

Le Grid nella comunità della Ricerca

www.garr.it/ricerca

Le Grid (o Griglie Computazionali) sono un insieme di servizi avanzati su Internet che permettono agli utenti di condividere risorse di calcolo, spazio di archiviazione, dati e strumentazione scientifica tra siti geograficamente dispersi ed appartenenti a domini amministrativi differenti.

Nate all'interno della comunità della ricerca, oggi le Grid sono utilizzate in moltissime applicazioni nei diversi settori della scienza e della tecnologia.

I servizi Grid sono integrati e coordinati in modo da fornire all'utente la possibilità di utilizzare tutte le risorse di cui ha bisogno, attraverso un'unica operazione di login, generalmente basata sull'utilizzo di certificati digitali.

Grazie alle Grid, infatti, un utente, in virtù della sua affiliazione ad una determinata Organizzazione Virtuale, sfrutta numerose funzionalità (accesso a dati, archiviazione, calcolo, strumenti di collaborazione) senza preoccuparsi di alcun dettaglio architetturale, logistico o amministrativo.

Nonostante si tratti di una tecnologia in costante evoluzione, nei prossimi anni le Grid avranno, per il mondo della ricerca, un impatto comparabile a quello del World Wide Web. Tutti conoscono il WWW ed ogni giorno milioni di persone lo utilizzano per condividere informazioni su Internet, ma non molti sono a conoscenza delle potenzialità delle Grid.

L'idea di base sta alla radice stessa del nome grid (letteralmente "griglia") con cui, in inglese, normalmente si indica la rete di distribuzione dell'energia elettrica: l'energia viene infatti fornita senza che l'utente sia tenuto a sapere come essa sia stata prodotta. Lo stesso si può fare per le risorse di calcolo e di archiviazione dati grazie alle Grid.

Con l'uso delle Griglie, in un prossimo futuro, i computer utilizzati quotidianamente per lavoro potranno diventare un'unica, grande risorsa computazionale a cui chiunque potrà accedere on demand in ogni momento: insomma ciascun utente potrà disporre dell'equivalente di un enorme supercomputer semplicemente connettendosi alla Grid dalla sua postazione.

Come funzionano le Grid

Il Grid computing è un esempio particolare di calcolo distribuito, basato sull'idea di condividere risorse su scala globale.

Le Grid sono costituite da:

- un software (il cosiddetto middleware) capace di gestire le risorse ed allocarle in maniera ottimale ai vari utenti ed alle applicazioni che ne fanno richiesta, un po' come il sistema operativo fa con i programmi che sono lanciati sul proprio PC;
- un meccanismo (che generalmente fa parte del middleware) capace di fornire un accesso sicuro alle risorse e in grado di garantire la confidenzialità e l'integrità dei dati (particolarmente importante in alcuni campi, come ad esempio quello biomedico, in cui la privacy è un aspetto fondamentale);
- una connessione di rete affidabile e a larga banda tra le risorse, che assicuri che il tempo impiegato per il trasferimento dei dati sia trascurabile se comparato ai benefici di una elaborazione più rapida, ottenuta grazie al calcolo distribuito effettuato su un gran numero di computer.



 Consortium
GARR

WWW.GARR.IT
info@garr.it



Le applicazioni

Le applicazioni scientifiche che ormai stabilmente utilizzano le Grid sono innumerevoli ed afferiscono ai domini più diversificati: dall'Astronomia alla Fisica delle Alte Energie, dalla Chimica computazionale alle Scienze della terra, dalla Biologia alla Genetica computazionale, dalla Medicina e diagnostica medica, alla Scienza dei materiali.

Nel settore della **Fisica delle Alte Energie**, ad esempio, ci si trova davanti ad una sfida computazionale molto impegnativa: l'analisi dei dati che verranno prodotti dal Large Hadron Collider del CERN di Ginevra. LHC è una macchina collisore protone-protone ad alta energia che si propone di investigare argomenti fondamentali per la fisica contemporanea. Per far questo è necessario disporre di enormi risorse di calcolo e di archiviazione dei dati da rendere accessibili alla comunità mondiale dei fisici. Per distribuire il più possibile queste risorse è stata creata W-LCG, la **Worldwide LHC Computing Grid** (<http://lcg.web.cern.ch>), una infrastruttura di Grid che complessivamente è in grado di fornire ai fisici risorse di CPU e di archiviazione pari ad oltre 100.000 processori collegati in 140 istituti in tutto il mondo.

