Tapporti tecnici 77

La nuova infrastruttura di acquisizione e distribuzione dati della Rete Integrata Nazionale GPS (RING)





Editorial Board

Andrea Tertulliani - Editor in Chief (INGV - RM1)

Luigi Cucci (INGV - RM1)

Nicola Pagliuca (INGV - RM1)

Umberto Sciacca (INGV - RM1)

Alessandro Settimi (INGV - RM2)

Aldo Winkler (INGV - RM2)

Salvatore Stramondo (INGV - CNT)

Gaetano Zonno (INGV - MI)

Viviana Castelli (INGV - BO)

Marcello Vichi (INGV - BO)

Sara Barsotti (INGV - PI)

Mario Castellano (INGV - NA)

Mauro Di Vito (INGV - NA)

Raffaele Azzaro (INGV - CT)

Rosa Anna Corsaro (INGV - CT)

Mario Mattia (INGV - CT)

Marcello Liotta (Seconda Università di Napoli, INGV - PA)

Segreteria di Redazione

Francesca Di Stefano

Tel. +39 06 51860068

Fax +39 06 36915617

Rossella Celi

Tel. +39 095 7165851

redazionecen@ingv.it

Lapporti tecnici 77

LA NUOVA INFRASTRUTTURA DI ACQUISIZIONE E DISTRIBUZIONE DATI DELLA RETE INTEGRATA NAZIONALE GPS (RING)

Luigi Falco, Gianpaolo Cecere, Ciriaco D'Ambrosio

INGV (Istituto Nazionale di Geofisica e Vulcanologia, Centro Nazionale Terremoti)

Indice

Introduzione	7
PARTE PRIMA: L'architettura del sistema di acquisizione e distribuzione dati	8
I sistemi di trasmissione dati utilizzati	8
2. Precedente infrastruttura di acquisizione e distribuzione dati	8
2.1 Il vecchio software CLINIC	9
2.2 GPSGIVING E GPSFREE	9
2.3 Il vecchio BANCADATI	10
3. La nuova infrastruttura: integrazione di CLINIC e BANCADATI	10
3.1 Requisiti utente	10
3.2 Il progetto	11
3.3 Il nuovo CLINIC e l'integrazione con BANCADATI	13
3.4 Il Permission Manager	15
3.5 L'astrazione del file system	15
PARTE SECONDA: L'applicativo web	16
4. L'applicativo web integrato CLINIC-BANCADATI	16
4.1 Home page utente	16
4.2 Home-> Le stazioni di una rete GPS	17
4.2.1 GridView	17
4.2.2 MapView	18
4.2.3 ListView	19
4.2.4 TaskView	19
4.3 Informazioni di dettaglio stazione GPS	20
4.3.1 Statistics Stazione	20
4.3.2 LOG FILE di Stazione	22
4.3.3 Informazioni aggiuntive di stazione	23
4.3.4 Monografia di stazione	23
4.3.5 Manutenzione stazione	24
4.3.6 Mappa dislocazione stazione	24
4.3.7 Grafici di stazione	24
4.4 Pannello Admin	25
4.4.1 Admin -> TeqcLog	25
4.4.2 Admin -> TeqcReprocess	26
4.5 Statistics	27
5. GSAC: integrazione archivio RING con le reti convenzionate UNAVCO	27
6. Conclusioni	29
Bibliografia	30

Introduzione

Il progetto CESIS, finanziato dal Ministero dell'Università e della Ricerca (Legge 488/92), oltre a prevedere la realizzazione di una nuova sede dell'Istituto Nazionale di Geofisica e Vulcanologia (INGV) a Grottaminarda, aveva come obiettivo il potenziamento della Rete Sismica Nazionale nel centro-sud Italia attraverso l'installazione di 60 nuove stazioni di monitoraggio multi-parametriche costituite da sismometro, accelerometro e GPS di precisione.

