

P.O.N. RICERCA E COMPETITIVITA' 2007-2013 - Azione I "Interventi di rafforzamento strutturale"
PONA3_00052, Avv. 254/Ric. - ReCaS (Rete di Calcolo per SuperB e altre applicazioni)



ABSTRACT sottoposto in risposta alla Call for Papers, sul tema "operations di grandi infrastrutture di calcolo distribuito"

Titolo: *Il progetto ReCaS – la nuova infrastruttura di calcolo per SuperB ed altre applicazioni*

Autori: L. Merola^(1,2), V. Boccia⁽²⁾, G. Russo^(1,2), on behalf of the ReCaS collaboration

(1) Università degli studi di Napoli Federico II

(2) Istituto Nazionale di Fisica Nucleare, sez. di Napoli

Abstract:

La collaborazione ReCaS (l'INFN con le sedi operative di Bari, Catania, Cosenza, Napoli, l'Università di Napoli Federico II e l'Università di Bari Aldo Moro) sta portando avanti l'omonimo progetto, iniziato ad ottobre 2011, per la realizzazione di una *e-infrastructure* di calcolo nel Mezzogiorno d'Italia. Tale infrastruttura integrerà e potenzierà quanto già esistente nelle varie sedi, in base a precedenti finanziamenti PON o a altri finanziamenti regionali, o a finanziamenti propri degli enti stessi.

L'aggregato di potenza elaborativa minima che il progetto si prefigge di acquisire è pari a 33 kHepSpec , e la quantità di storage che si prevede di mettere in linea è pari a 5,5 PByte. Considerata l'evoluzione della tecnologia, è molto probabile che le capacità di calcolo e di storage andranno al di là di quanto attualmente preventivato.

In tutte le sedi verrà realizzata 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. In tutte le sedi verrà installato un identico sistema di monitoring e di gestione, costituito da 4 server per sede. Il software di monitoring verrà realizzato in-house dagli enti partecipanti. La rete GARR-X del Consortium GARR sarà usata per l'interconnessione delle sedi e dei Data Center.

Ai fini di una economia di scala, molte gare saranno effettuate in modo congiunto, e comunque parallelamente, in modo da ottenere maggiori sconti. Di conseguenza, le specifiche degli apparati, sia di rete sia di calcolo sia di storage, saranno unificate.

All'interno delle attività necessarie nel progetto ReCaS, verranno realizzati 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, in particolare:

- Portali di accesso alle risorse;
- Integrazione di servizi grid e cloud in un'unica soluzione open e utilizzo di tecniche di virtualizzazione sia per le risorse di calcolo che per quelle di archiviazione;
- Ottimizzazione e semplificazione della gestione delle elaborazioni e del carico di lavoro sulle risorse;
- Offerta di servizi di storage virtualizzato e gestione di cataloghi e archivi integrati;
- Realizzazione di un sistema di monitoring e di telecontrollo degli impianti centralizzato.

Il *paper* presenterà le scelte architetturelle fatte per l'impiantistica, gli armadi, la rete, lo storage, i nodi per i servizi, i nodi di calcolo, e le scelte sul software, finalizzato alla creazione di una *e-infrastructure* accessibile sia con il paradigma del grid computing sia con il paradigma del cloud computing. Mentre la prima opzione è particolarmente interessante per un immediato utilizzo nell'ambito di collaborazioni scientifiche già attive, come gli esperimenti in corso presso il *Large Hadron Collider* del CERN, la seconda opzione è pensata per un uso più generale da parte della comunità scientifica nazionale, incluse le ricerche di soggetti privati che operano nel sistema produttivo italiano.

Le *strategie di deployment* adottate per la realizzazione dell'infrastruttura di produzione tengono conto:

- delle tematiche di interoperatività delle infrastrutture dei partner del progetto,
- dell'integrazione dei diversi paradigmi di calcolo parallelo e distribuito (HPC, Grid e Cloud)

In particolare si prevede la realizzazione e integrazione dei siti di produzione RECAS-NAPOLI, RECAS-BARI, RECAS-CATANIA e RECAS-COSENZA ciascuno dei quali renderà disponibili sul proprio sistema informativo (IS) risorse di calcolo e storage. I servizi di tutti i siti saranno integrati tra loro mediante servizi di alto livello, bilanciati ed in alta affidabilità.

Sulle risorse di calcolo, interconnesse da reti ad alte prestazioni, saranno disponibili strumenti utili alle applicazioni data intensive: file system distribuiti/paralleli ad alte prestazioni ed accessibili mediante protocolli di diversi livelli.

Inoltre, al fine di supportare efficientemente ed efficacemente le comunità HPC di utenti del progetto ReCaS, l'infrastruttura comprenderà sistemi di calcolo multinodo con un numero anche elevato di core per nodo e in taluni casi dotati anche di acceleratori GPU, coerentemente con le linee guida di altri progetti internazionali per la progettazione e realizzazione di sistemi di calcolo HPC per i prossimi dieci anni (ad es. *l'International Exascale Software Project - IESP*). Tale caratteristica renderà l'infrastruttura ReCaS particolarmente adatta ad ospitare anche applicazioni basate su differenti paradigmi di parallelizzazione (Message Passing, MultiThread e GPU-based) eventualmente anche combinati tra loro.

Sulle risorse di calcolo saranno resi disponibili ed opportunamente documentati strumenti software, PSE, toolkit e librerie che costituirà lo strato software di riferimento delle applicazioni ("*middleware applicativo*").

Al fine di rendere l'infrastruttura distribuita più "flessibile" e meno "statica" delle tradizionali infrastrutture *Grid-based* saranno integrati servizi Cloud.

Attualmente è in produzione un prototipo dell'infrastruttura nelle more dell'espletamento delle gare di acquisizione delle forniture hardware.

Dettagli sul progetto sono disponibili sul sito www.pon-recas.it.