Visualization

- 3D Visualization is a nice way to show and understand results
- Many options to make renders
  - Free software
    - Vapor
    - Visit
    - Paraview
    - IDV
  - Commercial software
    - Maya
    - Avizo Green



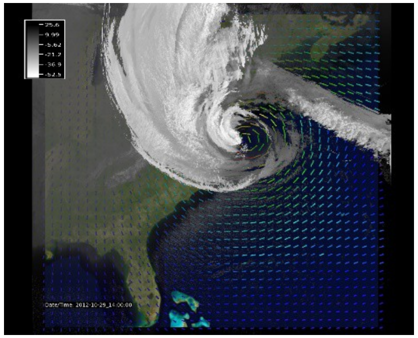
- Be careful, usually big amount of data is needed to get nice volumes (big amount of vertical layers).

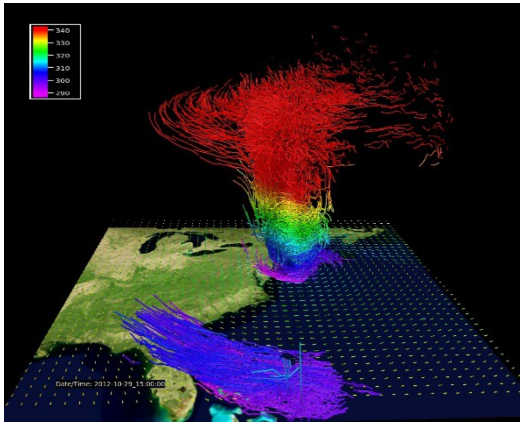
3D Visualization

- Data needs to be in high resolution
- Sometimes format is a mess
  - Some programs don’t understand sigma layers
  - Others needs coordinates in a specific format
- Conversion with NETCDF operators is sometimes required
- A powerful computer with a powerful graphic card (NVIDIA Quadro) is required

3D Visualization - Vapor

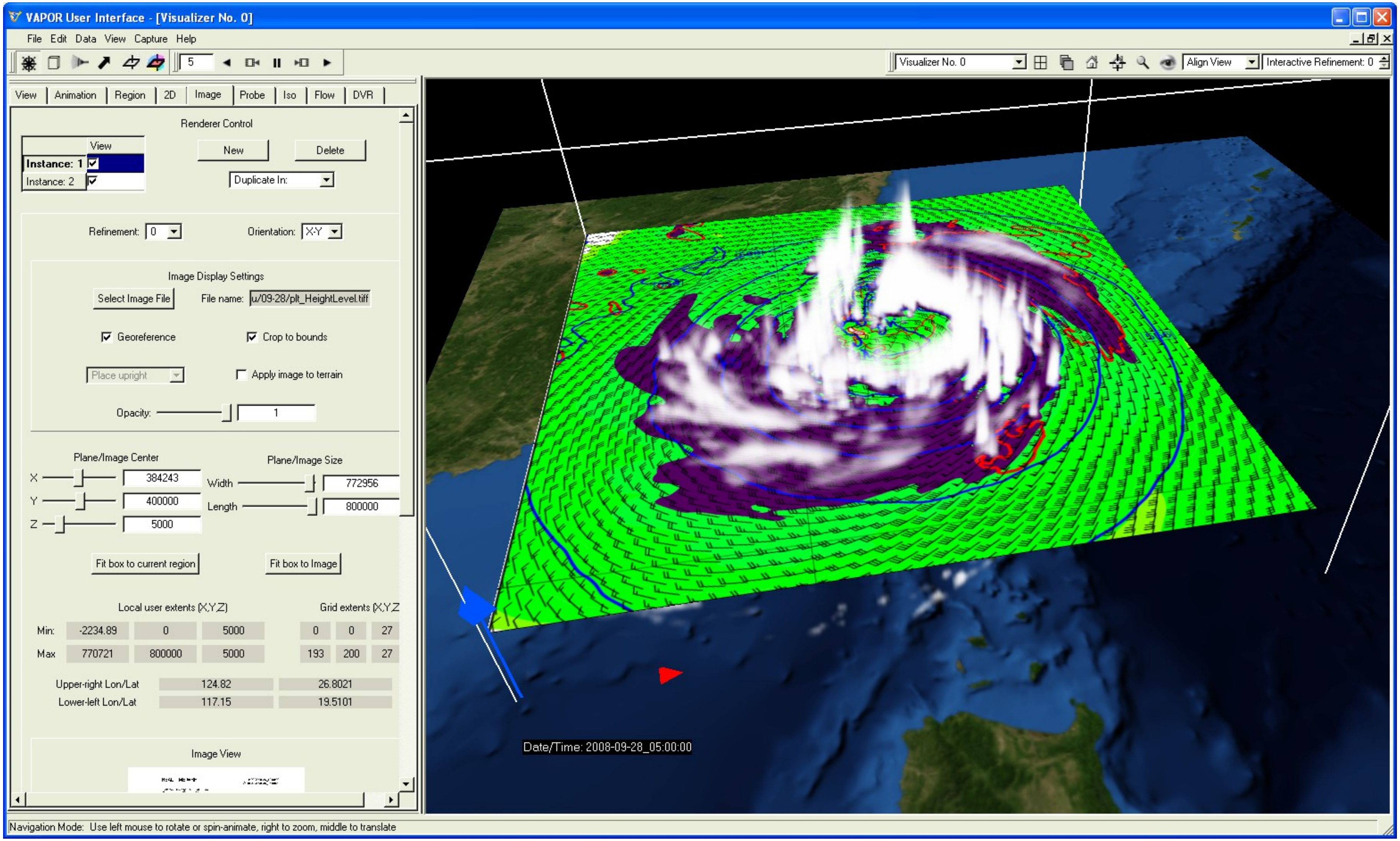
- Vapor is a free software designed to show 3D outputs
- Features
  - A visual data discovery environment tailored towards the specialized needs of the astro and geosciences CFD community
  - A desktop solution capable of handling terascale size data sets
  - Advanced interactive 3D visualization tightly coupled with quantitative data analysis
  - Support for multi-variate, time-varying data
  - Integrated with Python.
  - Support for 3D visualization of WRF-ARW datasets
- http://www.vapor.ucar.edu/





