

Interviste Newsletter

PNRR: NASCE IL CENTRO NAZIONALE DI SUPERCALCOLO

Intervista ad Antonio Zoccoli, presidente dell'INFN, neo-presidente della Fondazione ICSC Centro Nazionale di Ricerca in High Performance Computing, Big Data and Quantum Computing



Si sono insediati il 19 luglio gli organi direttivi della Fondazione ICSC* che gestirà uno dei cinque Centri Nazionali previsti dal PNRR Piano Nazionale di Ripresa e Resilienza. Nasce così il Centro Nazionale di Ricerca in High Performance Computing, Big Data e Quantum Computing che, proposto dall'INFN, conta 51 membri fondatori distribuiti su tutto il territorio nazionale, provenienti dai settori pubblico e privato, dal mondo della ricerca scientifica e dell'industria. Il nuovo Centro

Nazionale farà base al Tecnopolo di Bologna, una cittadella dell'innovazione promossa dalla Regione Emilia-Romagna, anche grazie a investimenti del Governo Italiano e della Comunità Europea, che già ospita il Data Center del Centro Meteo Europeo (ECMWF) e a breve accoglierà il supercalcolatore Leonardo gestito da CINECA, e il Centro di Calcolo dell'INFN. Il progetto metterà in rete e a sistema le specifiche conoscenze, competenze e risorse di realtà che operano in tutta Italia in molteplici ambiti, con l'obiettivo di costruire un'infrastruttura distribuita e trasversale che supporti la ricerca scientifica e il mondo produttivo nell'innovazione e digitalizzazione del Paese.

Per portare a compimento la sua missione, ICSC conterà su un finanziamento, su fondi Next Generation EU nell'ambito della Missione Istruzione e Ricerca del PNRR coordinata dal MUR Ministero dell'Università e della Ricerca, pari a euro 319.938.979,26, di cui il 41% sarà investito al Sud. In particolare, del finanziamento complessivo, oltre 100 milioni di euro saranno dedicati al personale, un investimento che viene considerato prioritario, con una partecipazione femminile di almeno il 40%, e con quasi 16 milioni di euro riservati a borse di dottorato e quindi all'alta formazione e alla carriera dei giovani. ICSC, coerentemente con gli obiettivi strategici del PNRR, realizzerà i suoi obiettivi specifici promuovendo le carriere dei giovani e iniziative per il superamento del divario di genere nelle carriere professionali e tra il Nord e il Sud del Paese. Nel corso della prima assemblea plenaria, che è stata ospitata nella sede della Regione Emilia Romagna lo scorso 19 luglio, i membri fondatori di ICSC hanno eletto Antonio Zoccoli, presidente dell'INFN, alla presidenza della Fondazione, e i dieci membri del Consiglio di Amministrazione**. Abbiamo chiesto ad Antonio Zoccoli, neo-eletto presidente della Fondazione ICSC, di raccontare il progetto.

L'INFN si è fatto promotore della proposta del centro nazionale di supercalcolo. La proposta è stata valutata e quindi selezionata dal MUR per essere realizzata nell'ambito del PNRR. Come si è arrivati a questo esito positivo?

Gli esperimenti di big science, da quelli di fisica delle alte energie a quelli di fisica astroparticellare, sono caratterizzati dalla produzione di big data, cioè grandi moli di dati che devono essere archiviate, condivise, analizzate. Lavorare nei principali progetti di big science a livello mondiale ha quindi consentito all'INFN di maturare nel tempo una lunga e consolidata esperienza dal punto di vista della gestione dei dati e delle infrastrutture necessarie a farlo. Porto un esempio su tanti, all'inizio degli anni 2000, la comunità della fisica delle particelle impegnata nei grandi esperimenti all'acceleratore LHC del CERN, che producono enormi quantità di dati, ha creato un'infrastruttura globale, chiamata Grid, che permetteva ai ricercatori di sfruttare al meglio tutte le risorse di calcolo disponibili. La Grid, nata dunque nell'ambito della ricerca in fisica fondamentale, è stata il precursore della cloud commerciale che oggi tutti noi abbiamo a disposizione sui nostri telefonini. L'INFN è l'unica istituzione italiana che può vantare una simile esperienza nella gestione dei big data e, grazie alle sue competenze e a quelle degli altri cinquanta partner del progetto in diversi altri settori, è stato possibile mettere a punto una proposta ben strutturata, solida e concreta, che è stata valutata positivamente e quindi approvata dal MUR, sotto la cui egida sono nati i cinque centri nazionali previsti dal PNRR.

