

# **L'HPC c'è!**

## **News dal sondaggio sul calcolo HPC nell'INFN**

---

R.Alfieri, S.Arezzini, E.Calore, D.Cesini, G.Donvito, S.Lusso, S.Parlati

Tutorial Days CCR - 2024

# Perche' fare un sondaggio sull'HPC nell'INFN

L'idea di organizzare i Tutorial Days 2024 sui temi dell'HPC è nata in una riunione di CCR di giugno scorso

Nella riunione è emerso che

- Più sedi INFN ospitano al loro interno cluster HPC, gestiti dal servizio calcolo o dai gruppi di ricerca
- Non c'è consapevolezza di questa situazione
- Non c'è una rete di comunicazione tra le persone che si occupano di calcolo HPC
- Non ci sono stati eventi o corsi di formazione su questo tema negli ultimi anni
- L'INFN ha progetti per la realizzazione di una rete di calcolo HPC, le HPC Bubbles, ma non sono ancora stati programmati corsi di informazione e formazione su questo argomento

# Perche' fare un sondaggio sull'HPC nell'INFN

Alessandro ha proposto di dedicare i Tutorial Days a questo tema con lo scopo di

- Avere una formazione specifica su cosa è il calcolo HPC in tutte le sue forme
- Ricevere formazione sulla future HPC Bubbles
- Scambiarsi informazioni tecniche ed organizzative sulla gestione di cluster HPC
- Avere informazioni dai produttori di HW e SW HPC sulle tecnologie attuali e sulla roadmap futura delle principali tecnologie
- Dare consapevolezza che esiste una comunità di persone che lavorano all'HPC nell'INFN

Il gruppo incaricato di organizzare i Tutorial Days 2024 ha pensato di partire **raccogliendo** dalle sedi quante più informazioni possibile circa l'utilizzo e l'interesse verso il calcolo HPC e far **emergere** le necessità di formazione in questo campo.

# Sondaggio: a chi si è rivolto e chi ha risposto

Per raccogliere il maggior numero di informazioni possibile, il sondaggio e' stato proposto a

- Responsabili dei servizi calcolo
- Rappresentanti in CCR
- Responsabili dei Tier1/Tier2
- Rappresentanti dei centri di calcolo ICSC
- Membri del C3SN

Chiedendo uno sforzo collaborativo per dare una risposta unica per sede.

