Performance portability across Arm microarchitectures

The Arm ISA provides functional portability across a range of microarchitectures from different Arm licensees.

Arm Performance Libraries provide the performance portable BLAS, LAPACK and FFT routines that the HPC ecosystem requires.

How do we quickly and effectively tune for all of these microarchitectures?

“Spotlight”
64-bit Arm Assembly Kernel Benchmark Reports

Fused multiply-add latency?
How many independent scalar FMADD instructions before we can reuse destination registers without performance penalty?

Fastest way to load 64 bytes (512 bits)?
Will two LDP instructions (each of which loads a pair of 128-bit vectors) outperform a single LD1 instruction (which loads four 128-bit vectors in one go)?

Performance penalty for unaligned loads?
In a loop, perform 20 FMLA instructions, then load 128 consecutive bytes from memory. What is the effect if we increment the base address by 0-16 bytes each time around the loop?

Acknowledgements: Chris Armstrong, Chris Goodyer and Joe Dobson from the Arm Performance Libraries team in Manchester, UK.