Barcelona Supercomputing Center
Centro Nacional de Supercomputación

3D Visualization - Vapor



VAPOR User Interface - [Visualizer No. 0]
File Edit Data View Capture Help
Visualizer No. 0
Align View
View Animation Region 2D Image Probe Iso Flow DVR
New
Delete
Duplicate In:
Select Image File
Fit box to current region
Fit box to Image
Image View

Interactive Refinement: 0

Renderer Control

View	
Instance: 1	☑
Instance: 2	☑

Refinement: 0 Orientation: X-Y

Image Display Settings

File name: u/09-28/plt_HeightLevel.tiff

☑ Georeference ☑ Crop to bounds

Place upright □ Apply image to terrain

Opacity: 1

Plane/Image Center

X 384243
Y 400000
Z 5000

Plane/Image Size

Width 772956
Length 800000

	Local user extents (X,Y,Z)			Grid extents (X,Y,Z		
Min:	-2234.89	0	5000	0	0	27
Max	770721	800000	5000	193	200	27

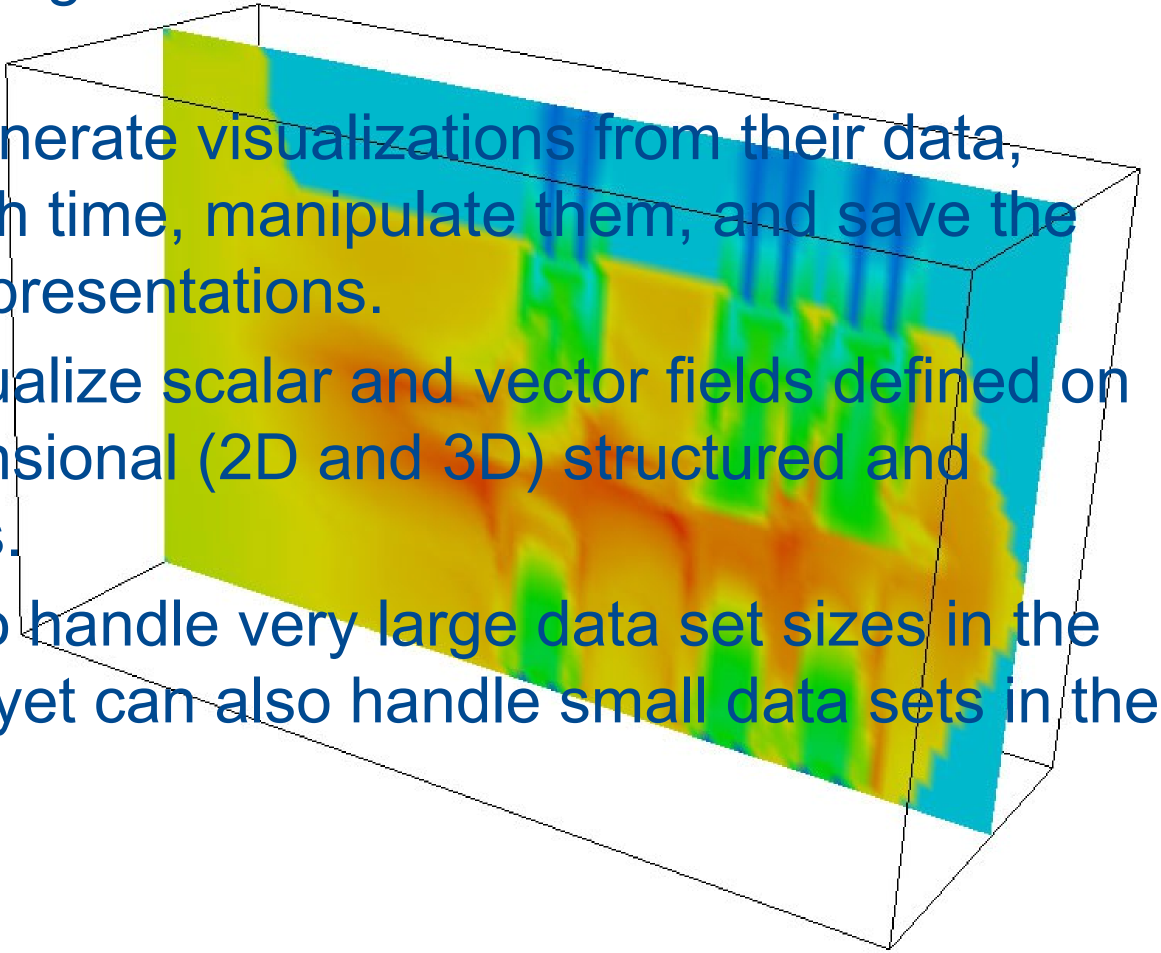
Upper-right Lon/Lat	124.82	26.8021
Lower-left Lon/Lat	117.15	19.5101

Date/Time: 2008-09-28_05:00:00

Navigation Mode: Use left mouse to rotate or spin-animate, right to zoom, middle to translate

