

Prospettive calcolo CSN2

Michele Punturo

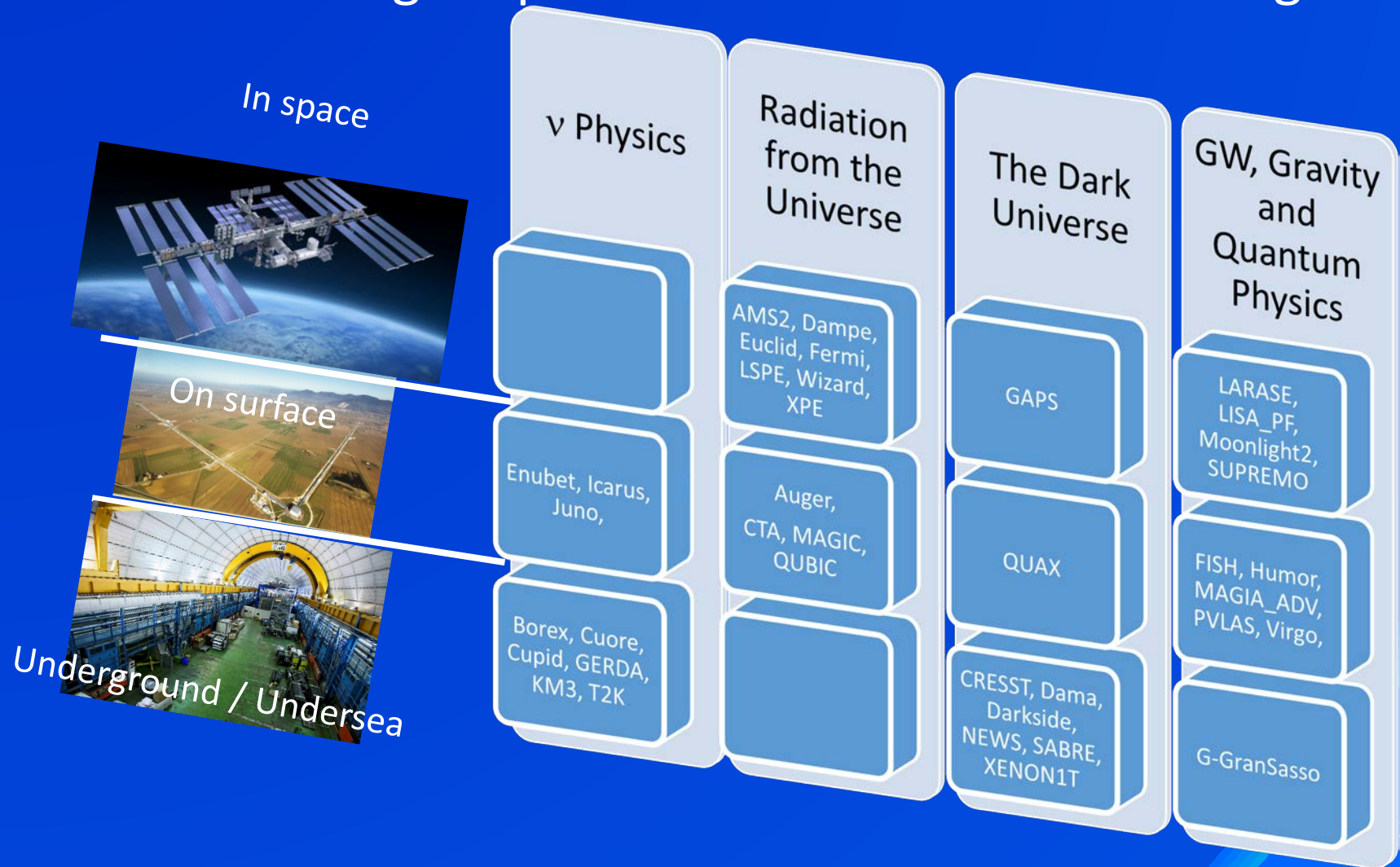
INFN Perugia

Keywords

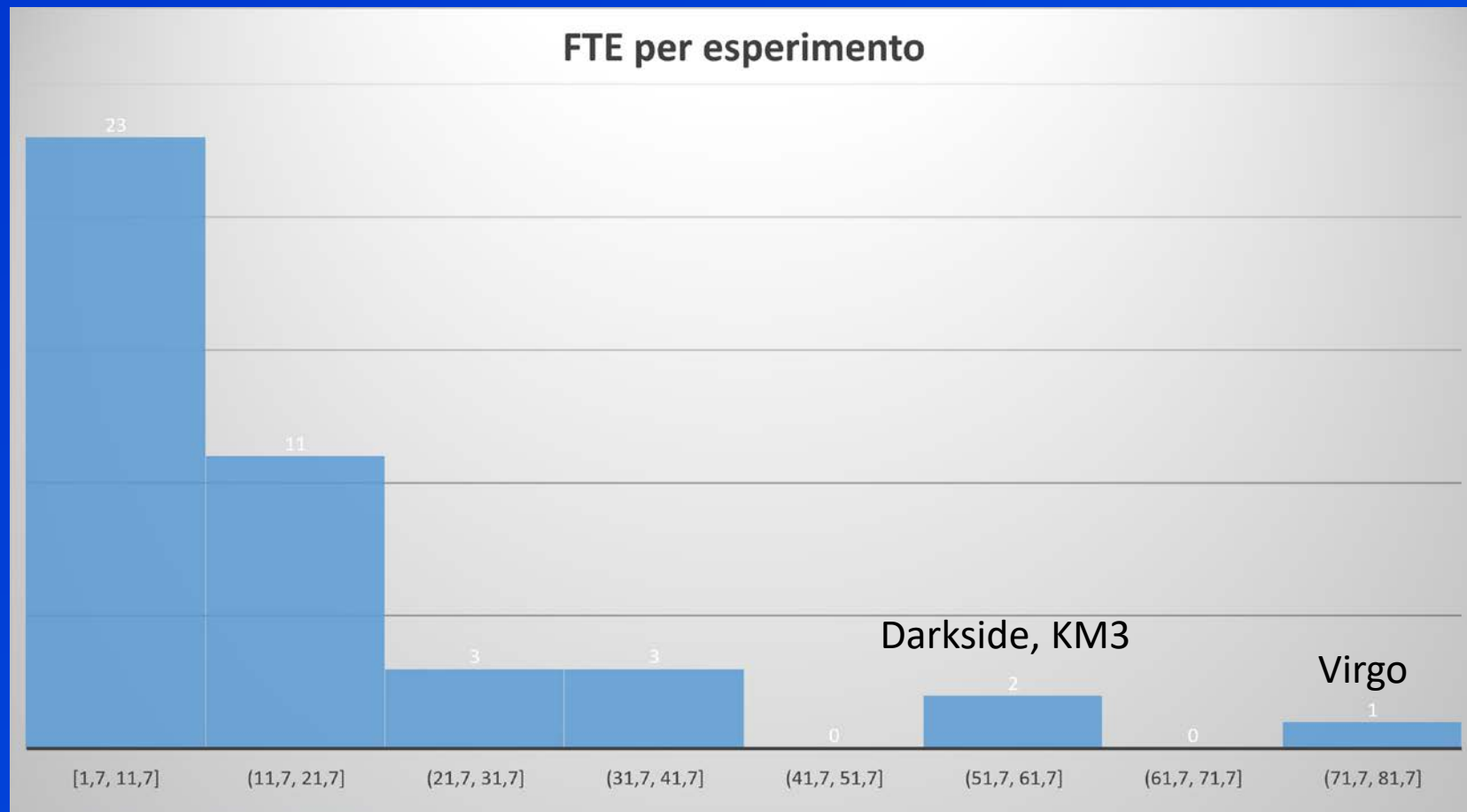
- Dipingere uno scenario per l'evoluzione del calcolo degli esperimenti relative alla CSN2 è un compito arduo
- Le parole chiave che descrivono il complesso del calcolo in CSN2 sono:
 - Eterogeneità
 - Sviluppo
 - Scarsità

Eterogeneità

- I Target scientifici, le metodiche, le collocazioni, le dimensioni degli esperimenti di CSN2 sono «eterogenei»:



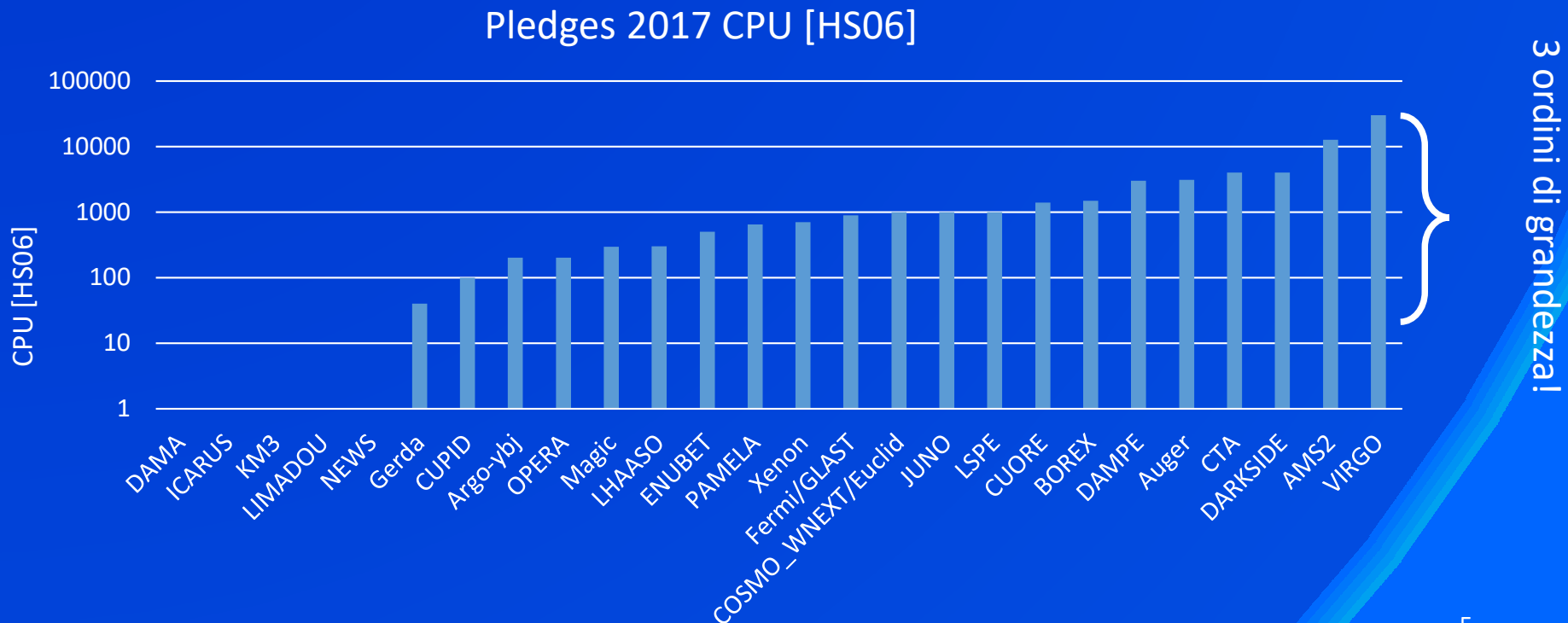
Eterogeneità: Dimensioni (FTE)*



* Database assegnazioni INFN 2017

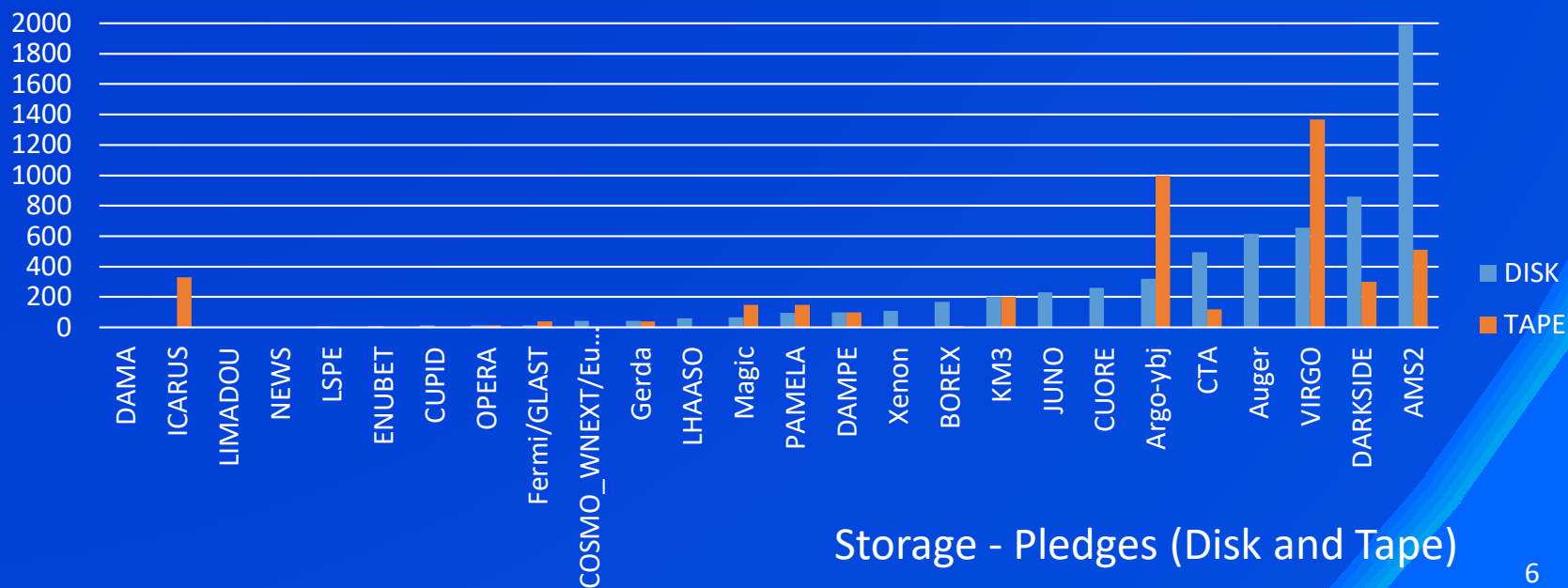
Eterogeneità nel calcolo: CPU

- La diversità tra gli esperimenti si trasforma in una diversità delle esigenze di calcolo, delle implementazioni dei modelli di calcolo (quando esistono) e quindi delle risorse allocate per esso:



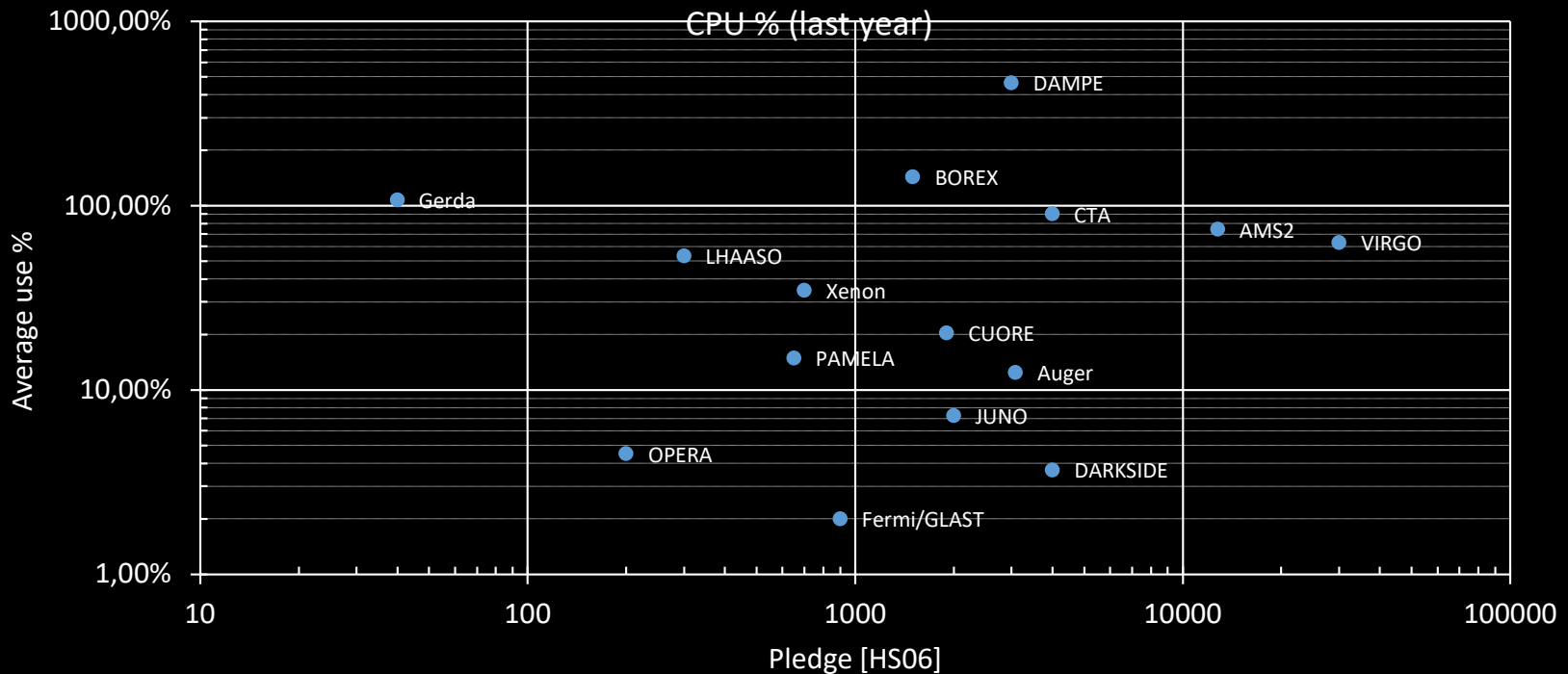
Eterogeneità: Storage

- Anche lo storage mostra l'eterogeneità degli esperimenti in CSN2
- Ma rivela anche un uso «disinvolto» del disco rispetto al tape forse a volte legato ad un computing model poco meditato



Efficienza nell'uso delle risorse

- Siamo sicuri che il modello di condivisione delle risorse implementato ora al CNAF soddisfi i bisogni della CSN2?
- Siamo efficienti?
- Non sarebbe meglio aver risorse comuni per i piccoli esperimenti di CSN2 e avere risorse allocate indipendenti solo per 3-4 esperimenti sopra-soglia?



