HPC features of CNRM-CM

S. Sénési, J-P. Piedelièvre – CNRM-GAME (Météo-France and CNRS)

E. Maisonnave - CERFACS

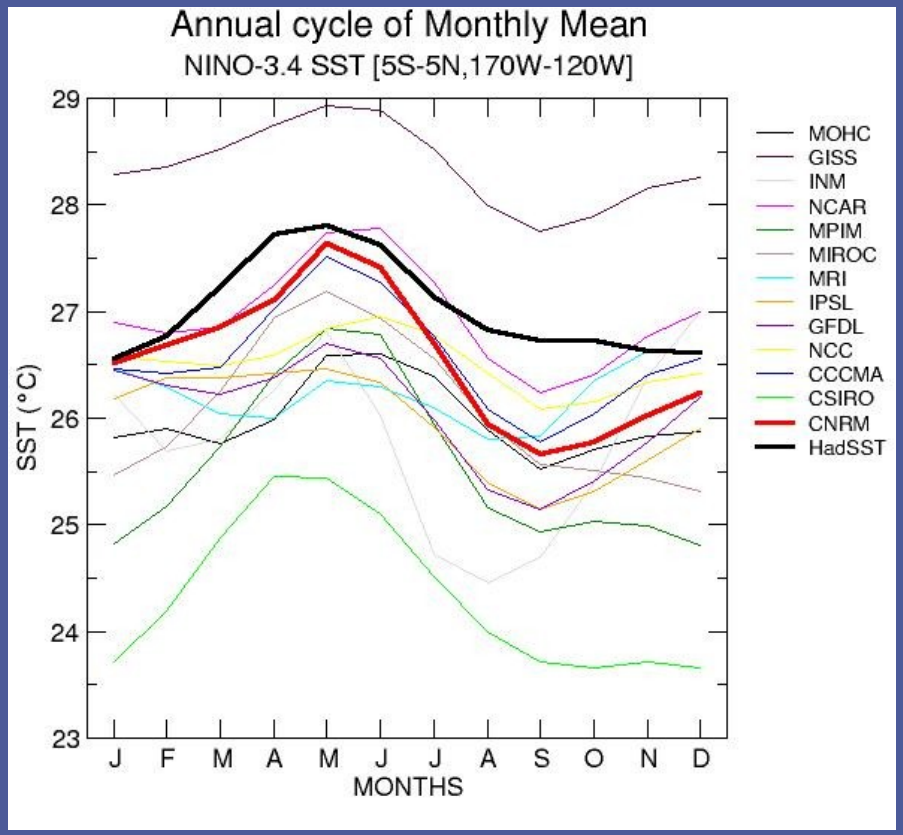
IS-ENES HPC Workshop – Toulouse- Jan 2013

