Architect of an Open World™

Weather and Climate simulations with BULL Technology

02/10/2012

Franck VIGILANT
HPC Consultant
Applications & Performances Team
Extreme Computing Business Unit

A group operating in over 50 countries

WITH ITS CENTER OF GRAVITY I N EUROPE…

…A STRONG BUSINESS IN HPC

eXtreme Computing

- **HPC started in 2004**
- **180 M€ Revenue in 2011**
- **600 people, including 200 people involved in the HW & SW R&D**

bullx

Some key realizations for eXtreme Computing

One step ahead

With **Tera100**, Bull provided to CEA the first Petaflopic Supercomputer in Europe

> Leading Innovation

Think for/with users

With 2 PFlops, **Curie** is one of the most powerful supercomputer in Europe, configured with **4GB/core, an IB Fat tree** topology, **250GB/s IO bandwidth** to cover any kind of applications.

> Understanding HPC Users

One HPC key player

Bull has installed more than 4,5 Pflops (peak) within **Tera100, Curie and Helios** systems.

> 3 TOP20 Configurations

TOP500 SUPERCOMPUTER SITES

And also…

KNMI

Loughborough University

IFP Energies nouvelles

Universidad Pablo de Olavide Sevilla

Cenaero

RWTH Aachen University

University of Leicester

mtc Manufacturing Technology Centre

wirthresearch COMPLETE PERFORMANCE ENGINEERING

AWE

A&P: Climate and Weather Forecasts Interests

Collaborations

Agence Nationale de la Recherche ANR

LOCEAN

CERFACS

Mercator Ocean
Ocean Forecasters

RFPs (not only specialised partners)

METEO FRANCE
Toujours un temps d'avance

Arpege, Arome, …

KNMI

RACMO,…

sara

POP

TECHNISCHE UNIVERSITÄT DRESDEN

COSMO

Support

CURIE

WRF, LMDZ, …
Expert level

R&D

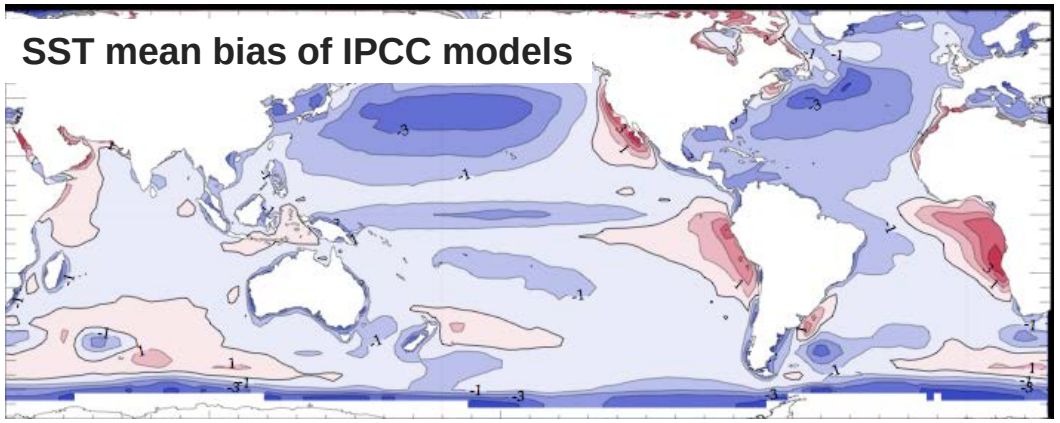
bullx

No better tools for WF/climate applications ! ☺

Pulsation

PULSATION: Peta scale mULti-gridS ocean-ATmosphere coupled simulatIONs

Climate modeling project from the PRACE program (PaRtnership for Advanced Computing in Europe). Started in february 2012

