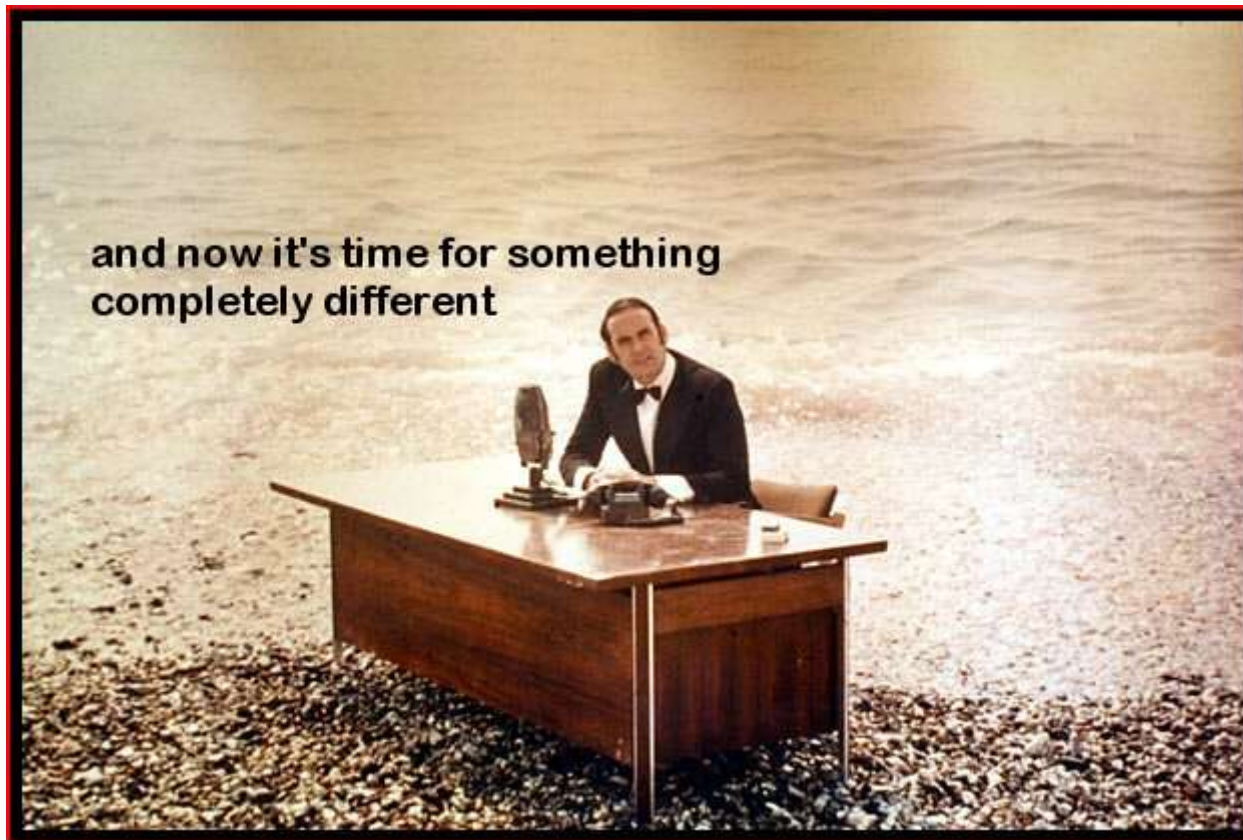


MAX DRIVING
THE EXASCALE
TRANSITION



H2020 EUROPEAN CENTRE OF EXCELLENCE _ EU GRANT # 676598



**and now it's time for something
completely different**



H2020 EUROPEAN CENTRE OF EXCELLENCE _ EU GRANT # 676598

Introducing DAVIDE

- DAVIDE is **D**eveloped of an **A**dded-**V**alue Infrastructure **D**esigned in **E**urope
- DAVIDE is the fruit of a PRACE Pre-Commercial Procurement of R&D services on «Whole System Design for Energy Efficient HPC»
- It is a solution based on IBM Power and Nvidia GPUs, integrated by E4
- Installed in Cineca from October 2017, it is still in a pre-production phase and will be opened to external users in January 2018



DAVIDE

- D.A.V.I.D.E. is composed by 45 nodes connected with an efficient Infiniband EDR 100 Gb/s networking, with a total peak performance of 990 TFlops and an estimated power consumption of less than 2kW per node.
- Each node is a 2 OU OCP form factor and hosts two IBM POWER8 Processors with NVIDIA NVLink and four Tesla P100 data center GPUs, with the intra-node communication layout optimized for best performance.

