

# HPC e BigData: una nuova digitalizzazione per il New Green Deal

L'Europa individua nelle ICT un ruolo chiave di *enabling technology* per raggiungere gli obiettivi del Green Deal per il clima e l'ambiente. Calcolo ad alte prestazioni, Intelligenza artificiale, 5G, cloud, IoT, possono contribuire ad accelerarne la realizzazione. L'ENEA con la Divisione ICT è attiva nella ricerca in questo settore, come evidenziano le esperienze e i risultati ottenuti in progetti europei e nazionali.

DOI 10.12910/EAI2020-046

di **Andrea Quintiliani, Massimo Celino, Marta Chinnici, Francesco Iannone, Angelo Mariano, Giovanni Bracco, Silvio Migliori**, Dipartimento Tecnologie Energetiche e Fonti Rinnovabili, Divisione ICT, ENEA (\*)

L'Europa pone al centro delle proprie strategie di sviluppo il nuovo Green Deal, inteso come programma per rispondere alla sfida di disaccoppiare la crescita economica e l'uso delle risorse, e promuovere la transizione energetica verso le fonti rinnovabili. Anche le tecnologie digitali sono individuate come uno dei fattori prioritari per il raggiungimento di questo obiettivo e, contemporaneamente, per il rilancio dell'economia. **La recente comunicazione della Commissione riconosce alle ICT un ruolo di *enabling technology* per il raggiungimento degli obiettivi del Green Deal europeo. Intelligenza artificiale, 5G, cloud computing, Internet of Things, sono fra le tecnologie che possono contribuire ad accelerare le politiche legate al cambiamento climatico e alla protezione dell'ambiente. In questo contesto giocano un ruolo di primo piano anche le grandi infrastrutture di calcolo ad alte prestazioni (HPC). In particolare, viene messa in primo piano la combinazione fra infrastrutture digitali, quali supercomputer e sistemi cloud, reti ad altissime prestazioni con sistemi per l'elaborazione e l'interoperabilità dei dati e algoritmi di intelligenza artificiale,**

e come questo insieme di tecnologie possa consentire l'elaborazione di modelli ad alta precisione, l'analisi di sistemi di dati eterogenei, la regolazione ed il management di sistemi complessi.

Per quanto concerne il supercalcolo, il programma di sviluppo mondiale è molto intenso. Si prevede un incremento di potenza di un ordine di grandezza nel giro di un paio d'anni. La prossima generazione di supercomputer, i cosiddetti computer *exascale*, potranno eseguire un miliardo di miliardi di operazioni al secondo ( $10^{18}$ , pari a 1 exaflops) già a partire dal prossimo anno. È notizia recente che il sistema giapponese Fugaku ha raggiunto prestazioni pari a 0,4 exaflops.

In questa fase di forte crescita e di sviluppo di tecnologie di nuova generazione, la Commissione Europea ha deciso di investire in modo molto massiccio, per mettere a disposizione di ricercatori, imprese e organismi pubblici europei l'accesso a risorse computazionali del più alto livello. Oggi l'Europa dipende fortemente nel settore ICT da tecnologie e strumentazione prodotta all'estero. **Nel nuovo Programma Quadro (Horizon Europe) essa si candida a recitare un ruolo di primo piano nello sviluppo di una nuova generazione di com-**

**puter, nuova in termini di componenti ed architetture, ma anche per il ridotto impatto ambientale.**

Il notevole miglioramento dell'hardware comporta un ripensamento di molti aspetti del ciclo di vita di una infrastruttura HPC, che comprende la riscrittura del software, l'utilizzo di nuovi algoritmi, l'adozione di nuove architetture del software (container) e lo sviluppo di nuovi modelli di programmazione che permettono di rilasciare software in maniera più veloce, più economica e con una scalabilità senza precedenti. Molti, nell'ecosistema HPC, sostengono un cambiamento rivoluzionario, in contrapposizione a quello evolutivo, in cui gli approcci tradizionali sono sostituiti da pratiche più moderne di progettazione coordinata (co-design). **La richiesta di servizi HPC è in continuo aumento, poiché accanto alla tradizionale area della simulazione numerica è sopraggiunta la nuova richiesta di analizzare velocemente le quantità enormi di dati resi disponibili dalle reti di comunicazione.** Questa nuova area per l'ecosistema HPC è l'analisi dei dati ad alte prestazioni o High Performance Data Analytics (HPDA), che descrive la convergenza tra le soluzioni HPC orientate all'analisi massiva di dati

