

CSN4 - Risorse di calcolo per il 2023

CALC4_Tier1

(a cura di L. Cosmai, M. Pepe)

Le sigle della CSN4 che fanno uso di risorse di calcolo HPC

● Lattice QCD

- LQCD123 (*F. Sanfilippo*)
- NPQCD (*M. D'Elia*)
- QCDSLAT (*L. Giusti*)
- SFT (*M. Panero*)
- SIM (*M.P. Lombardo*)
- GAGRA (*M. Papinutto*)

● Cosmology and Astroparticle Physics:

- INDARK (*M. Lattanzi*)
- NEUMATT (*R. De Pietri*)
- TEONGRAV (*B. Giacomazzo*)

● Nuclear Physics:

- MONSTRE (*N. Itaco*)
- NUCSYS (*M. Viviani*)

● Standard Model Phenomenology:

- QFTATCOL (*C. Carloni Calame*)

● Physics of the Complex Systems:

- BIOPHYS (*S. Morante*)
- ENESMA (*C. Presilla*)
- FIELDTURB (*G. Boffetta*)

● Quantum Information:

- QUANTUM (*S. Montangero*)

● Condensed matter:

- NEMESYS (*A. Sindona*)

Il consumo delle risorse di calcolo CPU al CINECA nel 2020 (*)

(*) ultimo anno con risorse MARCONI-A3 come da accordo attuativo per cofinanziamento da fondi HPC_HTC (CIPE)

CPU (MARCONI + GALILEO)	
account	consumo (corehours)
INF20_FBS	1,930,453
INF20_biophys	2,195,372
INF20_disorder	9
INF20_fldturb	10,176,062
INF20_indark	1,452,919
INF20_lqcd123	35,895,679
INF20_manybody	1,798,454
INF20_nemesys	528,331
INF20_neumatt	10,394,869
INF20_npqcd	35,990,061
INF20_qcdlat	35,383,721
INF20_qftatcol	0
INF20_quantum	516,347
INF20_sft	3,096,754
INF20_sim	5,987,693
INF20_strength	534,739
INF20_teongrav	13,574,603
INF20_test	551,823
TOTALE	160,007,889

Il 35% delle risorse può essere spostato su GPU nel 2023

Fabbisogno di CPU nel 2023: ~ 100 Mcorehours

intervista alle sigle sulla transizione alle GPU

sigla	GPU 2023
LQCD123	60% [1]
SIM	100%
ENESMA	0
BIOPHYS	70%
QFTATCOL	5%
TEONGRAV	0
NUCSYS	50%
FIELDTURB	0%
QCDLAT	15%
MONSTRE	15%
INDARK	1.5
NPQCD	60%
FIELDTURB	20%

Richieste fatte e proposte di assegnazione 2023

(CPU e GPU)

sigla	richieste 2023 per CPU (corehours)	proposta di assegnazione per CPU 2023 (corehours)	richieste 2023 per GPU (nodehours)	proposta di assegnazione per GPU 2023 (nodehours)
BIOPHYS	2,125,000	800,000	675,000	200,000
ENESMA	50,000	50,000	0	0
FIELDTURB	14,000,000	6,000,000	250,000	150,000
GAGRA	250,000	250,000	0	0
INDARK	2,400,000	1,500,000	1,600	5,000
LQCD123	15,000,000	12,000,000	600,000	520,000
MONSTRE	2,600,000	1,500,000	20,000	20,000
NEMESYS	1,200,000	400,000	600,000	20,000
NEUMATT	9,000,000	4,500,000	50,000	50,000
NPQCD	15,000,000	12,000,000	800,000	520,000
NUCSYS	1,000,000	350,000	200,000	100,000
QCDLAT	30,000,000	25,000,000	250,000	120,000
QFTATCOL	3,000,000	1,800,000	0	0
QUANTUM	300,000	300,000	25,000	25,000
SFT	4,500,000	2,800,000	4,000	20,000
SIM	4,500,000	1,000,000	3,000	100,000
TEONGRAV	15,000,000	9,750,000	100,000	150,000
TEST	0	0	0	0
Totale	119,925,000	80,000,000	3,578,600	2,000,000

