

Conferenza GARR_10

Selected papers



Welcome to the Future Internet!

La rete della ricerca e la sua comunità
oggi: servizi, applicazioni, idee di domani

Torino, 26-28 ottobre 2010

Conferenza GARR_10

Selected papers



Welcome to the Future Internet!

La rete della ricerca e la sua comunità
oggi: servizi, applicazioni, idee di domani

Torino, 26-28 ottobre 2010

ISBN 978-88-905077-1-7

Tutti i diritti sono riservati ai sensi della normativa vigente.

La riproduzione, la pubblicazione e la distribuzione, totale o parziale, di tutto il materiale originale contenuto in questa pubblicazione sono espressamente vietate in assenza di autorizzazione scritta.

Copyright © 2011 Associazione Consortium GARR

Editore: Associazione Consortium GARR

Via dei Tizii, 6, 00185 Roma, Italia

<http://www.garr.it>

Tutti i diritti riservati.

Curatori editoriali: Federica Tanlongo, Maddalena Vario, Carlo Volpe

Progetto grafico e impaginazione: Carlo Volpe

Prima stampa: Aprile 2011

Numero di copie: 1500

Stampa: Tipolitografia Fa.Ro.press, Via del Falco 28, 00193 Roma, Italia

Tutti i materiali relativi alla Conferenza GARR 2010 sono disponibili all'indirizzo: <http://www.garr.it/conf10>.

Indice



Introduzione	5
C. Allocchio	
L2L (Live to e-Learning)	7
M. Bertazzo	
WIDEN: implementazione di protocolli di Clinical Trials con servizi web	13
P. Cerello, A. Biggi, S. Chauvie, A. Gallamini, A. Stancu	
LinkO: una infrastruttura in fibra ottica per confronti remoti di tempo e frequenza ad alta accuratezza	17
C. Clivati, D. Calonico, F. Levi, A. Mura, G. A. Costanzo, A. Godone	
Utilizzo di Zabbix per il Monitoring di risorse di calcolo e sistemi di storage su WAN in ambito Grid Computing	23
A. Colavincenzo, A. Petricca, G. Bracco, S. Migliori, F. Simoni, S. Pecoraro, D. Giammattei, A. Rocchi, C. Sciò	
NEXPreS: Completamento ed evoluzione dell'e-VLBI	27
M. Nanni	
FARO: Accesso web a risorse remote per l'industria e la ricerca	32
A. Rocchi, S. Pierattini, G. Bracco, S. Migliori, F. Beone, C. Sciò, R. Ciavarella, A. Petricca	
WNoDeS: un servizio per la gestione di infrastrutture condivise Cloud e Grid	37
D. Salomoni, A. Chierici, A. Italiano, E. Ronchieri, P. Solagna	

Introduzione

Claudio Allocchio

Chair del Comitato di Programma della Conferenza GARR 2010



“Benvenuti nel futuro che è già qui!” rappresenta la frase più adatta ad inquadrare il tema della Conferenza GARR 2010, della quale raccogliamo in questa pubblicazione i lavori migliori.

Rispondere al tema della *call for papers* di questa edizione è stato sicuramente una sfida per gli autori: si chiedeva loro infatti di parlare da utenti della rete prima ancora che da ricercatori, focalizzando l'intervento sulla rete GARR, sui servizi e sulle caratteristiche che essa deve avere per supportarli proprio nel loro lavoro, piuttosto che concentrarsi sui temi scientifici e sui risultati che da quel lavoro derivano. Una sfida alla quale gli autori hanno però saputo rispondere molto bene, arricchendo ancor di più l'insieme di innovazioni che, noi che realizziamo la rete, dovremo mettere a loro disposizione.

Da paradigmi in cui esisteva una netta distinzione dei ruoli tra gli oggetti in grado di fare calcolo e i servizi e gli elementi che li interconnettono si è passati ad un sistema in cui, giustamente, l'utente può disinteressarsi dell'infrastruttura per occuparsi dei problemi di ricerca e applicativi, ottenendo da essa i servizi di cui ha bisogno in modo del tutto trasparente. Tale modello è reso possibile dalla rete stessa. Anche l'idea di “rete” come infrastruttura condivisa si modifica per diventare un servizio che fornisce, a richiesta, tante reti differenti, ognuna con le sue caratteristiche. Sino ad arrivare all'estremo dove è la rete stessa a diventare uno strumento di stu-

dio e misura (del tempo, come vedrete leggendo l'articolo su LinkO dell'INRiM) oppure un elemento fondamentale dell'apparato sperimentale (come nel caso di NEXPREs).

Una rete onnipresente, che permette di avere a disposizione il dato o la sua elaborazione, quando e dove serve, e che esce dal laboratorio di ricerca, per arrivare a servire tutti, specialmente quando si parla di applicazioni in campo biomedico o per la didattica. È questa forse la sfida più grande che i lavori della Conferenza hanno evidenziato: il saper produrre dei servizi di rete che escono dal modello al quale siamo abituati, per avvicinarsi a quello che le prossime generazioni ritengono ovvio e naturale: “la rete c'è, fornisce servizi e per interagire con essa non servono necessariamente un computer, una tastiera e un mouse”.

Non ho citato sinora la parola “Cloud”, volontariamente, dal momento che uno dei messaggi emersi più chiaramente dalla Conferenza è che non occorre partire da concetti astratti: il modo migliore di realizzare davvero “la nuvola”, anzi “le nuvole”, forse consiste proprio nel rispondere alle sfide che le esperienze raccontate in questi articoli ci presentano.

Nel 2005 il titolo della Conferenza GARR era stato “la Rete daPerTutto”: giocando con le parole, proponevamo quello che allora era ancora un obiettivo tutto da raggiungere, ma cinque anni dopo possiamo dire essere diventato realtà. Oggi, con la rete GARR-X ed i suoi servizi già disponibili o in fase di rapi-

da realizzazione, siamo in grado di rispondere a molte delle sfide che gli utilizzatori pongono: il futuro è già qui, ma allo stesso tempo arrivano anche nuove richieste, nuove sfide, che prima non erano nemmeno pensabili. La Conferenza del 2010 e gli articoli che leggerete mostrano che dobbiamo, ancora una volta, adeguare il modo con cui costruiamo le reti avanzate a ciò che gli utenti devono fare, in un circolo virtuoso che produce sempre nuovi ed interessanti risultati.

Buona lettura e... alla prossima Conferenza GARR!



L'entrata principale del Politecnico di Torino (sede dell'evento) e il manifesto della Conferenza GARR 2010

Comitato di programma

Claudio Allocchio - **GARR (Chair)**
 Maristella Agosti - **Università degli Studi di Padova**
 Isabella Castiglioni - **CNR / Osp. San Raffaele di Milano**
 Giovanni Erbacci - **CINECA**
 Luciano Gaido - **INFN Sezione di Torino**
 Laura Leone - **GARR**
 Marcello Maggiora - **Politecnico di Torino**
 Silvio Migliori - **ENEA**
 Gabriella Paolini - **GARR**
 Ciro Palermo - **Università IUAV di Venezia**
 Angela Re - **Università degli Studi di Torino**

Advisory Committee

Giuseppe Battistoni - **INFN Sezione di Milano**
 Mauro Campanella - **GARR**
 Massimo Cocco - **INGV**
 Silvano Gai - **CISCO**
 Mimmo Laforenza - **CNR-IIT**
 Maurizio Lancia - **CNR**
 Marco Mezzalama - **Politecnico di Torino**
 Daniela Perani - **CNR/Osp. S.Raffaele/Ministero Salute**
 Giovanni Ragone - **Università di Roma La Sapienza**

Comitato Organizzatore Locale

Marcello Maggiora - **Politecnico di Torino**
 Marco Mezzalama - **Politecnico di Torino**
 Angela Re - **Università degli Studi di Torino**

Tutte le presentazioni e maggiori informazioni sono disponibili sul sito dell'evento: www.garr.it/conf10

GARR PREMIA I GIOVANI INNOVATORI

Il Consortium GARR, nell'ambito delle attività di promozione della conoscenza, istituisce due premi destinati a proposte o lavori originali formulati da giovani ricercatori (nati a partire dal 1 gennaio 1981).

I premi sono dedicati al ricordo di due personalità di spicco che si sono spese nella costruzione della rete della ricerca italiana: il prof. Orio Carlini, che è stato uno dei padri fondatori della rete GARR e il prof. Antonio Ruberti, che è stato Rettore e Ministro dell'Università e della Ricerca Scientifica e Tecnologica.



PREMIO ORIO CARLINI

sui temi del Future Internet e delle tecnologie abilitanti al lavoro collaborativo in rete.

PREMIO ANTONIO RUBERTI

sui temi delle Next Generation Network e delle tecnologie di trasmissioni dati.

Ciascuno dei due premi in palio ha il valore di **5.000,00 euro** interamente finanziati dal Consortium GARR.

La scadenza per la sottomissione delle domande è fissata al **15 settembre 2011**.

Alla domanda dovranno essere allegati in formato pdf:

- a) Elaborato della proposta in italiano o in inglese, di massimo 35.000 caratteri inclusi gli spazi. Nel caso di articoli o tesi di dottorato che superino tale limite si richiede una sintesi;
- b) Curriculum vitae e scientifico del candidato.

Nel caso di domanda presentata da più candidati, ognuno dei quali dovrà soddisfare il requisito anagrafico, il punto b) si intende adeguatamente ripetuto; l'eventuale premio verrà suddiviso in parti eguali tra i candidati.

Le **modalità per l'invio della domanda** di partecipazione sono descritte alla pagina web dedicata al concorso: www.garr.it/premi.

Tra le proposte pervenute, una apposita Commissione sceglierà, con giudizio insindacabile, i lavori migliori per interesse ed originalità.

I vincitori verranno premiati nel corso di una cerimonia pubblica e saranno invitati ad illustrare il proprio lavoro durante la Conferenza GARR 2011.

MAGGIORI INFORMAZIONI: WWW.GARR.IT/PREMI

 Consortium
GARR

L2L (Live to e-Learning)

Matteo Bertazzo

CINECA



Abstract. Il servizio L2L (Live to e-Learning) consente di trasformare in modo semi-automatico le lezioni dal vivo in oggetti e-learning pronti per la pubblicazione e l'erogazione tramite una piattaforma e-learning. Il servizio è stato progettato per soddisfare le esigenze di un consorzio di università, cercando di venire incontro alle necessità delle singole realtà ed al contempo di contenere i costi.

Per la realizzazione di questo servizio è stato utilizzato un unico sistema di *Digital Asset Management* (DAM) centralizzato, basato sulla piattaforma open source MediaMosa.

Un plug-in rende possibile l'integrazione del sistema DAM con la piattaforma di e-learning open source Moodle. L'architettura complessiva è aperta e modulare e consente sia la diffusione di modelli di servizio differenti che il riutilizzo dei contenuti digitali in contesti diversi (ad esempio: piattaforme e-learning, portali web, webTV, digital library).

1. Introduzione

Nell'ultimo decennio le università hanno sentito in maniera crescente l'esigenza di integrare all'interno della didattica e-learning le registrazioni delle lezioni in aula. Tuttavia le video-lezioni sono ancora poco diffuse a causa dei costi e dei tempi necessari alla loro realizzazione.

Con L2L il CINECA, consorzio interuniversitario formato da 44 università, ha realizzato un servizio che viene incontro alle diverse esigenze delle università consorziate, cercando comunque di razionalizzare i costi e le risorse.

Il servizio si basa su un unico sistema *Digital Asset Management* (DAM) centralizzato, che serve tutte le università e racchiude le componenti di produzione ed erogazione. La centralizzazione del sistema assicura bassi costi di *start-up* per l'ateneo che volesse adottare il servizio, mentre l'approccio modulare consente di implementare servizi sia in modalità ASP (*Application Service Provider*), ovvero su sistemi informatici completamente esterni all'università, che in modalità interna o ibrida, ad esempio con erogazione tramite portale web o LMS dell'università.

Particolare attenzione è stata prestata anche agli strumenti per aumentare la produttività dei

docenti: è stata realizzata una modalità semi-automatica per la registrazione e la gestione dei metadati, mentre la post-produzione e l'erogazione sono state completamente automatizzate.

Il risultato è la diminuzione del lasso di tempo che passa dalla lezione in aula alla pubblicazione di un oggetto digitale "consistente" sulla piattaforma e-learning. Attraverso il sistema DAM, le lezioni e tutti i contenuti ad esse associate sono facilmente riutilizzabili in diversi contesti e servizi: MediaMosa, fornendo un insieme completo di servizi web e consentendo un'alta granularità sia nell'accesso sia nel controllo dei contenuti, permette la realizzazione di nuovi scenari di utilizzo.

L2L supporta la possibilità di registrare *offline* (ad esempio nelle aule senza connettività) e, grazie all'approccio modulare, permetterà di migliorare ulteriormente il sistema per rispondere ai futuri bisogni delle università e dei loro utenti.

2. Riferimenti ai lavori correlati

Nell'ultimo decennio i sistemi automatici di registrazione delle lezioni hanno riscosso l'interesse delle università e degli sviluppatori. Attualmente sono disponibili molti siste-

mi avanzati di registrazione delle lezioni in aula (*Lecture Recording System - LRS*) che permettono di ottenere buoni risultati con il minimo sforzo da parte degli utenti, come quelli proposti in [13, 6, 7, 2]. Tuttavia l'integrazione di questi sistemi con LMS standard è di solito complicata dall'integrazione delle fasi di erogazione e di presentazione nel sistema di acquisizione stesso.

Recentemente sono state realizzate diverse soluzioni in cui il sistema di registrazione è stato integrato con successo con l'LMS grazie all'intermediazione di una piattaforma DAM. In [9], un *Content Management System* (CMS) è utilizzato per memorizzare e visualizzare le lezioni dentro l'LMS. In [2], il sistema DAM non è direttamente responsabile della presentazione, che è gestita da uno speciale player integrato nel LMS grazie ad un apposito modulo.

L2L adotta un approccio simile, estendendo ulteriormente la struttura modulare: l'LRS non si basa sulla disponibilità del DAM per la registrazione, permettendo in tal modo la registrazione *offline* e l'*editing* delle lezioni. Recentemente Tele-Task ha introdotto uno strumento di ricerca semantico che identifica

automaticamente le parole chiave e le organizza in aree concettuali, con l'obiettivo principale di consentire agli utenti di eseguire ricerche in linguaggio naturale [8]. In L2L viene eseguita un'analogia analisi semantica sul testo delle slide per mezzo del *Concept Mapper* [11] in modo da permettere agli utenti di navigare la lezione utilizzando un indice di concetti.

3. Architettura del servizio

L'architettura del servizio L2L è basata sul modello *Service Oriented Architecture* (SOA) [3] e le interazioni tra i componenti del servizio avvengono principalmente attraverso l'uso di API REST (*Representational Transfer State web Application Program Interface*).

Come mostrato in figura 1, il cuore dell'architettura è la piattaforma MediaMosa, che è a sua volta basata sul modello SOA e permette interazioni attraverso servizi web, sia dall'esterno sia verso le proprie componenti interne. MediaMosa è stata estesa mediante la creazione di un servizio di "*ingestion*" (modulo L2L Ingester) al fine di aprire, validare ed elaborare tutti i pacchetti provenienti da stazioni L2L remote (depositi FTP o WebDAV).

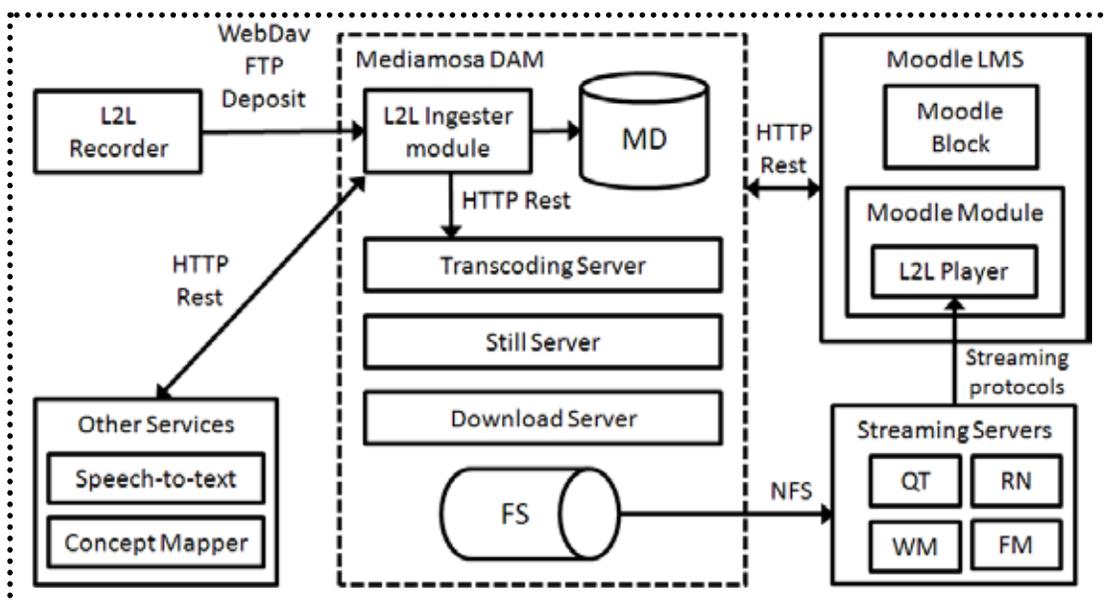


Fig. 1 La piattaforma MediaMosa, cuore dell'architettura del servizio L2L.

