

UNIVERSITÀ DEGLI STUDI DI NAPOLI
"PARTHENOPE"



INAUGURAZIONE
DELL' ANNO ACCADEMICO 2001-2002

Prolusione: Prof. Giulio Giunta

Supercalcolo e reti nella ricerca e nella formazione

Napoli, 28 Gennaio 2002

Aula Magna
Napoli - Via Acton

Supercalcolo e Reti nella Ricerca e nella Formazione

Giulio Giunta, 28 Gennaio 2002

Signor Ministro, Autorità, Magnifici Rettori, Colleghi, Studenti, Signore e Signori, questa prolusione è un'occasione per esporre alcune riflessioni sugli effetti che le attuali, enormi risorse di calcolo e la presenza diffusa di una infrastruttura di rete stanno avendo e avranno nel breve e nel medio termine sull'organizzazione della ricerca e, di riflesso, sulla formazione.

1. Il contesto attuale

Sia i fautori del determinismo tecnologico sia quelli dell'evoluzionismo sociale sono concordi nel ritenere che stiamo entrando nella cosiddetta *società dell'informazione*, il cui tratto fondamentale e distintivo è l'accesso istantaneo, per via telematica, alle informazioni e agli strumenti computazionali per la loro elaborazione.

La Rete (Internet e il Web), alimentata dall'innovazione nelle tecnologie dell'informazione e della comunicazione, è l'infrastruttura che rende possibile l'accesso all'informazione e la sua fruizione e sta diventando uno strumento importante nel campo dell'economia e delle interazioni sociali e anche nella ricerca scientifica e nella formazione.

Alcuni dati danno un'idea della situazione attuale e della rapidità della sua evoluzione. La Rete connette milioni di calcolatori e contiene miliardi di pagine Web; la sua crescita è esponenziale, con il traffico che raddoppia ogni 100 giorni; tutte le nazioni a tecnologia avanzata hanno reti nazionali a larga banda che collegano praticamente tutti i centri di

ricerca e le università; le università della maggior parte dei paesi usano la Rete per erogare in modo permanente vari livelli di servizi, sia didattici sia amministrativi; il reddito che la Rete produrrà nel 2002 è stimato in 1300 miliardi di Euro; il 93% delle imprese americane usa la posta elettronica (il 71% in Italia), il 76% delle imprese svedesi ha un sito web per il marketing (il 44% in Italia).

Anche nel campo della tecnologia dei sistemi di calcolo e delle discipline computazionali si assiste a crescite di tipo esponenziale. Ogni anno e mezzo raddoppia il numero di componenti elettroniche in un processore (è la cosiddetta *legge di Moore*, che può anche essere sintetizzata dicendo che i calcolatori raddoppiano la loro potenza, a parità di costo, ogni 18 mesi); l'efficienza degli algoritmi per la risoluzione dei problemi di base in svariate discipline ha registrato negli ultimi quaranta anni un incremento paragonabile a quello dell'hardware. Inoltre, negli ultimi quindici anni si sono realizzati sistemi di calcolo (*supercalcolatori*) in cui un gran numero di unità di elaborazione distinte (attualmente fino a diecimila) operano in parallelo in modo coordinato, fornendo potenze di calcolo superiori a quelle previste dalla legge di Moore. In tal modo, la comunità scientifica ha a disposizione risorse computazionali in grado di erogare, a costi contenuti, le cosiddette *potenze di supercalcolo*, che oggi consentono di effettuare da 50 miliardi a 5000 miliardi di operazioni al secondo.

Per dare un'idea concreta della rapidità dell'evoluzione in questo campo, basti pensare che in un'attuale automobile del segmento medio c'è più potenza di calcolo di quanta ce ne fosse a bordo dell'Apollo 11 (missione lunare, 1969) e che una normale piattaforma per videogiochi dell'ultima generazione (dal costo di circa 300 Euro) ha una potenza di calcolo superiore a quella erogata dal supercalcolatore più potente del 1988 (il cui costo era di circa 10 milioni di Euro).

