CATALYST UK @ EPCC

Professor Mark Parsons

EPCC Director
Associate Dean for e-Research

epcc

THE UNIVERSITY of EDINBURGH

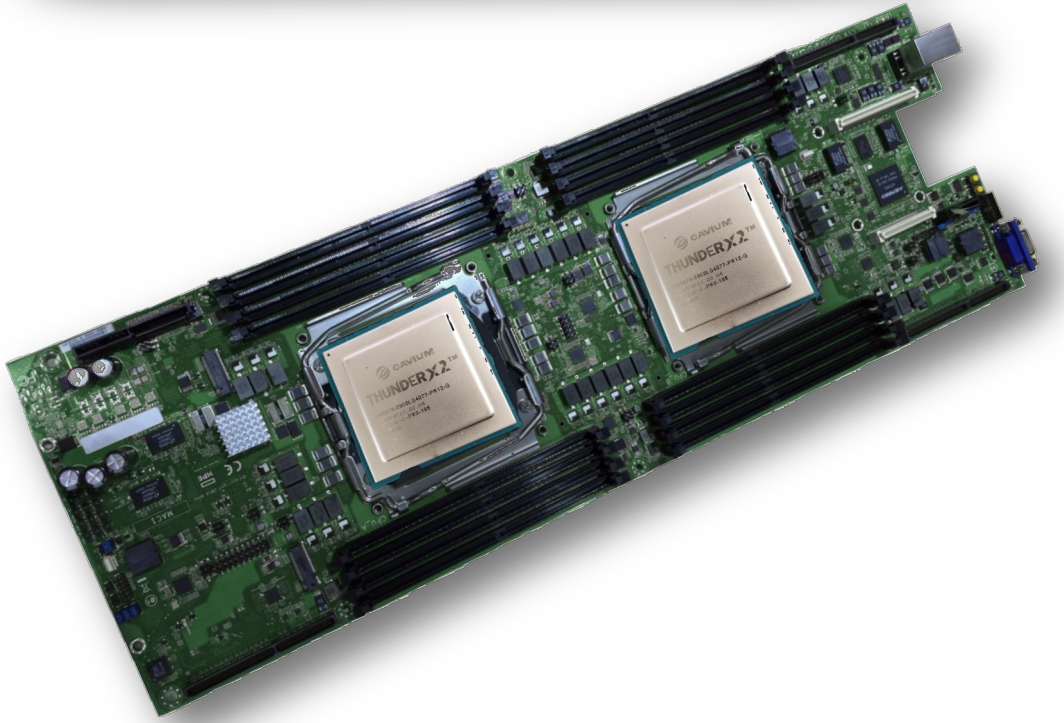
The aim of the Catalyst UK programme

- To investigate and showcase the potential of Arm-based HPC systems
- To further drive supercomputer adoption in the UK in general, and in the commercial sector in particular
- To cooperate with the UK industry to jointly develop critical applications and workflows to best exploit the Arm system capabilities
- To provide training for researchers to successfully and productively work with Arm-based systems in the future

Catalyst UK system outline

- HPE Apollo 70
- The clusters at each university will be largely identical
  - Designed, built and supported by HPE
- 64 nodes (i.e. 4096 cores) per cluster
  - 2 racks
  - Dual-socket nodes with Cavium ThunderX2 CPUs (32 cores per CPU)
  - 128GB (16 x 8GB) DDR4 per node
  - Mellanox EDR InfiniBand

HPE Apollo 70 platform



epcc

Porting plans

- Port most heavily used ARCHER and Cirrus packages
  - Make them available as modules
  - Built with different compilers, libraries, etc
- No specific performance optimisation in the first instance
  - Focus on making as many applications available as possible