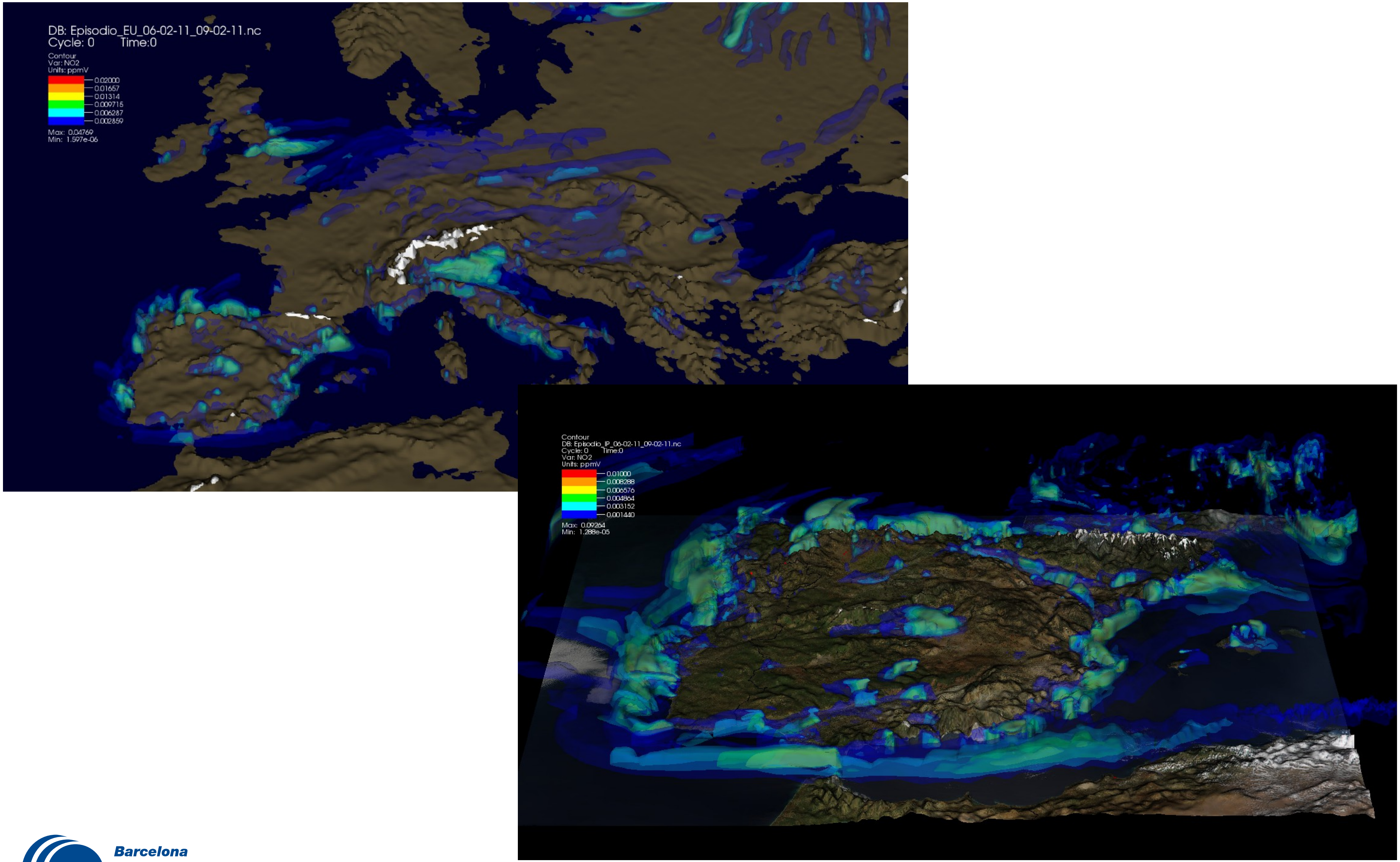
VisIt Software

- VisIt is a free interactive parallel visualization and graphical analysis tool for viewing scientific data on Unix and PC platforms.
- Users can quickly generate visualizations from their data, animate them through time, manipulate them, and save the resulting images for presentations.
- It can be used to visualize scalar and vector fields defined on two- and three-dimensional (2D and 3D) structured and unstructured meshes.
- VisIt was designed to handle very large data set sizes in the terascale range and yet can also handle small data sets in the kilobyte range.