Il supercalcolo è da alcuni anni considerato una risorsa strategica, tanto che il livello di sviluppo tecnologico di un paese è

quantificato anche in base a nuovi parametri, come il numero di supercalcolatori nella classifica detta dei *top 500* (i 500 supercalcolatori più potenti, così distribuiti: 230 in Usa, 59 in Germania, 57 in Giappone, 34 in Gran Bretagna, 23 in Francia, 16 in Corea, 11 in Italia, etc.) oppure la potenza di calcolo per abitante (264.000 operazioni al secondo in Usa, 22.000 in Italia).

Quali sono gli effetti visibili di questo contesto sulla ricerca scientifica?

Attualmente la Rete viene utilizzata dalla quasi totalità dei ricercatori di tutte le aree scientifiche.

Alcuni rilevano che il suo *caos enciclopedico* sia un pericolo e un freno alla conoscenza piuttosto che il *labirinto generatore di percorsi e di sapienza* che Diderot vedeva nella sua *Encyclopédie*.

Questo aspetto potenzialmente negativo della Rete può essere idealmente riassunto ricordando le splendide pagine in cui J.L. Borges descrive la Biblioteca di Babele. Borges immagina una *Biblioteca universale* in cui sono raccolti tutti i libri, non solo quelli scritti nella storia dell'umanità, ma anche tutti i possibili libri che possono essere scritti, ovvero quelli che si possono comporre effettuando tutte le possibili combinazioni delle lettere dell'alfabeto. Cito da Borges e probabilmente qualche utilizzatore della Rete riconoscerà una peregrinazione virtuale vissuta direttamente: [la Biblioteca comprende] *la storia minuziosa dell'avvenire, le autobiografie degli arcangeli, il catalogo fedele della Biblioteca, migliaia e migliaia di cataloghi falsi, la dimostrazione della falsità di questi cataloghi, la dimostrazione della falsità del catalogo autentico,..... Quando si proclamò che la Biblioteca comprendeva tutti i libri, la prima impressione fu di straordinaria felicità. Tutti gli uomini si sentirono padroni di un tesoro intatto e segreto. Non v'era problema personale o mondiale la cui soluzione non esistesse in qualche scaffale. ... Migliaia di ricercatori vagavano accapigliandosi nella speranza di trovare il libro della propria vita, ma quei ricercatori dimenticavano che la probabilità che un uomo trovi il suo libro è sostanzialmente prossima a zero.*

La metafora della Rete come Biblioteca suggerisce che la Rete è la naturale evoluzione delle tecniche di accesso alle informazioni e di trasmissione della conoscenza che hanno determinato lo sviluppo scientifico e tecnologico soprattutto da Gutenberg in poi, ma che in effetti hanno caratterizzato fin dalle origini la storia dell'uomo, a partire dalle prime pitture rupestri.

L'aspetto fondamentale della Rete è soprattutto quello di consentire alle persone di interagire tra loro, simile in questo alle infrastrutture classiche, dalle strade dei romani, al sistema ferroviario del diciannovesimo secolo, a quello telefonico del ventesimo.

2. L'impatto sulla ricerca

Nel campo scientifico, uno dei vantaggi che sta emergendo, tra quelli indotti dalla rapidità di interazione via Rete, è che non solo i ricercatori coinvolti in ricerche su tematiche simili possono dialogare e confrontarsi più velocemente, ma anche i ricercatori di aree diverse sono spinti alla cooperazione e riescono a collaborare nella risoluzione di problemi che sono intrinsecamente multidisciplinari. In questo senso la Rete sembra favorire quella *concatenazione delle scienze* che Diderot immaginava come la vera finalità della sua Enciclopedia.

D'altra parte, la comunità scientifica è fisiologicamente predisposta a operare nella dimensione globale di interazione, di confronto e di competizione che la Rete sta realizzando. Non a caso sia le prime reti sia il Web sono nati nei laboratori di ricerca sulla spinta dell'esigenza di rapidità e flessibilità di comunicazione tra gruppi di scienziati.

L'impatto delle metodologie computazionali e della potenza di calcolo sulla ricerca è ancora più profondo e sta modificando il metodo stesso dell'indagine scientifica.

