



investiamo nel vostro futuro

Il progetto ReCaS – infrastruttura distribuita di calcolo per SuperB ed altre applicazioni

L. Merola^x, V. Boccia^{*}, G. Russo^x on behalf of ReCaS collaboration

Workshop GARR su Calcolo e Storage Distribuito, 28-29 Novembre, Roma

Sessione: operations di grandi infrastrutture di calcolo distribuito

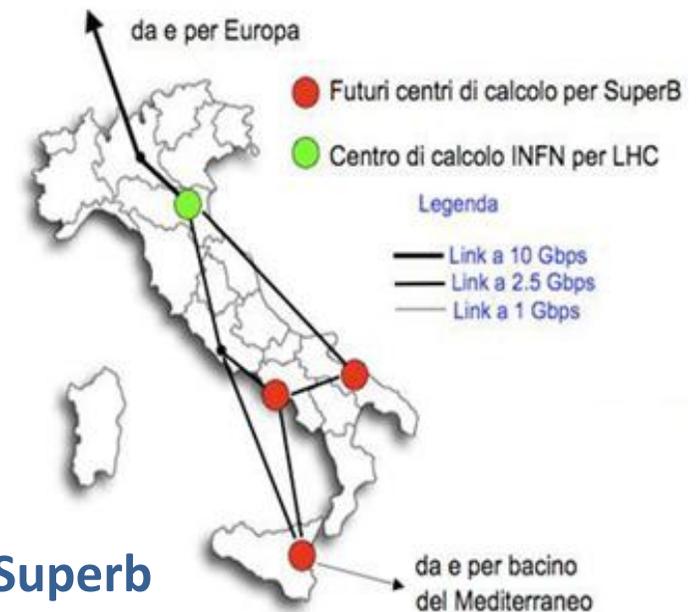
^x Università degli Studi di Napoli “Federico II” ed INFN Sezione di Napoli