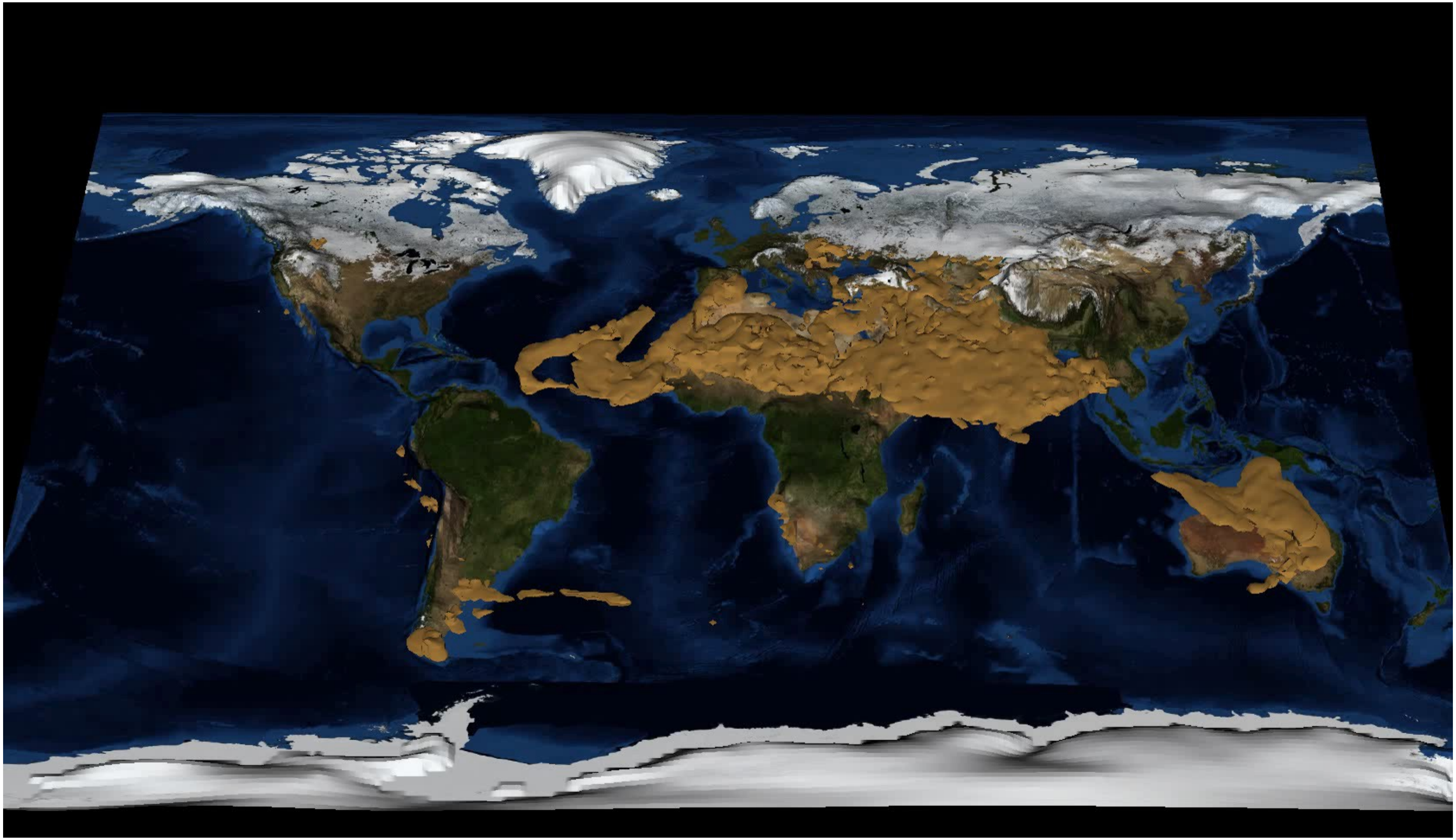


Barcelona Supercomputing Center
Centro Nacional de Supercomputación

VisIt with CMAQ CTM Model



VisIt with NMMB-BSC Dust



Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Session 3: analysis packages and hands on

Objective

- Numerical models produce a huge amount of data on a variety of formats
  - Binary
  - NetCDF
  - ASCII
  - HDF5
  - GRIB
  - ...

- We need tools to analyse and visualise them