Da Galileo in poi, l'indagine scientifica poggia sui pilastri dell'*approccio sperimentale* e dell'*approccio teorico*.

Oggi si è affermato in modo irreversibile un terzo approccio, l'*approccio computazionale*.

E' una metodologia che offre un ulteriore grado di libertà nello studio dei fenomeni, aprendo nuove prospettive all'indagine scientifica; essa taglia orizzontalmente tutte le discipline, dalle scienze della vita all'astrofisica, dalle scienze della Terra all'economia, dalle scienze sociali a quelle dell'ambiente.

Le metodologie e le tecniche computazionali consentono di simulare, cioè di approssimare, il comportamento di fenomeni e di sistemi reali.

I vantaggi di tale approccio sono molteplici. Si possono studiare sistemi su cui non è economicamente conveniente fare esperimenti, come nel caso dell'aerodinamica, dei giacimenti petroliferi, dei *crash test* di automobili, oppure sistemi su cui gli esperimenti sono estremamente pericolosi, sui pensi all'energia nucleare, ai motori a combustibile solido, oppure ancora sistemi su cui è impossibile fare esperimenti, come nel caso dello studio dell'atmosfera e degli oceani, del clima e della meteorologia terrestre, della struttura interna della Terra, dei sistemi economici e finanziari.

Le potenze di supercalcolo consentono di studiare fenomeni sempre più complessi, di aumentare il livello di dettaglio della descrizione di un sistema e quindi di pervenire a simulazioni più realistiche, di ottenere in tempi accettabili i risultati di più simulazioni dello stesso sistema al variare di alcune condizioni rilevanti, e permettono ai ricercatori di interagire direttamente, anche visivamente (attraverso le tecniche di visualizzazione scientifica), con il sistema simulato che è oggetto dell'indagine.

Si pensi a quanto questo sia importante in campo ambientale, per valutare quantitativamente gli effetti di diverse politiche di riduzione dell'emissione di inquinanti, o in campo industriale, per considerare varie alternative di progetto e di realizzazione di prodotti, ma anche

nell'indagine sperimentale, sia in fase di progetto dell'esperimento, sia in fase di analisi dei dati.

Di fatto si tratta di una rivoluzione che sta modificando la ricerca scientifica ma anche la pratica dell'attività tecnica in quasi tutti i settori della tecnologia.

E' ormai opinione diffusa che, in molte discipline scientifiche e in molti settori industriali, la potenza di calcolo e la rapidità di accesso ai dati siano il fattore più importante per l'avanzamento della conoscenza e per lo sviluppo tecnologico.

3. Un ipotizzabile scenario futuro

Qual è lo scenario ipotizzabile per la ricerca scientifica?

La Rete diventerà una connessione globale dove miliardi di persone interagiranno a vari livelli usando e generando informazione.

I sistemi di calcolo avanzato e le relative metodologie permetteranno di affrontare (entro il 2010) la risoluzione di problemi con potenze di supercalcolo dell'ordine di un milione di miliardi di operazioni al secondo, con la possibilità di immagazzinare miliardi di miliardi di informazioni elementari. Per avere un'idea concreta di questi ordini di grandezza, si consideri che si stima che gli uomini, dalla preistoria a oggi, abbiano pronunciato non più di un miliardo di miliardi di parole.

Una enorme quantità di dati, in misura che non ha precedenti nella storia della scienza, verrà generata grazie allo sviluppo delle attuali tecniche di osservazione *in situ* e *remote* (via satellite), che daranno la possibilità di osservare, anche in tempo reale, praticamente ogni fenomeno naturale.

Il supercalcolo consentirà di analizzare tali masse di dati a vari livelli, di aggregarli in modo associativo (generando quella che è tecnicamente definita *conoscenza*), di simulare la maggior parte dei fenomeni in modo sempre più accurato, anche in virtù della conoscenza estratta dai dati, di utilizzare dati e modelli computazionali per fare

previsioni affidabili in tutti i campi, dalla meteorologia alla climatologia, dai terremoti ai mercati finanziari.

