Isambard: tales from the world’s first Arm-based production supercomputer
'Isambard' is a new UK Tier 2 HPC service from GW4
Tier 0: international

Tier 1: national

Tier 2: regional

Tier 3

The tiered model of HPC provision

Tier 2 HPC CENTRES

Edinburgh

Cambridge

UCL

Loughborough

Bristol

Oxford

Understanding the Human Condition

Using deep learning and Summit's advanced supercomputing, researchers are mapping patterns in human proteins and cellular systems, seeking to understand the genetic factors that contribute to diseases such as Alzheimer's and conditions such as opioid addiction.

Combating Cancer

Using scalable deep neural networks, scientists are making strides in the fight against cancer. By pairing unstructured data with deep learning on Summit, researchers can uncover hidden relationships between genes, biological markers, and the environment.

Investigating Astrophysics Data

Exploding stars reveal clues about how heavy elements seeded the universe. With AI supercomputing on Summit, physicists can simulate these phenomena at unprecedented scale, thousands of times longer and tracking 12X more elements than previously possible.

Harnessing Fusion Energy

Fusion energy—the source of the sun's energy and a potential source of clean electricity—requires reliable reactors. With deep learning on Summit, scientists at the world's largest experimental fusion reactor can explore performance criteria and optimize operations before it comes online in 2025.

AI + HPC SUPERCHARGING EARLY SCIENCE PROJECTS

Summit combines HPC and AI techniques to automate, accelerate, and drive advancements in health, energy, and engineering. In fact, new breakthroughs are already underway in the Summit Early Science projects.

PERFORMANCE SPECIFICATIONS

MADE POSSIBLE

BY NVIDIA GPU ACCELERATION

NVIDIA

Volta is the revolutionary GPU architecture bringing today's moonshots within reach. Each Volta GPU is equipped with over 21 billion transistors, 640 Tensor Cores, and 125 teraFLOPS of deep learning performance. And there are over 27,000 of them powering Summit today. Imagine what's possible.

Discover new capabilities with GPU-accelerated AI and HPC.

www.nvidia.com/hpc

©2018 NVIDIA Corporation. All rights reserved. NVIDIA, NVIDIA logo, and NVIDIA Volta are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. All other trademarks and copyrights are the property of their respective owners.

APPLICATION

PERFORMANCE

200 petaFLOPS (Double Precision), 3.3 exaOPS (Tensor operations)

PROCESSORS

(407x)

4608

6 NVIDIA Volta: 27,648 GPUs

2 IBM POWER9: 9,216 CPUs

POWER CONSUMPTION

13 megawatts

NODE INTERCONNECT

300 GBps NVIDIA NVLink

OPERATING SYSTEM

Red Hat Enterprise Linux (RHEL) version 7.4

NUMBERS OF NODES

4608

NODE PERFORMANCE

49 teraFLOPS

TOTAL SYSTEM MEMORY

>10 PB DDR4 + HBM2 + Non-volatile

NVIDIA Volta-accelerated improvements over ORNL's previous system:

8X higher performance

4X fewer nodes

5X higher energy efficiency

INTRODUCING THE WORLD'S MOST POWERFUL SUPERCOMPUTER

A NEW AGE OF SCIENTIFIC DISCOVERY

Summit is Oak Ridge National Laboratory's (ORNL) newest leadership-class system and the world's smartest and most powerful supercomputer. With more than 27,000 NVIDIA Volta™ Tensor Core GPUs paired with 9,000 IBM Power9 CPUs, Summit is the world's largest GPU-accelerated system, purpose-built for AI and high performance computing (HPC) and designed to advance scientific discovery.

FUN FACTS

1 second

Summit can perform 200 quadrillion floating-point operations per second (FLOPS). If every person on Earth completed 1 calculation per second, it would take 1 year to do what Summit can do in 1 second.

185 mi

Summit is connected by 185 miles of fiber optic cables, or the distance from Knoxville to Nashville, Tennessee.

250 PB

Summit's file system can store 250 petabytes of data, or the equivalent of 74 years of high-definition video.

At over 340 tons, Summit's cabinets, file system, and overhead infrastructure weigh more than a large commercial aircraft.

5,600 ft

Occupying 5,600 square feet of floor space, Summit is the size of two tennis courts.

