ExCALIBUR H&ES overview

DiRAC Science Day 2023

Alastair Basden

DiRAC philosophy

- Bespoke systems tailored for science
  - Extreme Scaling
  - Memory Intensive
  - Data Intensive
- Hardware prototyping and benchmarking
  - ExCALIBUR is essential:
    - UK preparation-for-Exascale fund

ExCALIBUR H&ES

- The Hardware and Enabling Software component of ExCALIBUR
  - Preparation for Exascale fund
- £4.5m hardware (and enabling software) fund
  - Over 4.5 years, ending March 2024
  - Funding for final year has (mostly) been announced
- Novel hardware and supporting software
  - To help the ExCALIBUR working groups and software projects
  - And the wider community
- Enabling software:
  - Libraries, toolkits, digital assets, system configurations
    - etc

ExCALIBUR projects

- RISC-V
- FPGA
- AMD GPU
- BlueField
- Rockport
- Liqid
- Composable RAM system
- DAOS file system
- Graphcore
- Cerebras
- Exascale data testbed
- ARM+GPU
- Adaptable Cluster
- NextSilicon
- Benchmarking
- Visualisation
- Quantum credits

RISC-V testbed

- Various RISC-V processors
  - Open-source Instruction Set Architecture
  - Enabling porting of user codes
    - Standard Linux environment
    - Differences in ISAs

FPGA testbed

- Access to various FPGA accelerator systems
- RSE software stack effort
- Multiple FPGA families
  - Software stack in place

AMD GPU testbed

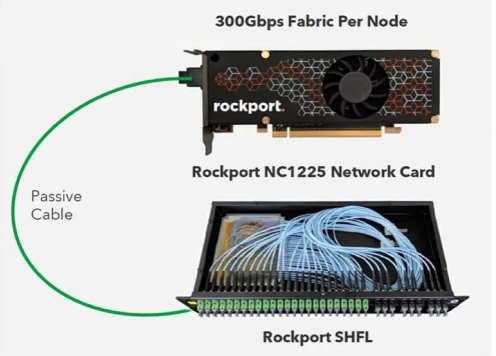
- 1x MI100 GPU
- 4x MI200 GPUs in 2 servers
- Software stack
- Ready for use
  - HIP, SYCL

NVIDIA BlueField

- Data Processing Unit (DPU)
- The DINE cluster
  - 24x NVIDIA/Mellanox BlueField-2 cards
    - Smart NIC technology
      - 200G InfiniBand
    - 24x host nodes with 32 cores, 512GB RAM
  - Allowing for offloading of algorithms and communications
  - An ARM computer on a NIC
    - 8 cores, 16GB RAM

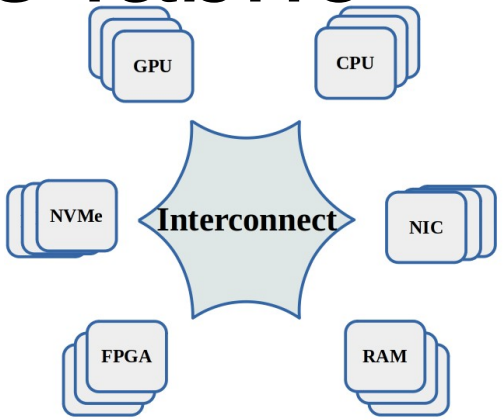
Rockport fabric

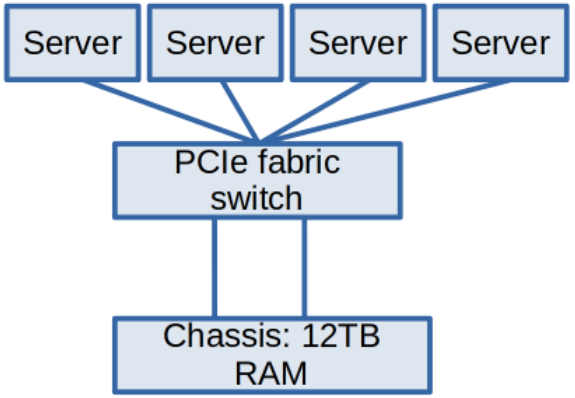
- 224 nodes of COSMA7 equipped with 100GBit/s Rockport fabric
- 6D torus topology
- Switchless network
- In production for DiRAC users
- Direct comparison with InfiniBand EDR
  - Another 224 nodes of COSMA7
- Performance has been good
  - Consistent low latency



Liqid composable fabric

- Hosting 3x NVIDIA A100 GPUs
  - Ability to move these between 4 servers in software
- And 12TB RAM
  - Move memory around at the click of a button
- Physical components no longer limited by the server
  - Dynamically provisioned servers