Nel campo della **Bioinformatica** e della **Genetica computazionale** le Grid sono utilizzate per moltissimi scopi come lo studio del genoma umano, l'analisi delle sequenze di DNA, il confronto tra sequenze note di codice genetico nello studio della proteosintesi. Molte applicazioni utilizzate nel settore genetico, come ad esempio BLAST o Autodock, sono state portate su Grid, fornendo ai ricercatori l'enorme vantaggio di poter accedere a quantità di risorse di calcolo e di dati fino ad ora mai raggiunte.

Un caso di particolare successo è il progetto **WISDOM** (<http://wisdom.eu-eggee.fr>) che ha applicato le Grid alle tecniche di docking *in silico*, riuscendo a dare un significativo contributo alla lotta alla malaria. Per ottenere gli stessi risultati conseguiti in due mesi, infatti, sarebbero stati necessari circa 80 anni con una tradizionale tecnologia.

Un altro progetto internazionale di ricerca è **BioinfoGrid** (www.bioinfoGRID.eu) che, grazie allo sfruttamento dell'elevata potenza di calcolo ha portato allo sviluppo di soluzioni ottimali per la scoperta di nuovi farmaci.

Nell'ambito della **Medicina** sono moltissimi i progetti già avviati. Le Grid sono state con successo interfacciate agli standard più comuni per la gestione delle analisi mediche, come per esempio DICOM. Questo permette oggi una migliore osservazione di immagini mediche (PET, TAC, Radiografia) e l'implementazione di interi workflow di analisi.

Attraverso le Grid è possibile gestire in maniera sempre più efficiente le cartelle cliniche rendendo le prestazioni mediche maggiormente personalizzate, nell'ottica di una medicina che gravita tutta intorno al paziente. L'utilizzo della rete e delle Grid inoltre permette il teleconsulto e la tele diagnosi soprattutto dove non vi sono impianti medici immediatamente a disposizione dei pazienti.

Negli ultimi anni, numerosi progetti internazionali si sono basati sull'utilizzo delle Grid. Un esempio è **Health-e-Child** che si occupa della definizione di una terapia pediatrica incentrata sulle specificità di ogni bambino, grazie all'integrazione verticale di tecniche diagnostiche e di analisi di dati medici con la possibilità di gestire tutta la cartella clinica dei piccoli pazienti. Altri importanti progetti di medicina riguardano il settore della Neuroscienza: **NeuGRID**, che utilizza le Grid per l'analisi delle immagini PET per la diagnosi precoce del morbo di Alzheimer e **Neuroinf**, il portale di Neuroinformatica italiano, con il quale, attraverso una semplice interfaccia web, vengono fornite risorse di calcolo ai medici che devono effettuare analisi delle immagini PET del cervello di pazienti affetti da Alzheimer o nei casi di sospetta malattia.



CASE STUDIES

Health-e-child

Progetto europeo che ha come scopo lo sviluppo di un sistema computerizzato di supporto clinico alla diagnosi e alla ricerca in pediatria, basato sull'integrazione verticale di dati e conoscenze di natura biomedica.

Il progetto è coordinato a livello clinico dall'Istituto G. Gaslini di Genova e ha tra i partner l'Ospedale Bambino Gesù di Roma e l'Università di Genova.

www.health-e-child.org

NeuGRID

Progetto europeo che prevede lo sviluppo di un'infrastruttura digitale per analizzare immagini tridimensionali di risonanza magnetica che consentono l'identificazione dei marcatori delle malattie neurodegenerative come l'Alzheimer.

Il progetto è coordinato dal Centro S. Giovanni di Dio-Fatebenefratelli di Brescia.

www.neugrid.eu

Neuroinf

Portale italiano delle Neuroscienze realizzato nell'ambito di un progetto internazionale che coinvolge 15 diversi paesi. L'obiettivo è fornire uno strumento per la condivisione delle informazioni in un panorama interdisciplinare.

Partner del progetto è la Fondazione San Raffaele del Monte Tabor.

www.neuroinf.it



Consortium
GARR