

INAF



INAF
DS - ICT

HPC: MoU INAF-Cineca 2017-2020

Report intermedio

U. Becciani, G. Mulas, G. Murante, P. Rossi

ICT Workshop
12 Settembre, 2018
Catania

Mou/Framework INAF - Cineca.

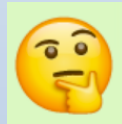
→ Accordo valido per 3 anni (con estensione di 6 mesi) operativo dal 1 Maggio 2017

→ Risorse annuali di calcolo per totali **50 Milioni di ore cpu/core Marconi KNL** ma utilizzabili su tutti i sistemi aperti del Cineca. **150 TB spazio**



- **Oltre tale risorsa si ottiene anche un ulteriore 10% qualora tutte le ore disponibili siano consumate. Quindi il potenziale totale è di 55 Mhours/anno**

→ **Flagship Project: mai attivati**



→ **Prevista una unità di personale INAF @ Cineca.**



➤ Sono state fatte due selezioni. Abbiamo avuto anche ottimi candidati e utilmente selezionati. Ma hanno rinunciato al momento della stipula del contratto

➤ Cambio di strategia: **in via di risoluzione!!**

Infrastruttura HPC al Cineca

Marconi

Partizione A1 (Broadwell), 2 PFlop/s

Nodi: 1.512, Processori per nodo: 2 x 18-cores Intel Xeon E5-2697 v4 (Broadwell) a 2.30 GHz RAM: 128 GB/node **54.432 core in totale**

Partizione A2 (Knights Landing), 11 PFlop/s.

Nodi: 3.600, Processori per nodo: 1 x 68-cores Intel Xeon Phi 7250 (Knights Landing) a 1.40 GHz, RAM: 96 GB/node **244.800 core in totale**

Partizione A3 (Skylake), 20 PFlop/s

Nodi: 1512, Processori per nodo: 2 x 24-cores Intel Xeon 8160 a 2.1 GHz, RAM: 193 GB/node **72.576 core in totale**

Rete di connessione basata su Omni-Path e filesystem GPFS

Galileo

Non più aperta per la ricerca ma dedicata alle applicazioni industriali (eccetto per particolarissime condizioni)

Nodi: 360, Processori per nodo: 2 x 18-cores Intel Xeon E5-2697 v4 (Broadwell) a 2.30 GHz RAM: 128 GB/node **15 nodi hanno 2 nVidia K80 GPU.**

Infrastruttura HPC al Cineca

D.A.V.I.D.E

(Development of an Added Value Infrastructure Designed in Europe)

Fa parte del progetto europeo Prace Pre-Commercial Procurement (PCP) per sviluppare un sistema completo per HPC ad alta efficienza energetica. Basato su server OpenPOWER. **Aperto su richiesta.**

Architettura: OpenPower NViDIA NVLink

Nodi: 45 x (2 Power8+4Tesla P100) + 2 (service&login nodes)

Rete Infiniband 100 Gb/s, Peak Performance: 1 Pflops

Pico-Cloud

Classi di applicazioni: "BigData", relative alla gestione e all'elaborazione di grandi quantità di dati, provenienti sia da simulazioni che da esperimenti.

Modelli supportati: **Urgent Computing** e (prevalentemente) **Cloud Computing – OpenStack.**

Compute/login node	66	Intel Xeon E5 2670 v2 @2.5Ghz	20 core/node	128 GB	
Visualization node	2	Intel Xeon E5 2670 v2 @ 2.5Ghz	20 core/node	128 GB	2 GPU Nvidia K40
Big Mem node	2	Intel Xeon E5 2650 v2 @ 2.6 Ghz	16 core/node	512 GB	1 GPU Nvidia K20
<u>BigInsight</u> node	4	Intel Xeon E5 2650 v2 @ 2.6 Ghz	16 core/node	64 GB	32TB of local disk

I Progetti Supportati

Il comitato di gestione ha definito e mantenuto 3 classi di progetto

- **TEST** : effettuare test di scalabilità, verifica dell'adeguatezza del proprio codice all'infrastruttura, effettuare brevi test con differenti configurazioni ecc
 - ❑ **Domanda a sportello**, assegnazione per durata di un mese (eventualmente prolungabile). A ciascun Progetto sono allocati massimo 100.000 ore

- **Classe B**: esecuzioni di programmi di **media richiesta** in termini computazionali.
 - ❑ **Domande a sportello**. A ciascun Progetto sono allocati massimo 1.000.000 di ore, durata prevista per massimo 6 mesi.