- Richiesta CPU 2023**

(massima riduzione che riteniamo possibile in base alla previsione di attività dei gruppi, ma saremmo in grado, in base allo storico, di usare più risorse CPU)
- Possiamo usare più risorse GPU di quante indicate nella proposta di assegnazione**

ALTRE RICHIESTE

Richieste per piccoli sistemi CPU e GPU per attività di R&D

- **16 nodi di calcolo con 128 cores/nodo connessi con rete InfiniBand 200Gb/s bidirezionale**
- **4 nodi di calcolo con 4 GPU/nodo**

Richieste per storage su nastro al CNAF

- **850 TB**

BACKUP SLIDES

HPC COMPUTING RESOURCES 2017-2021

CINECA 2017-2021

EXPERIMENTAL PHYSICS

1.8%

OTHERS

1.4%

COMPLEX SYSTEMS

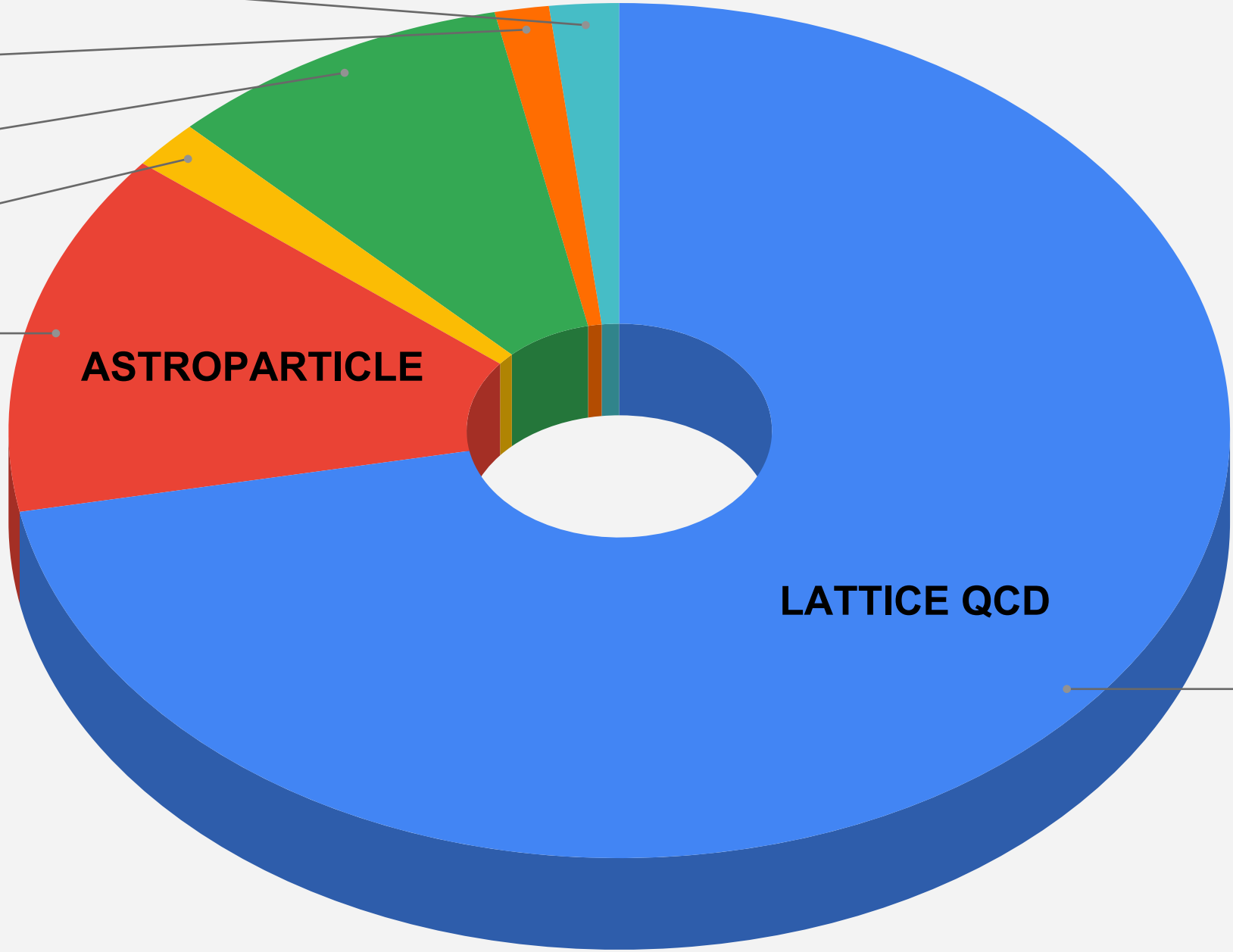
9.1%

NUCLEAR PHYSICS

1.8%

ASTROPARTICLE

13.8%

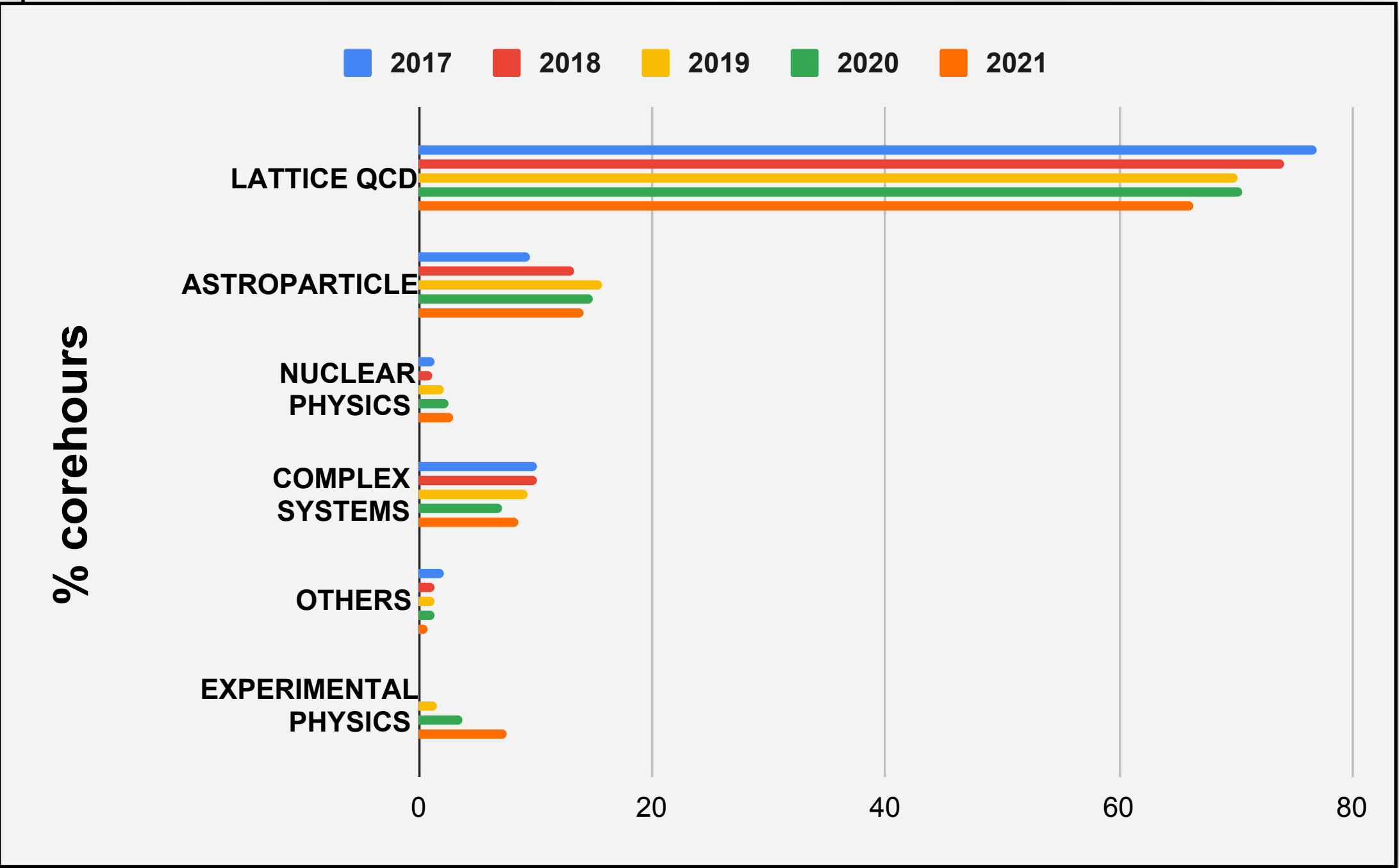


LATTICE QCD

LATTICE QCD
72.0%

AREA	corehours	%
LATTICE QCD	687,787,117	72.0
ASTROPARTICLE	131,826,951	13.8
NUCLEAR PHYSICS	17,546,676	1.8
COMPLEX SYSTEMS	87,396,389	9.1
OTHERS	13,724,017	1.4
EXPERIMENTAL PHYSICS	17,596,432	1.8
<i>TOTAL</i>	<i>955,877,582</i>	<i>100</i>

