

## Le risorse HPC e la chimica teorica e computazionale in Italia

### Premessa

Questo documento, espressione di una comunità scientifica dedita allo sviluppo ed all'applicazione di strumenti teorico-computazionali alla chimica e fisica dei sistemi molecolari, materiali, bio- e nano-sistemi, vuole richiamare l'attenzione della collettività e in particolare della direzione del CINECA, il maggiore centro di calcolo italiano, sull'importanza di coinvolgere l'utenza scientifica nella formulazione di una adeguata politica di gestione delle risorse calcolo e in particolare HPC del nostro paese, secondo linee di un modello partecipato e collaborativo, in modo da consentire alla comunità computazionale italiana di consolidare i risultati ottenuti e rimanere competitiva, in Europa e nel mondo.

I sottoscritti firmatari sono componenti di un gruppo di lavoro indicato all'unanimità dall'assemblea dei membri della Divisione di Chimica Teorica e Computazionale (CTC) della Società Chimica Italiana (SCI), tenutasi a Padova il 22 febbraio 2013, col mandato di porsi come interlocutore e trasmettere alla direzione del CINECA le istanze della comunità chimica computazionale italiana.

### Gestione dell'HPC e rapporti con la comunità scientifica

Uno dei requisiti necessari per la ricerca scientifica moderna è la capacità di raccolta, trasmissione ed elaborazione di dati. Il calcolo scientifico ad alte prestazioni (*high performance computing*, HPC) è lo strumento fondamentale di tutte le discipline scientifiche dalla fisica alla chimica e alla biologia, in tutte le loro possibili declinazioni tecnologiche, dall'ingegneria dei materiali alle nano- e bio-tecnologie. È evidente come una politica adeguata di gestione delle risorse HPC possa rappresentare una strategia virtuosa per lo sviluppo della ricerca scientifica: essa richiede, infatti, investimenti relativamente modesti mentre consente ricadute immediate e diffuse in tutti i settori di sviluppo primari, favorisce sia la crescita culturale (ricerca di base) che l'innovazione tecnologica.

La comunità dei chimici teorici-computazionali da vari decenni è stata tra i principali attori nell'impiego innovativo di strumenti HPC. È utile ricordare a questo proposito che la comunità chimico-teorica rappresenta una frazione significativa di quella computazionale italiana (15%). Negli ultimi tempi però la comunità ha vissuto con sempre maggiore disagio l'evoluzione delle politiche nazionali relative al calcolo scientifico. Vari gli elementi che hanno generato tale disagio, ma in ultima analisi tutti riconducibili alla trasformazione dell'offerta HPC da un modello di "servizio" alle varie comunità scientifiche ad un modello in cui l'aspetto decisivo è il prestigio o la massima utilizzazione dei sistemi di calcolo più che il mettere a disposizione strumenti computazionali avanzati all'utente reale.

Un esempio è dato dalla modalità di convergenza del CASPUR (Consorzio di calcolo del Centro sud) e del CILEA (Consorzio di calcolo del Nord) nel CINECA (Consorzio Interuniversitario di Calcolo Avanzato) che, pur rispondendo ad una precisa direttiva ministeriale di razionalizzazione, è stato di fatto repentino, non consentendo un'adeguata programmazione dell'uso delle risorse disponibili.

Soprattutto, si deve poi rilevare che le recenti scelte strategiche e le politiche di gestione del CINECA sono andate in una direzione diversa da quella di una integrazione reale con una gran parte della comunità scientifica. Il CINECA partecipa con risorse hardware comparabili a quelle di altre nazioni europee all'iniziativa PRACE, ma l'acquisizione del supercomputer Fermi a tecnologia *IBM Blue Gene/Q* e la sostituzione del precedente supercalcolatore *IBM SP6*, se ha permesso di riguadagnare molte posizioni nella famosa classifica dei Top 500 (i più veloci *supercomputer* del mondo), ha avuto ricadute limitate (se non addirittura limitanti in termini di offerta di varietà di piattaforme di calcolo) per alcune classi di applicazioni CTC. Nello specifico:

- Da un punto di vista tecnico, l'impiego di un'unica architettura basata su processori a bassa RAM per core e la necessità di utilizzare un numero elevato di core in parallelo (1024) non favorisce molte applicazioni nel campo della scienza dei materiali e delle bio- e nano-tecnologie, basate su applicativi che richiedono processori altamente performanti, se non al prezzo di una profonda ristrutturazione dei codici.
- Da un punto di vista gestionale si deve poi registrare negli ultimi anni una progressiva diminuzione di assistenza tecnica qualificata, il che ha pesantemente rallentato l'acquisizione da parte di molti utenti del nuovo sistema

e ha limitato sia lo sviluppo di applicazioni innovative degli strumenti informatici sia un insufficiente *porting* dei codici di più largo utilizzo e diffusione.