La condizione necessaria affinché questo scenario si concretizzi è che non solo le informazioni ma anche le risorse computazionali siano accessibili in modo pervasivo, standardizzato, sicuro, scalabile e a basso costo.

L'infrastruttura tecnologica globale che realizzerà questa evoluzione è la Griglia (*the Grid*).

La Griglia è un modo nuovo di pensare e utilizzare le informazioni e gli strumenti di calcolo. L'idea ispiratrice è di considerare la potenza di calcolo come una risorsa di pubblica utilità da distribuire in modo diffuso, analogamente ad altre risorse classiche come l'energia elettrica o l'acqua potabile.

Dalla stanza di ogni ricercatore e, a più lungo termine, dagli uffici e dalle case sarà possibile collegarsi a una presa che erogherà la potenza di calcolo necessaria per risolvere un qualunque problema computazionale. Per l'utilizzatore non avrà importanza sapere da dove provenga, come venga generata, da chi venga controllata la risorsa di calcolo che sta usando, come accade oggi quando si inserisce una spina nella presa elettrica per far funzionare un elettrodomestico e raramente ci si chiede da chi, come e dove sia stata generata l'energia elettrica che si utilizza.

La Griglia sarà costituita da molte griglie computazionali e si costruirà negli anni a partire dalle griglie che saranno sviluppate dalla comunità scientifica, finalizzate a problemi specifici, per esempio all'astrofisica o alla genetica, e poi dalle griglie computazionali regionali e nazionali.

La Griglia trasformerà l'insieme dei calcolatori collegati all'infrastruttura di rete in un unico supercalcolatore, virtuale e distribuito,

e sarà la realizzazione della *macchina universale* sognata da Babbage, il professore della cattedra *lucasiana* di Cambridge (che nel 1669 fu la cattedra di Newton e che nella famosa saga di fantascienza *Star Trek* è occupata nel 2200 da *Data*, un androide supercomputerizzato), che per primo intuì e cercò di progettare, nel secolo diciannovesimo, una macchina generale esecutrice di algoritmi per la risoluzione di problemi. Per inciso, l'idea di Babbage si scontrò contro il realismo del primo ministro inglese dell'epoca Benjamin Disraeli che, a proposito dell'utilità di una simile macchina, excepì testualmente che l'unica sua utilizzazione immaginabile era il calcolo del suo costo stratosferico.

Un eventuale scetticismo su questo scenario è comprensibile. D'altra parte, è ben noto che si tende a sovrastimare le aspettative sulle ricadute immediate della tecnologia e invece a sottostimare i suoi effetti a medio e lungo termine.

Alcuni esempi possono aiutare a comprendere che questo orizzonte di sviluppo, almeno per quanto riguarda la ricerca scientifica, è altamente probabile.

Attualmente, il più grande supercomputer mai utilizzato per la ricerca *non esiste*, nel senso che non esiste come un'unica macchina localizzata in un luogo geografico e con caratteristiche costanti nel tempo. Infatti, la massima potenza di calcolo è stata ottenuta utilizzando circa tre milioni di normali pc, sparsi negli uffici e nelle case del mondo, e facendoli operare sotto un controllo debolmente centralizzato. Si tratta del famoso progetto SETI. I pc sono messi volontariamente a disposizione dai proprietari e sono utilizzati durante i periodi di tempo in cui non vengono utilizzati, per esempio quando entra in funzione un salvaschermo oppure durante le ore notturne. Per concedere il proprio computer basta collegarsi al sito Web del progetto e riempire un semplice modulo elettronico. La finalità scientifica del progetto è quella di analizzare i dati del radiotelescopio di Arecibo, per cercare presenze di forme intelligenti e comunicanti nell'universo.

Nel progetto Folding, decine di migliaia di pc sono utilizzati, con modalità analoghe, per simulare la conformazione tridimensionale di proteine, a partire dai dati raccolti dal progetto Genoma Umano.