%corehours	2017	2018	2019	2020	2021
LATTICE QCD	76.97	74.09	70.06	70.62	66.31
ASTROPARTICLE	9.49	13.34	15.64	14.80	14.06
NUCLEAR PHYSICS	1.24	1.05	2.17	2.45	2.86
COMPLEX SYSTEMS	10.09	10.10	9.25	7.10	8.51
OTHERS	2.21	1.43	1.35	1.28	0.80
EXPERIMENTAL PHYSICS			1.53	3.75	7.46



Le risorse di calcolo

• Andamento storico (2012-2021)

RISORSE INFN SULLE MACCHINE CINECA A PARTIRE DA SETTEMBRE 2012 (Mcorehours)						
	FERMI	GALILEO	MARCONI (A1)	MARCONI (A2)	MARCONI (A3)	MARCONI100
Set 2012 - Dic 2013	115					
2014	110					
2015	100	15				
2016	100	15				
2017		15	18	120		
2018			9	120	164	
2019				120	164	
2020					164	10
2021					110	21

2018	ASSEGNAZIONE (Mcorehours)	CONSUMO (Mcorehours)	CONSUMO/ASSEGNAZIONE (%)
MARCONI (A1)	9.00	4.90	54.44%
MARCONI (A2)	120.00	119.00	99.17%
MARCONI (A3)	164.00	141.00	85.98%

2019	ASSEGNAZIONE (Mcorehours)	CONSUMO (Mcorehours)	CONSUMO/ASSEGNAZIONE (%)
MARCONI (A2)	120.00	119.20	99.33%
MARCONI (A3)	166.10	162.80	98.01%

2020	ASSEGNAZIONE (Mcorehours)	CONSUMO (Mcorehours)	CONSUMO/ASSEGNAZIONE (%)