Iniziamo spiegando che cosa significa high performance computing, big data e quantum computing.

L'high performance computing (HPC) è una tecnologia che permette di effettuare calcoli ad alte prestazioni, ovvero di fare calcoli paralleli che consentono di velocizzare e potenziare le operazioni. Prendiamo per esempio il caso in cui dobbiamo studiare il comportamento dell'atmosfera. Quello che facciamo è dividere l'atmosfera in tanti piccoli volumetti per capire come evolvono i vari parametri, come la temperatura o l'umidità, in ognuno di essi. Poiché tra questi volumetti vengono scambiate continuamente informazioni, per esempio la temperatura del volumetto A influenza la temperatura del volumetto B, per studiarne l'evoluzione, la temperatura di ogni singolo volumetto deve essere messa in correlazione con quella degli altri: è necessario cioè che il calcolo vada in parallelo, cosicché le informazioni vengano scambiate in contemporanea. Quindi, il calcolo parallelo implica avere tante CPU, tanti processori, che eseguono calcoli simultaneamente e si scambiano informazioni in tempo reale. Il calcolo parallelo è utile per risolvere determinati tipi di problemi, per esempio in fluidodinamica, appunto, per lo studio dell'atmosfera o del clima, ma anche per fare il training degli algoritmi di una disciplina emergente, l'Intelligenza Artificiale.

Mentre, non è necessario se devo lavorare una grande sequenza di dati, come quelli prodotti negli esperimenti a LHC, dove si analizza evento per evento. In questo caso utilizzo il cosiddetto l'high-throughput computing (HTC). L'HTC è la tecnologia richiesta per l'analisi di grandi moli di dati, i cosiddetti big data, cioè quantità di dati superiori al petabyte (milione di gigabyte). HPC e HTC sono, quindi, due tecnologie di calcolo differenti che servono a risolvere problemi diversi, e le migliori competenze nazionali nell'HPC e nell'HTC sono presenti in questo progetto, con il CINECA e l'INFN che sono le due istituzioni leader in Italia in questi due settori.

Infine, c'è una possibile tecnologia di calcolo futura, basata su un concetto nuovo: la presenza contemporanea di stati diversi. È il calcolo quantistico. Nel caso di un computer tradizionale, un bit può essere zero oppure uno, mentre in un computer quantistico un bit può essere zero, oppure uno, oppure zero e uno

contemporaneamente: è il qubit, il bit quantistico. Questo consente di avere più informazione. Macchine che si basano su bit quantistici permettono di avere grande potenza con pochi qubit. È una tecnologia molto promettente, tuttavia non è ancora sufficientemente matura per fornire un servizio. La difficoltà risiede nel mantenere la cosiddetta coerenza quantistica, nel gestire gli stati quantici, stati microscopici, a livello atomico, molto sensibili alle interferenze esterne. È necessaria una stabilità del sistema, per questo i computer quantistici tipicamente sono macchine criogeniche. È una tecnologia completamente diversa da quella dei computer ordinari. La sfida sarà dunque costruire un computer quantistico stabile, potente, con un numero di qubit sufficientemente elevato, e poi creare il software in grado di sfruttarlo al meglio.

In quali settori e come opererà il Centro di supercalcolo?