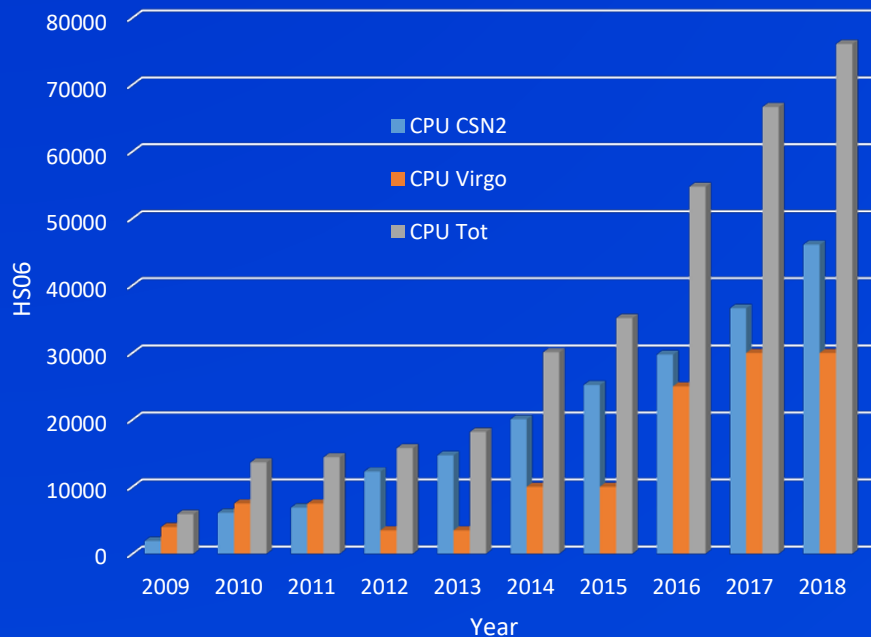
Eterogeneità: HPC vs HTC

- Il modello di calcolo con risorse condivise (à la LHC) non si adatta a tutti gli esperimenti
 - LIGO→Virgo adotta principalmente (per ora) un modello a farm dedicate
 - Nuovi esperimenti, di estrazione astronomica, necessitano di architetture HPC
 - Euclid
 - LSPE
 - QUBIC
 - Finora si è avuta una certa difficoltà a soddisfare queste richieste.

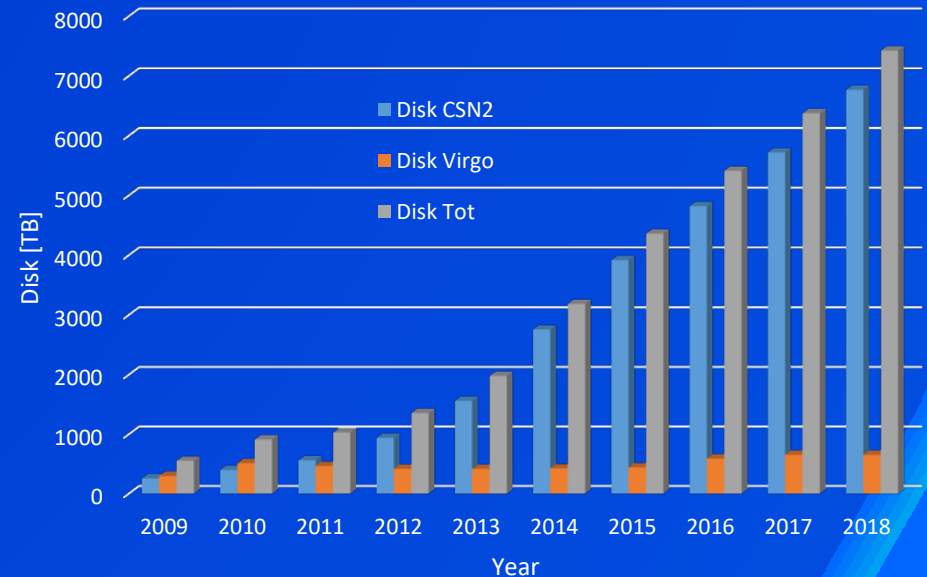
Sviluppo: passato

- Nonostante i problemi di maturità ed efficienza, il calcolo in CSN2 ha avuto un grande sviluppo negli anni passati

Pledges CPU (Pledges-2017, proposal 2018)



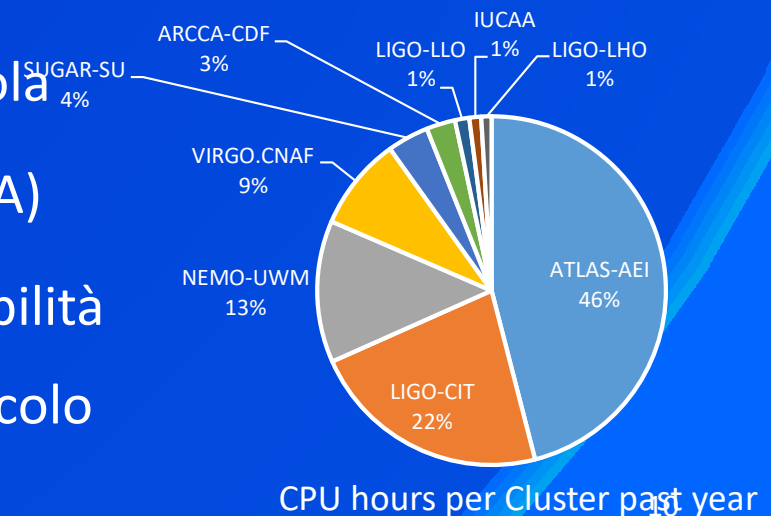
Disks (Pledges -2017, proposal 2018)