MARCONI (A3)	169.00	163.00	96.45%
MARCONI100	10.11	11.93	117.98%

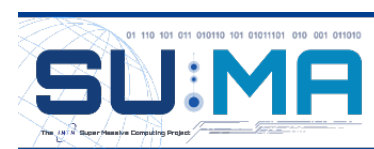
2021	ASSEGNAZIONE (Mcorehours)	CONSUMO (Mcorehours)	CONSUMO/ASSEGNAZIONE (%)
MARCONI (A3)	102.25	73.25	71.64%
MARCONI100	20.70	14.50	70.05%

Le macchine al CINECA (2012-...)

	processor	#cores/ node	#accelerators/ node	#nodes	#cores	(peak perf)/node [GFlop/s]	(peak perf)/core [GFlop/s]	peak performance [PFlop/s]	start date	end date
FERMI	IBM PowerA2	16		10,240	163,840	204.8	12.8	2.1	01-Sep-2012	18-Jul-2016
GALILEO	Intel Haswell	16		516	8,256	1,170.0	73.1	0.6	01-Feb-0205	20-Nov-2017
MARCONI-A1 (Broadwell)	Broadwell	36		1,512	54,432	1,300.0	36.1	2.0	04-Jul-2016	26-Sep-2018
MARCONI-A2 (KNL)	Knights Landing (KNL)	68		3,600	244,800	3,000.0	44.1	10.8	04-Jan-2017	01-Jan-2020
MARCONI-A3 (SKL)	Skylake	48		3,188	153,024	3,200.0	66.7	10.2	07-Aug-2017	
MARCONI100	IBM Power AC922	32	4 nVidia V100	980	31,360	32,000.0	1,000.0	31.4	04-May-2020	
GALILEO100	Intel Xeon Platinum	48	2 V100 (su 36 nodi)	554	26,592	3,530.0	73.5	2.0	15-Oct-2021	