Grazie ai finanziamenti di questo progetto, verso la fine del 2004, l'INGV avviò la realizzazione di una rete di ricevitori GPS permanenti a scala nazionale, sia integrando le esperienze preesistenti nelle diverse sedi INGV che sviluppando una strategia triennale di nuove installazioni. Nacque così la Rete Integrata Nazionale GPS (RING) caratterizzata dalla co-localizzazione di strumentazione sismologica e GPS; in particolare, sismometri a banda larga o larghissima (40 sec ÷ 240 sec) e accelerometri strong motion insieme a ricevitori GPS. L'obiettivo principale consisteva nel contribuire ad aumentare le conoscenze relative alla cinematica e alla tettonica attiva del territorio italiano, puntando soprattutto sulla condivisione delle esperienze lavorative dei vari gruppi di lavoro in particolare della sezione di Napoli, Catania, Bologna, Roma



Figura 2. Dislocazione geografica stazioni GPS.

Figura 1. Crescita numero stazioni negli anni.

e Grottaminarda dell'INGV [Avallone et al., 2010]. La RING è una multisensoriale con sistema trasmissione in tempo reale, costituita, oggi, da più di 170 stazioni di proprietà INGV dislocate in tutta la penisola e con maggiore densificazione nelle sismogenetiche più importanti. L'infrastruttura di acquisizione distribuzione dati che ci si appresta a descrivere archivia non solo i dati delle stazioni GPS della RING, ma anche quelli forniti da reti GPS permanenti gestite da altri Enti, Università o soggetti privati in convenzione con INGV. Le figure 1 e 2 mostrano l'incremento del

numero di stazioni acquisite dal 2003 ad oggi e la loro dislocazione sul territorio nazionale.

Il rapporto tecnico si sviluppa in due parti: la prima parte descrive il vecchio sistema di acquisizione e distribuzione dati analizzandone le criticità intervenute in seguito a nuove esigenze di servizio; il processo di migrazione dalla vecchia alla nuova architettura ed infine il processo di integrazione di due applicazioni

precedentemente separate, CLINIC e BANCADATI; una seconda parte che descrive l'applicativo web BANCADATI rinnovato nella sua veste grafica e arricchito di nuove funzionalità che consente, oggi, la gestione centralizzata dell'applicazione CLINIC anch'essa rivista e migliorata.

PARTE PRIMA: L'architettura del sistema di acquisizione e distribuzione dati

1. I sistemi di trasmissione dati utilizzati

I sistemi di trasmissione del dato, maggiormente utilizzati, sono prevalentemente di tipo satellitare, cellulare e Wi-Fi.

Il sistema satellitare sfrutta l'infrastruttura di acquisizione dati utilizzata dal sistema Libra VSAT di Nanometrics (sistema nato per la trasmissione del dato sismico) affiancando, al flusso prodotto dai sensori sismici, il flusso seriale prodotto dai ricevitori GPS; il dato viene campionato a 30s ed estratto dai NAQS Server (il server di acquisizione dei dati sismici prodotto da Nanometrics) ogni 24 ore. Il vantaggio di tale sistema è essenzialmente il basso costo in quanto sfrutta la connessione della stazione sismica ma, di contro, ha grossi limiti sia per quanto riguarda la disponibilità di banda (molto limitata) che per le operazioni di controllo remoto del ricevitore. Per i dettagli si veda [Falco, (2006)].

Il sistema di telecomunicazione "cellulare" è basato sulle connessioni mobili offerte dagli operatori telefonici; la velocità di collegamento di queste connessioni è molto variabile e dipende dalla copertura di rete offerta dal provider nella zona in cui è ubicata la stazione di monitoraggio; si va da connessioni abbastanza limitate di tipo GPRS/EDGE, che comunque consentono la trasmissione in tempo reale di un flusso dati a 30s, fino alle più veloci connettività UMTS/HSDPA di terza generazione. Purtroppo, la copertura di rete non sempre presente o sufficientemente stabile e le politiche di gestione degli indirizzamenti adottate dai vari operatori telefonici non sempre consentono di raggiungere facilmente il terminale remoto. Per i dettagli circa il funzionamento e la strumentazione utilizzata si veda il seguente Rapporto Tecnico INGV: [Falco (2008)]. È inoltre in corso di realizzazione un progetto con la società di telecomunicazioni Vodafone che consentirà l'attivazione di un APN (Access Point Name) completamente dedicato all'INGV e uno scambio dati diretto tra le stazioni remote e la sede di Grottaminarda tramite circuito MPLS (Multi Protocol Label Switching); questo tipo di collegamento assicurerà percorsi di instradamento diretto tra la sede e le stazioni remote su rete Vodafone.