340 tons
Isambard system specification

- **10,752** Armv8 cores (168 x 2 x 32)
  - Cavium ThunderX2 32 core 2.1GHz
- Cray XC50 Scout form factor
- High-speed **Aries** interconnect
- Cray HPC optimised software stack
  - CCE, CrayPAT, Cray MPI, math libraries, ...
- **Technology comparison:**
  - x86, Xeon Phi, Pascal GPUs
- Phase 1 installed March 2017
- Phase 2 (the Arm part) currently in bring-up
- £4.7m total project cost over 3 years
Cavium ThunderX2, a seriously beefy CPU

- 32 cores at up to 2.5GHz
- Each core is 4-way superscalar, Out-of-Order
- 32KB L1, 256KB L2 per core
- Shared 32MB L3
- Dual 128-bit wide NEON vectors
  - Compared to Skylake’s 512-bit vectors, and Broadwell’s 256-bit vectors
- 8 channels of 2666MHz DDR4
  - Compared to 6 channels on Skylake, 4 channels on Broadwell
  - AMD’s EPYC also has 8 channels
## Benchmarking platforms

<table>
<thead>
<tr>
<th>Processor</th>
<th>Cores</th>
<th>Clock speed</th>
<th>TDP</th>
<th>FP64</th>
<th>Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>GHz</td>
<td>Watts</td>
<td>TFLOP/s</td>
<td>GB/s</td>
</tr>
<tr>
<td>Broadwell</td>
<td>2 × 22</td>
<td>2.2</td>
<td>145</td>
<td>1.55</td>
<td>154</td>
</tr>
<tr>
<td>Skylake Gold</td>
<td>2 × 20</td>
<td>2.4</td>
<td>150</td>
<td>3.07</td>
<td>256</td>
</tr>
<tr>
<td>Skylake Platinum</td>
<td>2 × 28</td>
<td>2.1</td>
<td>165</td>
<td>3.76</td>
<td>256</td>
</tr>
<tr>
<td>ThunderX2</td>
<td>2 × 32</td>
<td>2.2</td>
<td>175</td>
<td>1.13</td>
<td>320</td>
</tr>
</tbody>
</table>

- **BDW 22c**  Intel Broadwell E5-2699 v4, $4,115 each (near top-bin)
- **SKL 20c**  Intel Skylake Gold 6148, $3,078 each
- **SKL 28c**  Intel Skylake Platinum 8176, $8,719 each (near top-bin)
- **TX2 32c**  Cavium ThunderX2, $1,795 each (near top-bin)
Key architectural comparisons (node-level, dual socket)

<table>
<thead>
<tr>
<th>Processor</th>
<th>Cores</th>
<th>TFLOPS/s</th>
<th>L1 bandwidth (agg. TB/s)</th>
<th>L2 bandwidth (agg. TB/s)</th>
<th>L3 bandwidth (agg. GB/s)</th>
<th>Memory bandwidth (GB/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broadwell 22c</td>
<td>2⇥22</td>
<td>2.2</td>
<td>1.55</td>
<td>3.46</td>
<td>2.14</td>
<td>537.6</td>
</tr>
<tr>
<td>Skylake Gold 28c</td>
<td>2⇥20</td>
<td>2.4</td>
<td>1.55</td>
<td>3.57</td>
<td>2.14</td>
<td>767.2</td>
</tr>
<tr>
<td>Skylake Platinum</td>
<td>2⇥8</td>
<td>2.1</td>
<td>2.14</td>
<td>3.76</td>
<td>2.14</td>
<td>131.2</td>
</tr>
<tr>
<td>ThunderX2 32c</td>
<td>2⇥32</td>
<td>2.2</td>
<td>2.14</td>
<td>3.57</td>
<td>2.14</td>
<td>253.4</td>
</tr>
</tbody>
</table>

**TABLE 1** Hardware information (peak figures)

**FIGURE 2** Comparison of properties of Broadwell 22c, Skylake 28c and ThunderX2 32c. Results are normalized to Broadwell.

that achieved the highest performance in each case was used in the results graphs displayed below. Likewise for the Intel processors, we used GCC 7, Intel 2018, and Cray CCE 8.5–8.7. Table 2 lists the compiler that achieved the highest performance for each benchmark in this study.

4.2 Mini-apps

Figure 3 compares the performance of our target platforms over a range of representative mini-applications.

