



# *Trend e requisiti emergenti: la visione dell'INFN*

*Claudio Grandi e Stefano Zani*

*INFN*



**INFN**

Claudio Grandi

Istituto Nazionale di Fisica Nucleare





# Outline

*La presentazione raccoglie informazioni relative ad alcuni **use case** particolarmente significativi per le operazioni dell'INFN*

*Cerchiamo di identificare il **trend** nel prossimo quinquennio in termini di **dimensioni, tecnologie e modalità operative***

*Contributo ad una indagine del CTS del GARR*







***USE CASE:***

***ACCESSO AI DATI SCIENTIFICI***

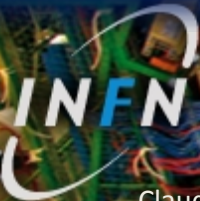
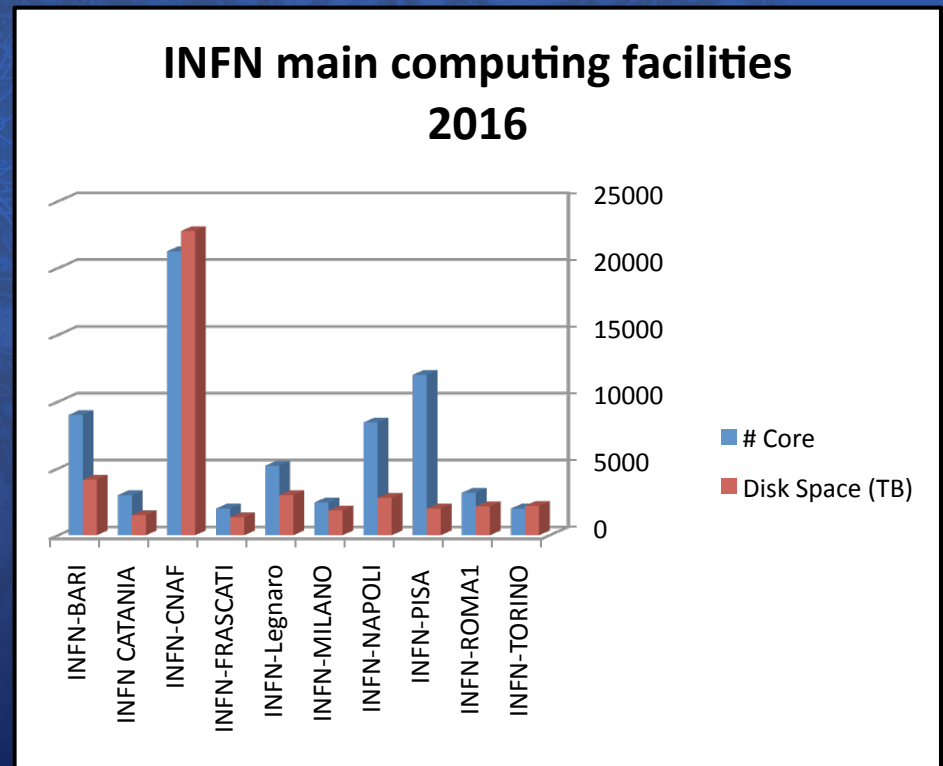




# Infrastruttura di calcolo scientifico

INFN fa parte della comunità **WLCG** e inoltre supporta il calcolo di tutti gli esperimenti in cui è impegnato.

Per svolgere questo compito ha costruito una infrastruttura essenzialmente basata su 1 centro di primo livello (TIER-1) e 9 centri di secondo livello (TIER-2) distribuiti sul territorio per una capacità di calcolo complessiva di circa **68.000 core**, **43 PB** di storage su **disco** e circa **50 PB** di storage su **nastro**.







Claudio Grandi

WWW.GARR.IT





# Infrastruttura calcolo scientifico

*La banda di connessioni fra i centri di calcolo dell'INFN è complessivamente di **220Gb/s**.*

*Il traffico fra i centri di calcolo coinvolti a livello mondiale, avviene principalmente attraverso due reti (VRF dedicati): **LHCOPN** e **LHCONE**.*

*L'accesso alle risorse di calcolo avviene prevalentemente attraverso interfaccia "**GRID**".*

*L'accesso ai dati avviene per quanto possibile al livello locale ma sta crescendo l'utilizzo di protocolli di accesso remoto che consentono di effettuare analisi su dati geograficamente distanti dai nodi calcolo con conseguente impatto sui pattern di collegamento geografico.*