Le connessioni di tipo Wi-Fi sfruttano l'infrastruttura progettata e sviluppata presso la Sede INGV di Grottaminarda e descritta nel seguente Rapporto Tecnico INGV: [Cardinale et al., 2010]. Si tratta, essenzialmente, di una rete ad alta velocità basata su connessioni radio a 5Ghz (Frequenze libere). Il grosso vantaggio di tale soluzione è soprattutto la velocità e la stabilità delle connessioni (6/11 Mb/s) oltre che il costo; non ci sono canoni annui, l'unico costo è stato quello di start up.

Infine esistono varie soluzioni che sfruttano connessioni internet offerte da enti ospitanti la strumentazione come Scuole, Comuni, e altro.

2. Precedente infrastruttura di acquisizione e distribuzione dati

La precedente piattaforma di gestione dati e metadati della RING si avvaleva di due distinte applicazioni: CLINIC e BANCADATI. CLINIC era una java application che eseguiva il controllo di qualità e l'archiviazione su file system dei dati GPS, mentre BANCADATI era una tipica web-application, finalizzata alla gestione e condivisione delle informazioni relative ai dati (metadati). L'intero processo di acquisizione, pre-elaborazione e archiviazione dei dati della RING era, quindi, suddiviso su diverse infrastrutture separate sia sul piano hardware che software. Tale separazione ha comportato in questi anni un aggravio di lavoro soprattutto in termini di configurazione dell'intero sistema di acquisizione dei file, dovendo, ad esempio, rendersi necessaria la replica delle credenziali di accesso ai sistemi per ogni utente, e la configurazione delle relative cartelle FTP, in funzione delle diverse reti esistenti. La gestione contemporanea di più reti GPS ha comportato una necessaria evoluzione del database, progettato inizialmente per una sola rete, riorganizzato in modo da poter ospitare e gestire più reti e relative utenze di accesso. Il tutto complicando ancor più la gestione e manutenzione dell'intera infrastruttura. La Fig. 3 mostra l'architettura della vecchia infrastruttura.

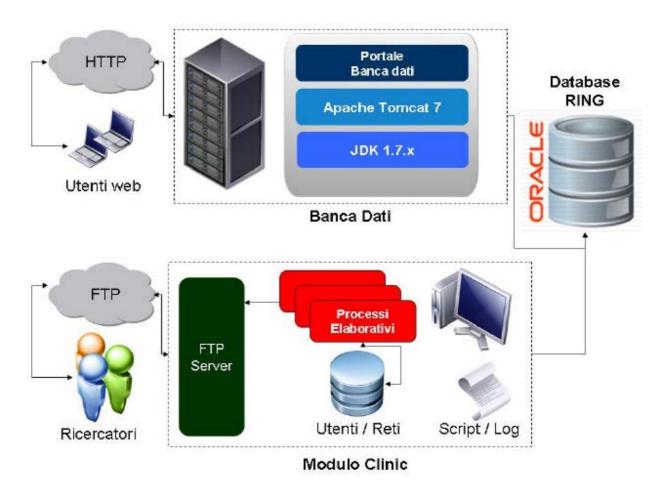


Figura 3. CLINIC e BANCADATI nella vecchia architettura del sistema.