Il nuovo Centro aggregnerà le comunità scientifiche italiane di eccellenza in dieci diversi ambiti, sarà strutturato su due colonne portanti di eguale rilevanza, le infrastrutture e le aree tematiche, e sarà organizzato secondo il modello Hub e Spoke. L'Hub avrà la responsabilità di validare e gestire i programmi di ricerca, le cui attività verranno condotte dagli Spoke e dalle realtà a loro affiliate, anche attraverso bandi aperti a istituzioni di ricerca e aziende esterne al Centro.

Gli Spoke saranno dieci e cureranno altrettante aree tematiche: Future HPC & Big Data, Fundamental Research & Space Economy, Astrophysics & Cosmos Observations, Earth & Climate, Environment & Natural Disaster, Multiscale Modeling & Engineering Applications, Materials & Molecular Sciences, In-Silico Medicine & Omics Data, Digital Society & Smart Cities, Quantum Computing. Gli Spoke Future HPC & Big Data e Quantum Computing saranno di carattere tecnologico e avranno come obiettivo di frontiera lo sviluppo di chip e microchip avanzati e di tecnologie emergenti come quelle per il calcolo quantistico.

Inoltre, il centro potrà supportare gli altri quattro Centri Nazionali finanziati nell'ambito del PNRR. Oltre al National Centre for HPC, big data and quantum computing, anche National Research Centre for Agricultural Technologies, il Sustainable Mobility Center, il National Biodiversity Future Center e il National Center for Gene Therapy and Drugs based on RNA Technology. Si prevede che i cinque centri lavorino laddove possibile in modo sinergico. Il nostro Centro potrebbe ospitare i dati degli altri Centri Nazionali. Non è previsto da progetto ma può essere previsto sulla base di accordi specifici, come per esempio già avviene per il Centro Nazionale dedicato alla medicina che si appoggerà a noi per la gestione dei suoi dati. D'altra parte, anche gli altri Centri Nazionali trovano corrispondenza con alcune nostre aree tematiche.

Quale sarà il ruolo dell'INFN nell'ambito del progetto?

L'INFN ha promosso e coordinato la proposta di progetto. Dopo la sua approvazione, è stata costituita la Fondazione ICSC che avrà il compito di attuarlo e di cui l'INFN è membro assieme ad altre cinquanta realtà provenienti sia dal mondo della ricerca scientifica sia dal mondo imprenditoriale. Nel corso della prima riunione dei soci sono stati nominati il presidente e il CdA, che saranno garanti che il progetto sia realizzato secondo il disegno proposto. L'INFN, avendo espresso il presidente della Fondazione, avrà in un certo senso un ruolo di guida, ma sarà anche esecutore, perché ha incarichi in alcune delle aree tematiche in cui esso si svilupperà. In particolare, l'INFN contribuirà all'area infrastrutturale, nell'ambito della quale saranno potenziate le risorse dedicate ai big data, e all'interfaccia cloud con cui sarà gestita tutta l'infrastruttura. Inoltre, lavorerà su alcune aree tematiche: quella della ricerca fondamentale, ambito dove svolge la propria missione, ma anche quella di

medicina, settore dove l'INFN ha competenze nella gestione delle immagini e nello sviluppo di modelli e algoritmi per l'analisi dei dati medici, e quella del quantum computing perché, se tra dieci anni la prossima macchina di calcolo sarà una macchina ibrida, quindi anche quantistica, è importante aver acquisito tutte le competenze necessarie per sfruttarla al meglio.

Quali saranno i primi obiettivi del Centro?

Muoveremo i primi passi a settembre, quando costruiremo il Centro dal punto di vista sia organizzativo sia amministrativo, con la definizione chiara di tutti i piani di sviluppo e l'allocazione delle risorse. Quindi, procederemo in parallelo con la realizzazione dell'infrastruttura, che richiederà almeno un paio di anni, ma probabilmente proseguirà tutti e tre gli anni della fase di start-up del progetto, e con lo sviluppo delle applicazioni che serviranno a sfruttare l'infrastruttura. Queste saranno sviluppate dalle persone più competenti che ci sono sul territorio nazionale negli specifici ambiti, e potranno essere utilizzate sia nel campo della ricerca sia dal sistema industriale. L'attività di ricerca che sarà svolta dal Centro sarà, quindi, un'attività di sviluppo di software: il Centro non deve fare ricerca scientifica, deve sviluppare l'infrastruttura e le sue applicazioni, grazie alle quali si farà ricerca, costruirà strumenti per la ricerca scientifica.