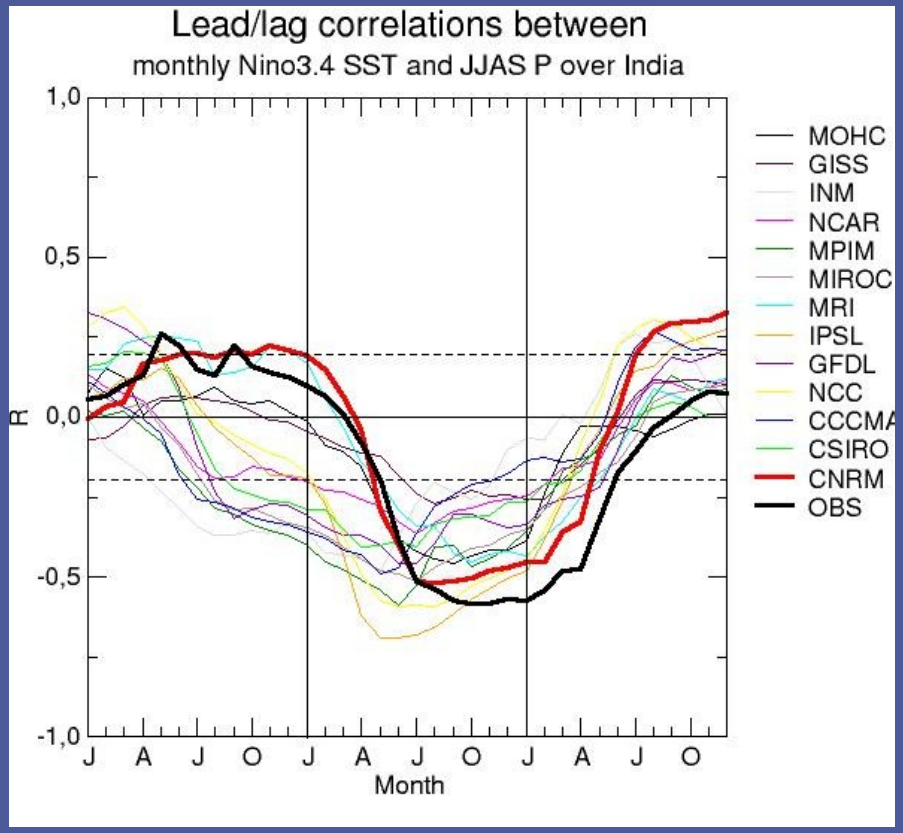
METEO FRANCE

Two words of science

CNRM-CM5 provided CMIP5 with a comprehensive set of useful experiments.

ENSO and teleconnections examples :





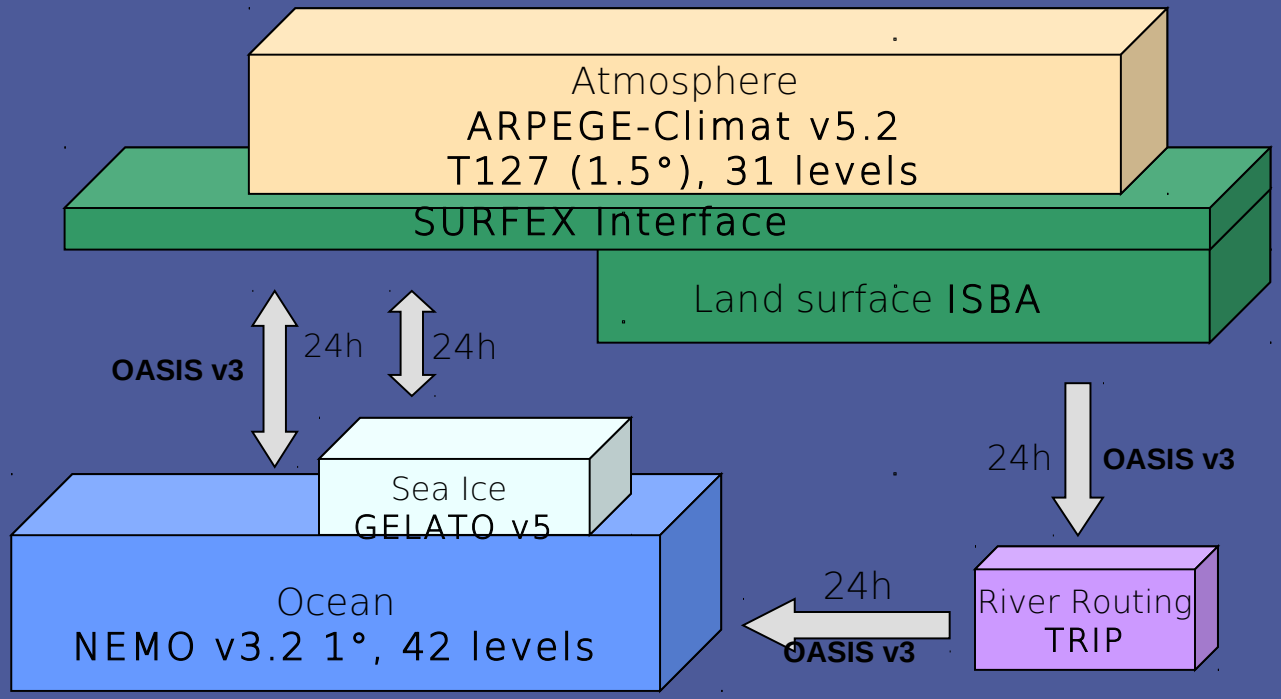
Douville, p.c.

