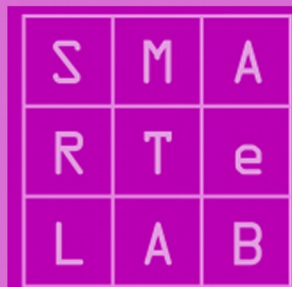


# SMART eLAB

VOLUME 3 - ANNO 2014



ISSN 2282 - 2259

## SOMMARIO

Vol. 3, 2014

## Articoli

- 1-5 **Ombretta Tarquini, Antonello Ranieri, Marcello Colapietro** *Allestimento di una postazione per la diffrazione dei raggi X per l'analisi di materiali cristallini. y*
- 6-16 **Giuseppe Confessore, Giuseppe Stecca, Marco Simonetti** *Progettazione di un Sistema Web per favorire il Trasferimento Tecnologico dei Prodotti della Ricerca*
- 17-21 **Augusto Pifferi, Andrea Lora, Giuseppe Nantista, Luca Ianniello** *Setup of a clustered antispam and antivirus service based on Mailcleaner®suite*
- 22-24 **Augusto Pifferi, Angelo De Simone, Giovanni Agostini, Luca Ianniello, Andrea Lora, Giuseppe Nantista** *Wifi in Fara, un progetto di access point federati per l'accesso a internet gratuito.*
- 25-32 **Luca Ianniello, Giuseppe Nantista, Andrea Lora, Augusto Pifferi** *Progetto Virtual Hosting in High-Availability per l'Area della Ricerca RM1 del CNR.*

Smart e-Lab: <http://smart-elab.mlib.ic.cnr.it>

A peer-reviewed online resource, published by the Istituto di Cristallografia (CNR-IC)

EDITORS-IN-CHIEF : Michele Saviano, Augusto Pifferi

ASSOCIATED EDITOR : Guido Righini

GRAPHIC DESIGN : Claudio Ricci

EDITORIAL ASSISTANT : Caterina Chiarella

CNR - Istituto di Cristallografia, Strada Provinciale 35/d, I-00015 Monterotondo, Italy

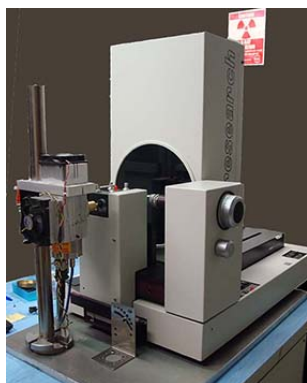


Creative Commons Attribuzione - Non commerciale - Condividi allo stesso modo 4.0 Internazionale



## Allestimento di una postazione per la diffrazione dei raggi X per l'analisi di materiali cristallini.<sup>†</sup>

Ombretta Tarquini,<sup>a</sup> Marcello Colapietro,<sup>a</sup> Antonello Ranieri<sup>a</sup>, Augusto Pifferi.<sup>a</sup>



Presentazione di una postazione per diffrazione X realizzata ricondizionando un diffrattometro a raggi X della Marresearch con rivelatore Image Plate ed equipaggiato con una sorgente a bassa potenza. Modifiche Hardware e applicazioni.

**Keywords:** XRD, Strumentazione per diffrazione, Imaging Plate.

### 1 Introduzione

Nel Laboratorio a Raggi X dell'Istituto di Cristallografia (UOS Montelibretti - RM) è stata assemblata una postazione per l'analisi di materiali cristallini basata su un diffrattometro a raggi X. La stazione è costituita da un diffrattometro a singolo asse della MarResearch (Modello Mar300) già presente nell'Istituto e da una sorgente di raggi X a bassa potenza opportunamente scelta.

Questo rapporto presenta la strumentazione con le modifiche hardware eseguite e le potenzialità dello strumento.

### 2 Lo Strumento

#### 2.1 La sorgente RX

La sorgente a raggi X è costituita da un generatore Spellmann (modello LG50P100) (Fig. 1) che è stato scelto per la sua caratteristica di essere raffreddato ad aria e risulta essere estremamente compatto. Il generatore è stato installato all'interno di un Rack 19" che è parte integrante della strumentazione.

Il tubo a Raggi X è il XTF5011 il della "Oxford Instruments" con anodo di molibdeno operante a 50kV, 1.0mA massimi, con macchia focale di 110 $\mu$ m ed angolo

di uscita di 25°.

Il raffreddamento del tubo è realizzato con un blocco di alluminio cavo, in cui esso è inserito, dotato di dissipatori passivi e di due ventole. Alla massima potenza di esercizio la temperatura si attesta intorno ai 45°, con una temperatura ambiente di 22°.



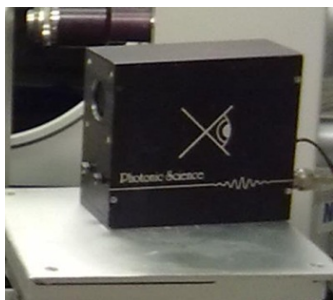
**Fig. 1** Generatore Spellmann.

Il fascio X in uscita è focalizzato da una semi lente Kumaikhov.<sup>1</sup> Questa soluzione ha dato origine a notevoli problemi di allineamento. Per risolvere queste complicazioni è stato realizzato un supporto in ottone che permette di centrare l'ottica attraverso un sistema di viti distanziate di 120°. Una camera video CCD per RX della Photonic Science ha permesso di controllare il fascio X in uscita dall'ottica focalizzante durante le operazioni di allineamento (Fig. 2).

<sup>a</sup> CNR - Istituto di Cristallografia, Strada Provinciale 35/d, Montelibretti, Italia

Creative Commons Attribuzione - Non commerciale - Condividi allo stesso modo 4.0 Internazionale

<sup>†</sup> rapporto tecnico IC 13/08 registrato con numero di protocollo IC/1871 del 19/12/20133



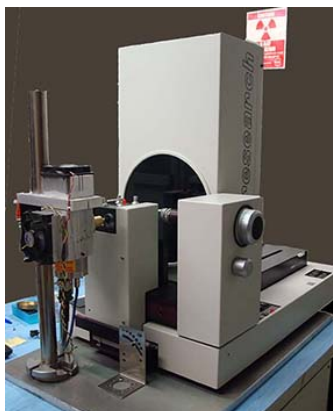
**Fig. 2** Fotocamera video CCD intensificata della Photonic Science.

## 2.2 Il diffrattometro

Il diffrattometro, Marresearch è costituito da un singolo asse rotante avente risoluzione millesimale (angolo  $\varphi$ ) sul quale è fissata la testina goniometrica con il campione da analizzare.

Una telecamera ottica integrata all'interno del diffrattometro permette di effettuare la centratura del campione.

Il rivelatore è costituito da una lastra Imaging Plate (IP) dal diametro di 300 mm. I rivelatori IP si basano sull'acquisizione di un'immagine latente ad opera di fosfori fotostimolabili della famiglia dei fluoroalidi di bario ( $BaFX : Eu^{+2}$ ). Tali composti quando vengono eccitati da radiazione X, intrappolano gli elettroni rilasciati da  $Eu^{2+}$  in stati metastabili nella regione proibita. Stimolando il rivelatore con una radiazione rossa (630-650nm) tali elettroni vengono promossi in banda di conduzione per poi tornare nello stato iniziale con l'emissione caratteristica a 390nm. Mentre il raggio di radiazione rossa esplora la lastra, la radiazione emessa dal fosforo è raccolta da un fotomoltiplicatore. Uno speciale filtro ottico, posto all'ingresso del fotomoltiplicatore, impedisce il passaggio della radiazione stimolante. Il fotomoltiplicatore converte i fotoni in segnali elettrici che vengono poi trasformati in dati digitali da un convertitore A/D ed inviati ad un PC. Un apposito software di gestione ricostruisce l'immagine digitale conoscendo la geometria di scansione della lastra.<sup>2</sup>

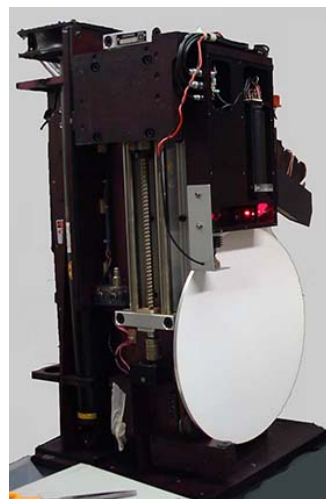


**Fig. 3** Diffrattometro IP della Marresearch.

## 3 Modifiche Hardware

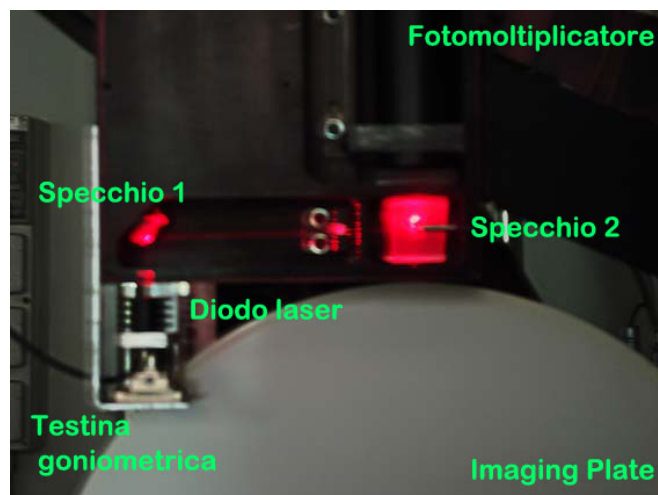
### 3.1 Sostituzione del laser

Nel sistema originale il raggio rosso di lettura dell'IP era generato da due tubi laser He-Ne (630nm) da 40 mW ciascuno. Questi erano posizionati in modo tale che, attraverso un sistema di specchi, i due raggi laser convergessero in un unico punto ( $\varnothing$  circa  $100\mu\text{m}$ ) sulla lastra. La lettura dell'immagine latente è realizzata con la rotazione del disco dell'IP insieme allo spostamento della testa. In questo modo la lunghezza del cammino ottico del laser variava durante la lettura.



**Fig. 4** Diffrattometro senza coperchio di protezione, è visibile il rivelatore IP ed il sistema di lettura.

Data l'irreperibilità sul mercato di uno dei tubi He-Ne risultato guasto, entrambi sono stati sostituiti con un diodo laser a stato solido ( $\lambda = 650\text{nm}$ , 200mW max). Il diodo è stato montato su una testina goniometrica di precisione che permette un accurato allineamento del fascio grazie a quattro movimentazioni: due per la traslazione e due per il "tilting" (Fig. 5). Per le ridotte dimensioni, la testina è stata montata direttamente alla testa di



**Fig. 5** Sistema di lettura con diodo laser.

lettura dove è installato il fotomoltiplicatore. In questo modo è stato ridotto drasticamente e mantenuto costante il cammino ottico del raggio laser.

### 3.2 Alimentazione e isolamento

L'alimentazione dei diodi laser è critica in quanto bastano brevissimi picchi di tensione (dell'ordine dei nano secondi) per trasformare in modo irreversibile il loro funzionamento a diodi LED. Inizialmente è stato utilizzato un alimentatore di tipo switching da 5V opportunamente stabilizzato in corrente. Questa soluzione è risultata non valida in quanto, dopo poche ore di funzionamento, il diodo si bruciava.

A seguito di indagini minuziose è emersa la presenza di picchi di tensione che causavano la rottura del laser.

Pertanto si è proceduto alla sostituzione dell'alimentatore switching con un alimentatore lineare ma anche in questo caso si arrivava alla rottura del diodo in breve tempo.

Allo stato attuale il sistema di alimentazione del diodo è formato da un circuito nel quale sono inseriti un trasformatore con "isolamento galvanico" in grado di fornire una tensione di 12V ed una corrente massima di 500mA. A questo è collegato un ponte di diodi ed un regolatore di tensione (MC7808) che stabilizza la tensione ad 8V DC. Un circuito integrato (LM317), con una serie di trimmer e resistenze, consente di aggiustare la corrente di alimentazione del diodo laser per avere una potenza emissione di circa 100mW. (Fig. 6)

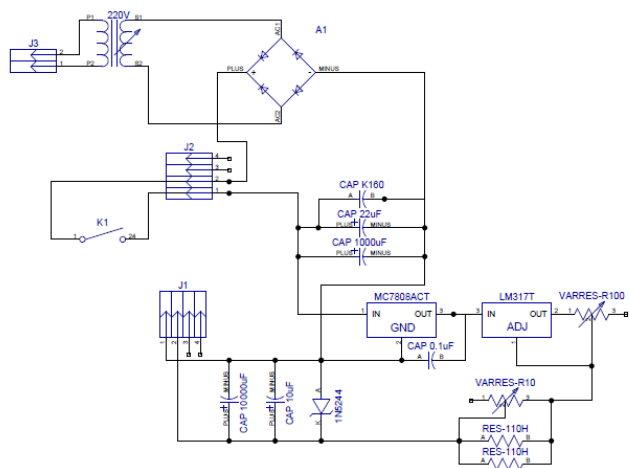


Fig. 6 Circuito di alimentazione del diodo laser.

### 3.3 Misure ed applicazioni scientifiche

Con tale implementazione si riescono ad ottenere delle figure di diffrazione ben risolte spazialmente (Fig. 7) e con un ottimo rapporto segnale/rumore (Fig. 8).

Il sistema è controllato da un PC remoto e da un programma di gestione. Una volta ottenuta l'immagine di diffrazione, questa viene rielaborata con il programma

*Fit2d*<sup>3</sup> per ottenere uno spettro delle intensità in unione dell'angolo di diffrazione ( $2\theta$ ) oppure di  $d$ . Gli spettri ottenuti vengono elaborati e confrontati con il programma *Match!* della *CRYSTAL IMPACT* utilizzando l'archivio COD (*Crystallography Open Database*) dell'*AMCSD* (*American Mineralogist Crystal Structure Database*, Versione 20/03/2012) e l'archivio PDF2 (*Powder Diffraction File*).

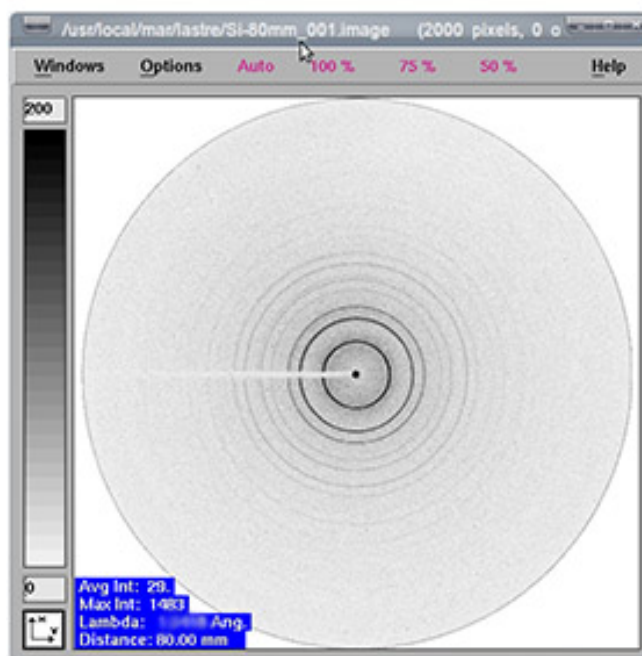


Fig. 7 Figura di diffrazione del silicio, 40kV, 1.3 mA, 300 s.

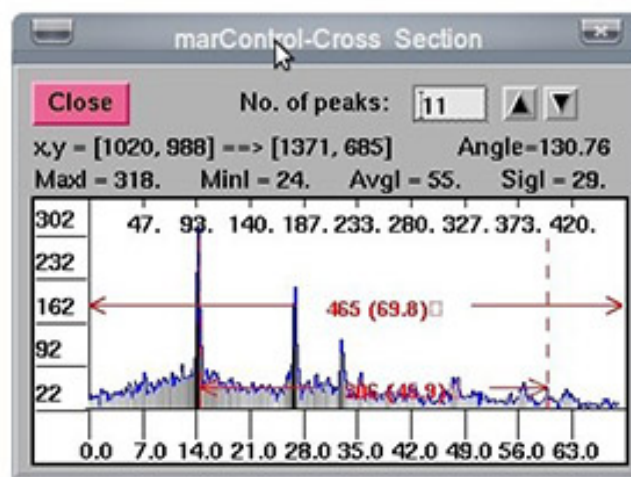
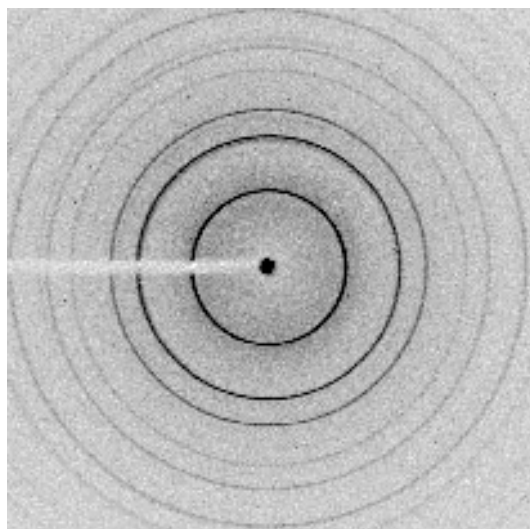


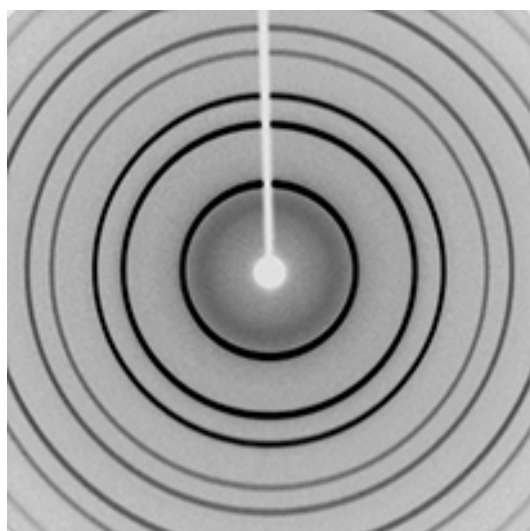
Fig. 8 Spettro di diffrazione relativo alla fig. 7. E' evidente l'ottimo rapporto segnale/rumore

L'immagine di diffrazione del Si standard ottenuta con la strumentazione in oggetto (Fig. 9) è stata confrontata con quella ottenuta con un diffrattometro a geometria K della Oxford Instruments con rivelatore CCD con taper di FOS (Fig. 10). Dal confronto è evidente la maggior risoluzione dello spettro ottenuto con l'IP tenendo in

considerazione, a parità di tempo di raccolta, la potenza utilizzata, circa 50W per l'IP e 2000W nel secondo caso.



