
PRACE

PRACE Application Benchmark Suite for Accelerators

Victor Cameo Ponz - cameo@cines.fr

CINES

PRACE Unified European Applications Benchmark Suite (UEABS)

- Lack of an APPLICATION benchmark suite
  - Lots of HPC oriented benchmarks [NERSC-STREAM][SPEC][RODINIA]
  - Non HPC application benchmark suite [PRINCETON-PARSEC]

- Previously PRACE and DEISA had an application benchmark suite featuring 29 codes
  - Not maintainable
  - Out of date codes

- Do we really need an application benchmark suite ?

PRACE Unified European Applications Benchmark Suite (UEABS)

- "Pure" performances (as I/O and interconnect speed, peak perf) are obscure to some people
- Production machines run applications at the end

- UEABS has be designed for:
  - HPC center when buying hardware
  - HPC users when choosing center
  - HPC benchmarker when testing performance regression of systems

- Trying to avoid previous mistakes
  - less codes than PRACE&DEISA suite
  - Documented/maintained/run with modern tools

PRACE Unified European Applications Benchmark Suite (UEABS)

- Code selection is the result of a carefully executed process [PRACE-2IP D7.4]:
  - publicly available
  - datasets publicly available.
  - not have any significant barriers to portability.
  - demonstrate good scalability.
  - active support by the developers
- Trying to cover as much scientific fields as possible
- Keep code selection open and flexible
- Maintained by PRACE through man power allocation to partners (BCOs)
- Online accessible
- Target European usage

	Standart UEABS	Accelerated UEABS	Science Area application
ALYA			CFD
Code_Saturne			CFD
CP2K			Quantum MD
GADGET			Astrophysics
GENE			Plasma Physics
GPAW			Quantum MD
GROMACS			Classical MD
NAMD			Classical MD
NEMO			Earth Sciences
PFARM			Astrophysics
QCD			QCD
Quantum Espresso			Quantum MD
SHOC			Benchmaking
Specfem3D_Globe			Earth Sciences

PRACE UEABS for accelerators

- Renewal interest of accelerators in HPC
- Different languages, librairies

- Dedicated task in the 4th implementation of PRACE to design this part of the suite
- Traget Intel Xeon Phi and NVIDIA GPUs

Design process: code and test cases definition

Call for codes to be included in the suite and belonging man power to run them. For each code accelerators specifics informations has been gathered [PRACE-4IP WP212]:

- Implementation description
- Test cases definition

- Finally selected: 10 applications and 1 synthetic benchmark, partly based on existing UEABS

- Expected to be merged the "regular" UEABS

	OpenMP	OpenCL	Cuda
ALYA			
Code_Saturne			
CP2K			
GPAW			
GROMACS			
NAMD			
PFARM			
QCD			
Quantum Espresso			
SHOC			
Specfem3D_Globe			

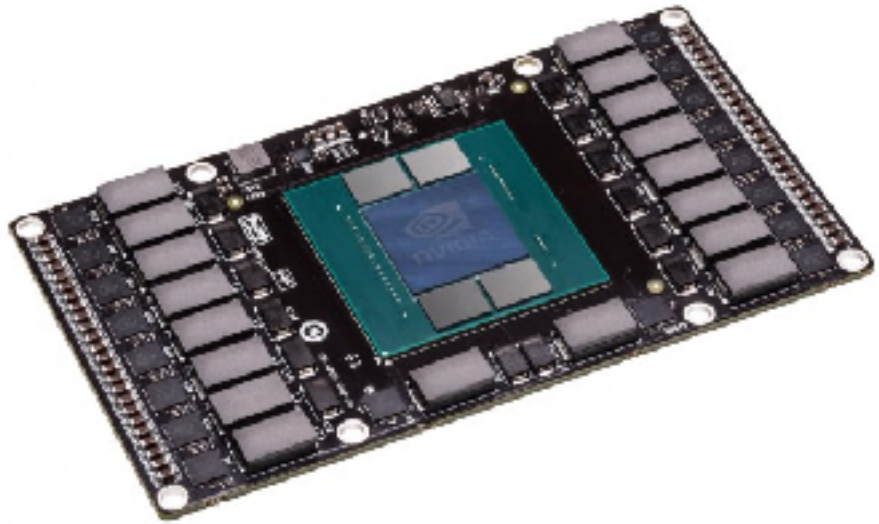
Design process: Initial runs and guides to compilation an execution