METEO FRANCE

CNRM-CM basics

A long lasting cooperation between CNRM-GAME and CERFACS

- V1 : Arpege (atm) coupled to OPA (ocean). (Terray et al., Cli. Dyn. 1995)
- V1b : includes Oasis Coupler. Took part in CMIP2
- V2 , 1999 : includes sea-ice model GELATO - 1999
- V3 , 2004 : includes river routing scheme TRIP. Took part in CMIP3
- V5 , 2010 : includes Surface scheme SURFEX (incl LSM). Took part in CMIP5. (Voldoire et al., Cli. Dyn., 2012)



Voldoire, 2011

METEO FRANCE

Atmosphere model ARPEGE-Climat

ECMWF 's IFS (the Integrated Forecast System)

- used in EC-EARTH
- **Spectral, gaussian reduced grid**
- **Semi-lagrangian advection scheme**

Arpege based on IFS

- adds a specific physics and stretched grid
- Used for NWP operations at Météo-France, at high resolution : T1198
  - LAM version Aladin

Arpege-Climat : Climate version for CNRM-CM

- climate-specific physics and Land Surface Scheme
- Used in EUROSIP (ensemble seasonal forecast , with ECMWF, NCEP, UK MetOffice)
- V6 based on IFS Cy37

Designed for MPI et OpenMP

- **2 axis MPI distribution both in spectral space (waves, level) and in physical space (lat / lon)**
- **Optimization of the 2D partitionning in physical space**
- **I/O : 1 axis MPI distribution ; proprietary output format ; single output file per snapshot**
- **Validated at ECMWF and in PRACE 1 IP for thousands of procs at high resolution**

Vector design : full control over leading dimension of all arrays in physical space, whatever the problem size

Typical resolution in CNRM-CM : T127 or T159; L31 to L70 - Single configuration for CMIP5 decadal and centennal

