AGENZIA NAZIONALE PER LE NUOVE TECNOLOGIE, L'ENERGIA E LO SVILUPPO ECONOMICO SOSTENIBILE

CRESCO8: il nuovo HPC Tier-0 di ENEA

PNRR DTTU - WP2 : Plasma Modelling & HPC

Portici 9 Aprile 2025

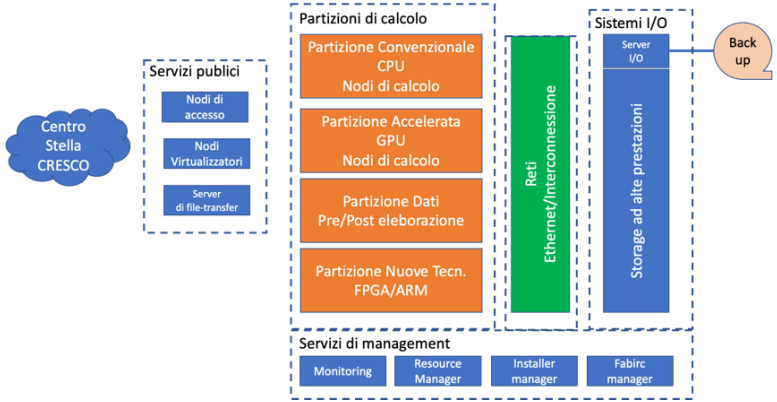
CRESCO / TERIN-ICT: F.Iannone

Il Quadro progettuale

"Rafforzamento e creazione di Infrastrutture di Ricerca" da finanziare nell'ambito del PNRR
Missione 4: "Istruzione e Ricerca"
Componente 2: "Dalla ricerca all'impresa"
Linea di investimento 3.1: "Fondo per la realizzazione di un sistema integrato di infrastrutture di ricerca e innovazione", finanziato dall'Unione europea - NextGenerationEU

Progetto: Divertor Tokamak Test Upgrade (**DTTU**)
WP2: Plasma modelling and HPC
Obiettivo: Potenziare l'infrastruttura di ricerca CRESCO

BANDO di GARA
CRESCO8



Aggiudicata al RTI

CRESCO8: i numeri

N. racks: 13 (12 SALA A + 2 SALA B)

N. nodi di calcolo: 790 (774 CPU + 15 GPU)

Potenza di picco aggregata: 9 PFlops (6 PF CPU + 3PF GPU)

Rmax (HPL): 6.5 PFlops (5.3 PF CPU + 1.2 PF GPU)

Memoria DDR5 totale: 414 TB

Storage ad alte prestazione: 10 PB

Potenza elettrica max (HPL): 800 kW

PUE: 1.045

DLC: Dry Cooler Adiabatico 1 MWf (inlet 45 °C outlet 35 °C)

CDU: 1350 kWf (95% calore rimosso)

UPS: 2 x 500 kW

Cablaggio: ~10 Km

CRESCO8: le risorse di calcolo

Partizione CPU

774 nodi = 758 EMR(DDR5 4 GB/core) + 16 SPH(DDR5 9GB/core + 128 GB HBM)

Partizione GPU

16 nodi = PonteVecchio (DDR5 4 GB/core)

Partizione Pre-Post Processing

2 nodi = 17 GB/core + Scheda Grafica Avanzata

Partizione Nuova Tecnologia

4 nodi = 2 x FPGA Xilinx (U55/U250/U280)