Rank	Code Name	Usage / Nh	% Usage	Jobs	% Jobs	Users	Mean Job Size / nodes	Median Job Size / nodes
1	VASP	451961	16.73	9055	20.58	87	43.27	17
2	Gromacs	188732	6.99	1976	4.49	31	20.93	21
3	cp2k	155805	5.77	3762	8.55	39	38.64	22
4	CASTEP	139383	5.16	6177	14.04	31	168.06	16
5	Quantum Espresso	116045	4.30	614	1.40	10	95.66	68
6	NEMO	107968	4.00	1256	2.86	12	63.95	25
7	iIMB	102601	3.80	350	0.80	3	41.50	27
8	LAMMPS	79910	2.96	2857	6.49	26	9.31	5
9	ONETEP	65543	2.43	68	0.15	5	52.04	60
10	MITgcm	44673	1.65	761	1.73	7	15.03	12
11	WRF	44028	1.63	99	0.23	7	42.14	4
12	EPOCH	40176	1.49	85	0.19	4	181.53	80
13	Oasis	38625	1.43	231	0.53	6	19.21	17
14	NAMD	37336	1.38	178	0.40	6	18.17	21
15	Met Office UM	36773	1.36	354	0.80	9	11.37	6
16	SBLI	36258	1.34	74	0.17	3	130.63	32
17	OpenFOAM	31593	1.17	231	0.53	19	29.62	20
18	Code_Saturne	23527	0.87	201	0.46	5	16.51	11
19	SENGA	22347	0.83	54	0.12	4	53.85	9
20	CASINO	18572	0.69	80	0.18	1	13.67	8
21	CRYSTAL	17013	0.63	386	0.88	8	11.98	9
22	Nektar++	15614	0.58	19	0.04	3	98.15	50
23	FHI aims	14238	0.53	157	0.36	1	40.97	20
24	ChemShell	9855	0.36	170	0.39	6	11.45	8
25	JOREK	8946	0.33	11	0.03	1	249.04	210
26	CESM	6636	0.25	92	0.21	4	5.70	3
27	OSIRIS	5357	0.20	16	0.04	5	413.89	324
28	HYDRA	4713	0.17	61	0.14	2	12.13	2
29	Nek5000	3520	0.13	6	0.01	1	80.00	80
30	GS2	3042	0.11	18	0.04	2	35.03	22

```
mwngio@eslogin005:~> module avail gromacs

---------------------- /opt/modules/packages-archer -----------------------
gromacs/4.6.5                   gromacs/5.1.2-plumed2.3b
gromacs/4.6.5-plumed            gromacs/5.1.4
gromacs/4.6.7-plumed            gromacs/5.1.4-plumed2.3.3
gromacs/5.0.2                   gromacs/5.1.4-plumed2.4.1-adjmat
gromacs/5.1                     gromacs/2016.3(default)
gromacs/5.1-plumed              gromacs/2016.4
gromacs/5.1.1-plumed            gromacs/2018.2
gromacs/5.1.2                   gromacs/2018.2-plumed2.4.2
mwngio@eslogin005:~>
```

```
[mweiland@cirrus-login0 ~]$ module avail netcdf

--------------------------- /lustre/sw/modulefiles ---------------------------
netcdf/4.4.1                        netcdf-parallel/4.5.0-intel17
netcdf-parallel/4.5.0               netcdf-parallel/4.5.0-intel17-mpt214
netcdf-parallel/4.5.0-gcc6-mpt214
[mweiland@cirrus-login0 ~]$
```

epcc

Documentation and feedback

- Documentation is key part of driving adoption
- We will **document build processes** as part of porting activity
  - Contribute to the Arm HPC Community on GitLab
- Porting will **unearth issues with compilers, libraries, network, ...**
  - Document the issues
  - Report them to HPE/Arm/Mellanox/SUSE
  - Work with the vendor to resolve them
  - Publish any build-level or application-level fixes/patches

Performance optimisation plans

- We will **pick a handful of codes** for **in-depth optimisation**
  - We have performance numbers for ARCHER and Cirrus
  - Poor performance will be discovered as part of the initial porting exercise
- Areas of particular interest are **engineering**, **computational chemistry** **and** **weather/ocean** **modelling**
  - Applications from those domains will be given priority
  - But it's not an exclusive club...