- In this section we will introduce some utilities freely available and widely used within the Earth Sciences community

- Many more are available…

Many Packages exists out there

- Visualization platforms
  - NCVIEW
  - PANOPLY
  - GRADS
  - NCL
  - MapGenerator

- Analytics packages
  - CDO
  - NCO
  - R
  - NCL

Analytics
CDO: Climate Data Operators

- Collection of command line operators to manipulate and analyse Climate and NWP model data

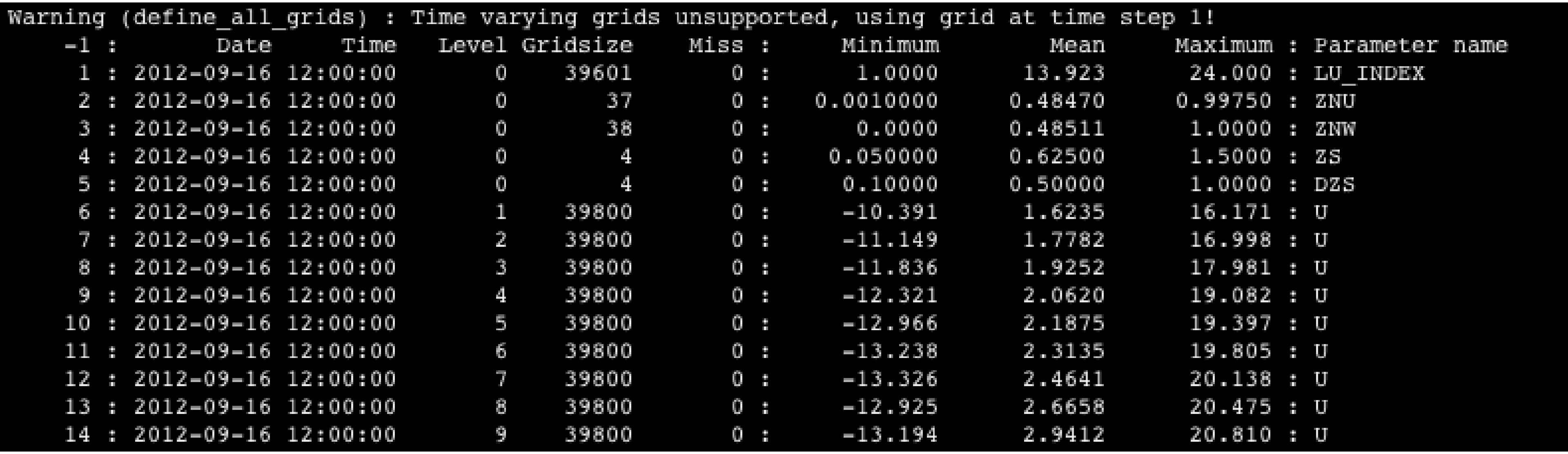
- Supports data formats GRIB1/2, netCDF 3/4, EXTRA and IEG