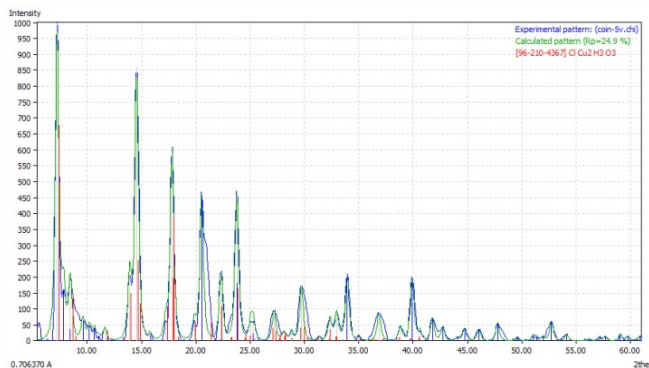
**Fig. 9** Figura di diffrazione del silicio, 40 kV, 1.3 mA e 300 s.



**Fig. 10** Figura di diffrazione del Silicio ottenuta con un diffrattometro a geometria K della Oxford instruments, rivelatore CCD con taper di Fos. (50 kV, 40 mA, 300 s)

Attualmente lo strumento viene utilizzato per raccogliere spettri di diffrazione da polvere al fine di riconoscere le fasi cristalline presenti in campioni provenienti da diverse tipologie di reperti archeologici.

Per i manufatti metallici tale tecnica è utilizzata per identificare i prodotti di corrosione che costituiscono la patina superficiale. La loro conoscenza è un valido indice per determinare l'autenticità dell'oggetto, per decidere sia il miglior trattamento conservativo, che l'ambiente più idoneo in cui custodirlo. Inoltre possono fornire informazioni sull'ambiente e sugli eventi che hanno trasformato il metallo. La rimozione, senza la loro conoscenza, potrebbe cancellare importanti informazioni sulla storia del reperto.<sup>4</sup>



**Fig. 11** Spettro di diffrazione del campione di patina prelevato nella moneta di Fig. 12. è stata individuata la fase cristallina della paratacamite ( $\text{Cu}_2(\text{OH})_3\text{Cl}$ ) [96-210-4367] con un FoM=0.83.

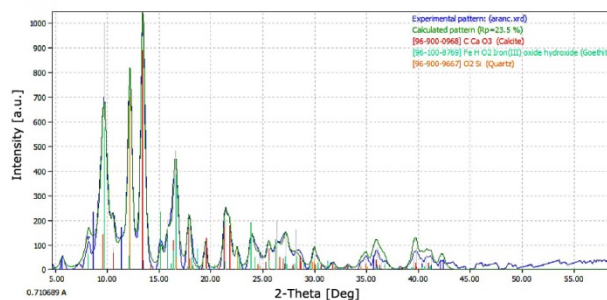


**Fig. 12** Moneta archeologica: è evidente l'attacco del cancro del bronzo.

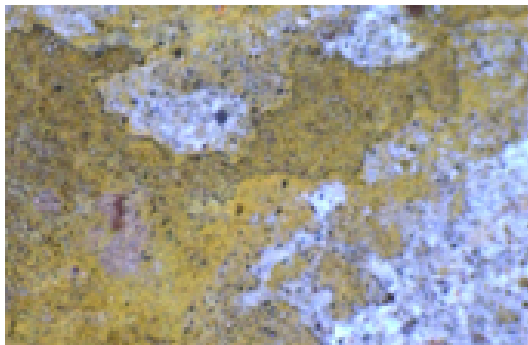
La diffrazione è anche un valido aiuto nell'identificazione dei pigmenti. In Fig. 13 è riportato lo spettro di diffrazione del pigmento giallo di Fig. 14.

## 4 Conclusioni

Concludendo la strumentazione assemblata permette di fare analisi accurate, in tempi brevi, su materiali policristallini e si è rivelata un valido supporto alle indagini archeometriche.



**Fig. 13** Spettro di diffrazione del pigmento giallo in cui sono state individuate le fasi cristalline della calcite ( $\text{CaCO}_3$  [96-900-0968] FoM=0.84), della goethite ( $\text{FeO}(\text{OH})$  [96-100-8769] FoM=0.78) e del quarzo ( $\text{SiO}_2$  [96-900-9667] FoM=0.71)



**Fig. 14** Foto microscopio di pigmento giallo su terracotta architettonica.

---

## Riferimenti

- 1 M. Kumakhov, Channeling of photons and new x-ray optics, *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 48 (1) (1990) 283 – 286. [doi:10.1016/0168-583X\(90\)90123-C](https://doi.org/10.1016/0168-583X(90)90123-C).
- 2 A. Kulik, C. Rosenfeld, Charged particle detection using storage phosphor crystals, in: *1996 IEEE Nuclear Science Symposium. Conference Record, Vol. 1, 1996*, pp. 539–543 vol.1. [doi:10.1109/NSSMIC.1996.591054](https://doi.org/10.1109/NSSMIC.1996.591054).
- 3 A. Hammersley, FIT2D: an introduction and overview, Tech. Rep. ESRF97HA02T, European Synchrotron Radiation Facility Internal Report (1997).
- 4 D. A. Scott, *Copper and Bronze in Art-Corrosion, colorants, conservation*, Getty Publications, Los Angeles, California, 2002.



## Progettazione di un Sistema Web per favorire il Trasferimento Tecnologico dei Prodotti della Ricerca<sup>†</sup>.

Giuseppe Confessore,<sup>a</sup> Marco Simonetti,<sup>b</sup> Giuseppe Stecca<sup>c</sup>

Il presente lavoro è redatto nell'ambito delle operazioni svolte nella prima parte del progetto GENESI, relativo al bando nazionale RIDITT (Rete Italiana per la Diffusione dell'Innovazione e il Trasferimento Tecnologico alle imprese) finanziato dal Ministero dello Sviluppo Economico (MiSE). Il progetto ha come obiettivo la creazione di una rete connettiva stabile tra CNR, Università, associazioni imprenditoriali e PMI per il trasferimento dei risultati di ricerca in campo meccanico e dei sistemi di produzione industriale, in particolare nel settore delle macchine, della robotica industriale, delle tecnologie per i processi produttivi e logistici. Il progetto si articola in attività legate all'analisi della domanda e dell'offerta di tecnologie meccaniche per i sistemi di produzione industriale, il trasferimento e il sostegno all'inserimento delle innovazioni nelle aziende. Il presente rapporto tecnico descrive le attività svolte nell'ambito del progetto riguardanti lo sviluppo di sistemi e tools che supporteranno le attività di analisi, di diffusione, di promozione e trasferimento dei prodotti della ricerca.

**Keywords:** Trasferimento Tecnologico, Reti Collaborative, Survey, Basi di Dati, Content management, programmazione web.



## 1 Introduzione

La realizzazione di una rete integrata di enti di Ricerca e Sviluppo e di imprese può favorire e gestire il trasferimento tecnologico in sistemi di produzione analizzando allo stesso tempo, dati di stress e/o deformazione insieme a dati economici. E' possibile così avere soluzioni progettuali industrialmente valide tenendo in considerazione nella progettazione di una struttura non solo delle caratteristiche meccaniche ma anche dei vincoli economici. Si può, inoltre, ottenere innovazione nel sistema organizzativo in generale e della produzione in particolare, affrontando con più forza e coesione la sfida della competitività. Per questo obiettivo risulta necessario una costante ristrutturazione degli impianti di produzione e una rivitalizzazione della pianificazione al fine di realizzare su larga scala tecnologie e attrezzature, migliorare la capacità di innovazione autonoma, e promuovere l'ottimiz-

zazione della struttura industriale e la ristrutturazione per rafforzare la competitività industriale.

In questo contesto il CNR può svolgere un importante ruolo avendo fra i suoi obiettivi di ricerca numerose attività nell'ambito del manifatturiero avanzato e delle sue tecnologie abilitanti. Nello specifico a partire dalle competenze dell'ex Dipartimento Sistemi di Produzione ora convogliato insieme al Dipartimento Energia e Trasporti e al Dipartimento ICT in un unico Dipartimento denominato Ingegneria, ICT e tecnologie per l'Energia e i Trasporti, si è dato vita ad un'iniziativa scientifica strutturata che focalizza gli sforzi coordinati, di un numero significativo di ricercatori afferenti a settori disciplinari diversi, sulle tematiche di ricerca nell'ambito di processo e di prodotto industriale. L'obiettivo generale, anche del nuovo Dipartimento DIITET, è quindi quello di confermare ed accrescere questo ruolo centrale, progettando ed attuando strategie per promuovere e realizzare, attraverso le proprie strutture di afferenza, sinergie ed attività comuni con l'impresa e l'Università.

Esso quindi contribuisce ad accrescere la competitività e la sostenibilità economica, sociale, ambientale ed energetica del sistema industriale Italiano attraverso attività di ricerca fondate su:

- l'integrazione di nuove tecnologie abilitanti nel-

<sup>a</sup> CNR - Istituto di Cristallografia, Strada Provinciale 35/d, Montelibretti, Italia

<sup>b</sup> CNR - Istituto di Tecnologie Industriali e Automazione, Strada Provinciale 35/d, Montelibretti, Italia

<sup>c</sup> CNR - Istituto di Analisi dei Sistemi ed Informatica, viale Manzoni 30, Roma, Italia

Creative Commons Attribuzione - Non commerciale - Condividi allo stesso modo 4.0 Internazionale

<sup>†</sup> rapporto tecnico IC 13/07 registrato con numero di protocollo IC/1848 del 18/12/2013



- lo sviluppo di nuovi prodotti/servizi e processi ad elevata qualità sostenibile e ad alto valore aggiunto;
- la riduzione del tempo di trasferimento tra ricerca;
- l'integrazione di diversi partner, lungo la catena del valore ricerca-innovazione industriale;
- lo sviluppo di nuova imprenditorialità basata sulla conoscenza.

Tre sono i filoni di ricerca che le strutture scientifiche coinvolte del CNR seguono a tale scopo:

- ricerche finalizzate all'innovazione industriale di prodotti e tecnologie di specifici settori produttivi;
- ricerche finalizzate all'introduzione di tecnologie abilitanti all'interno di tali settori;
- ricerche finalizzate allo sviluppo di strumenti per il miglioramento dei processi produttivi.

I settori affrontati dal Dipartimento, per il progetto GENESI, sono quelli relativi a:

- macchine operatrici e robot per lavorazioni in diversi settori
- macchine movimento terra
- industria tessile
- industria del legno
- industria della ceramica, tradizionale ed innovativa
- industria delle costruzioni (materiali, prodotti e sistemi).

Nel campo della ricerca sui sistemi di produzione si rivolge principalmente a:

- nuovi modelli di processo produttivo
- nuovi modelli di business
- trasferimento e adeguamento di modelli di processo tra settori produttivi diversi.

Inoltre, le strutture scientifiche del CNR, per la loro natura e per l'evoluzione della ricerca, hanno maturato specifiche competenze di ricerca e sviluppo tecnologico a supporto dell'automazione di fabbrica. In tal contesto, le principali attività riguardano:

- la modellazione, progettazione e simulazione dei sistemi di controllo logico e modulante per fabbriche riconfigurabili;
- la definizione di architetture di controllo logico distribuito basato sulla modellazione ad oggetti;
- la generazione di approcci per la riconfigurazione dei sistemi produzione che integrino la modellazione tecnologico-funzionale delle risorse operatrici con la loro caratterizzazione energetica;
- la modellazione e simulazione dell'impatto ambientale di prodotti e processi manifatturieri ad alto valore aggiunto;

- lo sviluppo di approcci di supervisione e manutenzione per la riconfigurazione di fabbrica in tempo reale;
- l'utilizzo di tecnologie emergenti per l'implementazione di soluzioni di controllo logico di sistema attraverso l'integrazione dell'automazione di processo con quella di fabbrica.
- la modellazione meccatronica dei macchinari, integrando aspetti legati alle strutture meccaniche.
- sviluppo di modelli numerici per la simulazione dei processi.
- verifiche sperimentali su prototipi (risposta dinamica, risposta degli assi controllati, analisi modale sperimentale, prove di lavorazione) e confronto con i risultati forniti dai corrispondenti modelli numerici.
- analisi energetica di macchinari, sviluppo di metodologie per l'analisi energetica.
- sviluppo di sistemi di monitoraggio per macchine utensili e mandrini.

Le Tecnologie trattate che sono oggetto di un possibile trasferimento sono quelle, come detto legate all'ambito dei sistemi di produzione meccanici, anche in relazione alle richieste dei partner industriali con cui il progetto viene sviluppato che sono i seguenti:

- ZIPA (consorzio Zone Imprenditoriali della Provincia di Ancona),
- Consorzio per lo sviluppo industriale della provincia di Rieti,
- FICEI (Federazione Italiana Consorzi Enti Industrializzazione),
- FEDERLAZIO, associazione piccole e medie imprese del Lazio;

inoltre oltre ai prodotti proposti dal CNR è stato coinvolto anche un'altra importante struttura che eccelle anche nell'ambito della ricerca industriale:

- Università degli Studi di Roma Tre, Dipartimento di Informatica e Automatica (D.I.A.).

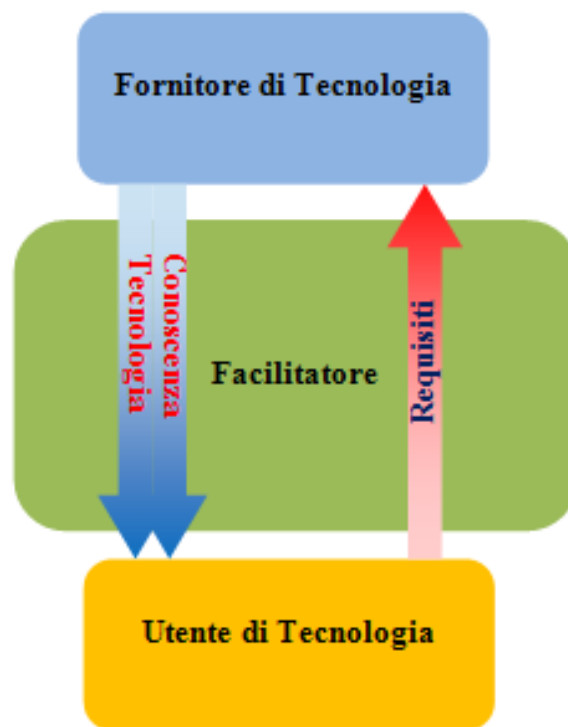
Vista la composizione del partenariato, si è posta l'attenzione, nello specifico, in alcuni ambiti di forte attrattività industriale in cui i partner della ricerca potessero portare innovazione a valore aggiunto come:

- Processi industriali high-tech: metodi e strumenti: Il progetto analizza i processi che, nella correlazione che unisce filiere verticali di produzione, caratterizzano il ciclo di vita della Fabbrica nei suoi principali macro-livelli (macchina, cella/sistema ed impresa) ed individua i processi più importanti (cioè, progettazione, gestione, riconfigurazione) ai fini del conseguimento di un effettivo vantaggio competitivo basato sulla conoscenza.

- **Microsistemi embedded:** l'obiettivo generale consiste nella progettazione, realizzazione, studio ed applicazione di materiali e sistemi innovativi di interesse per il sistema di produzione nazionale. Fra gli argomenti in fase di sviluppo si segnalano: sistemi di monitoraggio ad alte prestazioni (rivelatori gas, di radiazioni ad alta energia, sensori chimici, ...); attuatori e trasduttori per applicazioni in meccanica, domotica e in sistemi manifatturieri; materiali innovativi per generazione e trasporto di energia.
- **Robot e sistemi integrati di produzione:** il progetto integra competenze riconducibili alla meccatronica, robotica e automazione ed ha ricadute su imprese che operano in diversi settori produttivi con tecnologie medio-alte. Esso è indirizzato verso la concezione e sviluppo di sistemi integrati di produzione, macchine utensili, robot, automazione e sistemi e componenti high-tech per diverse applicazioni.
- **Le attività interessano lo sviluppo integrato,** la scelta dei materiali, la simulazione, la prototipazione, la brevettazione, l'industrializzazione e l'adeguamento normativo di sistemi per produrre e dei relativi processi.
- **Sistemi di monitoraggio, controllo e sicurezza** nei contesti produttivi e d'uso: sistemi autonomi intelligenti di monitoraggio, controllo e sicurezza in contesti applicativi specifici con sviluppo di metodologie di progettazione alternative a quelle attuali in grado di rispondere alle esigenze innovative dei contesti applicativi coinvolti.
- **Sistemi per movimentazione e lavorazione** in ambienti non strutturati: tematiche di interesse applicativo per comparti industriali rilevanti: sviluppi avanzati delle macchine e delle loro catene di potenza; sviluppo degli azionamenti meccatronici e delle relative architetture di controllo (a livello software e hardware); simulazione avanzata di sistemi complessi; sviluppo e applicazione dei requisiti di salute e comfort sia dal punto di vista ergonomico che dal punto di vista della sicurezza attiva e passiva delle apparecchiature (rischio fisico e EMC); intensimetria acustica per la caratterizzazione di sorgenti complesse.
- **Prodotti industriali high-tech:** il progetto si pone l'obiettivo di progredire nello sviluppo di prodotti knowledge based per il manifatturiero cosiddetto maturo, tipico del Made in Italy, al fine di migliorare, rafforzare e valorizzare i punti di forza delle imprese nazionali nei mercati esteri. In particolare, opera sullo sviluppo di nuovimateriali anche nano strutturati, nell'ambito di settori industriali tradizionali – calzature, tessile, industria del legno, dei materiali ceramici (innovativi e tradizionali) e bioceramici, componentistica elettronica.