Front end system

2+2 nodi uguali alla partizione CPU/GPU

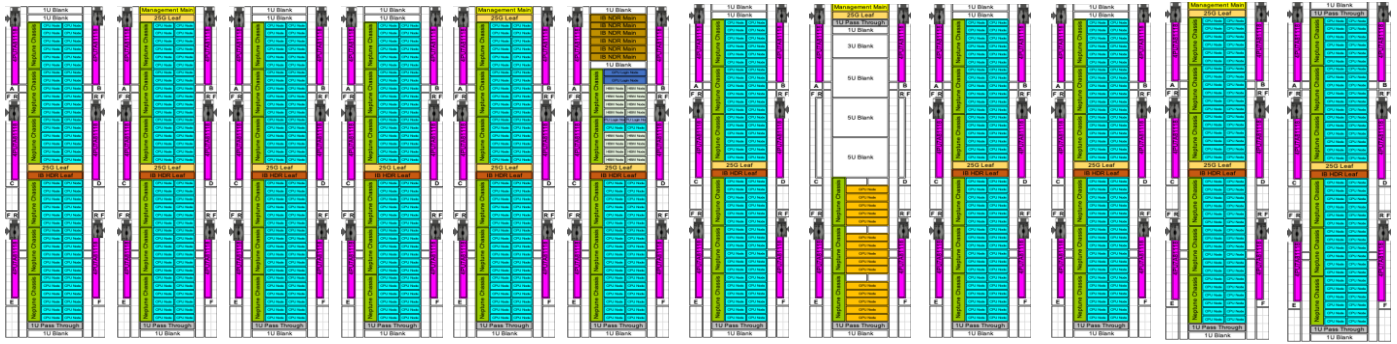
Virtualizzazione

3 nodi inclusa l'area di storage

Servizi

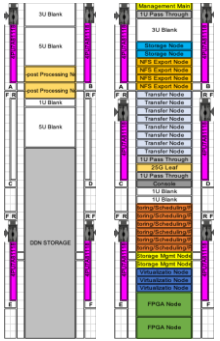
14 nodi partizione IO + 8 nodi management

SALA 3A



SALA 3B

STORAGE
DDN ES400NVX2
n.5 enclosure SS9024
n.8 porte NDR
n.4 OSS VMs
n.450 HS SAS 22TB
n.24 HS NVME 7.6 TB
10 PB raw
R/W: 65 – 50 GB/s
LUSTRE PFS

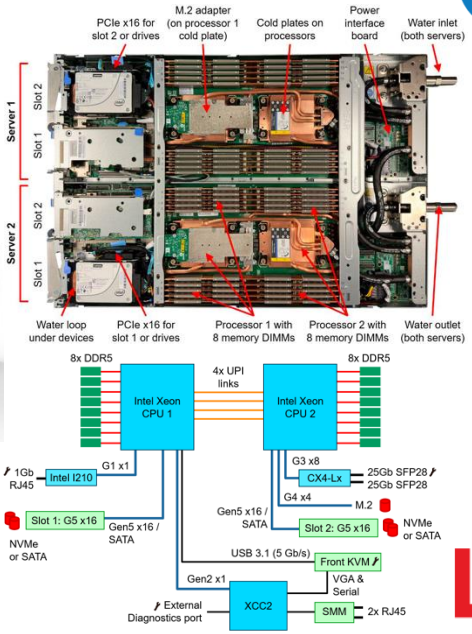


CRESCO8: i nodi di calcolo (lenovo/intel)

PARTIZIONE CPU Emerald rapids

Lenovo Thinksystem SD650 V3

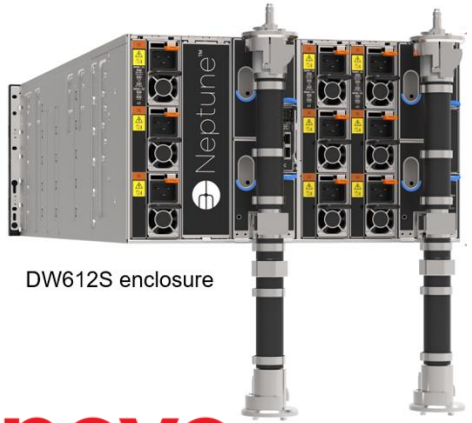
2x Intel Xeon 8592+ 64 cores, 1.9 GHz (128 core/nodo)
512 GB DDR5 5600 MHz (16 nodi con 1 TB+128 GB HBM)
1 NDR200 adapter
2x 1U - 1 kW/nodo
DLC: 95% heat removal



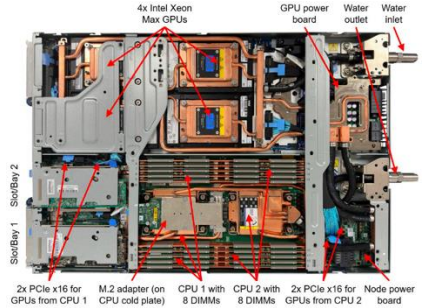
PARTIZIONE GPU Pontevecchio

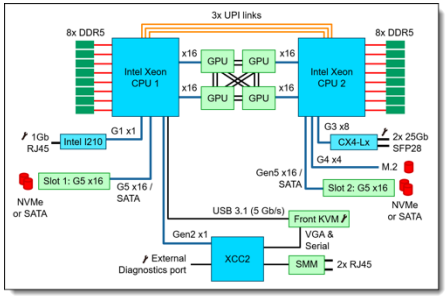
Lenovo Thinksystem SD650-I V3

2x Intel Xeon 8592+ 64 cores, 1.9 GHz (128 core/nodo)
4x Intel Max Series 1550 (PonteVecchio)
512 GB DDR5 5600 MHz
1 NDR200 adapter
1x 1U – 3 kW/nodo
DLC: 95% heat removal