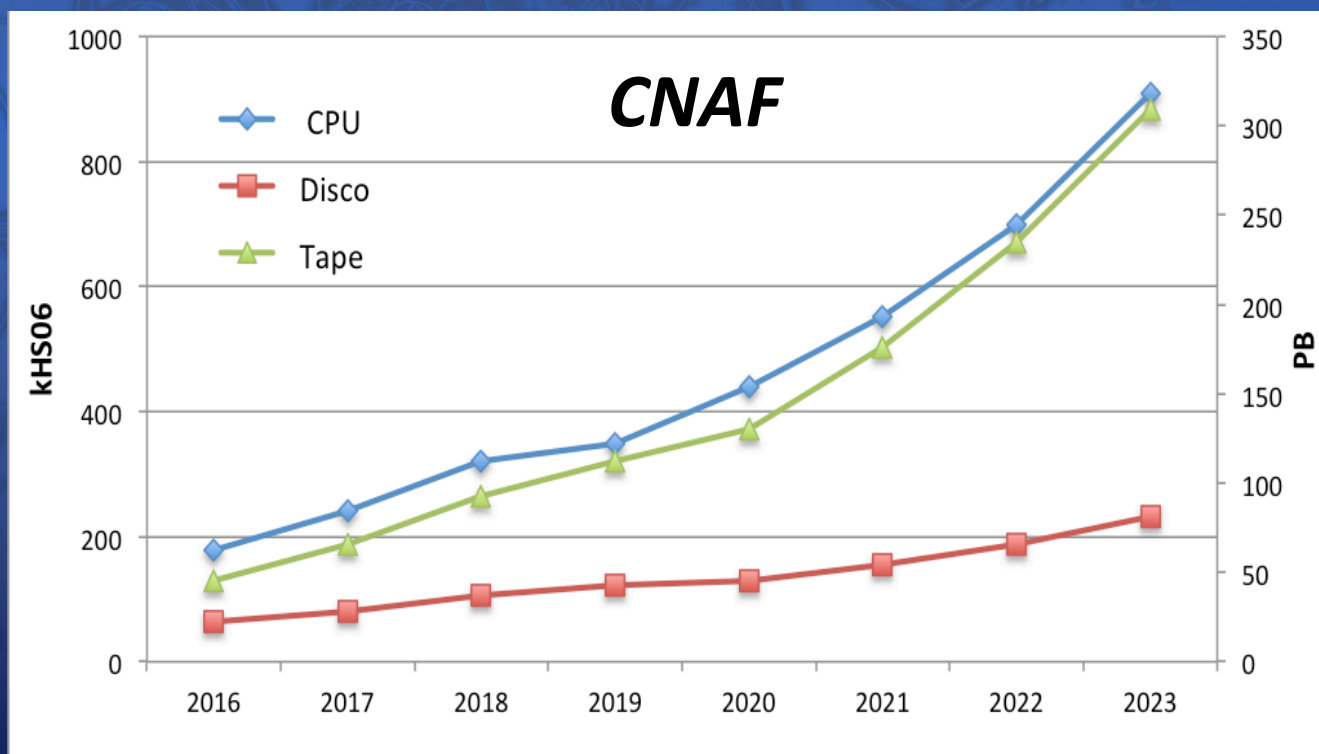




# Previsioni di crescita

L'aumento di risorse di calcolo previsto è di un fattore 2 nel prossimo triennio ed un **fattore circa 4 nei prossimi 5 anni.**

*Ci si aspetta nei prossimi anni un aumento delle necessità di banda proporzionale all'aumento delle risorse di calcolo.*





# Previsioni sulle modalità operative

*Ci si aspetta un aumento di accesso alle risorse di calcolo tramite interfacce “Cloud”*

*Le risorse saranno più **dinamiche**: crescerà la necessità di estendere i centri di calcolo su risorse esterne quali per esempio risorse di partner scientifici, risorse opportunistiche o commerciali, sia in modo statico che dinamico.*

*Ci si aspetta che l'**accesso remoto** ai dati aumenti e sarà quindi indispensabile lo sviluppo di tecnologie per l'ottimizzazione (ad esempio **caching** dei dati) in modo da non sovraccaricare la rete.*







# Previsioni sulle tecnologie

Per soddisfare queste esigenze sarà utile disporre di tecnologie di rete in grado di **attivare rapidamente connessioni** ad alta capacità fra siti diversi fornendo anche all'utente gli strumenti per gestire (per configurazione e monitoring) tali connessioni.

Nell'ambiente HEP si sta investigando sull'utilizzo di piattaforme quali Spark e/o Hadoop per la trattazione di grandi moli di dati utilizzando tecnologie di “**Machine Learning**”.







# *USE CASE:* *SERVIZI ICT*







# *Situazione attuale*

*INFN utilizza servizi ICT realizzati all'interno dei propri centri di calcolo .*

*I servizi includono sia **sistemi informativi** a supporto delle attività amministrative sia **strumenti collaborativi**.*

*Ospitando anche servizi strategici, sono stati implementati meccanismi di **Disaster Recovery**.*







# Previsioni di evoluzione

INFN sta rivedendo le infrastrutture per l'erogazione dei servizi ICT con l'obiettivo di implementare **Business Continuity** e **Disaster Recovery** per tutti i servizi fondamentali e razionalizzare gli investimenti.

Per la realizzazione di tali obiettivi sarà necessario creare connessioni dedicate (fisiche o virtuali) fra i centri di calcolo che ospiteranno le risorse ed utilizzare nuovi strumenti che consentano di **gestire infrastrutture virtuali in modo centralizzato** consentendo anche la migrazione dei servizi fra le sedi fisiche in modalità manuale o automatica.

Si prevede l'implementazione di **servizi federati** che possano essere condivisi con collaboratori anche esterni all'Ente (Per esempio Cloud storage, Open access repository, ...)







# Commenti conclusivi

*E' necessario aumentare le conoscenze su tecnologie per la **gestione "software" delle reti.***

*Gestione federata di **AuthN/AuthZ** e in generale di accesso alle risorse, condivisione dinamica delle risorse (Grid → Cloud) ed aspetti di sicurezza legati a tali metodologie di condivisione.*

*Il costo economico della **sicurezza** sulla rete è alto. E' indispensabile sviluppare tecnologie per ridurre i costi senza pregiudicarne l'efficacia.*