Il progetto, quindi, ha come obiettivo quello di formare una rete stabile di connessione tra CNR, Università e aziende delle aree sottoutilizzate di Lazio e Marche e di aumentare in modo considerevole il numero delle connessioni Industria – Settore scientifico (Industry-Science Links) per le aziende del campo meccanico e dei sistemi di produzione; il numero di tali link può essere esplicitato in termini di:

- Start-up e aziende ad alta tecnologia formate con la collaborazione di ricercatori;
- Collaborazioni di ricerca tra aziende e ente di ricerca sia di tipo bilaterale che basate su consorzi;
- Contratti di ricerca e consulenze (o convenzioni) commissionate dall'industria;
- Sviluppo di strategie per la protezione dei diritti di proprietà non limitate allo sviluppo di brevetti;
- Attività di trasferimento di conoscenza basate training, formazione, disseminazione;
- Attività congiunte (ad esempio in laboratori);
- Scambio sistematico di staff tra enti di ricerca e aziende.



**Fig. 1** Trasferimento Tecnologico

La fattibilità dei risultati attesi è funzionale e proporzionale alla capacità di recepire il Trasferimento Tecnologico, e più precisamente ai tre elementi: 1) lo stock

di tecnologia posseduta inizialmente dall'agente che acquisisce tecnologia; 2) il flusso di tecnologia che viene trasferito dall'agente leader; 3) la variazione dello stock reale di tecnologia posseduta dall'agente che lo ha ricevuto.

Per fare ciò le operazioni descritte nel report, e cioè la realizzazione dei tools e delle piattaforme di scambio dati e informazioni, sono fondamentali e hanno un ruolo centrale per far conoscere le richieste e le offerte di tecnologia e per permettere il loro corretto trasferimento dalla ricerca alle imprese.

## 2 Attività svolte dal CNR

Nel partenariato sopra descritto, il CNR e nello specifico il nostro gruppo di lavoro, anche in relazione al ruolo di coordinazione dell'intero progetto, ha una serie di attività (o work Package - WP) fondamentali e centrali che risultano poi essere gli input per successive azioni.

Di seguito verranno indicati, in una tabella, tutte le attività svolte in questa prima parte di lavoro.

AZIONE	ATTIVITÀ
A) Studio e Analisi	A1) Screening tecnologico A2) Analisi del bacino di utenza A3) Modello di ottimizzazione della rete di competenze A4) Studio di analisi delle performance tecniche A5) Attività di formazione per il gruppo di lavoro
B1) Promozione B3) Dimostrazione	B1.2) Sviluppo Sistema Web Based B3.1) Audit tecnologici
C3) Sostegno alla Creazione di Nuove Imprese	C3) Sportello all'impresa

Per quanto riguarda questo report verranno descritte le attività di tipo tecnico, come la realizzazione dei sistemi informativi (che comprendono la progettazione e realizzazione delle schede prodotto, la realizzazione della piattaforma web e del survey).

### 2.1 Infrastruttura di rete e sviluppo del sistema Web Based

Tra i compiti principale assegnati al gruppo CNR uno dei più importanti è sicuramente la progettazione dei sistemi informativi che hanno il compito di raccogliere tutti i dati e di fornire una visione dei prodotti della ricerca già censiti, secondo le schede prodotto progettate, e disponibili oltre che permettere l'inserimento dei nuovi.

Il punto di partenza quindi risulta coincidere con un sito web di progetto che oltre a fungere da sistema

divulgativo e promozionale del progetto si pone come accentratore delle informazioni e dei dati prodotti.

Per la realizzazione del sito web, ospitato (in hosting) su una macchina virtuale sita all'interno del CED dell'Area della Ricerca Roma 1 di Montelibretti e parzialmente in gestione all'Istituto delle Tecnologie Industriali e Automazione (ITIA) - sezione di Roma, si è deciso di utilizzare Joomla ( <http://www.joomla.org/> ) nella stable release più recente (v. 3). La scelta di utilizzare una macchina virtuale, gentilmente concessa dal Servizio Reti d'Area (SRA) dell'Area della Ricerca RM1, è giustificata dall'infrastruttura presente e disponibile presso il CED nuovo appena completato.

Infatti la macchina web che gestisce il sito di progetto, raggiungibile all'indirizzo web <http://genesi.mlib.cnr.it> , non è una semplice macchina fisica, ma una macchina virtuale (o meglio un pool di macchine virtuali) facente parte di un'infrastruttura di tipo VMWare licenziato "Standard" che tra le varie opzioni prevede alcuni servizi molto utili per garantire un elevato livello di controllo e affidabilità come il vMotion (opzione che permette di eliminare i downtime dei sistemi per situazioni di migrazioni e maintenance e quindi aumenta il livello di SLA dei servizi da noi offerti).

Senza dilungarsi particolarmente sull'infrastruttura, in tutta la sua composizione e complessità, possiamo brevemente dire che il web server virtuale è situata in 2 lame HP Blade BL460c equipaggiate con 2 processori Xeon 2.9ghz Esa-core e 48GB di RAM. Lo storage fisico è un HP Left-hand P4000 con 14.4TB di capacità lorda, configurato per l'HA (High-Availability) con 2 chassis in mirror. I sistemi Blade e Storage colloquiano tra di loro alla velocità di trasmissione dati di 10GB attraverso switch ridondati anch'essi HP.

Inoltre il server web è ridondato, cioè fa parte di un pool di macchine, chiusa in una rete DMZ (DeMilitarized Zone); le macchine web, tutte uguali tra loro, hanno 12 processori Intel(R) Xeon(R) CPU E5-2667 0 @ 2.90GHz, 6GB di vRAM e 1TB di hard disk condiviso sfruttando la tecnologia NFS (Network File System), che ha il compito di gestire la consistenza dei dati su tutte, e utilizzano un sistema round robin di risposta alle query esterne, visto che sfruttano tutte una NAT (Network Address Translation) sullo stesso indirizzo IP pubblico (questa logica viene usata per distribuire il carico e il traffico sui vari siti web e servizi attivi).

Per concludere questa breve analisi sui sistemi ICT, va aggiunto che il tutto è sito all'interno del nuovo CED di Campus, dotato di sistemi di connessione in Fibra Ottica con l'esterno e l'interno della rete d'Area, in un edificio dedicato e servito sia da un gruppo di continuità, in grado di sopportare un down time elettrico fino all'intervento di uno dei gruppi elettrogeni dell'Area.

Tornano al sistema informativo che si è scelto di uti-

lizzare, Joomla! è un software di Content Management System (CMS) per la realizzazione di siti web, scritto in linguaggio PHP, pubblicato con licenza open source GNU GPL v.2 e distribuito sotto forma di pacchetto compresso.

La sua installazione è molto semplice e veloce è sufficiente scompattare l'archivio in una cartella pubblica di un server Web dotato di supporto a PHP ed avere a disposizione un database MySQL per i dati del programma. Dopo un processo di installazione di pochi minuti, che spiegheremo molto rapidamente di seguito (sono reperibili molte guide online che aiutano l'utente anche meno esperto ad installare il sistema), il sito è operativo.

Per installare Joomla! basta eseguire queste semplicissime operazioni:

- Scaricare l'ultima versione di Joomla! dal repository ufficiale (<http://www.joomla.org/download.html>)
- Decomprimere la cartella sul pc e caricarla via FTP sul web server per copiarci i script di Joomla! (o effettuare tutto direttamente sul server web che ospita il nostro sito).
- Iniziare l'installazione guidata digitando sul browser "www.nostrosito.xx/NomeCartellaCaricata", inserendo le informazioni richieste (attenzione ai permessi di scrittura relativi alla cartella e alla necessaria implementazione del database di servizio che è necessario al CMS, inoltre è necessario alla fine dell'installazione cancellare o rinominare la cartella relativa ai file d'installazione!)

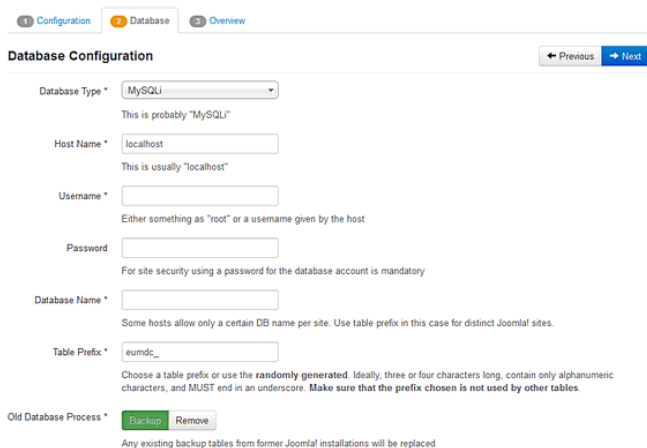


Fig. 2 Installazione di Joomla.

### 3 Progettazione scheda e database del prodotto di ricerca

Come detto inizialmente lo scopo del nostro lavoro, e del progetto GENESI stesso, è quello di trasferire i prodotti della ricerca, che nel nostro caso sono proposti da CNR e Uniroma3, verso un network di imprese e enti per migliorare il comparto industriale del nostro paese.

Per fare ciò l'operazione prioritaria da svolgere era quella di censire i prodotti a disposizione degli enti al fi-

ne di proporli alle aziende e instaurare un primo interesse sulle ricerche; è stato necessario pensare e prevedere un sistema che potesse censire tutta la produzione scientifica trasferibile sul mercato industriale legato ai sistemi di produzione e che con le informazioni strettamente necessarie (evitando di fornire troppe informazioni anche per proteggere la proprietà intellettuale delle ricerche) facesse comprendere l'ambito d'interesse e il potenziale (anche in relazione ai costi e le tempistiche di relazione) del prodotto della ricerca.

Il nostro lavoro è stato quindi quello di pensare a dei templates che potessero raccogliere le informazioni necessarie a caratterizzare un prodotto e che avessero il giusto livello di completezza d'informazione da far comprendere il potenziale ed attrarre potenziali investitori.

Sono state progettate due tipologie di schede, una ad un livello macro, cioè che descrive sommariamente il prodotto (attraverso un TITOLO e un breve ABSTRACT) e inquadra l'ambito di applicazione (con KEYWORDS e CAMPI DI APPLICAZIONE specifici) e un'altra molto più completa, da presentare in un secondo momento ad interesse dimostrato.

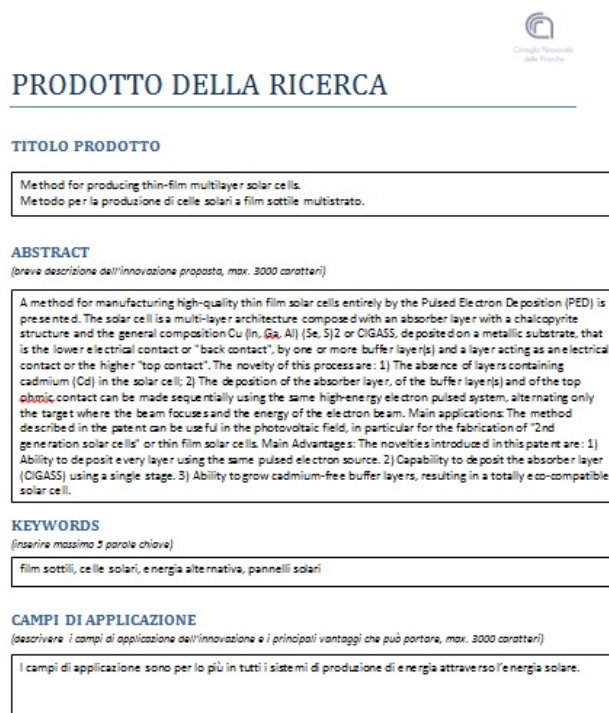


Fig. 3 Scheda prodotto breve.

La scheda completa, oltre a comprendere tutto quello presente su quella breve (Figura 3), aggiunge molte informazioni e campi che servono a specificare il prodotto:

- il PROPRIETARIO o REFERENTE del prodotto con tutti i suoi contatti (ISTITUTO/DIPARTIMENTO DI

AFFERENZA, NOMINATIVO, TELEFONO, MAIL) e il relativo GRUPPO DI RICERCA;

- il TRL (Technology Readiness Level) attuale e atteso alla fine della collaborazione;
- lo STATO DELL'ARTE (attuale e atteso in relazione ad un piano di investimenti studiato e concordato);
- la presenza di BREVETTI correlati al prodotto;
- il POTENZIALE DI BUSINESS e i POSSIBILI VANTAGGI COMPETITIVI;
- il PREZZO/COSTO INDICATIVO, cioè il prezzo di acquisto/produzione o l'eventuale dell'investimento previsto per la prototipazione e/o l'industrializzazione/ingegnerizzazione dello stesso;
- le PUBBLICAZIONI CORRELATE che danno valore alla ricerca così come il NUMERO DI APPLICAZIONI (qualora fossero già presenti);
- gli eventuali PARTNERS scientifici o industriali, diversi dall'ente proponente, coinvolti e la POTENZIALITA' DI SPIN/OFF (indicati su una scala di valori che va da 0=impossibile a 5=facilmente realizzabile).

Lo schema di scheda è mostrato in Figura 4.

DETTAGLIO OFFERTA DI TECNOLOGIA	
titolo	titolo ed Abstract del progetto
Categoria	ICI, IRI/ICTM
ASSEGNO / Ente	Università Roma IHC
Livellamento / settore	nanos
PRODOTTO	
Campi di applicazione	lettura ambientale, sistemi di sicurezza
Altri possibili campi di applicazione	possibile ambito di ricoperta industriale
Competenza specifica	algoritmi, modelli, applicazioni e test sperimentali
ICL tecnologia / materiali	
ICL tratta atteso	
stato attuale	
stato attuale	stato attuale della ricerca nel settore specifico ed eventuali applicazioni industriali dei risultati della ricerca già presenti sul mercato
Ulteriore stato del prodotto	Ulteriore stato del prodotto di innovazione, rispetto allo stato attuale, dei risultati della ricerca ottenuti
risultati tecnici tecnologici e/o eventuali vincoli all'operatività	Ulteriore qualitativa, rischi tecnologici ammessi in progetto ed eventuali vincoli all'operatività (necessità di laboratori attrezzati, di servizi ad hoc da sviluppare, etc.)
eventuali brevetti correlati	
possibili vantaggi competitivi	Ulteriore qualitativa dei possibili vantaggi competitivi potenzialmente offerti (ingegnerizzazione, prototipazione, mini-centri di gestione, etc.)
potenziale di business	individuale, in linea generale, opportuno incoraggiamento del prodotto/processo/ servizio offerto (sistema della domanda potenziale)
Principale/altro indicativo (di)	rispetto principalmente al prodotto, nel caso in cui si possa trattare di servizio, idoneo al stato del mercato previsto per la prototipazione e/o l'industrializzazione/ingegnerizzazione dello stesso
note	
referente	Contatto responsabile gruppo di lavoro
nominativo	nominativo team di progetto
versione	versione precedente

Fig. 4 Scheda prodotto completa

Tutte le considerazioni fatte, sono poi state riportate sui DB (Data Base), creati all'interno dei due server SQL replicati (uno Master e l'altro Slave) utilizzati dal progetto, realizzandone di diversi per i templates descritti e contenenti anche ulteriori informazioni utili per filtrare le schede e per implementare il matching offerta tecnologica / ricerca tecnologica.

Questi database verranno poi popolati di volta in volta dalle immissioni che avvengono dai form realizzati per censire i prodotti della ricerca e pubblicati sul sito web di riferimento; inoltre saranno anche la base per le pagine dinamiche realizzate (e spiegate di seguito) che mostreranno i risultati della ricerca, alle aziende interessate.

## 4 Realizzazione interfacce web

Lo scopo del sistema di web è quello di supportare l'azione di facilitazione dell'incontro tra domanda ed offerta di tecnologia. Riprendendo lo schema in Figura , il si-

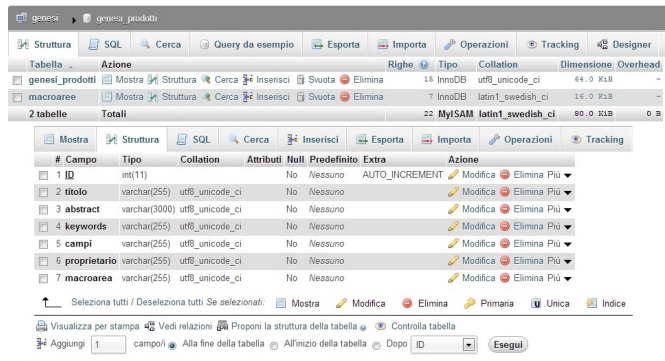


Fig. 5 Data Base, maschera di PhpMyAdmin.

stema web supporta l'azione del box verde "Facilitatore". Chiaramente un sito web statico non permette di svolgere tale attività e va detto subito che il sistema di content management progettato e sviluppato va visto come sistema che coadiuva le altre azioni di facilitazione previste all'interno del progetto. Il sistema web ad esempio raccoglierà gli spunti e sintetizzerà gli incontri tra domanda ed offerta tramite altri canali quali quelli della survey, dei meetings e degli sportelli all'impresa. Il sistema web inoltre ospiterà un'altra azione di facilitazione prevista dal progetto che si sintetizza nell'algoritmo di matching progettato ed implementato nella seconda parte del progetto. Nel seguito di questa sezione verranno descritti i sistemi di inserimento e di visualizzazione dei prodotti della ricerca che sono stati progettati proprio nell'ottica di supporto alle azioni di incontro tra domanda ed offerta di ricerca. I componenti dell'applicazione di content management descritti sono raggiungibili a partire dall'indirizzo web <http://genesi.mlib.cnr.it/> e navigando nella scheda del menu denominata "tech offer".