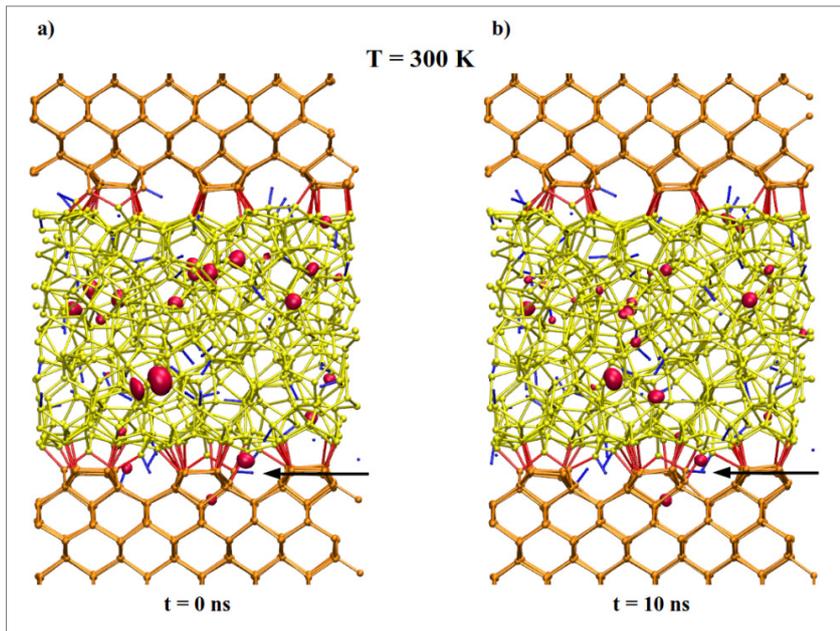


Fig. 1 Studio di un difetto all'interfaccia silicio cristallino - silicio amorfo idrogenato per celle fotovoltaiche ad eterogiunzione [1]

e il mercato dell'analisi commerciale di fascia alta. HPDA rappresenta, dal punto di vista infrastrutturale, la convergenza fra i tradizionali metodi di modellazione e simulazione utilizzati nelle applicazioni HPC e i più recenti metodi basati su tecniche di intelligenza artificiale e deep learning che sono sempre più utilizzati per l'analisi massiva e per l'estrazione di valore dai dati (BigData).

### Calcolo ad alte prestazioni

L'ENEA ha giocato da sempre un ruolo importante nel calcolo avanzato e ha raggiunto uno status rilevante a livello nazionale già a partire da fine anni 2000 con il progetto CRESCO, cofinanziato dal Ministero della Università e della Ricerca nel quadro dei Fondi FESR. Negli anni successivi, l'aggiudicazione di altri progetti e una feconda collaborazione con CINECA ha permesso di potenziare ripetutamente le risorse di calcolo dislocate nella sede di Portici, valorizzando l'investimento complessivo e consolidando un'attività basata sulla disponibilità di sistemi all'avvan-

guardia e la creazione di un importante gruppo di competenza.

Attualmente, il maggior sistema in produzione è il cluster CRESCO6, un sistema allo stato dell'arte dotato di più di 400 nodi con complessivamente più di 20.000 core. CRESCO6 si è posizionato al 420esimo posto nella graduatoria mondiale al novembre 2018, e rimane il secondo sistema italiano nel mondo della ricerca pubblica, dopo quelli di CINECA.