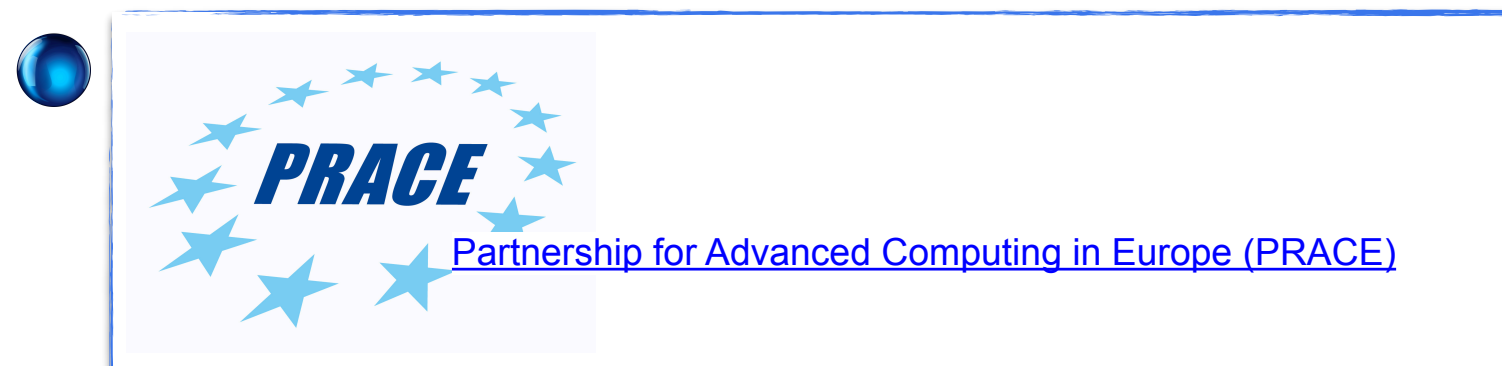
Le risorse di calcolo

Le risorse computazionali oltre l'Accordo CINECA-INFN (e.g. periodo 2017-2019)



agreement CINECA-INFN dal 2012 parzialmente finanziato da:

- progetto premiale SUMA (0.5 Meuro: GALILEO)
- progetto HPC_HTC (CIPE) (1.5 Meuro: MARCONI-SKL)



Italian SuperComputing Resource Allocation - ISCRA

	# progetti	corehours
CINECA-INFN AGREEMENT (2017-2019)		730,000,000
ISCRA-C	45	20,615,000
ISCRA-B	40	205,440,950
PRACE	30	925,928,640
ALTRE RISORSE INTERNAZIONALI	11	367,000,000
TOTALE		1,518,984,590
GRAN TOTALE		2,248,984,590
Lavori pubblicati (con uso di risorse HPC)	315	

● Moltiplicatore per accesso a risorse ISCRA e PRACE: $2,250,000,000 / 730,000,000 = 3$

Le risorse di calcolo HPC

(in via di approvazione)

- **Accordo Quadro CINECA-INFN 2022-2026**
- **Accordo Attuativo dell'Accordo Quadro CINECA-INFN**

- **MARCONI-A3**
- **MARCONI100**
- **GALILEO100**



	CPU (mhz,core, ...)	Total cores / Total Nodes	Memory per node	Accelerator
MARCONI-A3	Intel SkyLake 2x Intel Xeon 8160 @2.1GHz 24 cores each	48*3216 / 3216	192 GB	-
MARCONI100	IBM Power9 AC922 @3.1GHz 32 cores HT 4 each	32*980 / 980	256 GB	4x NVIDIA Volta V100 GPUs per node, NVlink 2.0 16 GB
GALILEO100	2 x Intel CascadeLake 8260 @2.4 GHz 24 cores each,	48*554 / 554	384 GB 3.0 TB	34 nodes with 2x NVIDIA V100 per node, PCIe3