Quale sarà l'impatto del centro a livello nazionale e internazionale e in quali settori?

Se riusciremo a condurre al successo questo nostro ambizioso progetto, l'Italia sarà il primo Paese dotato di un'infrastruttura di calcolo nazionale di tipo "datalake", che significa una rete di trasmissione dati veloce, l'accesso cloud a questa rete, e infrastrutture su tutto il territorio che, grazie alla cloud, saranno assegnate in maniera dinamica e virtuale: questo consentirà di sfruttare nel modo più efficace le risorse disponibili in tutto il Paese. Gli utenti non utilizzeranno più un computer fisico per conservare i propri dati e una CPU fisica per analizzarli, ma avranno a disposizione macchine virtuali che saranno assegnate in base alla loro disponibilità in modo da massimizzare l'efficienza del processo. Questo significa flessibilità e ottimizzazione delle risorse, che diventano così pressoché illimitate. La realizzazione di questo progetto produrrebbe un unicum a livello internazionale. Riuscire in questa impresa significherebbe mettere a disposizione del Paese un'infrastruttura d'avanguardia in Europa, anzi io credo nel mondo. Saremo i primi a portare a compimento questo esercizio fino in fondo: ad oggi, infatti, molti Paesi dispongono di tanti centri sparsi sul loro territorio, che però non sono coordinati tra loro, anzi sono spesso in competizione. Noi vogliamo invece mettere tutte le risorse nazionali a sistema. In questo contesto è importante anche il ruolo delle aziende private. Nei prossimi anni - e in realtà sta già accadendo - tutte le discipline scientifiche produrranno una quantità di dati senza precedenti. Dove risiederanno questi dati? Come saranno analizzati? E come verranno utilizzati? Questo è un tema caldo per la ricerca scientifica, ma è di grande interesse anche per l'industria: dal monitoraggio degli impianti allo sviluppo di nuovi materiali, dall'impatto del clima sull'agricoltura di precisione ai trasporti. È quindi fondamentale che l'industria sviluppi le competenze per usare in modo efficiente ed efficace le infrastrutture di calcolo. Da qui il forte coinvolgimento delle imprese.

Il progetto è quindi ambizioso e potrebbe far diventare l'Italia un polo di attrazione e un punto di riferimento in Europa.

Assolutamente. Tant'è che l'Europa sta investendo. La potente macchina pre-exascale che il CINECA porterà nel Centro è sostenuta con finanziamenti non solo italiani ma anche europei. E anche la macchina quantistica di cui doteremo il Centro Nazionale sarà cofinanziata con fondi europei.

Inoltre, saremo attrattivi anche per il mondo della ricerca e per quello imprenditoriale. Sono, infatti, previsti 32 milioni di euro di *open call*, cioè bandi per il finanziamento di progetti in cui potranno essere coinvolti soggetti che non fanno già parte del Centro, sia pubblici sia privati. Mentre, altri 32 milioni saranno a disposizione per i soggetti del Centro con l'obiettivo di promuovere l'innovazione, aumentando il TRL (Technology Readiness Level, ossia il grado di maturità tecnologica) di nuovi prodotti per finalizzare il processo, in modo da far compiere loro il cosiddetto ultimo miglio fino al loro ingresso sul mercato.

Veniamo al personale che opererà il Centro, come sarà reclutato? Quali saranno le possibilità per le carriere dei giovani e delle donne, che il PNRR vuole sostenere?