**Al sondaggio hanno risposto 27 sedi!**

# Contenuti del sondaggio

---

Overview del calcolo HPC nelle sedi

---

Calcolo HPC su risorse esterne

---

Calcolo HPC su risorse locali

---

Uno sguardo al futuro

---

Necessita' di formazione durante i Tutorial days

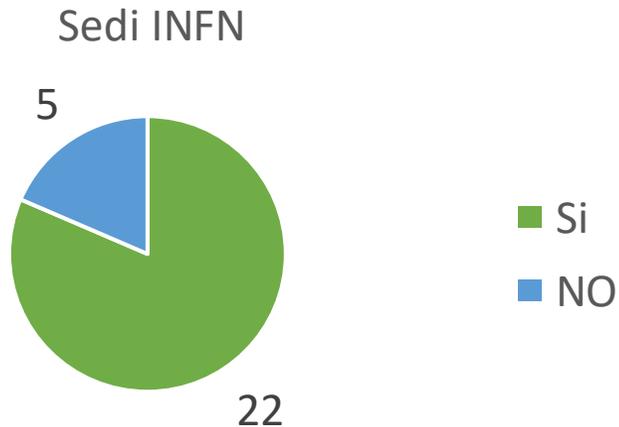
---

Proposte di formazione e suggerimenti

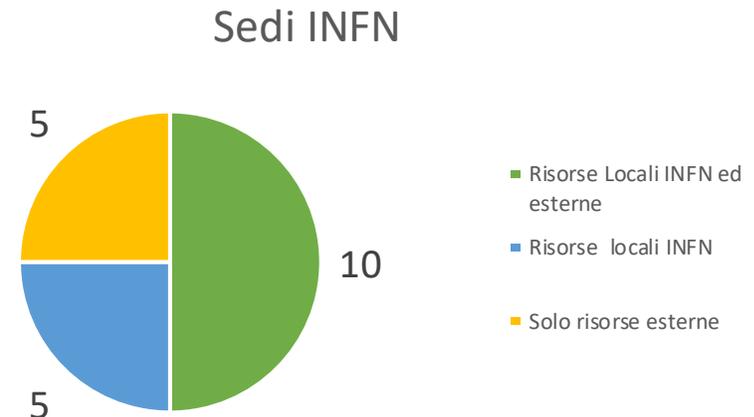
# Overview del calcolo HPC nelle sedi

Per risorsa HPC intendiamo nodi connessi con rete ad alta velocità e bassa latenza con eventuale filesystem parallelo e/o con acceleratori computazionali (GPU, FPGA..)

Gli utenti della vostra sede utilizzano il calcolo HPC?

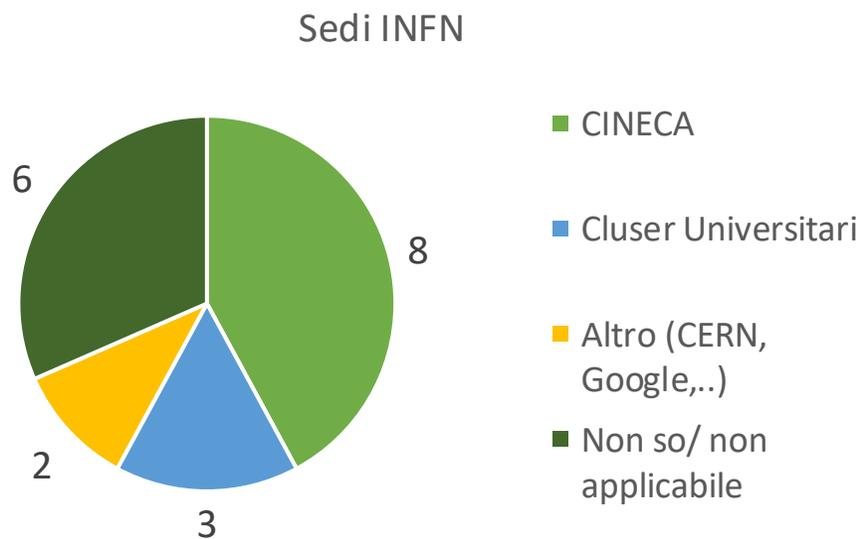


Se sì potete indicare se utilizzano risorse locali e/o risorse esterne?

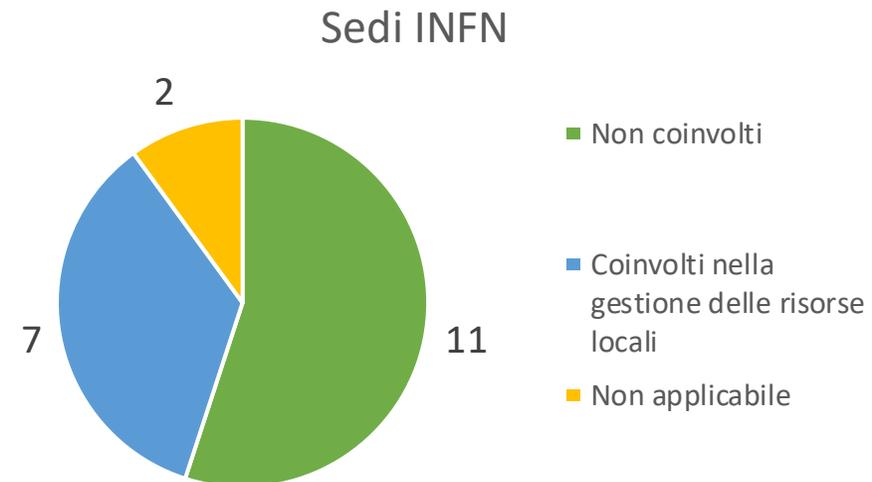


# Overview del calcolo HPC nelle sedi: risorse HPC esterne

Se usano risorse esterne indicate quali :



Siete coinvolti nella gestione delle risorse (esterne)? Come?