Nel più ampio contesto della Joint Undertaking EuroHPC, creata in ambito europeo per sostenere un ecosistema europeo integrato che copra tutti i segmenti della catena del valore (hardware, software, applicazioni, servizi, interconnessioni e competenze), hanno un ruolo cruciale le applicazioni scientifiche dell'HPC per affrontare le prossime sfide industriali, scientifiche e sociali. Su questo tema la Comunità Europea ha costituito dieci Centri di Eccellenza (CoE) europei fra i quali è presente EoCoE (The European Energy oriented Centre of Excellence in computing applications, [www.eocoe.eu](http://www.eocoe.eu)) rifinanziato

a inizio 2019 per un secondo triennio. L'ENEA è partner del progetto EoCoE, che mira allo sviluppo di un numero selezionato di codici di calcolo relativi a cinque Scientific Challenge di interesse nel settore energia (nuovi materiali per il fotovoltaico e le batterie, meteorologia, eolico, idrologia e fusione) per portarli a performance in grado di sfruttare le infrastrutture *exascale* di prossima realizzazione<sup>1</sup>. Ad esempio, nel settore fotovoltaico è possibile calcolare proprietà elettroniche ed ottiche partendo da modelli atomistici per comprendere l'origine microscopica di comportamenti fisico-chimici macroscopici (Figura 1).

Altri risultati interessanti sono stati ottenuti nel campo dell'energia eolica, dove sono stati realizzati modelli completi di rotori di pale eoliche (Figura 2) e nei sistemi di modellazione idrologica ad alta risoluzione (3 km di risoluzione) su scala continentale.

**Il tema della sostenibilità è diventato di primaria importanza nelle ICT, un settore in cui i consumi energetici crescono ad un tasso del 9% annuo.** Ad esempio, la crescita nelle prestazioni dei grandi datacentre si accompagna ad una crescita dei consumi energetici, con potenze che nelle più grandi infrastrutture possono superare i 10 MWatt. Per questo motivo, per i supercomputer *exascale* ad obiettivi prestazionali sono associati obiettivi di efficienza energetica altrettanto stringenti. **L'organizzazione**

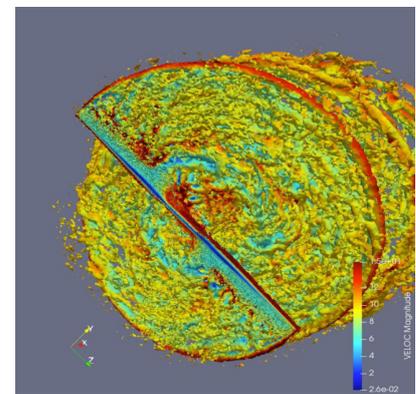


Fig. 2 Modello di mesh rotante che incorpora la rotazione delle pale [2]

di un centro di supercalcolo energeticamente efficiente esige attento studio, capacità di ottimizzazione delle risorse e investimenti importanti sul fronte tecnologico.

In questi anni, la Divisione ICT si è occupata dello sviluppo di strumenti quali metriche e Key Performance Indicator per valutare i consumi energetici e termici di un centro di calcolo, ed allo stesso tempo per incrementarne l'efficienza senza alterare le prestazioni. Ad esempio, ha sviluppato nuove metriche di produttività che mettano in relazione gli effettivi carichi di lavoro al consumo energetico, testandole nel proprio centro di calcolo di Portici, e nuove metodologie per l'analisi dei dati energetici e termici applicando tecniche di *machine learning* e di Intelligenza Artificiale.

Piattaforme per la gestione dati

Sul versante dei dati, l'ENEA è partner di EERAdata ([www.eeraddata.eu](http://www.eeraddata.eu)), progetto europeo che si propone di sviluppare

una infrastruttura FAIR europea che consenta una gestione trasparente e integrata dei dati energetici, assicurandone qualità e disponibilità, per poter rendere possibile ideare, attuare e monitorare nuovi percorsi di transizione sostenibili e offrire opportunità oggi impensabili di innovazione (Figura 3).