Il Centro conterà su personale senior proveniente dai partner della Fondazione: possiamo contare su oltre 1500 professionisti di alta competenza che dedicheranno parte del loro tempo lavorativo al progetto. Ma punteremo sui giovani, il loro coinvolgimento è una priorità: sia le Università sia gli Enti di Ricerca coinvolti nel Centro bandiranno borse di dottorato (200 in totale) e borse post-doc (altre 200) per diversi profili. E almeno il 40% delle persone che lavoreranno nel progetto saranno donne: stiamo studiando iniziative da attuare per incentivare la loro partecipazione, in modo da raggiungere gli obiettivi del PNRR. Tra questi obiettivi, vi sono anche gli investimenti al Sud, per un valore pari al 41% dei fondi: abbiamo nel nostro programma il potenziamento della rete già esistente e la costruzione di nuovi centri. Sarà per esempio potenziato il centro della fondazione CMCC Centro Euro-Mediterraneo sui Cambiamenti Climatici a Lecce, e saranno costruiti nuovi centri a Frascati, al Gran Sasso, a Napoli dove sarà realizzato un centro di calcolo parallelo gestito dal CINECA. Ma sarà anche potenziato il centro di calcolo del CINECA e il Tier1 dell'INFN entrambi a Bologna. Inoltre, sarà potenziata tutta la rete del GARR, la rete italiana di trasmissione dati per la ricerca, che attualmente funziona tra i 10 e i 100 Gigabit per secondo. È perciò già una rete molto veloce, ma quello che noi vorremmo fare, è renderla ancora più veloce e performante: vogliamo portarla a multipli del Terabit per secondo, quindi velocizzarla di un fattore 100 rispetto ad ora, o anche 1000: questo permetterebbe di realizzare davvero delle performance straordinarie. Questi sono alcuni degli obiettivi del nostro ambizioso ma assolutamente realistico progetto.

** La fondazione ICSC: INFN Istituto Nazionale di Fisica Nucleare, CINECA, GARR, CNR Consiglio Nazionale delle Ricerche, INAF Istituto Nazionale di Astrofisica, INGV Istituto Nazionale di Geofisica e Vulcanologia, IIT Istituto Italiano di Tecnologia, CMCC Centro Euro-Mediterraneo sui Cambiamenti Climatici, FBK Fondazione Bruno Kessler, ENEA Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile, CRS4 Centro di Ricerca e Sviluppo e Studi Superiori in Sardegna, OGS Istituto Nazionale di Oceanografia e di Geofisica Sperimentale, Università di Bologna, Università di Ferrara, Università di Bari, Università di Milano Bicocca, Sapienza Università di Roma, Università degli Studi di Roma Tor Vergata, Università di Trieste, Università di Padova, Università di Pavia, Università di Trento, Università di Torino, Università dell'Aquila, Università Federico II di Napoli, Università di Pisa, Università di Firenze, Università di Catania, Università della Calabria, Università del Salento, Università di Modena e Reggio Emilia, Università di Parma, Politecnico di Bari, Politecnico di Milano, Politecnico di Torino, SNS Scuola Normale Superiore, SISSA Scuola Internazionale*

Superiore di Studi Avanzati, Gruppo Autostrade, Engineering Ingegneria Informatica, Eni, Ferrovie dello Stato, Fincantieri, Fondazione Innovazione Urbana, IRCCS Istituto Clinico Humanitas, IFAB International Foundation Big Data and Artificial Intelligence for Human Development, Intesa Sanpaolo, Leonardo, Sogei, Thales Alenia Space Italia, Terna, UnipolSai Assicurazioni.

***Componenti del CdA neo-eletti: Emilio Fortunato Campana del CNR, Paolo Maria Mancarella dell'Università di Pisa, Francesco Scarcello dell'Università della Calabria, Donatella Sciuto del Politecnico di Milano, Simona Tondelli dell'Università di Bologna, Francesco Ubertini del CINECA, Matteo Laterza di UnipolSai, Elisabetta Oliveri di Autostrade per l'Italia, Andrea Quacivi di Sogei, Francesca Zarri di Eni*