OpenFOAM in HPC centers -
CRIANN’s experience feedback

1e Journée française des utilisateurs d’OpenFOAM, 18/05/2016

Patrick BOUSQUET-MELOU, CRIANN
patrick.bousquet-melou@criann.fr
HPC in France

HPC pyramid

- Regional (Tier-2) computing centers, « mésocentres »
  - Up to 12 000 CPU cores today
  - CRIANN is Normandy’s « mésocentre »

- National centers (GENCI - Grand Equipement National de Calcul Intensif)

- European centers (PRACE - Partnership for Advanced Computing in Europe)
CRIANN HPC center
Today’s technical resources

- « Antarès » supercomputer
- IBM iDataPlex cluster
  - 3068 CPU cores owned by CRIANN
  - 624 CPU cores by ECN
  - Many standard compute nodes
    - MPI oriented resources
  - Specialized compute nodes
    - I/O intensive applications (quantum chemistry, solid state chemistry)
  - GPU
  - Visualization, pre-/post-processing
ANTARÈS cluster: varied job types and servers

But main cluster part used by MPI jobs, especially in CFD

| 125 «Nehalem» nodes (8 cores, 24 GB RAM) |
| «standards» MPI jobs (≤ 3 GB memory / core) |
| of medium size |
| ≤ 128 cores during 24 hours |
| or 256 cores during 12 hours |
| CFD, molecular dynamics, materials sciences, etc. |

| 169 «Westmere» nodes (12 cores, 48 or 96 GB RAM) |
| 12 «disks» nodes: intranode I/O intensive jobs in quantum chemistry |
| 3 «memory» nodes (96 GB RAM / node): solid state chemistry: MPI 20 cores, 8 GB memory / core |

| 13 GPGPU nodes |
| 141 standard nodes (4 GB RAM / core): standard MPI jobs (4 GB RAM / core) of big size (> 128 cores and ≤ 1032 cores) during 24 hours, or of medium size (≤ 144 cores) during 48 hours, in CFD, molecular dynamics, materials sciences, etc. |
Scientific topics
Modeled with OpenFOAM at CRIANN

· CORIA UMR 6614
  – Atomization LES, cavitation in injectors, SPH-FVM coupling
  – Fire (with FireFOAM solver)
  – Buoyant destabilization in wet granular media
  – Heat transfer and flow computations for eyeglass-framed thermo-electric converter design

· M2C UMR 6143 (Morphodynamique Continentale et Côtière, Caen)
  – Hydrodynamic impact of barnacles colonization on marine turbine blades
  – Sedimentary bed / hydrodynamic turbulence interaction
  – Impact of offshore wind turbines foundations on hydro-sedimentary system
Industrial example
INNOSEA

- Prototype of a services offer for wind farms development
  - OpenFOAM derived aerodynamics LES solver
- Project certified by HPC-PME
  - GENCI - INRIA - OSEO
OpenFOAM software resources
at CRIANN

- Installed OpenFOAM versions
  - Mutualized free versions
    - 1.7.1, 2.0.1, 2.1.0, 2.1.1, 2.2.0, 2.2.1, 2.3.0, 2.3.1, 2.4.0, 3.0.1, 3.0+
    - Optional solvers / utilities added on demand: wave2Foam, CFMesh
  - Private versions
    - Foam-extend 3.1 + Naval Hydro Pack

- Academic researchers often fix on a given version, so do not often ask for updates

- Industries internally qualify each new version
  - Then ask CRIANN each new version installation on Antarès cluster
OpenFOAM users support
at CRIANN

- General documentation and (half a day) class on Antarès cluster usage
- General 3 days parallel programming training with 1,5 day on MPI

- Support to new OpenFOAM users at CRIANN
  - Advice on splitting pre - processing (decomposePar), solver and post-processing (reconstructPar) in 3 batch jobs
  - Pre - / post - processing with serial jobs and as memory as necessary (usually < 40 GB, sometimes up to 90 GB)
  - Solvers with parallel jobs and less than 4 GB of memory per core (MPI task)
OpenFOAM users support (2)

at CRIANN

• Adaptation of disk user space
  - Individual permanent folders created in the scratch partition of the parallel file system (GPFS)
  - Avoids copies from /home to scratch and scratch to /home, as the input data itself may be big for OpenFOAM LES applications

• Technical information given to developers on how to install OpenFOAM in their user space
  - CRIANN experienced and solved issues in compiling OpenFOAM x.x.x
  - Then if a developer needs OpenFOAM x.x.x in its own user space, CRIANN provides him the fixes
  - How to correctly couple OpenFOAM MPI library to the batch scheduler (LoadLeveler)
OpenFOAM batch submission
With LoadLeveler at CRIANN

# @ job_name = damBreakFine

# Job time (hh:mm:ss)
# @ wall_clock_limit = 12:00:00

# OpenFOAM parallel processes
# @ total_tasks = 32

# Parallel process maximum memory (mb, gb)
# @ data_limit = 2000mb

... # @ queue

# OpenFOAM 3.0+ environment
# ——————————————————————————-
module load openfoam/3.0+

# Path of customized mpirun command (coupled to LoadLeveler)
export PATH=/soft/OpenFOAM-3.0+/crihan_tools:$PATH
# ——————————————————————————-

mpirun interFoam -case ${LOADL_JOB_NAME} -parallel > ${LOADL_JOB_NAME}.log
Remote visualization
Meshing with Salome, post-processing with Paraview

- Interactive session obtained after a specific batch job for visualization
  - A TurboVNC client is needed on the user’s local desktop
OpenFOAM I/O
Can stress today’s CRIANN’s parallel file system

- The I/O traffic periodically saturates Antarès parallel file system (GPFS)
  - 2009 I/O configuration (4.5 GB/s aggregated bandwidth (for 3000 cores))
  - Issue of high numbers of files read or written simultaneously

*Reads and writes during a month period (iop/s) on Antarès parallel file system*
OpenFOAM I/O (2)

Can stress today’s CRIANN’s parallel file system (2)

· Can arise due to OpenFOAM parallel solvers writing with each MPI task at each time step or frequently (one file per process per variable per time step or frequently)
  - Then post-processings can stress the parallel file system

· Good practice: use of function objects for creating post-processing images of data subsets
  - http://cfd.direct/openfoam/user-guide/function-objects/#x31-1940044
  - Limiting time and spatial extent for post-processing function objects
I/O good practices with OpenFOAM

No more I/O than necessary

- Check PurgeWrite in order to limit the overall number of files
  - limit the number of time directories that are stored (cyclic overwrite)
  - default value: purgeWrite 0 => no overwrite
  - in steady state calculation purgeWrite 1 => overwrite previous iteration
- Check data writing parameters in controlDict
  - Is it necessary to write at every time step?
- Carefully choose the appropriate time step (not too small)
- Is data reading mandatory at each time step?
  - check value of runTimeModifiable
  - by default yes => re-read dictionaries at beginning of each time step
I/O good practices with OpenFOAM (2)

No more I/O than necessary (2)

• Source of the previous slide
  - http://www.openfoam.com/documentation/user-guide/controlDict.php#x19-106003r4
  - https://wiki.calculquebec.ca/w/OpenFoam/en