Dal Settembre 2017 ad Aprile 2018 la sottomissione a sportello è stata sospesa per over-subscription.

- **Classe A**: esecuzioni di programmi di **elevata richiesta** in termini computazionali, associati ad obiettivi scientifici di rilevanza internazionale
 - ❑ **Domande in risposta a call for application**. A ciascun Progetto sono allocati massimo 6.000.000 di ore. Durata prevista per massimo 12 mesi.

NOTA: TUTTE le sottomissioni avvengono attraverso il sito ICT/computing sezione Cineca:

<https://www.ict.inaf.it/computing/cineca/>

I Progetti Supportati

Sono state emesse fino ad oggi 3 Call for Application

Aprile 2017, Novembre 2017, Maggio 2018

L'assegnazione delle risorse è stata fatta tenendo con un **continuo monitoraggio dell'effettivo utilizzo delle risorse**. Evitando quindi l'assegnazione di risorse che qualora non consumate non sono più riutilizzabili.

- **Classe A:** L'assegnazione è stata tipicamente suddivisa in 3 periodi di 4 mesi e per circa il 33% delle risorse richieste per ogni periodo
- **Classe B:** L'assegnazione è stata tipicamente suddivisa in 2 periodi di 3 mesi e per circa il 50% delle risorse richieste per ogni periodo

Call 1 - Aprile 2017: Approvati 23 Progetti

Risorse Assegnate

N. 7 Progetti di Classe A → 39 Mhours assegnate

N. 9 Progetti di Classe B → 7.7 Mhours assegnate

N. 7 Progetti TEST → 1 Mhours assegnate

Totale 47.9 Mhours assegnate
(attesi 25 Mhours → Over-subscription fattore circa 2)

Risorse Allocate e Utilizzate

N. 7 Progetti di Classe A → 29 Mhours

N. 9 Progetti di Classe B → 7 Mhours

N. 7 Progetti TEST → 1 Mhours

Totale 37 Mhours assegnate
(fattore di allocazione e utilizzo complessivo 99%)