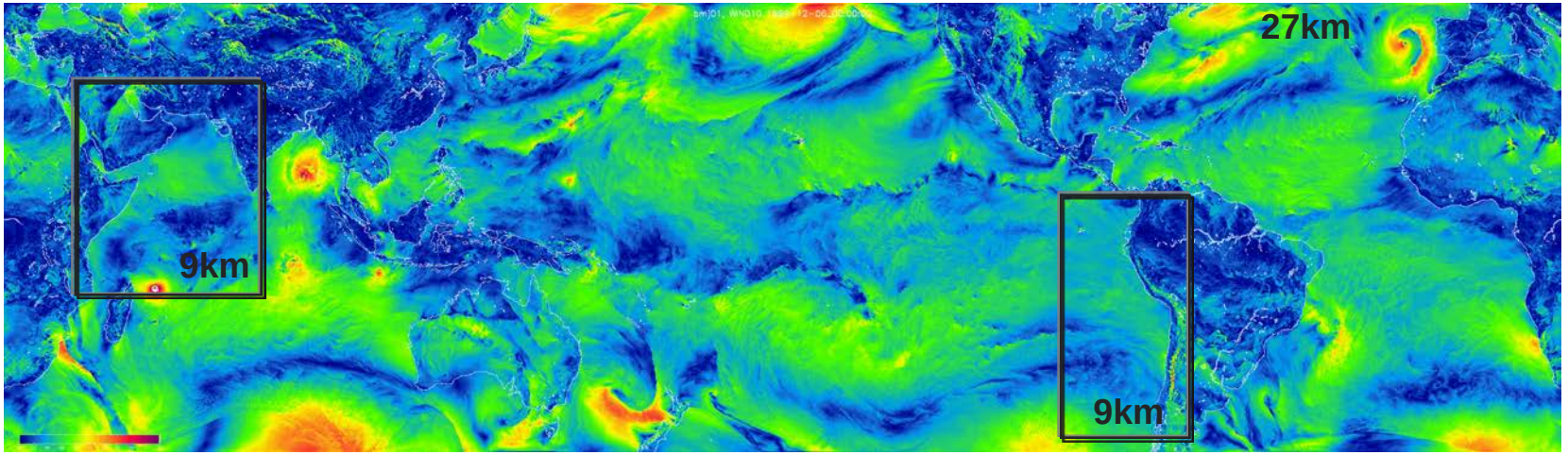


- **Interest:** upscalling processes in coastal upwelling areas
- **Objectives**
  - Quantify the impact of small processes on global climate
  - Reduce large scale and recurrent biases in climate simulations
- **Method:** multi scales coupled models
- **Collaboration**

Pulsation

PULSATION benefits form the 'Grand Challenge' phase of CURIE

- runnning very large scale simulation
- enables researchers to achieve major scientific advances.



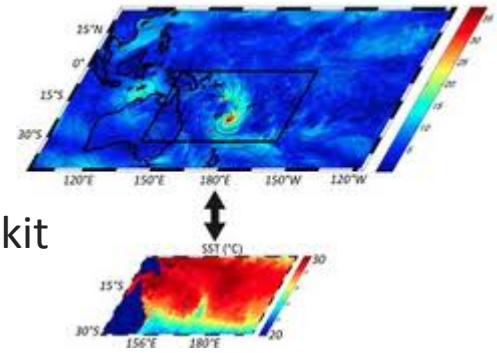
Target: Tropical channel - embedded zooms on both componnents (ocean and atmosphere) - 27km tropical grid + 9km zooms

Courtesy: S. Masson, LOCEAN-IPSL

- Coupled simulation based on **NEMO, WRF & OASIS**
- **19 millions of CPU hours**
- **140 To** of data to be expected
- Simulations up to **16 000 cores**

Pulsation

MODEL, combinaison of state-of-the-art and popular models:



- **Atmosphere:** Weather Research and Forecasting (**WRF**)
- **Ocean:** Nucleus for European Modeling of the Ocean (**NEMO**)
- **Coupler:** Ocean Atmosphere Sea Ice Soil and Model Coupling Toolkit (**OASIS3-MCT**)

- **Horizontal resolutions:** from 27 km down to 9 km, tropical channel (45°S-45°N)
- **2 years simulation** (x2 configurations)

WRF and NEMO are among the very few models able to combine:

- high-resolution simulations at global scale
- 2-way embedded zoom

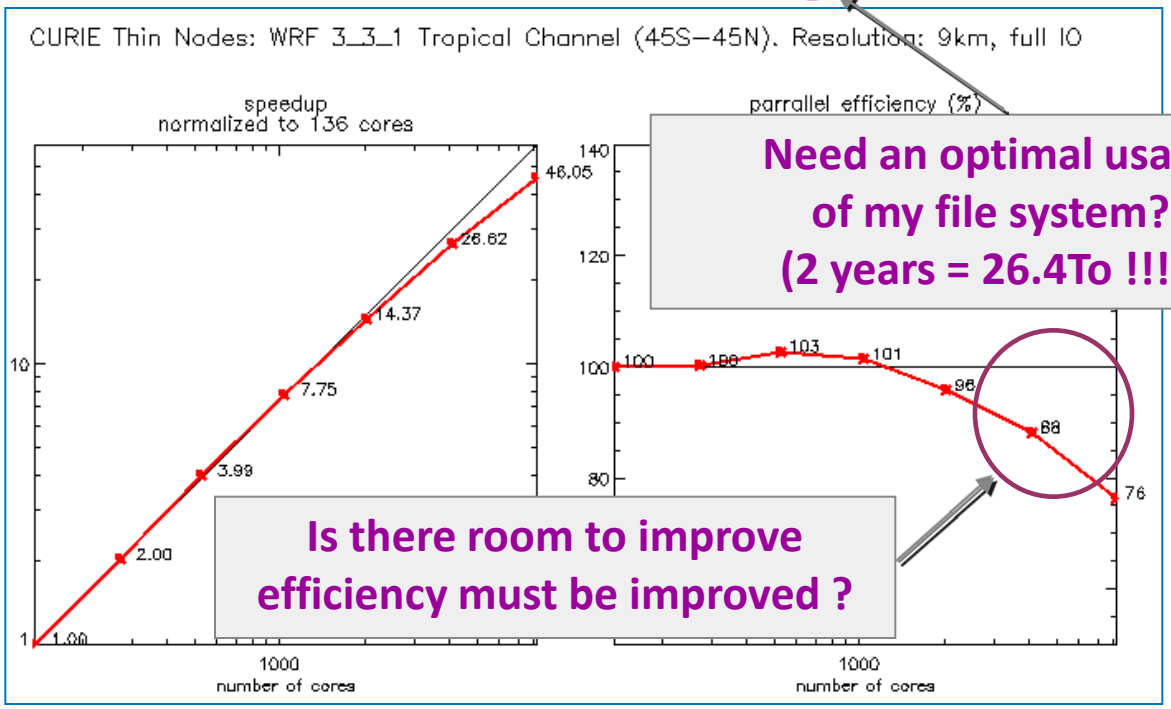
OASIS3-MCT is used as library, so only 2 executables are needed

Pulsation: Focus on Scalability

- **global resolution 9km**
- Running on **8 204 cores**

WRF

- **6h elapsed / 1 month** of sim.
- **1.1To / month of simulation**



Courtesy: S. Masson, LOCEAN-IPSL

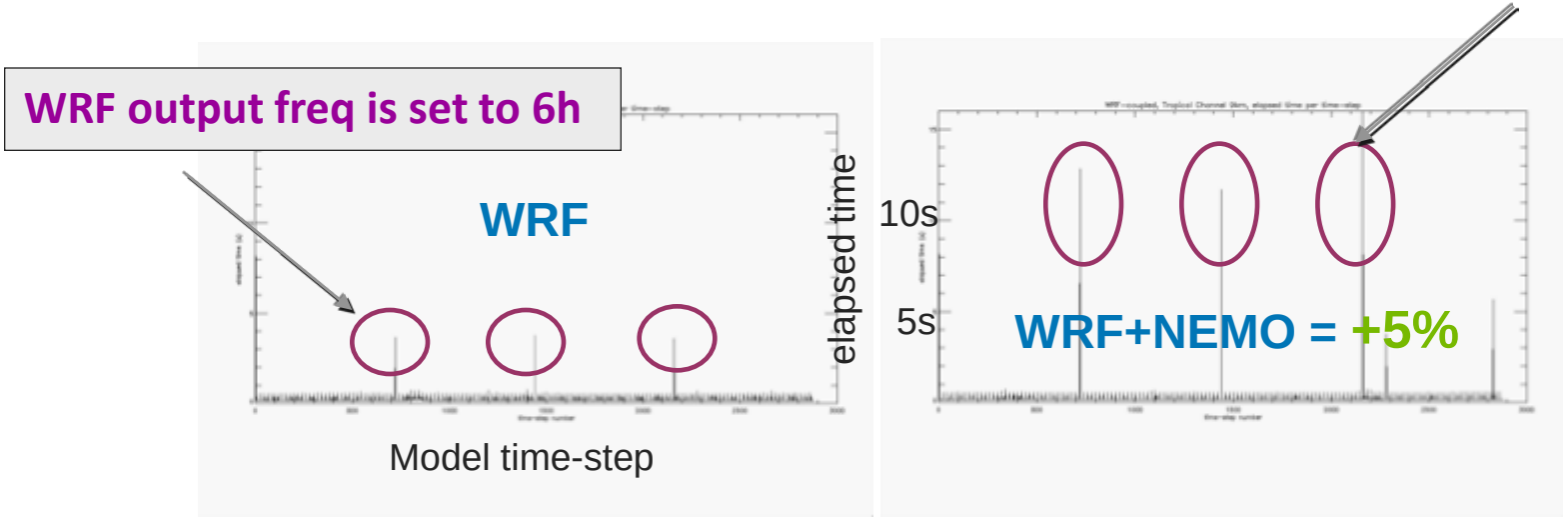
76% of parallel efficiency on 8 204 cores (based on 136 cores run)

Pulsation: Timing & IO Performances

WRF + NEMO:

- **global resolution 9 km**
- WRF @ 8204 cores – NEMO @ 512 cores

Coupling stresses the 'burst' phenomenon



Courtesy: S. Masson, LOCEAN-IPSL

Coupling: only 5% degradation w.r.t WRF alone ☺

So what ?

any bottlenecks in my code ?