STREAM: The STREAM benchmark measures the sustained memory bandwidth from the main memory. For the processors tested, the available memory bandwidth is essentially determined by the number of memory controllers. Intel Xeon Broadwell and Skylake processors have four and six memory controllers per socket, respectively. The Cavium ThunderX2 processor has eight memory controllers per socket. The results in Figure 3 show a clear trend that Skylake achieves a 1.64⇥ improvement over Broadwell, which is to be expected, given Skylake's faster memory speed.
Isambard’s core mission: deploying Arm in production HPC

Starting by porting/benchmarking/optimizing codes from the top 10 most heavily used on Archer:

• **VASP, CASTEP, GROMACS, CP2K, UM, NAMD, Oasis, SBLI, NEMO**
• Most of these codes are written in FORTRAN

Additional important codes for project partners:

• **OpenFOAM, OpenIFS, WRF, CASINO, LAMMPS, ...**
Performance on heavily used applications from Archer

![Performance chart]

https://github.com/UoB-HPC/benchmarks
Performance summary

• Performance is competitive with contemporary Intel processors
  • ThunderX2 is faster when memory bandwidth is critical
  • ThunderX2 is slower when FLOP/s and L1 cache bandwidth matters
  • Even in the worst case, only drops ~30% performance versus Broadwell

• Next-gen Arm CPUs will increase FLOP/s + cache bandwidth
  • Introduction of SVE will allow vector width of up to 2048-bits
  • Fujitsu A64FX chip unveiled recently with 512-bit SVE
  • Expecting 512-bits to be a common choice for server chips
<table>
<thead>
<tr>
<th>Benchmark</th>
<th>ThunderX2</th>
<th>Broadwell</th>
<th>Skylake</th>
</tr>
</thead>
<tbody>
<tr>
<td>STREAM</td>
<td>Arm 18.3</td>
<td>Intel 18</td>
<td>CCE 8.7</td>
</tr>
<tr>
<td>CloverLeaf</td>
<td>CCE 8.7</td>
<td>Intel 18</td>
<td>Intel 18</td>
</tr>
<tr>
<td>TeaLeaf</td>
<td>CCE 8.7</td>
<td>GCC 7</td>
<td>Intel 18</td>
</tr>
<tr>
<td>SNAP</td>
<td>CCE 8.6</td>
<td>Intel 18</td>
<td>Intel 18</td>
</tr>
<tr>
<td>Neutral</td>
<td>GCC 8</td>
<td>Intel 18</td>
<td>GCC 7</td>
</tr>
<tr>
<td>CP2K</td>
<td>GCC 8</td>
<td>GCC 7</td>
<td>GCC 7</td>
</tr>
<tr>
<td>GROMACS</td>
<td>GCC 8</td>
<td>GCC 7</td>
<td>GCC 7</td>
</tr>
<tr>
<td>NAMD</td>
<td>Arm 18.2</td>
<td>GCC 7</td>
<td>GCC 7</td>
</tr>
<tr>
<td>NEMO</td>
<td>CCE 8.7</td>
<td>CCE 8.7</td>
<td>CCE 8.7</td>
</tr>
<tr>
<td>OpenFOAM</td>
<td>GCC 7</td>
<td>GCC 7</td>
<td>GCC 7</td>
</tr>
<tr>
<td>OpenSBLI</td>
<td>CCE 8.7</td>
<td>Intel 18</td>
<td>CCE 8.7</td>
</tr>
<tr>
<td>UM</td>
<td>CCE 8.6</td>
<td>CCE 8.5</td>
<td>CCE 8.7</td>
</tr>
<tr>
<td>VASP</td>
<td>GCC 7.2</td>
<td>Intel 18</td>
<td>Intel 18</td>
</tr>
</tbody>
</table>
Comparison of compilers on Arm