- More than 600 operators available
  - Information on datasets
  - Selection of specific data from a file
  - Comparison
  - Modification of attributes, names, variables
  - Arithmetic operations
  - Statistical values
  - Regression
  - Interpolation
  - Transformations

- It runs on Linux, Unix, MAC OS, Windows

CDO: examples

- http://www.nco.ncep.noaa.gov/pmb/codes/nwprod/sorc/rtofs_cdo-1.4.0.1.fd/cdo-1.5.0/doc/cdo_refcard.pdf
- Syntax: cdo Operator1 [-Operator2] filein.nc fileout.nc
- Example: cdo info wrfout_d01_2012-09-16_12:00:00



```
Warning (define_all_grids) : Time varying grids unsupported, using grid at time step 1!
   -1 :       Date     Time   Level Gridsize    Miss :     Minimum        Mean     Maximum : Parameter name
    1 : 2012-09-16 12:00:00       0    39601       0 :      1.0000      13.923      24.000 : LU_INDEX
    2 : 2012-09-16 12:00:00       0       37       0 :   0.0010000     0.48470     0.99750 : ZNU
    3 : 2012-09-16 12:00:00       0       38       0 :      0.0000     0.48511      1.0000 : ZNW
    4 : 2012-09-16 12:00:00       0        4       0 :    0.050000     0.62500      1.5000 : ZS
    5 : 2012-09-16 12:00:00       0        4       0 :     0.10000     0.50000      1.0000 : DZS
    6 : 2012-09-16 12:00:00       1    39800       0 :     -10.391      1.6235      16.171 : U
    7 : 2012-09-16 12:00:00       2    39800       0 :     -11.149      1.7782      16.998 : U
    8 : 2012-09-16 12:00:00       3    39800       0 :     -11.836      1.9252      17.981 : U
    9 : 2012-09-16 12:00:00       4    39800       0 :     -12.321      2.0620      19.082 : U
   10 : 2012-09-16 12:00:00       5    39800       0 :     -12.966      2.1875      19.397 : U
   11 : 2012-09-16 12:00:00       6    39800       0 :     -13.238      2.3135      19.805 : U
   12 : 2012-09-16 12:00:00       7    39800       0 :     -13.326      2.4641      20.138 : U
   13 : 2012-09-16 12:00:00       8    39800       0 :     -12.925      2.6658      20.475 : U
   14 : 2012-09-16 12:00:00       9    39800       0 :     -13.194      2.9412      20.810 : U
```

CDO: examples

- Select a variable from a netcdf file (selvar,*var*):

```
cdo selvar,T2 wrfout_d01_2012-09-16_12:00:00 T2.nc
```

- Compute the temporal mean of a variable (tim*mean*):

```
cdo timmean T2.nc T2timmean.nc
```

- Compute the spatial mean over a domain of a variable (fld*mean*):

```
cdo output –fldmean T2timmean.nc
```

  - Execute the above three example at once (Operator1 –Operator2 -…):

```
cdo selvar,T2 –timmean –fldmean wrfout_d01_2012-09-16_12:00:00 out.nc
```

CDO: examples

- Apply a mask over a field:
- Select mask field from file:

```
cdo selvar,LANDMASK wrfout_d01_2012-09-16_12:00:00 mask.nc
```

  - Apply the mask over temperature at 2 m (multiply – 1land 0sea):

```
cdo mul T2.nc mask.nc T2_land.nc
```

  - Set a constant as missing value:

```
cdo setctomiss,0 T2_land.nc T2_land.nc
```

