



Non-Volatile Memory for Next Generation I/O

Michèle Weiland
m.weiland@epcc.ed.ac.uk



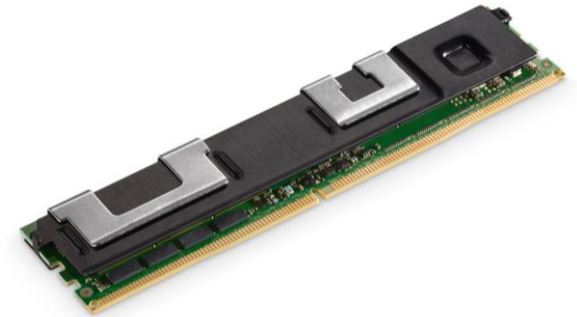
TECHNISCHE
UNIVERSITÄT
DRESDEN



NEXTGenIO project



- Research project focussed on building HPC platform with NVRAM on node
 - Chosen technology is Intel's "Optane DC Persistent Memory"
 - Can be both storage or memory
 - Or a combination of the two...
- Co-design of hardware AND system software
 - Key design requirement: applications must be able to use the system without any changes



Prototype delivered and installed

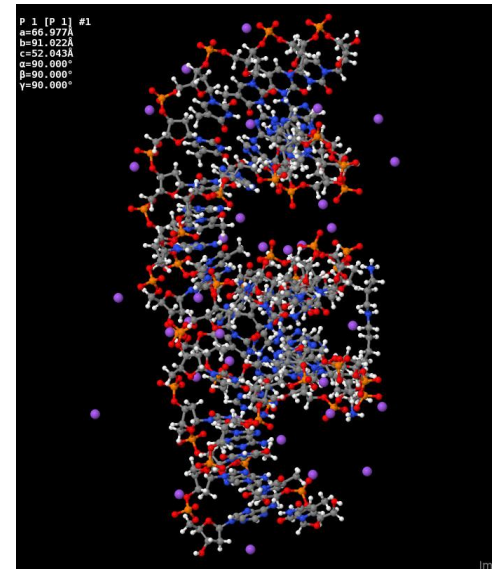


- System built by Fujitsu using bespoke motherboard
- 34 compute nodes - node configuration
 - Dual socket, 2 24-core Intel Xeon Platinum 8260M CPUs
 - 192GB DDR4 DRAM
 - 3TB DCPMM
- Omni-Path interconnect
- 270TB Lustre file system
- Total memory capacity
 - 6.5TB DRAM
 - 100TB NVRAM
- World's first large-scale deployment of DCPMM!

Memory mode



- NVRAM is transparent to user, DRAM is LLC
- No application changes required – not even recompilation
- Example: CASTEP with DNA
 - UK HPC benchmark
 - Large memory requirement
 - Fits into single node on NEXTGenIO
 - High memory watermark: 1.7TB



App Direct mode



- NVRAM has to be addressed directly
 - PMDK, system software, direct loads/stores
- Example: OpenFOAM workflow
 - Read/write between workflow steps to NVRAM
 - Create ephemeral distributed file system using GekkoFS
 - Integrated with SLURM – job creates and destroys file system
 - Data movers copy data off DCPMM after job completes



Two more points...

- SC19 paper will have detailed results

“An Early Evaluation of Intel’s Optane DC Persistent Memory Module and its Impact on High-Performance Scientific Applications”

- Fully-funded PhD studentship in Exascale storage & I/O available at EPCC
 - Focus on I/O challenges of multi-scale & multi-physics engineering simulations