LEONARDO

Sustained performance: 249.4 PFlop/s
Peak performance: 322.6 PFlop/s



Specifiche tecniche

- Il supercomputer Leonardo è costituito da 3 parti:
- La prima, chiamata Booster, è formata da oltre 3.000 nodi basati sui sistemi a raffreddamento diretto a liquido BullSequana XH2000, che integrano una CPU Intel Xeon SP di terza generazione e quattro GPU Nvidia Ampère A100, per un totale di quasi 14.000 acceleratori;
 - La seconda, chiamata Data Centric Partition, è formata da oltre 1.500 nodi basati sugli stessi BullSequana XH2000, ma che ospitano processori Intel Xeon di generazione futura.
 - L'intero sistema è collegato ad un terzo modulo di servizio e visualizzazione, ed ad un'area di archiviazione multi-tier ad alte performance, con una capienza aggregata di oltre 100 PetaByte, basata sui prodotti DDN. Tutte le parti del sistema comunicano tra di loro grazie un'interconnessione a bassa latenza InfiniBand HDR a 200Gb/s per collegamento, allocata su tecnologia Nvidia Networking.
- I processori Intel Xeon di generazione futura sono ottimizzati per eseguire carichi di lavoro computazionalmente intensi in sistemi di calcolo ad alte prestazioni consentendo computing con avanzate capacità di accelerazione AI integrate.
- Più di 136 BullSequana XH2000 rack di raffreddamento diretto a liquido
 - 250 PFLOPs HPL Linpack Performance (Rmax) aggregata
 - 3+PB RAM
 - 10 ExaFLOPs di prestazioni FP16 AI
 - 3456 server equipaggiati con Intel Xeon Ice Lake e GPU con architettura Nvidia Ampère
 - 1536 server con processori Intel Xeon SP di prossima generazione
 - 5PB di storage ad alte prestazioni
 - 100PB di storage a grande capacità
 - Larghezza di banda di interconnessione minima di 200Gb/s per singolo link InfiniBand HDR
 - 9MW di Potenza elettrica aggregate
 - Oltre 140 km di cablaggi in fibra ottica

Le risorse di calcolo HPC

● Accordo Quadro CINECA-INFN 2022-2026

ACCORDO QUADRO DI COLLABORAZIONE PER LO SVOLGIMENTO DI ATTIVITÀ DI RICERCA E SVILUPPO NEL SETTORE DEL CALCOLO SCIENTIFICO AD ALTE PRESTAZIONI (HPC) IN AMBITO DI FISICA DELLE ALTE ENERGIE, FISICA ASTROPARTICELLARE, FISICA NUCLEARE

Art. 2 – Oggetto e finalità dell'Accordo

Con il presente Accordo Quadro, le Parti instaurano una collaborazione rivolta allo

- sviluppo e sperimentazione di sistemi di calcolo ad alte prestazioni e ad alta capacità;
- sviluppo di algoritmi e applicazioni relative all'ambito della modellazione e della simulazione numerica e dell'analisi dei dati in aree di interesse della fisica fondamentale, in particolare utilizzando architetture innovative dei processori, delle reti d'interconnessione e delle strutture di input-output;

● Accordo Attuativo dell'Accordo Quadro CINECA-INFN

ACCORDO ATTUATIVO DELL'ACCORDO QUADRO DI COLLABORAZIONE PER LO SVOLGIMENTO DI ATTIVITÀ DI RICERCA E SVILUPPO NEL SETTORE DEL CALCOLO SCIENTIFICO AD ALTE PRESTAZIONI (HPC) IN AMBITO DI FISICA DELLE ALTE ENERGIE, FISICA ASTROPARTICELLARE, FISICA NUCLEARE

● **MARCONI-A3:** 60 Mcorehours/anno fino a dismissione

● **MARCONI100:** 60 Mcorehours/anno fino a dismissione

● **GALILEO100:** 6 Mcorehours/anno fino a dismissione

● **LEONARDO**

PARTIZIONE	2023	2024	2025	2026
General Purpose (CPU)	165 nodi	300 nodi	300 nodi	300 nodi
Booster (GPU)	3 Mnodehours	3 Mnodehours	3 Mnodehours	3 Mnodehours

su 1536 nodi *Intel Sapphire Rapids (48 cores)*

su 3456 nodi *Intel Ice-Lake (32 cores)*

Le risorse di calcolo HPC

ISCRA: Italian SuperComputing Resource Allocation



CINECA, through the **Italian SuperComputing Resource Allocation - ISCRA**, releases Call for Proposals.

