

Abstract per una Presentazione Orale.

ESPERIMENTI DI CLOUD COMPUTING IN ENEA-GRID

G. Ponti, A. Secco¹, F. Ambrosino, G. Bracco, R. Ciavarella, A. Colavincenzo², P. D'Angelo, M. De Rosa, A. Funel, G. Guarnieri, D. Giammattei, S. Migliori, S. Pecoraro, A. Petricca, S. Pierattini, S. Podda, A. Rocchi, C. Sciò.

ENEA – Centro Ricerche Portici – Piazzale Enrico Fermi 1, Portici (NA)

1) Nice srl Asti, 2) Kelyon srl Napoli

ENEA-GRID [1] è l'infrastruttura distribuita che realizza l'integrazione delle risorse computazionali dedicate al calcolo scientifico in ENEA. Tali risorse sono orientate sia al calcolo parallelo ad alte prestazioni che all'esecuzione massiva di elaborazioni seriali e al momento comprende circa 5000 core per una potenza integrata totale di oltre 40 Tflops, distribuiti in gran parte su architetture Linux x86_64 ed AIX. Tali risorse sono distribuite su 6 siti collegati tra loro dalla rete GARR. Il sito di maggior rilevanza è quello del Centro Ricerche ENEA di Portici (NA) che ospita il sistema di calcolo CRESCO, che al momento mette a disposizione più di 3300 core. ENEA-GRID opera da più di un decennio e la sua architettura è stata sviluppata secondo il paradigma del GRID computing in modo da fornire all'utenza un ambiente omogeneo ed unificato, ma nello stesso tempo capace di incorporare nel tempo sistemi multi-piattaforma di architettura diversa e distribuito sull'intero territorio nazionale, seguendo le esigenze degli utilizzatori e l'evoluzione esplosiva della tecnologia informatica.

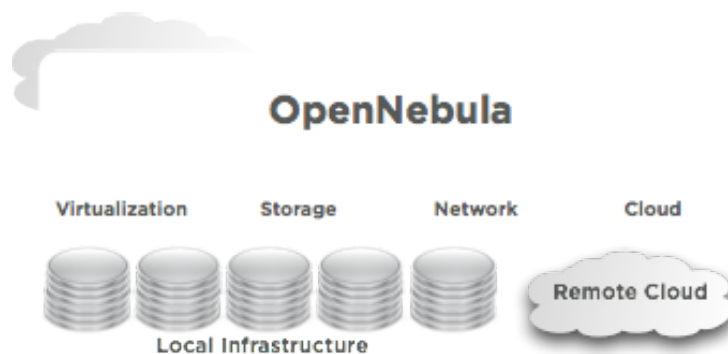
Tale processo è stato reso possibile dall'adozione sin dall'inizio di una architettura basata su componenti software maturi, sia di tipo proprietario, come il gestore di risorse LSF Multicluster che Open Source come il file system distribuito OpenAFS. Tali componenti, insieme a Kerberos 5 e NX integrato nel portale FARO [2], costituiscono il middleware di ENEA-GRID, ovvero lo strato software abilitante alle funzionalità GRID.

Negli anni recenti, al concetto di GRID Computing si è affiancato quello di Cloud Computing che si focalizza sulla fornitura di servizi all'utente. Tali servizi possono consistere nell'accesso a applicativi specifici o ad interi sistemi customizzati secondo le richieste specifiche e vengono forniti da centri di calcolo di taglia importante, capaci di grossa economia di scala, attraverso in generale l'utilizzo di risorse di calcolo virtualizzate. I due concetti di GRID Computing e Cloud Computing possono, quindi, in qualche maniera integrarsi all'interno di un'infrastruttura come ENEA-GRID che, pur essendo distribuita geograficamente, presenta all'utente un ambiente unificato ed integrato. In particolare, al suo interno sono stati sviluppati da tempo i Laboratori Virtuali [3], che offrono servizi specifici a comunità omogenee di utenti e per i quali sono stati sviluppati portali tematici.

Per quel che riguarda invece i servizi virtualizzati, sono state recentemente effettuate sperimentazioni che hanno permesso di introdurre all'interno di ENEA-GRID alcune delle caratteristiche più specifiche del Cloud Computing. Questo lavoro in particolare illustrerà in dettaglio alcune di queste sperimentazioni.

Nella prima di tali sperimentazioni si è risposto all'esigenza di un committente esterno ad ENEA che aveva espresso la necessità di elaborazione massiva di dati multimediali tramite software disponibile solo per la piattaforma Windows. E' stato quindi realizzato un sistema di macchine virtuali Windows, ospitate su sistemi Linux standard di ENEA-GRID tramite il virtualizzatore KVM [4] e gestite in termini di risorse da un cluster LSF Windows virtualizzato. In questo modo, le macchine virtuali condividono in tale situazione la stessa architettura di gestione di ENEA-GRID basata su LSF ed OpenAFS. In tale progetto si è affrontato anche il problema dell'accesso a licenze di rete disponibili esternamente all'ENEA e all'accounting relativo alle elaborazioni effettuate.