Typical time steps for T127: 30 min for L31; 10 min for L70

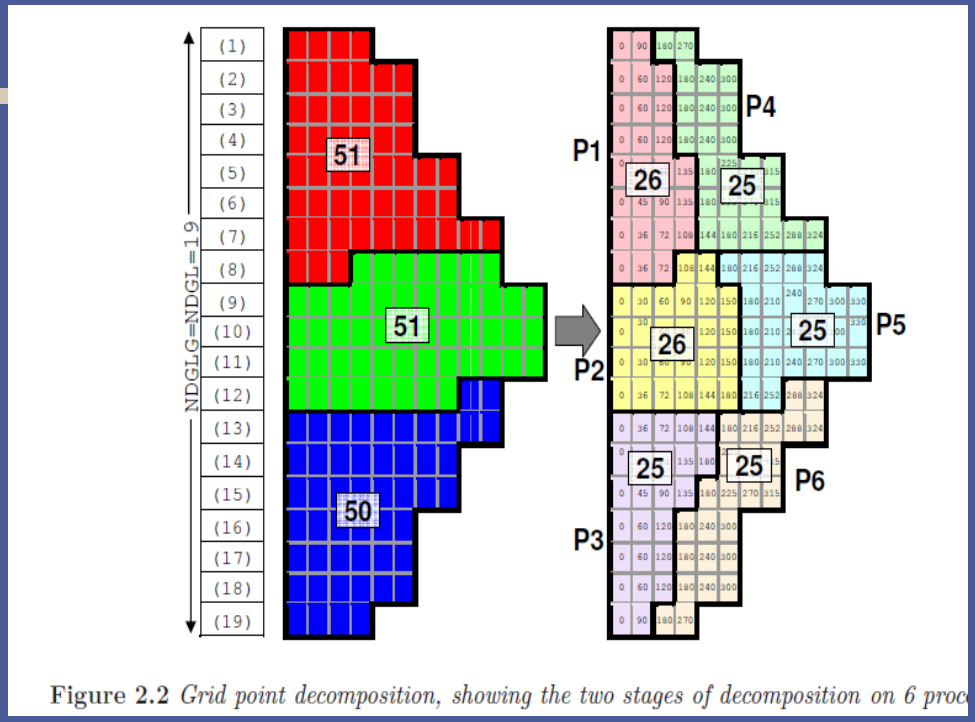
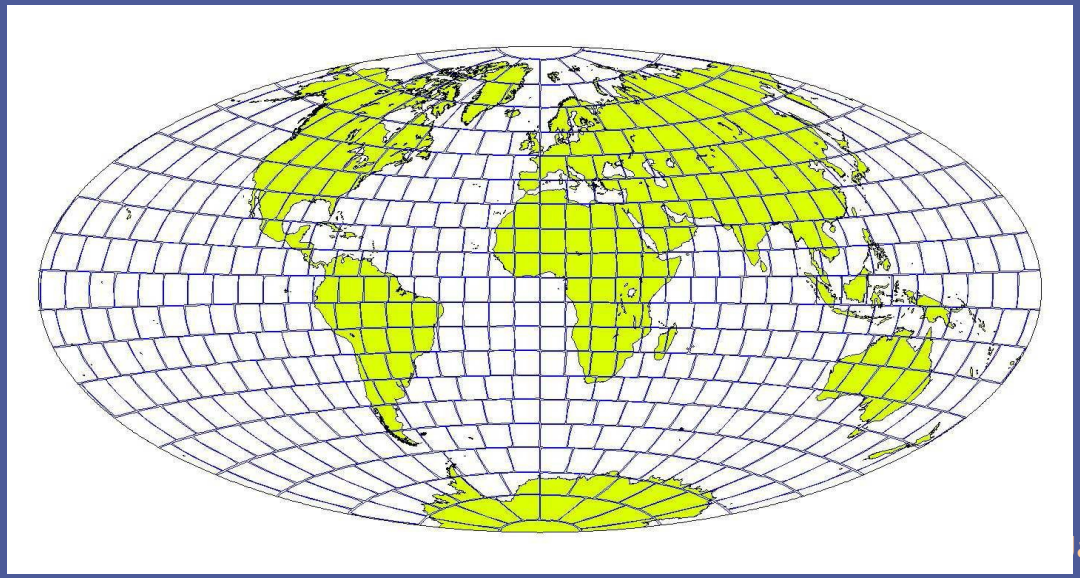


