

CRESCO EDOC (EDucation On Cloud): cloud computing a servizio della formazione e della didattica per il calcolo scientifico avanzato

Angelo Mariano¹

A. Funel, D. Abate, F. Ambrosino, G. Aprea, T. Bastianelli, F. Beone, R. Bertini, G. Bracco, M. Caporicci, B. Calosso, M. Chinnici, R. Ciavarella, A. Colavincenzo, A. Cucurullo, P. D'Angelo, M. De Rosa, P. De Michele, G. Furini, D. Giammattei, S. Giusepponi, R. Guadagni, G. Guarnieri, S. Magagnino, G. Mencuccini, C. Mercuri, S. Migliori, P. Ornelli, F. Palombi, S. Pecoraro, A. Perozziello, S. Pierattini, S. Podda, F. Poggi, G. Ponti, A. Quintiliani, G. Santomauro, A. Scalise, F. Simoni

¹ENEA Centro Ricerche Brindisi
SS 7 Appia km706, 72100 Brindisi
angelo.mariano@enea.it

L'unità di ICT (Information & Communication Technology) di ENEA ha elaborato un'iniziativa volta ad integrare nella propria infrastruttura di calcolo ENEAGRID una piattaforma open source di cloud computing i cui servizi siano esplicitamente dedicati al supporto della formazione per le scuole medie superiori per il calcolo avanzato. Questa iniziativa si inserisce all'interno del progetto PON EDOC@WORK 3.0 [EDOC@WORK].

1. Introduzione

L'ENEA, l'Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile, persegue i suoi obiettivi di ricerca e innovazione tecnologica anche nei settori del calcolo scientifico avanzato. ENEAGRID [Bracco et al, 2012] è l'infrastruttura che integra tutte le risorse di calcolo dell'ENEA. I cluster HPC (High Performance Computing) CRESCO [Migliori et al, 2009] fanno parte di ENEAGRID e sono dedicati al calcolo scientifico ad alte prestazioni. L'iniziativa CRESCO EDOC (EDucation On Cloud) nasce nell'ambito del PON R&C 2007-2013 EDOC@WORK 3.0 (Education and work on Cloud) ed intende integrare nell'infrastruttura ENEAGRID una piattaforma di cloud computing localizzata nel centro di ricerche ENEA di Brindisi, dedicata ad offrire servizi di tipo educational focalizzati in particolar modo sul calcolo scientifico ad alte prestazioni a disposizione delle scuole medie superiori. CRESCO EDOC si propone di favorire l'apprendimento delle tecnologie ICT

consentendo di operare concretamente su un sistema di calcolo avanzato e di promuovere l'acquisizione di nozioni computazionali, con l'ausilio di personale esperto che collaborerà con il corpo insegnante nella fase di sperimentazione e anche oltre. L'utilizzo di una piattaforma cloud consente di specializzare il contenitore che offrirà questi servizi in base alle specifiche richieste delle istituzioni scolastiche e di modellare l'utilizzo delle risorse in base al carico effettivo presente sull'infrastruttura sottostante.

2. ENEAGRID e CRESCO EDOC

ENEAGRID integra in maniera affidabile in un'unica infrastruttura tutte le risorse di calcolo dell'ENEA e fornisce un ambiente di lavoro uniforme rendendo la modalità di utilizzo dell'infrastruttura indipendente dalla localizzazione geografica della risorsa hardware/software richiesta. I siti ENEAGRID sono connessi dalla rete GARR, la rete italiana utilizzata dalle università e dagli enti di ricerca pubblici, come illustrato in fig. 1. ENEAGRID fornisce supporto ad un'ampia gamma di applicazioni in diversi campi di ricerca tra cui energia, scienza dei materiali, ambiente e clima, e rende possibile l'interazione con apparati sperimentali (microscopio TEM, sequenziatore DNA, microscopio ad Elio, nanotomografo) per il controllo da remoto e l'acquisizione dei dati. Esperienza importante sono i laboratori virtuali [Beone et al, 2010], organizzati per aree tematiche ed accessibili da web, che costituiscono uno spazio di lavoro in cui gli utenti possono collaborare condividendo dati e accedendo ad applicazioni relative a uno specifico campo di interesse comune. I sistemi di calcolo più importanti di ENEAGRID sono i cluster HPC CRESCO, con la parte più consistente nel sito di Portici (NA), con ~8K core e una potenza computazionale di ~130 TFlops.

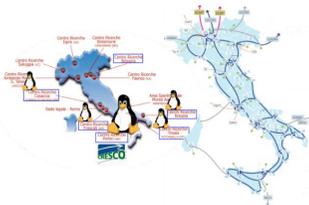


Fig.1 – A sinistra la rete dei centri ENEAGRID, dove con il bordo azzurro sono indicanti i centri di calcolo, a destra la rete GARR che connette i centri ENEA

All'interno del progetto EDOC@WORK 3.0, che si propone di integrare l'uso delle nuove tecnologie IT nel sistema scuola, nelle università e nella formazione professionale, ENEA ha inteso sviluppare la propria iniziativa CRESCO EDOC, convertendo ad infrastruttura cloud le proprie risorse di calcolo collocate nel

centro di ricerche di Brindisi e avviando la sperimentazione con le scuole della Regione Puglia. CRESCO EDOC, avvalendosi dell'esperienza già maturata da qualche anno nell'ambito del cloud computing [Ponti et al, 2012] e nel settore educational [Funel et al, 2013] permette di offrire un sistema cloud educational IaaS (Infrastructure as a Service) per le scuole superiori in ambito HPC tarato sulle specifiche esigenze delle istituzioni scolastiche ed interoperabile con eventuali altri sistemi cloud privati e/o pubblici, perfettamente integrato con l'infrastruttura ENEAGRID.