I sistemi di supercalcolo realizzati da questi due progetti, in cui la potenza di calcolo è altamente diffusa e frazionata, sono utilizzabili efficacemente per risolvere problemi scientifici che possono essere formulati come un gran numero di sottoproblemi indipendenti.

Per molti problemi la potenza di supercalcolo deve essere erogata da un unico sistema (o al più da un piccolo insieme di sistemi cooperanti) di calcolo avanzato; quindi, i sistemi diffusi non sostituiranno i supercalcolatori, che saranno componenti essenziali delle griglie computazionali.

Il progetto TeraGrid, recentemente finanziato dalla National Science Foundation americana, collegherà con connessioni molto veloci (circa 25 volte più veloci della più veloce connessione attualmente operativa in Italia) i supercalcolatori presenti in quattro centri di ricerca, creando il primo esempio di griglia computazionale su scala nazionale. Si tratta della più grande infrastruttura in termini di quantità di risorse computazionali e umane e di vastità di competenze mai messa in campo per la ricerca scientifica. In Teragrid i ricercatori coopereranno per la risoluzione di problemi (dinamica molecolare, modelli per il controllo e la previsione ambientale, visualizzazione scientifica avanzata,...), agendo in gruppi virtuali in cui c'è solo l'unità di azione ma non quella di luogo e di tempo.

Estrapolando nel medio termine, la Griglia globale sarà lo strumento principale che la comunità scientifica userà per generare conoscenza.

La Griglia contribuirà a rendere facilmente e rapidamente accessibili anche alle imprese e alla società i risultati e i prodotti dell'attività di ricerca e costituirà, per tutte le società avanzate, la principale infrastruttura per la raccolta, l'approvvigionamento e lo sviluppo delle competenze e del sapere. In questo senso, è ragionevole

ritenere che diventerà uno strumento strategico per la crescita economica e civile.

L'evoluzione dalla Rete alla Griglia viene descritta dagli studiosi di scienze sociali come passaggio dalla società dell'informazione alla *società della conoscenza*, perché il suo tratto distintivo e caratteristico è che la potenza di calcolo, opportunamente trasformata in conoscenza e competenze, diventa un servizio fruibile da tutti.

4. Prospettive per la formazione

Le prospettive per la formazione sono strettamente correlate con quelle per la ricerca.

Nel quadro di questo cambiamento che si preannuncia profondo e rapidissimo è necessario ripensare soprattutto agli obiettivi della formazione. Se è vero che l'Università deve dare risposte non tanto a quello che la società chiede ma soprattutto a quello di cui la società ha realmente bisogno, allora l'Università e la comunità scientifica devono guidare coerentemente e pianificare con tempestività la trasformazione dei processi formativi.

Innanzitutto, i futuri ricercatori devono essere formati in modo adeguato, poiché saranno gli attori dello sviluppo scientifico che prima è stato delineato. Essi alimenteranno l'innovazione e determineranno la rapidità di crescita della società dell'informazione e della conoscenza. Saranno la chiave del circolo virtuoso nel quale la formazione sostiene l'innovazione tecnologica che a sua volta migliora lo sviluppo di competenze.

Le metodologie e gli strumenti dell'approccio computazionale devono essere presenti, a vari livelli, in tutto il percorso di formazione. Lo scopo non è la cosiddetta *alfabetizzazione informatica*, ovvero il semplice addestramento all'utilizzo, per così dire domestico, del calcolatore e di alcuni programmi di produzione individuale, ma una formazione che educi anche alla visione algoritmica dei problemi, alla

sperimentazione individuale mediante l'interazione con simulazioni computazionali di fenomeni concreti e che trasmetta competenze operative per l'utilizzo effettivo degli strumenti di calcolo avanzato.

In altre parole, si tratta di educare a una sorta di *bilinguismo*, che metta in grado i ricercatori di *parlare fluentemente* sia il *linguaggio* del proprio settore disciplinare sia il *linguaggio* computazionale.

Poiché la Rete e la Griglia renderanno sempre più naturale l'accesso alle informazioni e alle risorse computazionali, una delle finalità della formazione dei ricercatori deve essere quella di possedere la capacità di muoversi efficacemente nella ricchezza di queste potenzialità e di utilizzarle concretamente per costruire competenza individuale. Sotto certi aspetti ciò estende il ruolo tradizionale del processo formativo perché deve aiutare a *imparare come imparare*.