Total number (racks)	3
Total number of nodes	45 (compute) + 2 (service & login nodes)
Compute node form factor	2 OU
SoC	2xPOWER8 NVlink
GPU	4xNVIDIA Tesla P100 HSMX2
Network	2xIB EDR, 2x 1GbE
Cooling	SoC and GPU with direct hot water
Heat exchanger	Liquid-liquid, redundant pumps
Max performance (per node)	22 TFLOPs (double precision), 44 TFLOPs single precision
Storage	1xSSD SATA
Power	DC power distribution



Compute node

- Derived from the IBM® POWER8 System S822LC (codename Minsky).
- 2 OU 21" Open Rack Enclosure with integrated piping & power distribution.
- Power8-based node in OCP form-factor, with leading edge features specifically engineered for HPC workloads.
- Two IBM POWER8 with NVlink and four NVIDIA Tesla P100 HSXM2.
- Differently from Minsky, DAVIDE uses direct liquid cooling for CPUs and GPUs.
- Each compute node has a peak performance of 22 TFLOPS and a power consumption of less than 2kW.

Cooling solution

- Direct hot-water cooling (about 27 °C) for the CPUs and GPUs.
- Extremely flexible and requiring minor modifications of the infrastructure.
- Each rack has an independent liquid-liquid or liquid/air heat exchanger unit with redundant pumps.
- The system has internal pumps on the GPUs. Each Rack has its CDU.
- The compute nodes are connected to the heat exchanger through pipes and a side bar for water distribution.

The accelerator

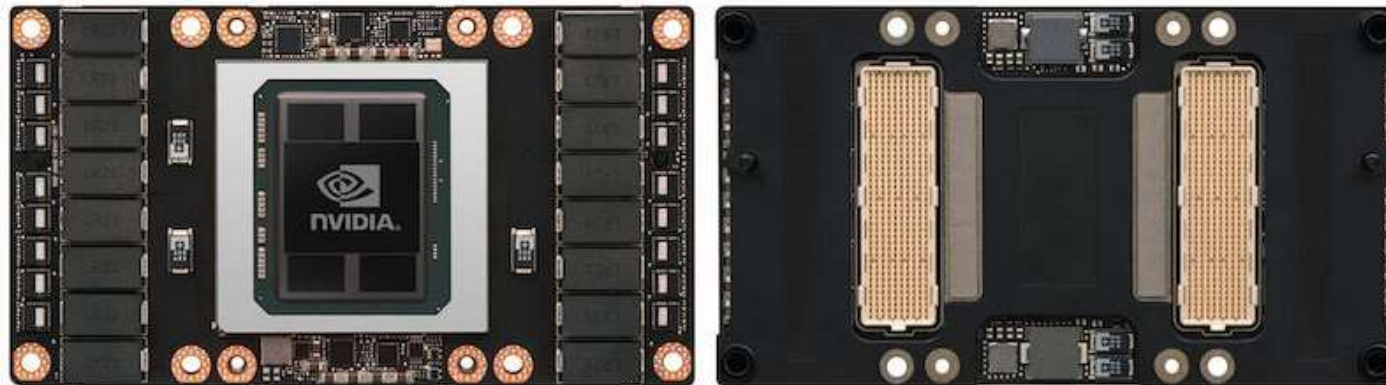
- The system is coupled with four NVIDIA Tesla P100 HSMX2 per node with NVLINK interconnect, to deliver performance for the most demanding compute applications, providing:

5.3 TFLOPS of double precision floating point (FP64) performance

10.6 TFLOPS of single precision (FP32) performance

21.2 TFLOPS of half-precision (FP16) performance

- A single link supports up to 40 GB/s of Bidirectional Bandwidth. The NVLink implementation in NVIDIA Tesla P100 supports up to four links, enabling ganged configurations with aggregate maximum bidirectional bandwidth of 160 GB/sec.



Energy monitoring

A key feature of D.A.V.I.D.E. is an innovative technology for measuring, monitoring and capping the power consumption of the node and of the whole system, through the collection of data from the relevant components (processors, memory, GPUs, fans) to further improve energy efficiency. The technology has been developed in collaboration with the University of Bologna.



	TOP500					
Rank	Rank	System	Cores	Rmax (TFlop/s)	Power (kW)	Power Efficiency (GFlops/watts)
18	440	DAVIDE - E4 OP206 cluster, IBM Power8+ 8C 3.26GHz, Infiniband EDR, NVIDIA Tesla P100 SXM2 , E4 Computer Engineering S.p.A. E4 Computer Engineering Italy	10,560	615.4	78	7.861