

Report 3 Novembre 2022

# Outline

- Stato gare e tempistica
- Stato accordi e tavoli CINECA
- Open dallo scorso meeting:
  - Situazione/test Skyway
  - Manutenzione e Commissioning
- News da PNRR /progetti INFN / preventivi 2023
- Project management news

# Stato gare e tempistica

- Gare Cablaggio e Rete
  - Aperte, sono arrivate le prime offerte e richieste di chiarimenti. Termine presentazione offerte: 21 novembre
- Acquisto storage 2021-2022 (tecnicamente non CNAF reloaded)
  - Prezzo di aggiudicazione: **95.86 €/TB-N**
  - Nostre stime nel TDR per il 2022: **96 €/TB-N**
  - **Il problema sembra rientrato oltre che per CPU (visto a settembre) anche per lo storage!**
- AQ storage 2023-24
  - ~~Approvato in GE 18 Ottobre 2022~~
  - Delibera annullata per permetterne uso per PNRR → GE 11/11/22

**Model vs reality ...**  
**“so far so good...”**

| Year              |  | 2021           | 2022  | 2023 | 2024  | 2025   | 2026    |
|-------------------|--|----------------|-------|------|-------|--------|---------|
| Disk (Eur/TB-Net) |  | 120            | 96    | 77   | 61    | 49     | 39      |
| Disk (Eur/TB-Net) |  | no procurement | 95.85 |      |       |        |         |
| CPU (Eur/HS06)    |  | 8              | 6.4   | 5.12 | 4.096 | 3.2768 | 2.62144 |
| CPU (Eur/HS06)    |  | no procurement | 6.82  |      |       |        |         |

# Variante: costo

- Alcune variazioni lato nostro (elettrizzazione porte, piccoli lavori edili aggiunti)
- Abbiamo numero finale per il costo (bollinato da CINECA, RUP e Direttore dei lavori) – IVA da aggiungere
  - In linea con i numeri forniti al DG a luglio

|                     |  | <b>Rev04 - finale</b> |
|---------------------|--|-----------------------|
| STACCHI TUBAZIONI   |  | 44,929.13 €           |
| EDILE               |  | 283,408.70 €          |
| ASSISTENZA IMPIANTI |  | 49,562.83 €           |
| TOTALE ELETTRICO    |  | 1,426,708.87 €        |
| TOTALE MECCANICO    |  | 1,150,558.35 €        |
| TOTALE RACK         |  | 1,110,019.47 €        |
| SICUREZZA IMPIANTI  |  | 95,106.63 €           |
| <b>TOTALE</b>       |  | <b>4,160,293.98 €</b> |

# Variante: tempistica

- Avevamo detto 30/6 data di consegna Botte B5 (compresi rack)
- Avevamo ventilato la possibilità che questo si anticipasse a Fine Marzo, in linea con la data di consegna prevista pre-variante
- **Questa ipotesi è rientrata:** situazione fra CINECA (appaltatore) e ICM (lavori edili) decisamente tesa – non per colpa della nostra variante
  - ICM insiste per avere più tempo → 30/06 data confermata con RUP

# Stato Accordi

- Accordo utilizzo risorse FINALIZZATO
  - La parte tecnica lo era da tempo, erano rimasti dettagli da definire fra i due uffici legali CINECA / INFN
  - Risolto 17/10 → in approvazione al CD di ottobre novembre
- Tavolo Tecnico Attuativo
  - Prossimo passo essenziale è costituire gruppo di lavoro per seguire le fasi operative
  - Riunione comitato di coordinamento 11/10: assegnato a tale tavolo il compito di seguire
    - Capitolati manutenzione
    - Modalità di commissioning

## Art.3

### (Tavolo tecnico attuativo)

Le Parti concordano di istituire, altresì, un Tavolo tecnico attuativo, di seguito indicato come “Tavolo tecnico”, avente facoltà attuativa in merito alle decisioni assunte dalle Parti su proposta del Comitato, in virtù del presente Accordo.