### 4.1 Sistema di inserimento prodotti della ricerca

La funzionalità di inserimento prodotti della ricerca è accessibile solo ad utenti autorizzati in possesso di credenziali di accesso.

Tramite l'accesso a tale funzionalità, mostrata in Figura 6, i prodotti possono essere agevolmente inseriti nel database del sito. La funzionalità replica ed implementa la scheda prodotti della ricerca di tipo base. I prodotti inseriti possono avere la provenienza CNR o Uniroma 3 (i partner di ricerca del progetto GENESI).

## 5 Sistema di visualizzazione delle schede prodotti

Uno dei problemi più noti nella classificazione delle innovazioni e dei prodotti della ricerca sono le classificazioni di tali prodotti. Spesso esistono termini simili per rappresentare la stessa cosa e chi fa una ricerca può o non trovare nulla oppure trovarsi di fronte a elenchi interminabili. Spesso lo stesso termine può riferirsi a campi totalmente diversi e quindi la ricerca del prodotto di in-

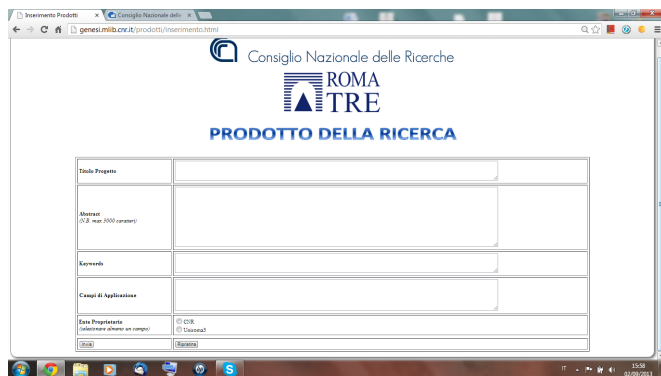


Fig. 6 Funzionalità di inserimento prodotti.

teresse può essere faticosa. In effetti dare degli strumenti per ricercare in modo efficace e semplice ciò che si cerca è di fondamentale interesse per facilitare il trasferimento tecnologico. A questo aggiungiamo che spesso chi fa la ricerca non è un ricercatore di professione e non ha una idea precisa di cosa ricerca.

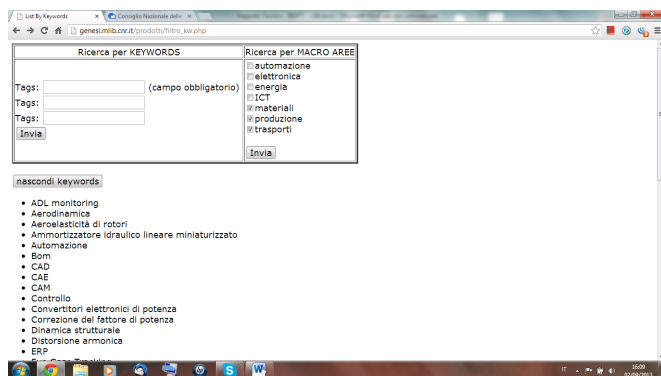


Fig. 7 Funzionalità di ricerca. Dettaglio elenco parole chiave.

Per superare queste problematiche è stato progettato un sistema che possa agevolare la ricerca sia a parola libera che tramite schemi di classificazione. Il sistema si fonda sul concetto di campo chiave per la ricerca libera e su quello di macroarea per la ricerca su classificazione. Tali concetti sono stati implementati come campi del database. A ciascun prodotto possono essere associati più keyword. Chi fa l'inserimento può inserire più parole chiavi che descrivono il proprio prodotto, ciascuna separata da un delimitatore “,”.

Il sistema di ricerca è stato implementato in PHP, javascript, html e tecnologia ajax. La pagina di ricerca legge tutte le parole chiavi che sono state inserite nel database e crea un'unica lista di keyword (i duplicati vengono eliminati). Tale lista può essere visualizzata agevolmente tramite il tasto *mostra/nascondi keywords* (vedi Figura 7).

La pagina di ricerca presenta due aree. Nell'area di sinistra è implementata la funzione di ricerca per keyword. È possibile ricercare inserendo fino a 3 keywords. La particolarità del sistema è quella che esso permette di “guidare” la ricerca libera. Ciascun campo di input infat-

ti contiene la lista caricata dinamicamente dal database, delle keywords contenute in tutti i prodotti della ricerca. Il campo di ricerca è un campo html *autocompletamento* sviluppato con la tecnica ajax; quindi, quando si inizia a scrivere una keyword il sistema aiuta l'inserimento presentando dinamicamente l'elenco delle parole chiavi che contengono la parola che si sta scrivendo. L'esempio in Figura 8 mostra il funzionamento del sistema. Si vuole ricercare un prodotto che abbia come parola chiave *vehicle routing* e non appena si scrivono le lettere *ve* viene mostrato l'elenco delle parole chiavi che contengono la sequenza di lettere *ve*.

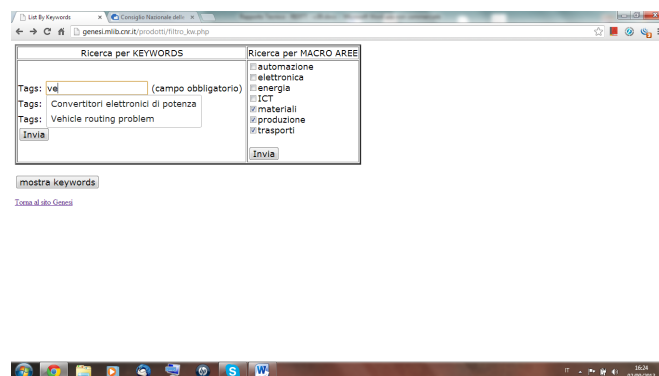


Fig. 8 Funzionalità di ricerca autocompletamento e macroaree.

Tale sistema elimina completamente la possibilità di ricercare parole chiavi inesistenti o grammaticalmente errate. Inoltre suggerisce quali parole chiavi ricercare se l'operatore si trova nel dubbio di cosa ricercare. La parte destra della schermata implementa invece la classificazione in macroaree. Tale funzionalità è più classica ma per questo non meno importante. In particolare funge da completamento alla ricerca libera per parola chiave. Nel database, un campo macroaree permette di inserire la classificazione in macroaree. Anche in questo caso la classificazione può essere non univoca e chi inserisce il prodotto può inserire più aree di classificazione separandole con un delimitatore “,”. Le macroaree sono le seguenti:

- Automazione
- Elettronica
- Energia
- ICT
- Materiali
- Produzione
- Trasporti.

La lista è caricata in modo automatico da una tabella del database chiamata “macroaree” tramite php e visualizzata tramite il costrutto html *input type='checkbox'* che permette una selezione multipla delle classificazioni (come si vede da Figura 7).

L'implementazione dinamica basata su tabella di database permette di poter inserire, modificare, o cancellare macroaree agendo semplicemente sui record della tabella. L'azione sul tasto invia permette di sottoporre la query composta a seconda delle selezioni dell'operatore e di mostrare i risultati in tabella come da immagine in Figura 9. Da tale pagina si possono mostrare tutti i dettagli di un singolo prodotto semplicemente cliccando sul nome (Figura 10).

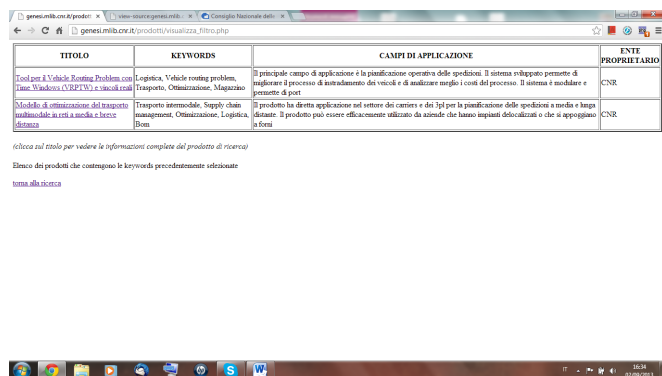


Fig. 9 Elenco prodotti della ricerca filtrato da selezione.

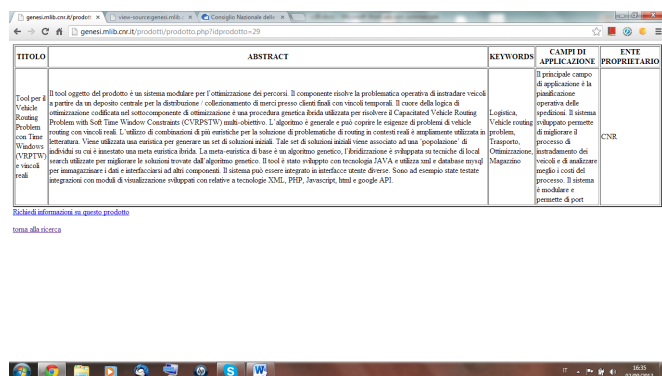


Fig. 10 Dettaglio di un prodotto della ricerca.

Il ciclo di collegamento tra offerta e domanda si completa con la funzionalità di contatto tra offerta e domanda. Nella pagina di visualizzazione del singolo prodotto della ricerca esiste un link che permette di richiedere via email informazioni dettagliate sul prodotto della ricerca. Come si vede dalla Figura 11 il link genera una email di contatto con tutti i dettagli del prodotto selezionato.

## 6 Newsletter

Per portare avanti le azioni di divulgazione delle informazioni, dei risultati del progetto e delle collaborazioni ricerca-impresa che si realizzano, è stata prevista la realizzazione di una newsletter che potesse dare ampio spazio ai progetti e alle esperienze di lavoro svolte.

Un ottimo prodotto open source che svolge la funzione di publishing software è **Scribus** (<http://www.scribus.net/>).

Scribus è un software dedicato al mondo dell'editoria: si tratta infatti di un'applicazione per il Desktop Publi-

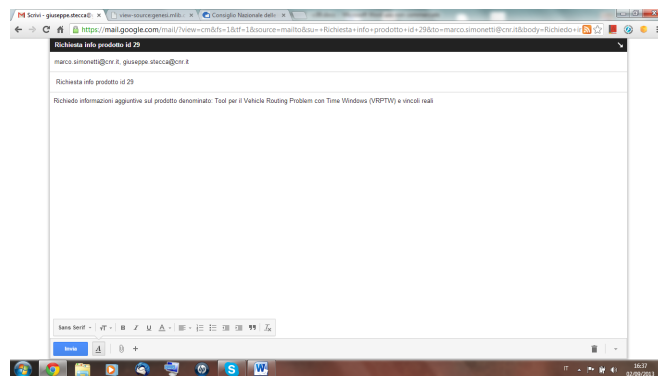


Fig. 11 Scheda email di contatto a valle della visualizzazione di un prodotto.

shing rilasciata con licenza open source ed utilizzabile gratuitamente, sia in ambienti Windows che Linux. Grazie a questo programma è possibile creare libri, newsletter, brochure, volantini e numerose altre tipologie di contenuti editoriali utilizzando strumenti appositamente studiati per realizzare versioni digitali di materiale da trasferire successivamente su carta.

Uno dei principali vantaggi di un software di Desktop Publishing è infatti proprio la disponibilità di tool ed accessori dedicati prettamente all'editoria. Grazie a queste tipologie di applicazioni è possibile curare svariati dettagli non accessibili invece in software per la produttività e l'ufficio, come ad esempio le numerose suite che offrono componenti per la scrittura digitale.

## 7 Survey

Una parte fondamentale di tutto il procedimento di trasferimento tecnologico, è quella che a monte prevede l'analisi della stato dell'arte dei sistemi industriali e dell'indagine sugli argomenti di maggiore interesse per le imprese e industrie del territorio in considerazione (o anche nazionale visto che l'indagine mira a fotografare lo stato nazionale).

Per fare ciò si è studiato e articolato un questionario o survey, congiuntamente con gli altri partner, che potesse raccogliere e mettere a fattor comune gli interessi mostrati dalle aziende nell'ambito dei prodotti trattati dal progetto GENESI.

Il questionario, logicamente è stato previsto in formato sia cartaceo che elettronico per poter raggiungere il più ampio bacino di utenti anche con livelli di competenze e possibilità informatiche differenti. Per permettere una corretta analisi e manipolazione dei dati, i questionari cartacei verranno poi riportati sui sistemi informativi del progetto in modo d'avere una situazione il più possibile identificativa delle necessita e bisogni del comparto industriale.

A questo fine abbiamo studiato molti prodotti, tutti di natura open source, come spirito adottato dal gruppo di lavoro e al fine di non pesare sul budget di progetto, ed

abbiamo individuato un'applicazione che facesse al caso nostro; la scelta, anche in base alle molte references sia in ambito educational che government (basti pensare che lo stesso ISTAT - Istituto nazionale di statistica, se n'è servito in più di un'occasione oltre a numerose università e istituti di alta formazione) è ricaduta su **LimeSurvey** (<http://www.limesurvey.org>).



Fig. 12 LimeSurvey website.

LimeSurvey è un programma che permette di creare in modo semplice ed efficace indagini di tipo online, a cui possono partecipare decine di migliaia di intervistati. Il software permette di creare indagini auto-guidate per facilitare la compilazione da parte dell'intervistato, anche in caso di percorsi logici più complessi; inoltre oltre alla creazione dei questionari è pensato anche per la raccolta dei dati, l'analisi statistica, il tutto fatto in modo molto semplice utilizzabile anche senza essere esperti di programmazione, in modo da dare possibilità di visione a tutti i partners del progetto. Il software è scritto in PHP e supporta sia MySQL che Microsoft SQL Server utilizzai sia per il sistema stesso che per i dati da memorizzare.

L'installazione della piattaforma è piuttosto facile; LimeSurvey richiede caratteristiche minime, di server: circa 9 MB di spazio libero sul disco rigido per gli script e la presenza sulla macchina di destinazione dei servizi PHP, MySQL (o Microsoft SQL Server), quindi è un sistema molto "snello".

Essendo un sistema open source e per giunta supportato con interesse dalla comunità, esistono anche molti plugins interessanti come ad esempio le estensioni estensione GD-library (per le captcha nelle form di registrazione) e l'estensione LDAP-library (per importare i token attraverso LDAP).

Per installare la piattaforma è necessario collegarsi all'area Download del sito ufficiale e scaricare la versione più aggiornata del software (quella presente sul nostro sistema attualmente è la 2.0).

Visto l'interesse che il sistema ha riscosso si è deciso di realizzare il sistema, non solo per il solo lavoro di progetto, ma di installare un servizio su dei sistemi centralizzati e in utilizzo anche per tutta l'Area della Ricerca Roma1, creando diverse istanze specifiche per ogni utilizzatore.

Il sistema LimeSurvey, configurato e funzionante è raggiungibile all'indirizzo <http://survey.mlib.cnr.it>, dove sono presenti diversi questionari, tra cui quello legato al progetto GENESI (compilabile cliccando su Questionario Aziende Progetto GENESI).



Fig. 13 Questionario Aziende progetto GENESI.

Dopo aver scaricato il pacchetto compresso, occorre estrarre i file in una cartella e poi modificare il file *config.php*, utilizzando il proprio editor di testo preferito; questo file contiene infatti le impostazioni di configurazione della piattaforma: per procedere con l'installazione è necessario modificare i parametri di configurazione base, inserendo le impostazioni di connessione al database (che abbiamo creato nella nostra macchina o nel nostro database server).

E' possibile anche scaricare il pacchetto direttamente sulla macchina web di destinazione, usando un programma FTP, copiare l'intero contenuto della cartella di installazione decompressa, e poi successivamente lanciare attraverso web l'installazione guidata.

Ipotizziamo ad esempio di aver caricato il pacchetto di installazione nella cartella "limesurvey" dello spazio relativo al generico dominio "www.miodominio.com", per procedere con l'installazione basta collegarsi all'indirizzo "http://www.miodominio.com/limesurvey/admin/install/index.php". Se tutto è stato configurato correttamente ci verrà richiesta l'autorizzazione a creare il database e/o creare le tabelle al suo interno. A questa richiesta occorre rispondere affermativamente, avviando così di fatto l'installazione. Al termine del processo è necessario rinominare o cancellare la cartella "install".

Il sito ufficiale di LimeSurvey mette a disposizione degli utenti una guida completa all'interno dell'area documentazione, parzialmente tradotta anche in lingua italiana. La procedura di installazione è tra i contenuti tradotti ed è descritta in maniera molto chiara ed esaustiva. Terminata l'installazione, basta lanciare il browser e inserire l'indirizzo per raggiungere il file *admin.php* (nel nostro esempio occorre digitare "http://www.miodominio.com/limesurvey/admin/admin.php").

In questo modo si aprirà la pagina di login in cui inserire le credenziali di accesso (le credenziali di default sono: *user=admin; password=password*).

Nell'area amministrazione è presente una barra di menu che consente di creare utenti e gruppi, gestire data-



base e template e creare indagini. È presente anche una piccola guida in linea, attivabile cliccando sul simbolo del punto interrogativo.

L'interfaccia di LimeSurvey è semplice e intuitiva: la funzione basilare del software è relativa alla creazione e gestione di indagini. Per cominciare a lavorare basta premere sul pulsante "crea nuova indagine. Comparirà una schermata che ci chiederà di specificare lingua, titolo, breve testo descrittivo, messaggio di benvenuto e amministratore di riferimento.