2.1 Il vecchio software CLINIC

L'applicazione CLINIC sovrintendeva al processo di elaborazione e valutazione della qualità dati dei file di tipo RINEX [Receiver Indipendent Exchange Format]. L'elaborazione dei dati partiva dall'analisi dei file ricevuti (via FTP) nella area "incoming" del server. L'alberatura del filesystem di ricezione era configurata in funzione delle diverse tipologie di reti gestite. Le utenze FTP di accesso a tale alberatura erano gestite separatamente rispetto alle utenze (ed autorizzazioni) già definite sulla piattaforma web BANCADATI. Il funzionamento di CLINIC era sostanzialmente basato su configurazione e lancio di N script (shell) separati, che si occupavano di gestire i sottocasi di acquisizione che richiedevano, ad esempio, una diversa configurazione dei parametri in ingresso di TEQC, tool a riga di comando utilizzato per l'analisi e l'elaborazione dei file RINEX sviluppato da UNAVCO. Grazie a questo tool per ogni RINEX venivano estratti parametri di qualità come MP1 (rms giornaliero riferito al multipath sulla frequenza L1), MP2 (rms giornaliero riferito al multipath sulla frequenza L2), SLPS (salti di ciclo legati al tracking dei satelliti), GAP (numero di campioni non acquisiti nelle 24 ore di riferimento), HRS (percentuale di completezza del dato giornaliero) ed archiviati in un database relazionale Oracle 10 XE pronti per essere visualizzati nell'applicazione web BANCADATI. I file così processati venivano successivamente "spostati" in apposite cartelle FTP di distribuzione ("outgoing"). Il processo era inoltre in grado di inviare notifiche (e-mail) che riportavano eventuali problemi di elaborazione dati agli utenti che ne avevano effettuato l'upload sul sistema.

2.2 GPSGIVING E GPSFREE

GPSgiving (GPSgiving.gm.ingv.it) era il server FTP che acquisiva ed archiviava tutti i dati delle stazioni GPS della RING e delle stazioni GPS di altre reti in convenzione con INGV. Il software FTP utilizzato era VsFtpDaemon su distribuzione Slackware Linux. La capacità di storage di GPSGiving era estesa da un NAS esterno di circa 2 TB. L'accesso a questo server era limitato alle sole sedi INGV. Era su questo server che l'applicazione CLINIC riceveva i dati RINEX da processare.

La politica di distribuzione dati della RING prevedeva di rendere pubblici solo i dati di un numero limitato di stazioni RING. A tal proposito quando CLINIC analizzava i dati RINEX di queste stazioni ne verificava anche la loro politica di distribuzione e se previsto ne archiviava una copia su GPSFREE (GPSfree.gm.ingv.it), il server FTP accessibile pubblicamente. Anche questo server utilizzava il software VsFtpDaemon su distribuzione Slackware Linux.

2.3 Il vecchio BANCADATI

BANCADATI era una piattaforma web di Knowledge Management per la gestione della Rete Integrata Nazionale GPS. Permetteva sia la centralizzazione delle informazioni in un'unica banca dati comune, che la creazione di un unico servizio di accesso ai dati fruibile da tutte le diverse sedi INGV mediante la rete internet. Il sistema era pensato secondo un'architettura divisa in tre livelli, nella quale i dati o la loro rappresentazione, i tool di gestione e i servizi di portale erano mantenuti separati. Questa separazione logica permetteva l'integrazione delle diverse tecniche di gestione delle basi di dati per ben adattarsi allo sviluppo delle nuove tecnologie legate soprattutto al mondo della ricerca. La realizzazione di tale infrastruttura ha rappresentato la base di partenza per lo sviluppo di nuove soluzioni di knowledge per i diversi gruppi di lavoro dell'ente, in modo da poter cooperare sinergicamente alla realizzazione di un'unica risorsa di dati per l'intera comunità.

Il vecchio sistema consentiva la gestione di tutti i metadati delle stazioni GPS della RING ed in particolare l'aggiornamento costante dei log file di stazione [IGS Site Log Standard]; tramite il suo portale web era possibile verificare in tempo reale i controlli di qualità sui dati RINEX acquisiti effettuati dal software CLINIC col quale ha sempre lavorato in continua sinergia pur rimanendo entrambe due applicazioni indipendenti.

La revisione di BANCADATI e la successiva integrazione col software CLINIC hanno notevolmente migliorato la gestione di tutta l'infrastruttura tecnologica di acquisizione e distribuzione dati della RING e oggi anche di altre reti in convenzione con l'INGV. Pur mantenendo le funzionalità di base per cui era stato pensato, BANCADATI oggi è in grado di fornirne altre in maniera più efficiente ed intuitiva.