First result has been obtained on PRACE partners systems

- Access granted to NVIDA K40 and Intel KNC
  - On Cartesius @SURFsara
  - And on Curie @CEA
- User guides were issued
  - Based on the runs on those machines
  - Cross-reviewed

Design process: runs on leading edge systems

- Access to leading edge GPU and Xeon Phi system has been granted to BCOs
  - On Ouessant @IDRIS
  - On Frioul @CINES
  - On various systems based on benchmark owners ressources (PizDain, Marconi...)



Reports & publications

- Benchmark descriptions, tests cases and results has been compiled into the PRACE paper [PRACE-4IP D7.5]
- The benchmark suite is available online in the [PRACE CodeVault] repository, including:
  - Instructions to run the benchmark
  - Test cases
  - Sample results

Accelerated UEABS on OpenPower

- Example codes that ran on OpenPower platforms
  - Code_Saturne, GROMACS, QCD, Specfem3D_Globe
  - benchmark owner (BCO)
  - some details about physics addressed
  - some details about the implementation
- Present some results
  - Mainly on OpenPower+GPU
  - Sometimes on KNL

Code_Saturne

- Benchmark run by Charles Moulinec, STFC, Daresbury UK
- Code general description
  - CFD software package
  - EDF R&D since 1997, open-source since 2007
  - finite volume method approach
- Implementation
  - Hybrid OpenMP/MPI
  - PETSc CUDA is plugged in to solve operation on sparse matrices
  - 350 000 lines of code, 37% Fortran, 50% C, 13% Python
- Test case: Flow in a 3-D lid-driven cavity (tetrahedral cells)
  - Simple geometry
  - Modular and the mesh size can easily been varied (here about 13 million cells)

Code_Saturne

IBM POWER 8 / S822LC	IBM POWER 8 / S824L
2 processors with 10 cores	2 processors with 12 cores
8 logical cores	8 logical cores
4 on-chip memory controllers (SCM)	8 on-chip memory controllers (DCM)
256 GiB RAM/node	256 GiB RAM/node
~2.92 GHz	~3.00GHz
2 NVIDIA K80	2 NVIDIA K40
Each K80: 2 GPU GK210/K80	-
2,496 stream processors	2,880 stream processors
12 GiB RAM	12 GiB RAM

Code_Saturne

IBM POWER8 S822LC				IBM POWER8 S824L			
#C	T (s)	T (s)	SP	#C	T (s)	T (s)	SP
1	1022.18	630.54	1.62	1	1087.75	637.86	1.71
2	621.20	337.12	1.84	2	659.71	327.81	2.01
4	263.61	173.95	1.51	4	267.82	*189.04*	*1.42*
20	76.38	*109.75*	*0.70*	24	57.81	*104.82*	*0.55*

- Blue cells => cheaper to run on CPU than using CPU+GPU
  - same amount of data transferred by 20 or 24 slices generates system administration costs, latency costs