DW612S enclosure



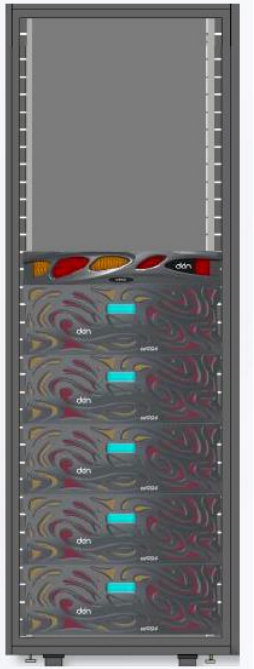


CRESCO8: lo storage ad alte prestazioni (DDN/LUSTRE)

DDN Exascaler 10 PB LUSTRE

ES400NVX2E - DDN X2 Scale out Hybrid Platform

5x SS9024 enclosures
8x IB NDR port (200 Gpbs)
4x Embedded OSS VMs
22 rack units Power ~ 8 kW

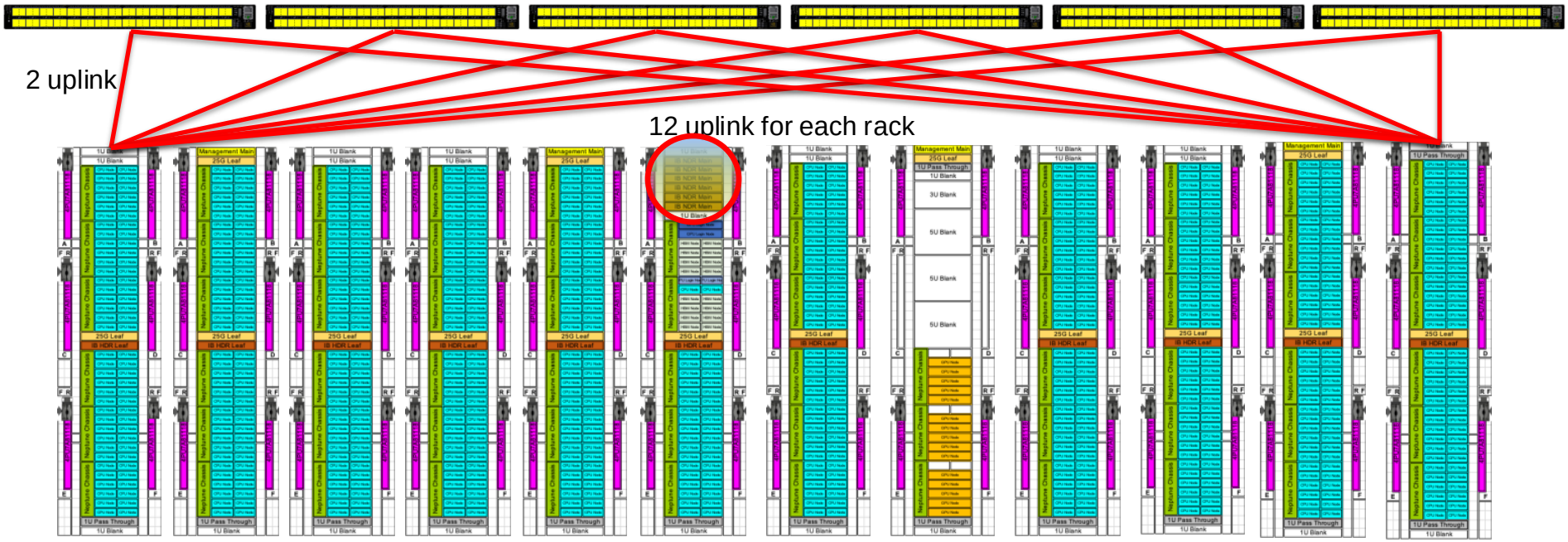


Embedded Lustre Parallel Fileserver

TOTAL RAW Capacity: **10.01 PB**
Addressable HDD Capacity: **7.66 PB**
Addressable Flash Capacity: **133.86 TB**
Number of files (Customer/Max): **4.5/12 Bilion**
Seq.Read (Flash/HDD): **~65/55 GB/s**
Seq.Write (Flash/HDD): **~50/45 GB**
Total HDDs: **450 x 22 TB**
Total NVMEs: **24 x 7.68 TB**