CINECA, the Italian most powerful HPC center, twice a year **will directly award in excess of 100 millions core hours**, to ensure an adequate supply to scientists and engineers for HPC-related research.

CINECA infrastructure offers different HPC resources to its users. The available resources are divided in three categories:

- The TIER-0, top level computing resources, which is the new **MARCONI100** machine and can be accessed through class B and C projects.
- The TIER-1 level, is **GALILEO100** (BROADWELL) and can be accessed through class C and B projects.
The Big Data resources and new state-of-the-art public CLOUD.
The data infrastructure is available for data analysis, visualization, post-processing, bio-informatics applications.
The **CLOUD** infrastructure integrates and completes the HPC ecosystem, providing a tightly-integrated infrastructure that covers both high performance and high flexible computing. The flexibility of the CLOUD will better adapt to the diversity of user workloads, while still providing high-end computing power.
- The DGX resources:
Available through ISCRA-C: these projects aim to support researchers in the domain of Machine Learning and Artificial Intelligence and their applications.
- The Quantum Computing Resources.
Available through ISCRA-C, CINECA makes quantum computing resources of various kinds available to its users. It will be possible to request machine time on our Tier-0 system to emulate complex quantum circuits (both general purpose and special purpose type) and directly request quantum computing resources (currently of quantum annealing type).

Class B projects are received twice a year. They go under peer-review evaluation and a 3 month delay is expected before your project gets access to HPC resources. For each user it is allowed to have only one class B project each 6 months as Project Investigator.

Class C projects are received through continuous submission and reviewed once per month. An average period of about 15 days is required for activating the project. For each user it is allowed to have only one class C project each 6 months as Project Investigator.

Le risorse di calcolo HPC

PRACE, EuroHPC JU (Joint Undertaking)



- International, not-for-profit
- 25 member countries - pan-European supercomputing infrastructure enabled by 5 Hosting Members
 - BSC (Spain), CINECA (Italy), ETH/CSCS (Switzerland), GSC (Germany), GENCI (France)
- Access on basis of peer-reviewed (open) science & industry proposals
- Additionally, schools, workshops, PRACE IP ...



- Based on 2 EU Council Regulations (2018 & 2021)
- Objectives
- Build and operate a world class integrated HPC and data infrastructure
 - Enable member states to improve HPC competency
 - Foster HPC skills, education and training
 - Develop HPC core technology
 - *European Processor Initiative (EPI)*
 - *Energy efficient HPC*
 - *Quantum Computing*

- Enable high-impact scientific & engineering discovery and R&D.
- Tier-0 computing and data management resources and services through **competitive peer review**.
- Training & education via schools, workshops, seminars.
- Enabled by 5 Hosting Centres.
- **No procurement, no technology development.**

In 2021 EuroHPC planned resources and timelines

Infrastructure and Operations

- **5 x petascale machines** @ 15-30 petaflops. In operation by Q1 2021. Bulgaria, Czech Republic, Luxembourg, Portugal, Slovenia.
- **3 x pre-exascale machines**. In operation Q2/3 2021:
 - CSC - Finland *Cray: 375PF sustained, 552PF peak; GPU, x86, data analytics, cloud container partitions*
 - Cineca - Italy *ATOS-Bull Sequana 249PF sustained, CPU-GPU, DDR5 & local NVM for data analysis*
 - BSC - Spain
- 2 x exascale in operation by 2022-2023. In coordination with EPI (European-based technology) for at least one machine?
- Data I/O, management, storage and security federation for impact: GEANT, EOSC etc

Access

There is an access policy at <https://eurohpc-ju.europa.eu/access-our-supercomputers>

Talk S. Ryan @ GGI Mini Workshop "Phase transitions in particle physics"