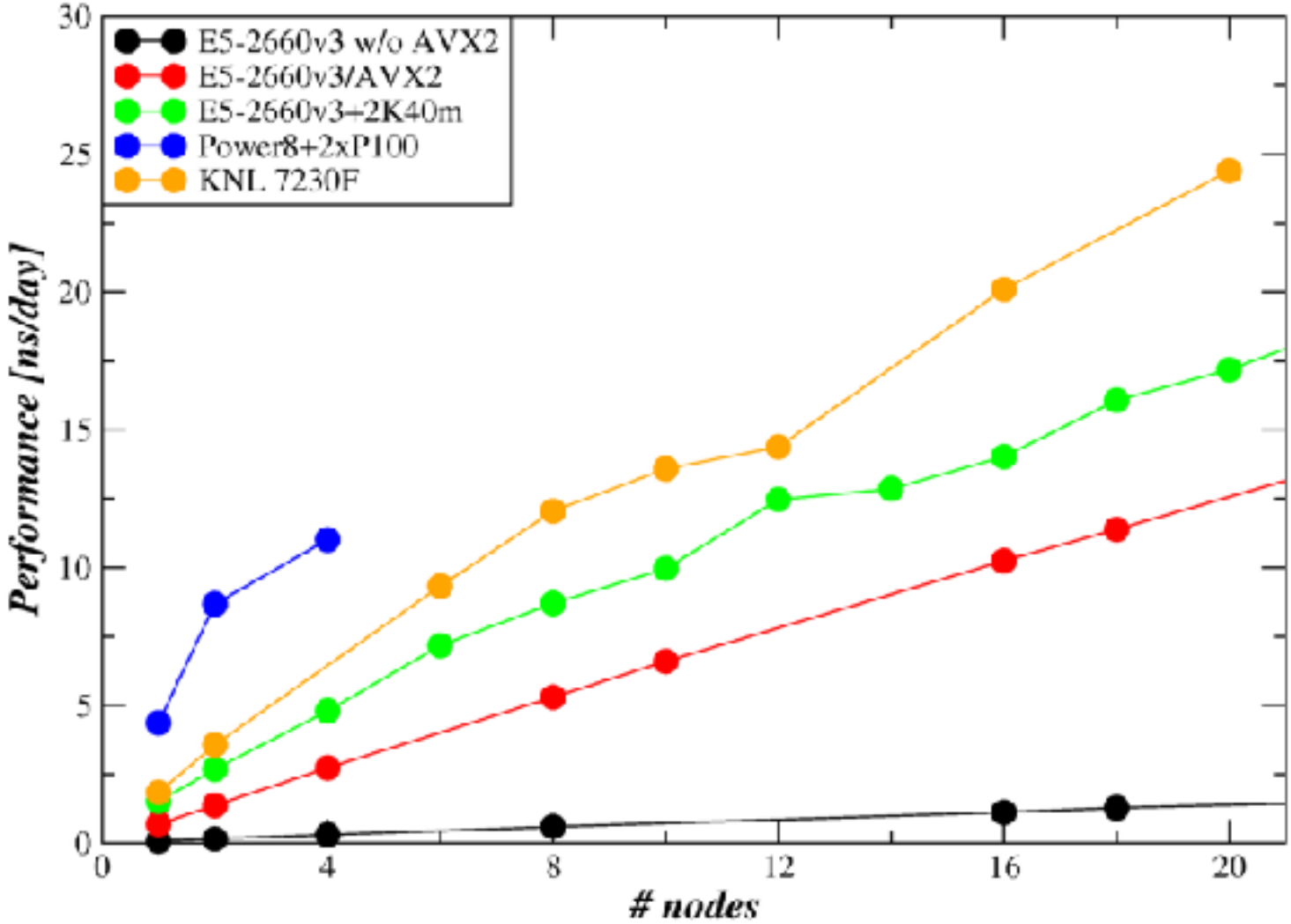
GROMACS

- Benchmark run by Dimitris Dellis, GRNET, Athen Grece
- Code general description
  - versatile package to perform molecular dynamics
  - biochemical molecules like proteins, lipids and nucleic acids
  - also used for non-biological systems research, e.g. polymers
- Implementation
  - MPI + OpenMP (Xeon Phi) and Cuda (GPU)
  - GROMACS is written in C/C++ and freely available under the GPL license
- Test case: Lignocellulose
  - 3.3 million atoms
  - Inhomogeneous system
  - scale well up to 10 000 cores on X86