Il progetto nasce dalla constatazione che la maggior parte dei database disponibili in Europa non sono stati pensati per la distribuzione open dei dati tramite sistemi informativi avanzati con la conseguente perdita di innumerevoli opportunità tecnologiche ed economiche derivanti dalle nuove soluzioni ICT. Tuttavia, il crescente grado di complessità e incertezza dovuto all'integrazione di diversi tipi di flussi di dati e alle numerose interconnessioni richiede nuove metodologie per i dati e le relative infrastrutture. È necessario individuare, insieme agli attori principali del settore energetico, metodologie e infrastrutture affinché tutti i database aderiscano ai principi dei dati aperti e FAIR (reperibilità, accessibilità, intero-

perabilità, riutilizzabilità) e diventino il motore di nuovi percorsi di innovazione. L'infrastruttura FAIR sarà ospitata e gestita dall'ENEA e si candida a diventare in EOSC (The European Open Science Cloud), il portale di ingresso ai dati europei nel settore delle energie rinnovabili.

La piattaforma si collegherà alle piattaforme esistenti, fornendo un accesso unificato e senza soluzione di continuità ai relativi archivi di dati energetici. Essa pertanto non duplicherà le risorse, ma stabilirà un'interfaccia tecnica per accedere a una varietà di risorse utilizzando formati web standard.

In ambito energetico-ambientale sono di particolare interesse le piattaforme per la gestione di BigData provenienti da infrastrutture smart. In un ambiente urbano intelligente, l'esplosiva crescita del volume, della velocità e della varietà di dati prodotti ogni giorno richiede un aumento continuo delle velocità di elaborazione di server e intere infrastrutture di rete, piattaforme e nuovi modelli di gestione delle risorse.

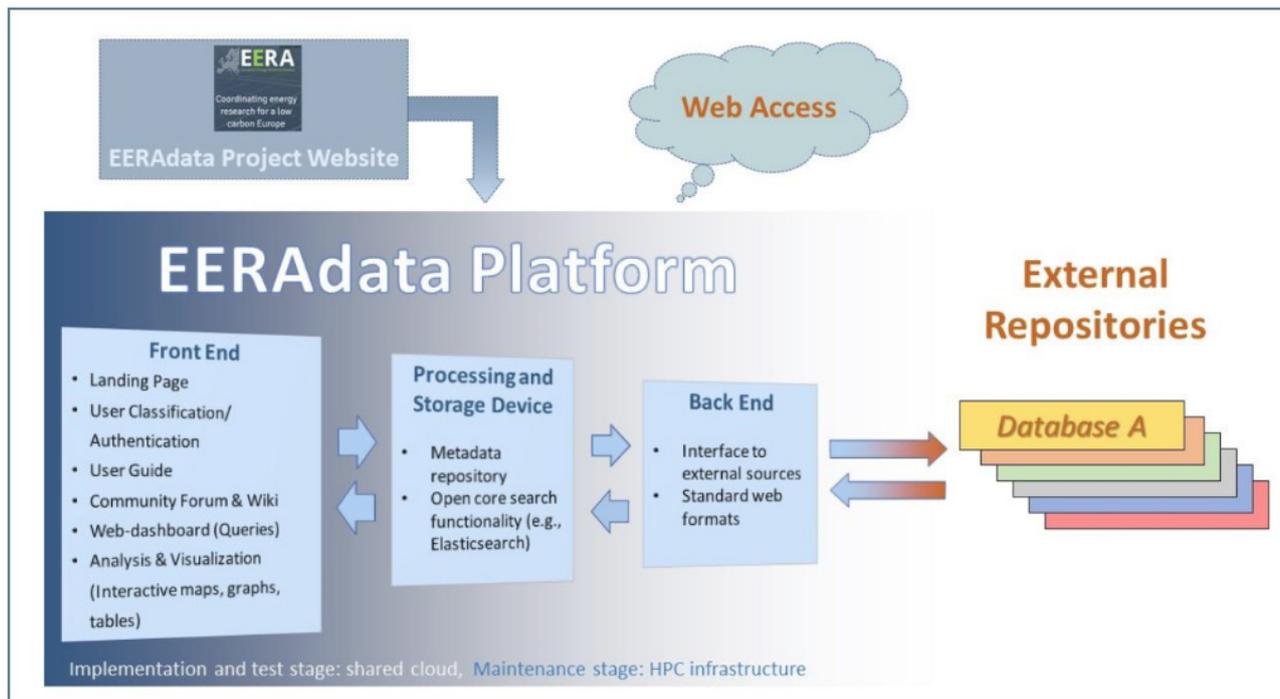


Fig. 3 Funzionalità della piattaforma EERAdata, punto di ingresso ai dati nel settore energia per la ricerca in Europa [4]