Il Tavolo tecnico dovrà, pertanto, tenere costantemente informato il Comitato e le Parti in relazione ad eventuali problematiche che dovessero insorgere nello svolgimento delle attività previste e a possibili azioni correttive da assumere.

Il Tavolo tecnico è composto da sei membri, tre per ciascuna Parte, due con competenze tecnico-scientifiche e una con competenze amministrativo-finanziarie, e da un Presidente, che dura in carica 3 anni senza possibilità di svolgere mandati consecutivi, nominato dai suoi componenti nel rispetto del principio dell'alternanza rappresentativa tra le Parti e dell'alternanza con la presidenza del Comitato.

Abbiamo proposto al management di nominare lato INFN:

- Luigi Scarponi (resp. Infrastruttura)
- Daniele Cesini (resp. Tier-1)
- Martina Allegro (resp. amministrativo INFN-CNAF/BO)

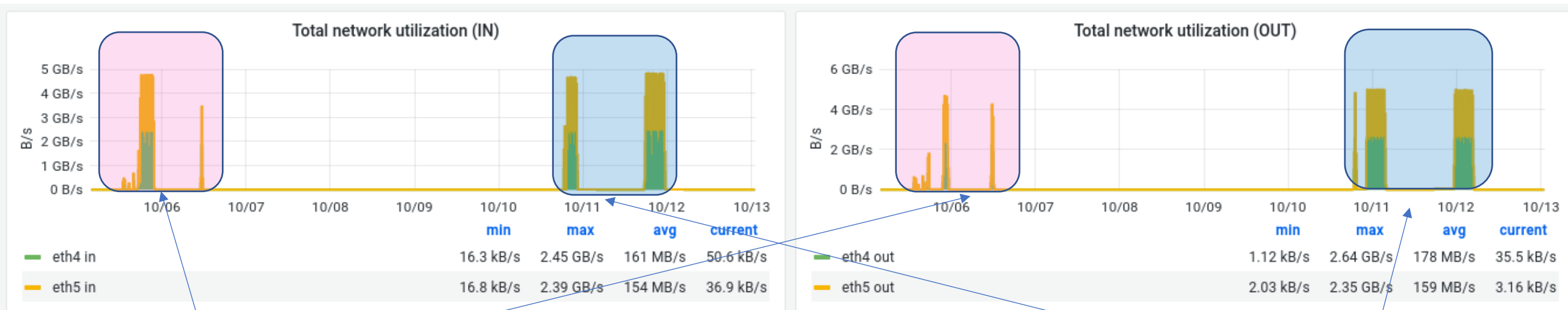
Il comitato è comunque già operativo, in contatto con CINECA

# Skyway test

- Allo scorso meeting: anche in assenza di Skyway testabili, simulare comportamento su scala significativa
  - Reminder: la perplessità maggiore è dovuta non tanto agli Skyway *per se*, ma al loro limite di MTU a 4000 Byte, diverso dal valore di 9000 usato nel data center
  - Il dubbio era soprattutto su tenuta/performance di GPFS in queste condizioni
- Test di scala effettuati
  - Con test specifici GPFS per testare stabilità del sistema. Fino a 40 nodi, 2400 jobs
  - Con jobs di produzione CMS (stessi 2400 job slots)

# Test GPFS (lato server)

- Script specifico, server GPFS con MTU=9000 e 40 wn con dapprima MTU=9000 (per definire baseline) e poi MTU=4000
  - “MTU negotiate = on” su tutti i sistemi
- Switch intermedio limitato a 60 Gbit/s