L'ambito di tale sperimentazione è stato poi ampliato attraverso l'installazione su macchine ENEA-GRID di OpenNebula [5], un progetto open-source che fornisce una serie di strumenti ad-hoc per creare un ambiente di cloud computing. Le motivazioni che hanno portato alla scelta di tale sistema sono legate alle caratteristiche di OpenNebula di garantire la gestione di infrastrutture complesse ed eterogenee, esportando i servizi all'utente in modo semplice e omogeneo. Sostanzialmente, OpenNebula definisce un nuovo livello architetturale capace di gestire e coordinare diversi aspetti di un sistema complesso, quali la virtualizzazione, lo storage, la rete, etc., ottimizzando le risorse e fornendo un'interfaccia semplice e allo stesso tempo capace di esportare al meglio tutti i servizi definiti nell'infrastruttura. Lo schema concettuale multi-tier di OpenNebula e del modo in cui coopera con i servizi dell'infrastruttura è illustrato nella seguente figura:



Nello specifico di ENEA-GRID, OpenNebula è stato inizialmente utilizzato al fine di garantire la gestione dinamica di macchine virtuali esportate da KVM. In tale contesto è stata poi sviluppata l'integrazione dell'accesso a tali macchine attraverso il portale FARO di ENEA-GRID in modo da fornire all'utente la possibilità di accedere a sistemi operativi ed applicazioni specifiche per le sue esigenze. Risulta anche semplice capire come un sistema di questo tipo sia del tutto robusto e flessibile rispetto sia all'infrastruttura locale che alle applicazioni e ai servizi esportati. Infatti, upgrade architetturali e nuovi servizi e applicativi software sono facilmente supportati e configurabili al livello di OpenNebula, il tutto per semplificare il lavoro dell'utente, il quale non deve più preoccuparsi della configurazione ma soltanto dell'utilizzazione dei servizi esportati dall'interfaccia.

Referenze

- [1] S. Migliori, G. Bracco, P. D'Angelo "L'architettura di ENEA-GRID, infrastruttura distribuita per la condivisione di risorse informatiche per il calcolo scientifico"
Conferenza GARR 2005 "La rete daPERTutto", Pisa, 10-13/5/2005
- [2] A. Rocchi, S. Pierattini, G. Bracco, S. Migliori, F. Beone, C. Sciò, A. Petricca "Faro: accesso WEB

a risorse remote per l'industria e la ricerca"

Conferenza GARR 2010, Politecnico di Torino, 26-28 Ottobre 2010

[3] F. Beone, A. Rocchi, G. Mencuccini, M.L. Mongelli, I. Roselli "I Laboratori Virtuali di ENEA"

Conferenza GARR 2010, Politecnico di Torino, 26-28 Ottobre 2010

[4] <http://www.linux-kvm.org>

[5] <http://opennebula.org/>

Esperimenti di Cloud Computing in ENEA-GRID

Biografie degli autori

Giovanni Ponti è ricercatore ENEA dal 2010. E' laureato in Ingegneria Informatica e, nel 2010, ha conseguito il titolo di Dottore di Ricerca in Ingegneria dei Sistemi e Informatica. Le sue attività di ricerca riguardano il GRID computing e i sistemi HPC, le Basi di Dati e il Data Mining. In queste tematiche, è coautore di diversi articoli su riviste scientifiche, capitoli di libro, articoli presentati in conferenze nazionali e internazionali, e rapporti tecnici. Svolge inoltre attività di revisione di articoli per riviste scientifiche e conferenze del settore informatico.

Alessandro Secco è professionista nell'ambito del calcolo distribuito, collabora con Nice dal 2000, per la quale segue alcuni dei principali clienti, tra cui ENEA. Nel 2006 ha partecipato al progetto Beingrid tramite il partner Domina, per poi dedicarsi, nel 2008 a nuovi progetti di Nice.

Fiorenzo Ambrosino collabora con ENEA in aree di ricerca attinenti calcolo distribuito e applicazioni di fluidodinamica computazionale dal 2007. E' laureato in Ingegneria Aerospaziale e studente di dottorato in Ingegneria Chimica.

Giovanni Bracco è ricercatore ENEA dal 1981. E' laureato in Fisica e, dopo una breve esperienza in IBM Italia, ha iniziato la sua attività di ricerca nel campo della fusione nucleare in Francia al CEA Fontenay-aux-Roses e, successivamente, presso l'ENEA. Nel 2003 entra a far parte del gruppo ENEA di informatica e, dal 2010, è responsabile del Laboratorio Infrastrutture per il Calcolo Scientifico e ad Alte Prestazioni. Le sue aree di ricerca riguardano il GRID computing e l'HPC.