^{*} INFN Sezione di Napoli – progetto speciale IGI



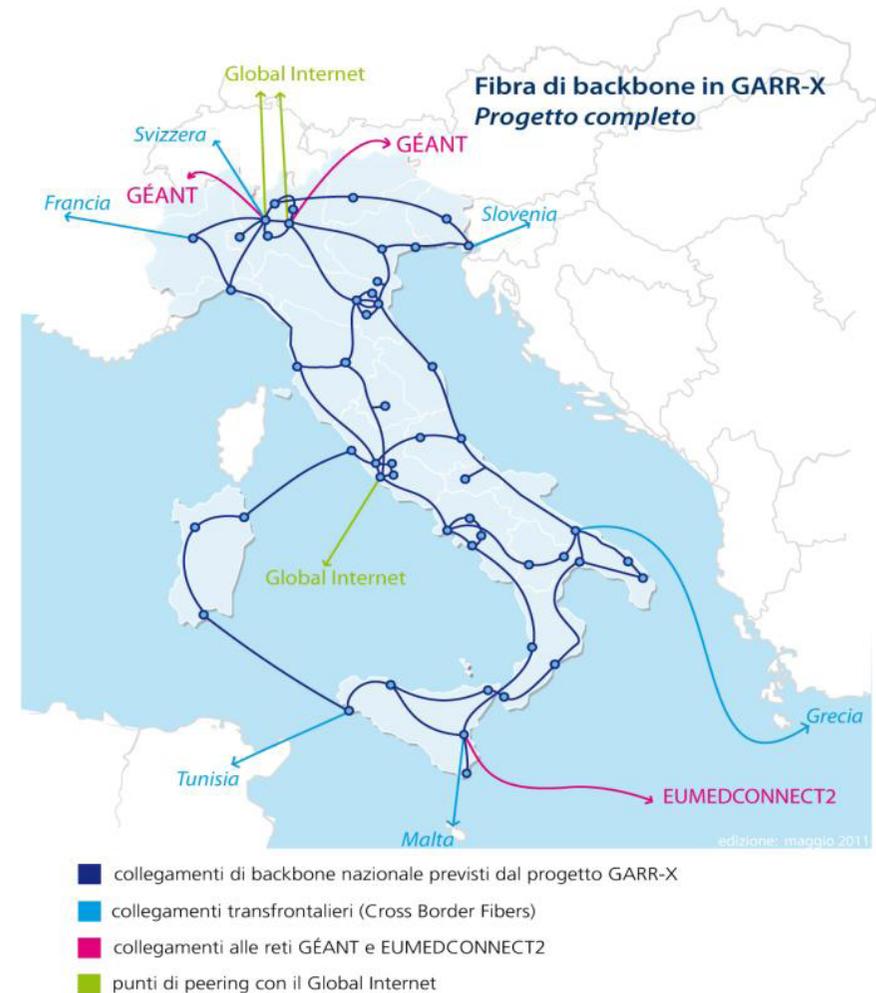
Obiettivi generali del progetto

- Realizzazione di una *e-infrastruttura* distribuita di calcolo e storage nel Mezzogiorno d'Italia che colleghi i 4 Data Center dei partner di progetto:
 - Bari (partner INFN e UNIBA)
 - Napoli (partner INFN e UNINA)
 - Catania (partner INFN)
 - Cosenza (partner INFN)
- Supportare su tale infrastruttura:
 - la comunità del **progetto-bandiera INFN Superb e di LHC (ATLAS, CMS, Alice, LHCb)**
 - altre comunità scientifiche, con particolare riguardo alla **Fisica, alla Medicina, alle Biotecnologie, e all'Ambiente.**



Obiettivi quantitativi (rete)

- In tutte le sedi verranno realizzate:
 - una rete locale in tecnologia 10 GbE per il trasporto dati ed
 - una rete locale in tecnologia 1 GbE per il monitoring, il telecontrollo, la gestione
- La rete GARR-X del Consortium GARR sarà usata per l'interconnessione delle sedi dei Data Center.



Obiettivi quantitativi (calcolo e storage)

- **Calcolo:**
 - L'aggregato di potenza elaborativa minima che il progetto si prefigge di acquisire è di 33 kHepSpec
- **Storage:**
 - mentre la quantità di storage che si prevede di mettere in linea è di 5.5PByte.

	Potenza elaborativa (kHepSpec)	Storage (PByte)
UNINA	6	0,8
INFN-NA	2	0,3
UNIBA	10	2,5
INFN-BA	3	0,5
INFN-CT	7	0,8
INFN-CS	5	0,6
TOTALI	33	5,5

Contestualizzazione Nazionale ed Europea

- Il progetto andrà a contestualizzarsi nelle roadmap Europee e Nazionali delle infrastrutture, quindi nei piani ESFRI, e-IRG, nella roadmap Italiana - come richiesto dal Bando - e dovrà favorire l'internazionalizzazione specialmente a sostegno dell'area MED.



Attività in ambito SuperB

- Attualmente attività di simulazione (FastSim, FullSim)
- A valle della rimodulazione subita dal progetto si potranno soddisfare le esigenze iniziali per il calcolo e storage di SuperB fino al 2016
- Circa il 50% delle attrezzature saranno nuove acquisizioni tramite il progetto. Il resto è già disponibile nei Data Center già esistenti grazie a precedenti PON 2000-2006.