- Incremento CPU annuo medio negli ultimi 5 anni: 25%

Sviluppo: Futuro

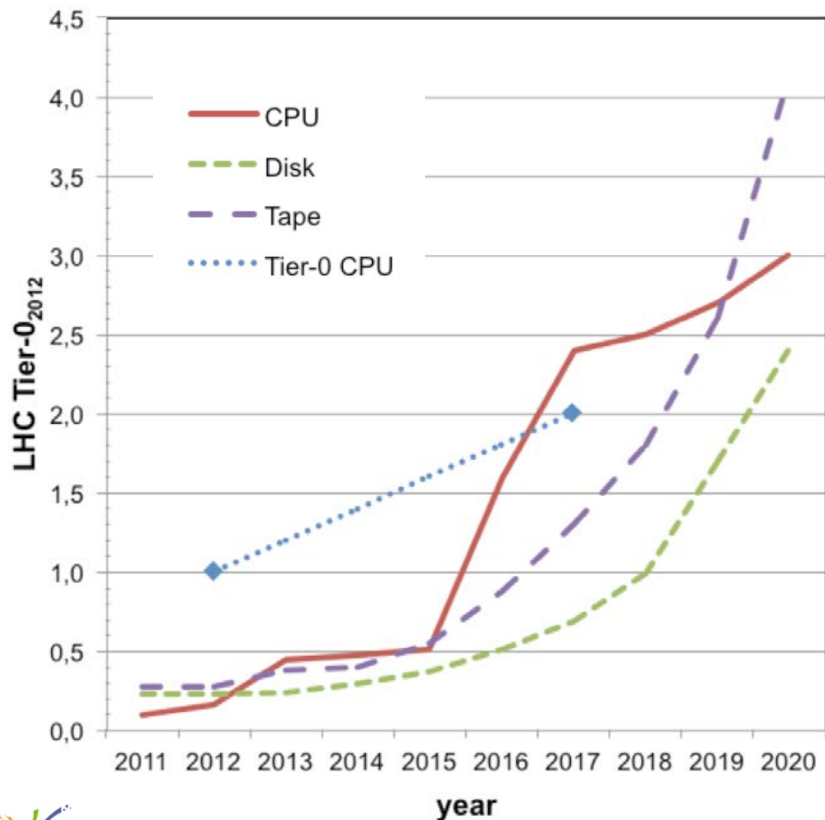
- Gli sviluppi futuri saranno dettagliati nei prossimi talks, che descrivono tra i maggiori consumers di risorse di calcolo in CSN2
- Qui possiamo vedere i punti chiave:
 - GW: Il network di detector di onde gravitazionali completa la sua configurazione minima (3 detectors) quest'anno con l'ingresso di Virgo. Attualmente il calcolo CNAF è un 10% del calcolo totale LIGO-Virgo (e Virgo contribuisce per un 15% circa)
 - Le necessità di calcolo per una buona frazione delle pipelines di analisi cresceranno di un fattore 2-3 per la sola presenza continuata di Virgo.
 - Nel 2020 entrerà il 4° detector (KAGRA) con un'ulteriore crescita a scalino
 - L'aspettato miglioramento della sensibilità a bassa frequenza determinerà una crescita notevole delle esigenze di calcolo delle analisi basate su template



Sviluppo: Altri esperimenti

- Attualmente AMS2 è il secondo consumer di CPU in CSN2 con circa 13kHS06.
 - È prevista una crescita costante nei prossimi anni (2-3kHS06/year) a cui si affianca una crescita parallela e tumultuosa di Dampe
- Relativamente grandi users saranno in futuro:
 - Darkside, Juno, Icarus, ...
- CTA presenta una grossa incognita:
 - Se il CNAF diventa uno dei data center di CTA le richieste di hardware (CPU, Disk & Tape) genereranno una discontinuità
- Anche escludendo il DC-CTA, le esigenze di CSN2 (incluso Virgo) potrebbero raddoppiare entro il 2020

Sviluppo: Valutazione APPEC (2015)