Il modulo di *ingestion* utilizza i servizi REST di MediaMosa per memorizzare, analizzare e codificare tutti i contenuti all'interno di un pacchetto L2L. L'*ingerer* si occupa anche dell'archiviazione dei metadati descrittivi (il DAM supporta l'estensione allo schema Qualified Dublin Core - QDC) e richiama il servizio Concept Mapper come descritto nella sezione 5.

In base alle chiamate REST provenienti dal modulo *ingerer*, il DAM esegue una serie di azioni su ogni media correlato alle registrazioni. Una volta che l'analisi del media e l'estrazione dei metadati tecnici è stata completata, questo viene elaborato dal server di transcodifica. Il server di transcodifica è basato su FFmpeg ed opera in base ad un insieme predefinito di profili e al tipo di lezione L2L.

Il modulo Moodle permette l'integrazione del DAM con il LMS: utilizzando questo modulo la piattaforma Moodle diventa un client di MediaMosa che utilizza i servizi REST per gestire e accedere alle risorse. Il blocco Moodle permette ai docenti e agli amministratori di controllare e di gestire le loro lezioni, sia a livello di attività Moodle che a quello del DAM

4. L2L: componenti e tecnologia

Le tre componenti principali del servizio L2L sono: L2L Recorder, piattaforma DAM e modulo di Moodle (che permette l'integrazione del DAM con il LMS).

Il registratore L2L è basato su un PC, desktop o portatile, e su un software in grado di registrare e sincronizzare una sorgente audio e due fonti video: il video del docente più la cattura dello schermo del computer, oppure una sorgente video esterna come una lavagna interattiva. Il client L2L abilita e aiuta l'insegnante a definire e descrivere la lezione.

Per quanto riguarda l'hardware della stazione di L2L, esso si basa su un PC con una scheda di acquisizione video, una telecamera Pan Tilt Zoom (PTZ), un dispositivo di cattura VGA2USB e un microfono Lavalier.

Il software L2L è un'applicazione MS Windows .NET e i flussi audio/video sono codificati in alta qualità nel formato Windows Media Video.

Il processo di acquisizione della lezione prevede i seguenti passaggi:

- autenticazione del docente basata su *Security Assertion Markup Language* (SAML);
- recupero delle informazioni personali degli insegnanti da Moodle tramite servizi web (SOAP);
- recupero delle informazioni descrittive dei corsi;
- associazione automatica alla lezione dei metadati specifici relativi al docente e alle sue materie di insegnamento;
- selezione della modalità di registrazione: il client L2L supporta diversi tipi di registrazione (solo audio, audio-video del docente, schermo del computer, audio-video e un video secondario), al fine di meglio soddisfare i bisogni del docente;
- inserimento allegati;
- inserimento delle informazioni descrittive della lezione: titolo, abstract e tag;
- registrazione della lezione;
- controllo qualità: l'insegnante è in grado di controllare la qualità di registrazione visualizzando l'anteprima;
- caricamento della lezione sul *repository*: un processo automatico di caricamento gestisce l'archiviazione in modo trasparente per il docente.

Grazie ad un specifico plug-in MS Office, L2L registra i tempi di cambio di diapositiva, inserisce nel pacchetto della lezione gli *snapshot* delle slide e un file XML contenente i testi delle stesse. Il registratore supporta la registrazione *offline*, per soddisfare l'esigenza diffusa di registrare le lezioni tenute in aule prive di connettività o per i docenti che preferiscono registrare una serie completa di lezioni prima di caricarle sul *repository*.

Il DAM integrato nella piattaforma L2L è basato sulla piattaforma MediaMosa, softwa-

re gratuito e open source per la realizzazione di servizi *web oriented* di gestione ed erogazione di media.

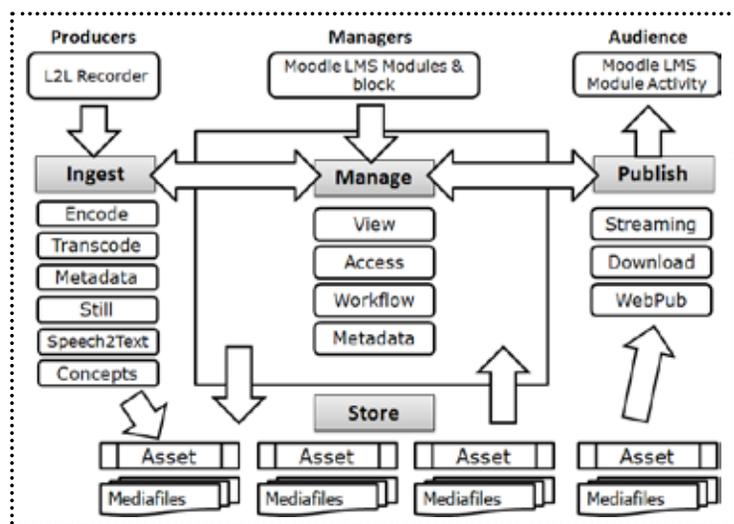


Fig. 2 Schema logico del sistema DAM

La scelta di MediaMosa per la piattaforma DAM permette di avere a disposizione funzionalità per la ricerca, il caricamento, la transcodifica, la fruizione, così come un'alta granularità nel controllo dell'accesso ai contenuti per e da parte degli utenti.

Come mostrato nella figura 2, che rappresenta lo schema logico del sistema DAM, MediaMosa implementa le quattro principali funzionalità del sistema - *ingestion*, memorizzazione, gestione e pubblicazione - attraverso l'utilizzo di chiamate REST.

Il modello MediaMosa si basa sui concetti di *asset* e *mediafiles*: un asset è composto da una serie di mediafiles (il file originale e tutte le versioni transcodate). In questo modo un asset può essere fruito usando una serie di protocolli e formati diversi.

L'integrazione tra il *repository* e la piattaforma e-learning, che consente di gestire, pubblicare e distribuire le lezioni L2L, è stata realizzata

segundo il paradigma Moodle, con la creazione un apposito modulo, ovvero una nuova attività di apprendimento. Il modulo L2L

può essere utilizzato congiuntamente alle altre risorse e attività di un corso e la stessa lezione può essere utilizzata in diversi corsi e installazioni Moodle che condividono lo stesso *repository*. Anche la comunicazione tra il *repository* e il modulo Moodle di L2L avviene tramite REST.

Il modulo consente al docente di definire le attività di apprendimento all'interno del corso, di cercare una lezione tra tutte le registrazioni memorizzate nel *repository* e di limitare la visibilità delle attività L2L in base al ruolo e al profilo degli utenti.

Un docente può inoltre ricercare le lezioni da associare alle attività in base ai metadati in esse contenuti. Lo stesso modulo L2L permette agli studenti di accedere alla lezione digitale.

Dato che una lezione di L2L è un insieme di risorse multimediali e descrittori XML, il modulo L2L di Moodle ha bisogno di uno specifico *player* per la visualizzazione della lezione. Il *player* L2L offre diverse funzionalità: ri-

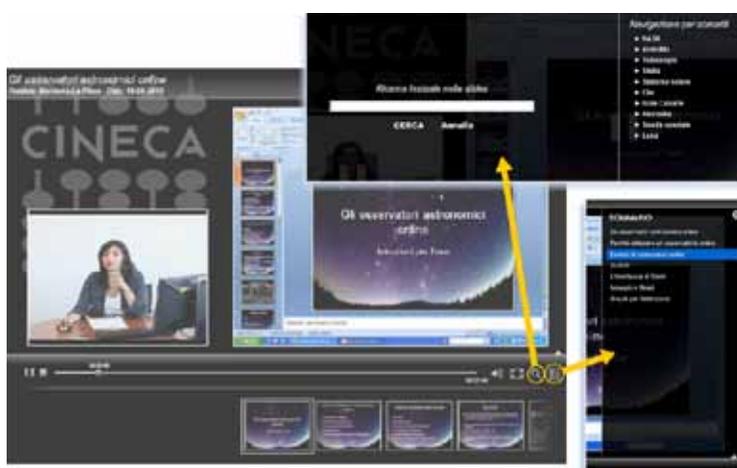


Fig. 3 Schermata di una lezione con il player L2L in cui viene evidenziata la possibilità della navigazione per concetti

produzione sincronizzata, ricerca *full-text*, navigazione della lezione.

Una lezione L2L ha una struttura complessa e non è adatta ad essere erogata utilizzando il paradigma del podcasting, per questo motivo è stato aggiunto un servizio automatico di podcasting. Questo servizio, basato sul framework multimediale GStreamer, effettua la post-elaborazione automatica, trasformando le lezioni L2L in un'unica risorsa audio-video a seconda del tipo di lezione.

5. Concetta: fruizione delle lezioni tramite concetti

L2L fornisce agli studenti la possibilità di navigare la lezione utilizzando un elenco di concetti, raggiungendo lo specifico istante in cui il singolo concetto viene discusso all'interno della lezione.

Nella maggior parte dei casi una registrazione L2L dura più di 45 minuti e quindi creare manualmente annotazioni e sintesi sarebbe estremamente dispendioso. Per questo motivo L2L ha integrato lo strumento Concept Mapper, precedentemente sviluppato da CINECA nell'ambito del progetto Papyrus, finanziato dalla Comunità europea [11].

Il Concept Mapper si basa sullo strumento WikipediaMiner, che permette di utilizza-

re Wikipedia sia come fonte di metadati che come risorsa linguistica per disambiguare parole chiave o fornire sinonimi di parole chiave (le ancore, ovvero il testo che gli utenti utilizzano per linkare in Wikipedia) e traduzioni (utilizzando informazioni sugli *interwiki link*) [10]. Il Concept Mapper sfrutta le informazioni di Wikipedia per identificare automaticamente concetti rilevanti nei testi e il processo complessivo può essere riassunto nei seguenti passaggi logici:

- identificazione delle possibili parole chiave;
- ricerca delle ancore di tutte le parole chiave candidate;
- disambiguazione della parola chiave;
- selezione delle parole chiave rilevanti.

L'output del processo consiste in un elenco di concetti, individuati con i titoli delle pagine di Wikipedia che li descrivono, dove per ogni concetto viene fornita la misura di correlazione (al documento o a un dominio specifico), una breve descrizione (il primo paragrafo della pagina di Wikipedia), i sinonimi (ancore e *redirect*), le categorie e le traduzioni.

Come mostrato in figura 4, la piattaforma L2L alimenta il Concept Mapper con i testi estratti dal diapositive MS-Power Point e il Concept Mapper identifica i concetti.

Dato che ogni diapositiva è associata ad un intervallo di tempo specifico, i concetti identificati possono essere utilizzati per fornire agli utenti un indice, che può essere utilizzato in alternativa alla ricerca standard *full-text* o la tabella di navigazione dei contenuti (*fig.3*). I concetti estratti vengono utilizzati anche per ricercare tra tutte le lezioni L2L presenti nel *repository* (aggregazione per concetti).

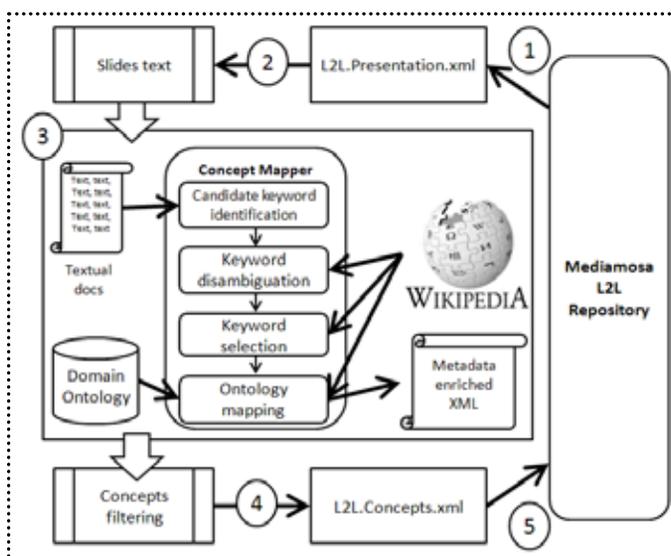


Fig. 4 Schema di funzionamento della piattaforma L2L

6. Sviluppi futuri

Dato che le università telematiche italiane devono rispettare specifiche norme relative all'erogazione e al-

la conservazione dei dati di tracciamento dei corsi e-learning, è in fase di sviluppo l'adattamento della fase di fruizione delle lezioni L2L in modo che siano compatibili con un *player Sharable Content Object Reference Model (SCORM)* [1].

In questo momento le funzionalità più avanzate di L2L (ricerca *full-text*, estrazione concetti) si basano sull'utilizzo di presentazioni Microsoft PowerPoint. Si stanno valutando alcune tecniche più generali per le presentazioni, come quelli proposti in [13, 6, 7, 2] o in [8], in modo da fornire le stesse funzionalità anche quando il presentatore utilizza formati diversi.

Sono invece appena iniziati studi per l'integrazione di servizi *speech to text* (basato su IBM Attila Speech recognition Toolkit II [12]), per aggiungere un nuovo player HTML5 e per adottare standard open di codifica audio-video.

Riferimenti bibliografici

- [1] Advanced Distributed Learning. ADL guidelines for creating reusable content with SCORM 2004. Technical report, 7 2008. Available from: <http://www.adlnet.gov>
- [2] D. H. Anh, G. G. D. Nishantha, Y. Hayashida, and P. Davar. A flash-based lecture recording system and its integration with lms. In ICACT'10: Proceedings of the 12th international conference on Advanced communication technology, pages 1425–1429, Piscataway, NJ, USA, 2010, IEEE Press
- [3] D. K. Barry. Web services and service-oriented architectures: the savvy manager's guide. Morgan Kaufmann, San Francisco, 2003
- [4] CCSDS. Reference model for an open archival information system (OAIS). Blue book. Technical Report 1, 2002. <http://public.ccsds.org/publications/archive/650x0b1.pdf>
- [5] S. Chandra. Lecture video capture for the masses. SIGCSE Bull., 39(3):276–280, 2007. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.71.8856&rep=rep1&type=pdf>
- [6] A. Dong and H. Li. Ontology-driven annotation and access of presentation video data. Journal of Theoretical and Applied Information Technology, 4(9):844–848, 9 2008. Available from: <http://www.jatit.org/volumes/research-papers/Vol4No9/11Vol4No9.pdf>
- [7] Q. Fan, K. Barnard, A. Amir, A. Efrat, and M. Lin. Matching slides to presentation videos using sift and scene background matching. In MIR '06: Proceedings of the 8th ACM international workshop on Multimedia information retrieval, pages 239–248, New York, NY, USA, 2006, ACM
- [8] S. Linckels, S. Repp, N. Karam, and C. Meinel. The virtual tele-tASK professor: semantic search in recorded lectures. SIGCSE Bull., 39(1):50–54, 2007. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.79.5237&rep=rep1&type=pdf>
- [9] R. Mertens, N. Birnbaum, M. Ketterl, and R. Rolf. Integrating lecture recording with an lms: An implementation report. In C. Bonk, M. Lee, and T. Reynolds, editors, Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2008, pages 1067–1074, Las Vegas, Nevada, USA, 11 2008. University of Osnabrück, AACE. Available from: http://www.inf.uos.de/papers_pdf/2008_06.pdf
- [10] D. Milne. An open-source toolkit for mining Wikipedia. Published online, 2009. Accessed 16.02.2010, 14.00h. Available from: <http://www.cs.waikato.ac.nz/~dnk2/publications/AnOpenSourceToolkitForMiningWikipedia.pdf>
- [11] G. Paci, G. Pedrazzi, and R. Turra. Wikipedia-based approach for linking ontology concepts to their realisations in text. In N. Calzolari (Conference Chair), K. Choukri, B. Maegaard, J. Mariani, J. Odjik, S. Piperi-

dis, M. Rosner, and D. Tapias, editors, Proceedings of the seventh International Language Resources and Evaluation (LREC'10), Valletta, Malta, may 2010. European Language Resources Association (ELRA). Available from: http://www.lrec-conf.org/proceedings/lrec2010/pdf/132_Paper.pdf

[12] H. Soltau, B. Kingsbury, L. Mangu, D. Povey, G. Saon, and G. Zweig. The IBM 2004 Conversational Telephony System for Rich Transcription. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, volume 1, pages 205–208, 2005

[13] P. Ziewer. Flexible and Automated Production of Full-Fledged Electronic Lectures.

PhD thesis, Technischen Universit at Munchen, 2006



Matteo Bertazzo

m.bertazzo@cineca.it

Laureato in informatica, lavora al CINECA come tecnologo.