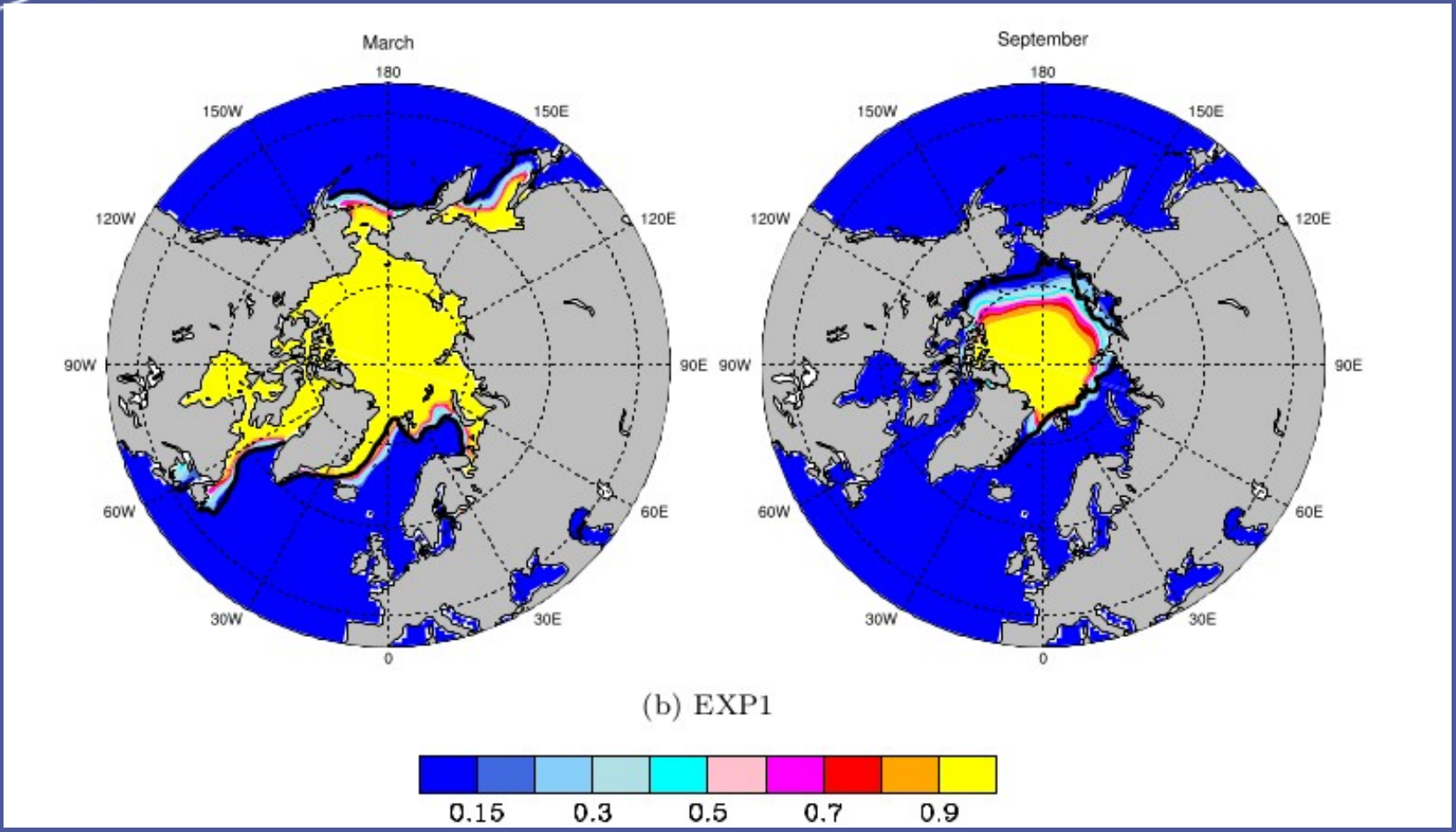
Figure 2.2 *Grid point decomposition, showing the two stages of decomposition on 6 proc*