Oltre a queste informazioni base, per ogni indagine è possibile specificare varie opzioni aggiuntive, scegliendo ad esempio se visualizzare una pagina per ogni domanda o tutte le domande in una sola pagina; indicando se è possibile dare risposte anonime o meno; specificando i testi delle email di invito, sollecito e registrazione all'indagine; formalizzando la lingua di riferimento e la scadenza della ricerca, e così via.

Dopo aver creato l'indagine comparirà una barra di menu aggiuntiva contenente tutte le funzioni connesse alla predisposizione e pubblicazione del questionario stesso. Il software è strutturato in modo da associare ad un'indagine un gruppo di utenti e poi le relative domande.

Dobbiamo quindi premere sull'apposito pulsate per creare il gruppo di utenti a cui è indirizzato il questionario e poi premere sul pulsante "aggiungi domanda".

Si aprirà una pagina in cui inserire la domanda specificando codice, testo, eventuale nota di spiegazione e tipo di domanda.

Rispetto a quest'ultima caratteristica la piattaforma offre la possibilità di inserire varie tipologie di domande (aperte, a scelta multipla con elenco delle modalità di risposta in forma di menu a tendina o pulsante radio, a scelta multipla con commento, ad attribuzione di punteggio secondo varie scale numeriche, con classificazione delle modalità di risposta predefinite ecc.).

Solo dopo aver aggiunto almeno una domanda è possibile attivare l'indagine premendo sull'apposito pulsante. L'attivazione è un'operazione da compiere quando non sono previste ulteriori modifiche sostanziali al questionario. Una volta attivata l'indagine, infatti, non sarà più possibile aggiungere o eliminare i gruppi di utenti e alcune tipologie di domande.

Una volta attivata l'indagine la piattaforma consente anche di personalizzare e inviare l'email di invito a partecipare e l'eventuale email di sollecito ai destinatari, qualora si avesse già una lista di possibili interessati.

Per gestire l'indirizzario, LimeSurvey prevede l'inserimento manuale dei dati dei destinatari e/o l'importazione (e poi all'occorrenza l'esportazione) di un database in formato .CSV. È inoltre possibile assegnare manual-

mente o generare automaticamente codici e password di identificazione per gli utenti.

Per quanto riguarda la somministrazione del questionario è possibile ottenere, cliccando sull'apposito pulsante, una versione stampabile dello stesso. La piattaforma genera anche una maschera di inserimento manuale dei dati. Grazie a queste utili funzioni LimeSurvey è utilizzabile anche per la gestione di indagini tradizionali offline.

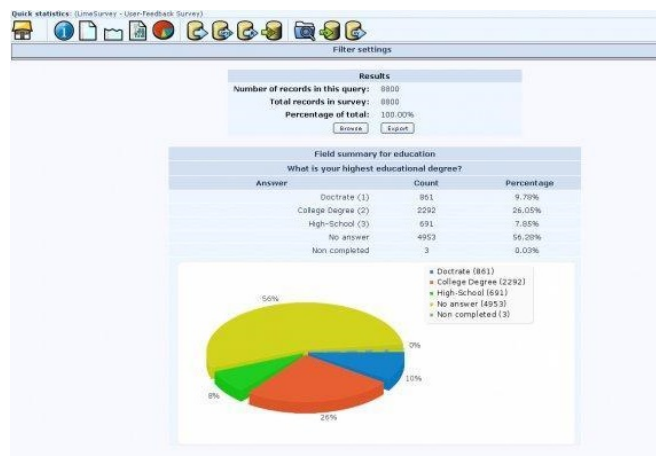


Fig. 14 LimeSurvey: Risultati e Statistiche.

Una volta attivata l'indagine è possibile visionare le risposte fornite dagli utenti e le relative statistiche di sintesi premendo sul pulsante "sfoglia risposte al questionario".

Per quanto riguarda la rappresentazione grafica delle statistiche LimeSurvey integra anche la libreria JGraph. Grazie a queste semplici operazioni le statistiche saranno disponibili sia in formato numerico sia in formato grafico e per giunta è possibile tenere traccia anche dei questionari non completati (anche al fine di comprendere eventuali problematiche/difficoltà su specifiche domande) e del tempo impiegato nel rispondere.

LimeSurvey è quindi uno strumento molto utile per gestire ricerche di mercato online. La piattaforma è ricca di funzioni e strumenti che supportano l'intero processo di realizzazione della ricerca: dal questionario all'elaborazione dei dati.

## Riferimenti

- 1 Joomla! <http://www.joomla.org/>.
- 2 LimeSurvey <http://www.limesurvey.org/>.
- 3 MySQL <http://www.mysql.it/>.
- 4 PHP <http://php.net>.
- 5 PhpMyAdmin <http://www.phpmyadmin.net>.
- 6 Java <http://www.java.com/it/>.
- 7 Scribus <http://www.scribus.net/>.

## 8 Ringraziamenti

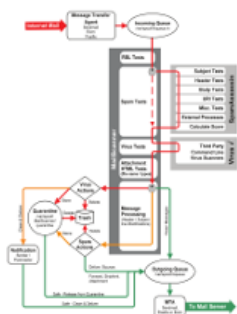
Il progetto GENESI, relativo al bando nazionale RIDITT (Rete Italiana per la Diffusione dell'Innovazione e il Trasferimento Tecnologico alle imprese) finanziato dal Ministero dello Sviluppo Economico (MiSE).

Si ringrazia per la collaborazione il sig. **Luca Ianniello**, gestore della piattaforma web dell'Area della Ricerca Roma 1 di Montelibretti per la disponibilità e l'ausilio dato durante l'installazione della piattaforma LimeSurvey sui sistemi centralizzati di Campus e il dott. **Salvatore Fiorino** per il supporto allo sviluppo di parti del codice di programmazione utilizzato nei vari sistemi informativi.



## Setup of a clustered antispam and antivirus service based on Mailcleaner<sup>®</sup> suite<sup>†</sup>.

Andrea Lora,<sup>a</sup> Luca Ianniello,<sup>a</sup> Augusto Pifferi.<sup>a</sup>



In order to provide an effective antispam and antivirus service to support the email servers delivered at C.N.R. - Area della Ricerca RM1 we setup a clustered installation of Mailcleaner, an open source suite based on MailScanner. Here we will describe the setup activity and the configuration of all components. For more information about Mailcleaner and MailScanner refer to the websites of both projects <sup>1,2</sup>.

**Keywords:** Antispam, antivirus, spamassassin, mailcleaner.

### 1 Introduction

On late summer 2012 the research campus “Area della Ricerca RM1” of C.N.R. didn’t yet had its own email server, all the research institute inside the Area were using an external email server hosted at “Area della Ricerca di Tor Vergata”. Because of the growth of the number of used mailboxes, we started to deploy a new email server. The first step was to evaluate different open source antispam and antivirus services available on the web.

We choose Mailcleaner, for its detailed ajax admin interface, its clustering capabilities, the availability of a quarantine area and because it’s open source!

### 2 Architecture

We planned to setup a redundant architecture dual-frontend dual-backend, in order to implement HA (High Availability) and load balancing. The redundant design allows us to operate maintenance of a single machine without putting the entire system offline. At frontend layer we simply exploit the capability of DNS to define more than one MX records for each domain, in case of system outage of the first server, email delivery is granted via the second one. At backend layer we used load balancing techniques issued by our firewall. The overall architecture is shown in Fig. 1

The balancer is not a single point of failure (SPoF) because it’s implemented with two nodes, ensuring maximum availability, on the other hand the balancing is not configured per session but per source ip address, so often a mailhub keeps many more sessions than the other one, as shown in Tab. 1.

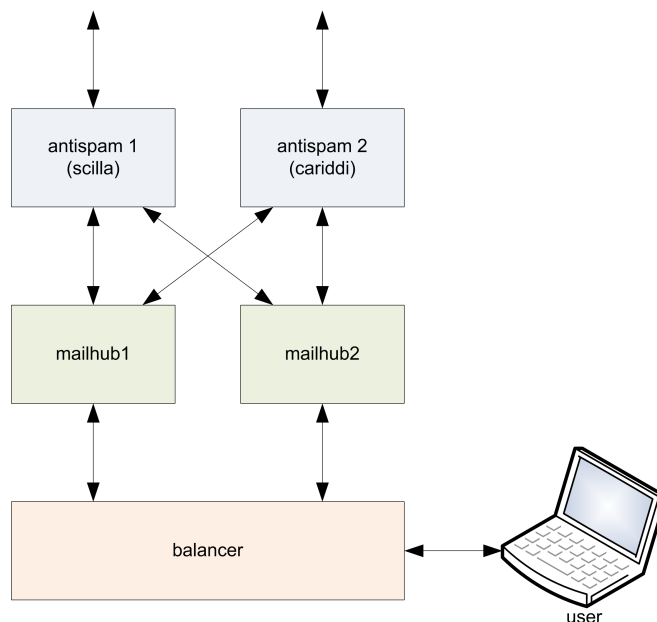


Fig. 1 Overall architecture

### 3 Installation

Mailcleaner is distributed as a self-installable distribution, based on Debian with latest cutting-edge kernel re-

<sup>a</sup> CNR - Istituto di Cristallografia, Strada Provinciale 35/d, Montelibretti, Italia

Creative Commons Attribution - Non commerciale - Condividi allo stesso modo 4.0 Internazionale

<sup>†</sup> rapporto tecnico IC 13/10 registrato con numero di protocollo IC/1880 del 20/12/2013

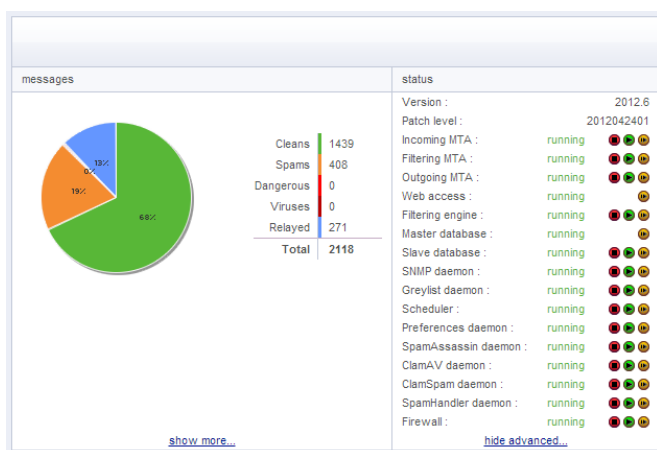
<pre> root@mailhub1:~# vnstat Database updated: Mon Oct 28 10:47:48 2013 eth0 since 10/01/13 rx: 86.68 GiB tx: 45.48 GiB total: 132.16 GiB </pre>	<pre> root@mailhub2:~# vnstat Database updated: Mon Oct 28 10:47:44 2013 eth0 since 10/01/13 rx: 152.85 GiB tx: 96.05 GiB total: 248.90 GiB </pre>
---	--

**Table 1** Network usage.

lease. The setup process is quite easy. Once installed on the first machine we installed it again on a second machine in order to setup a clustered deployment.

Clustering feature is really powerful and well described on the official wiki.

Once finished, the system is up and running, the only package we had to install as separate process is ClamSpam. We installed it using the unofficial signatures released by Sanesecurity.<sup>3</sup> A status window of the overall system is immediately available, see Fig. 2. It is useful to get an overview on the status of processes, messages handled, number of spam or dangerous messages received and resources used.



**Fig. 2** Status windows.

## 4 Initial Configuration

The first configurations include IP addressing, DNS and time server configuration. Then we started to configure all those domains whose antispam and content filtering system should be on our machines. We called our servers Scilla and Cariddi, as the two epic monsters which generated storms between Sicily and Italy destroying passing ships.

Mailcleaner sets a value for spam score of an email, and when it's greater than the spam threshold, it is possible to take different actions: tag, quarantine or delete the message. We opt to simply tag email and deliver it, in order to avoid that false positives can cause excessive delay in email delivery or loss of messages. Quarantine action will be taken into account in case of a dangerous

content inside the message.

This action will isolate dangerous messages inside the Mailcleaner system, avoiding to deliver them to user's mailbox. In the meanwhile the user receives a notification that messages are retained in his quarantine area, so he can decide to review and optionally relay them.

Another feature we choose to not use is greylisting. This is a method to defend email users by spam. A Mail Transfer Agent (MTA) using greylisting will "temporarily reject" any email from a sender it does not recognize. If the email is legitimate the originating server will try again to deliver it after a delay and, if sufficient time has elapsed, the email will be accepted. However this introduce an unacceptable delay that we choose to avoid.

In order to secure all connections to web interface for administrators and users while they browse their quarantine, we also installed a SSL certificate.

## 5 Description of Components.

To understand the architecture of mailcleaner we need to describe the activities to be done by a spam server to do its job.

The first task is to receive the email from the external sources. This is the reason why the spam servers are often configured as MX records for a domain. During this task the spam server stores the incoming email into a queue. The second task is to check the queue of inbound messages and process them. At this stage further actions are evaluated for the spam retention. The third task is to delivery the processed message to the last server.

The first task is accomplished by the incoming queue, managed from an instance of the Exim daemon named Stage1, it is responsible for accepting messages received from an external internet MTA and verifying some basic rules, such as consistency of sender domain and recipient mailbox. Stage1 can also operate some network checks, for example it can check if the sender MTA is publically blacklisted.

The second task is carried out by the Filtering queue, another instance of Exim named Stage2. It's probably the most important process because it must check the content of every received email in order to mark spam messages. This decision is taken through different algorithms that take into account lots of parameters such as known signatures, blacklists checks or content of the body message of the email. We'll look into that deeply in chap. VI.

The last task is the final delivery. Another instance of Exim, named Stage4, is responsible of delivering messages to the recipient email servers according to the result of the filtering action.

## 6 The Filtering Stage

It is possible to apply up to six different filters to distinguish spam from ham messages, which can be ordered

according to settings applied by administrators.

1. Trusted sources: filter that accept email originated by specific trusted SMTP servers or SPF capable domains. It can also use information from DNS white list services.
2. Bayesian filters: the one used by mailcleaner is Bogofilter. It is an email filter that classifies messages as spam or ham (non-spam) by a statistical analysis of the message's header and content (body). The program is able to learn from the user's classifications and corrections.
3. Clamspam: it is a signature-based filter that can identify phishing attempts, dangerous links or contents, malware, fakes and other generic spam. We used signatures from Sanesecurity.<sup>3</sup>
4. PreRBL: it is a Real-time Blackhole Lists system based on a network of databases containing IP addresses of servers used to originate or forward spam messages. Lists are kept up to date in order to avoid false positives, however our implementation only add a hit (positive value on the overall score) when a RBL match is found.
5. UriRBL: it's another Real-time Blackhole Lists system, based on databases of web sites uri that have appeared in unsolicited messages. Unlike most lists, UriRBLs are not lists of message senders, because web sites seen in unsolicited messages tend to be more stable than the rapidly changing botnet IP addresses used to send the vast majority of them.
6. Spamassassin: it's probably the most strong engine for spam filtering. It applies lots of checks on incoming messages including RBLs previously discussed and:

- Distributed Checksum Clearinghouse (DCC), Razor and Pyzor, three similar Hash-Sharing Systems. Every time a spam message is detected, a checksum of the message is calculated and reported to an online database. When an email server queries that online database about the hash of a message, it compares the hash with all strings contained and returns a result.
- SPF: Sender Policy Framework (SPF) is an email validation system designed to prevent spam by detecting email spoofing, a common vulnerability, by verifying sender IP addresses. SPF allows administrators to specify which hosts are allowed to send email from a given domain by creating a specific SPF record (or TXT record) in the Domain Name System (DNS). Email exchangers use the DNS to check that email from a given domain is being sent by a host sanctioned by that domain's administrators. Sender Policy Framework is defined in IETF publication RFC 4408.<sup>4</sup>

- DKIM: DomainKeys Identified Mail (DKIM) is a method for associating a domain name with an email message, thereby allowing a person, role, or organization to claim some responsibility for the message. The association is set up by means of a digital signature which can be validated by recipients. Responsibility is claimed by a signer, independently of the message's actual authors or recipients, by adding a DKIM-Signature: field to the message's header. The verifier recovers the signer's public key using the DNS, and then verifies that the signature matches the actual message's content. DomainKeys Identified Mail is defined in IETF publication RFC 4871.<sup>5</sup>

### 6.1 Antivirus and Content Protection.

Mailcleaner can also process every inbound and outbound email message for virus identification and removal. Many antivirus software can be configured inside Mailcleaner, we choose to use the default one, Clam AV. Configuration is quite simple, automatic update is executed by default every hour in order to ensure maximum protection against virus threats.

Mailcleaner can also check for IFrame objects, web bugs and other malformed HTML inside the body of the emails and also scan the content of a compressed archive. In case of a positive the infected email will be cleaned (attachment removed and html sanitized) and a notification will be delivered to the recipient. The infected but cleaned message will be kept in quarantine for 15 days, so that recipients can check it and, optionally, release from quarantine to their mailboxes.

A flow chart of every single email that cross our servers is reported in Figure 3.

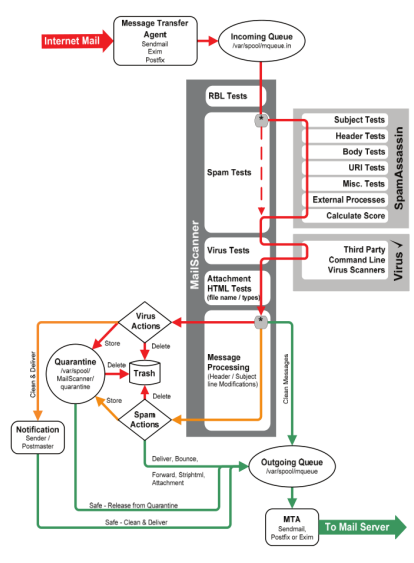


Fig. 3 Mail flow inside mailcleaner.