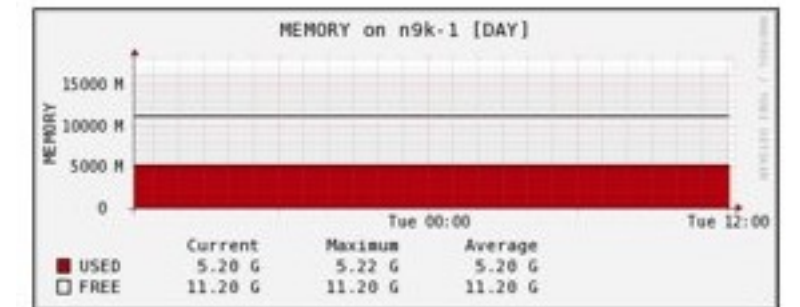
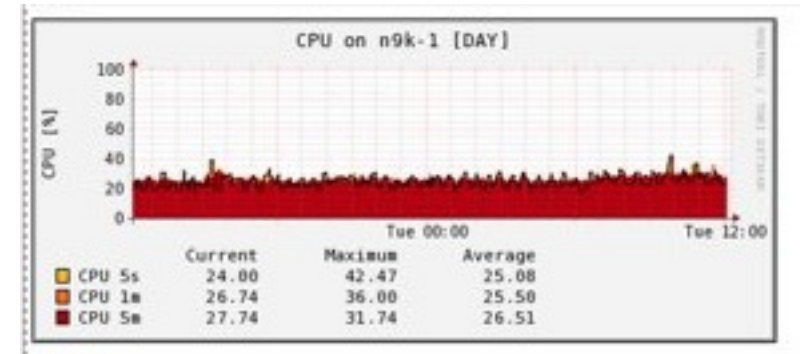
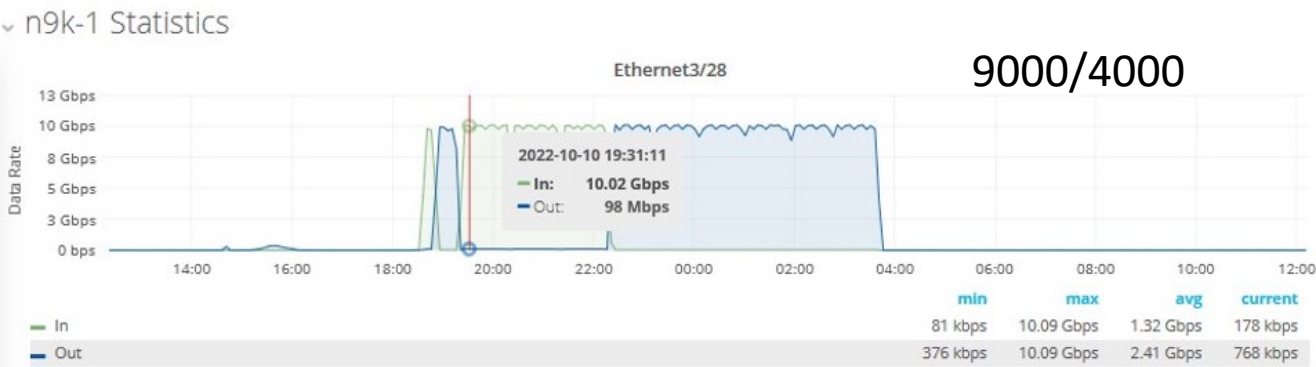
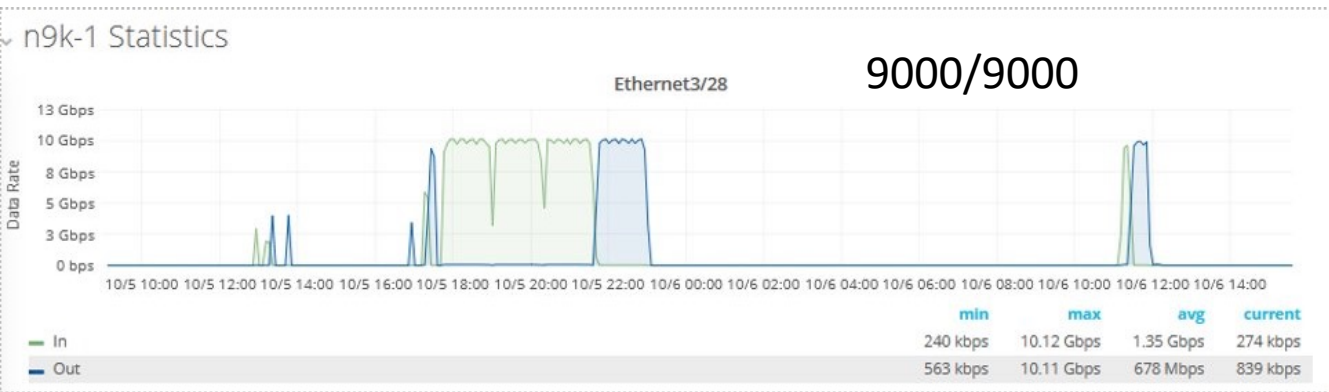
9000/9000: baseline