CRESCO8: la rete ad alte prestazioni (IB NDR 200 Gbps)

NVIDIA QM9700 64 port 400 Gbps: 6 spine / 11 leaf : 1.6:1 blocking Fat-Tree



CRESCO8: i benchmarks

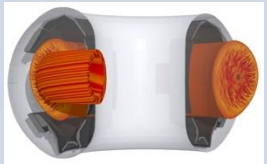
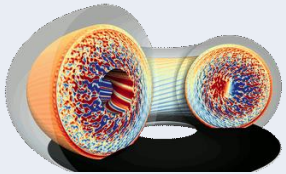
BASE

	CPU-EMR (758 nodi)	CPU-SPR-HBM (16 nodi)	GPU-EMR (15 nodi)
HPL [PFLOPS/s]	**4.5**	**0.08**	**1.18**
HPCG [TFLOPS/s]	**70.4**	**2.6**	**15.72**
STREAM [GB/s]	**552.6** 1 nodo	**1200** 1 nodo	**552.6** 1 nodo
MPI_B_EFF [MB/s]	**600000** 38 nodi	**79000** 2 nodi	**120000** 2 nodi
MPI_OSU_BW [MB/s]	**24000** 2 nodi	**24500** 2 nodi	**48000** 2 nodi
MPI_OSU_LATENCY [µs]	**2.65** 2 nodi	**1.4** 2 nodi	**1.4** 2 nodi

CRESCO8: i benchmarks su CPU

Gyro-kinetic Plasma Physics models for Nuclear Fusion

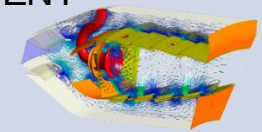
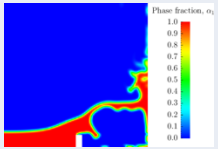
Time To Solutions

		CPU-EMR (N MPI x M Thread)		CRESCO6 (N MPI x M Thread)	
GENE		**192 s**	**329 s**	**247 s**	**376 s**
		16 nodi	8 nodi	16 nodi	8 nodi
		768 x 1	384 x 1	768 x 1	384 x 1
GYSELA		**40 s**	**65 s**	**144 s**	**227 s**
		16	8	16	8
		64 x 32	32 x 32	32 x 48	16 x 48

CRESCO8: i benchmarks su CPU (II)

Fluidodinamica numerica

Time To Solutions

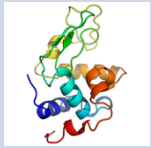
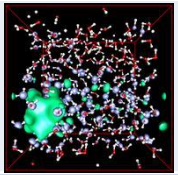
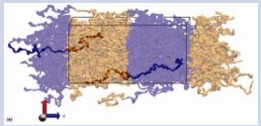
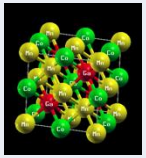
	CPU-EMR (N MPI x M Thread)		CRESCO6 (N MPI x M Thread)	
ANSYS FLUENT combustor 	**87 s**	**160 s**	**380 s**	**707 s**
	4 nodi	2 nodi	4 nodi	2 nodi
	512 x 1	256 x 1	192 x 1	96 x 1
OPENFOAM Dam break 	**72 s**	**76 s**	**193 s**	**205 s**
	16	8	16	8
	256 x 8	64 x 16	16 x 1	8 x 1

CRESCO8: i benchmarks su CPU (III)