## 7 Custom Filters Setup

It is possible to fully customize all the settings of the system, included which checks operate and in which stage apply them. In order to avoid a too aggressive behavior against external servers not well configured or not well maintained, we choose to keep a more conservative, permissive profile.

RBL checks can be applied both at incoming or filtering stage, the main difference is that at incoming stage a message delivered from a blacklisted server is immediately rejected, while if operating the same check at filtering stage will result as a score assigned to every single message received. If the overall score of the message is greater than the system threshold then the message will be marked as spam, but delivered to the recipient.

We applied all the dynamic checks at spamassassin level, while, at incoming stage, we only check for some basic tests, such as the existence of the sender domain and the delivery mailbox.

## 8 Interaction with the Email Server

There are many interactions between Scilla and Cariddi and the email servers they deliver to.

First of all the way to balance all messages passing from a system to the other: mailcleaner allows to set many delivery server for any domain, so we added both mailhub1 and mailhub2 in load-balancing mode; inside the two hubs we installed HA-proxy, a load balancing daemon, in order to balance also the outbound traffic to the content filtering system.

The content filtering system is also used to scan outgoing messages, protecting our server from being blacklisted worldwide; users access their quarantine areas and their personal set of antispam rules authenticating themselves via an IMAP callout to the email servers; messages whose recipient does not exists are immediately rejected, this is possible via a SMTP query to the servers.

Another interaction between those two systems is about the notification of false positives and false negatives. Every user can notify the administrator that a received message has been incorrectly rated, so that administrators can instruct the bayesian filter (bogofilter) to correctly detect messages similar to the reported one.

This system has been also used to manually adjust some filters in order to avoid false positives, i.e. a spamassassin score is assigned every time the “soma” word or some variations of the word occurs in the body of an email. Unfortunately the italian word “somma” means “sum”, so we had lots of false positive matching of that rule. We identified that mistake and correct it manually lowering the score of the FRT\_SOMA rule.

## 9 The Learning System

As described before in this document Mailcleaner includes Bogofilter as a statistical analysis filter against spam. The statistical technique used by Bogofilter is known as the bayesian technique and its use for spam detection was described by Graham and Robinson.<sup>6–8</sup>

Configuring Bogofilter requires continuous training with huge amount of both spam and ham messages. At the present we didn't configure an effective fully-automated learning system, it will be argument for further enhancements of the system.

We would like to implement a system that automatically learn as ham any messages reported by user as false positives, and learn as spam messages reported as false negatives. Furthermore it should learn as spam (ham) messages whose spamassassin score is greater (less) than a threshold.

## 10 Monitoring

Mailcleaner provides itself a built-in console to overview the status of the whole system, including cpu load, disk and memory usage, number of messages in spools, status of all vital processes and a pie diagram that represents the number of messages processed by the system.

All elements of the architecture are automatically monitored via SNMP by our Zabbix platform,<sup>9</sup> in order to identify an issue as soon as possible helping troubleshooting and limiting the downtime and unavailability of a service.

At the present we monitor the number of messages in all three queues, the reachability of both servers and the usage of resources.

The most important OID to analyze that way is 1.3.6.1.4.1.2021.8.1.101.6. It returns an integer list containing the number of queued messages for the three stages ( e.g. :|190|4|26). Since the information about queues is relative to the local server we should remind to monitor both servers in the Mailcleaner cluster.

Choosing alert thresholds isn't immediate due to the different job every queue is in charge of. Incoming queue can grow larger after massive attacks or when distribution list messages are sent. Filtering spool contains messages processed by the antispam and the antivirus engines. These jobs are heavily cpu and io related so this queue can grow up easily if the system is in resources starvation. Outgoing spool is in charge to deliver to the final destination the email. If the messages cannot be delivered immediately ( e.g. temporary failure of destination host) they are kept here.

## 11 Conclusion

An effective antispam and content filtering system is the base of a well working email server architecture, be-

cause on one hand it keeps users safe from virus and free from spam, on the other hand it makes possible to save storage space and internet bandwidth, limiting the amount of resources wasted to handle useless messages.

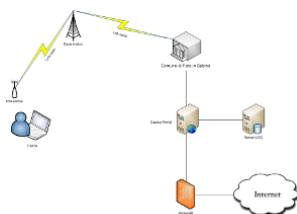
## References

- 1 Mailcleaner official online documentation, available at <http://www.mailcleaner.org/doku.php>.
- 2 Mailscanner official online documentation, available at <http://www.mailscanner.info/documentation.html>.
- 3 Sanesecurity project website <http://sanesecurity.com>.
- 4 M. Wong, W. Schlitt, **Sender Policy Framework (SPF) for Authorizing Use of Domains in E-Mail, Version 1**, Request for Comments: 4408.  
URL <https://tools.ietf.org/html/rfc4408>
- 5 E. Allman, J. Callas, M. Delany, M. Libbey, J. Fenton, M. Thomas, **DomainKeys Identified Mail (DKIM) Signatures**, Request for Comments: 4871.  
URL <https://tools.ietf.org/html/rfc4871>
- 6 P. Graham, "A plan for Spam", August 2002 .  
URL <http://www.paulgraham.com/spam.html>
- 7 P. Graham, "Better Bayesian Filtering", January 2003.  
URL <http://www.paulgraham.com/better.html>
- 8 G. Robinson, "Spam Detection", September 2002.  
URL <http://radio-weblogs.com/0101454/stories/2002/09/16/spamDetection.html>
- 9 A. Pifferi, G. Nantista, L. Ianniello, A. Lora, M. Simonetti, **Analisi e Implementazione di Sistemi per il Monitoraggio della Rete Wireless Relativa al Progetto ADD (Anti Digital Divide) e delle Infrastrutture di Campus AdR RM1., SMART eLAB 2 (2013) 1–9.** doi: [10.30441/smart-elab.v2i0.46](https://doi.org/10.30441/smart-elab.v2i0.46).



## “Wifi in Fara”, un progetto di access point federati per l’accesso a internet gratuito.<sup>†</sup>

Angelo De Simone,<sup>a</sup> Giovanni Agostini,<sup>a</sup> Luca Ianniello,<sup>a</sup> Andrea Lora,<sup>a</sup> Giuseppe Nantista,<sup>a</sup> Augusto Pifferi.<sup>a</sup>



Il presente documento descrive il progetto e la realizzazione di un sistema di punti di accesso wifi gratuito per i cittadini del Comune di Fara in Sabina. Illustreremo le linee guida di progetto, le postazioni selezionate e l’architettura di sistema.

**Keywords:** Captive Portal, Access Point, Mikrotik, VPN

### 1 Premessa

Un accesso ad internet disponibile gratuitamente per i cittadini è il primo passo che una amministrazione locale può compiere per abbattere il Digital Divide, il gap cioè che si pone fra i cittadini che hanno accesso e sanno utilizzare internet e quelli che non hanno queste possibilità.

In quest’ottica il Comune di Fara in Sabina, a fronte di numerose collaborazioni svolte in passato con il servizio reti dell’Area della ricerca RM1 del CNR, ha affidato allo stesso gruppo il compito di predisporre una piattaforma unica per consentire ai cittadini delle sue numerose frazioni di usufruire di questo servizio. Il progetto ha visto la luce nel luglio 2013 con il nome di “Wifi in Fara”.

### 2 Le postazioni

Su indicazione dell’amministrazione pubblica sono state predisposte postazioni nelle frazioni di:

Passo Corese, Prime Case, Corese Terra, Coltodino, Canneto, Borgo Quinzio, Fara In Sabina. Le postazioni sono sommariamente evidenziate nella mappa di figura 1.

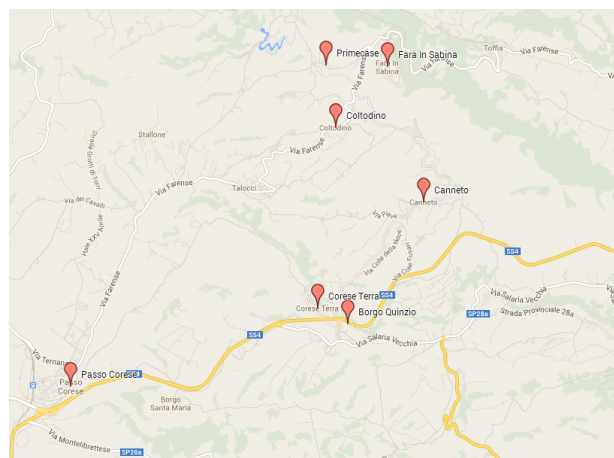


Fig. 1 Mappa delle postazioni.

### 3 Postazione tipo e rete di raccolta

In ogni postazione sono stati installati una CPE per la connessione alla rete wireless CNR e un Access Point per la distribuzione locale del segnale “Wifi in Fara”. Il primo collegamento viene realizzato in standard Hiperlan 802.11a, nell’intorno della frequenza 5Ghz, il secondo invece in standard Wifi 802.11b/g nell’intorno della frequenza 2.4Ghz. Lo schema di connessione di rete, rappresentato in Fig.2 prevede che da ogni postazione venga instaurato un tunnel VPN verso la sede principale, dove gli utenti, previa accettazione delle condizioni di servizio, possono navigare verso internet liberamente.

<sup>a</sup> CNR - Istituto di Cristallografia, Strada Provinciale 35/d, Montelibretti, Italia



Creative Commons Attribuzione - Non commerciale - Condividi allo stesso modo 4.0 Internazionale

<sup>†</sup> rapporto tecnico IC 13/09 registrato con numero di protocollo IC/1879 del 20/12/2013



L'accettazione delle condizioni avviene tramite il meccanismo del captive portal, come già discusso nel precedente articolo,<sup>1</sup> ma grazie alla nuova normativa in vigore, il D.L. 21 giugno 2013 n. 69, è sufficiente che le connessioni effettuate da un utente siano riconducibili ad esso esclusivamente tramite il mac-address dell'apparato con cui si connette alla rete. Ogni dispositivo può usufruire di una banda massima di 2Mbit/s in download e può scaricare un massimo di 300 MByte di dati al giorno.

Per raccogliere tutte le connessioni provenienti dai vari hotspot è stato realizzato un unico dominio di broadcast, usando le funzioni messe a disposizione dagli apparati Mikrotik utilizzati per le funzioni di CPE e concentratore. Ogni CPE instaura una connessione pptp verso il concentratore e successivamente i tunnel così instaurati vengono collegati in bridge lato CPE con l'access point e lato concentratore con il server di Captive Portal.

Configurazione delle CPE:

Instaurazione del tunnel pptp:

```
[admin@c_farasabina.wifi_primecase] > int pptp-client add name="pptp-out1" \
connect-to=10.10.235.2 user="fara2-primecase" password=***** \
profile=default-encryption add-default-route=no dial-on-demand=no \
allow=pap,chap,mschap1,mschap2
```

Creazione del tunnel eoip:

```
[admin@c_farasabina.wifi_primecase] > int eoip add name="eoip-tunnell" \
remote-address=172.31.31.1 tunnel-id=3112
```

Creazione del bridge:

```
[admin@c_farasabina.wifi_primecase] > int bridge add name="bridge1"
[admin@c_farasabina.wifi_primecase] > int bridge port add bridge=bridge1 \
interface= ether1
[admin@c_farasabina.wifi_primecase] > int bridge port add bridge=bridge1 \
interface= eoip-tunnell
```

Lato concentratore sono presenti le regole gemelle di quelle appena descritte, il motivo per cui sono stati creati anche dei tunnel ethernet over ip (eoip) è che Mikrotik non consente di aggiungere ad una interfaccia virtuale bridge un tunnel pptp.

## 4 Il log dei dati sensibili

Come da normativa vigente i dati relativi alle connessioni effettuate dagli utenti, così come anche la corrispondenza mac-address - indirizzo IP, vengono salvati per un periodo di sei mesi (salvo comunicazioni dell'autorità giudiziaria) su un sistema di logging basato su syslog-ng.

Configurazione del syslog:

```
source s_net { udp (); };
destination df_captive { file("/var/log/captive.log"); };
filter f_captive { host( "IP_CAPTIVE_PORTAL_ROUTER" ); };
```

Configurazione Del Router Mikrotik Che Invia i Log

```
[admin@CAPTIVE_PORTALS] > ip firewall mangle add chain=prerouting action=log \
connection-state=new src-address=10.0.0.0/8
```

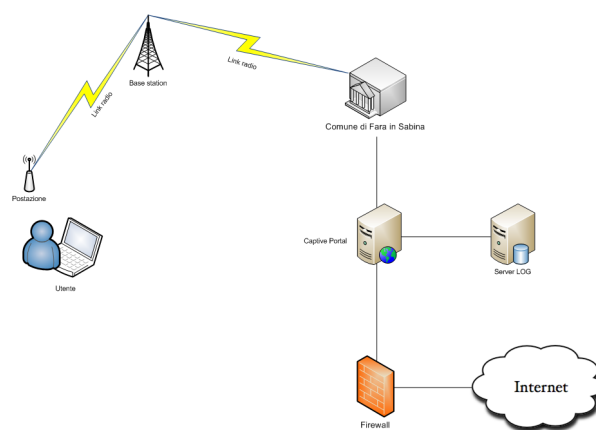


Fig. 2 Schema della connessione di rete.

## 5 Il monitoraggio del sistema

Tutti gli apparati di rete impiegati dal sistema sono stati posti sotto monitoraggio tramite la piattaforma Istituto di Zabbix in uso presso l'Area della Ricerca RM1.<sup>2</sup> In caso di irraggiungibilità di uno di essi, un avviso via mail raggiunge lo staff tecnico, al fine di intervenire tempestivamente nella diagnosi e nella risoluzione del guasto.

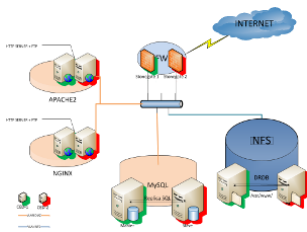
### Riferimenti

- 1 A. Pifferi, G. Nantista, L. Ianniello, C. Ricci, L. Rossi, M. Simonetti, Un Captive Portal per l'autenticazione su Reti Wifi dedicate agli Internet Access Point Liberi, SMART eLAB 2 (2013) 15–19. doi:10.30441/smart-elab.v2i0.55.
- 2 A. Pifferi, G. Nantista, L. Ianniello, A. Lora, M. Simonetti, Analisi e Implementazione di Sistemi per il Monitoraggio della Rete Wireless Relativa al Progetto ADD (Anti Digital Divide) e delle Infrastrutture di Campus AdR RM1., SMART eLAB 2 (2013) 1–9. doi:10.30441/smart-elab.v2i0.46.



## Progetto Virtual Hosting in High-Availability Area Della Ricerca RM 1.<sup>†</sup>

Luca Ianniello,<sup>a</sup> Giuseppe Nantista,<sup>a</sup> Andrea Lora<sup>a</sup> and Augusto Pifferi<sup>a</sup>



Gestione dei servizi web 2.0 in High Availability: il numero dei servizi di rete da gestire aumenta di pari passo con le necessità dell'utenza dell'Area di Ricerca di Montelibretti. Tra la varietà dei servizi richiesti risulta evidente la necessità di requisiti minimi comuni per l'esecuzione del software di turno, ad esempio un database Mysql o un CGI (Common Graphic Interface), in grado di eseguire il codice dell'evenienza PHP, PERL, PYTHON, JAVA etc.. La soluzione adottata mira a garantire la massima disponibilità per i dati e continuità di servizio h24.

**Keywords:** Virtual Hosting, Web Services, Linux, Gentoo, Backup, High Availability, HTTP Server, Apache, Nginx, CGI.

### 1 Introduzione

La varietà e il numero, per diffusione e accessibilità, dei dati offerta dal Web fa sì che sia quasi impossibile che esista un unico software in grado di fornire servizi in modo esclusivo e inimitabile. Il quasi è dovuto esclusivamente da eventuali limitazioni hardware, imposte dal Brand di produzione ed applicate a speciali dispositivi accompagnate da software proprietario (**Adobe Flash**, **SAN**, **AXIS** Camera Station etc..), ottenendo di fatto l'esclusività di un prodotto.

Nel mondo dell'Open Source, invece, dove le limitazioni quasi scompaiono, è possibile trovare diverse soluzioni per le stesse necessità, si è quindi in grado di scegliere quale software utilizzare a seconda del servizio richiesto.

La pagina di un sito Web, allo stato pratico, è un'interfaccia grafica descritta e modellata da una serie di comandi predefiniti, che cambiano sintassi a seconda del linguaggio di programmazione utilizzato, rinchiusi generalmente in uno o più files che saran-

no processati dall'apposito CGI nel momento in cui l'utente (client) ne fa espressa richiesta.

Il **CGI** è l'interfaccia che genera la pagina Web attraverso i comandi contenuti nei programmi che risiedono all'interno del sito consultato.

Per rendere fruibili questi contenuti a tutti i client del World Wide Web è necessario un ulteriore software l'**HTTP Server** che svolge il compito di tenere aperta una porta di comunicazione in attesa di richieste di consultazione.

I servizi del Web 2.0, grazie al loro incessante progresso, diventano sempre più user-friendly e graficamente piacevoli a discapito della potenza di calcolo e della complessità della struttura di supporto, la caratteristica intrinseca del Web è la fruibilità h24 dei contenuti, sostenere questo status (on-line) nel tempo non è cosa banale soprattutto per via del continuo aggiornamento, necessario o richiesto, dai software utilizzati.

Per poter fornire un servizio in alta disponibilità occorre creare una struttura ridondata sia a livello hardware che software con procedure di fall-back pronte a superare eventuali blocchi o malfunzionamenti occasionali a tutti i livelli.

