

Dawn - new system in Cambridge

Kacper Kornet

Research Computing Services, University of Cambridge

History of GPU systems in Cambridge

2013 Wilkes, 128 nodes with 2x NVidia Tesla K40c

2017 Wilkes 2, 90 nodes with 4x NVidia Pascal P100

2021 Wilkes 3, 80 nodes with 4x NVidia Ampere A100

Accelerated computing in Cambridge

- Wilkes3
 - 90 nodes
 - 2x 64 cores AMD Milane (Zen3) CPU
 - 4x NVidia A100 (Ampere) GPUs
 - 80GB of HBM memory per GPU
 - NVLink connections between GPUs
 - 2x HDR 200 Gb/s per node
 - Number 100 on Top 500 in June 2021 (4.12 PFlops)
 - ~ 60 TFlops per node

Dawn phase 1

- Part of Artificial Intelligence Research Resource (AIRR)
- Federated service with future Isambard-AI based in Bristol
- Equally suitable for AI and simulation workloads, or combination of both
- Fastest Intel GPU system in Europe, second in the world
- 3 weeks from node delivery to Top500 runs



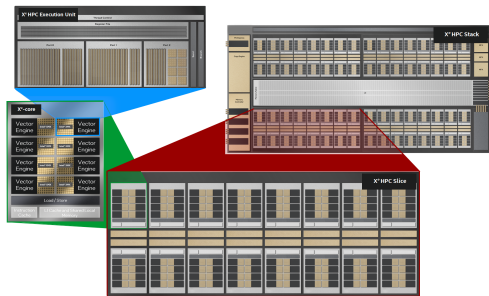
Copyright: Joe Bishop

Dawn phase 1 hardware overview

- 256 Dell PowerEdge XE9640 nodes with DLC to CPUs and GPUs
- 2x 48 cores Intel Xeon 4th gen (Sapphire Rapids) CPU
- 4x Intel Data Center Max GPU 1550 (Ponte Vecchio) GPUs
- 4x HDR 200Gb/s per node (2x in October benchmarks)
- Number 41 on Top 500 in November 2023 (19.46 PFlops) with 243 nodes
- ~ 100 TFlops per node

Intel Data Center Max GPU 1550

- 128 Xe cores in 2 stacks
- 8x 512-bit Vector Engines per Xe core
- 900MHz base frequency, 1600 MHz max frequency
- 230GB/s in both directions between stacks
- 8x Xe links per stack (26.5GB/s per link in each direction)

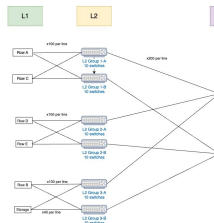
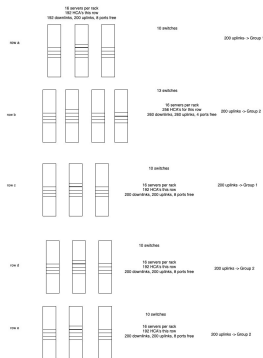


Source: Intel

- 128GB of HBM memory with 3.2TB/s of BW
- PCIe gen5 (63 GB/s)

Fabric topology

- Based on HDR 200
- Fully non-blocking fat tree (12.8TB/s of bisection BW)
- 53 L1 switches
- 100 L2 and L3 switches
- Custom cable layout to spine switches inspired by TACC
- 34km of Infiniband cables in total



Programming models

- Libraries: Intel oneAPI libraries (MKL, AI frameworks)
- Directives: OpenMP offload
- Direct programming: SYCL
- Language based parallelism (in future)

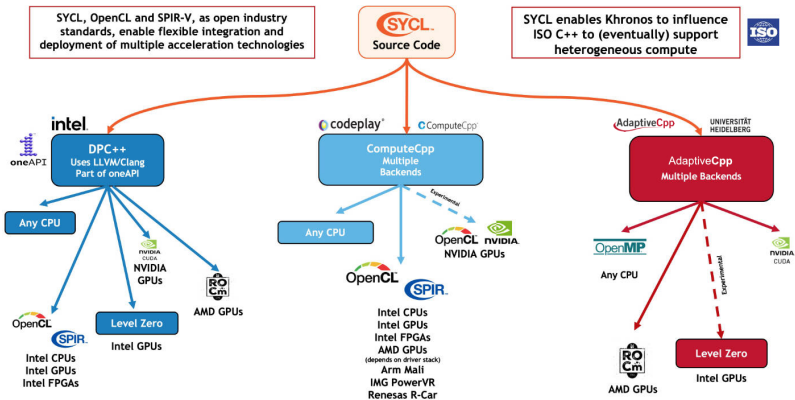
OpenMP offload

- Part of OpenMP standard since 4.0
- Much improved in recent standards
- Supports both C/C++ and Fortran
- Multiple implementations on multiple hardware

SYCL

- Open standard maintained by Khronos Group
- C++ based (mainly C++ templates)
- Multiple data models: explicit (USM) or implicit (buffers/accessors)
- Can be used GPUs from all main vendors

SYCL ecosystem



Transition from CUDA

- SYCLOmatic
 - Open source tool for one off migration from CUDA to SYCL
 - It is supposed to do 80% to 90% of work
 - DiRAC RSE report "Development of Portable Scientific Applications using SYCL"

Cambridge oneAPI Centre of Excellence

- Part of Cambridge Open Zettascale Lab
- Offers oneAPI workshops, courses and tutorial (the most recent one at CIUK'23)

Early access do Dawn

- Currently 16 nodes accessible for early access
- 256 nodes available by the end of January
- 7 invited research groups starting in December
- University of Cambridge wide call with deadline 8 Jan 2024
- Ask if you are interested (kk562@cam.ac.uk)