GROMACS

- Run on Ouessant at IDRIS
  - Technological watch group of GENCI

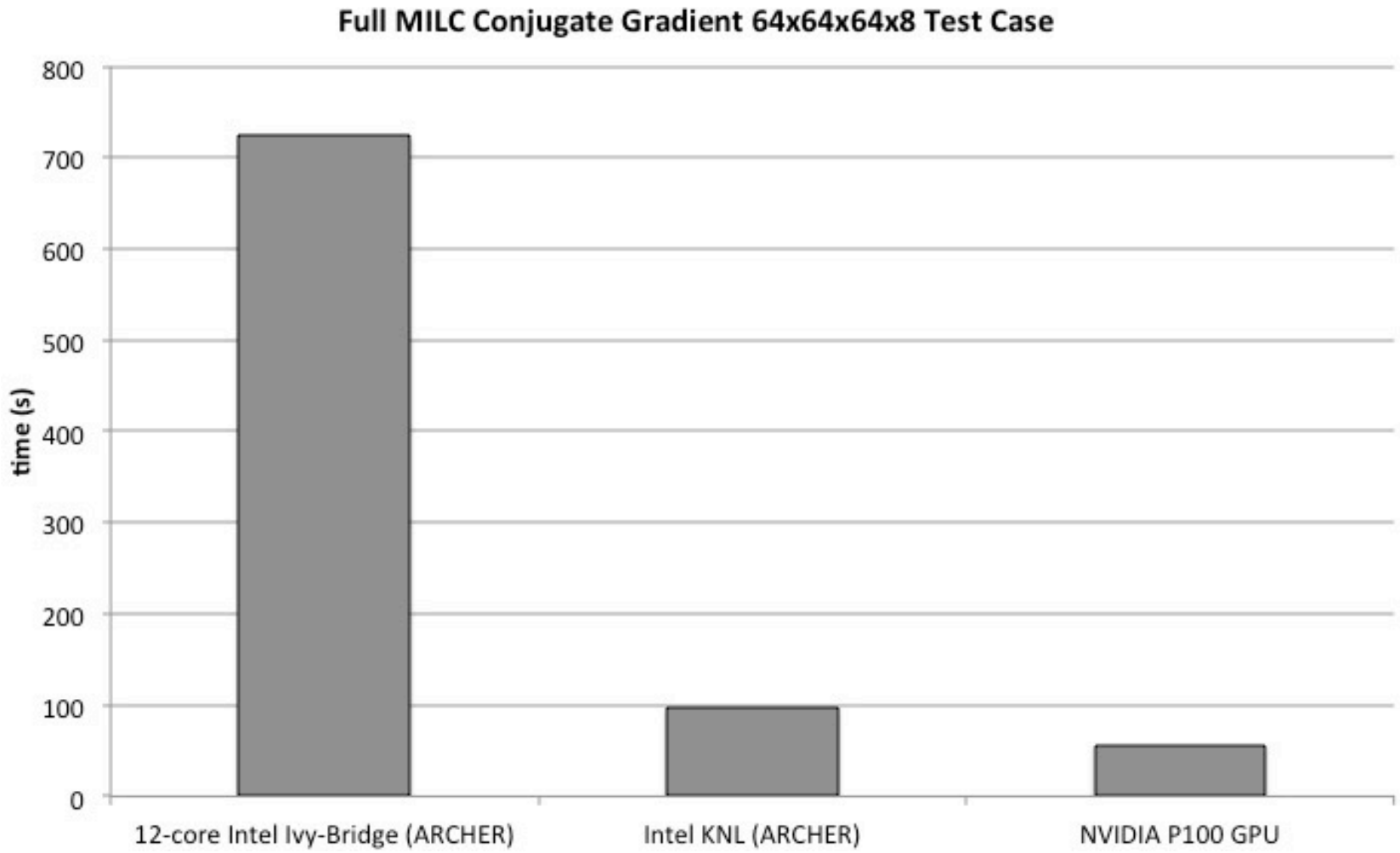
- Compute nodes
  - POWER8+ sockets, 10 cores
  - 128 GB of DDR4 memory
  - 4 NVIDIA's new generation Pascal P100 GPU
- Interconnect
  - NVLink interconnects GPUs
  - A Mellanox EDR InfiniBand interconnects nodes



- speed up of 2-5x with respect CPU only was achieved with GPU.

QCD kernel E

- Benchmark run by Alan Gray, EPCC, Edinburgh UK
- Code general description
  - Quantum Chromo Dynamics (QCD)
  - quarks and gluons interaction to form nucleon
  - compute intensif
- Implementation
  - C + MPI as base
  - “targetDP” programming model (abstraction model that allow multiple platforms using CUDA or OpenMP)
- Test case: MILC 64x64x64x8
  - Very modular, allows strong as weak scaling



- Single node performances on Ouessant (similar to GROMACS) and one KNL,64-core 7210 model
- KNL is 7.5X faster / the Ivy-bridge, the Pascal is 13X faster and the Pascal is 1.7X faster than the KNL.

