# Grid performance summary: Benchmark_dwf

- Day 1: Baseline performance results
  - Machines configured in NPS4
  - Best results results obtained with
  - One MPI process per NUMA domain (1 quadrant/2CCDs)
  - Each MPI process running 16 threads (one per CPU)
  - 2.86 TFLOPS (compute-only)
  - 1.1 TFLOPS (including comms)
  - Single-node, 16^3 x 32 Lattice
  - Evenly distributed MPI (2 x 2 x 2 x 2)
- Day 1: observations
  - Diagnosing where bandwidth losses occur non-trivial
  - AOCC21 vs AOCC22 minimal improvement (~2%?)

Day 2 - Infinity

- Rome ∞ 32/16 bit Read/Write
  - Naples ∞ /16 bit Read/Write
- Expectations:
  - We expect to saturate DRAM
    - … but we don't
    - 90GB/s we see here equivalent to 180GB/s on DRAM, i.e. approx. 50% DRAM peak
  - We expect GET to be faster than PUT
    - But it's not (when we hit cache)
    - GET=read remote, write local
- Findings
  - Streaming intrinsics marginally worse
  - Up to L3 cache size
    - High throughput
    - (S)PUT many x faster than (S)GET (why ?)
 - Above cache size
    - (S)GET 10%? faster than (S)PUT

Bidirectional comms bandwidth vs Buffer Size
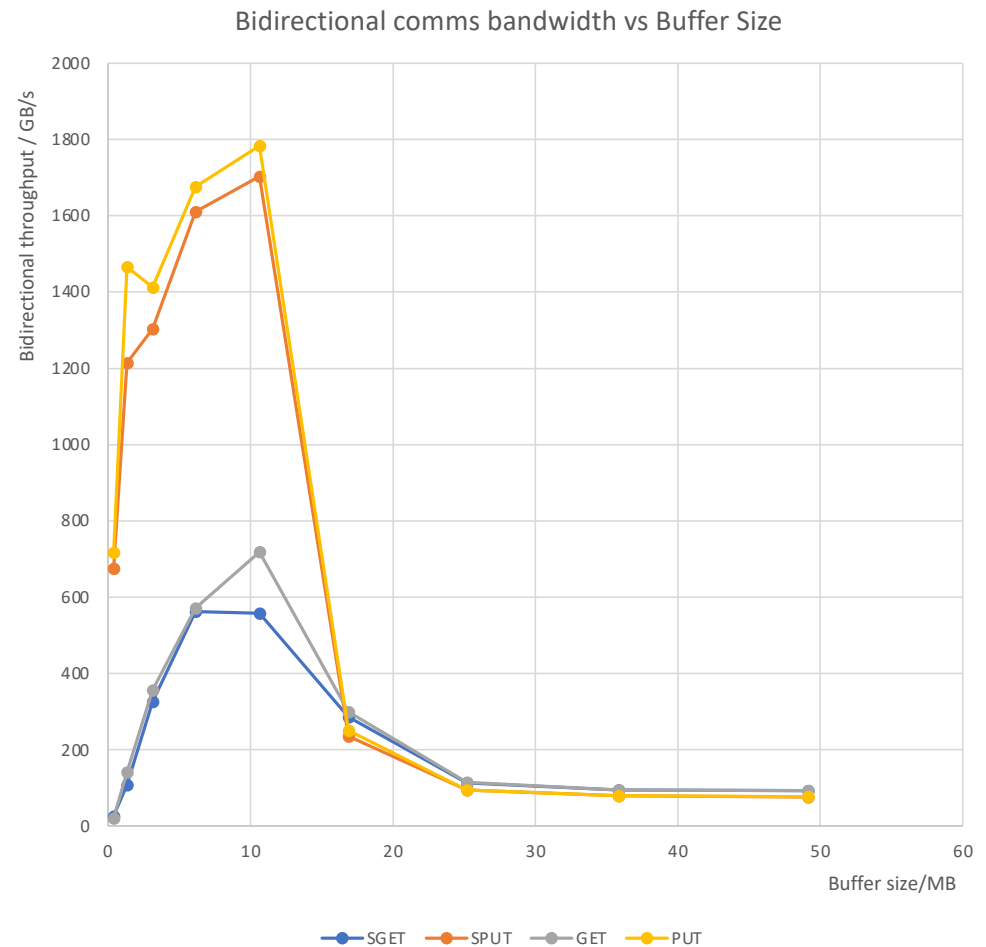
Day 2 - Other