**STESSE PERFORMANCE,  
NESSUN ERRORE NEI  
LOG**

9000/4000: Skyway-like setup

# Test GPFS (lato rete)

- Visione lato switch intermedio (2x in load balancing)



**STESSE PERFORMANCE, NESSUN ERRORE NEI LOG** CPU e memoria dello switch flat durante il periodo

# Conclusione test GPFS

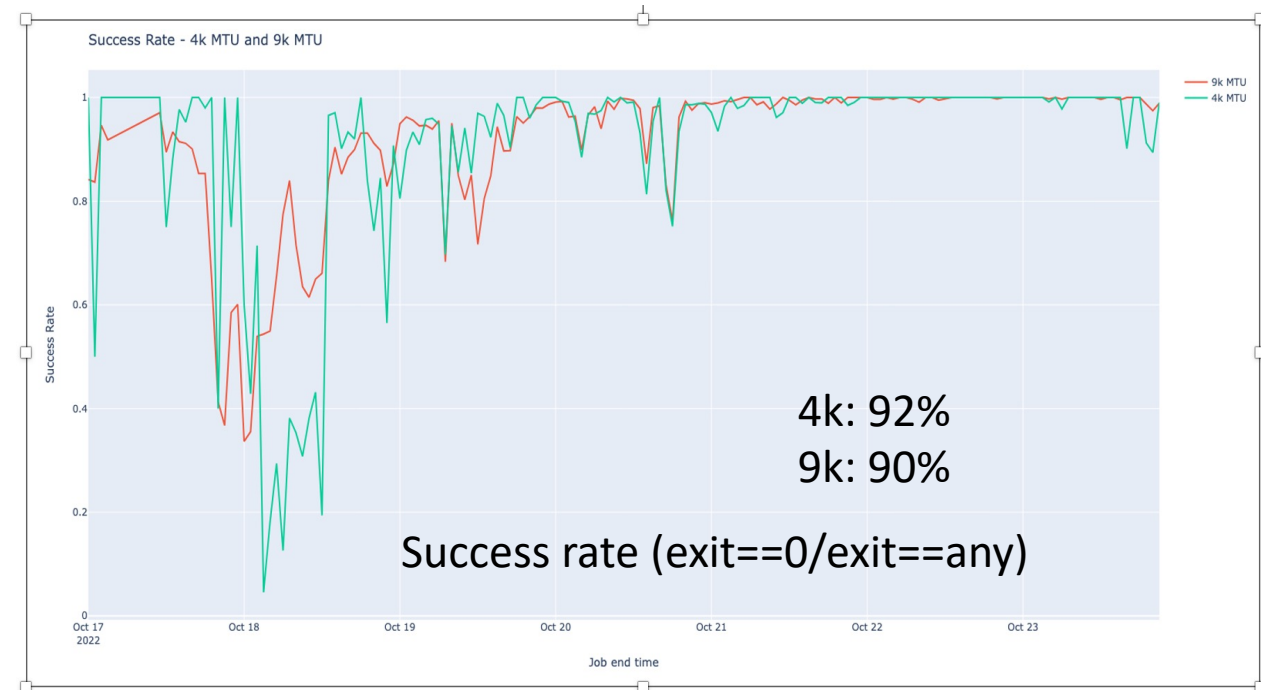
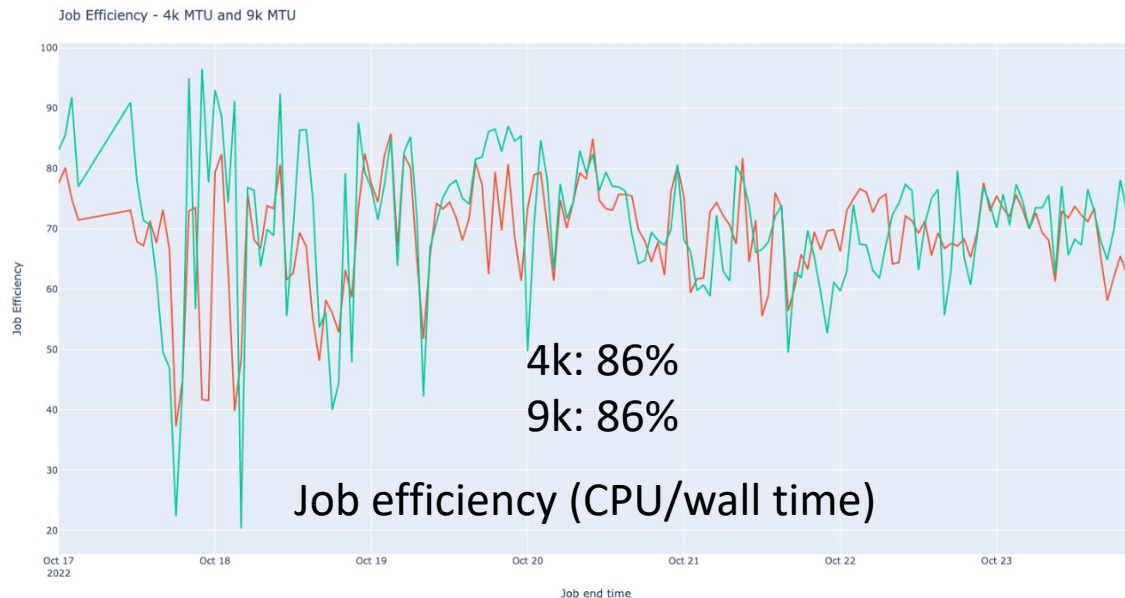
- GPFS non ha problemi a gestire MTU diversi connessione per connessione, configurate in autodiscovery
- Non abbiamo evidenza di calo di prestazioni o di aumento di uso di risorse di nessun tipo
- **Next step:** test di produzione con job veri (CMS)
  - Riconfigurato un rack completo (2400 CPU cores) a MTU = 4000
  - Reso disponibile solo a job di CMS
  - Confrontati i job CMS che girano su quel rack rispetto a quelli che girano su altri rack standard CNAF. Metriche di interesse:
    - Success rate (per mostrare eventuali problemi dovuti a servizi che non gradiscono la configurazione di rete)
    - CPU efficiency (per evidenziare se il nuovo setting peggiora le performance dei job)
- →