Specfem3D_globe

- Benchmark run by Victor Cameo Ponz, CINES, Montpellier France
- Code general description
  - simulates three-dimensional global and regional seismic wave propagation
  - spectral-element method
- Implementation
  - Fortran + MPI as base
    - OpenMP and/or CUDA.
- Test case
  - 3D shear-wave speed model (S362ANI)
  - Mesh varies with size of the problem and the targeted MPI task number

Specfem3D_globe

- Run on Ouessant, as QCD and GROMACS
- 24 MPI tasks
  - Power
    - 1 dual socket P8 (MPI:24): 707.276s
    - 2 n dual socket P8 (MPI:12/12): 404.410s
  - Power + GPU
    - 1 dual socket P8 + 2P100 (shared) : 52.659
    - 1 dual socket P8 + 4P100 (shared) : 37.836
    - 6 dual socket P8 + 24P100 (exclusive) : 27.111s
- Troubles using NVIDIA profiler and GPU sharing tool from NVDIA (MPS)

Remarks regarding runs and results

- OpenPower compilation stack is currently maturing
- Some test cases should be made bigger to scale on GPU
- Meanwhile good result on some applications

Current state

Most of benchmark applications evolved and matured due to the quick evolution of accelerators cards

- Guides evolved with the latest run on NVIDIA Pascal and Intel KNL
  - need cross-review again
- Some benchmark suffered going to last generation hardwares
  - KNL dropping OpenCL lost SHOC availability on this system
  - GPAW couldn't be ran on multiple P100
- Running the suite on lead edge machines showed its limits but also proved its maturation

Future plans

We have plenty of ideas to make the suite evolve in a better way

- Enhance CodeVault git repository
- Automate guide review
  - make guides smaller & executable
  - Setup a simulated environment so that review is CI
- More run on more different architectures
  - showcase where bottleneck of each application are
  - Run on PRACE-3IP machines to exploit energy metrics

Energy metrics on PRACE machines

PRACE-3IP is ending this year with the delivery of 3 machines, each featuring energy probes.
Goal is to measure energy to solution and energy efficiency.

- E4 computer engineering
  - 45 nodes with x2 IBM POWER8 processors and x4 NVIDIA P100 GPU
  - CPU and GPU liquid cooling
- Atos/Bull
  - Xeon Phi @CINES, France
  - 168 KNL nodes, fully liquid cooled
- MAXELER
  - 4 MPC-H servers including 2x MAX5 DFE
  - Small but innovative and promising

Acknowledgements

- Lots of people !
- Adem Tekin, ITU, Alan Grey, EPCC, Andrew Emerson, CINECA, Andrew Sunderland, STFC, Arno Proeme, EPCC, Charles Moulinec, STFC, Dimitris Dellis, GRNET, Fiona Reid, EPCC, Gabriel Hautreux, INRIA, Jacob Finkenrath, CyI, James Clark, STFC, Janko Strassburg, BSC, Jorge Rodriguez, BSC, Martti Louhivuori, CSC, Philippe Segers, GENCI, Valeriu Codreanu, SURFSARA

- This work was financially supported by the PRACE project funded in part by the EU's Horizon 2020 research and innovation programme (2014-2020) under grant agreement 653838.

Bibliography

- [PRACE-2IP D7.4] Unified European Applications Benchmark Suite - Mark Bull, EPCC - 07/2013
- [PRACE-4IP WP212] Description of the initial accelerator benchmark suite - Gabriel Hautreux, GENCI - 02/2016
- [PRACE-4IP D7.5] Application performance on accelerators - Victor Cameo Ponz, CINES - 04/2017
- [NERSC-STREAM] Stream Benchmark provided by NERSC
- [SPEC] Standard Performance Evaluation Corporation - http://spec.org/
- [PRINCETON-PARSEC] The Princeton Application Repository for Shared-Memory Computers - http://parsec.cs.princeton.edu/
- [RODINIA] Rodinia Benchmark Suite - https://github.com/JuliaParallel/rodinia
- [PRACE CodeVault] Official PRACE git repository: https://gitlab.com/PRACE-4IP/CodeVault & https://gitlab.com/MisterFruits/CodeVault

THANK YOU FOR YOUR ATTENTION

www.prace-ri.eu