

# Introduzione

Giuseppe Attardi, Vincenzo Vagnoni

*Chair del Comitato di Programma del Workshop GARR CSD 2012*



Le infrastrutture distribuite di calcolo e storage, siano esse griglie computazionali (Grid) o Cloud, sono oggi strumenti indispensabili nell'ambito di un sistema evoluto di ricerca. Tutti i grandi Paesi al mondo si sono già dotati, o stanno dotandosi, di sistemi in grado di soddisfare le crescenti richieste di calcolo e storage poste dalla ricerca di frontiera in tutti i campi.

Il calcolo di tipo Grid ha efficacemente soddisfatto nell'arco dell'ultimo decennio le necessità di talune comunità, come i fisici del CERN e dell'INFN, dotate al loro interno di elevate competenze nel settore ICT e caratterizzate da una struttura di grande scala. Nel contempo, tuttavia, questo modello ha non poco faticato a penetrare in molti altri ambiti. Di fatto, le comunità di ricerca più piccole, per arrivare sino ai singoli gruppi o ai singoli ricercatori, che non hanno potuto permettersi investimenti in formazione, personale e mezzi, sono rimaste al di fuori del circuito, e hanno continuato a lavorare con strumenti tradizionali senza poter avere accesso alla capacità di eseguire complicati calcoli in tempi brevi e memorizzare ingenti quantità di dati. Uno dei limiti principali della Grid risiede infatti nella complessità di utilizzo da parte degli utenti, che ha portato alla necessità di progettare nuove soluzioni.

Col termine di Cloud si identificano oggi una serie di soluzioni finalizzate a mettere a disposizione servizi di elaborazione in maniera flessibile e scalabile: piattaforme (*Platform as a Service, PaaS*), infrastrutture (*Infrastruc-*

*ture as a Service, IaaS*), software (*Software as a Service, SaaS*) e alcune combinazioni per offrire servizi di storage (*Storage as a Service, STaaS*) o di condivisioni di dati (*Data as a Service, DaaS*).

Il Cloud Computing rappresenta quindi un insieme di tecnologie e di interfacce sviluppate al fine di conseguire maggior dinamismo nell'utilizzo di risorse di calcolo distribuito, siano esse CPU, dispositivi di storage o infrastrutture di rete. Sebbene la metafora del Cloud Computing abbia avuto origine degli anni '60, il modello ha cominciato a diffondersi solo negli ultimi 6 anni, dapprima attraverso servizi commerciali, e in seguito attraverso la sua adozione anche in ambiti di ricerca pubblica.

In termini commerciali, la Cloud rappresenta un modello di business per la vendita di servizi e sistemi informatici on-demand da parte di grandi e piccoli provider di risorse private. La stessa modalità può però essere applicata all'interno di un'organizzazione di dimensioni significative, per condividere e rendere disponibili le risorse evitando sprechi e duplicazioni.

In alternativa a mettere in comune servizi di base, per le esigenze della ricerca scientifica si può pensare a rendere disponibili come servizi accessibili via Web raccolte di funzionalità specifiche che rispecchiano le necessità e le consuetudini d'uso di ciascuna comunità di ricerca. È l'approccio degli *Science Gateway*, tramite i quali è possibile eseguire tutta una serie di operazioni, come eseguire applicativi

su infrastrutture Grid o Cloud, ricercare informazioni all'interno di database remoti, movimentare grandi quantità di dati, ecc., senza la necessità di dover imparare un linguaggio di programmazione specifico e senza avere alcuna particolare conoscenza del funzionamento complessivo dell'infrastruttura stessa.

Nell'ambito del workshop GARR Calcolo e Storage Distribuito, nato dall'idea a lungo perseguita da Enzo Valente e dal GARR di riunire le tante anime presenti nelle varie e variegata comunità della ricerca Italiana al fine di esplorare e possibilmente mettere in campo sinergie e condivisioni di competenze e risorse, sono stati presentati 26 contributi da parte di ricercatori e tecnologi impegnati nei vari settori ICT, includendo due keynote speech, da parte di Fabrizio Gagliardi (Direttore del Technical Computing, Microsoft Research), e Ignacio Llorente (Direttore del progetto OpenNebula). Nella fase conclusiva del workshop sono stati anche presentati tre position paper da parte di ENEA, INAF e INFN, contenenti la loro visione strategica per il calcolo e lo storage distribuiti, con l'obiettivo di stimolare una discussione aperta ai circa 130 partecipanti all'evento. Da questa discussione è emersa come fondamentale la necessità di perseguire un sistema integrato di e-Infrastructure della ricerca Italiana, che possa capillarmente raggiungere con i suoi benefici anche la vastissima comunità di piccoli gruppi di ricerca e singoli ricercatori. I grandi attori del settore della ricerca Italiana, come CNR, ENEA, INAF, INFN, INGV insieme ovviamente al sistema universitario e al GARR, possono fare un grande lavoro di condivisione di competenze e infrastrutture che possa finalmente modernizzare gli strumenti ICT, rendendoli fruibili a tutti, e conseguentemente le modalità di fare ricerca nel nostro Paese.

In questo volume sono stati selezionati i contributi ritenuti più significativi da parte del comitato di programma del workshop.

Buona lettura!



#### Chair del Workshop

Giuseppe ATTARDI - Università di Pisa (co-chair)

Vincenzo VAGNONI - INFN (co-chair)

#### Comitato di programma

Roberto BARBERA - INFN

Giovanni BRACCO - ENEA

Luca DELL'AGNELLO - INFN

Marco PAGANONI - INFN

Tutte le presentazioni e maggiori informazioni sono disponibili sul sito dell'evento:

<http://www.garr.it/ws-garr-csd>