Roberto Ciavarella lavora in ENEA dal 2007. E' laureato in Ingegneria Elettronica. Collabora al progetto CRESCO occupandosi di di GPU e di aspetti di grafica 3D.

Antonio Colavincenzo è studente di Informatica. Fin dai primi anni di università, con il Linux User Group di Napoli, promuove il software libero in conferenze e workshop. In ENEA si occupa, da freelance, dell'amministrazione di varie infrastrutture informatiche, tra cui il sistema HPC Cresco ed ENEA-GRID.

Pietro D'Angelo è tecnologo ENEA dal 1976. Perito industriale in elettronica, fin dall'inizio ha lavorato nel campo dell'informatica, attualmente collabora al progetto CRESCO ed ENEA-GRID occupandosi prevalentemente della gestione e della sicurezza dei dati.

Matteo De Rosa è ricercatore ENEA. E' laureato in Scienze dell'Informazione e dal 1983 ha sempre operato nel settore informatico dell'Istituto, prima come sistemista/analista di sistemi centralizzati IBM, poi di sistemi eterogenei, per concentrarsi ad oggi su problematiche di networking in ambienti distribuiti. Ha maturato esperienza anche nella progettazione/sviluppo/amministrazione di Basi di Dati relazionali e Web, nella sicurezza perimetrale (FireWalling, Antivirus ed URL filtering, Intrusion Detection/Prevention System), nel VoIP. L'attività prevalente di oggi è quella di Network Administrator del Centro Enea di Portici e di tutti i Sistemi ivi ospitati (HPC Cresco, gLite, EFDA, VideoComunicazione Enea).

Agostino Funel è ricercatore presso l'ENEA dal 2010. E' laureato in Fisica e, nel 2006, ha conseguito il titolo di Dottore di Ricerca in Fisica. La sua attività di ricerca concerne lo sviluppo di architetture GRID innovative, il porting di codici su sistemi ad alto parallelismo, lo sviluppo e l'ottimizzazione di codici numerici per la fluidodinamica computazionale.

Guido Guarnieri è un assegnista di ricerca ENEA dal 2007. E' laureato in Ingegneria Informatica - ramo automazione industriale. Collabora al progetto CRESCO occupandosi dell'integrazione di infrastrutture hardware e software per sistemi di calcolo ad alte prestazioni

Dante Giammattei è ricercatore ENEA dal 1985. Si è laureato in Ingegneria Elettronica presso l'Università Federico II. Ha lavorato su modelli matematici nel settore della sicurezza nucleare. Si è sempre occupato di protocolli di comunicazione (in particolare SNA, Decnet e TCP/IP), nonché delle problematiche legate all'infrastruttura di networking dell'Ente. Ha progettato e diretto il cablaggio strutturato della rete del Centro di Portici, di cui è responsabile.

Silvio Migliori è direttore dell'Unità Tecnica Sviluppo Sistemi per l'Informatica e l'ICT dal 2010. E' laureato in Ingegneria Nucleare e, nel 1981, ha lavorato presso Italcable come responsabile in un team di sviluppo di sistemi gestionali. Nel 1983 ha ricoperto ruoli di responsabilità nel progetto per la fusione termonucleare controllata. Dal 1995, lavora nel settore del calcolo scientifico e, dal 2001, ricopre il ruolo di vice direttore, dove ha promosso lo sviluppo di nuove tecnologie nel campo del GRID e dell'HPC.

Samuele Pierattini è Perito Informatico, assunto nel 1989 presso il Centro Elaborazione Dati dell'ENEA di Bologna, con compiti di progettazione, realizzazione, gestione e supporto di sistemi informatici avanzati per la visualizzazione grafica 3D. Riveste dal 2007, per il progetto CRESCO e per la ENEA-GRID il ruolo di coordinatore tecnologico per le infrastrutture e le applicazioni di grafica 3D.

Salvatore Podda è ricercatore ENEA dal 1982. E' laureato in Fisica e ha una sostanziale esperienza in Fusione Nucleare avendo partecipato al progetto, operazione e data analisi e interpretazione dei sistemi di diagnostica neutronica per gli esperimenti FT, FTU e JET e più recentemente alle stesse attività nei

gruppi di fisica delle operazioni e nei laboratori di radiofrequenza degli stessi esperimenti. Attualmente è Senior Scientist del Laboratorio Infrastrutture per il Calcolo Scientifico e ad Alte Prestazioni dell'ENEA e i suoi compiti includono il progetto, la gestione e l'operazione di sistemi di calcolo ad alte prestazioni

Alessio Rocchi lavora dal 2005 in ENEA. E' Ingegnere informatico e attualmente si occupa di GRID e Cloud computing. I suoi interessi scientifici riguardano I sistemi distribuiti, con particolare riguardo alla security e alla fault tolerance.