TUTTI I PROGETTI SONO GIA' CONCLUSI

Call 2 - Novembre 2017: Approvati 11 Progetti

Call 2. Progetti ancora in corso: Risorse Assegnate

N. 6 Progetti di Classe A → 17 Mhours assegnate

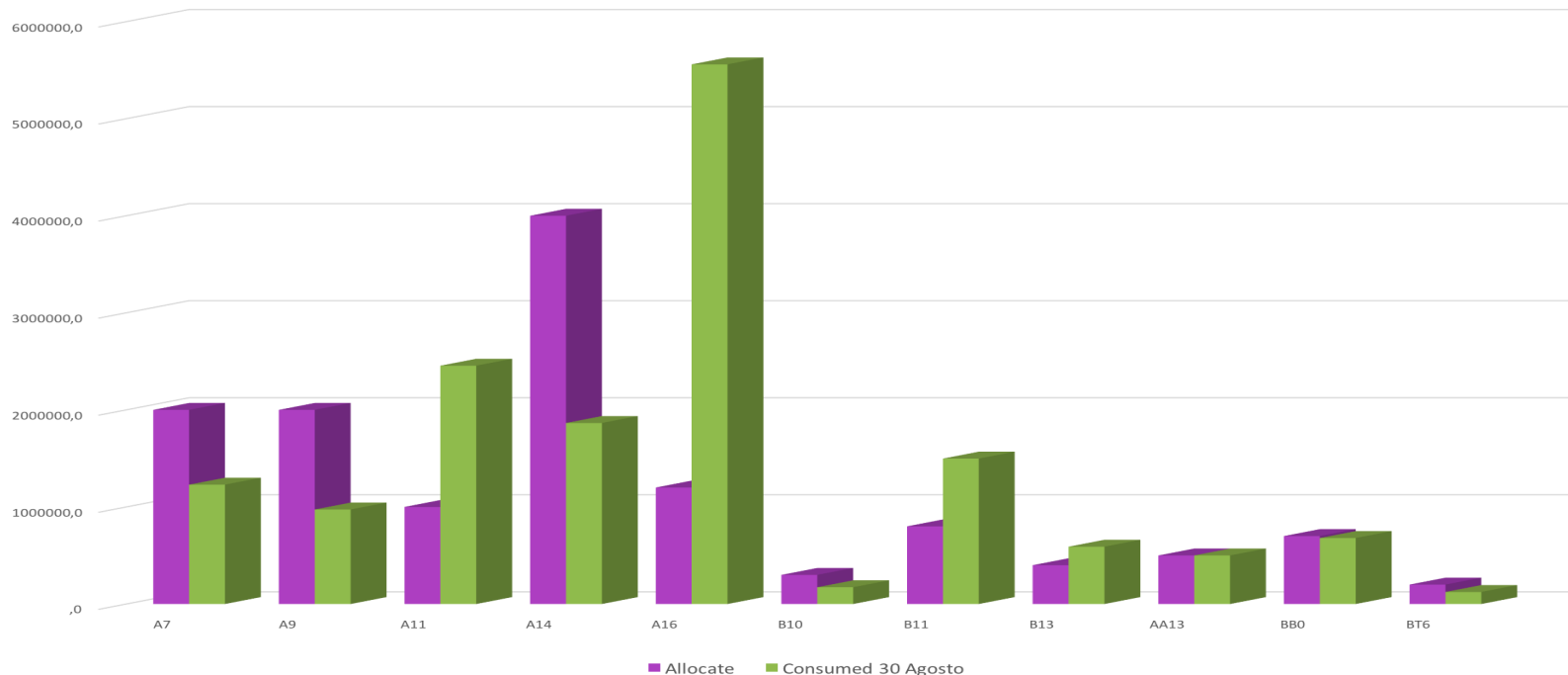
N. 4 Progetti di Classe B → 2.6 Mhours assegnate