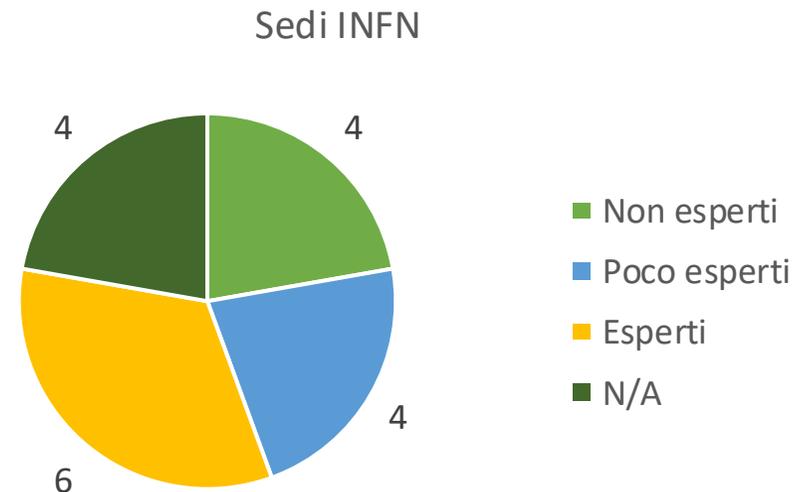


# Overview del calcolo HPC nelle sedi: risorse HPC esterne

Conoscete le applicazioni dei vostri utenti? Potete indicare le principali? Sono sviluppate in autonomia o sviluppate da community o sono applicazioni commerciali?

- 7 sedi non conoscono in dettaglio le applicazioni dei proprio utenti su risorse esterne HPC
- Le altre indicano:
  - Principalmente sviluppate localmente (CUDA), in autonomia, (Machine Learning)"
  - "Le applicazioni principali riguardano il machine learning
  - "Monte Carlo o altro software sviluppato dalla community
  - "Fisica teorica (Cosmologia, fisica gravitazionale, Fisica Statistica, LatticeQCD); Einstein toolkit (da community) + programmi sviluppati in autonomia"
  - "Il 90% dei casi d'uso riguardano uso di software scientifici di terze parti."
  - "Applicazioni principalmente open-source per calcoli teorici (qcd su reticolo, ed altro). Alcuni casi di utilizzo di software compilato con Matlab."
  - "Non conosco i prodotti in dettaglio. Per l'HPC sono applicazioni open source per simulazioni meteorologiche o in ambito biofisico."

Quanto vi sentite esperti nella gestione o nell'utilizzo di queste risorse ?