Le sue attività sono focalizzate sullo studio, l'implementazione e l'integrazione di servizi e-learning e multimediali (DAM). Partecipa a vari progetti di ricerca europei sull'utilizzo di tecnologie innovative e dei Virtual World a scopi educativi.

WIDEN: implementazione di protocolli di *Clinical Trials* con servizi web

P. Cerello¹, A. Biggi², S. Chauvie^{1,2}, A. Gallamini², A. Stancu²

¹INFN, Sezione di Torino, ²Ospedale S. Croce e Carle di Cuneo



Abstract. Nei protocolli per i “*trials*” clinici (TC), che rivestono un ruolo essenziale nella medicina moderna, sempre più spesso si sceglie di utilizzare tecniche di *imaging* come elemento essenziale per le decisioni relative ai singoli pazienti partecipanti, operando in modalità multicentrica, con molte strutture che contribuiscono al reclutamento dei pazienti.

La piattaforma WIDEN, utilizzando servizi web per la raccolta, la gestione e l’analisi dei dati e delle immagini, consente di gestire al meglio il *workflow* corrispondente a quanto previsto dal protocollo approvato, garantendo in tempo reale l’accesso alle immagini per i *reviewers*, la notifica dei risultati ai responsabili dei siti che reclutano i pazienti e la conoscenza delle statistiche relative al *trial* da parte dei responsabili del loro andamento complessivo.

1. Introduzione

I “*trials*” clinici (TC) rivestono un ruolo essenziale nella medicina moderna. Infatti, prima di essere reso disponibile alla popolazione, ogni nuovo approccio terapeutico deve andare incontro ad un processo di valutazione dell’efficacia e della tolleranza da parte dei pazienti.

Nel registro mantenuto presso il National Institute of Health (www.nih.gov) sono attualmente iscritti 90.686 TC (32.951 attivi) in 172 Paesi. Sempre più frequentemente, i protocolli dei *trials* prevedono l’uso delle tecnologie di *medical imaging* e di conseguenza richiedono l’analisi di quantità progressivamente crescenti di immagini, spesso raccolte in vari luoghi nel mondo nell’ambito dello stesso studio.

La PET (*Positron Emission Tomography*), storicamente utilizzata nella diagnosi, stadiazione, ristadiazione e *follow-up* delle più svariate patologie tumorali, da alcuni anni è divenuta importantissima anche nella valutazione della risposta alla terapia: in questo caso, il suo utilizzo è anche quantitativo.

Ad esempio, la PET è stata introdotta anche nel corso del trattamento (*interim-PET*), con funzione di indicatore della risposta alla terapia e quindi di strumento per modulare l’inten-

sità del trattamento in base al suo andamento. Tuttavia, dal momento che non esistono regole di interpretazione per la *interim-PET* riconosciute dalla comunità scientifica internazionale, si è creata la necessità di disporre di un *panel* di esperti medici nucleari preposti alla interpretazione delle immagini. Questo contesto rivoluziona la filosofia dell’*imaging* nei TC in relazione a due aspetti fondamentali:

- ai singoli centri diagnostici PET periferici si sostituisce un *panel* di esperti medici nucleari per la interpretazione delle immagini per ogni TC;
- le immagini devono essere interpretate rapidamente per permettere ai singoli centri clinici di non perdere tempo prezioso nella terapia dei pazienti inseriti nel TC.

Per rendere possibile tutto ciò è necessario condividere le immagini e diagnosi a distanza. WIDEN (*Web-based Image and Diagnosis Exchange Network*) affronta il problema utilizzando tecnologie software open source di maturità ampiamente dimostrata dal loro uso su scala mondiale (PHP, Perl, MySQL, JavaScript) e mette a disposizione la piattaforma ai centri partecipanti ad un TC, il cui protocollo è implementato attraverso un portale web dedicato. Durante lo sviluppo sono stati mante-

nuti come criteri fondamentali i requisiti sulla sicurezza e scalabilità del sistema.

In questo contesto, il ruolo dell'infrastruttura di rete è fondamentale, in quanto consente il trasferimento delle immagini (circa 150-200 MB per un tipico studio PET-CT) in tempo reale tra i centri di reclutamento dei pazienti ed il server WIDEN e tra questo e le stazioni di lavoro dei *reviewers* che effettuano la refertazione.

2. WIDEN

WIDEN è stato disegnato ed implementato a partire dalle funzionalità richieste dal *trial* clinico HD0607 che corrispondono sostanzialmente agli elementi funzionali di base di ogni protocollo di *trial* clinico/farmaceutico (fig. 1): la raccolta di dati e immagini dei pazienti reclutati, la notifica dei nuovi casi ai *reviewers* deputati alla valutazione delle immagini, la ricezione dei risultati delle revisioni da parte degli esperti, l'elaborazione dei risultati per costruire il risultato sintetico relativo al paziente, la notifica del risultato al centro di reclutamento del paziente, in modo tale che in funzione del risultato medesimo, si possa adeguare il prosieguo della terapia.

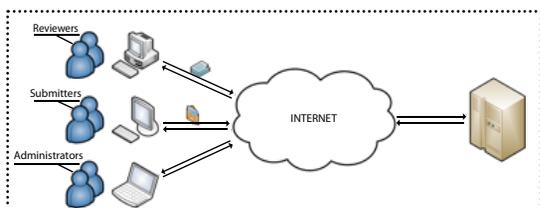


Fig. 1 L'accesso ai servizi WIDEN è basato su una topologia a stella, in cui il portale WIDEN interagisce direttamente con tutti i nodi che partecipano al *trial* e agisce da centro di raccolta, gestione ed analisi dei dati.

WIDEN è utilizzato da circa due anni nel *trial* HD0607, il cui protocollo è schematizzato in figura 2: sono stati abilitati più di cento utenti con credenziali di accesso e ruoli diversi e registrati circa 400 casi di pazienti affetti da linfoma.

Agli utilizzatori di WIDEN, siano essi i re-

sponsabili dei centri partecipanti al TC o i *reviewers*, non è richiesta alcuna installazione di software dedicato, nè modifiche alla configurazione della sicurezza del sito, in quanto tutti gli aspetti legati alla sicurezza di accesso al sistema e di trasferimento dei dati sono garantiti dall'uso del protocollo https. L'accesso ai servizi, effettuato tramite un internet browser, consente in ogni momento la disponibilità di tutte le informazioni relative alle immagini e (nella sola sede di reclutamento) ai pazienti cui queste sono associate.

Nella fase più recente dello sviluppo, poiché HD0607 ha acquisito una quantità di dati sufficiente ad impostare un'analisi dei risultati, nell'interfaccia di WIDEN sono stati inseriti moduli di analisi statistica del contenuto del database. Il modulo di base calcola alcuni parametri generali, quali il numero di utenti, di centri coinvolti, di pazienti reclutati, ecc. Inoltre, poiché i casi devono essere valutati da più di un radiologo, è necessario verificare la variabilità "inter-osservatore": per questo è stato implementato un modulo che calcola il livello di accordo per ogni coppia di revisori (Cohen's Kappa) e quello complessivo di tutto l'insieme di revisori (Krippendorff's Alpha).

Tutte le quantità calcolate, essendo i moduli integrati nel sistema, sono aggiornate in tempo reale, ogni volta che un nuovo caso o una nuova refertazione vengono registra-

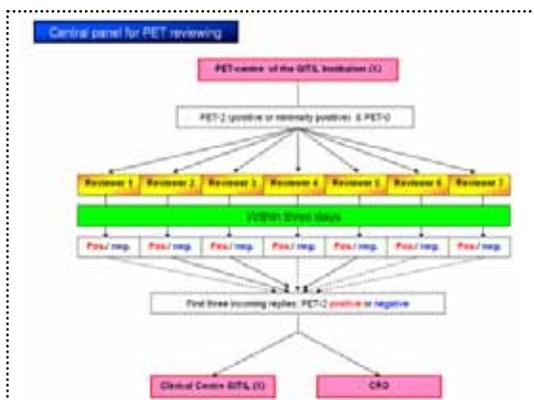


Fig. 2 Lo schema del protocollo previsto per il trial clinico HD0607

ti: questa caratteristica è stata particolarmente apprezzata dal centro di coordinamento di HD0607, che in ogni momento ha a disposizione tutte le informazioni statistiche che permettono la valutazione dei risultati parziali rispetto agli obiettivi del *clinical trial*. Questa è una novità fondamentale per un *trial* clinico in cui di solito si effettuano le valutazioni statistiche ad intervalli predeterminati, con il rischio di prostrarre inutilmente uno studio che abbia già prodotto i risultati attesi o di non terminarne in tempo uno di cui si dimostri, in itinere, l'inutilità.

WIDEN impone inoltre, tramite l'interfaccia di caricamento dei dati e delle immagini, che i centri aderenti osservino rigorosamente il protocollo approvato per il TC.

L'esperienza degli utilizzatori del gruppo GITIL (Gruppo Italiano Terapie Innovative dei Linfomi), che ha organizzato il *trial* HD0607 è stata essenziale nel promuovere l'uso di WIDEN presso altri studi, sempre nel campo dello studio dei linfomi. Attualmente sono quattro i progetti che, come "pilot user", utilizzano WIDEN.

3. I vantaggi di WIDEN

WIDEN è stato disegnato per facilitare il più possibile le operazioni agli utenti. I suoi vantaggi, dal loro punto di vista, sono così sintetizzabili:

- semplicità d'uso: una volta acquisite le credenziali di accesso ed effettuato il login al sito dedicato al *trial* a cui si partecipa, l'interfaccia di WIDEN guida l'utilizzatore attraverso tutti i passi previsti dal protocollo approvato per il *trial*;
- nessuna necessità di acquisto, installazione e configurazione di hardware o software dedicato;
- velocità (dipende naturalmente dalle caratteristiche della connessione di rete dell'utente, ma il tempo medio di caricamento di un caso - cioè 2 studi PET/CT, intorno ai 100-150 MB ciascuno, è di circa 4 minuti);

- sicurezza: uso del protocollo https per l'autenticazione e l'accesso al sito, dal momento che i dati sono trasferiti criptati grazie alla codifica a 256-bit tipo SSL;

- nessun vincolo geografico per i *reviewers*: l'uso di tecnologie web open source garantisce la possibilità di accesso a WIDEN da qualunque terminale, anche pubblico, da uno *smartphone* (iPhone, BlackBerry, Palm, Android) o da un *tablet* (iPad).

Come già sottolineato, l'unico requisito per l'utente è l'installazione di un web browser. Per i *reviewers*, inoltre, è necessaria la disponibilità di un programma di visualizzazione delle immagini. A questo proposito, WIDEN non impone nessuno strumento comune ai revisori, consentendo loro di operare con i programmi che utilizzano abitualmente. In questo modo, si ottiene un duplice ulteriore vantaggio: non è necessario che i revisori concordino sulla scelta di un unico strumento, non è necessario che modifichino le proprie abitudini e di conseguenza è presumibile che operino nelle condizioni per loro migliori per giungere ad una valutazione ottimale delle immagini.

L'accesso ai servizi è basato su una topologia a stella, in cui il portale web interagisce direttamente con tutti i nodi che partecipano al *trial* e agisce da centro di raccolta e smistamento dei dati e delle immagini.

4. WIDEN e la rete

Dal punto di vista dell'uso della rete, il tipo di applicazioni e soprattutto la dimensione tipica di uno studio multicentrico (10-100 centri) determina un impatto ridotto in termini di flusso di dati.

A titolo di esempio, possiamo ancora una volta considerare il TC HD0607 che, in circa due anni di operatività in 38 sedi partecipanti (in 3 continenti), ha coinvolto 236 pazienti, per ciascuno dei quali sono state trasferite 2 PET/CT di dimensione media di 160 MB, corrispondenti ad un upload complessivo di circa 38 GB ed un download di circa 230 GB (i re-

viewers del panel sono 6).

Dunque, anche nell'ipotesi che in futuro tutti i TC in corso che utilizzano tecniche di *imaging* adottassero questo approccio, abbandonando il trasferimento delle immagini su supporto fisso (CD, DVD), l'impatto sul traffico in rete sarebbe dunque trascurabile: a livello mondiale, si può stimare un flusso di dati dell'ordine di 20-50 TB l'anno.

Tuttavia, il vero valore aggiunto dell'approccio presentato, per cui l'infrastruttura di rete è indispensabile, è la velocità di valutazione delle immagini da parte di un panel di esperti non localizzati in una unica sede, un requisito necessario per poter variare rapidamente la terapia nei casi opportuni. Nel caso di HD0607, sono 32 i pazienti ai quali è stata modificata la terapia in conseguenza del risultato delle revisione della *interim-PET*: un approccio tradizionale non lo avrebbe consentito, riducendo l'efficacia della terapia e le prospettive di vita dei pazienti.

In sintesi dunque, grazie alla funzionalità messa a disposizione dalla rete ed alle caratteristiche di sicurezza garantite dal protocollo https, con un impatto estremamente ridotto sui flussi complessivi di dati sarebbe possi-

bile un significativo miglioramento nella procedura di validazione di nuovi farmaci e/o approcci terapeutici in termini di tempi, costi e di funzionalità dei protocolli.

5. WIDEN e le prospettive future

Il team di WIDEN si propone, nell'immediato futuro, di proporre a nuovi studi in fase di progettazione l'approccio sviluppato per HD0607 e già adottato da altri quattro TC.



Piergiorgio Cerello
cerello@to.infn.it

Laureato in fisica nel 1989, è ricercatore presso l'INFN dal 1995.

Tra i fondatori del progetto ALICE, ha contribuito a definirne il *computing model*. Coordina dal 2004 il progetto MAGIC-5, che sviluppa algoritmi di ricerca automatica di anomalie in immagini mediche.

LinkO: un'infrastruttura in fibra ottica per confronti remoti di tempo e frequenza ad alta accuratezza

C.Clivati², D.Calonico¹, F.Levi¹, A.Mura¹, G.A.Costanzo², A.Godone¹

¹Divisione di Ottica, Istituto Nazionale di Ricerca Metrologica-INRiM, ²Politecnico di Torino



Abstract. L'articolo descrive l'esperimento LinkO dell'Istituto Nazionale di Ricerca Metrologica INRiM di Torino, volto alla realizzazione di una struttura in fibra ottica per il trasferimento su lunga distanza di segnali ultrastabili di frequenza. Si tratta di una tecnologia innovativa e di largo impatto per i sistemi di disseminazione dei riferimenti di tempo e frequenza. LinkO necessita di un'infrastruttura in fibra ottica con un'architettura di rete come la futura GARR-X; allo stesso tempo costituisce per GARR-X un'applicazione innovativa, con notevoli ricadute a livello scientifico, tecnologico ed industriale.

In questo lavoro si descrive in particolare la realizzazione della sorgente laser a 1542 nm con stabilità relativa di alcune parti in 10^{-14} a 1 s e del sistema di compensazione del rumore di fase, aspetti fondamentali per il raggiungimento delle prestazioni desiderate. LinkO ha dimostrato finora una stabilità relativa di 10^{-15} a 1 s sul trasferimento di una frequenza ottica lungo 50 km di fibra in laboratorio.

1. Introduzione

Nella maggior parte dei processi tecnologici sono necessari accurati riferimenti di tempo e frequenza, pertanto ogni miglioramento nella realizzazione e nella disseminazione di questi ha un impatto a largo spettro sull'innovazione scientifica e tecnologica, in particolare in aree come le telecomunicazioni (velocità/capacità di trasmissione dati) o la navigazione satellitare (GPS, Galileo).

Gli Istituti Metrologici Nazionali, tra cui l'Istituto Nazionale di Ricerca Metrologica INRiM di Torino, possiedono approfondite conoscenze nella realizzazione di orologi atomici, nella generazione della scala di tempo e nella disseminazione dei riferimenti di tempo/frequenza. Gli orologi ottici di nuova generazione mostrano la capacità di scendere nettamente sotto la soglia di accuratezza relativa dei tradizionali campioni primari a microonde, basati su orologi a fontana di cesio, che è pari ad alcune parti in 10^{-16} . Invece, le possibilità di sincronizzazione remota di segnali di tempo/fre-

quenza hanno il loro stato dell'arte nei metodi satellitari a due vie (in particolare il *Two Way Satellite Time and Frequency Transfer - TWSTFT*), che al massimo consentono confronti a livello di parti in 10^{-16} dopo più di 20 giorni di misure continue e che quindi sono totalmente inadatti a sfruttare le potenzialità dei nuovi orologi atomici.

La necessità di sviluppo di nuovi metodi di sincronizzazione e confronto remoto che raggiunga adeguati risultati in termini di incertezza è stato riconosciuto quale obiettivo primario ai più alti livelli metrologici internazionali ed il solo metodo finora in grado di raggiungere gli obiettivi prefissati è il link ottico di frequenza su fibra ottica.

Progetti pionieri hanno dimostrato eccellenti risultati in Francia, Stati Uniti, Germania [1,2]. Lo sviluppo di network in fibra per il confronto di frequenza tra Istituti di Metrologia e Laboratori di Spettroscopia atomica sarà una priorità per gli Istituti Metrologici nel prossimo futuro, come già riconosciuto dall'Asso-

ciazione Europea degli Istituti Metrologici Nazionali (EURAMET) e già sono presenti concrete proposte per la costituzione di network transnazionali basati su fibre commerciali [3].