3. Approfondimento tecnico su CRESCO EDOC

L'infrastruttura di cloud computing CRESCO EDOC, localizzata nel centro di ricerche di Brindisi, si avvale della disponibilità di 6 nodi dotati di processore a 4 socket 4 core Intel Xeon X7350 2.93 Ghz e 64GB di RAM con interconnessioni G Ethernet e InfiniBand, che svolgono la funzione di nodi di tipo compute all'interno del sistema OpenStack. In fig.2a è possibile vedere una schermata di esempio della dashboard di amministrazione di CRESCO EDOC.

Per esigenze educational sono quindi disponibili fino a 96 core totali per la definizione degli ambienti virtuali su cui le scuole potranno testare le funzionalità e le potenzialità dell'infrastruttura ENEAGRID. Previa richiesta di account educational all'Unità Tecnica ICT dell'ENEA [UTICT], è possibile realizzare l'accesso ai sistemi CRESCO EDOC in due modalità: protocollo ssh verso un client di frontend CRESCO EDOC; protocollo NX/FARO [Rocchi et al, 2010] con l'uso dell'applicazione visuale FARO Educational, illustrata in fig.2b.



Fig.2 – (a) Schermata di amministrazione delle risorse CRESCO EDOC; (b) Schermata di FARO Educational per CRESCO EDOC

Tutti gli utenti, inclusi quelli di CRESCO EDOC, condividono lo stesso ambiente di lavoro (middleware di ENEAGRID), quindi compilatori e sistema di gestione dei job basato su LSF (Load Sharing Facility). I file system disponibili sono: AFS (Andrew File System), file system geograficamente distribuito utilizzato da tutti gli utenti di ENEAGRID; GPFS (General Parallel File System) di IBM, ad alte prestazioni, ottimizzato per l'accesso parallelo ai dati.

Durante la sperimentazione prevista per la fine del 2015, gli utenti di CRESCO EDOC potranno accedere a tutta una serie di software open source e richiedere modifiche ai nodi virtuali dedicati sul cloud, in accordo con i docenti, grazie all'utilizzo di strumenti avanzati di supporto ed helpdesk forniti all'utenza.

Si prevede anche di implementare degli strumenti automatici di gestione e controllo del cloud che permettano di aumentare gli ambienti di esecuzione dei codici di calcolo scientifico in corrispondenza di situazioni di criticità.

Tra gli elementi di spicco che caratterizzano e qualificano l'iniziativa di cloud computing CRESCO EDOC possiamo citare quindi: integrazione **seamless** con l'infrastruttura di grid computing ENEAGRID; creazione di ambienti didattici **flessibili** per il calcolo scientifico ad alte prestazioni; utilizzo della tecnologia **ad alta velocità** Infiniband a 20Gb/s per la comunicazione di rete tra i nodi fisici; possibilità di rendere l'infrastruttura **interoperabile** con altre infrastrutture cloud mediante lo standard OCCI (Open Cloud Computing Interface).

3. Conclusioni

CRESCO EDOC, l'iniziativa di ENEA nata per rendere disponibile un sistema cloud privato ad uso delle scuole superiori per il calcolo HPC, si propone di introdurre innovazione nel processo di apprendimento di nozioni computazionali e di ICT avanzato tramite un sistema flessibile ed interoperabile.

Bibliografia e sitografia

[Beone et al, 2010] Beone F. et al., I Laboratori Virtuali di ENEA, Poster Conferenza GARR 2010, Torino 26-28 ottobre 2010

[Bracco et al, 2012] Bracco G. et al, ENEA_GRID Infrastructure, Report – High Performance Computing on CRESCO Infrastructure: Research Activities and Results 2010-2011, July 2012, 9-16.

[EDOC@WORK] www.edocwork.it.

[Funel et al, 2013] CRESCO Educational: una proposta per formare e orientare i ragazzi delle scuole medie superiori all'utilizzo dei sistemi per il calcolo scientifico avanzato e alle tecnologie ICT, Proceedings - 27a DIDAMATICA 2013: Tecnologie e Metodi per la Didattica del Futuro, p. 1073-1077, May 2013.

[Migliori et al, 2009] Migliori S. et al, CRESCO HPC System Integrated into ENEA-GRID Environment, Proc. Of the Final Workshop of the GRID Projects of the Italian National Operational Programme 2000-2006 (Call 1575), February 2009, 151-155.

[Ponti et al, 2012] Ponti G. et al., Cloud Computing in ENEA-GRID: Virtual Machines, Roaming Profile and Online Storage, Pres. Workshop GARR Calcolo e Storage Distribuito, MIUR Roma 29-30 November 2012.

[Rocchi et al, 2010] Rocchi A. et al., FARO: Accesso web a risorse remote per l'industria e la ricerca, Proc. Conferenza GARR 2010, Torino 26-28 ottobre 2010.

[UTICT] www.utict.enea.it.