N. 2 Progetti TEST → 0.2 Mhours assegnate

Totale 19.8 Mhours assegnate . Al 30 Agosto consumate XXX Mhours

**MA LE RICHIESTE SONO STATE MOLTO PIU' ALTE== > 50 MHours
DRASTICO TAGLIO! Over-subscription fattore 2 mantenuto!**

Call 2 Risorse Allocate e Consumate al 30 Agosto 2018



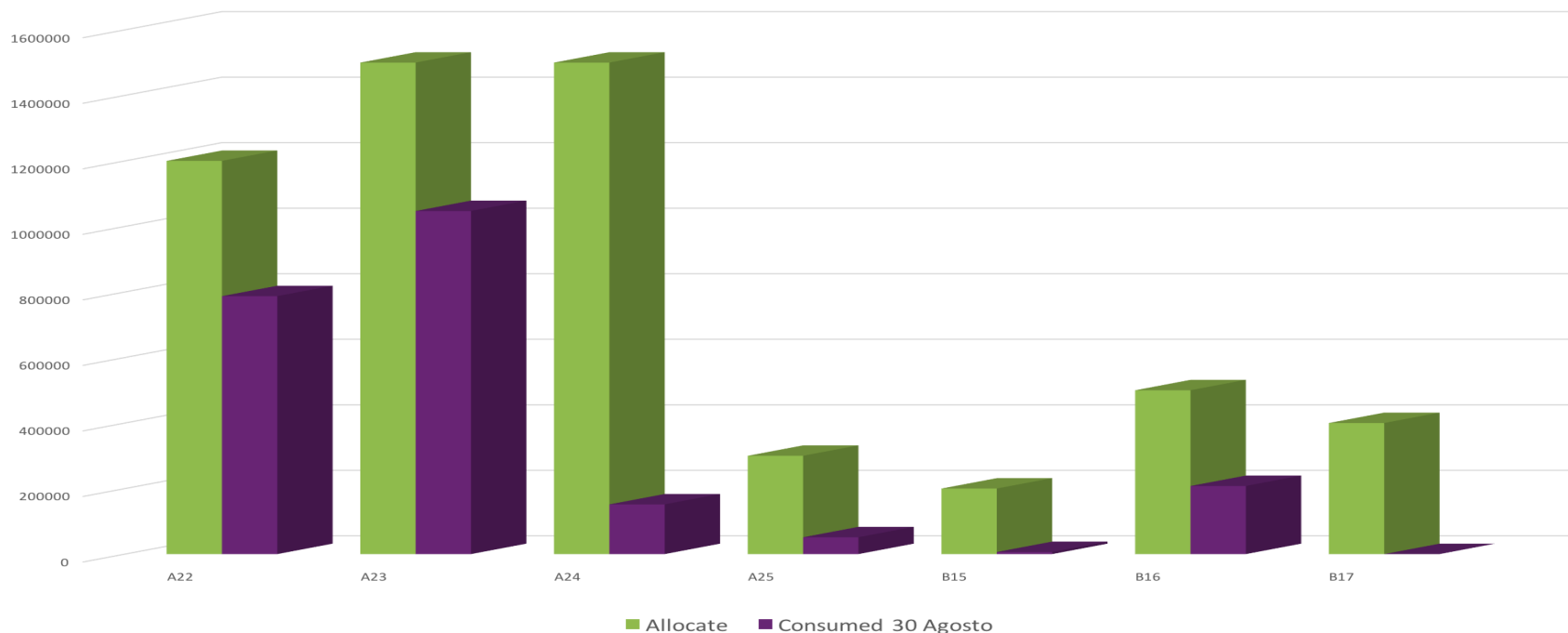
Call 3 – Maggio 2018: Approvati 9 Progetti

Call 3. Progetti ancora in corso o in fase di inizio: Risorse Assegnate

- N. 4 Progetti di Classe A → 18 Mhours assegnate
- N. 3 Progetti di Classe B → 1.8 Mhours assegnate
- N. 2 Progetti TEST → 1 Mhours riservate

**Totale 20.8 Mhours assegnate (in linea con le risorse disponibili).
Al 30 Agosto consumate 2.25 Mhours (34% dell'allocato)**

Call 3 Risorse Allocate e Consumate al 30 Agosto 2018



Come è andata?

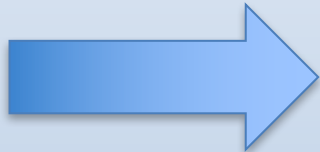
- Assistenza **Help desk** e supporto alla convenzione: **elevata e ottima**.
- **Tempi di risposta del comitato** alle richieste di assegnazione e contatti con gli utenti (massimo 15 gg... il più delle volte **1-2 giorni**)
- **Tempi di attesa** sulle code: **Marconi A1** si fa coda, **A2** mediamente poca attesa, **A3** quasi nessuna, pochissima coda.
- **L'uscita di produzione di Galileo e PICO** (old style) hanno prodotto la necessità di alcuni cambiamenti (non drammatici)
- **Abbiamo avuto qualche "stop!"**: Gennaio-Aprile... da PBS a Slurm e poi conseguente blocco e quindi sovraccarico di A1 e A2
- Spazio disco: **150 TB non sono tanti ma ci siamo rientrati**... non senza qualche taglio!
- **Monitoraggio continuo delle risorse...** (auto-responsabilità degli utenti: chiedere l'adeguamento delle risorse appena finita l'assegnazione temporanea)
- Se avessimo avuto **più risorse** le avremmo utilizzate.