Dinamica Molecolare

Scienza dei Materiali

Time To Solutions

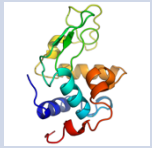
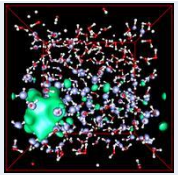
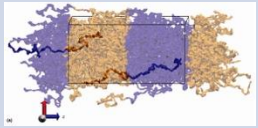
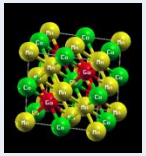
		CPU-EMR (N MPI x M Thread)		CRESCO6 (N MPI x M Thread)	
GROMACS lignocellulose		**265 s**	**463 s**	**1659 s**	**2502 s**
		16 nodi	8 nodi	16 nodi	8 nodi
		2048 x 1	1024 x 1	128 x 6	64 x 6
CP2K H2O-2048		**1425 s**	**2562 s**	**8971 s**	**-**
		8 nodi	4 nodi	16 nodi	-
		512 x 2	256 x 2	192 x 4	-
LAMMPS Lennard-Jones		**256 s**	**-**	**1160 s**	**2269 s**
		16 nodi	-	16 nodi	8 nodi
		1024 x 1	-	394 x 24	192 x 24
QUANTUM ESPRESSO tantalum pentoxide 552 atoms		**376 s**	**-**	**3187 s**	**6540 s**
		16 nodi	-	16 nodi	8 nodi
		1024 x 1	-	192 x 4	96 x 4

CRESCO8: i benchmarks su GPU (IV)

Dinamica Molecolare

Scienza dei Materiali

Time To Solutions

		GPU-EMR (N MPI x M Thread)		M100 / XCRESCO (N MPI x M Thread)	
GROMACS lignocellulose		**858 s**	**-**	**1406 s**	**1717 s**
		1 nodo	-	4 nodi	2 nodi
		8 x 28	-	16 x 32	8 x 32
CP2K H2O-2048		**1892 s**	**3578 s**	**26818 s**	**6482 s**
		8 nodi	4 nodi	16 nodi	16
		256 x 6	128 x 3	192 x 4	256 x 2
LAMMPS Lennard-Jones		**128 s**	**322 s**	**2904 s**	**5316 s**
		8 nodi	2 nodi	16 nodi	8 nodi
		1024 x 1	128 x 1	256 x 8	128 x 8
QUANTUM ESPRESSO tantalum pentoxide 552 atoms		**4980 s**	**-**	**1880 s**	**3322 s**
		1 nodo	-	8 nodi	4 nodi
		8 x 14	-	32 x 8	16 x 8

CRESCO8 : Stack Software

Ambienti operativi

Sys.Op. Linux Redhat 9.2
Parallel Filesystem : Lustre
Home/software folder : AFS
Autenticazione : Kerberos 5
Batch System : SLURM/AUKS
GUI : FARO/thinlic – OpenOnDemand

Ambienti di sviluppo

Compilers: INTEL ONEAPI – GNU
MPI: INTEL MPI – INTEL OPENMPI
Librerie Scientifiche: (MKL – SCALAPACK – PETSC – MUMPS – HDF5 – NETCDF)
Debugger: Intel Distribution for gdb - Totalview
Gestori: Moduli - Spack

Migrazione filesystems in 2 mesi (ottimistica)

Transfer filesystems & sync. unidirezionale: /gporq3 (2.5 PB) e gporq2/clima_0 (0.5 PB)
/gporq2/scratch no transfer

Ringraziamenti

ENEA:

- TERIN-ICT: G. Ponti, S.Migliori – CRESCO teams: Portici – TERIN-ACP: D.Remoli – Evento: B.Calosso, T.Procaccini
- Dir. Lavori: S.Cavaliere (B.Mastroianni)
- Direzione Centro ENEA Portici: F.D'Onza

RTI:

- Ricca IT: S.Ricca, M.Catalano, F.Brandi, D. Plantone, S.Mammola, D.Pagliaro (DoIT)
- IRTET: A.Vaghi & team
- M.C.M.: D. Sorrentino, G.Bene

Vendors:

- Lenovo: D.Schampers, A.De Bartolo
- Vertiv: M.Giordano
- INTEL: A. Luiselli, C.Bellini

TAVOLA ROTONDA

Trasformazioni in corso effetto delle transizioni (digitale ed energetica)

HPC e AI Factory stanno trasformando i data center.

Sostenibilità energetica ➡ PUE < 1.1 – DLC – Riuso del calore – Fonti rinnovabili

Sostenibilità ambientale ➡ Dimensioni – Alta densità – Modularità – Riqualificazione

Nuove frontiere dell'HPC: dalla modellistica numerica all'Intelligenza Artificiale

<u>HPC con CPU convenzionali</u>

HPC CRESCO8 (ENEA) e PITAGORA (EUROFUSION)

HPC ECMWF, CMCC, ENEA

DATI SIMULAZIONI NUMERICHE ➡ IA ➡ SIMULAZIONI SURROGATE