Il ruolo del maggiore centro di calcolo nazionale si è quindi progressivamente ridotto a quello di fornitore di un servizio tecnico, per di più attribuito su basi premiali fondate sul massimo ritorno dell'investimento finanziario in termini di pura contabilità numerica (ore calcolo) più che sull'intrinseco grado di innovazione dei progetti. Questo risulta evidente anche semplicemente dalla forma poco flessibile assunta dalle recenti call progettuali di accesso alle risorse HPC (ISCRA) in cui una parte eccessiva è rappresentata dalla parte tecnica in cui si deve ogni volta dimostrare la scalabilità degli applicativi e che si potrebbe evitare facilmente se si sviluppasse un pool di applicazioni "certificate" su cui sia possibile stimare facilmente il consumo a progetto, consentendo di snellire le richieste di *grant* e di concentrarsi sulla parte scientifica.

In sintesi, sembra sia venuta meno quella interazione feconda tra il singolo ricercatore, che può fornire idee, intuizioni originali ed applicazioni innovative, e la grande struttura nazionale di calcolo, di recente sempre più orientata a fornire una soluzione tecnologica unica e rigida. Da ciò la recente situazione paradossale di crescente disaffezione della nostra comunità (come si vede dal grafico nell' ALLEGATO relativo alla recente indagine condotta dalla Divisione CTC sino ad oggi per quasi un 40% utenti CINECA e per un quasi un 30% utenti CASPUR o CILEA), specialmente preoccupante nella prospettiva di un futuro passaggio a sistemi "ibridi", sempre più limitanti per i codici ad esempio quanto-meccanici utilizzati da molti chimici computazionali.

Per fare un confronto concreto con la situazione europea: il ricercatore tedesco che necessita di usufruire di metodi HPC può interagire in modo immediato con il Centro di Jülich, collaborando allo sviluppo di nuove soluzioni algoritmiche, e disponendo di assistenza tecnica e varie architetture hardware, fino eventualmente all'accesso al sistema PRACE per applicazioni estese. Il ricercatore italiano ha perso negli ultimi anni, avendone disposto in modo significativo fino a pochi anni fa, queste soluzioni intermedie.

## **Conclusioni**

In conclusione, gli scriventi propongono la costituzione di una interfaccia tra la comunità CTC italiana e i resource provider (tra i quali in primis il CINECA ma anche i gestori di altre risorse quali enti, università, etc.) che abbia voce in capitolo nelle scelte decisionali sullo sviluppo del sistema HPC e HTC (High Throughput Computing) nazionale per poter promuovere le esigenze e le istanze di tale comunità.

In questo spirito, ad esempio, si propone:

- 1) la costituzione di laboratori di calcolo regionali dotati di piattaforme di tipo Tier1 con forte caratterizzazione tematica (eventualmente coordinati dal CINECA) finalizzati all'affiancamento ai sistemi hardware attuali di altre architetture tecnologiche, mirate alle necessità specifiche di alcuni settori scientifici;
- 2) lo sviluppo di strumenti di *grid* e *cloud computing*, come suggerito da una parte della nostra comunità al fine di meglio indirizzare in un contesto maggiormente collaborativo le *facility* di calcolo locali nazionali e internazionali di cui negli ultimi anni vari dipartimenti universitari, centri di ricerca, singoli laboratori si sono dovuti dotare, proprio per rispondere alla crescente domanda di interoperabilità su piattaforme eterogenee finalizzate a meglio rispondere alla domanda diversificata delle comunità degli utenti;
- 3) il potenziamento delle competenze tecniche, soprattutto mediante l'inserimento in organico di personale tecnico specializzato nell'interfaccia con il mondo della ricerca operanti a livello nazionale ma collocati anche presso i siti di cui al punto 1;
- 4) una riforma della politica di distribuzione delle risorse computazionali nazionali che porti all'integrazione delle expertise dei laboratori e centri di ricerca universitari e regionali per la creazione di un vero centro di ricerca nazionale HPC interdisciplinare, analogo ai grandi centri di ricerca europei ed internazionali.

ALLEGATO (Sorgente: indagine della Divisione CTC della SCI – 2012)

APPARTENENZA DELLE RISORSE COMPUTAZIONALI USATE DAI MEMBRI DELLA DIVISIONE CTC