  - Compute the field mean over land:

```
cdo output –fldmean T2_land.nc
```

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

CDO: practice

- From WRF Exercice-1 output:
  - Compute the maximum temperature at 2m (T2) at 18UTC 16/9/2012
  - Compute total precipitation (RAINNC) over land
    - Compute wind speed at 10m (U10, V10)
    - Compute maximum wind speed at 10th model layer (U,V)
  - Useful operators:
    - cdo expr,’VAR1=VAR2+3;’ filein.nc fileout.nc
    - seltimestep, selvar, sellevel, fldmax

CDO: practice

From WRF Exercice-1 output:

- Compute the maximum temperature at 2m (T2) at 18UTC 16/9/2012

```
cdo fldmax –seltimestep,7 -selvar,T2 wrfout.nc MAXT2.nc
```

- Compute total precipitation (RAINNC) over land

```
cdo seltimestep,7 –selvar,RAINNC wrfout.nc RAINNC.nc
cdo selvar,LANDMASK wrfout.nc MASK.nc
cdo mul RAINNC.nc MASK.nc RAIN_LAND.nc
cdo output –fldsum RAIN_LAND.nc
```

- Compute wind speed at 10m (U10, V10)

```
cdo selvar,U10,V10 wrfout.nc out1.nc
cdo expr,’spd=sqrt(U10*U10+V10*V10);’ out1.nc SPD.nc
```

- Compute maximum wind speed at 10th model layer (U,V)

```
cdo selvar,U,V –sellevel,10 wrfout.nc out1.nc
cdo expr,’spd=sqrt(U*U+V*V);’ out1.nc SPD.nc
```

Analytics
NCO: netCDF Operator

- Comprise a dozen standalone, command-line programs that operates netCDF files
- Several operations (e.g., derive new data, average, print, hyperslab, manipulate metadata)
- Output results to screen or files in text, binary, or netCDF formats
- NCO aids manipulation and analysis of gridded scientific data
- The shell-command style of NCO allows users to manipulate and analyze files interactively, or with simple scripts
- Available programs:
  - ncap2 netCDF Arithmetic Processor
  - ncatted netCDF ATTribute Editor
  - ncbo netCDF Binary Operator (includes ncadd, ncsubtract, ncmultiply, ncdivide)
  - ncea netCDF Ensemble Averager
  - ncecat netCDF Ensemble conCATenator
  - ncflint netCDF FiLe INTerpolator
  - ncks netCDF Kitchen Sink
  - ncpdq netCDF Permute Dimensions Quickly, Pack Data Quietly
  - ncra netCDF Record Averager
  - ncrcat netCDF Record conCATenator
  - ncrename netCDF RENAMEer
  - ncwa netCDF Weighted Averager

- http://nco.sourceforge.net/

netCDF Operators (NCO) Software Stack

NCO: examples

- Select first level of a file
  - ncks -d bottom_top,1 wrfout.nc out.nc

- Select 6 hours of a file
  - ncks -d Time,1,6 wrfout.nc out.nc

- Calculate T at 2m in celsius.
  - ncap -s "T2=T2-273.15" wrfout.nc celsius.nc

- Change variable name
  - ncrename -h -O -v T2,T2_celsius celsius.nc

- Subsetting a region
  - ncea -d west_east,min_gridpoint,max_gridpoint -d lon, ,min_gridpoint,max_gridpoint in.nc out.nc
  - IN WRF: Iberian Peninsula: ncea -d west_east,45,95 -d south_north,64,110 celsius.nc iberian.nc

- http://jisao.washington.edu/data/nco/#example1

Analytics
R: the R Project for Statistical Computing

- Free software environment for statistical computing and graphics
- It includes:
  - an effective data handling and storage facility,
  - a suite of operators for calculations on arrays, in particular matrices,
  - a large, coherent, integrated collection of intermediate tools for data analysis,
  - graphical facilities for data analysis and display either on-screen or on hardcopy, and
  - a well-developed, simple and effective programming language which includes conditionals, loops, user-defined recursive functions and input and output facilities.