L'INRiM, che realizza e mantiene in Italia le unità campione del Sistema Internazionale della Convenzione del Metro, di cui l'Italia è firmataria, ha avviato l'esperimento LinkO, finalizzato a realizzare la disseminazione di frequenza basata sull'uso di infrastrutture in fibra ottica.

In considerazione di questi aspetti, GARR-X si configura come un'infrastruttura fondamentale per LinkO, che necessita di una rete completamente ottica per ottenere il trasferimento dei segnali. Infatti, la rete attuale GARR-G contiene nodi di trasduzione ottico-elettrico-ottico che distruggerebbero la coerenza di fase del segnale, impedendo il trasferimento di segnali di frequenza accurati. LinkO rappresenta dunque uno dei servizi innovativi che saranno offerti dalla nuova rete. Le applicazioni di un sistema fibrato di confronto e sincronizzazione vanno infatti ben oltre la metrologia e si estendono all'ambito fisico ed informatico, con un largo impatto anche sulle attività industriali ad alta tecnologia.

Il sistema di link ottico è composto da una stazione laser trasmittente, dall'infrastruttura in fibra ottica e dalla stazione di ricezione. Nella stazione trasmittente, si genera una radiazione laser ultrastabile, la cui frequenza è misurata rispetto a quella di un orologio atomico campione, nel dominio delle microonde, per mezzo di un pettine ottico di frequenza [4].

Una volta misurata la frequenza del laser ultrastabile, essa è inviata al laboratorio remoto mediante collegamento in fibra. Qui è ricevuta e convertita, ancora tramite un pettine ottico, in un segnale a radiofrequenza di uso comune nell'industria o nei laboratori, ad es. 10 MHz. La conversione eseguita dal pettine ottico non degrada le proprietà di accuratezza e stabi-

lità del laser in arrivo, consentendo di trasferire al segnale a radiofrequenza generato le prestazioni di un primario come se fosse effettivamente presente in laboratorio.

La procedura è illustrata nella figura sottostante.

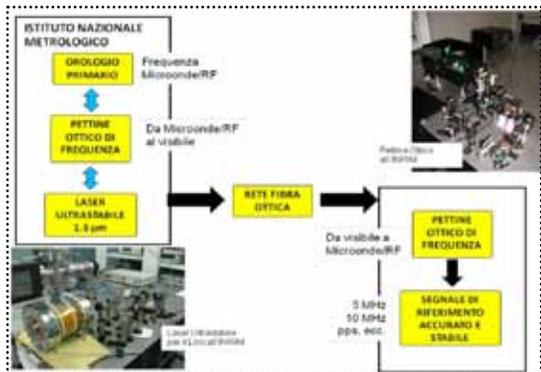


Fig. 1 Schema concettuale della disseminazione di segnali ultrastabili di frequenza con link in fibra ottica: la radiazione è misurata attraverso un pettine ottico rispetto ad un orologio primario e quindi trasferita al laboratorio remoto. Qui, sempre attraverso un pettine ottico, può essere riconvertita al dominio delle microonde ed utilizzata.

Per ottenere il trasferimento di frequenze ottiche in fibra a distanza di centinaia di km con il livello di stabilità richiesto, è necessario compensare anche il rumore di fase aggiunto dalla fibra stessa sulla portante ottica, in seguito a vibrazioni meccaniche e variazioni di temperatura e pressione ambientali. Il sistema di compensazione è basato sulla riflessione di parte del segnale che ha raggiunto il laboratorio di destinazione; nel doppio passaggio in fibra la radiazione accumula un rumore di fase che è corretto intervenendo sulla frequenza della radiazione inviata. Con questo sistema, la radiazione al laboratorio remoto risulta stabilizzata al livello del femtosecondo.

È necessario che anche la radiazione inviata in fibra abbia una stabilità relativa di frequenza al livello di 10^{-15} , per evitare che il rumore del laser si aggiunga a quello della fibra ottica, compromettendo l'efficacia della compensazione.

Le richieste di stabilità sulla frequenza della sorgente diventano tanto più stringenti quanto maggiore è la distanza a cui si intende

trasferire il segnale. Pertanto, un passo fondamentale del progetto LinkO è la realizzazione della sorgente laser.

2. La sorgente ultrastabile

La sorgente è costituita da un laser in fibra a 1542 nm, la cui frequenza è stabilizzata su uno dei modi di risonanza di una cavità ottica Fabry-Pérot, mediante la tecnica di Pound-Drever-Hall [5]. Essa consiste nella modulazione in fase del segnale incidente sulla cavità, tramite un modulatore elettro-ottico (EOM). Il segnale riflesso, incidente su un fotodiodo (PD) è proporzionale allo scostamento della frequenza della sorgente dal modo risonante e quindi, alla correzione da apportare alla frequenza del laser per renderla uguale al valore di riferimento. Opportunamente filtrato ed integrato, il segnale è utilizzato per correggere la frequenza laser attraverso un modulatore acusto-ottico (AOM) pilotato da un oscillatore controllato in tensione (VCO) e da un attuatore piezoelettrico. Lo schema del sistema è riportato in figura 2.

Dal momento che la frequenza del modo di cavità è inversamente proporzionale alla lunghezza $L=10$ cm della stessa [6], per raggiungere la stabilità relativa di 10^{-15} è necessario che le variazioni di L si mantengano al di sotto di 10^{-16} m. Per questo, la cavità deve es-

sere isolata termicamente e meccanicamente dall'ambiente esterno. Essa è costituita in vetro a bassissimo coefficiente di espansione termica ed è posta all'interno di una camera da vuoto (pressione residua: 10^{-6} Pa). Uno schermo in rame favorisce un migliore isolamento termico (fig. 3) ed un sistema di controllo elettronico mantiene le variazioni di temperatura al di sotto di $1 \mu\text{K}$ agendo su due riscaldatori posti sulle pareti esterne della camera da vuoto.

L'isolamento dal rumore sismico è molto più critico da ottenere: poiché non è possibile isolare completamente la cavità dalle vibrazioni, si cerca di renderla insensibile ad esse agendo sul sistema di sospensione. La forma della cavità (fig. 3) è tale da minimizzare l'influenza del rumore sismico in direzione verticale; inoltre, essa è poggiata su particolari punti, detti di Airy, ovvero i punti di sospensione in grado di ridurre al minimo l'influenza delle vibrazioni sismiche sulla lunghezza dell'asse ottico [7]. Dimensione e posizione ottimali sono state determinate effettuando delle simulazioni software agli elementi finiti. Sistemando i piedini nella posizione ottimale si può ricavare una sensibilità alle vibrazioni verticali sufficientemente bassa da consentire la stabilizzazione al livello relativo di 4×10^{-16} a 1 s, ampiamente sufficiente a raggiungere gli obiettivi richiesti.

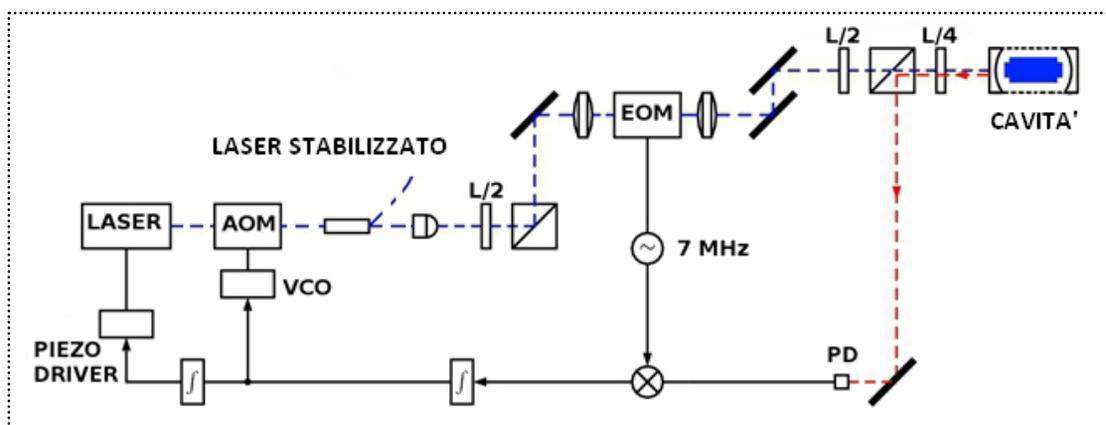


Fig. 2 Schema del circuito di controllo Pound-Drever-Hall. Il tratteggio indica i tratti di cammino ottico, la linea continua quelli di cammino elettronico. La radiazione è modulata in fase attraverso un EOM e fatta incidere sulla cavità. Il segnale riflesso incide su un fotodiodo, è demodulato, integrato ed inviato come correzione alla frequenza laser attraverso un AOM pilotato da un VCO. L/2 e L/4 identificano le lamine polarizzatrici utilizzate per modificare la polarizzazione della radiazione incidente e riflessa dalla cavità.

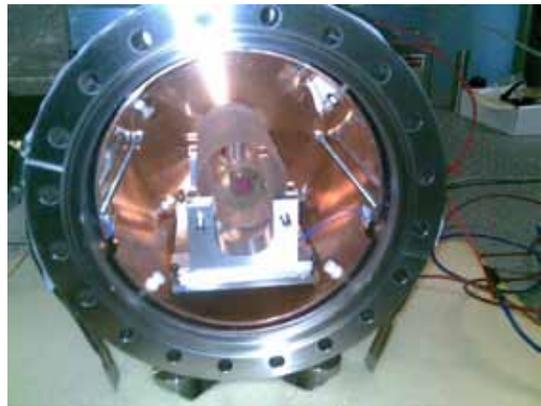
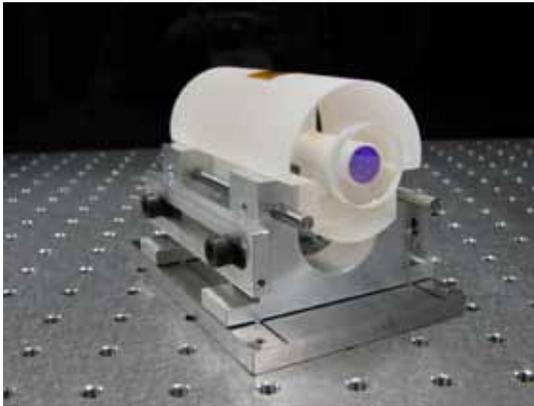


Fig. 3 La cavità posizionata sul sistema di supporto (a sinistra). La cavità all'interno della struttura di isolamento (a destra).

Per verificare le prestazioni ottenute, sono stati allestiti due sistemi indipendenti e si è analizzata la nota di battimento dei due laser stabilizzati sulle rispettive cavità. La larghezza di riga di questa è di 30 Hz; la stabilità relativa è pari ad alcune parti in 10^{-14} a 1 s in termini di deviazione di Allan, comunemente utilizzata come indicatore statistico in ambito metrologico [8] (fig. 4) ed è stata misurata in assenza del controllo attivo di temperatura. È atteso un miglioramento una volta che esso sarà attivato. Ulteriori miglioramenti riguarderanno l'accoppiamento ottico del modo laser a quello del risonatore e l'incremento della banda di aggancio dei laser alle rispettive cavità.

3. Il Link ottico

La radiazione stabilizzata in frequenza sulla cavità è destinata alla trasmissione in fibra al

laboratorio remoto. Qui costituisce un riferimento stabile, a condizione che sia compensato il rumore che perturba la stabilità della portante in seguito al passaggio in fibra. Al laboratorio di destinazione, per ottenere la compensazione, una parte della radiazione viene riflessa utilizzando un circolatore e fatta interferire con la radiazione emessa, secondo lo schema riportato in figura 5. Il modulatore acusto-ottico (AOM2) all'uscita della fibra ottica è utilizzato per spostare la frequenza della radiazione riflessa, al fine di distinguerla dalle possibili riflessioni spurie ai vari connettori.

Il segnale di battimento, sul fotodiodo PD1, contiene l'informazione relativa al rumore di fase accumulato durante il passaggio all'interno della fibra ottica in seguito a variazioni di temperatura e pressione e vibrazioni ambientali che ne alterano la lunghezza;

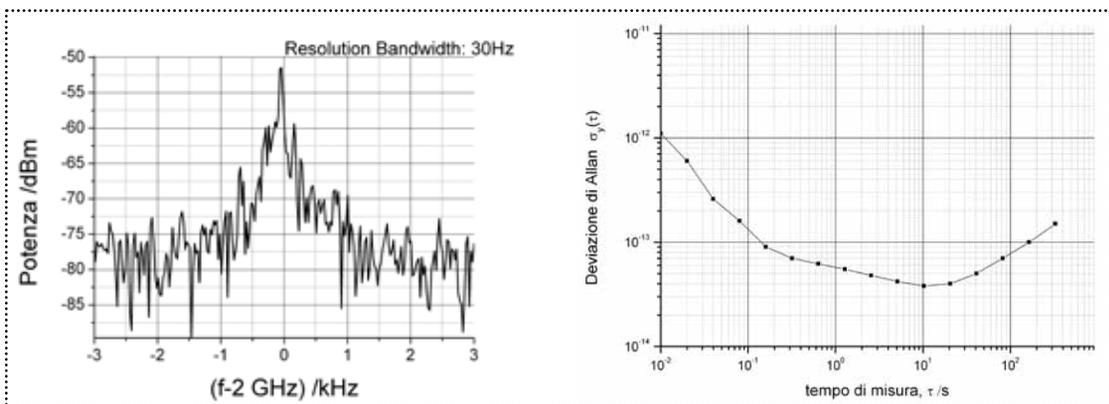


Fig. 4 A sinistra, la nota di battimento delle due cavità: la larghezza di riga è 30 Hz. A destra, la stabilità relativa della nota di battimento dei due laser stabilizzati sulle rispettive cavità.

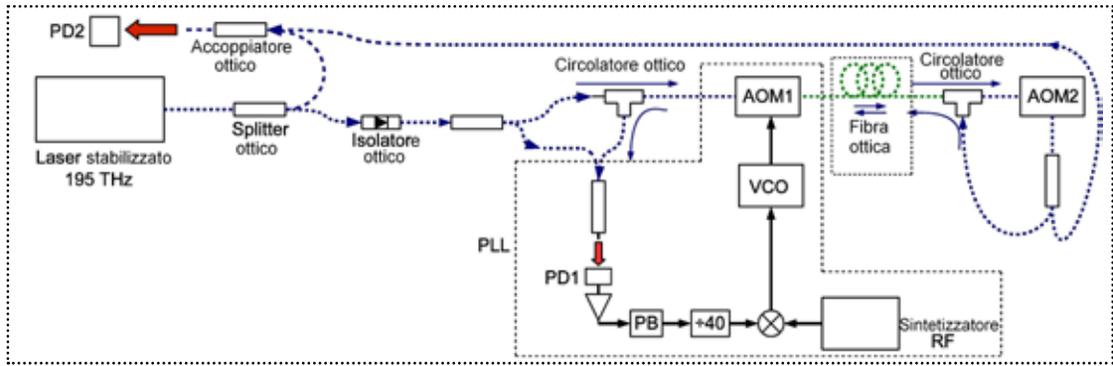


Fig. 5 Schema del link ottico. La radiazione viene inviata al laboratorio remoto; qui parte viene riflessa dopo essere stata spostata in frequenza grazie all'AOM2 e confrontata con la radiazione immessa in fibra. La nota di battimento ottenuta sul fotodiodo PD1 è amplificata, filtrata (PB indica il filtro passa-basso) e divisa. Dopo essere convertita alla tensione continua è inviata come correzione alla frequenza laser attraverso l'AOM1, pilotato da un VCO.

in particolare, utilizzando un divisore di frequenza N , in grado di mantenere il valore efficace del rumore di fase al di sotto di 1 rad, è possibile fare in modo che il segnale sul fotodiodo PD1 sia proporzionale al rumore stesso. Convertito alla tensione continua e opportunamente filtrato ed integrato, esso è utilizzato come correzione alla frequenza inviata in fibra tramite un AOM pilotato da un VCO. Per verificare le prestazioni raggiunte, si è ottenuta sul fotodiodo PD2 la nota di battimento fra la radiazione immessa in fibra e quella trasmessa all'altro capo, simulando in tal modo il laboratorio remoto.

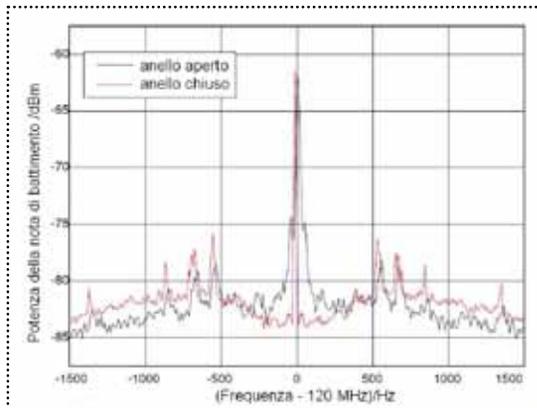


Fig. 6 Spettro della nota di battimento ottenuta sul fotodiodo PD2 ad anello aperto, in nero, e chiuso, in rosso, su bobina di fibra di 25 km (Larghezza di banda di risoluzione: 10 Hz). La riduzione del rumore attorno alla portante è evidente.