Come è andata?

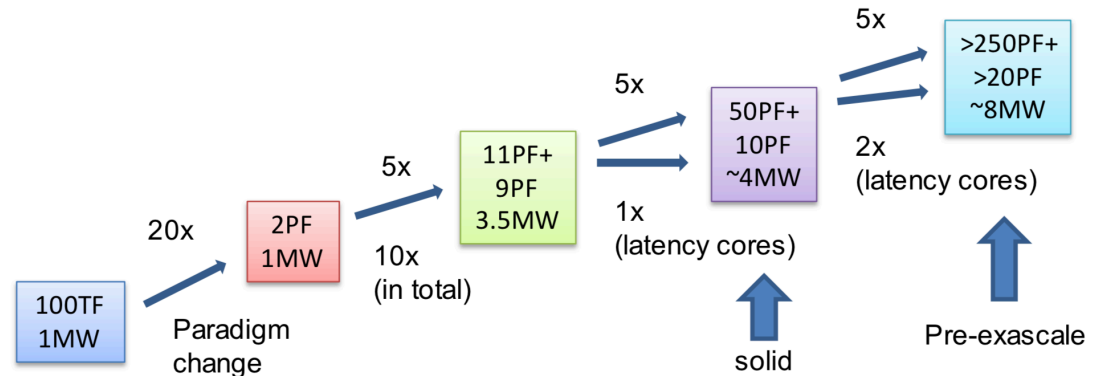
- Abbia fatto 3 Calls: abbiamo assegnato circa **90 milioni di ore.**
- Abbiamo attivato **43 progetti di cui 30 nuovi** di produzione
- Oltre **100 unità di personale di ricerca INAF** ed associati coinvolto
- Da una indagine preliminare svolta, **saranno circa 40 gli articoli che verranno prodotte con le ore già assegnate** (alcuni già inviati/pubblicati) e alcune tesi di PhD che utilizzeranno i risultati ottenuti.
- **Report sulle ricerche:** in corso la prima raccolta. Contributi di 4 pagine che stiamo già raccogliendo e che pubblicheremo a breve sul sito. Al termine dei 3 anni pensiamo di fare una pubblicazione di un volume dedicato.
- **Ricerca di personale... siamo in via di risoluzione:** una persona selezionata dal Cineca ci verrà dedicata. Siamo in fase di definizione di come fare il trasferimento fondi al Cineca.
- MoU INAF-GAIA: da non perdere di vista!
- Alcune interazioni con il progetto CHIPP sono state anche registrate (ma poche!)
- **Alto livello di soddisfazione generale**

Cosa ci aspetta!

Cosa succederà nell'immediato futuro.



Cineca “sustainable” roadmap toward exascale



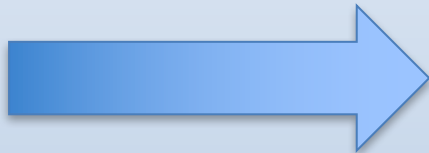
2009	2012/2013	2016/2017	2019/2020	2021/2022
IBM SP6 Power6	Fermi IBM BGQ PowerA2	Marconi Xeon + KNL	Marconi + PPI4HPC + ICEI (PPI-HB)	EuroHPC

PPI4HPC

In its communication on the [“European Cloud Initiative”](#) the European Commission in April 2016 stressed the need for **building an European Data Infrastructure including high-end supercomputers**, which will around 2022 reach the exascale performance level. In this proposal, a group of **leading European supercomputing centres** decided the formation of a **buyers group** to execute a joint Public Procurement of Innovative Solutions (PPI) for the first time in the area of high-performance computing (HPC). The co-funding by the European Commission (EC) will allow for a significant enhancement of the planned **pre-exascale HPC infrastructure from 2019**.