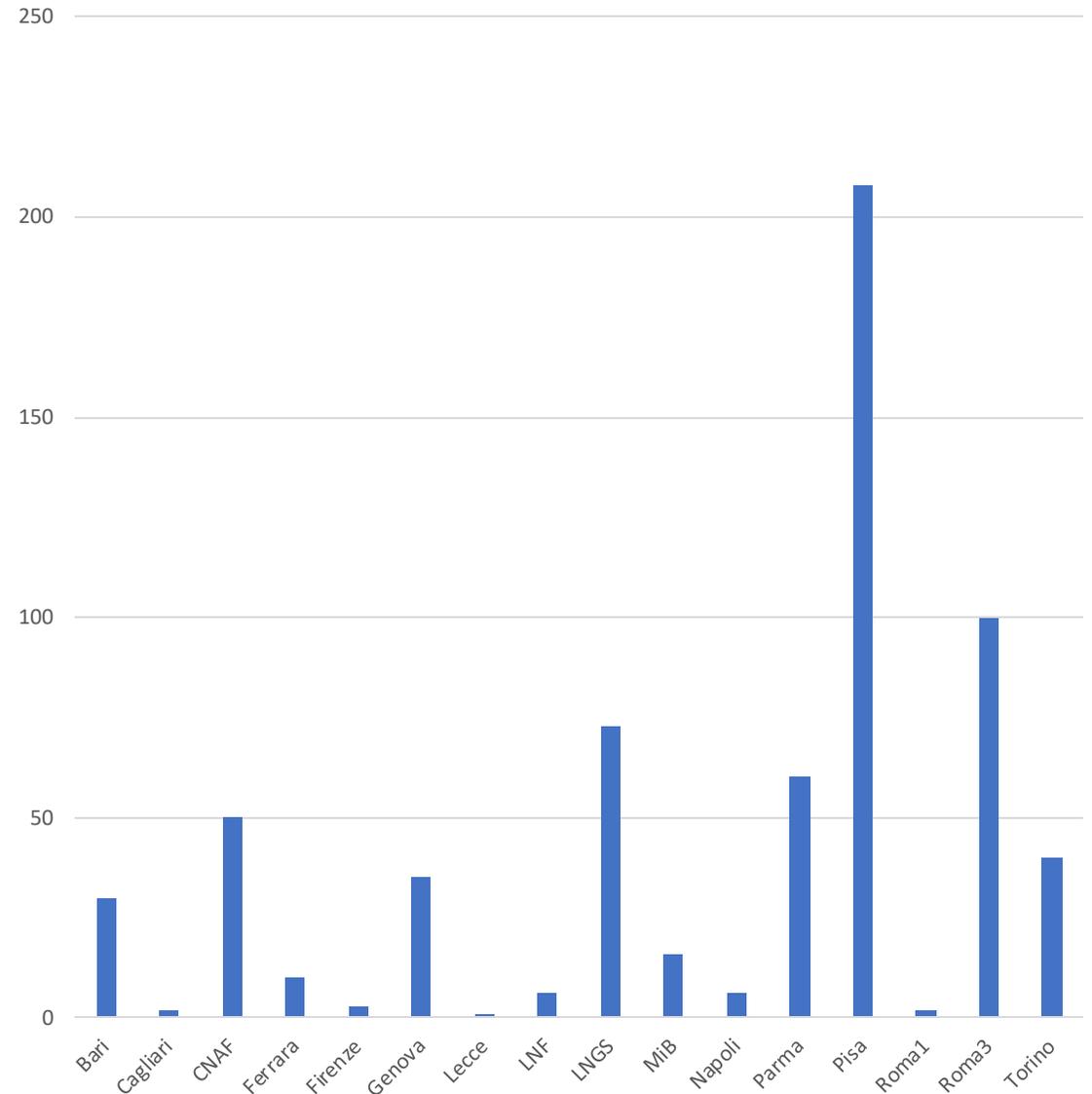


# Calcolo HPC su risorse locali

Se nella vostra sede si utilizzano risorse HPC locali gestite direttamente (dal servizio o da altro personale) potete fornire una descrizione dettagliata?

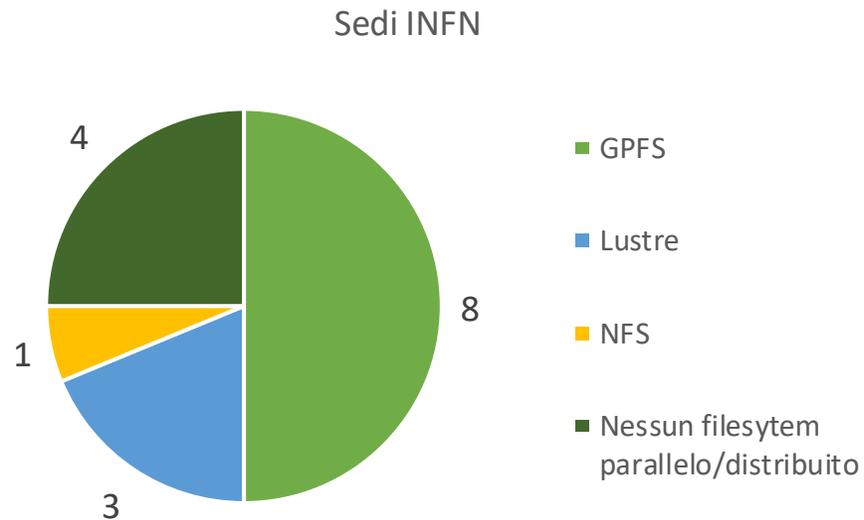
## Numero di nodi

In generale le risorse sono disomogenee tra le sedi, acquisite o acquistate in momenti diversi.