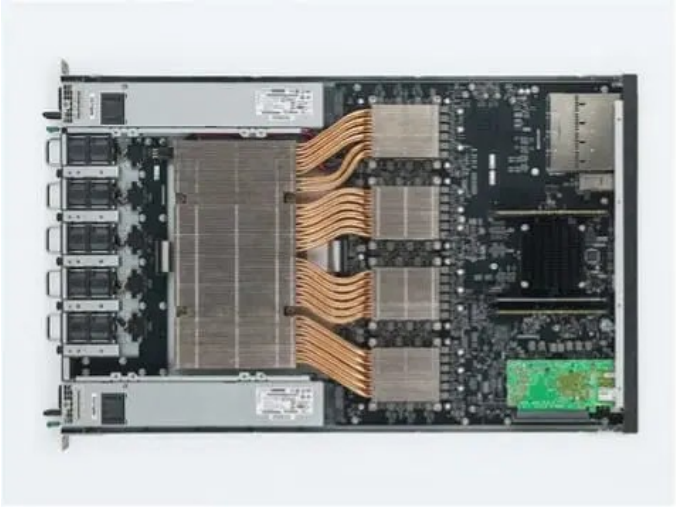


DAOS file system

- Small Intel DAOS file system
  - Attached to DINE
  - 4 storage servers
  - Total capacity about 100TB
  - User-mountable

Graphcore

- Novel AI architecture
  - Massively parallel processor
- 4x Intelligent Processing Unit (IPU)
  - ~1PFLOP Float16
  - Also Float32



Cerebras testbed

- Wafer-scale (large) processor
  - Neural network training
  - TensorFlow and PyTorch
  - Graph compiler, optimised kernels
  - Software environment CPU-GPU-FPGA combination

Exascale data testbed

- NVMe-based storage testbed
  - Variety of file systems tested
    - DAOS, BeeGFS, Lustre
  - Monitoring
  - I/O profiling

ARM+GPU cluster

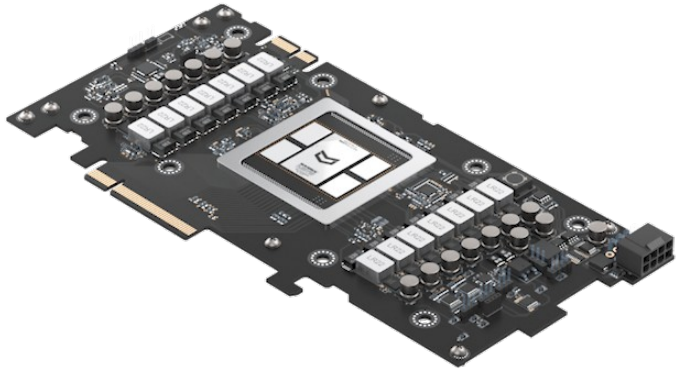
- NVIDIA ARM HPC Dev kit system
  - 4 nodes with Ampere Q80-30 processor
    - 512GB RAM
    - 2x A100GPU
    - 2x BlueField2 DPU

Adaptable Test cluster

- HDR/Ethernet comparison
  - Including BlueField DPU
- Switch-level computation
- Host offload to NIC

NextSilicon testbed

- Novel processor architecture
  - Distributed RAM
  - Cross between CPU and FPGA
  - Intelligent iterative hardware optimisation

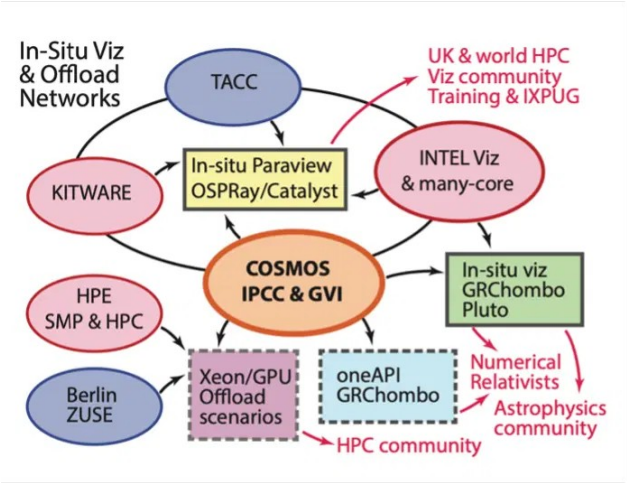


Benchmarking

- Automated benchmarking project using ReFrame
- Visualisation of complex performance portability data

Visualisation

- In-situ visualisation
- Code porting using OneAPI
- Updates to ParaView software



Quantum computing

- DWAVE Quantum annealing credits
  - Monthly quota on DWAVE system

How to use?

- excalibur.ac.uk

Future projects

- ExCALIBUR-2