- NPS 4 tests continued …
- Compiler flags –march=znver2
  - 16 x 16 x 32 x 32 Lattice
  - 1 x 2 x 2 x 2 MPI
  - 3.1 TFLOPS (compute-only)
  - 1.2 TFLOPS (including comms)
- Optimising MPI Layout
  - 16 x 16 x 16 x 64 Lattice
  - 1 x 1 x 1 x 8 MPI… cheating
  - 2.4 TFLOPS (compute-only)
  - 1.5 TFLOPS (including comms)

- NPS1
  - 2 MPI processes 64 threads/proc
  - 16 x 16 x 16 x 64 Lattice
  - MPI layout 1 x 1 x 1 x 2 Lattice
  - 2.4 TFLOPS (compute-only)
  - 1.9 TFLOPS (including comms)
  - Needed:
    - KMP_AFFINITY=verbose,granularity=fine,compact
- Observations
  - NPS1 looks to be optimal
    - Because we avoid interior comms
    - … i.e. buffer copies
  - Infinity link working
    - Might see 2 x boost of interior comms?
  - Cache-resident performance excellent

Final Summary::Highest performance obtained under NPS1::GFLOP/s rating of=1900

Appendix - AMDuProf
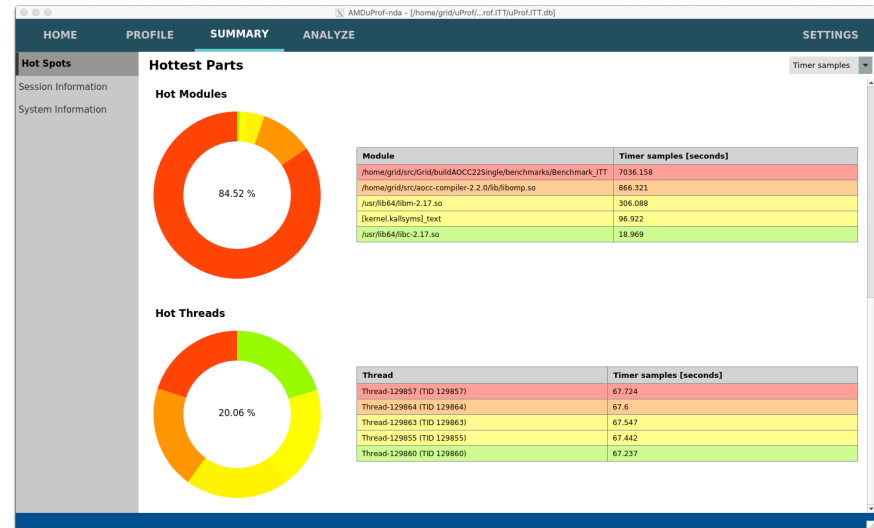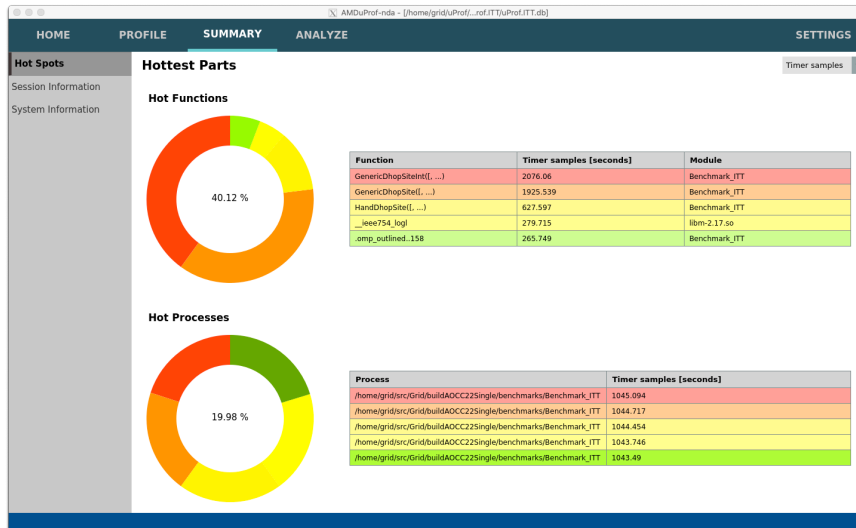
- We tried various tools, including the profiler
- We appreciated the technical support during the Hackathon