La vasta varietà delle fonti complica la gestione dei dati a causa del loro differente formato, soprattutto quando si vuole archivarli, trasformarli o scambiarli. L'utilizzo di infrastrutture e piattaforme cloud è diventata una necessità ed allo stesso tempo rappresenta una criticità, poiché le prestazioni complessive delle applicazioni, soprattutto quelle urbane, dipendono fortemente dalle proprietà del servizio di gestione dei dati.

**In quest'ottica l'ENEA, attraverso le Divisioni ICT e SEN, ha sviluppato una piattaforma urbana (Smart City Platform – SCP) che ha la capacità di:**

- rendere smart i dati e le informazioni, per le imprese, le persone ed i decisori politici riprogettando le relazioni

tra i vari attori;

- garantire sinergie e interoperabilità all'interno e tra i diversi sistemi e domini delle politiche cittadine (ad es. trasporti, energia, istruzione, sanità ecc.);
  - promuovere l'innovazione, ad esempio attraverso i cosiddetti open data, i laboratori viventi e gli hub tecnologici.
- La piattaforma è di tipo modulare, multi-layer e scalabile, in grado di affrontare il problema dell'interoperabilità tra gli ampi e differenti domini nel contesto delle applicazioni urbane. ENEA-SCP è implementata seguendo il paradigma Software-as-a-Service (SaaS), e sfrutta le risorse di calcolo dell'infrastruttura ENEA in ambiente cloud ospitata presso il Centro di Ricerca ENEA di Portici. Gli esempi citati non esauriscono il pa-

norama delle iniziative ENEA di impiego delle tecnologie di punta delle ICT nel campo delle energie rinnovabili e nella mitigazione del cambiamento climatico. Il consolidamento nel tempo di un'infrastruttura di calcolo di livello nazionale ha contribuito allo sviluppo di competenze specialistiche in grado di partecipare con un ruolo di primo piano ai prossimi programmi di Ricerca e Sviluppo nazionali ed europei.

(\*) *Massimo Celino, Marta Chinnici, Angelo Mariano, Silvio Migliori, Divisione ICT; Francesco Iannone, Laboratorio Calcolo ad alte prestazioni, Andrea Quintiliani, Giovanni Bracco, Collaboratori esterni*

<sup>1</sup> ENEA partecipa anche al progetto FocusCoE che ha l'obiettivo di coordinare le azioni dei Centri di Eccellenza costituendo, fra l'altro, un apposito organismo associativo denominato HPC3

## BIBLIOGRAFIA

1. Ab initio study on localization and finite size effects in the structural, electronic, and optical properties of hydrogenated amorphous silicon. P Czaja, M Celino, S Giusepponi, M Gusso, U Aeberhard. *Computational Materials Science* 155 (2018) 159–168
2. Herbert Owen, Barcelona Supercomputing Center, *EoCoE Newsletter* n. 2, 2019
3. A 3 km spatially and temporally consistent European daily soil moisture reanalysis from 2000 to 2015. Bibi S. Naz, et al. *Scientific Data* 7:111, 2020
4. Open science in Energy system research: current practices and future directions. A. Wierling, W.J. Schwanitz, M. Paier, M. Celino. Open Science Conference, Berlin, March 2020
5. IoT Meets Opportunities and Challenges: Edge Computing in Deep Urban Environment. M. Chinnici, S. De Vito. In “Dependable IoT for Human and Industry”. River Publishers Series in Information Science and Technology, 2018
6. Smart City Platform Specification: A Modular Approach to Achieve Interoperability in Smart Cities. A. Brutti, P. De Sabbata, et al., in *The Internet of Things for Smart Urban Ecosystems* pp 25-50, Springer, 2018
7. The role of medium size facilities in the HPC ecosystem: the case of the new CRESCO4 cluster integrated in the ENEAGRID infrastructure. G. Ponti et al., *Proceedings of the 2014 International Conference on High Performance Computing and Simulation*