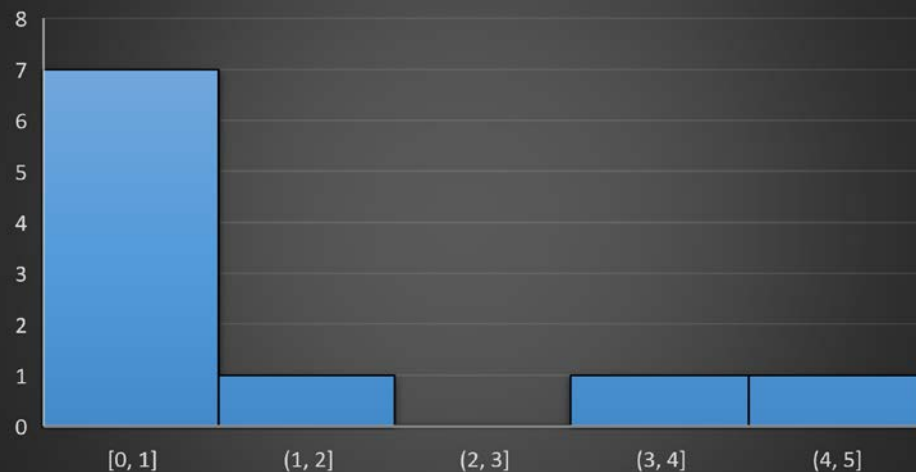


- Necessità espressa in LHC Tier0₂₀₁₂:
 - 65000 cores (650kHS06)
 - 30PBytes disk storage
 - 30PBytes tape storage
- CPU “scalano” con fabbisogno GW

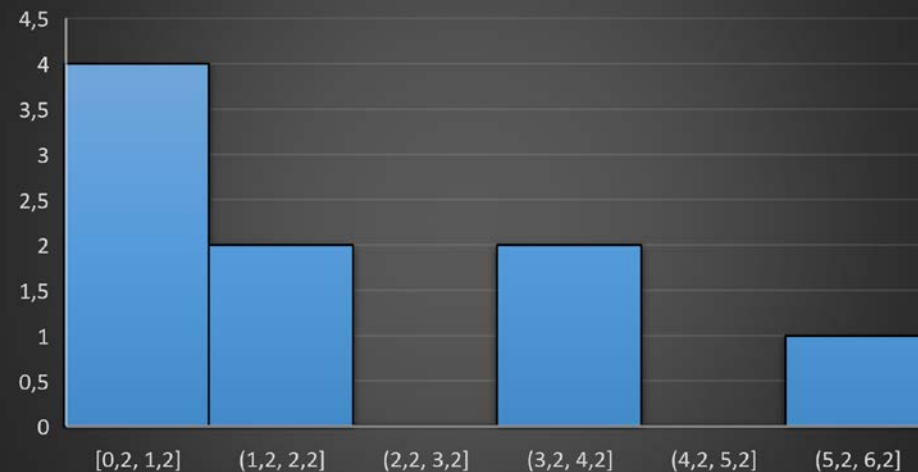
Scarsità: risorse umane

- Gli esperimenti di CSN2 hanno una totale mancanza di personale dedicato interamente o parzialmente al calcolo
 - Alcuni di essi faticano anche a capire cos'è il personale di calcolo

Totalmente dedicati al calcolo



FTE



Scarsità: Competenze

- Gli esperimenti di CSN2 in genere devono sviluppare un maggiore know-how in termini di computer science
- Spesso il calcolo è la cenerentola degli investimenti negli esperimenti
- Occorre uno sforzo nell'INFN per sostenere gli esperimenti di CSN2 in questo processo
 - Risorse umane
 - Competenze e Consulenze
 - Nuove soluzioni
- Decisione strategica (e draconiana) della CSN2:
 - Solo il calcolo al Tier1 è raccomandato (e appoggiato)
 - Soluzioni «in house» non vengono finanziate

Conclusioni

- La ricerca in astroparticelle/gravitazione/neutrini sta affrontando una grande evoluzione
 - Crescita di dimensioni
 - Multidisciplinarietà
 - Crescita di risultati
- Il calcolo è parte integrante di questa evoluzione
 - La CSN2 sta puntando ad una gestione più organizzata e professionale delle esigenze di calcolo
 - Investimenti (di varia natura) a supporto saranno (e già sono) necessari