



# Next Generation I/O for the Exascale

Professor Mark Parsons  
Project Coordinator

Dr Michèle Weiland  
Project Manager

EPCC, The University of Edinburgh

# I/O is the Exascale challenge

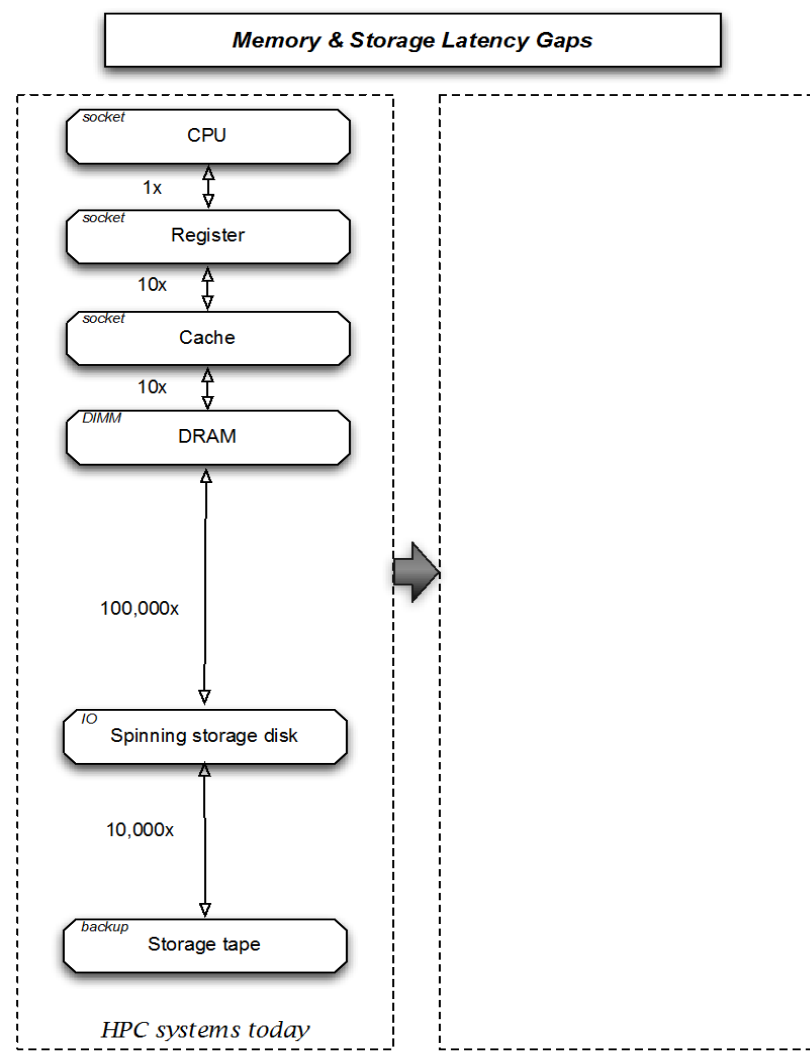


- Parallelism beyond 100 million threads demands a new approach to I/O
- Today's Petascale systems struggle with I/O
  - Inter-processor communication limits performance
  - Reading and writing data to parallel filesystems is a major bottleneck
- New technologies are needed
  - To improve inter-processor communication
  - To help us rethink data management and processing on capability systems

# A new hierarchy



- Next generation NVRAM technologies will profoundly changing memory and storage hierarchies
- HPC systems and Data Intensive systems will merge - HPDA
- Profound changes are coming to ALL data centres
- ... but in HPC we need to develop software – OS and application – to support their use

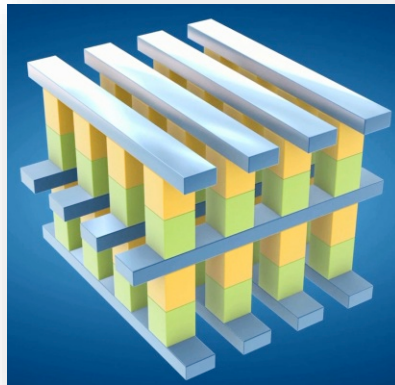


# Intel/Micron 3D XPoint Memory



## Features

- Transistorless
- Very fast compared to NAND flash
- Low power (no DRAM refresh)
- Non-volatile
- Very large
- ... and close to the CPU



## NEXTGenIO objectives

- Develop new server architecture based on next gen Intel Xeon and 3D XPoint technologies
- Investigate how best to use it in HPC – develop the software stack
- 3D XPoint is very versatile and will transform HPC

# NEXTGenIO project



## Project

- Research & Innovation Action
- 36 month duration
- €8.1 million
- Approx. 50% committed to hardware development
- Prototype system available from Month 27

## Partners

- EPCC
- INTEL
- FUJITSU
- BSC
- TUD
- ALLINEA
- ECMWF
- ARCTUR



# Key project components



- Development of a new server by Fujitsu using Intel Xeon and 3DXpoint technologies
- Development of an I/O Workload Simulator to predict how the introduction of new storage and I/O technologies will impact
- Data locality aware scheduler and energy aware scheduler e.g. temp aware job placement
- Data scheduler e.g. prefetch, writeback etc
- Application demonstrators



# How will we use this?



- Main options
  - As memory – volatile or non-volatile
  - As a file system
  - As a combination of the above
- Different use models
  - Check pointing of applications
    - Resiliency
    - Power efficiency
  - High performance parallel data storage
    - During job execution
    - Within a workflow
  - Very large memory applications

# An example: 'Hibernating' an Exascale system



- A key Exascale challenge relates to electricity costs
- Early systems will require > 50Megawatts
- NV-DIMMs give us the opportunity to
  - 'Barrier' an entire system
  - Save all DRAM data to NV-DIMM
  - Power down during a peak period e.g. dinner time
  - Restart in a matter of seconds
- Easy to negotiate lower electricity pricing with this operational mode



# Final words



- NEXTGenIO will be the first project to develop solutions using the 3D XPoint technology
- Very exciting mix of hardware and software development
- Strong team of partners
- Making good progress
- First architectural designs completed
- This may be one of the most transformational projects any of us will ever work on