Nella struttura che abbiamo realizzato il requisito

<sup>a</sup> CNR - Istituto di Cristallografia, Strada Provinciale 35/d, Montelibretti, Italia

Creative Commons Attribuzione - Non commerciale - Condividi allo stesso modo 4.0 Internazionale

<sup>†</sup> Rapporto tecnico IC-RM 2013/11 protocollato in data 20/12/2013 n. IC 1881

di alta disponibilità è stato applicato anche ai dati ospitati, quindi anche tutti i dati sono ridondati fisicamente con procedure di sincronizzazione e backup automatizzati.

## 2 Software per i servizi web

L' HTTP Server è la porzione di software che si occupa di rendere reperibili dati e le informazioni presenti nelle pagine Web.

Agli inizi il contenuto di una pagina Web era composto di solo codice in linguaggio **HTML** (*HyperText Markup Language*); attraverso questo potente linguaggio è stato possibile creare contenuti Web rappresentati in maniera statica e graficamente molto più gradevole (testo e immagini) rispetto al classico terminale (solo testo). L'evoluzione del WWW ha reso possibile incorporare porzioni di codice di linguaggi differenti dall'HTML, come Javascript, Python, Php etc., all'interno delle pagine Web rendendo possibile la produzione di contenuti graficamente dinamici e strutturalmente molto più complessi (**multi-tier architecture**).

Con l'**HTML5** e le sue nuove API è possibile generare contenuti graficamente dinamici alla stregua di **Adobe Flash**, il software che per anni è stato eccellenza per le animazioni Web. L'HTML è uno standard e può essere interpretato universalmente: tutti i dispositivi muniti di un browser in grado di richiedere e rappresentare contenuti HTML possono visualizzare lo stesso dato.

Questa considerazione, seppur ovvia, sta spingendo i maggiori produttori di dispositivi e di software verso l'uso del Web per l'interazione tra l'utente finale e i servizi offerti in quanto oggi, con l'avvento e la diffusione degli smartphone e i tablet, il Web è il primo mezzo utile attraverso il quale è possibile controllare o consultare lo stato di apparecchiature di uso comune da remoto (interruttori telecomandati, telecamere di sorveglianza, condizionatori, router etc.). Ne consegue che una gran parte dei dispositivi elettronici oggi presenti sul mercato globale contengono un, seppur minimo, HTTP Server consultabile da un browser.

Ovviamente ogni dispositivo ha il suo sistema operativo e per ogni sistema c'è un HTTP Server differente con una differente capacità di lavoro, dovendo costruire una struttura in grado di ospitare n- siti web abbiamo installato su tutti e 4 gli HTTP Ser-

ver Gentoo-Linux su architettura x64 come sistema operativo di base.

Come riportato dalle pagine wiki, il software che per anni è stato la colonna portante tra gli HTTP Server è **Apache** e la versione da noi utilizzata al momento della stesura è la 2.2.25 (Unix).

Apache2 presenta un'architettura modulare, quindi ad ogni richiesta del client vengono svolte funzioni specifiche da ogni modulo di cui è composto, come unità indipendenti. Ciascun modulo si occupa di una funzionalità, ed il controllo è gestito dal core.

Questo tipo di approccio orientato ai thread o ai processi nella gestione delle richieste risulta abbastanza oneroso in termini computazionali ed è per questo che è stato completamente stravolto con **Nginx**, successore di Apache2 (n.d.r.), un altro HTTP Server che è stato sviluppato con la prerogativa di dover servire centinaia di milioni di richieste giornaliere utilizzando un approccio asincrono basato su eventi nella gestione delle richieste in modo da ottenere prestazioni più prevedibili sotto stress.

**Nginx**, fondamentalmente, è single-threaded ma può utilizzare più processi e più core. C'è un evento principale ciclico che attende istruzioni dal sistema operativo, ad esempio la lettura di un file, a questo punto il file viene letto in un buffer ed elaborato. Il singolo thread può servire in maniera molto efficiente decine di migliaia di connessioni simultanee (mentre il modello thread-per-connection non riesce molto bene in questo a causa della enorme contenuto di dati contestuale che dovrebbe gestire con un grande consumo di memoria, poiché ogni thread necessita di un proprio stack).

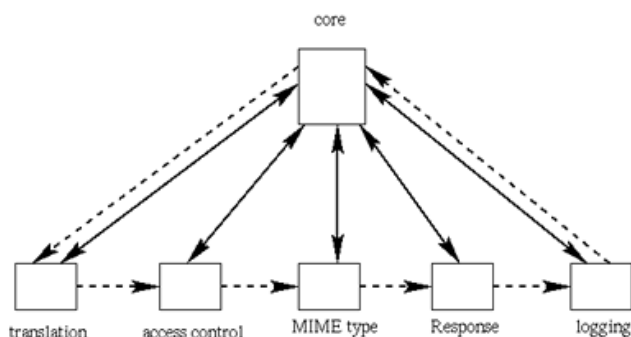
I siti più comuni hanno pagine generate da sistemi di gestione dei contenuti i **CMS** (*Content Management System*), questi sistemi oltre all'HTTP Server, per poter funzionare, hanno come requisito un database di dati e generalmente una CGI PHP da qui nasce l'acronimo **LAMP** che sta a indicare la struttura composta da Linux server con Apache come HTTP Server, Mysql per il database dei dati e l'interprete dei comandi Php.

## 3 Descrizione di High Availability

All'interno dell'Area di Ricerca Roma 1 esistono numerosi siti Web costruiti su piattaforma LAMP e proprio per favorirne e migliorarne l'utilizzo è stata creata una infrastruttura LAMP in High Availability

	HTML5 standard	HTML web browser features	Adobe Flash features
<b>Date started</b>	Work began in 2003 Working Draft as of 2011	N/A	Work began in 1996 Version 1 released in 1997
<b>Desktop operating systems</b>	N/A	Microsoft Windows, Apple Mac OS X, Linux Windows Phone 8+; Android 2.3+; Apple iOS 6+; Symbian Belle+; BlackBerry OS 7+	Microsoft Windows, Apple Mac OS X, Linux (PPAPI-only after 11.2), Solaris Up to Android 4.0 (unofficially for Android 4.1), Windows RT
<b>Mobile operating systems</b>	N/A	Xbox 360, [10][11] Sony PlayStation 3, Nintendo Wii	Sony PlayStation 3 (Flash 9 only), Nintendo Wii (Flash Lite only)
<b>Video game consoles</b>	PlayStation Vita	Limited access to web camera, microphone, accelerometer, GPS	Full access to web camera, microphone, accelerometer, GPS
<b>Devices support</b>	N/A		
<b>Market penetration</b>	N/A	~96% of internet connected PCs are CSS2/3 ACD compliant, ~99% support JavaScript 1.5 or above, ~84% of devices support basic SVG	~95% of internet connected PCs, (~83% have Flash Player 11, ~17% have 10 and below)
<b>Vector graphics formats</b>	N/A	Scalable Vector Graphics (SVG) supported on ~84% of browsers	SWF with embedded graphics
<b>Bitmap effects</b>	Yes	Varying support of Filters in CSS3 (eg. Glow, Blur, Drop Shadow, Sepia)	Yes, applied to text or graphics (eg. Glow, Drop Shadow, Bevel)[23]
<b>Vector text display</b>	Yes	Yes	Yes, with Saffron Type System
<b>Font support</b>	N/A	Installed fonts and custom fonts using CSS3 web fonts	Installed fonts and embedded fonts
<b>Text anti-aliasing</b>	N/A	Yes, implemented in most browsers, for system and custom fonts	Yes, in most cases
<b>Text tab stops</b>	No	Only supported inside "pre" tags	Yes, with Text Layout Framework
<b>Liquid text layout</b>	Yes	Yes, using the "div" tag and CSS styling	No, but text fields can be resized in ActionScript
<b>Tabular data</b>	Yes	Yes, using the "table" tag	No, but text fields can be arranged into a grid
<b>Linked text frames</b>	No	No	Yes, with Text Layout Framework
<b>Programming languages</b>	Depends	JavaScript	ActionScript, Pixel Bender
<b>C++ support</b>	Cross-compiling of C++ code to JavaScript using Emscripten	Native code execution with Google Native Client in Google Chrome only	Cross-compiling of C++ code to run in Flash Player using FlasCC
<b>Source code format</b>	Plaintext	Plaintext JavaScript with limited obfuscation	Compiled bytecode, can be obfuscated
<b>Data formats</b>	Depends	CSS3, HTML, XML, JSON	JSON, XML, Subset of CSS 1
<b>Data compression</b>	No	GZIP compression for HTML, JS and CSS files (on supported servers)	LZMA or DEFLATE for SWF files
<b>Image formats</b>	Depends	PNG, JPEG, Animated GIF	PNG, JPEG, JPEG-XR, Single-frame GIF
<b>Video formats</b>	Depends	Varying support of H.264, WebM and Ogg Theora (see HTML5 video)	H.264, Sorenson Spark, and On2 VP6
<b>Streaming video</b>	No	Pseudo-streaming only of WebM and Ogg Theora using video tag[33][33]	Flash Video, H.264 and partial support for MP4
<b>Audio formats</b>	Depends	Varying support of MP3, Ogg Vorbis, WAV PCM, AAC and WebM Vorbis (see HTML5 audio)	MP3, WAV and AAC audio files or embedded sound
<b>Fullscreen support</b>	No	Supported on some browsers	Yes, with warning displayed
<b>Encryption/DRM</b>	With obfuscation	No, all files being plaintext, except for obfuscation	Yes, being binary formatted files, unless decompiled
<b>File system access</b>	No	Varying support of single file upload, and drag and drop of files onto browser	Support for single file upload and generation, AIR only; full create/read/write access to file system
<b>Bitmap manipulation</b>	No	Varying support for HTML5 "canvas" element	Yes, using the BitmapData class
<b>Binary manipulation</b>	No	No, unless JavaScript Arrays are used for the same	Yes, using the ByteArray class
<b>Large binary data</b>	No	Using Web Sockets to stream binary or XML data	Yes, embedded or streaming binary data
<b>Offline storage</b>	Depends	Using Web storage or cookies to store binary or XML data	Using Local Shared Objects to store AMF-formatted data
<b>Metadata</b>	Meta tags	Can be included in meta tags	Extensible Metadata Platform

Comparazione HTML5 vs Flash. estratta da [wikipedia](#)



**Fig. 1** Schema funzionamento Apache2 thread-per-connection

a disposizione di tutti gli istituti afferenti all'Area. L' High Availability è la capacità, di una risorsa o di un servizio fornito da un sistema, di essere sempre disponibile.

La diponibilità di un servizio è generalmente espressa in percentuale in base al tempo di diponibilità effettiva (uptime) nell'arco di un anno.

La tabella riportata di seguito ([http://en.wikipedia.org/wiki/High\\_availability](http://en.wikipedia.org/wiki/High_availability)) riporta il calcolo delle percentuali di disponibilità in base ai tempi di uptime e i diversi periodi di downtime:

Per ottenere l'High Availability si usa ridondare i sistemi (in ingegneria raddoppio delle componenti):

**Tabella 1** Estratta da [wikipedia](#)

Availability %	Downtime per year	Downtime per month*	Downtime per week
90%("one nine")	36.5 days	72 hours	16.8 hours
95%	18.25 days	36 hours	8.4 hours
97%	10.96 days	21.6 hours	5.04 hours
98%	7.30 days	14.4 hours	3.36 hours
99%("two nines")	3.65 days	7.20 hours	1.68 hours
99.5%	1.83 days	3.60 hours	50.4 minutes
99.8%	17.52 hours	86.23 minutes	20.16 minutes
99.9%("three nines")	8.76 hours	43.8 minutes	10.1 minutes
99.95%	4.38 hours	21.56 minutes	5.04 minutes
99.99%("four nines")	52.56 minutes	4.32 minutes	1.01 minutes
99.999%("five nines")	5.26 minutes	25.9 seconds	6.05 seconds
99.9999%("six nines")	31.5 seconds	2.59 seconds	0.605 seconds
99.99999%("seven nines")	3.15seconds	0.259 seconds	0.0605 seconds

la ridondanza può essere attiva o passiva.

La ridondanza attiva viene utilizzata soprattutto per sopperire ad elevati livelli di carico di lavoro, bilanciando i task tra i componenti attivi mentre nella ridondanza passiva una componente rimane allineata e pronta ad intervenire in caso di failure della componente attiva garantendo continuità al servizio.

La scelta di utilizzare l'una o l'altra soluzione va presa in base al carico di lavoro da supportare ed in base all'hardware a disposizione oltre che dal tipo di servizio richiesto.

Nella struttura realizzata sono stati applicati entrambi i criteri per la ridondanza (Attiva/Passiva) in base alle necessità incontrate.

#### 4 Infrastruttura delle piattaforme

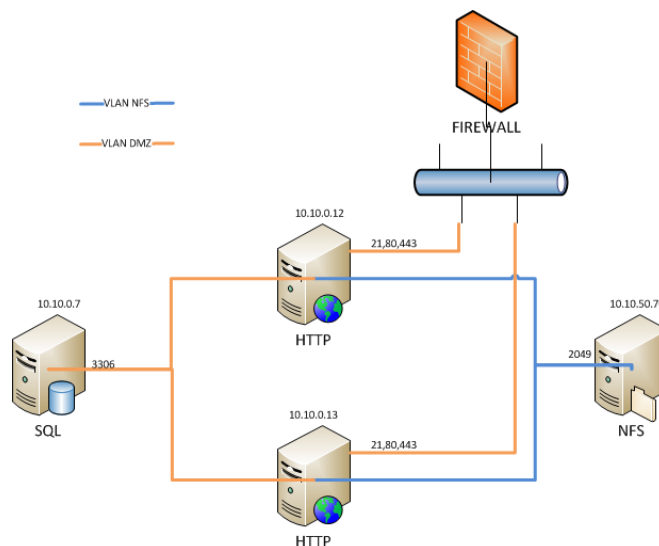
Per ridondare un'infrastruttura di tipo LAMP è possibile utilizzare diversi approcci, quello riportato di seguito è focalizzato verso l'isolamento dei singoli servizi necessari alla struttura.

Sono stati preparati, dunque, un host per ogni singolo servizio

- 1 host HTTP Server + CGI PHP
- 1 host per il database MySQL

Essendo i due host separati e distinti è stato possibile applicare ridondanze specifiche per ogni servizio.

Per permettere a due o più HTTP Server differenti di indirizzare la stessa porzione di files, abbiamo posizionato i files dei siti ospitati, su condivisioni NFS. Questa soluzione permette di scalare orizzontalmente la struttura nel caso vi fosse necessità di un incremento delle performance semplicemen-



**Fig. 2** Schema della struttura LAMP

te aggiungendo un ulteriore Host per il medesimo servizio.

I problemi incontrati con questa soluzione non sono risolvibili in maniera immediata: la prima difficoltà, che porta con se una struttura così duplicata, è stata quella di impostare gli stessi UID e GID ai demoni che interagiscono con i files dei siti (Apache, Ftp, SAN) in modo da non creare eventuali errori di scrittura o negazione di permessi. Il secondo problema da risolvere con questa struttura è legata al demone FTP. Ritenendo che utilizzare un Host per il solo servizio FTP fosse eccessivo, abbiamo aggiunto il demone **Proftpd** sulle macchine con a bordo l'HTTP Server e, per non dover duplicare le credenziali d'accesso su entrambi i Server, abbiamo optato per l'utilizzo di un database comune come registro per le utenze FTP.

Un altro problema, conseguenziale al raddoppio degli HTTP Server, è la gestione delle connessioni in ingresso: anche qui si possono applicare diverse soluzioni che possono andare dal semplice bilanciamento di carico per mezzo del ROUND ROBIN tra multipli A Record nel DNS ad un eventuale load-balancer in grado di ripartire il traffico in ingresso tra gli Host a disposizione o in caso di configurazione Active/Passive, di indirizzare il traffico verso l'Host primario con l'incarico di passarlo tutto all'Host secondario nel momento in cui il primario risulta off-line.

In Figura 2 è riportato lo schema della struttura LAMP con i servizi isolati e raddoppiati per la parte HTTP Server.

Le connessioni in arrivo verso i server Web sono state bilanciate a monte dal Firewall di frontiera in modalità Active/Active.

## 5 Load Balancing e Fail-over su Stonegate

Il Firewall di frontiera verso la rete Internet di Montelibretti è di tipo software (Stonegate), questa applicazione mette a disposizione degli amministratori un Monitoring Agent proprietario, installabile su qualsiasi server, in grado di intervenire sulle decisioni del Firewall per l'instradamento dei pacchetti attraverso le informazioni raccolte dall'Agent in base a parametri prestabiliti. Anche il Firewall è ridondato in modalità Active/Active: 2 Host in bilanciamento di carico per la parte di Front-End più un ulteriore Host per la parte di Management come richiesto e previsto dalle best practice Stonegate.

```
# StoneGate Monitoring Agent global configuration file
# Package generated default configuration
begin host genweb4
config boot-delay 120
end

alert-interval 240
load-index 1500 load-average-1
config load-index-action exclude 10

test http_answering
interval 30
action exclude
command httpsaver 4 3500 /www.mlib.cnr.it/ /ARRM1v3/menu/menu.doc 80

test "port listening test"
interval 2:00
action exclude
host genweb4
recovery always
command portlistening 80 10.10.0.13

filesystem /dev/sda3
1024000
```

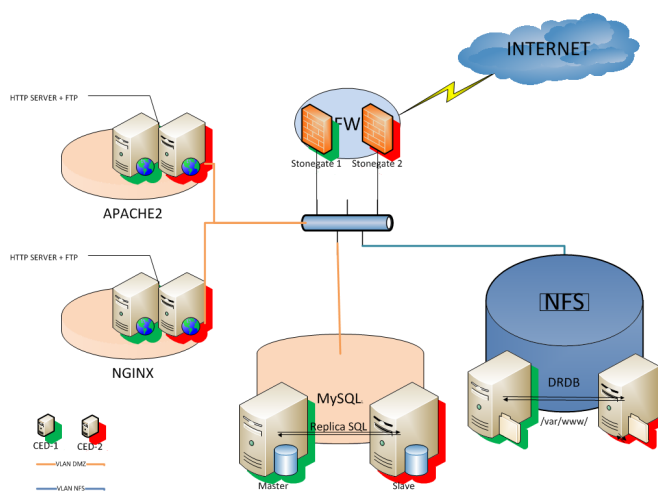
**Fig. 3** sgagent.conf file di configurazione dell'Agent Stonegate

Attraverso l'Agent è possibile gestire anche il meccanismo di Fail-over tra due o più Host: nella sezione "test http\_answering" della configurazione dell'Agent si imposta la verifica della presenza di un determinato file ogni 30 sec. Nel caso in cui questo controllo ritorni esito negativo l'Host viene automaticamente escluso dalla ricezione di nuove richieste HTTP che saranno quindi girate all'Host successivo. Oltre al meccanismo di Fail-over, gestito nella sezione "test" della configurazione dell'Agent, è stato abilitato un tetto massimo di carico per il bilanciamento delle connessioni tra i nodi Web.