- It runs on UNIX, Linux, Mac OS, Windows
- http://www.r-project.org/

R: examples

```
############
#  Read, operate and plot wrf data
############

library(RNetCDF) #open.nc, var.get.nc

wrf<-open.nc('wrfout_d01.nc')

timeWRF<-var.get.nc(wrf,"Times")
latWRF <- var.get.nc(wrf,"XLAT")
lonWRF <- var.get.nc(wrf,"XLONG")

u10 <- var.get.nc(wrf,"U10")
v10 <- var.get.nc(wrf,"V10")

dim(u10)

u10[2,2,]

spd<-sqrt(u10*u10+v10*v10)
dir<-atan2(u10,v10)*57.2957795+180.
dir[spd<0.5]<-NA
```

Arrays: var[x,y,z]
Vectors: var[x]

Create a variable: var<-value
var<-c(1,2,3,4)

Variable dimension: dim(var)

Print variable value: var[10:15]

R: examples

```
# Plots: scatter plot, time evolution, histogram, boxplots
par(mfrow=c(3,2))
plot(spd[40,40,],spd[80,80,], main="Scatterplot of spd at 40,40 vs. spd at 80,80")
plot(dir[40,40,],main="Direction time evolution",xlab="Timestep axis")
hist(spd[,,5], main="Histogram of spd")
boxplot(spd[,20,1:4], main="Boxplot of spd")
```

```
# Image plots
image(spd[,,3])
```

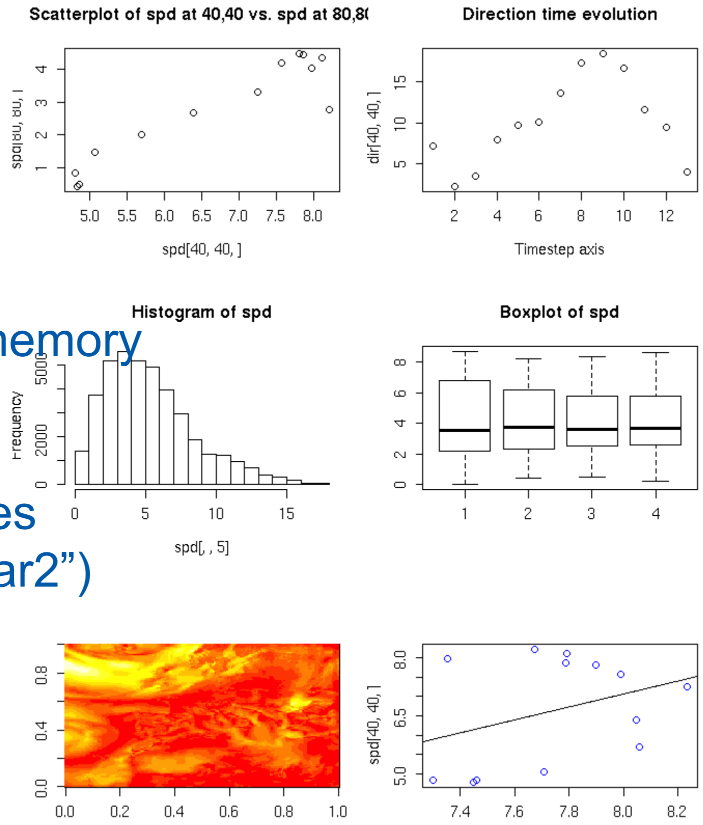
```
# Fit line to scatter plot
plot(spd[20,20,],spd[40,40,],col="blue")
myline.fit <- lm(spd[40,40,] ~ spd[20,20,])
abline(myline.fit)
```

```
# Statistics
cor(spd[20,20,1:7],spd[40,40,1:7])
mean(spd[20,20,7])
```

```
# Variables on memory
ls()
```

```
# Delete variables
rm(list=“var1”,”var2”)
```

```
# Close session
quit()
```



Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Session 4: free hands-on

Practice on HPC environment, application, visualization and analysis

- Exercise 1: Run a WRF simulation of 12h with 24 cpus
- Exercise 2: Create an animation over time of T2 with GrADS with model output of Exercise 1
- Exercise 3: compute 12h-average temperature from model simulation of Exercise 1
- Exercise 4: find the total amount of precipitation over the ocean from Exercise 1
- Exercise 5: compute the correlation of T at the first and fifth model layer from Exercise 1 output

www.bsc.es

Barcelona Supercomputing Center

Centro Nacional de Supercomputación

Thank you!

For further information please contact

oriol.jorba@bsc.es

georgios.markomanolis@bsc.es

kim.serradell@bsc.es

Fill the questionnaire of the course:

http://events.prace-ri.eu/confDisplayEvaluation.py/display?confId=194