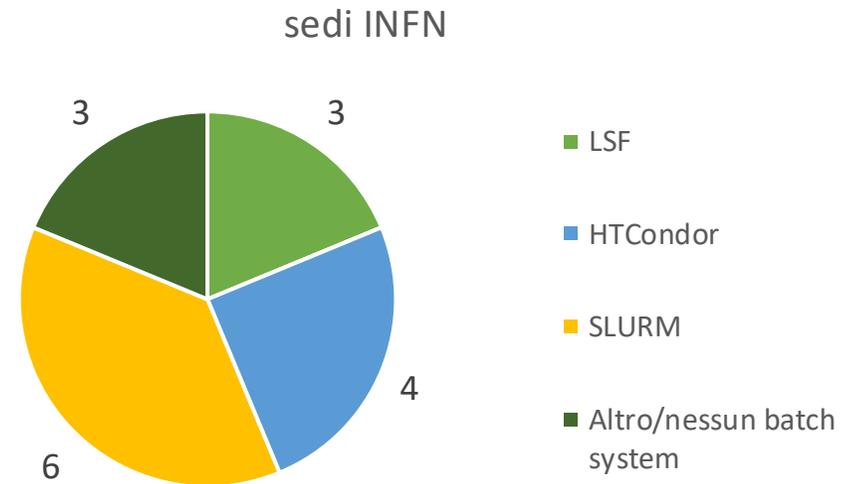


# Calcolo HPC su risorse locali

## Tipologia di filesystem

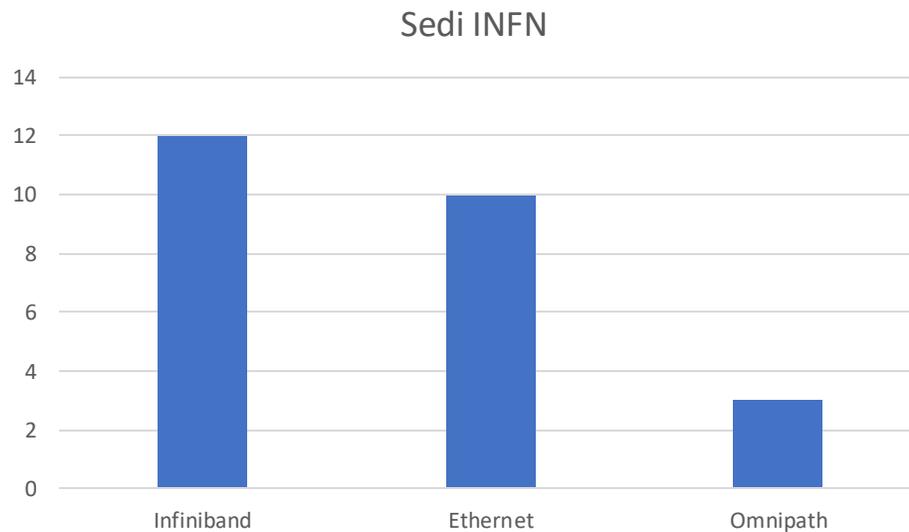


## Tipologia di Batch system



# Calcolo HPC su risorse locali

## Tipologia di rete di interconnessione



## Modalità di accesso alle risorse

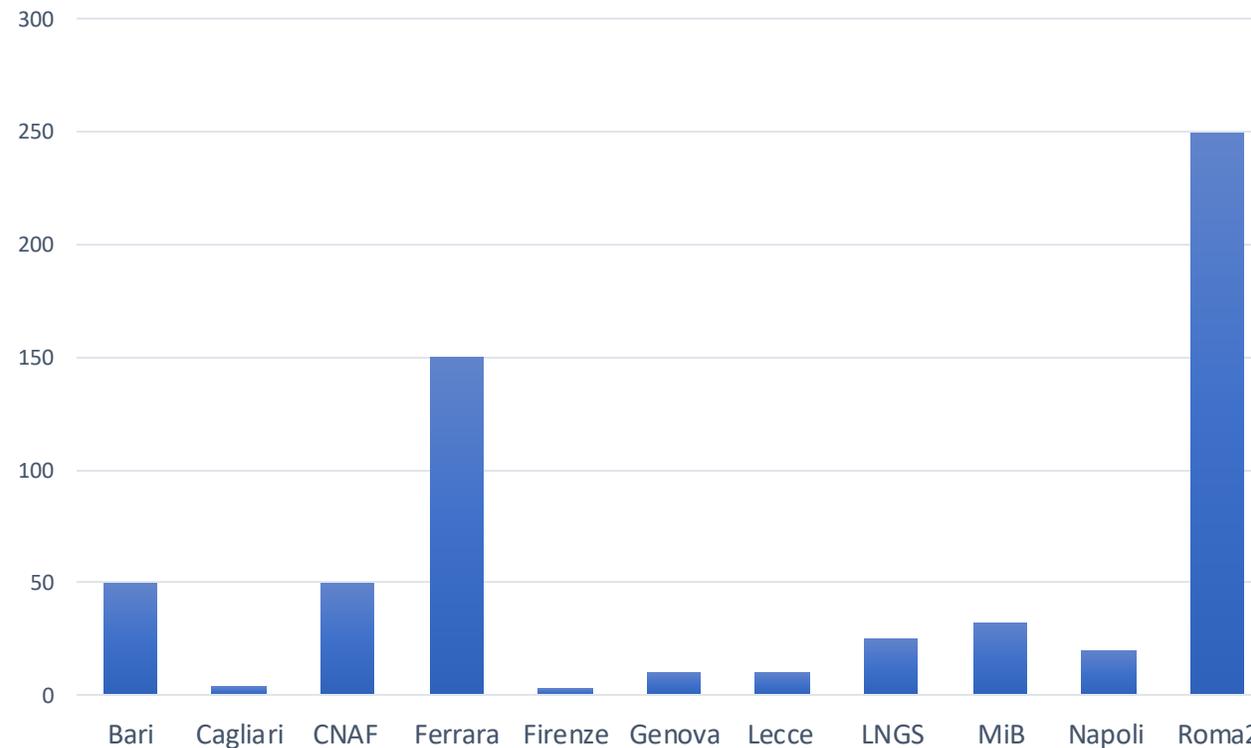
quasi tutte le sedi indicano accesso ssh su macchine di login dedicate

11 intervistati (65%) hanno risposto ssh a questa domanda.

accessi macchine UI Sottomissione batch richiesta  
macchina front end user interface cloud locale  
cloud nazionale accesso **ssh** login ex Galileo Sapienza  
orchestratore docker chiave progetti speciali  
modalità guest Gestione comitato tecnico/scientifico gruppi

# Calcolo HPC su risorse locali

## Numero di utenti



Roma1, Parma, LNF, Pisa e Torino servono grandi comunità : il numero di utenti e gruppi non è stato quantificato nel sondaggio.

# Calcolo HPC su risorse locali

---

## Eventuali benchmark o tool utilizzati per la verifica dell'efficienza di un codice:

**CNAF:** per CPU: HPL, NAMD apoa1/stmv. Per rete IB: microbenchmark installati con i driver

**Roma3:** script sviluppati in sede per la valutazione "grossolana" dell'efficienza.

**Ferrara:** per CPU/GPU: benchmark sintetici, SPEChpc, mini-apps, profilers e modello Roofline; per rete: OSU benchmarks..