# Monitoraggio jobs CMS

- Numero di job CMS nel periodo
  - 70k su WN MTU=9000
  - 11k su WN MTU=4000

- Non si osservano differenze apprezzabili tra i job di CMS, ne' per success rate ne' per efficiency
- Fluttuazioni trascurabili probabilmente dovute alla bassa statistica (una settimana di test)

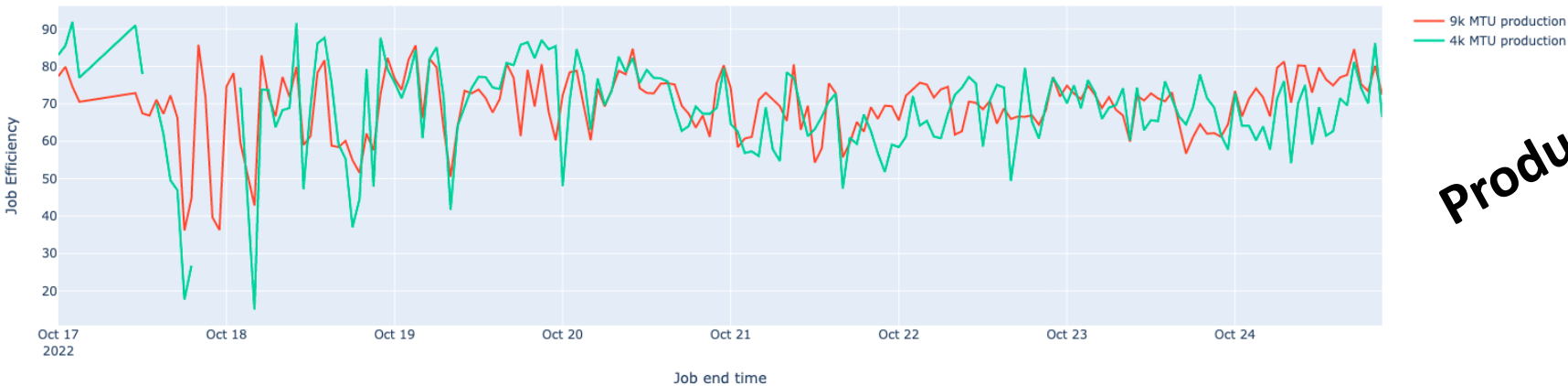
## Confronto job efficiency



# Analysis vs Production jobs

- Confronto interessante perché usano pattern completamente diversi di accesso allo storage
- ... di nuovo nessuna differenza (rosso vs verde)

Job Efficiency (PRODUCTION ALL) - 4k MTU and 9k MTU



Production

Job Efficiency (ANALYSIS) - 4k MTU and 9k MTU



Analysis

# Conclusioni lato “Skyway”

- La nostra preoccupazione principale (instabilità / cattive performance di CMS con MTU più basso del resto del cluster) è risolta
  - Non usando uno Skyway, non ancora consegnato, ma un settaggio funzionalmente equivalente
- A essere pignoli, saremmo più tranquilli se avessimo in mano uno Skyway per vedere se questo si comportasse in modo “non documentato”
  - Richiesta reiterata su base circa settimanale, ma se non c'è non c'è ....

**IL TEST EFFETTUATO (PUR ESSENDO UNA SIMULAZIONE) HA DATO  
ESITO SODDISFACENTE**

# News da PNRR/referaggio/fondi etc

- PNRR:

- ICSC + TeRABIT fondi 7.4 M€ per risorse al CNAF
- Piano di acquisto 3<sup>a</sup> tape library, dopo scrutiny dei referees (fondi ICSC)
- Stima al rialzo delle risorse da installare
  - 140 €/TB-N e 15 €/HS06 non più realistici (riduzione dei prezzi del 30-55%)
- Nuova stima comporta carico elettrico di ~260 kW
  - ~ 18 PB + Tape library + 500 kHS06

- Fondi

- Fondi Tecnopolo per variante (infrastruttura + rack) e rete (apparati+cablaggio)
- Da incasellare tutte le altre spese (i.e., pledge) per tenere conto del PNRR
- Cominciata anche discussione sui fondi per la fase 2
  - Costi delle soluzioni elettriche (tentativo di piano B: alta tensione direttamente nel Tecnopolo per CINECA e INFN, fino a 20 MW entro fine 2024 senza ridondanza e 2026 con ridondanza)
  - Upgrade dell'infrastruttura (cooling, etc..)

## PROJECT INDICATORS - EARNED VALUE MANAGEMENT

| CPU            |  | RACKS          |  |
|----------------|--|----------------|--|
| PV             |  | PV             |  |
| AC             |  | AC             |  |
| % Avanzamento  |  | % Avanzamento  |  |
| EV             |  | EV             |  |
| SPI            |  | SPI            |  |
| CPI            |  | CPI            |  |
| General Status |  | General Status |  |

| DISKS          |  | NETWORK        |  |
|----------------|--|----------------|--|
| PV             |  | PV             |  |
| AC             |  | AC             |  |
| % Avanzamento  |  | % Avanzamento  |  |
| EV             |  | EV             |  |
| SPI            |  | SPI            |  |
| CPI            |  | CPI            |  |
| General Status |  | General Status |  |

| TAPES          |  |
|----------------|--|
| PV             |  |
| AC             |  |
| % Avanzamento  |  |
| EV             |  |
| SPI            |  |
| CPI            |  |
| General Status |  |