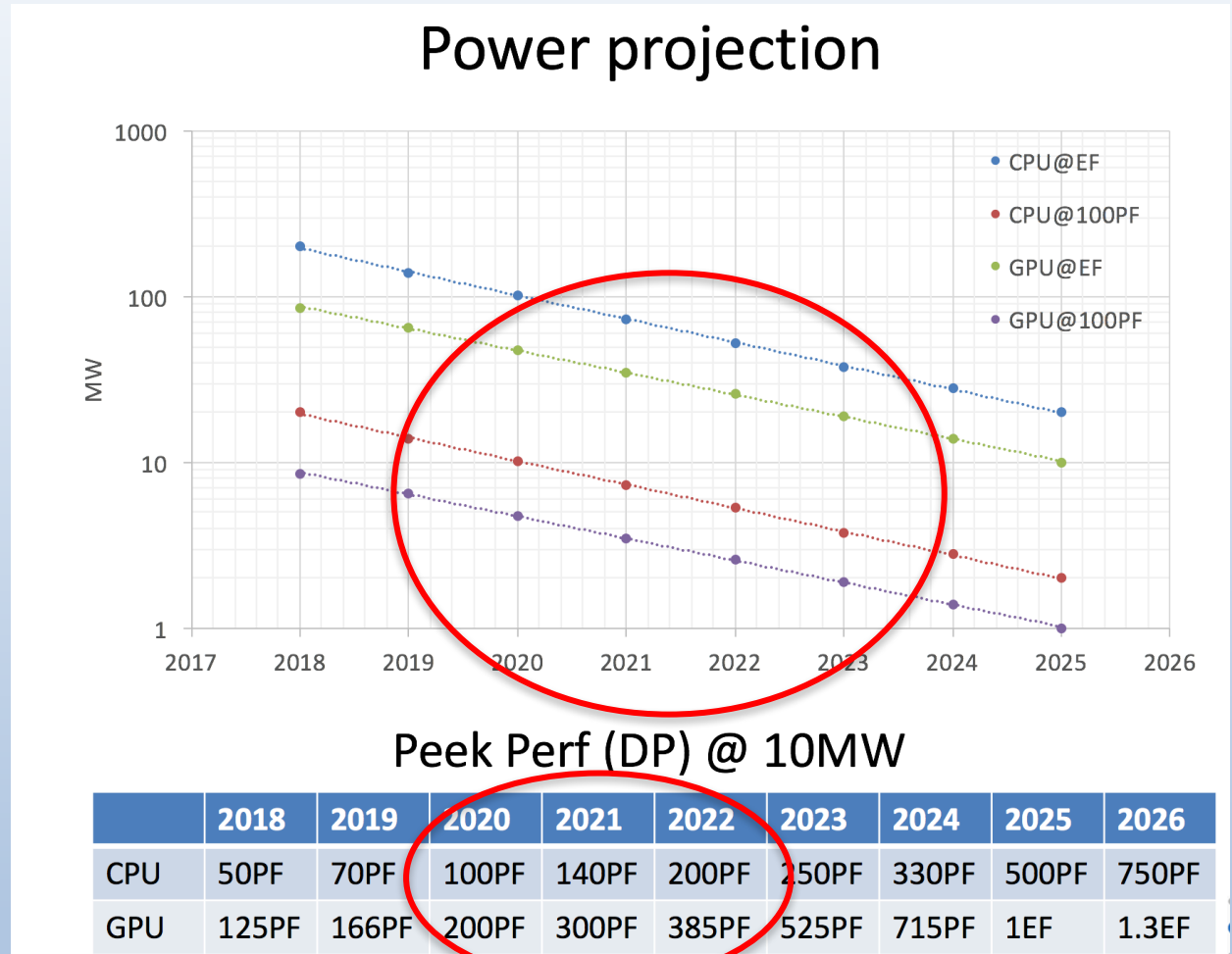
Cosa ci aspetta!

Cosa succederà nell'immediato futuro.