What about MPI communications ?

IO optimization needed ?

Vectorisation ?

Cache miss ?

Better understand your application easily….

bullx Prof

- **Main features:**
  - able to characterized a code over a large panel of functions such as:
    - ✓ **Function profiling**
    - ✓ **MPI profiling**
    - ✓ **Hardware counter profiling**
    - ✓ **IO profiling**
  - no dependance on MPI layer (OpenMPI, bullxMPI, Intel MPI)
- **light tool**
- **easy to handle**
- **fit for petaflopic tools**

Better understand your application easily….

bullx Prof

- **Only one command line:**

$ **bullxpprof** [bxprof-args] **program** [prog-args]

Or

$ **mpirun** [mpirun-args] **bullxpprof** [bxprof-args] **program** [prog-args]

- **Outputs:**
  - **USER report**
  - **MPI profile**
  - **I/O profile**
  - **HWC report**

Application analysis with bullxprof

Timing report - summary

A global overview of the simulation

```
Process walltime     :            50.772 sec
Number of processes  :               64
Estimated speed-up   :            21.39
Estimated efficiency :            33.43 %

-------------------------------------------------------------------------------------------------------------
| Region   | Min Time [rank]                  | Max time  [rank]                  | Avg time              | % walltime   |
-------------------------------------------------------------------------------------------------------------
| ALL      |           50.603 sec [     20 ] |           50.772 sec [      0 ] |           50.681 sec |      100.00 % |
| USER     |           15.490 sec [      8 ] |           17.512 sec [     45 ] |           16.308 sec |       32.18 % |
| MPI      |           32.955 sec [     29 ] |           35.141 sec [      1 ] |           33.741 sec |       66.57 % |
| I/O      |            0.110 sec [      1 ] |            0.841 sec [     60 ] |            0.632 sec |        1.25 % |
-------------------------------------------------------------------------------------------------------------
```

Timing report – USER region

Focus on USER main routines

```
Region : USER
-------------------------------------------------------------------------------------------------------------------------
| Min Time [rank]                  | Max time  [rank]                  | Avg time              | % region | % walltime | Functions
-------------------------------------------------------------------------------------------------------------------------
|      12.462 sec [     19 ] |      14.108 sec [     45 ] |      13.248 sec |   81.23 % |     26.14 % | conj_grad_
|       2.834 sec [     56 ] |       3.404 sec [     45 ] |       3.060 sec |   18.77 % |      6.04 % | MAIN__
|       0.000 sec [      0 ] |       0.000 sec [     57 ] |       0.000 sec |    0.00 % |      0.00 % | pow
|       0.000 sec [      1 ] |       0.000 sec [      0 ] |       0.000 sec |    0.00 % |      0.00 % | print_results_
|       0.000 sec [      0 ] |       0.000 sec [      5 ] |       0.000 sec |    0.00 % |      0.00 % | timer_clear_
|       0.000 sec [      1 ] |       0.000 sec [      6 ] |       0.000 sec |    0.00 % |      0.00 % | timer_read_
-------------------------------------------------------------------------------------------------------------------------
```

Application analysis with bullxprof

Timing report – MPI profiling

Region : MPI

Min Time [rank]	Max time [rank]	Avg time	% region	% walltime	Functions
0.002 sec [16]	0.004 sec [58]	0.003 sec	0.01 %	0.01 %	MPI_Barrier
1.094 sec [16]	1.186 sec [1]	1.126 sec	3.34 %	2.22 %	MPI_Init
0.048 sec [29]	1.496 sec [1]	0.426 sec	1.26 %	0.84 %	MPI_Irecv
0.000 sec [1]	0.019 sec [0]	0.008 sec	0.02 %	0.02 %	MPI_Reduce
4.887 sec [0]	24.091 sec [11]	19.043 sec	56.44 %	37.57 %	MPI_Send
7.810 sec [24]	26.302 sec [63]	13.135 sec	38.93 %	25.92 %	MPI_Wait

Timing report – I/O region

Region : I/O

Min Time [rank]	Max time [rank]	Avg time	% region	% walltime	Functions
0.001 sec [49]	0.001 sec [27]	0.001 sec	0.12 %	0.00 %	close
0.000 sec [33]	0.000 sec [29]	0.000 sec	0.00 %	0.00 %	pipe
0.089 sec [33]	0.114 sec [16]	0.105 sec	16.69 %	0.21 %	read
0.004 sec [1]	0.732 sec [60]	0.526 sec	83.18 %	1.04 %	write

And also…

- GFLOPS
- Floating point operations
- Cache misses L1, L3

What's next ?

bullx Prof

Apply those tools to IFS ! ☺

Architect of an Open World™