## PROJECT INDICATORS - MILESTONES ADVANCEMENT

### MILESTONES ADVANCEMENT

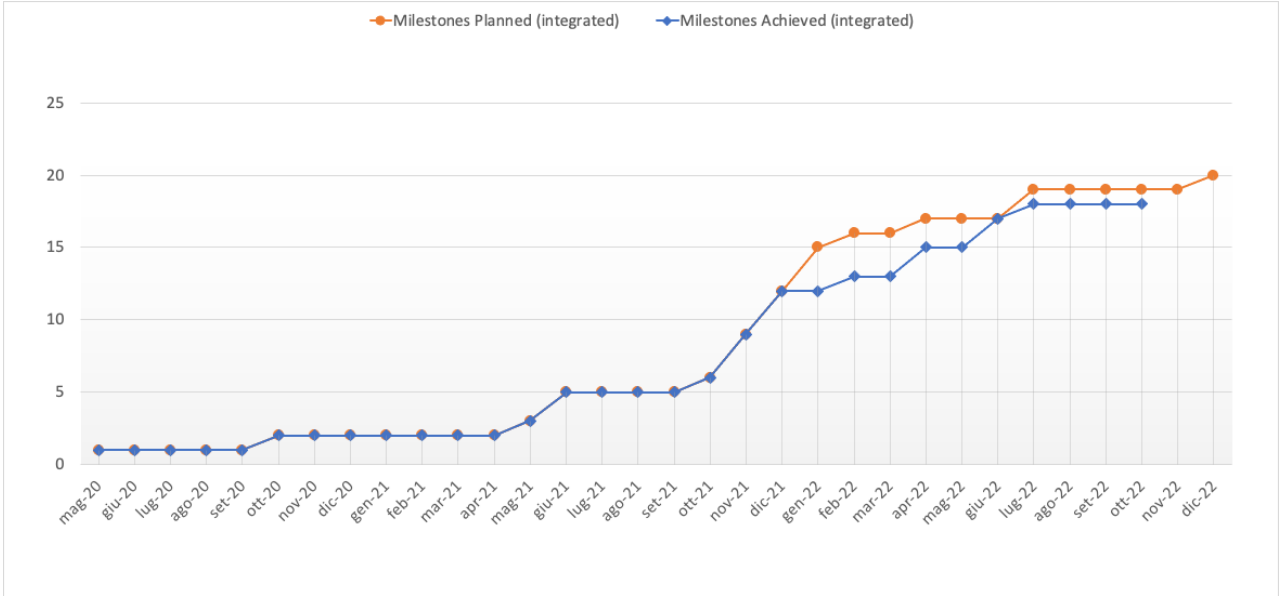
- #Milestones Planned = 19  
#Milestones Achieved = 18
- 2 of them partially achieved (50% each):
    - WP2-M12 - monitoring system
    - WP2-M13 - provisioning system

Milestone accomplishment:

MPI=94.7%

### Milestones update:

- Some milestones have been reshuffled and some of them will require additional tests to be executed in March 2023 (additional milestones).



# Prossimi passi (prima di Natale)

- Prox settimana: meeting con DG (e poi spero con GE) per quadro fondi (ex PNRR, Tecnopolo, FOE)
- 13/12: meeting con focus su WP 2.1.5 (EPIC) per migliore integrazione nel contesto Tecnopolo/CNAF
- TBD: meeting con focus su WP 2.2 (AAI) e 2.3 (IT security)
- TBD: review completa dei task alla luce di quanto deciso a Rimini ed elaborato durante il mese di ottobre
  - Possibile amendement al TDR con inserimento nuove milestone/deliverable

# Foto recenti

- [2022-10-26](#)