Dobbiamo prepararci a delle grandi novità nei modelli e paradigmi di progettazione e programmazione.

IMPATTO ANCHE FONDAMENTALE SUI GRANDI PROGETTI.

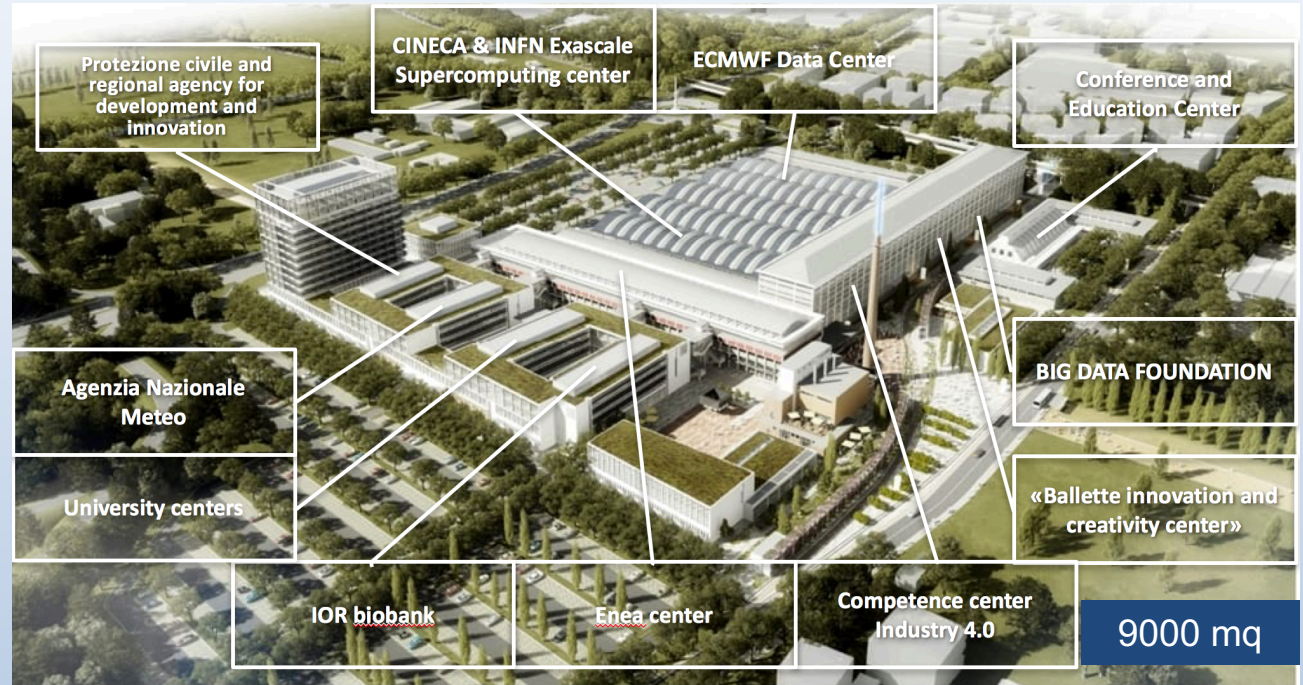


E' necessario cominciare subito a creare competenze e re-ingegnerizzare codici e pipeline che utilizzano o utilizzeranno sistemi HPC



MoU e prospettive: Verso l'Infinito e oltre

ECMWF – Bologna Science Park



La presenza di ECMWF (centro meteo europeo) fa diventare Bologna Science Park tra i più grandi poli di Supercalcolo e Big Data d'Europa. (ex Manifattura Tabacchi)

Cineca assume nel polo un ruolo più rilevante.

Fondamentale per INAF è essere un interlocutore sempre più attivo e principale

Abbiamo la consapevolezza di essere all'altezza delle sfide in gioco

Evitare di disperdere le eccellenze raggiunte delegando ad altri soggetti di rispondere alle nostre necessità attuali e future