Hands-on:
CloverLeaf_OpenACC reference run

Markus Geimer
Jülich Supercomputing Centre

Tutorial exercise objectives

- Familiarise with usage of Score-P & Cube
- 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
  - …

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_bm1024_short.in input deck: 30720x30720 cells, 87 timesteps (full-length clover_bm1024.in input deck does 2955 timesteps to reach validated solution)
  - Run with 4 MPI tasks-per-node so that each has a dedicated GPU
  - Run on 1 or more nodes to see its strong scaling performance: 2 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

Setup for CloverLeaf exercises

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

```
% ssh -4 -X <yourid>@jureca.fz-juelich.de
```

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

```
% source /p/project/training2443/env.sh
```

Needed in each new shell

- Copy tutorial sources to your WORK directory (reachable via `$HOME/natESM` symlink)

```
% jsc-material-sync
% cd $HOME/natESM/PA-Course/CloverLeaf_OpenACC
```

CloverLeaf_OpenACC source directory

```
% ls
CHANGE_LOG                clover.F90                parse.f90
InputDecks/               clover_leaf.F90           read_input.f90
Makefile                  data.f90                  report.f90
PdV.f90                   definitions.f90           reset_field.f90
PdV_kernel.f90            field_summary.f90         reset_field_kernel.f90
README.md                 field_summary_kernel.f90  revert.f90
accelerate.f90            flux_calc.f90             revert_kernel.f90
accelerate_kernel.f90     flux_calc_kernel.f90      start.f90
advec_cell_driver.f90     ftocmacros.h              timer.f90
advec_cell_kernel.f90     generate_chunk.f90        timer_c.c
advec_mom_driver.f90      generate_chunk_kernel.f90 timestep.f90
advec_mom_kernel.f90      hydro.F90                 update_halo.f90
advection.f90             ideal_gas.f90             update_halo_kernel.f90
bin/                      ideal_gas_kernel.f90      update_tile_halo.f90
bin.scorep/               initialise.f90            update_tile_halo_kernel.f90
build_field.f90           initialise_chunk.f90      viscosity.f90
calc_dt.f90               initialise_chunk_kernel.f90 viscosity_kernel.f90
calc_dt_kernel.f90        pack_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 -fastsse -acc -Minfo=acc -gpu=cc80 \
    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=CloverLeaf             # Job name
#SBATCH --nodes=2                         # Total number of nodes requested
#SBATCH --ntasks-per-node=4               # MPI tasks per node (one per GPU)
#SBATCH --cpus-per-task=1                 # Threads per MPI task
#SBATCH --time=00:05:00                   # Max. wall-clock time (hh:mm:ss)
#SBATCH --partition=dc-gpu                # Job partition
#SBATCH --account=training2443            # Project account to be charged
#SBATCH --reservation=training2443_day1   # Reservation
#SBATCH --output=%x.%j.out                # Output files
#SBATCH --error=%x.%j.out

srun ./clover_leaf

% sbatch run.sbatch
```

- Check jobscript
- then submit

CloverLeaf_OpenACC reference execution

```
% cat CloverLeaf.<job_id>.out
 MPI rank     0 (0) using device 0/4 on jrc0287
 MPI rank     1 (1) using device 1/4 on jrc0287
 MPI rank     2 (2) using device 2/4 on jrc0287
 MPI rank     3 (3) using device 3/4 on jrc0287
 MPI rank     4 (0) using device 0/4 on jrc0384
 MPI rank     5 (1) using device 1/4 on jrc0384
 MPI rank     6 (2) using device 2/4 on jrc0384
 MPI rank     7 (3) using device 3/4 on jrc0384

    Clover Version 1.300
       MPI Version
   OpenACC Version 201711
   Task Count      8

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

 Step    1 time   0.0000000 control sound timestep  1.93E-04   1, 1 x  1.63E-04 y  1.63E-04
 Step    2 time   0.0001926 control sound timestep  1.18E-04   1, 1 x  1.63E-04 y  1.63E-04
 [...]
 Step   87 time   0.0155351 control sound timestep  1.83E-04   1, 1 x  1.63E-04 y  1.63E-04

 Wall clock     17.04349994659424
 First step overhead   4.0761232376098633E-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