IFS documentation
Cy37r2

METEO FRANCE

Sea Ice model Gelato – Sea ice cover

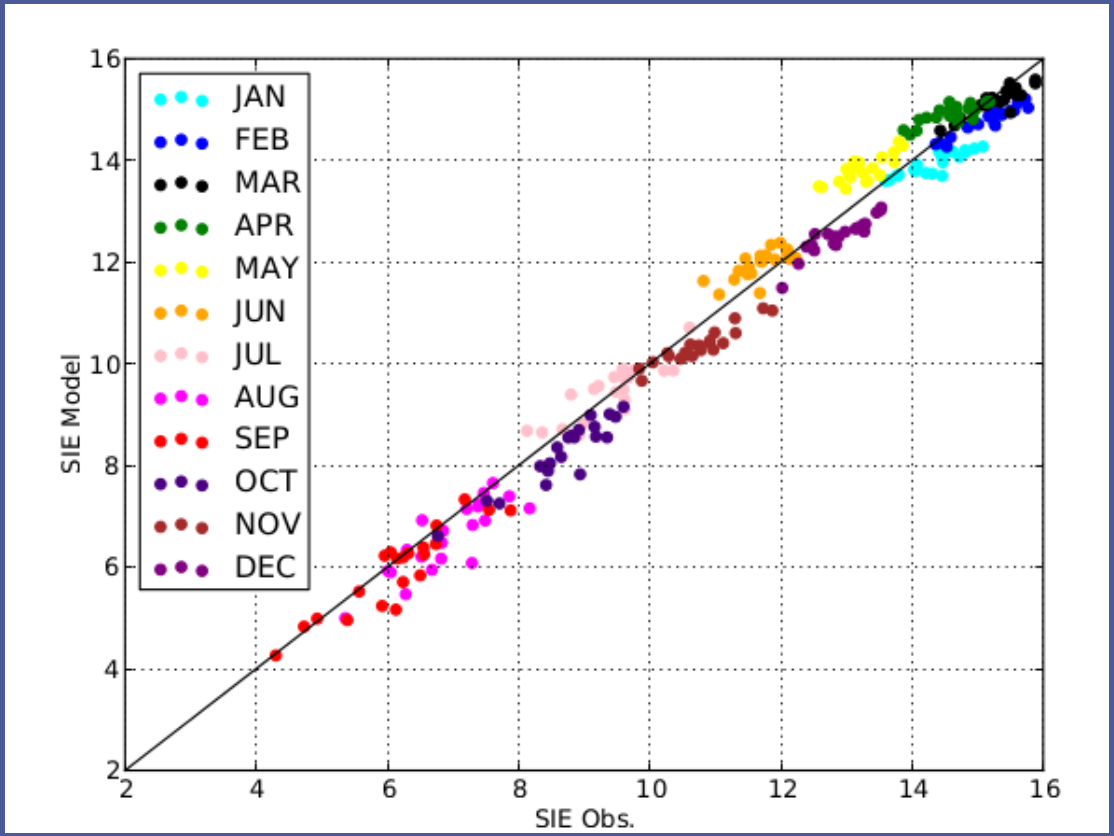


(b) EXP1

Chevallier, 2012

METEO FRANCE

Sea Ice model Gelato – Sea ice extent variability



Chevallier, 2012

METEO FRANCE

Sea Ice model Gelato

Salas y Melia (2002)

- Multiple ice depth categories.
- Thermodynamics : Cp=Cp(T,S) -> expensive iterative solving of heat transfer
- Transport and dynamics :
  - Now: Elastic Viscous Plastic rheology, semi-lagrangian
  - 2014 : JFNK dynamics (Jacobian Free Newton Krylov sheme, McGill) -> 1D

Embedded in Nemo - Uses host model grid

Distribution : MPI only

- **2010 - CMIP5 : thermodynamics only (50% CPU)**
- **2012 : dynamics and transport**
- **2013 : inherits host model distribution**

I/O : by master MPI task; yet proprietary formats

Typical time steps : 1 day for thermodynamics, 6h for dynamics (with 100 iterations)

Integration in Surfex under development - allow for more realistic atmospheric-forced experiments (1D framework, with OML)

Other components

Nemo : V 3.2 – Efficient MPI distribution – Using 1st version of IO server in CMIP5

Surface Scheme SURFEX : Land-Surface scheme ISBA + Sea flux scheme ECUME

- Involved in all major LSMIPs
- Used in operational NWP LAMs and in research meso-scale model MESO-NH
- Here embedded (linked) in Arpege. Same grid
- **Basically 1D**
- **Distribution :**
  - **Inherited from Arpege physical space distribution**
  - **OpenMP under optimization**
- **I/O by Arpege**