Per le misure effettuate finora sono state utilizzate bobine di fibra ottica di 25 km e 50 km all'interno del laboratorio, in previsione

di estendere il link su lunghezze maggiori ed all'esterno, su fibre stese. Le misure dimostrano una significativa riduzione del rumore. In figura 6 è riportata in nero la nota di battimento ottenuta sul fotodiodo PD2 ad anello aperto utilizzando una bobina di fibra di 25 km: la larghezza di riga è determinata dal rumore di fase aggiunto dalla fibra ottica. In rosso si riporta la nota di battimento quando viene attivato il circuito di controllo: la compensazione del rumore è evidente e la stabilità relativa raggiungibile è pari ad alcune parti in 10^{-17} in termini di deviazione di Allan (fig. 7).

In figura 7 si riporta in rosso la stabilità relativa ottenuta con bobine di 25 km, in blu la stabilità con bobine di 50 km. La stabilità relativa massima raggiungibile è superiore a 10^{-16} ; essa consente pertanto il confronto di sorgenti

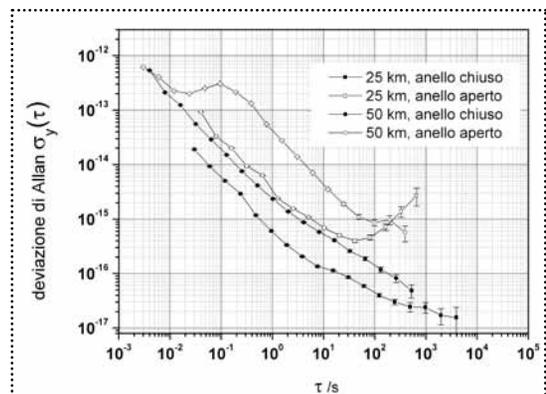


Fig. 7 Stabilità relativa del link ad anello aperto e chiuso su bobine di fibra di 25 km e 50 km. Quando il controllo viene attivato la stabilità ultima raggiungibile è superiore al livello relativo di 10^{-15} .

con questo livello di accuratezza.

4. Conclusioni e prospettive future

Si è descritto l'allestimento dell'esperimento LinkO presso l'INRiM ed i primi risultati ottenuti. Attualmente, la stabilità della sorgente consente di ottenere la compensazione del rumore sul link ottico fino al livello relativo di alcune parti in 10^{-17} .

Per un ulteriore miglioramento occorre in primo luogo stabilizzare la sorgente fino al livello relativo di 10^{-15} . Inoltre, sarà prioritario isolare termicamente e meccanicamente i rami dell'interferometro in fibra, in quanto si è dimostrato [9] che tratti di fibra non percorsi in doppio passaggio dalla radiazione, il cui rumore non è quindi compensato, possono limitare la stabilità del sistema. Inoltre, si prevede di adottare alcuni accorgimenti relativi all'ottica e all'elettronica del sistema, quali la compensazione della birifrangenza della fibra al fine di ridurre il rumore di ampiezza e la riduzione della conversione fra rumore di ampiezza e di fase.

Parallelamente all'ottimizzazione ed alla completa caratterizzazione del sistema realizzato, sarà possibile dedicarsi all'estensione di questa tecnica su fibre commerciali.

Riferimenti bibliografici

- [1] Kefelian et al., High-resolution optical frequency dissemination on a telecommunications network with data traffic, *Opt. Lett.* 34, 10, 2009
- [2] Grosche et al., Optical frequency transfer via 146 km fiber link with 10^{-19} relative accuracy, *Opt. Lett.* 34, 15, 2009
- [3] Lopez et al., Cascaded multiplexed optical link on a telecommunication network for frequency dissemination, *Optics Express*, 18,16, 2010
- [4] Cundiff et al., Colloquium: Femtosecond optical frequency combs, *Rev. Mod.Phys.* 75, 325–342, 2003

[5] Black, An Introduction to Pound-Drever-Hall laser frequency stabilization, *Am. J. Phys.* 69, 2001

[6] Yariv, Introduction to optical electronics, Holt, Rinehart and Winston, 1971

[7] Chen et al., Vibration induced elastic deformation of Fabry-Pérot cavities, *Phys. Rev. A* 74, 053801, 2006

[8] Allan et al., Statistics of atomic frequency standards, *Proc. Of the IEEE*, 54, 2,1966

[9] Williams et al., High stability transfer of an optical frequency over long fiber-optics links, *J. Opt. Soc. Am. B* 25, 8, 2008



Cecilia Clivati

c.clivati@inrim.it

Laureata in fisica delle tecnologie avanzate nel 2010, sta svolgendo il Dottorato di Ricerca presso la Divisione Ottica dell'INRiM, lavorando

alla realizzazione di un link in fibra ottica a basso rumore di fase che permetta il trasporto coerente di una portante a 1550 nm.

Utilizzo di Zabbix per il Monitoring di risorse di calcolo e sistemi di storage su WAN in ambito Grid Computing



A.Colavincenzo¹, A.Petricca², G.Bracco², S. Migliori², F.Simoni², S.Pecoraro², D.Giammattei², A.Rocchi², C.Sciò²

¹KELYON S.r.l., ²ENEA, Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile

Abstract. In questo articolo viene presentato il sistema di monitoring operante in ENEA-GRID [1], un'infrastruttura che integra le risorse per calcolo scientifico di ENEA, usando il paradigma del Grid Computing. Tale sistema è stato realizzato utilizzando il software Zabbix [2] che con le potenzialità della sua organizzazione gerarchica è in grado di coprire i vari siti di ENEA-GRID distribuiti su WAN (Wide Area Network). Particolare attenzione è stata riservata al monitoring delle componenti di rete, delle risorse e dei servizi utilizzati dagli utenti del calcolo scientifico. In tale contesto sono state sviluppate metodologie di monitoring dedicate alla misura della qualità di servizio delle componenti di trasferimento e condivisione dati, con particolare attenzione ai file system distribuiti.

1. Grid Computing su Rete WAN

Il concetto di Grid Computing si fonda sulla condivisione organizzata di risorse di calcolo distribuite geograficamente e appartenenti a domini amministrativi differenti. Per ottenere ciò deve essere definita un'architettura capace di gestire la condivisione di risorse informatiche tra diversi sistemi di calcolo, focalizzando l'attenzione sugli aspetti di sicurezza e confidenzialità, sulla condivisione di risorse e di dati e sulla gestione interna del sistema. In un sistema di Grid Computing sono fondamentali il controllo, la gestione e le performance della rete informatica distribuita. La funzionalità della Grid infatti è in relazione con le prestazioni fornite dalla rete stessa.

Il Grid Computing, applicato al concetto di calcolo scientifico su Wide Area Network, permette l'allocazione, la gestione e il processamento di una grande quantità di dati. Ciò comporta, quindi, la necessità da parte degli amministratori della Grid di possedere strumenti per controllare e rendere accessibili le risorse di calcolo nel miglior modo possibile.

2. ENEA-GRID

ENEA-GRID, il sistema di Grid Computing di ENEA, è composto dalle risorse che i 6 centri ENEA situati in tutta la penisola hanno a disposizione localmente e costituisce un unico sistema virtuale con una capacità globale di più di 30 TFlops distribuita su sistemi paralleli multi-piattaforma¹. Attualmente, la maggiore risorsa di calcolo di ENEA è il sistema HPC CRESCO [4] situato nel Centro Ricerche di Portici. La rete di connessione tra le varie sedi è fornita da GARR e i centri maggiori di Portici e Frascati hanno rispettivamente un PoP GARR² di 2Gbps e 1Gbps.

Le componenti principali che permettono ad ENEA-GRID di fornire all'utente finale questo unico sistema virtuale sono essenzialmente quattro:

- il gestore delle risorse LSF (*Load Sharing Facility*) Multicluster [5] per i meccanismi di scheduling dei job e della gestione delle risorse;

.....

¹ AIX/IBM SP, Linux x86/IA64/x86_64/, IRIX/SGI, Mac OSX/Apple G5.

² *Point of Presence*, indica uno dei nodi della rete telematica GARR

- il file system distribuito AFS (Andrew File System) attualmente disponibile in ambito open source sotto il nome di OpenAFS [6], strumento essenziale per la condivisione di dati su WAN;
- un'interfaccia grafica di accesso basata su NX [7];
- un'infrastruttura di autenticazione Kerberos 5 integrato con OpenAFS.

Recentemente è stato introdotto su diversi siti ENEA il file system parallelo GPFS [8], come strumento complementare ad AFS per la gestione dei dati, utilizzato attualmente in ambiente locale ma di cui è previsto lo sviluppo su WAN.

3. Esigenza del monitoring

La rete è in continua evoluzione, vi è quindi la necessità di garantire in ambito Grid la qualità dei servizi offerti all'utenza.

Il monitoring degli apparati diventa fondamentale per avere sia il controllo generale dell'infrastruttura che lo strumento per individuare parametri critici con la finalità di migliorare le prestazioni e le funzionalità del sistema. Il servizio reso dal monitoring dedicato deve quindi sia controllare la normale attività della rete, dei nodi di calcolo dell'infrastruttura e dei sistemi di storage utilizzati, sia verificare e valutare le prestazioni del completo sistema di Grid Computing. Questo semplifica la risoluzione dei problemi e le operazioni di gestione della rete.

4. La scelta di Zabbix

La scelta del software di monitoring Zabbix

è stata fatta poiché si adatta perfettamente alle caratteristiche richieste dall'ambientazione Grid Computing su WAN, infatti, una volta individuati i parametri prestazionali chiave di un determinato servizio offerto, la modalità "hierarchical distributed monitoring" (DM) di Zabbix [3] permette di avere una descrizione dettagliata del comportamento dei vari dispositivi, mediante la centralizzazione delle informazioni su di un Server-Master, che in ambiente server multi-livello si presta ad un processamento dei dati più funzionale.

Le principali caratteristiche di Zabbix al servizio della gestione degli apparati su rete WAN sono:

- monitoring distribuito e real time;
- scalabilità;
- monitoring via web;
- flessibilità (implementazione di script esterni);
- monitoring di apparati "agentless";
- capacità di adattamento a diversi sistemi operativi disponibili su ENEA-GRID: Linux, AIX HP-UX, MacOS, Solaris e Win32;
- possibilità di supportare diversi database: MySQL, PostgreSQL, SQLite e Oracle.

I vantaggi offerti da Zabbix risiedono nella possibilità di adattarsi perfettamente all'infrastruttura, agli apparati e alle tecnologie usate per il sistema Grid. Si possono creare ed importare degli script *ad hoc* e i MIB (Management Information Base) relativi agli apparati del sistema, componenti che vengono poi integrati nei template utilizzati da Zabbix per analizzare dettagliatamente gli aspetti critici dell'infrastruttura (fig.1). Tali template devono essere unificati e appartenenti ad un

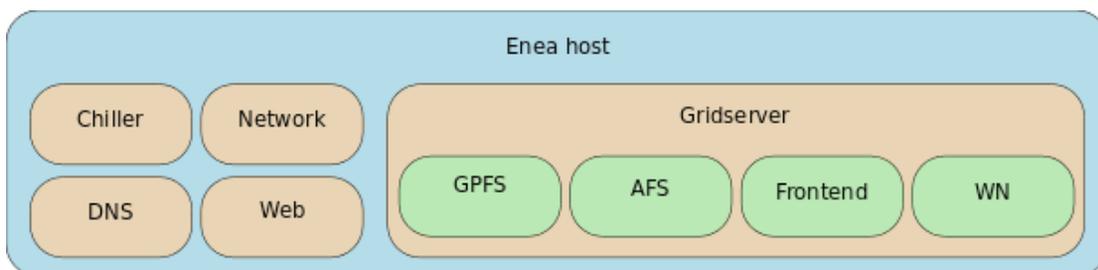


Fig. 1 Suddivisione logica dei template creati su Zabbix

determinato Host-Group (per esempio: front-end, dbserver, porte-IB, nodi di calcolo AFS o GPFS), indifferentemente dalla locazione dei siti dove viene installato il software Zabbix. Quest'ultima condizione infatti è necessaria per poter abilitare la modalità *distributed monitoring*.

Caratteristica importante di Zabbix è la possibilità di impostare delle visualizzazioni (nei *dashboard* e negli *screens*) differenti in base alla tipologia di eventi che si vuole monitorare, semplificando in questo modo notevolmente l'individuazione e lo studio dei problemi riscontrabili.

5. Monitoring dei file system paralleli (GPFS) e distribuiti (AFS) in ambito ENEA-GRID

L'affidabilità dei file system è una delle caratteristiche più importanti per la griglia ENEA, e Zabbix è stato lo strumento di aggregazione per i dati forniti dalle utilities di analisi prestazionali fornite dai due file system AFS e GPFS.

Host group	Disaster	High	Average	Warning	Information	Not classified
CRESCO - Fontend	0	1	0	7	10	0
CRESCO - GridServer	0	0	0	0	0	0
CRESCO - Management	0	1	0	0	0	0
CRESCO - Sensori	0	0	0	1	0	0
CRESCO - Sezione 1	0	2	0	4	34	0
CRESCO - Sezione 2	0	1	0	0	284	0
CRESCO - Sezione Speciale	0	1	0	2	11	0
Discovery CRESCO	0	5	0	13	339	0

Fig. 2 Status generale degli *host group* del Centro Ricerche ENEA di Portici

L'output delle *utilities* è stato integrato incapsulando gli strumenti a riga di comando in alcuni script di analisi che forniscono al software un pool di informazioni immediatamente fruibili, restituite all'utente attraverso schermate di facile consultazione (fig.2 e fig.3).

La criticità di tali sistemi si riflette nella specificità dei controlli effettuati e nella frequenza del loro utilizzo per garantire alla griglia ENEA il massimo standard possibile di disponibilità delle risorse per il calcolo scientifico

6. Punto di accesso unico e integrazione Zabbix-GINS di GARR

La possibilità di configurare Zabbix in "*distributed monitoring mode*" permette all'amministratore di rete e di calcolo di avere sotto controllo l'intera infrastruttura da un unico punto di accesso sotto un ambiente server multi-livello (Master & Child). Ciò permette di localizzare in maniera esaustiva e puntuale eventuali problematiche di rete e monitorare le prestazioni del calcolo su Wide Area, eliminando le condizioni di ambiguità tra i possibili fattori di errore o rallentamento del sistema di Grid Computing.

Valore aggiunto al sistema di monitoring Zabbix è la possibilità di integrazione con GINS [10], la suite software GARR che include gli strumenti per la diagnostica ed il *tracking* dei problemi di rete, i sistemi di acquisizione e visualiz-

Hosts	Availability	Disk	GPfs	GPFS Mem0	GPFS Mem1	GPFS Mem2	GPFS Mem3	GPFS Proj	GPFS Temp1	GPFS User	GPFS User0
creco-gpf1.portici.enea.it											
creco-gpf0.portici.enea.it											
creco-gpf3.portici.enea.it											
creco-gpf4.portici.enea.it											
creco-gpf5.portici.enea.it											
creco-gpf6.portici.enea.it											
creco-gpf7.portici.enea.it											
creco-gpf8.portici.enea.it											

Fig. 3 Triggers show di un pool di macchine GPFS

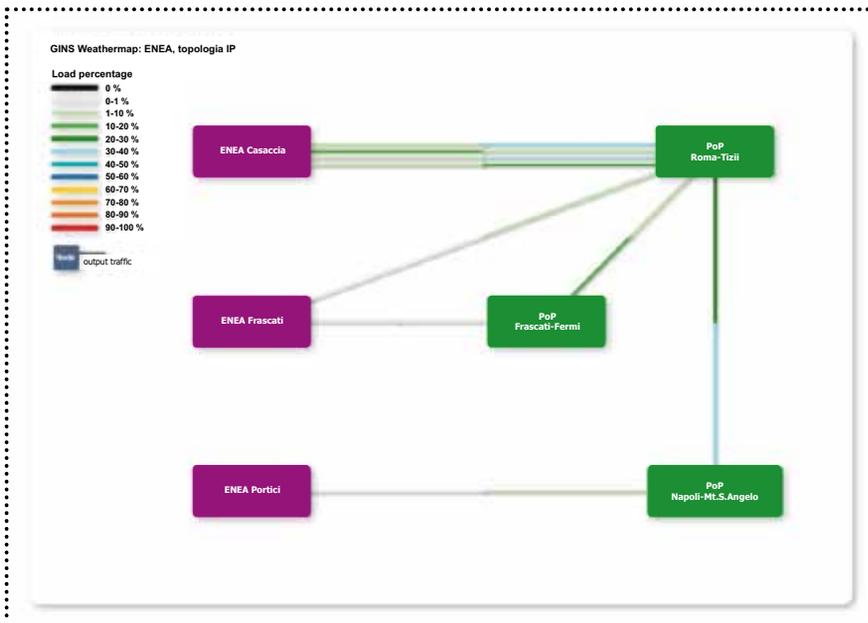


Fig. 4 Screen delle weathermaps GINS su Zabbix

zazione delle statistiche di traffico e di reportistica dell'attività di rete. L'interoperabilità tra i sistemi permette di integrare le *backbone traffic weathermaps* di GINS nelle viste locali di Zabbix e consente di controllare il grado di servizio dell'intera infrastruttura, sia dal punto di vista delle interconnessioni tra i centri ENEA sia per quanto riguarda la connessione dei centri ENEA con la rete GARR (fig.4).