**Napoli:** Python Profilers, Valgrind, Perf, Gprof, tools autoprodotti

**Pisa:** Confronto utilizzando casi standard nel caso di software commerciali

# Calcolo HPC su risorse locali

## Anno della prima installazione di risorse HPC in sede

2000 Pisa

2010 MiB

2014 Bari, CNAF

2015-2017 Ferrara, Torino,  
Roma3, Parma

2019 Genova, Lecce

2022 Napoli, LNGS, Roma1

2023 Cagliari

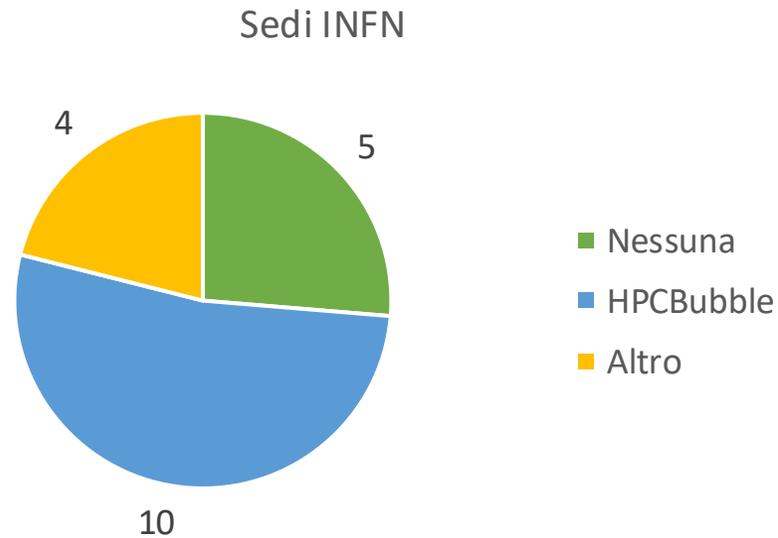
## Quanto siete o vi sentite esperti nella gestione di queste risorse ?

Alcune sedi hanno una buona esperienza, la maggior parte esperienza solo sulla gestione sistemistica del cluster, non sulle applicazioni

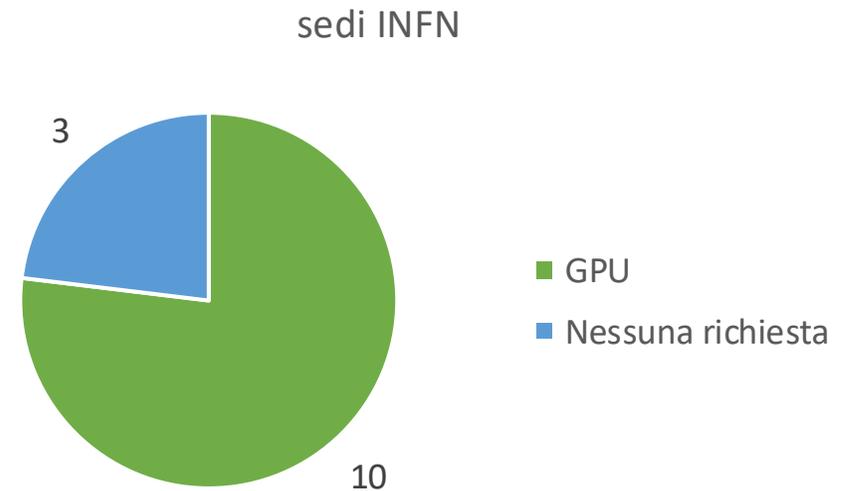


# Uno sguardo al futuro

Se nella vostra sede sono in arrivo risorse HPC potete indicare quali?  
(es. Terabit HPC Bubble, altro?)



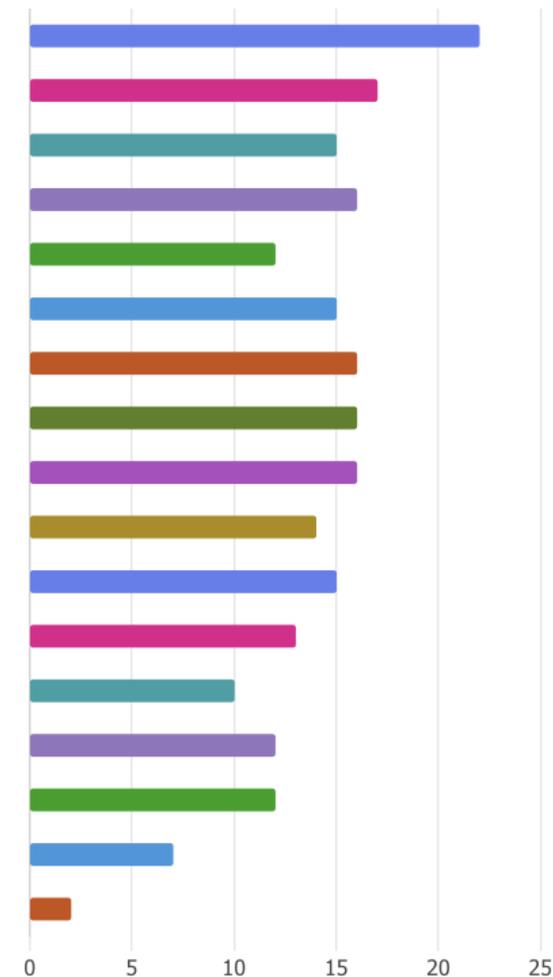
Se gli utenti della vostra sede non usano risorse HPC: c'è comunque richiesta, interesse verso questo tipo di calcolo?