Industry and non-HPC applications

- EPCC works with many companies
- Academic/open source codes will be ported first, but industry applications will also be considered
- We will **work with existing industry contacts** to port their applications to the Catalyst system
  - This **includes ISV codes**
  - We will be able to provide small amounts of support to such users
- Our work is also **not limited to traditional HPC**
  - Will also investigate AI/ML (Tensorflow, Caffee, ...), data analytics (Hadoop, Spark, ...), and modern programming languages (Python, Julia, ...)

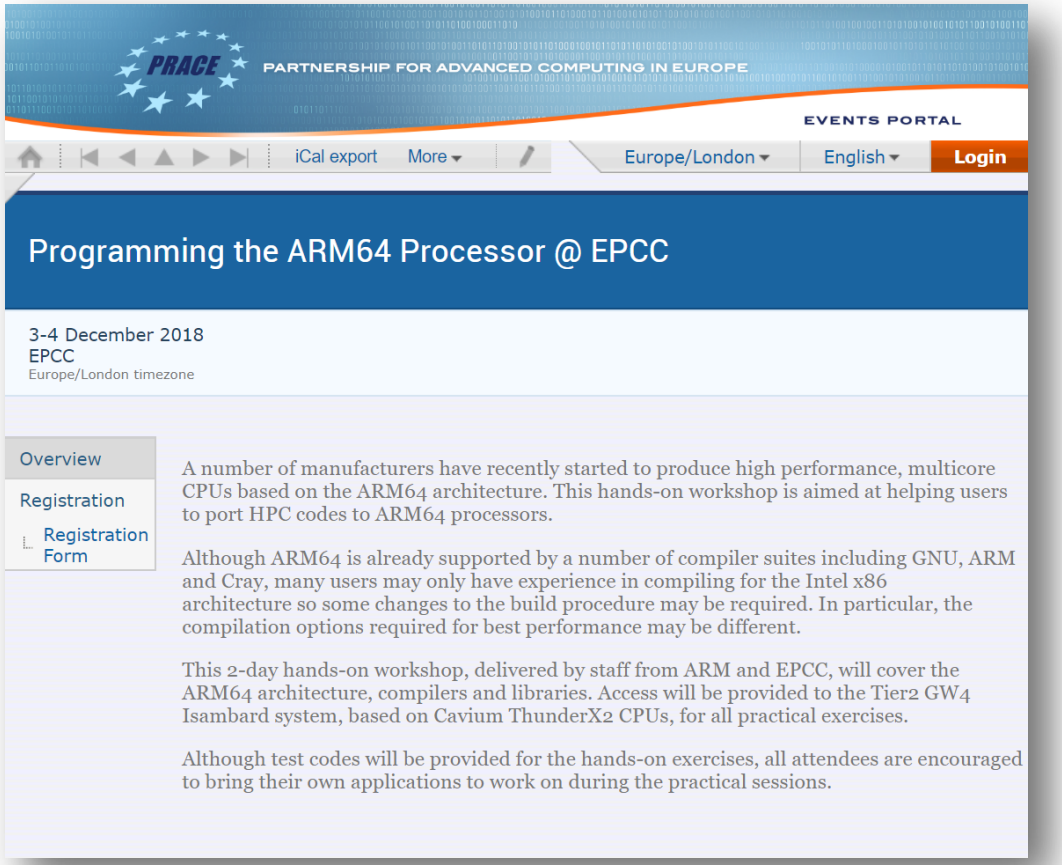
In-depth investigations

- Our in-depth investigations will be informed by porting exercise
- We will **target commonly used libraries** that underpin a wide range of applications
  - Scientific libraries such as **PETSc** and **FFTW**
  - I/O libraries such as **NetCDF, HDF5** and **MPI-IO**
  - Communications libraries
- Areas of interest are
  - Data/memory layout optimisation
  - Vectorisation, in particular the impact of SVE

Training opportunity

- PRACE Advanced Training Centre course on 3rd & 4th December
- Day 1: Lectures & Practicals
  - ThunderX2 overview, system architecture & software stack, compiling codes for ARM, hints and tips
- Day 2: Hands-on
  - Port your own code, supported by EPCC staff, Arm and HPE

https://events.prace-ri.eu/event/804/



epcc