Ottenere l'interazione tra sistemi di monitoring dell'ISP (Internet Service Provider) che fornisce la rete verso l'esterno e l'utilizzatore finale (ENEA) è di fondamentale importanza per avere il controllo globale dell'infrastruttura su cui si opera.

7. Infrastruttura in produzione

ENEA-GRID ha attualmente in funzione, nei principali centri di Frascati, Portici, Casaccia e Bologna, un sistema completo di monitoring dedicato tramite il software Zabbix (presenti circa 800 tra *host* di calcolo e apparati di rete, 4600 *items* e 5300 *triggers*), con sonde specifiche sia per gli apparati di rete e di calcolo (*availability*, *disk(free%)*, *load*, servizi di base, *share/free memory*) che per la valutazione prestazionale delle architetture utiliz-

zate (GPFS, AFS, LSF, InfiniBand). Un esempio è il monitoring del parametro RTT (Round Trip Time), che nel caso del file system distribuito AFS, a causa delle caratteristiche del suo protocollo di rete Rx [9], risulta essere l'elemento che determina le prestazioni del sistema.

L'implementazione del software di monitoring Zabbix mette in condizione gli amministratori e gli utilizzatori finali della griglia di avere sostanziali vantaggi in termini di immediatezza delle analisi di rete.

Uno strumento di monitoring fatto in questo modo permette agli amministratori di rete di individuare le criticità dell'infrastruttura e le cause degli eventuali problemi che pregiudicano le prestazioni. Gli strumenti dedicati agli utenti permettono loro invece di individuare gli scenari ottimali di esecuzione delle applicazioni secondo le loro richieste di traffico di dati. L'utilizzo, infatti, di risorse di calcolo distribuite nell'intera infrastruttura oppure concentrate in un singolo sito può dipendere criticamente dalla richiesta di accesso e/o dalla produzione dei dati.

I risultati ottenuti sono potenzialmente generalizzabili alle varie tecnologie impiegate nei sistemi di Grid Computing.

Riferimenti bibliografici

- [1] ENEA-GRID - www.eneagrid.enea.it
- [2] Zabbix - www.zabbix.com
- [3] Zabbix Manual - www.zabbix.com/documentation/1.8/complete

- [4] Progetto CRESCO - www.cresco.enea.it
- [5] LSF- www.platform.com/workload-management/high-performance-computing/lp
- [6] OpenAFS - www.openafs.org
- [7] NX Technology - <http://nomachine.com>
- [8] GPFS (IBM) - www-03.ibm.com/system/software/gpfs/index.html
- [9] Rx protocol specification draft - <http://web.mit.edu/kolya/afs/rx/rx-spec>
- [10] GINS - www.gins.garr.it



Antonio Colavincenzo

antonio.colavincenzo@enea.it

Nato a Napoli, ha studiato (e studia ancora) 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, di monitoraggio e gestione del sistema HPC Cresco e di ENEA-GRID.

NEXPreS: completamento ed evoluzione dell'e-VLBI

Mauro Nanni

INAF – Istituto di Radioastronomia di Bologna



Abstract. Le reti a larga banda stanno dimostrando la loro efficacia nelle osservazioni radioastronomiche eseguite con la tecnica dell'interferometria su lunghissima distanza, per questo la comunità radioastronomica internazionale si sta organizzando per sfruttare al meglio le opportunità offerte dalle reti della ricerca. Dopo il successo del progetto EXPreS, che ha mostrato la fattibilità tecnica dell'e-VLBI su base continentale, è stato ora finanziato il progetto NEXPreS che si pone l'obiettivo di completare la rete delle antenne e studiare nuovi sviluppi tecnici ed organizzativi. L'articolo illustra gli scopi del progetto e le attività che l'INAF intende svolgere grazie alla rete GARR-X.

Dalle prime osservazioni delle onde radio di origine extraterrestre, nel 1933 ad oggi, la radioastronomia si è impegnata nella costruzione di antenne sempre più sensibili e sempre più "direttive", capaci cioè di individuare con grande precisione la regione del cielo da cui provenivano i debolissimi segnali radio.

Fino agli anni '60 si è cercato di costruire antenne singole sempre più grandi, come il radiotelescopio di Arecibo di 300 metri di diametro o la "Croce del Nord" di Medicina di 500 metri di lunghezza, poi si è passati alla realizzazione di interferometri, cioè strumenti composti da più antenne che osservano simultaneamente la stessa zona di cielo. In questo modo è stato possibile creare radiotelescopi dalle dimensioni di alcuni chilometri che permettono di ottenere delle immagini radio del cielo con risoluzioni del decimo di secondo d'arco.

Utilizzando i cavi coassiali, si potevano posizionare le antenne alla distanza massima di alcuni chilometri e, con le tecnologie analogiche in uso negli anni '70, già su queste distanze si incontravano enormi problemi di interferenze e di attenuazione dei segnali. Per superare questi limiti è stato necessario realizzare una tecnica di interferometria off line che prevede di acquisire e registrare in formato di-

gitale i dati prodotti da ogni singola antenna e quindi "risincronizzarli" in un secondo momento utilizzando un computer.

Questa tecnica, che permette di costruire interferometri senza disporre di un collegamento fisico, presenta però problemi che sono legati alla grande quantità di dati prodotti da ogni antenna (flussi dell'ordine di centinaia di Megabit al secondo), alla necessità di introdurre nelle registrazioni dei marcatempi con precisioni esprimibili in parti su 10^{16} (secondi su milioni di anni), e alla conseguente richiesta di potenza di calcolo, richiesta che cresce in progressione geometrica con l'aumentare del numero di antenne che realizzano l'interferometro.

I primi esperimenti di "interferometria senza filo" sono stati condotti sul finire degli anni '60 ed hanno permesso di collegare antenne su continenti diversi, ottenendo immagini radio con risoluzioni dell'ordine del millesimo di secondo d'arco, risoluzioni che permettono di studiare con grande dettaglio i fenomeni che avvengono in regioni remote dello spazio. Dal '70 vi è stato poi un notevole sviluppo dei sistemi di memorizzazione passando dai nastri ai dischi magnetici, sono stati acquisiti orologi atomici sempre più stabili e precisi e costruiti correlatori (ovvero com-

puter dedicati alla sincronizzazione dei dati) sempre più potenti.

Gli ottimi risultati ottenuti con l'interferometria su lunghissima distanza (Very Long Baseline Interferometry) hanno favorito la collaborazione internazionale tra gli Istituti per definire standard operativi e risolvere i problemi tecnici, organizzativi e logistici. Sono stati creati consorzi su scala continentale come il VLBA negli Stati Uniti, l'ATNF-VLBI in Australia e l'EVN (European VLBI Network) e anche singole nazioni come Cina, Giappone, Corea, Sud-Africa hanno costruito antenne e correlatori partecipando alle osservazioni globali.

Già negli anni '90 la radioastronomia utilizzava le reti telematiche nelle osservazioni VLBI, soprattutto per distribuire ai sistemi di controllo le "schedule osservative", cioè i dati ed i parametri necessari al puntamento delle antenne ed all'esecuzione simultanea delle osservazioni. Nel 2000, con lo sviluppo delle dorsali in fibra ottica, si apriva la possibilità di trasmettere i dati acquisiti direttamente sulla rete ed effettuare la correlazione in tempo rea-

le. Si era ritrovato "il filo in grado di collegare le antenne" e nasceva l'e-VLBI o e-VLBI.

I vantaggi dell'e-VLBI rispetto al sistema tradizionale sono evidenti. La correlazione in tempo reale elimina gran parte dei problemi organizzativi e logistici relativi al trasporto dei supporti magnetici, problemi che portavano a ritardi di settimane o anche di mesi tra l'osservazione e la disponibilità dei risultati finali. La correlazione in tempo reale permette di riscontrare e correggere prontamente eventuali errori di configurazione delle antenne o dei ricevitori, di organizzare tempestivamente campagne osservative a fronte di fenomeni inaspettati quali esplosioni di supernove o gamma ray burst, fenomeni per i quali bisogna valutare rapidamente se e come effettuare ulteriori osservazioni anche con altri strumenti e su altre bande. Inoltre il collegamento permanente delle antenne alla rete si presta a campagne di monitoraggio che richiedono osservazioni brevi ma molto frequenti, come è il caso delle osservazioni geodinamiche.

Il fatto di disporre di apparati di trasmis-



Fig. 1 Rappresentazione grafica della maratona di 33 ore di osservazione congiunta organizzata in occasione della cerimonia di apertura dell'anno internazionale dell'astronomia nel mese di gennaio 2009.

sione realizzati su standard di mercato, apparati cioè che hanno avuto e continuano ad avere prestazioni sempre migliori a prezzi decrescenti, permette di acquisire i segnali su un numero sempre maggiore di canali radio, cioè di bande di frequenze. Così facendo si va ad aumentare la sensibilità delle stesse antenne, che è proporzionale alla radice quadrata della larghezza di banda totale: da qui deriva l'interesse della radioastronomia di disporre di bande trasmissive dell'ordine di 2-4 Gbps per singola antenna.

La comunità radioastronomica internazionale, ed europea in particolare, ha quindi deciso di impegnarsi sull'e-VLBI e, nel quadriennio 2006-2009, ha ottenuto i finanziamenti per il progetto EXPReS (*Express Production Real-time e-VLBI Service*) [1]. Con EXPReS ci si è focalizzati sul collegamento in fibra ottica delle antenne alle dorsali della ricerca ed allo studio di protocolli che garantissero trasferimenti costanti di flussi di dati con velocità di almeno 512 Mbps. In Italia è stato collegato il radiotelescopio di Medicina che partecipa regolarmente alle sessioni e-VLBI dal maggio 2006. Nel 2009, l'anno internazionale dell'astronomia è stato inaugurato da una sessione e-VLBI di 33 ore durante le quali i radiotelescopi americani, europei, sud africani, cinesi e australiani hanno osservato una radiosorgente inviando dati via rete a 512 Mbps e 1 Gbps al correlatore JIVE in Olanda (*fig. 1*).

Visto il successo di EXPReS il consorzio EVN ha proposto nel 2010 un nuovo progetto che è stato finanziato per 3,5 milioni di euro dalla Comunità Europea. Il progetto, chiamato NEXPReS (*Novel EXplorations Pushing Robust e-VLBI Services*) [2], si pone l'obiettivo di potenziare l'e-VLBI completando la rete dei radiotelescopi, incrementando la banda trasmissiva, studiando tecniche di correlazione distribuita per sostenere un maggiore carico computazionale e realizzando sistemi di archiviazione dei dati presso le antenne per poter rieseguire ed ottimizzare le correlazioni via rete.



Fig. 2 Il radiotelescopio SRT in costruzione a Pranu Sanguni, località nel comune di San Basilio, a 35 km da Cagliari.

L'Istituto Nazionale di Astrofisica (INAF) dispone dei radiotelescopi di Medicina e Noto, mentre una nuova antenna di 64 metri di diametro (SRT) è in avanzata fase di costruzione in Sardegna [3]. Queste antenne sono state costruite lontane dai grandi centri abitati per evitare il rumore elettromagnetico delle città e delle industrie e questo rende più difficile e costoso realizzare i collegamenti in fibra ottica alla rete GARR. Per l'antenna di Medicina si è avuta la fortunata coincidenza della realizzazione della rete regionale Lepida che ha messo a disposizione gratuitamente una fibra per il collegamento al PoP di Bologna, nodo che nel 2006 era già in grado di offrire collegamenti a larga banda. Con la realizzazione di GARR-X, l'INAF intende acquisire il collegamento in fibra di 135 Km tra l'antenna di Noto ed il PoP di Catania, mentre la Regione Sardegna ha stanziato i fondi per il collegamento di 40 Km tra SRT ed il PoP di Cagliari.

Il progetto NEXPReS, oltre a incrementare il numero delle antenne che partecipano alle osservazioni e-VLBI e a sperimentare bande trasmissive fino a 4 Gbps, che è la velocità offerta dalla nuova generazione di ricevitori digitali Digital BaseBand Converter, si pone l'obiettivo di studiare nuove metodologie osservative e nuove tecniche di elaborazione dei dati.

Quando si effettua una osservazione VLBI si va a produrre un'immagine di una regione del cielo dalle dimensioni di alcuni centesi-

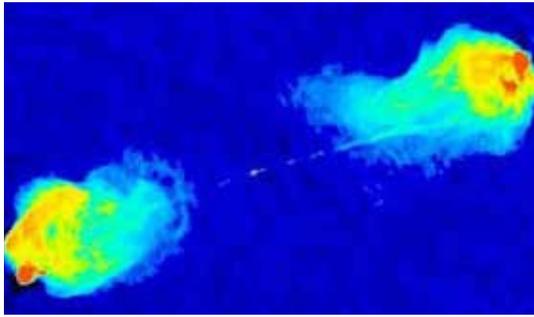


Fig. 3 Getti radio nella radiosorgente 3C120

mi di secondi d'arco, mentre le antenne hanno registrato i segnali radio da aree del cielo molto maggiori, aree dalle dimensioni di primi d'arco. Cambiando i parametri di correlazione, dallo stesso set di dati si potrebbe ottenere un diverso puntamento elettronico. È il processo di correlazione che seleziona la zona della quale ottenere un'immagine ad altissima risoluzione all'interno dell'area, molto più ampia, acquisita dalle antenne. Ripetendo più volte la correlazione con differenti parametri, si possono ottenere più immagini ed è questo il concetto che sta alla base del "multi beam" elettronico.

Un'osservazione di oggetti deboli può durare decine di ore producendo più di 10 Terabyte al giorno per antenna; la correlazione è un processo molto oneroso dal punto di vista computazionale e gli attuali correlatori VLBI sono costosi supercomputer dedicati in grado di correlare in tempo reale al massimo una decina di antenne per volta. Negli ultimi anni si stanno sperimentando correlatori software (come DiFX [4]) in grado di girare su cluster di calcolo paralleli. Questi sistemi sono molto più economici ma so-

no, in quanto a prestazioni, confrontabili con gli attuali "correlatori hardware". Se si vogliono studiare più oggetti dalla stessa osservazione, ottenendo il "multi beam in real time", sono necessari correlatori estremamente potenti o un sistema di elaborazione basato su cluster distribuiti che operano in parallelo ognuno su un campo diverso. Oppure si può semplicemente ripetere la correlazione più volte, ma in questo caso è necessario abbandonare la filosofia del real-time e salvare su disco i set dei dati per poterli rielaborare.

Il progetto NEXPREs prevede di allestire alcuni centri di storage di alta capacità (dell'ordine del Petabyte) e ad alta velocità (fino a 4 Gpbs) presso le antenne e presso i maggiori correlatori. Nello stesso tempo si studieranno nuove tecniche di correlazione per l'elaborazione dei dati in parallelo, valutando se sia più opportuno distribuire su diversi elaboratori i differenti canali (bande di frequenza), oppure le diverse frazioni di tempo osservativo o ancora sottogruppi di antenne. Si tratta cioè di capire come realizzare una griglia per la correlazione distribuita sulla rete a larga banda.

L'impegno dell'INAF nel progetto è quel-

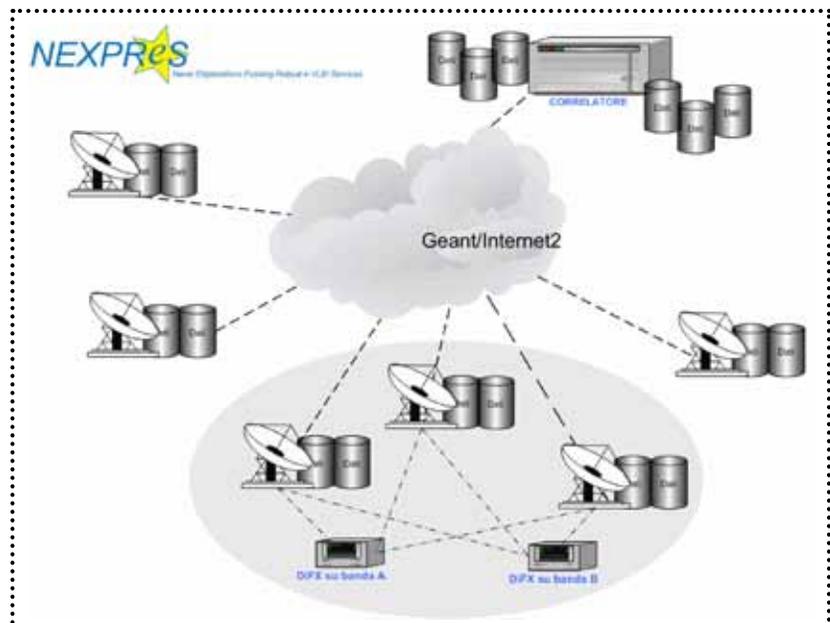


Fig. 4 Il progetto NEXPREs

lo di completare la rete nazionale della radioastronomia, di allestire un sistema di storage ad alta capacità ed un correlatore software in grado di servire 4-5 antenne e quindi partecipare allo studio ed allo sviluppo dei sistemi per il trasferimento, la gestione e l'elaborazione dei *dataset* osservativi prodotti dalle antenne. Un impegno che ci porta ad affrontare le problematiche relative all'efficienza dei supporti e dei file system, allo sviluppo di protocolli di trasmissione adeguati ai nuovi ricevitori digitali e anche all'analisi degli strumenti e dei concetti che il GRID e IPv6 mettono a disposizione.