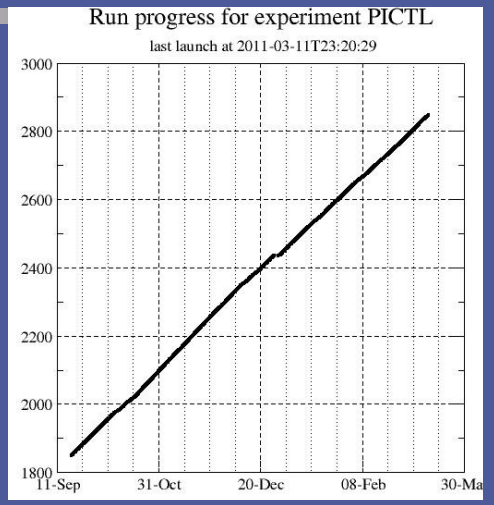
River routing : TRIP (Oki and Sud, 1998)

- 1° grid – simple physics - small cost – no performance issue

Coupler : OASIS3

- Ensures synchronization + interpolation and exchange of fields between models
- Includes global conservation of fields
- Coupling time step 24h
- Single process – Small cost - No performance issue

CNRM-CM production rates



CMIP5 production on **NEC SX8 tori : 2 nodes * (8 procs + 60 GB)**

- 9 procs for atm+land / 6 procs for ocean + sea-ice
- 1 proc for ocean I/O server + coupler + river routing

– PiControl on dedicated nodes : 1000 years ~ 6 month - or **5.75 y/d**.

– hardware and system reliability

COMBINE project production on non-dedicated nodes

– **NEC SX8 tori : 1 node** - 8 procs (5 atm : **peak 4.75 y/d**

– **IBM Power 6** vargas@IDRIS : 32 procs per node, 4 GB/proc

- CMIP5 settings : **1 node - peak 3.6 y/d**
- Going to 60 levels: time-step 30' -> 10' for stability - **2 nodes** - **peak 3 y/d**
- Adding the chemical scheme : **peak 0.8 y/d**