# Necessita' di formazione durante i Tutorial Days 2024

## 24. Quali argomenti vorreste venissero trattati durante i Tutorial Days?

● Panoramica su calcolo parallelo con MPI, calcolo su GPU e su FPGA	22
● Interconnessioni veloci e a bassa latenza per calcolo HPC	17
● Infiniband: come funziona, come si configura e come si gestisce	15
● Storage per cluster HPC e filesystem paralleli (GPFS, Lustre,..)	16
● GPU: Nvidia CUDAcore e SensorCore; interconnessione con NVLink; partizionamento...	12
● Panoramica sul software per cluster HPC: librerie MPI, OpenMP, CUDA e alternative a CUDA;...	15
● Batch system per cluster HPC; batch system e GPU	16
● Uso di container su cluster HPC (es. singularity/apptainer, altro..)	16
● HPC Bubble: descrizione del progetto e del modello di calcolo; tipo di storage e di accesso...	16
● HPC Bubble: configurazione dei server con GPU e con FPGA; configurazione dei nodi di storage...	14
● HPC Bubble e federazione: interfacce di accesso comuni ma gestione differenziata nelle varie se...	15
● Accounting di risorse HPC e come fare accounting dell'uso di GPU	13
● HPC e AI (ottimizzazione dei sistemi per applicazioni AI)	10
● Applicazioni utente: come ottimizzare i sistemi, esperienza utenti, strumenti per monitorare...	12
● Informazioni generali sulla gestione di risorse HPC (sistemi di installazione e gestione...	12
● Futuro dell'HPC, tendenze per le applicazioni e l'HW del futuro; futuri data center per l'HPC	7
● Altro	2



# Proposte di formazione e suggerimenti

---

**Siete interessati a presentare qualche argomento o attività che conoscete bene?**

**CNAF:** gestione slurm sulla base dell'esperienza fatta sugli attuali cluster

**Bari:** Batch system per cluster HPC e GPU 2. Uso di container su cluster HPC.

**Ferrara:** Seminario sull'esperienza raccolta in 10 anni di amministrazione del cluster HPC di Ferrara,

**LNGS:** Overview di Lustre opensource, esempio di configurazione

**Pisa:** 1) HPC evoluzione dei sistemi di virtualizzazione e ambiente utente.

2) Troubleshooting in ambiente HPC.

3) Il futuro dell'HPC fra il modello Grid e il problema della latenza.

## **Suggerimenti:**

**LNF:** Tutti gli argomenti di cui al punto 24 sono estremamente interessanti, ma ci si augura che vengano approfonditi in appositi corsi future

**Padova:** Per i tutorial days siamo interessati soprattutto agli aspetti di installazione e gestione dell'infrastruttura, piu' che al suo utilizzo.

**Torino:** Breve descrizione facility HPC OCCAM: Il funzionamento della facility HPC occam si basa sui docker container.

# Comitato organizzatore del sondaggio e dei Tutorial Days

- Roberto Alfieri - Parma
- Silvia Arezzini - Pisa
- Enrico Calore - Ferrara
- Daniele Cesini - CNAF
- Giacinto Donvito - Bari
- Stefano Lusso - Torino
- Sandra Parlati - LNGS