La comunità radioastronomica sarà impegnata nel prossimo decennio nella realizzazione di SKA (Square Kilometer Array), un interferometro costituito da centinaia di singole antenne che avranno una superficie totale di un chilometro quadrato [5]. Le antenne saranno collocate in aree desertiche dell'Australia o del Sud-Africa su distanze di migliaia di chilometri e collegate con fibre ottiche. I problemi che oggi si stanno affrontando per l'e-VLBI sono in gran parte gli stessi problemi che dovranno essere risolti per SKA ed è per questo diciamo che NEXPREs è anche un

progetto che prepara la radioastronomia europea a SKA.

Riferimenti bibliografici:

- [1] www.expres-eu.org
- [2] www.nexpres.eu
- [3] www.srt.inaf.it
- [4] <http://astronomy.swin.edu.au/~adeller/software/difx/>
- [5] www.skatelescope.org



Mauro Nanni
nanni@ira.inaf.it

Laureato in fisica presso l'Università di Bologna nel 1980. Lavora all'Istituto di Radioastronomia dell'INAF. Ha sviluppato software per l'analisi dei dati del radioastronomici, *packages* grafici e database astronomici. Dal 2001 lavora alla realizzazione dell'infrastruttura di rete per la radioastronomia, partecipando ai progetti EXPReS e NEXPREs di cui è coordinatore nazionale.

FARO: accesso web a risorse remote per l'industria e la ricerca

A. Rocchi, S. Pierattini, G. Bracco, S. Migliori, F. Beone, C. Sciò, R. Ciavarella, A. Petricca

ENEA, Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile



Abstract. Lo scopo principale del sistema software FARO (*Fast Access to Remote Objects*) è quello di centralizzare l'accesso ad impianti, risorse e servizi cruciali consentendone la fruizione attraverso il web. FARO è stato sviluppato nell'ambito di un'infrastruttura di ricerca con possibili applicazioni anche in ambito industriale ed è il risultato di un processo di *software integration*, in cui le componenti coinvolte sono in parte open source e in parte sviluppate internamente.

Il sistema consente di accedere velocemente all'intera infrastruttura ICT, permettendo l'interconnessione con qualsiasi piattaforma di ricerca e sviluppo anche distribuita geograficamente e garantendo la sicurezza dei dati coinvolti nelle transazioni.

Un ulteriore fattore di rilievo è costituito dalla semplicità di utilizzo: l'unico strumento necessario per l'accesso a FARO è un comune web browser.

Nell'articolo saranno descritti l'architettura del sistema e i dettagli su *case study* di implementazioni di produzione in importanti ambiti internazionali.

1. Generalità sull'infrastruttura di calcolo ENEA-GRID

FARO è stato ideato, progettato ed implementato all'interno dell'infrastruttura di Grid Computing di ENEA, l'Agenzia nazionale per le nuove tecnologie, l'ambiente e lo sviluppo economico sostenibile. Tale sistema complesso, denominato ENEA-GRID/CRESCO [1], mette a disposizione di ricercatori e soggetti industriali e/o istituzionali anche esterni un'ampia serie di *facilities* per calcolo scientifico e servizi.

La potenza di calcolo, che complessivamente si può valutare in circa 30 TFlops di picco, è erogata grazie all'armonizzazione di risorse eterogenee sia hardware che software (Linux, IBM AIX, SGI, Apple Mac, Microsoft Windows, ecc.), orientate sia ad HPC (*High Performance Computing*) che ad HTC (*High Throughput Computing*).

In particolare le risorse HPC fanno uso di interconnessioni ad alta banda e bassa latenza

come Infiniband [2].

La distribuzione delle informazioni è effettuata su Wide Area Network (12 centri di ricerca, 6 Centri Elaborazioni Dati) e la garanzia dell'omogeneità dell'ambiente computazionale e della sicurezza nel trasporto dei dati sono fornite grazie alla presenza del file system distribuito OpenAFS [3], integrato con un'infrastruttura di autenticazione basata su Kerberos 5 [4].

All'interno dell'articolo verrà descritto in dettaglio come FARO si integra all'interno di ENEA-GRID.

2. Architettura

Le componenti software principali che interagiscono allo scopo di realizzare gli obiettivi funzionali di centralizzazione dell'accesso all'infrastruttura ICT sono:

- la tecnologia NX;
- una *Graphical User Interface* (GUI), scritta in linguaggio Java, personalizzata per l'uten-

za specifica;

- alcuni moduli accessori che implementano la comunicazione effettiva con le risorse esposte.

2.1 La tecnologia NX

La tecnologia NX, sviluppato dalla software house italiana NoMachine [5], usa un set di sistemi proxy che, operando su diversi livelli del protocollo X11 tramite tecniche di compressione dei dati, ottimizzazione della TCP window, e gestione adattativa della velocità di trasferimento, promette una *user experience* di accesso remoto caratterizzata da tempi di risposta minimi, anche in condizioni di rete limitate [6].

NX lavora in modalità client/server, e il protocollo sottostante a tutte le implementazioni è SSH (*Secure Socket Shell*). Tutte le implementazioni di FARO si appoggiano sulla versione open source del server (FreeNX [7]), e realizzano l'accesso web tramite il client NX-Web Companion, distribuito dalla stessa NoMachine sotto forma di applet Java [8].

2.2 La GUI di FARO

La GUI Java realizza il centro operativo attraverso il quale l'utente impartisce comandi e accede alle risorse esposte dall'infrastruttura ICT.

Ulteriori moduli software si occupano della gestione della molteplicità architetturale: risorse omologhe implementate su sistemi differenti richiedono *policy* di accesso e invocazione diversificate. L'opportuna gestione di questa varietà computazionale garantisce omogeneità e trasparenza rispetto alla complessità dell'infrastruttura.

2.3 Working model

Per chiarire meglio, consideriamo il *working model* di FARO riportato in figura 1. Si evince come FARO operi su tre livelli di astrazione principali:

- un utente che voglia usufruire di una qualunque risorsa esposta dall'infrastruttura di calcolo effettua la sottomissione delle sue credenziali di accesso, collegandosi ad un

portale di accesso web che pubblichi il *Web Companion*.

Le credenziali utente vengono trasferite dal *Web Companion* a un *Master Server*, che le propaga verso uno *slave server* (scelto in *round robin* tra un pool di macchine poste reciprocamente in configurazione *load balanced*); l'autenticazione tra *Web Companion* e *Master Server* avviene tramite SSH/*Public Key Exchange*, mentre l'accesso agli *slave* usa effettivamente la combinazione di parametri login/password fornita dall'utente (sempre appoggiandosi sul protocollo SSH, e, in particolare, sulla sua versione open source OpenSSH).

Se la procedura di login si completa con successo, lo *slave* scelto provvede a lanciare la GUI e a creare un percorso di ritorno per i pacchetti trasferiti dagli *end-point* coinvolti nella comunicazione. L'interfaccia di FARO, che è dunque in esecuzione sullo *slave*, viene recapitata sullo schermo dell'utente grazie a questo canale di collegamento.

- La GUI Java presenta all'utente le risorse disponibili nell'infrastruttura, eventualmente in combinazione con meccanismi di *black/white list*. Le applicazioni, i dispositivi e le piattaforme disponibili nell'infrastruttura sono elencate staticamente all'interno di appositi file di configurazione, che contengono le informazioni necessarie per identificare la metodologia di accesso sup-



Fig. 1 Working model di FARO

portata dalla risorsa in oggetto.

La presenza di tali file garantisce la possibilità di fornire interfacce mirate per target di utenza.

- Infine, la GUI si preoccupa di fornire all'utente opportuno feedback, relativamente allo stato della richiesta della risorsa desiderata.

La gestione fattiva dell'accesso ad un device viene gestita, in accordo con le direttive di configurazione, da alcuni moduli software accessori: risorse eterogenee richiedono politiche e strategie di invocazione differenziali (ad esempio: software complessi possono aver bisogno di utilizzare funzionalità avanzate di opportuni *job scheduler*, altrimenti non necessarie). Tali componenti software concorrono ad informare la GUI sullo stato della richiesta di una risorsa da parte dell'utente.

L'interfaccia, in un esempio di come viene visualizzata dall'utente, è mostrata in figura 2.



Fig. 2 FARO Graphical User Interface

2.4 NX Watchdog

La tecnologia NX supporta nativamente alcune funzionalità avanzate come quella di *session resume*: tramite complessi meccanismi di *process reparenting*, un utente può essere in grado di disconnettersi dalla sua sessione attuale e riprenderla in qualunque altro momento, trovandola esattamente nello stato in cui essa si trovava all'atto della disconnes-

sione. Un *drawback* collegato a questa funzionalità riguarda possibili incoerenze di stato tra server e client, che rendono impossibile effettuare connessioni successive. Per rimediare, nell'ambito dello sviluppo di FARO è stato progettato e implementato un tool, NX Watchdog, che consente di mantenere la coerenza di processi e sessioni sul server.

2.5 Software integration

È importante denotare, ancora una volta, che FARO è composto da moduli software eterogenei, caratterizzati da *licensing* misto. FreeNX costituisce l'implementazione open source del server NX. Il codice sorgente è stato sottoposto ad una serie di modifiche, in modo da perfezionare l'integrazione con l'infrastruttura.

La comunicazione con i servizi FreeNX è gestita tramite l'*applet* proprietario NX Web Companion, distribuito da NoMachine secondo una *policy closed source*. Il protocollo nativo di trasporto dell'informazione è SSH.

La *User Interface* di FARO, *entry point* dell'utenza verso l'infrastruttura ICT aziendale, è stata sviluppata internamente in ENEA. Analogamente, anche NX Watchdog è il risultato di un'attività interna ENEA di *software design & implementation*.

3. Esempi di implementazioni

FARO è attualmente in produzione in più di un'organizzazione e viene utilizzato in vari ambiti che spaziano dalla chimica computazionale all'archeologia quantitativa. Esporremo brevemente le esperienze più significative.

3.1 Implementazioni di produzione

3.1.1 ENEA

All'interno di ENEA, FARO offre ai ricercatori una piattaforma intuitiva, immediata ed efficace per l'accesso all'infrastruttura, la sottomissione di job di calcolo, l'elaborazione grafica remota e l'utilizzo in genere di una vasta gamma di architetture hardware e

software.

Grazie all'utilizzo dei meccanismi di personalizzazione dell'interfaccia, FARO ha consentito ad ENEA di promuovere e potenziare lo sviluppo dei cosiddetti laboratori virtuali [9].

Un laboratorio virtuale è una sorta di *meeting point* costituito da aree dedicate ai ricercatori afferenti ad un determinato progetto. La GUI di FARO presenta, in questi casi, solo le risorse ritenute utili all'espletamento dei task specifici di ricerca.

3.1.2 EFDA

Un'ulteriore versione di produzione, che ha già ricevuto ottime valutazioni nella sola fase di analisi preliminare, è resa disponibile ai ricercatori EFDA (*European Fusion Development Agreement* [10]). EFDA usa FARO per lavorare su attività che spaziano dall'analisi alla simulazione, fino allo sviluppo software.

3.2 Progetti di ricerca

3.2.1 CPMD Job Submitter per G-Lite

È stata sviluppata una versione dimostrativa di FARO per la gestione di job di calcolo CPMD [11] su sistemi Grid Computing basati sul *middleware* G-Lite.

L'oggetto del lavoro [12] ha previsto la realizzazione di un'estensione alla GUI di FARO (fig.3), in modo che fosse possibile seguire le varie fasi del processo (dalla sottomissione fino al recupero degli output) in maniera intuitiva e intelligente.

3.2.2. ARK3D

ARK3D [13] (*The ENEA-GRID infrastructure for the Remote 3D*) è l'infrastruttura realizzata con il progetto CRESCO e integrata nella GRID ENEA per la grafica 3D remota.

Utilizza la tecnologia FARO e il "*Remote Visualization tools*" per eseguire le applicazioni grafiche (librerie OpenGL) sull'acceleratore grafico 3D (schede grafiche NVIDIA) per il *rendering*.

L'utente accede all'applicazione e al modello 3D (generici) senza dover eseguire in-

stallazioni e pesanti download, garantendo la protezione della proprietà dell'applicazione e del dato.

Utilizzando l'infrastruttura ARK3D sono stati sviluppati alcuni progetti tra i quali 3DWS [14], un repository di modelli 3D del patrimonio culturale, ad alta definizione, integrato con un database documentale accessibile tramite web.

I modelli 3D sono visualizzabili in alta definizione tramite l'infrastruttura ARK3D proteggendone il dato originale. "*Remote 3D Rendering Project - Rome Reborn on CRESCO Machine*" [15] è uno studio sperimentale per la visualizzazione Remota di un modello 3D, elaborato per rappresentare la città di Roma del 320 d.C. sviluppato utilizzando l'infrastruttura ARK3D.

4. Conclusioni

FARO è un software fortemente orientato a garantire centralizzazione nell'accesso alle risorse dell'infrastruttura ICT, introducendo semplicità ed immediatezza nell'esecuzione di task anche complessi.

Il paradigma secondo cui FARO è stato progettato è quello definito *Anything, Anywhere*: ovvero un sistema in grado di proiettare l'utilizzatore, qualunque sia la sua posizione geografica, all'interno dell'infrastruttura aziendale, consentendogli di usufruire di qualsiasi ri-

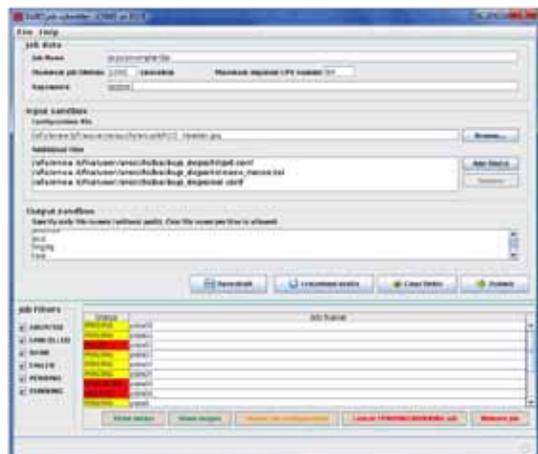


Fig. 3 FARO e griglie computazionali GLite

sorsa esposta dall'ambiente aziendale stesso, indipendentemente dalla sua complessità e garantendo una *user experience* positiva.

Il risultato dell'integrazione software, che coinvolge sistemi standard e open source, garantisce massima compatibilità nell'ambiente computazionale esistente e sicurezza nel trasporto dei dati.

FARO si adatta all'attività di gruppi di lavoro, permettendo all'utenza di specializzare l'interfaccia con le sole risorse d'interesse ai fini dell'esecuzione dell'attività lavorativa. In tal senso, FARO costituisce un elemento naturale di sistemi che integrano ed implementano paradigmi di Grid e Cloud Computing.

Riferimenti bibliografici

- [1] www.eneagrid.enea.it - www.cresco.enea.it
- [2] www.intel.com/technology/infiniband
- [3] www.openafs.org
- [4] <http://web.mit.edu/kerberos>
- [5] www.nomachine.com
- [6] www.nomachine.com/documentation/html/intr-technology.html
- [7] <http://freenx.berlios.de>
- [8] www.nomachine.com/download-companion.php?os=linux
- [9] I Laboratori Virtuali di ENEA – F. Beone, A. Rocchi et al. – Poster Conferenza GARR 2010, Torino

[10] www.efda.org

[11] www.cpmid.org

[12] User-friendly access to computational services with FARO portal for EGEE and ENEA-GRID - A. Rocchi et al. - 5th EGEE User Forum, Uppsala, Sweden

[13] www.ark3d.enea.it

[14] 3DWS - 3D Web Service Project - Abate et al. - CAA 2010 Computer Applications and Quantitative Methods in Archeology.

[15] www.afs.enea.it/project/graf3d_roma



Alessio Rocchi

alessio.rocchi@enea.it

Laureato in ingegneria informatica. Dopo circa 10 anni spesi come ingegnere del software freelance, approda in ENEA dove svolge il ruolo di *software architect* e *system integrator*. I suoi interessi di ricerca vertono sui paradigmi di GRID e Cloud computing.