<table>
<thead>
<tr>
<th></th>
<th>GCC</th>
<th>Arm</th>
<th>CCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>STREAM</td>
<td>97%</td>
<td>100%</td>
<td>99%</td>
</tr>
<tr>
<td>CloverLeaf</td>
<td>92%</td>
<td>95%</td>
<td>100%</td>
</tr>
<tr>
<td>TeaLeaf</td>
<td>99%</td>
<td>95%</td>
<td>100%</td>
</tr>
<tr>
<td>SNAP</td>
<td>74%</td>
<td>87%</td>
<td>100%</td>
</tr>
<tr>
<td>Neutral</td>
<td>100%</td>
<td>94%</td>
<td>85%</td>
</tr>
<tr>
<td>CP2K</td>
<td>100%</td>
<td>BUILD</td>
<td>CRASH</td>
</tr>
<tr>
<td>GROMACS</td>
<td>100%</td>
<td>91%</td>
<td>CRASH</td>
</tr>
<tr>
<td>NAMD</td>
<td>83%</td>
<td>100%</td>
<td>BUILD</td>
</tr>
<tr>
<td>NEMO</td>
<td>-</td>
<td>-</td>
<td>100%</td>
</tr>
<tr>
<td>OpenFOAM</td>
<td>100%</td>
<td>97%</td>
<td>BUILD</td>
</tr>
<tr>
<td>OpenSBLI</td>
<td>-</td>
<td>-</td>
<td>100%</td>
</tr>
<tr>
<td>Unified Model</td>
<td>84%</td>
<td>72%</td>
<td>100%</td>
</tr>
</tbody>
</table>

https://github.com/UoB-HPC/benchmarks

Exact same issues on x86
Enabling co-design of future architectures with cycle-accurate simulation

- We’ve developed a new configurable cycle accurate simulator in Bristol
- Within ~5-10% of TX2 hardware
- Highly configurable to almost any design of HPC CPU:
  - Planning a Post-K / A64fx version
  - Already supports SVE binaries
- Plan future support for x86, RISC-V, co-processors, ...
Conclusions

• Results show ThunderX2 performance is competitive with current high-end server CPUs, while performance per dollar is compelling

• The software tools ecosystem is already in good shape

• The full Isambard XC50 Arm system is coming up now, we’re aiming to have early results to share at SC18

• The signs are that Arm-based systems are now real alternatives for HPC, reintroducing much needed competition to the market
For more information

Comparative Benchmarking of the First Generation of HPC-Optimised Arm Processors on Isambard
S. McIntosh-Smith, J. Price, T. Deakin and A. Poenaru, CUG 2018, Stockholm

Bristol HPC group: https://uob-hpc.github.io/

Isambard: http://gw4.ac.uk/isambard/

Build and run scripts: https://github.com/UoB-HPC/benchmarks
Comparing performance per Dollar

• Hard to do this rigorously
  • RRP is not what anyone pays
  • Whole system cost has to be taken into account
  • Purchase price vs. TCO

• However, we can form some useful intuition
  • The following charts were generated by taking the performance results, dividing by the official published list prices of the CPUs only, then renormalizing to Broadwell
Performance per Dollar for applications

<table>
<thead>
<tr>
<th>Application</th>
<th>Broadwell 22c</th>
<th>Skylake 20c</th>
<th>Skylake 28c</th>
<th>ThunderX2 32c</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP2K</td>
<td>1.73</td>
<td>1.72</td>
<td>1.56</td>
<td>2.64</td>
</tr>
<tr>
<td>GROMACS</td>
<td>1.56</td>
<td>1.31</td>
<td>1.31</td>
<td>2.64</td>
</tr>
<tr>
<td>NAMD</td>
<td>1.31</td>
<td>1.31</td>
<td>0.57</td>
<td>2.65</td>
</tr>
<tr>
<td>NEMO</td>
<td>1.31</td>
<td>1.93</td>
<td>0.78</td>
<td>2.65</td>
</tr>
<tr>
<td>OpenFOAM</td>
<td>1.31</td>
<td>2.11</td>
<td>0.79</td>
<td>3.41</td>
</tr>
<tr>
<td>OpenSBLI</td>
<td>1.31</td>
<td>1.87</td>
<td>0.81</td>
<td>4.28</td>
</tr>
<tr>
<td>Unified Model</td>
<td>1.31</td>
<td>2.12</td>
<td>1.42</td>
<td>3.87</td>
</tr>
<tr>
<td>VASP</td>
<td>1.31</td>
<td>1.77</td>
<td>1.74</td>
<td>2.12</td>
</tr>
<tr>
<td>Geometric Mean</td>
<td>1.31</td>
<td>1.77</td>
<td>1.74</td>
<td>2.63</td>
</tr>
</tbody>
</table>

Performance Per Dollar (normalized to Broadwell)