Research Version

– **IBM Power 7** c2a@ECMWF : 64 procs per node, 1GB/proc

- T159 L62 PCMT physics dt=15' : 1 node (oce : 4 procs), 40 GB : **peak 3y/d**

METEO FRANCE

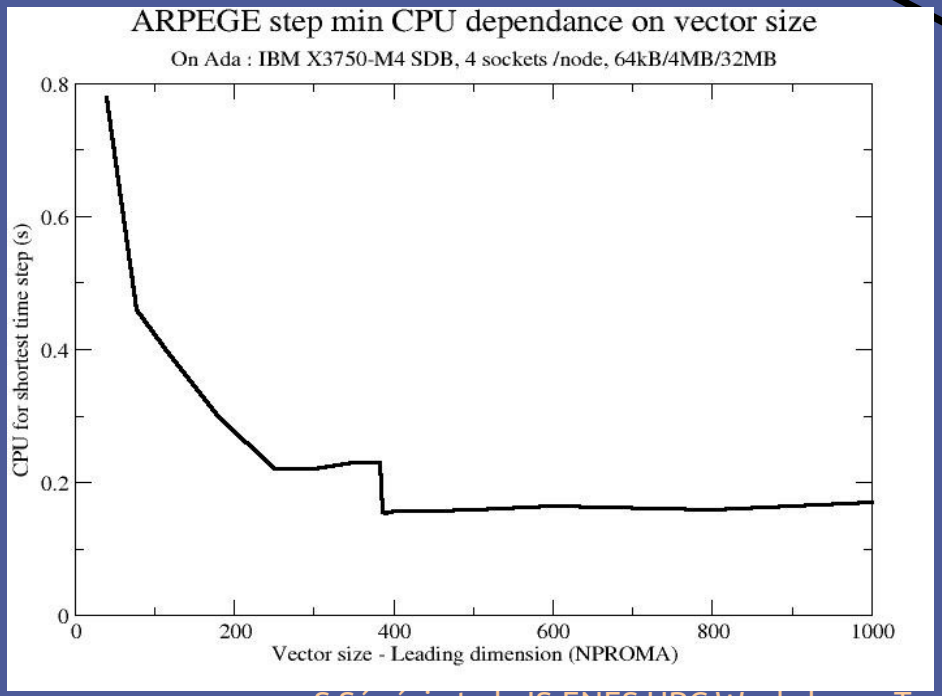
Arpege – Scalability and Vectorization

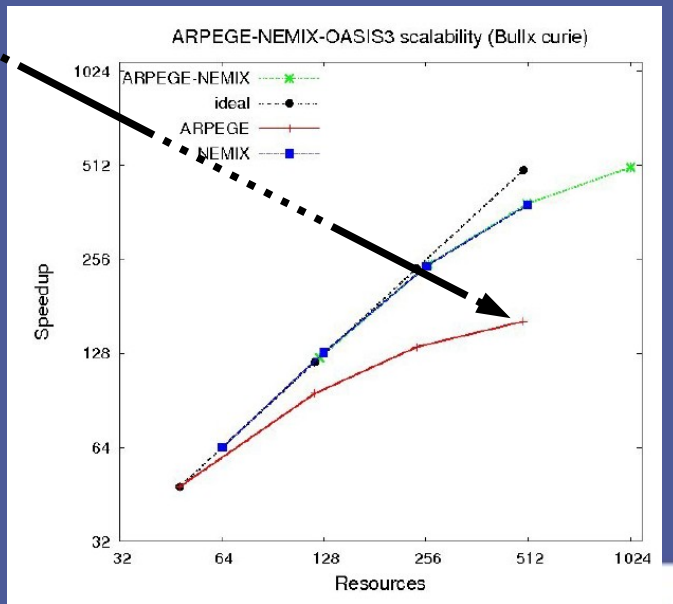
Issues with collective communications for matrix transposition -> logarithmic scalability drop with some subjective limit at :

- 128 / 256 procs for T127
- 512 procs for T359

(figures on SandyBridge, without I/O)

Large impact of arrays leading dimension





METEO FRANCE

Workflow for simulations

ECLIS : Environment for CLImate Simulations

- Allows for coupled and single-component experiments (ATM, LSM, OCE)
- Old-fashioned techniques : scripts, no GUI

Tools

- Proprietary, portable multi-step – handles PBS, LoadLeveler and SLURM
- Job loop tool : « relan » ; allows for easy monitoring and 'resume loop'

User interface

- Experiment design in **one single parameter file** (shell syntax).
- One-command install and run
- 'Plugins' : can provide shell commands to be added at given places
  - Standard plugins for
    - Atmospheric/oceanic LAM forcing
    - Atmospheric/oceanic nudging
    - Ocean bio-geochemistry / atmosphere chemistry
  - User plugins : easy
- Mail on crash
- A run monitor, with job re-launch
  - Analyze logs, may change namelists for butterfly effect for unstable configurations

Perspectives

Upgrade to : Nemo 3.4 or 3.5 / Oasis-MCT

Use XIOS in all components (link to EC-EARTH)

Distribution

- Activate OpenMP in Arpege-Climat and Surfex (link to EC-Earth)
- Validate and tune Gelato SeaIce model last distribution scheme

Tune models to new machines : IvyBridge at Météo-France (and SDB at IDRIS)

- 2 * ~1000 nodes * 2 procs * 12 cores
- Pflops peak : 2 * 0.5 (phase 2 : 2 * 2.8)

Tune CNRM-CM V6, for CMIP6 : choose resolution according to scalability for target 10 y/d

Build a hybrid coupling : stretched atmosphere grid + nested ocean grids - with LPO

Assess scientific value of increased T359 atmosphere resolution : seasonnal and decadal forecast - SPECS 7 FW project backed by PRACE/SPRUCE ressources

Integrate more graphics in the workflow