WNoDeS: un servizio per la gestione di infrastrutture condivise Cloud e Grid

D. Salomoni¹, A. Chierici¹, A. Italiano¹, E. Ronchieri¹, P. Solagna²

¹INFN-CNAF, ²EGI.eu



Abstract. Nei Centri di Calcolo di medie e grandi dimensioni l'ottimizzazione dell'utilizzo delle risorse è essenziale. In questo articolo vengono descritte le motivazioni e l'implementazione di un sistema integrato di accesso a infrastrutture Grid e Cloud tramite tecniche di virtualizzazione, detto WNoDeS (*Worker Nodes on Demands Service*). WNoDeS, sviluppato dall'Istituto Nazionale di Fisica Nucleare, crea in modo dinamico e trasparente Virtual Machines (VMs) opportunamente configurate per soddisfare le richieste dell'utenza. Tale sistema è progettato per essere strettamente accoppiato con un Local Resource Management System (LRMS); WNoDeS assicura quindi la scalabilità e le performance tipiche dei moderni LRMS e permette una completa integrazione con le politiche di allocazione delle risorse, di monitoraggio, di *accounting* e di sicurezza già presenti in un Centro di Calcolo. Queste VMs possono essere usate sia per l'esecuzione di *batch jobs* tradizionali sia per la fornitura di risorse di tipo Cloud.

1. Grid e Cloud Computing

Le infrastrutture di calcolo distribuito devono fornire tipicamente soluzioni per l'identificazione, la classificazione, la selezione e l'utilizzo delle risorse. Devono inoltre fornire e rispettare contratti relativi al livello di servizio (*Service Level Agreements*) e assicurare che l'accesso alle risorse sia controllato da metodi appropriati per l'autenticazione e l'autorizzazione degli utenti.

Nelle infrastrutture Grid [1] un concetto chiave è quello di Organizzazione Virtuale, o Virtual Organization (VO). Ogni utente che voglia accedere alla Grid deve fare parte di una VO e identificarsi attraverso un certificato digitale in formato X.509. Le risorse vengono poi normalmente utilizzate attraverso la sottomissione di jobs; la tipologia di risorse richieste viene specificata attraverso un linguaggio chiamato Job Description Language (JDL). Architetturalmente, la Grid pone quindi enfasi sulla condivisione globale di risorse a livello di VO, con un approccio d'uso tipicamente di tipo batch. Le infrastrutture Grid sono state adottate principalmente nel mondo scientifico, in particolare da comunità fortemente aggregate come quella della fi-

sica delle alte energie, mentre l'adozione di questo modello da parte dell'industria è risultata marginale.

Nel Cloud Computing [2], d'altra parte, normalmente un utente che desideri utilizzare risorse distribuite identifica prima di tutto uno specifico Service Provider (SP). Questo SP rende disponibile le risorse desiderate, che possono essere di calcolo, di storage o di rete, spesso attraverso interfacce web e utilizzando un modello di pagamento a tariffazione d'uso (il cosiddetto modello *pay-as-you-go*). Dal punto di vista dell'architettura, il Cloud Computing pone quindi enfasi sulla facilità di accesso alle risorse per utenti singoli, con relativamente poco interesse sulla federazione di differenti SP. Fornisce inoltre diversi livelli di astrazione per l'identificazione delle risorse, che possono essere viste attraverso modelli di servizio quali *Software as a Service*, *Platform as a Service* o *Infrastructure as a Service*. Il Cloud Computing ha avuto un'adozione di grande successo in ambito privato, con manifestazioni di forte interesse anche da parte di comunità scientifiche.

Dal punto di vista dell'utenza, c'è da una

parte forte interesse a mantenere accesso a risorse distribuite attraverso interfacce Grid (in particolare per comunità scientifiche massicciamente distribuite e con investimenti significativi nello sviluppo di software legato a Grid); dall'altra, c'è il desiderio di utilizzare interfacce semplificate e nuovi servizi tipici del Cloud Computing.

Parallelamente, dal punto di vista dei Service Provider c'è interesse a fornire entrambe le modalità di accesso (Grid e Cloud) ma in maniera ottimizzata, scalabile e senza allocazioni statiche di risorse. È in questo contesto, tenendo presente le richieste di utenti e Service Provider allo stesso tempo, che è nata la soluzione descritta in questo articolo.

2. WNoDeS

WNoDeS [3] è un software sviluppato per realizzare una integrazione efficiente e scalabile per l'accesso a risorse attraverso interfacce Grid e Cloud tramite tecniche di virtualizzazione. WNoDeS crea in modo dinamico e trasparente delle VM opportunamente configurate per soddisfare le richieste dell'utenza. Queste VM possono essere utilizzate sia per l'esecuzione di *batch jobs* tradizionali (sottomessi localmente oppure attraverso interfacce Grid), sia per la fornitura di risorse di tipo Cloud.

3. L'architettura di WNoDeS

Dal punto di vista architetturale WNoDeS è organizzato come indicato in figura 1.

L'autenticazione degli utenti è internamente basata su certificati digitali X.509. WNoDeS prevede tuttavia che utenti che usino altri metodi di autenticazione, quali ticket Kerberos, IDP Shibboleth, oppure meccanismi locali di tipo *credit-based*, possa-

no accedere alle risorse tramite un *Authentication Gateway*. Tale *gateway* utilizza credenziali digitali di breve durata (attraverso uno *Short-Lived Credential Service* o SLCS). Grazie a questo sistema è possibile consentire accesso trasparente alle risorse Grid per gli utenti che non appartenengono a Virtual Organization Grid e, simmetricamente, alle risorse Cloud per gli utenti in possesso di certificati digitali X.509. Lo strato di autorizzazione è invece gestito tramite gLite Argus [6].

Gli utenti possono accedere alle risorse di calcolo virtualizzate e personalizzate attraverso l'utilizzo di quattro tipi di interfacce:

- *interfaccia Locale*, per la sottomissione di *batch job*;
- *interfaccia Grid*, per la sottomissione di *batch job* in modo distribuito;
- *interfaccia Cloud OCCI* per istanziazioni personali delle risorse;
- *interfaccia Virtual Interactive Pool (VIP)*, per permettere agli utenti locali di un centro di calcolo di gestire autonomamente risorse

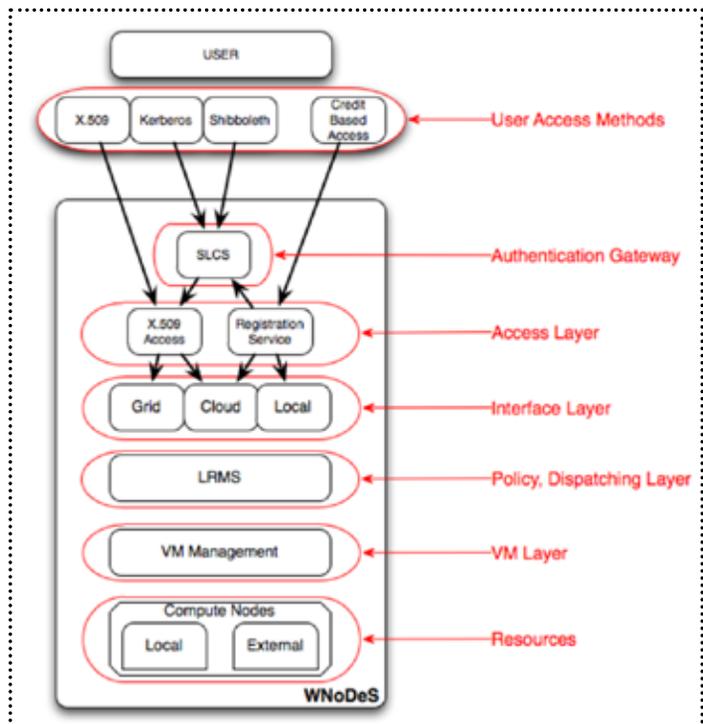


Fig. 1 L'architettura di WNoDeS

di calcolo.

L'utilizzo di interfacce locali e Grid è totalmente trasparente rispetto agli usi tradizionali. Una differenza fondamentale, tuttavia, è costituita dalla possibilità per gli utenti di specificare il tipo di VM sulla quale i propri job debbano essere eseguiti. A livello Grid, questo si realizza attraverso l'aggiunta di un parametro nello script di descrizione dei requisiti del job attraverso il JDL. Per la sottomissione del job vengono utilizzati i normali tool esistenti in infrastrutture Grid come la European Grid Infrastructure (EGI) [4].

La richiesta e l'uso di risorse Cloud, in particolare, può avvenire attualmente attraverso due meccanismi:

- un Web Service che implementa l'API Open Cloud Computing Interface (OCCI) [5];
- la WNoDeS Web Application, una applicazione che consente la allocazione di risorse attraverso un web browser.

Tutte le risorse fornite da WNoDeS, indipendentemente dall'interfaccia di accesso (Grid, Cloud o altro), appartengono ad un pool comune. Questo consente una gestione uniforme delle risorse e la massimizzazione del loro uso.

Il componente chiave per la gestione efficiente di migliaia di nodi in maniera concorrente è il LRMS. Ogni richiesta di risorse viene trasformata, in maniera trasparente per l'utente, in un job sottomesso al LRMS; WNoDeS gestisce quindi questi job creando risorse virtuali (VM) on demand utilizzando KVM [7]. Queste VM possono essere configurate in modo diverso (ad esempio per quanto riguarda il sistema operativo presente) a seconda delle richieste degli utenti, possono accedere a storage virtuale ed appartenere a diverse Virtual LAN. La flessibilità tipica dei moderni LRMS consente da una parte di applicare complesse *policy* di distribuzione e monitoring delle risorse; d'altra parte, la stretta integrazione di WNoDeS con il LRMS fa sì che non

sia necessario cambiare sostanzialmente la configurazione dell'LRMS per gestire WNoDeS.

Un altro componente chiave di WNoDeS è rappresentato dalla infrastruttura virtuale. Come *hypervisor* di Virtual Machine viene usato Linux Kernel Virtual Machine (KVM) installato su ciascuno nodo del centro della farm che supporti WNoDeS (potenzialmente, ma non necessariamente, tutti i nodi di un *cluster*). Ciascuna risorsa di calcolo esegue una speciale VM detta *bait* (o *esca*), che ha il compito di presentare al LRMS le capacità del nodo fisico su cui gira. Il *bait* viene rappresentato al *cluster* LRMS come una risorsa normale, e pubblica ad esempio il numero massimo di *slot* di calcolo disponibili. Questo numero è dinamico e può dipendere da vari parametri, come la RAM richiesta dal job o dalla VM. Il LRMS invia il job al *bait*, che non lo esegue direttamente, ma crea (attraverso l'*hypervisor* presente sulla macchina) un appropriato nodo virtuale sulla sua macchina fisica, al quale il job viene poi inviato per l'esecuzione. I nodi virtuali sono KVM VMs definite dall'amministratore di sistema o selezionate dagli utenti. La comunicazione tra LRMS, *hypervisor* e *bait* avviene tramite programmi Python che in sintesi:

- accettano i job inviati dal LRMS al *bait*;
- provocano la creazione, distruzione o sospensione di un KVM VM via *hypervisor*;
- mantengono le informazioni dello stato del job tra il nodo virtuale e il *bait*.

Tipo Interfaccia	Sviluppo	Impiego
Local	Stabile	In produzione
Grid	Stabile	In produzione
Cloud	Stabile	Test
VIP	In sviluppo	Prototipo

Tab. 1 Stato di sviluppo e impiego di WNoDeS

Molto lavoro è stato fatto per l'integrazione di Grid e Cloud, ed uno degli obiettivi di

Piena Integrazione con Grid e Cloud	Authentication Gateway
Selezione VM Immagine	Supporto di libguestfs
Multi-core VM	Supporto di Virtual LAN
Cloud: OCCI API	Integrazione con file system distribuiti (INFN Tier-1: IBM GPFS)
Cloud: Web Console	
Supporto di VirtIO	

Tab. 2 Principali funzionalità di WNoDeS

WNoDeS è propriamente di rendere questa integrazione il più trasparente possibile sia per gli utenti che per i fornitori di servizi, mirando allo sfruttamento ottimale delle risorse ed al mantenimento delle procedure esistenti e dei relativi flussi lavorativi.

4. Stato di WNoDeS

Lo stato dello sviluppo software e del suo impiego è schematizzato nella tabella 1. Le funzionalità ad oggi supportate sono riassunte nella tabella 2.

Il software WNoDeS, che attualmente si interfaccia con il LRMS LSF prodotto da Platform Computing [8], è da più di un anno in produzione presso il Centro di Calcolo Tier-1 dell'INFN sito al CNAF di Bologna [9]. Il Tier-1 dell'INFN è il centro di calcolo nazionale dell'INFN ed uno degli undici centri mondiali che partecipano alla gestione dei dati prodotti dall'acceleratore Large Hadron Collider [10] al CERN di Ginevra. È inoltre in corso l'installazione di WNoDeS presso i Laboratori Nazionali di Legnaro (LNL) dell'INFN e presso altre sedi INFN.

Le VM create da WNoDeS possono essere definite staticamente per utente o gruppi di utenti, oppure possono essere specificate direttamente dagli utenti attraverso i meccanismi supportati per la allocazione di risorse via Grid o via Cloud. WNoDeS viene attualmente utilizzato da circa 20 comunità scientifiche e gestisce una media di diverse decine di migliaia di job al giorno. Al momento sono disponibili circa 10 immagini virtuali differenti che possono essere scelte per istanziare macchine virtuali.

Presso il Tier-1 INFN sono attualmente attive (*on demand*) circa 2000 VMs in produzione, corrispondente a circa il 20% delle risorse di calcolo disponibili. Alcune di queste immagini sono usate per supportare alcuni sistemi operativi (attuali o *legacy*) come Scientific Linux 4 o 5, mentre altre sono state create per supportare richieste specifiche di utenti o di gruppi di utenti che necessitano per esempio di particolari ambienti di esecuzione.

Le risorse virtualizzate sono principalmente usate per l'esecuzione di *batch jobs* provenienti da sottomissioni locali oppure tramite interfacce Grid, mentre sono in fase di test quelle usate per l'esecuzione di job provenienti da interfacce Cloud e VIP.

È previsto che nel prossimo futuro le VM gestite da WNoDeS si incrementino al Tier-1 INFN di diverse altre migliaia, fino ad includere la maggior parte delle risorse presenti (se non tutte). WNoDeS è un progetto ambizioso, che pensiamo possa rappresentare una soluzione valida e scalabile per la virtualizzazione e la fornitura di servizi innovativi per centri di calcolo anche di grandi dimensioni.

Riferimenti bibliografici

- [1] I. Foster, What is the Grid? A Three Point Checklist, 2002, <http://www.mcs.anl.gov/itf/Articles/WhatIsTheGrid.pdf>
- [2] U.S. National Institute of Standards and Technology Definition of Cloud Computing, cfr. <http://csrc.nist.gov/groups/SNS/cloud-computing/index.html>

- [3] <http://web.infn.it/wnodes>
- [4] www.egi.eu
- [5] www.occi-wg.org/doku.php
- [6] <https://twiki.cern.ch/twiki/bin/view/EGEE/AuthorizationFramework>
- [7] www.linux-kvm.org
- [8] www.platform.com
- [9] www.cnaf.infn.it
- [10] <http://lhc.web.cern.ch/lhc>



Davide Salomoni
davide.salomoni@cnaf.infn.it

Davide Salomoni ha lavorato con l'INFN dal 1991 al 1998. Dal 1999 al 2001 ha coordinato il gruppo rete a SLAC, USA. Dal 2001 al 2003 è stato Internet Group Manager a COLT Telecom ad Amsterdam. Dal 2003 al 2005 ha lavorato su Grid a NIKHEF in Olanda. Nel 2005 è ritornato all'INFN e attualmente gestisce la parte calcolo del Tier-1 presso il CNAF di Bologna.

Conferenza GARR_10

Selected papers

L2L (Live to e-Learning)

M. Bertazzo

WIDEN: implementazione di protocolli di Clinical Trials con servizi web

P. Cerello, A. Biggi, S. Chauvie, A. Gallamini, A. Stancu

LinkO: una infrastruttura in fibra ottica per confronti remoti di tempo e frequenza ad alta accuratezza

C. Clivati, D. Calonico, F. Levi, A. Mura, G. A. Costanzo, A. Godone

Utilizzo di Zabbix per il Monitoring di risorse di calcolo e sistemi di storage su WAN in ambito Grid Computing

A. Colavincenzo, A. Petricca, G. Bracco, S. Migliori, F. Simoni, S. Pecoraro, D. Giammattei, A. Rocchi, C. Sciò

NEXPreS: Completamento ed evoluzione dell'e-VLBI

M. Nanni

FARO: Accesso web a risorse remote per l'industria e la ricerca

A. Rocchi, S. Pierattini, G. Bracco, S. Migliori, F. Beone, C. Sciò, R. Ciavarella, A. Petricca

WNoDeS: un servizio per la gestione di infrastrutture condivise Cloud e Grid

D. Salomoni, A. Chierici, A. Italiano, E. Ronchieri, P. Solagna

ISBN 978-88-905077-1-7



9 788890 507717