Per parlare più in generale di formazione, è opportuno considerare due dati di fatto: calcolatore, accesso alla rete e competenze computazionali sono entrati a far parte della definizione di alfabetismo; in una società e in una economia avanzata è necessario un livello più elevato e più diffuso di formazione.

I soggetti erogatori di formazione devono quindi ripensare e rimodulare la propria offerta formativa in funzione del fatto che i destinatari della formazione costituiranno una platea molto più vasta di quella tradizionale degli studenti per così dire istituzionali e che il momento della formazione e dell'aggiornamento sarà un'esigenza che si ripeterà più volte durante le carriere professionali.

In prospettiva futura, la formazione sarà uno dei servizi fondamentali a cui tutti potranno accedere attraverso la Griglia. L'Università non potrà non essere tra i soggetti sviluppatori e gestori di questo servizio.

È convinzione diffusa che la Rete/Griglia e i relativi paradigmi di comunicazione possano plasmare in profondità i meccanismi di apprendimento e addirittura la coscienza della realtà.

Già Platone manifestava analoga convinzione a proposito della comunicazione scritta.

E' lecito quindi considerare questo cambiamento, che la società si appresta a vivere, come una evoluzione naturale, ma la portata della trasformazione è simile a quella prodotta nel passaggio all'era tipografica della carta stampata, la *repubblica delle lettere* di Galileo, oppure all'era televisiva del *villaggio globale* di McLuhan.

Ne discende un aspetto non trascurabile degli obiettivi della formazione. Nel contesto globale realizzato dalla Rete/Griglia gli studenti, e in generale tutto il lavoro intellettuale, avranno una esposizione al confronto e alla cooperazione con gli altri studenti e alla comparazione con altri standard che si proporrà su una scala mai sperimentata dall'uomo. Il processo formativo deve aiutare a creare l'attitudine psicologica sia a misurarsi e confrontarsi con ampie platee e standard elevati sia a cooperare con gli altri e a sostenere la fatica del lavoro in gruppo.

E-learning, net-learning, telelearning, formazione a distanza, multimedialità e teleimmersione sono parole d'ordine ormai di uso comune e forse addirittura sovrautilizzate. E' convinzione di molti e anche mia che gli strumenti ottimali per la formazione in Rete emergeranno e si imporranno esclusivamente in base alle scelte dei loro utilizzatori e alla loro capacità di ampliare l'interazione docenti-studenti.

E' noto che ci sono molte perplessità sull'efficacia delle nuove tecnologie come strumento di supporto alla formazione.

Mi sia consentito sintetizzarle e, indirettamente, confutarle parafrasando le parole che Platone fa dire al suo Socrate (nel *Fedro*) a proposito degli effetti della scrittura nell'elaborazione e nella trasmissione del sapere, da lui considerati intrinsecamente peggiorativi rispetto alla tradizione orale: i nuovi strumenti di comunicazione *avranno per effetto il produrre la dimenticanza nelle anime di coloro che li usano,*

perché fidandosi di tali strumenti si abitueranno a ricordare dal di fuori mediante segni estranei e non dal di dentro e da sé medesimi.

Molti docenti sono concordi su un punto, ovvero che già oggi una formazione che utilizzi anche la Rete consente agli studenti di non essere solo consumatori passivi di informazioni ma di diventare anche produttori di informazioni facilmente utilizzabili da altri studenti e di arricchire le competenze e l'esperienza dei docenti e delle Università.

A ben pensarci, si tratta ancora di una evoluzione naturale: le *comitive*, ovvero i gruppi di docenti e studenti che elaboravano e agivano insieme nell'Universitas medievale, si ripropongono come aggregazioni libere da vincoli di vicinanza geografica.

Termino qui le mie considerazioni sulle prospettive future per la ricerca e la formazione e ricordo che, in definitiva, il modo migliore per predire il futuro è operare per costruirlo.