Avendo monitorato a lungo i livelli di stress a cui sono stati sottoposti i Web Server preesistenti, si è desunto che, sia per quantità di carico di lavoro che per dimensione dei singoli database, con un unico Host è possibile soddisfare ogni esigenza di servizio per l'intera Area della Ricerca di Montelibretti, que-

sto anche in virtù del fatto che, l'Host in questione, deve farsi carico di un solo servizio rispetto alla soluzione "all-in-one" che prevede l'installazione dell'intera struttura LAMP su un unico Host. Questa conclusione ci ha portato a scegliere di utilizzare, per la gestione dei database, una ridondanza di tipo Active/Passive ampiamente supportata dal demone **MySQL**.

La porzione dei dati relativi alle pagine Web è montata all'interno dei web server come share NFS. Questo ci ha consentito di implementare una pila separata per l'alta affidabilità dello storage. Dal punto di vista dei server web l'unica direttiva necessaria è il mountpoint che indirizza all'IP del cluster NFS. Tale cluster è di tipo active/passive, offre la share NFS tramite un IP flottante che viene preso in carico dal server NFS attivo. Ognuno dei due server NFS monta una LUN ISCSI offerta da due diverse SAN, su queste LUN vengono scritti i dati riguardanti la share NFS. Con l'uso del software DRBD, che si frappone tra il layer disco e quello di filesystem, le due LUN vengono mantenute identiche in maniera sincrona a livello di blocco, garantendo la coerenza dei dati su due diversi backend disco. In caso di guasto del nodo attivo i meccanismi di clustering provvedono ad eseguire una procedura di STONITH sul nodo fallito per preservare l'integrità dei dati e a migrare il server NFS sul nodo in stand-by. Un terzo nodo partecipa al cluster per portare il quorum a 3, rendendo impossibili situazioni di split-brain.



**Fig. 4**

Avendo raddoppiato sia a livello fisico sia a livello applicativo l'intera struttura, ed avendo adeguati spazi a disposizione, è stato adottato il criterio previ-

sto dal **Disaster Recovery**, ossia abbiamo posizionato gli Host gemelli in sale CED differenti ottenendo una struttura rappresentabile come riportato nella figura di seguito.

Va però detto che per essere pienamente compliant con le direttive imposte a una struttura con Disaster Recovery le sale CED dovrebbero essere distanti almeno 20 Km per poter prevenire e superare i casi di calamità naturale di vasta entità o di bombardamento.

Confidando in un futuro privo di calamità possiamo dire che la struttura qui rappresentata è una struttura in alta affidabilità pronta per un Disaster Recovery.

## 6 Virtual Hosting

Il Virtual Hosting è il metodo utilizzato per ospitare più siti all'interno di un singolo Host (o gruppi di Host) in modo da sfruttare al massimo le risorse messe a disposizione dal Server ospitante (memoria RAM, CPU, IP).

Il Virtual Hosting può essere di tre tipologie: Name-based, IP-based o Port-based. Il Name-based Virtual Hosting è la tecnica che permette di associare, grazie al lavoro di traduzione svolto dall'HTTP Server, contenuti Web ad un nome di dominio: nella richiesta di contenuti effettuata dai client il nome di dominio viene anteposto come prima informazione da elaborare subito dopo la necessaria procedura di **handshake**, l'HTTP Server che riceve questa informazione si occupa di tradurla fornendo i corretti contenuti associati a quel determinato nome di dominio. Questa procedura, permette di risparmiare indirizzi IP in quanto è possibile associare N nomi di dominio ad un singolo IP. Va però tenuto presente che anche questo tipo di soluzione ha i suoi limiti: nel momento in cui si vogliono implementare certificati di sicurezza per il traffico criptato su più di un nome a dominio si crea un'ambiguità nel servizio per cui l'HTTP Server non sa quale certificato utilizzare per instaurare la connessione al momento dell'**handshake**. Esistono diversi work-around noti che permettono comunque di affrontare queste evenienze ma possono risultare inefficaci nel momento in cui il client non è in grado di recepire determinati accorgimenti utilizzati. Oltre al limite appena descritto c'è da aggiungere che, stante la peculiarità dei nomi di dominio, il sito non sarebbe più rag-

giungibile nel caso in cui la traduzione iniziale effettuata dal DNS non dovesse avvenire correttamente. Nell'implementazione IP-based ogni sito ha un suo indirizzo IP quindi sull'Host che farà da Server andranno configurati i relativi indirizzi. Questo tipo di implementazione ha il vantaggio di escludere il client dal processo di traduzione, su accennato, e quindi esente da problemi di compatibilità. L'altra faccia della medaglia dell'approccio IP-based è l'elevata complessità di amministrazione che si viene a creare oltre al consumo di un indirizzo IPv4 per ogni sito. Per questo motivo questa modalità è utilizzata solo in casi particolari. L'ultima tipologia la Port-based sfrutta la possibilità, fornita dai software di Web services, di fornire contenuti Web diversi a seconda della porta di comunicazione utilizzata. La porta di comunicazione comunemente usata per i contenuti Web è la 80, nel caso si utilizzi la metodologia Port-based si arriverà ad avere (utilizzando diverse porte es. 8080,81 etc.) siti raggiungibili solo da chi conosce la giusta porta. Questo tipo di implementazione viene tipicamente utilizzata nelle intranet per i servizi di amministrazione dove i contenuti Web devono essere consultati da un numero ristretto di utenti di provenienza nota. Pertanto la modalità da noi utilizzata, la Name-based, risulta essere la più conveniente in termini di risorse sia fisiche che amministrative. Sia APACHE2 che NGINX supportano il Virtual Hosting e considerando le peculiarità dei due software si è optato per l'utilizzo di APACHE2 per i siti degli Istituti dell'Area che hanno generalmente un carico di lavoro elevato ma distribuito nel tempo, mentre i Server con a bordo NGINX vengono utilizzati per le piattaforme di e-learning che generano un traffico massivo per brevi periodi.

In tabella 2 l'elenco dei siti presenti sui server APACHE2 dell'Area di Montelibretti.

All'interno di questo elenco è possibile trovare siti web statici e siti web dinamici tutti comunque prevedono PHP come interprete dei comandi.

La versione dell'interprete PHP installato sulla struttura APACHE2 è PHP 5.3.27. anche il PHP è in continuo aggiornamento tanto che al momento della stesura di questo documento la versione disponibile è la 5.5.7. Per ogni nuova versione supportata di PHP ce ne è una che diventa DEPRECATED (non supportata) rendendo l'aggiornamento inevitabile; è solo questione di tempo.



**Tabella 2** Elenco siti presenti sui server Apache 2.

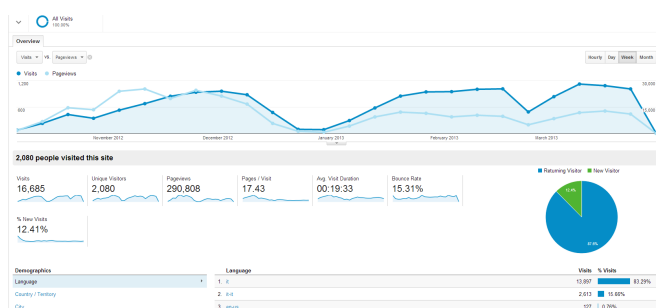
URL	Descrizione
http://biblio.mlib.cnr.it/	Biblioteca dell'Area della Ricerca di Roma 1
http://calliope.mlib.cnr.it/	Editoria Digitale a cura dell'I.C.
http://durc.mlib.cnr.it/	Portale per censimento DURC – Intranet Mlib
http://dvr.mlib.cnr.it/	Portale per il censimento DVR – Intranet Mlib
http://dyna.imip.cnr.it/	I.M.I.P. - Progetto Dynamics at nano scale
http://genesi.mlib.cnr.it/	Progetto GENESI – Rete integrata per il trasferimento di tecnologie e metodologie per la generazione di sviluppo industriale in sistemi di produzione.
http://gire.mlib.cnr.it/	I.B.A.F. - Gruppo Italiano Resistenza Erbicidi
http://greenability2011.mlib.cnr.it	I.T.A.B.C - Forum Internazionale sulla sostenibilità e la conservazione delle città storiche
http://grs-lab.itabc.cnr.it/	I.T.A.B.C - Ground Remote Sensing LAB
http://helpdesk.mlib.cnr.it/	S.R.A. – Front End telematico per la gestione dell'HelpDesk per l'Area di Ricerca RM1
http://icm2003.mlib.cnr.it/	I.M.C - Istituzionale
http://icmontelibretti.gov.it/	Istituto Comprensivo di Montelibretti – Progetto Scuole
http://impreseaimo.mlib.cnr.it/	Portale Imprese CAIMO
http://link.ism.cnr.it/	I.S.M.N - Laboratories for Integrated Nanoscale Knowledge
http://magnet.mlib.cnr.it/	CNR/INFN, CNISM, INSTM – I Convegno nazionale di Magnetismo
http://nano2010.mlib.cnr.it/	I.S.M – Conferenza Nano 2010
http://nmr.imc.cnr.it/	I.M.C. - Laboratorio Risonanza Magnetica
http://portaleucbs.mlib.cnr.it/	Unione di Comuni della bassa Sabina a cura dell' I.C.
http://postesuap.mlib.cnr.it/	Poste Italiane - Sportello SUAP
http://progetto.mlib.cnr.it/	Progetto Anti Digital Divide a cura dell'I.C.
http://www.16issssrome2010.it	I.M.C. - 16th International Symposium on Separation Sciences
http://www.appia.itabc.cnr.it/	I.T.A.B.C - Archaeological Park of ancient via Appia
http://www.ibaf.cnr.it/	I.B.A.F - Istituzionale
http://www.ibba.mlib.cnr.it/	I.B.B.A – Istituzionale sez. Montelibretti
http://www.ic.cnr.it/	I.C - Istituzionale
http://www.icfpm.mlib.cnr.it/	I.C.F.M.P. – Istituzionale sez. Montelibretti
http://www.imc.cnr.it/	I.M.C. - Istituzionale
http://www.imip.cnr.it/	I.M.I.P. - Istituzionale
http://www.isc.mlib.cnr.it/	I.S.C. – Istituzionale sez. Montelibretti
http://www.iscima.cnr.it/	I.S.C.I.M.A. - Istituzionale
http://www.isma.cnr.it/	I.S.M.A. - Istituzionale
http://www.itabc.cnr.it/	I.T.A.B.C - Istituzionale
http://www.iwp_rixs2014/.mlib.cnr.it/	International Workshop on Photoionization and Resonant Inelastic X-Ray Scattering
http://www.manfor.eu/	I.B.A.F. - Management of Forests, Carbon and Biodiversity
http://www.mlib.cnr.it/	Area della Ricerca di Montelibretti RM1
http://www.phototech.eu/	Photosynthetic proteins for technological applications
http://www.rstfen.iscima.cnr.it/	I.S.C.I.M.A. - Rivista di Studi Fenici
http://www.sra.mlib.cnr.it/	Servizio Reti Area – RM1
http://xenios.mlib.cnr.it/	Gestione telematica dei convegni scientifici a cura dell'I.C.

In tabella 3 l'elenco dei siti ospitati dalla struttura NGINX.

Come riportato da numerosi test e benchmark disponibili online NGINX risulta essere l'HTTP Server migliore, per velocità nell'offrire contenuti statici e

**Tabella 3** Elenco siti presenti sui server NGINX.

URL	Descrizione
http://minerva.mlib.cnr.it/	e-learning Formazione a Distanza
http://verdeorienta.mlib.cnr.it/	e-learning in collaborazione con ITCCG "E. Fermi"
http://fermi.mlib.cnr.it/	e-learning in collaborazione con: Istituto Tecnico Commerciale Geometri "Enrico Fermi" Istituto Tecnico Industriale "Alessandro Volta" Istituto d' Istruzione Superiore "Via Falisca snc" Istituto d'Istruzione Superiore "Via Pedemontana" Istituto d' Istruzione Superiore "G.Quarenghi"
http://romaforma.mlib.cnr.it/	e-learning in collaborazione con Roma Capitale
http://puentes.mlib.cnr.it/	e-learning in collaborazione con Universidad Mundial e Università degli Studi di Teramo
http://openopportunity.mlib.cnr.it/	Gestione Meeting Open Opportunity
http://catullo.mlib.cnr.it/	e-learning in collaborazione con il Liceo Classico e Linguistico "Catullo"
http://peano.mlib.cnr.it/	e-learning in collaborazione con Liceo Scientifico Statale "G. Peano" di Monterotondo



**Fig. 5** Confronto andamento tra le visite registrate e il numero di pagine richieste.

per maggior quantità di connessioni servite rispetto ad APACHE2. Il divario diminuisce se si confrontano i 2 software su contenuti dinamici generati da codice PHP in quanto il linguaggio utilizzato nei siti dinamici aggiunge la medesima componente a carico della CPU che dovrà farsi carico di interpretare il codice e generare il contenuto. A questo va aggiunto che la differenza di prestazioni, a parità di hardware tra i due HTTP Server, la si può apprezzare solo con grandi carichi di lavoro (dell'ordine di milioni di richieste al giorno, qualche centinaio al secondo). Pur essendo lontani da così alti carichi di lavoro l'utilizzo di NGINX sul fronte dell'e-learning si è rivelata, fino ad oggi, una scelta corretta. Di seguito è riportato un estratto (vedi figura 5) dei grafici del carico di lavoro sopportato dalla sola piattaforma Roma-Forma allestita per i corsi di formazione del personale di Roma Capitale. I grafici fanno riferimento a un periodo di 6 mesi: dal 7 Ottobre 2012 al 30 Marzo 2013.

## 7 Conclusioni

La struttura così come descritta si colloca, ad oggi, al livello “four nine” della tabella dell’ High Availability sopra riportata, il costante lavoro dell’IC mira ad attestare la propria struttura almeno al “five nine” ovvero a mantenere lo status “on-line” per un intero anno con al massimo 5 min di downtime.

Le soluzioni tecniche e il software utilizzato per la realizzazione della struttura in High Availability si sono dimostrate efficienti e rispondenti agli obiettivi del progetto. La qualità tecnica dell’infrastruttura e i tempi di risposta ottenuti nell’erogazione del servizio hanno consentito di soddisfare tutte le richieste giunte al servizio reti dell’Area di Ricerca di Roma 1. Si prevede che i futuri sviluppi dell’infrastruttura qui descritta sarà orientata ad all’implementazione della virtualizzazione degli Host. L’implementazione dell’utilizzo di macchine virtuali, attraverso i potenti tool disponibili nell’ambiente VSphere (come la clonazione di un’intera Virtual Machine a caldo o la possibilità di effettuare una snapshot di un’intera Virtual Machine che consente il ripristino della stessa, riportandola al momento dello snapshot, con tempi brevissimi), rafforza ulteriormente l’intera struttura.

### Riferimenti

- 1 [http://en.wikipedia.org/wiki/High\\_availability/](http://en.wikipedia.org/wiki/High_availability/).
- 2 <http://www.php.net/>.
- 3 <http://www.mysql.com/>.
- 4 <https://httpd.apache.org/>.
- 5 [http://wiki.dreamhost.com/Web\\_Server\\_Performance\\_Comparison/](http://wiki.dreamhost.com/Web_Server_Performance_Comparison/).
- 6 [http://www.wikivs.com/wiki/apache\\_vs\\_nginx/](http://www.wikivs.com/wiki/apache_vs_nginx/).

## 8 Glossario

**Browser:** un programma che consente di usufruire dei servizi di connettività in Internet, o di una rete di computer, e di navigare sul World Wide Web

**Client:** indica una componente che accede ai servizi o alle risorse di un’altra componente detta server  
**Disaster Recovery:** l’insieme delle misure tecnologiche e logistico/organizzative atte a ripristinare sistemi, dati e infrastrutture necessarie all’erogazione di servizi a fronte di gravi emergenze che ne intacchino la regolare attività.

**DRDB:** Distributed Replicated Block Device

**HTTP Server:** software di gestione del server Web.

**Multi-tier architecture:** è un architettura client-server nella quale le funzioni di presentazione, elaborazione delle applicazioni e gestione dati sono logicamente separate.

**MySQL:** è un sistema di gestione relazionale dei database, composto da un pacchetto client e da un pacchetto server.

**PHP:** linguaggio di programmazione interpretato, concepito per la programmazione Web e ora utilizzato principalmente per sviluppare applicazioni web dal lato del server.

**SAN:** Storage Area Network è una parte di una rete ad alta velocità di trasmissione costituita esclusivamente da dispositivi di memorizzazione di massa, Il suo scopo è quello di rendere tale storage disponibili per qualsiasi computer connesso ad essa.

**VSphere:** una suite potenziata di strumenti per la gestione del clud computing utilizzato da VMware ESX/ESXi

**VMWare:** software che provvede alla virtualizzazione e al cloud computing.