

Hands-on: *Turpan* (Ampere Altra Q80-30 + 2 x A100) CloverLeaf_OpenACC

VI-HPS Team

Tutorial exercise objectives

- Familiarise with usage of VI-HPS tools
 - complementary tools' capabilities & interoperability
- Prepare to apply tools productively to *your* applications(s)
- Exercise is based on a small portable benchmark code
 - unlikely to have significant optimisation opportunities
- Optional (recommended) exercise extensions
 - analyse performance of alternative configurations
 - investigate effectiveness of system-specific compiler/MPI optimisations and/or placement/binding/affinity capabilities
 - investigate scalability and analyse scalability limiters
 - compare performance on different HPC platforms
 - ...

Setup for exercises

- Connect to your training account on Turpan (with X11-forwarding)

```
% ssh -X <yourid>@turpanlogin.calmip.univ-toulouse.fr
```

- Set account and default environment (NVHPC + OpenMPI) via helper script

```
% source /tmpdir/vi-hps/opt/setup.sh  
WORK=/tmpdir/$USER
```

Needed in each new shell

- Copy tutorial sources to your WORK directory

```
% cd $WORK  
% tar zxvf /tmpdir/vi-hps/material/handsons/CloverLeaf_OpenACC.tar.gz  
% cd CloverLeaf_OpenACC
```

Case study: CloverLeaf_OpenACC

- HPC mini-app developed by the UK Mini-App Consortium
 - Solves the compressible Euler (partial differential) equations in 2D using an explicit, second-order finite-volume method on a regular Cartesian grid
 - Part of Mantevo project and SPEC HPC benchmark suite
 - Available on GitHub: <https://uk-mac.github.io/CloverLeaf/>
- OpenACC-enabled MPI version written in Fortran90, run using short testcase
 - clover_bm256_short.in input deck: 15360x15360 cells, 87 timesteps (full-length clover_bm256.in input deck does 2955 timesteps to reach validated solution)
 - Run with 2 MPI tasks-per-node so that each has a dedicated GPU
 - Run on 1 or more nodes to see its strong scaling performance: 3 nodes set for exercise
- Provided version of Makefile and sources customized for this tutorial
 - builds **clover_leaf** executable in separate binary directory when using instrumentation



CloverLeaf_OpenACC source directory

```
% ls
CHANGE_LOG
InputDecks/
Makefile
PdV.f90
PdV_kernel.f90
README.md
accelerate.f90
accelerate_kernel.f90
advec_cell_driver.f90
advec_cell_kernel.f90
advec_mom_driver.f90
advec_mom_kernel.f90
advection.f90
bin/
bin.scorep/
build_field.f90
calc_dt.f90
calc_dt_kernel.f90

clover.F90
clover_leaf.F90
data.f90
definitions.f90
field_summary.f90
field_summary_kernel.f90
flux_calc.f90
flux_calc_kernel.f90
ftocmacros.h
generate_chunk.f90
generate_chunk_kernel.f90
hydro.F90
ideal_gas.f90
ideal_gas_kernel.f90
initialise.f90
initialise_chunk.f90
initialise_chunk_kernel.f90
pack_kernel.f90

parse.f90
read_input.f90
report.f90
reset_field.f90
reset_field_kernel.f90
revert.f90
revert_kernel.f90
start.f90
timer.f90
timer_c.c
timestep.f90
update_halo.f90
update_halo_kernel.f90
update_tile_halo.f90
update_tile_halo_kernel.f90
viscosity.f90
viscosity_kernel.f90
visit.f90
```

43 Fortran90 modules (3 with pre-processor macros),
16 for kernels (1 timer module in C)

CloverLeaf_OpenACC: Makefile

```
#Crown Copyright 2012 AWE
#
# This file is part of CloverLeaf.
#
# CloverLeaf is free software...
#
# Agnostic, platform independent Makefile for the CloverLeaf benchmark code.
# It is not meant to be clever in any way, just a simple build script.
#
# this works as well:-
#
# make COMPILER=PGI # == NVHPC
#
...

#PREP=scorep --openacc --cuda --user

MPI_COMPILER=mpif90
C_MPI_COMPILER=mpicc

clover_leaf: *.f90 Makefile
    $(PREP) $(MPI_COMPILER) $(FLAGS) ...
...
```

No instrumentation by default

Building clover_leaf

```
% make
mpicc -c timer_c.c
mpif90 -O3 -acc=gpu -ta=nvidia \
  data.f90 definitions.f90 pack_kernel.f90 clover.F90 report.f90 timer.f90 \
  parse.f90 read_input.f90 initialise_chunk_kernel.f90 initialise_chunk.f90 build_field.f90 \
  update_tile_halo_kernel.f90 update_tile_halo.f90 update_halo_kernel.f90 update_halo.f90 \
  ideal_gas_kernel.f90 ideal_gas.f90 start.f90 generate_chunk_kernel.f90 generate_chunk.f90 \
  initialise.f90 field_summary_kernel.f90 field_summary.f90 viscosity_kernel.f90 viscosity.f90 \
  calc_dt_kernel.f90 calc_dt.f90 timestep.f90 accelerate_kernel.f90 accelerate.f90 \
  revert_kernel.f90 revert.f90 PdV_kernel.f90 PdV.f90 flux_calc_kernel.f90 flux_calc.f90 \
  advec_cell_kernel.f90 advec_cell_driver.f90 advec_mom_kernel.f90 advec_mom_driver.f90 \
  reset_field_kernel.f90 reset_field.f90 hydro.F90 clover_leaf.F90 visit.f90 \
  timer_c.o \
  -o bin/clover_leaf
```

CloverLeaf_OpenACC jobscript for reference execution

```
% cd bin
% cat run.sbatch

#!/bin/bash
#SBATCH --job-name=CL_ACC           # Job name
#SBATCH --nodes=3                   # Total number of nodes requested
#SBATCH --gres=gpu:2                 # Number of GPUs per node
#SBATCH --ntasks-per-node=2         # Number of MPI tasks per node (one per GPU)
#SBATCH --time=00:05:00             # Max. wall-clock time (hh:mm:ss)
#SBATCH --partition=small           # Queue partition

module load openmpi/nvidia/4.1.4-gpu

mpiexec ./clover_leaf

% sbatch run.sbatch
```

- Check jobscript
- then submit

CloverLeaf_OpenACC reference execution

```
% cat slurm-<job_id>.out
+ mpiexec ./clover_leaf
MPI rank    0 (0) using device    0/2 on turpancomp0
MPI rank    1 (1) using device    1/2 on turpancomp0
MPI rank    4 (0) using device    0/2 on turpancomp2
MPI rank    5 (1) using device    1/2 on turpancomp2
MPI rank    2 (0) using device    0/2 on turpancomp1
MPI rank    3 (1) using device    1/2 on turpancomp1

Clover Version 1.300
MPI Version
OpenACC Version 201711
Task Count      6

Output file clover.out opened. All output will go there.

Step    1 time    0.0000000 control sound timestep  3.85E-04   1, 1 x  3.26E-04 y  3.26E-04
Step    2 time    0.0003852 control sound timestep  2.35E-04   1, 1 x  3.26E-04 y  3.26E-04
Step    3 time    0.0006203 control sound timestep  2.99E-04   1, 1 x  3.26E-04 y  3.26E-04
[...]
Step   87 time    0.0310703 control sound timestep  3.66E-04   1, 1 x  3.26E-04 y  3.26E-04

Wall clock      6.195748090744019
First step overhead  2.0323038101196289E-002
```

- Verify the reported execution configuration and that the test execution passed

Hint: save the benchmark output (or note the run time) to be able to refer to it later