Attività in ambito LHC

- Il “commissioning” finale, e gli “stress test”, saranno eseguiti soprattutto sui dati reali raccolti dagli *esperimenti in corso all’ LHC al CERN (ATLAS, CMS, ALICE, LHCb)*.
- Il sito di Napoli sarà integrato con le risorse del TIER2, utilizzato essenzialmente dall’esperimento ATLAS
- Il sito di Bari sarà integrato con le risorse del TIER2, utilizzato essenzialmente dall’esperimento CMS
- Le risorse di ReCaS saranno importanti per la comunità LHC

Attività in ambito nazionale

- Nel progetto e nella rimodulazione è stato previsto che *si realizzino strumenti e servizi di "alto livello"*, in modo da rendere possibile *l'integrazione delle risorse fornite dai 4 centri di calcolo nelle infrastrutture di Grid/Cloud nazionale ed internazionale*
- Le attività si possono così riassumere:
 - **Portali** di accesso alle risorse
 - Integrazione di servizi Grid e Cloud in **un'unica soluzione open**
 - Ottimizzazione e semplificazione della gestione delle elaborazioni e del carico di lavoro sulle risorse
 - Offerta di servizi di **calcolo e storage virtualizzato** e gestione di **cataloghi** e archivi **integrati**
 - Realizzazione di un sistema di **monitoring e di telecontrollo** degli impianti centralizzato verso una postazione di management in ogni sede operativa di RECAS.

Altre collaborazioni

- Collaborazioni con il mondo produttivo nazionale ed internazionale, in particolare con i progetti Smart Cities
- Promozione dell'infrastruttura verso i Paesi dell'area mediterranea (in collaborazione con EUMED) , del Sud America, Asia e Africa (in collaborazione con CHAIN)
- Contestualizzazione nello Spazio Europeo della Ricerca (IGI, EGI, ESFRI).

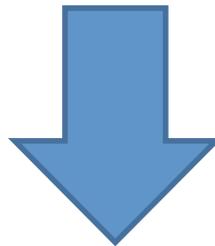
La e-Infrastruttura di produzione: strategie di deployment

- Prevista la creazione e l'integrazione di 4 siti:
 - *RECAS-BARI*
 - *RECAS-CATANIA*
 - *RECAS-COSENZA*
 - *RECAS-NAPOLI*

- Particolare attenzione sarà riservata:
 - alle tematiche di **interoperatività** delle infrastrutture dei partner del progetto,
 - all'**integrazione** dei diversi paradigmi di calcolo parallelo e distribuito (**HPC, Grid e Cloud**)

Strategie comuni per la scelta dei nodi di calcolo

- Numero anche elevato di core per nodo
- Presenza di acceleratori GPU
- Presenza di nodi con storage locale di grandi dimensioni
- Reti di interconnessione tra i nodi ad alta velocità



- Supporto alle applicazioni su differenti paradigmi di parallelizzazione (Message Passing, MultiThread e GPU-based) eventualmente anche combinati tra loro
- Supporto alle applicazioni data intensive

Altre strategie di deployment

- Sulle risorse di calcolo saranno resi inoltre disponibili:
 - strumenti software, PSE, toolkit e librerie che costituirà lo strato software di riferimento delle applicazioni (“*middleware applicativo*”)
 - File system paralleli e distribuiti (attualmente in sperimentazione)
- Al fine di rendere l'infrastruttura distribuita più “flessibile” e meno “statica” delle tradizionali infrastrutture *Grid-based* saranno integrati servizi Cloud.

Il prototipo della e-Infrastruttura

- Attualmente è in produzione un prototipo dell'infrastruttura nelle more dell'espletamento delle gare di acquisizione delle forniture hardware
- Tuttavia già da ora sono implementate gran parte delle strategie di deployment descritte che caratterizzeranno l'infrastruttura di produzione finale del progetto.

Sostenibilità

- Il core business del Progetto ReCaS è l'offerta dei servizi di calcolo/storage avanzati per la ricerca scientifica
- Si prevede inoltre di poter acquisire **nuovi fruitori dei Data Center**, istituzionali e non, attraverso l'opera di Promozione e Comunicazione prevista
- Infine si ipotizza anche la fornitura di servizi di calcolo/storage di base, per **clienti privati**, con esigenze non immediatamente risolvibile dai players già presenti sul mercato.