Per maggiori informazioni sulla precedente architettura di BANCADATI si veda [Cecere (2007)].

3. La nuova infrastruttura: integrazione di CLINIC e BANCADATI

Le versioni precedenti di BANCADATI e di CLINIC non erano nate da uno specifico progetto ma si sono sviluppate insieme alla rete mano a mano che le esigenze, di processamento ed archiviazione dei dati, lo richiedevano. Non essendoci, quindi, alle spalle un vero e proprio studio delle esigenze e degli obiettivi da raggiungere si è proceduto aggregando, di volta in volta, i vari moduli che le componevano (Server Ftp pubblico e privato, aree riservate ed aree pubbliche, software CLINIC per l'analisi del dato in ingresso e suo interfacciamento con il database Oracle).

Vista la sempre maggiore diffusione di reti di monitoraggio geodetico per fini di posizionamento in tempo reale istituite soprattutto da Regioni, associazioni professionali ed altri enti commerciali, appartenenti anche al settore privato, è nata la necessità di predisporre una nuova infrastruttura. In tale ottica è stata data la possibilità, a tali enti, di aggiornare i log delle loro stazioni e di gestire lo stato della strumentazione, sia attuale che pregresso.

3.1 Requisiti utente

Alla luce di quanto detto è maturata la necessità di snellire l'intera infrastruttura pregressa, integrando le varie funzionalità sia a livello software che hardware. I vantaggi sono stati molteplici, sia a livello di gestione che di costi.

In particolare la nuova architettura doveva prevedere:

- Possibilità di accettare in ingresso dati GPS non solo nel formato RINEX compresso ed hatanakato [Yuki Hatanaka. *A Compression Format and Tools for GNSS Observation Data*] come avveniva in precedenza ma anche dati in formato binario.
- La corretta associazione del nome di stazione al file RINEX deve tener conto del contenuto informativo del file e non semplicemente del suo nome.
- Quando un file RINEX viene acquisito ed elaborato la riscrittura del suo header deve tener conto delle informazioni di sito presenti nel database alla data di riferimento del RINEX processato e non facendo uso di file di configurazione ".cfg" come avveniva in precedenza.

- Possibilità di associare ad ogni utente del sistema determinati privilegi di accesso sia alle pagine web dell'applicativo che all'archivio dei dati GPS accessibili via FTP.
- Sistema di notifiche via mail in caso di riprocessamento dei dati o errori nell'elaborazione di un RINEX.
- Possibilità di graficare i parametri di qualità di ciascuna stazione GPS.
- Gestione dei guasti e delle manutenzioni delle stazioni di monitoraggio GPS.
- Possibilità di tenere traccia dall'applicativo web di tutte le fasi di elaborazione dei RINEX analizzati da CLINIC.
- Unico archivio dati eliminando la gestione dei 2 precedenti FTP Server (GpsGiving e GPSFree) con l'introduzione di un livello software di astrazione del file system.

3.2 Il progetto

Principalmente le attività di ristrutturazione del sistema hanno riguardato il livello dati ed il livello di elaborazione. Per quanto riguarda il database si è effettuato un porting dell'RDBMS da Oracle 10 XE al database open source MySQL 5.5. Più in dettaglio la ristrutturazione ha consentito di:

- Rendere funzionali le strutture delle tabelle rispetto alla attuale evoluzione del sistema;
- Eliminare possibili ambiguità nelle informazioni gestite;
- Migliorare i meccanismi di relazione ed integrità tra le tabelle.

Sono state altresì eseguite attività di normalizzazione della base dati allo scopo di eliminarne le ridondanze e rendere qualitativamente migliore lo schema del database riprogettato.

La strategia di migrazione del database ha previsto una prima fase di analisi delle tabelle, delle relazioni e dei relativi vincoli di integrità. Successivamente, in funzione delle esigenze di evoluzione individuate, è stata ridefinita la nuova base dati in MySQL (con i relativi script), seguita da una fase di *export* dei dati (in forma testuale) presenti del database Oracle tramite *tool* appositi di estrazione dati. I dati estratti sono stati opportunamente filtrati e modificati per poter poi essere reinseriti nella nuova base dati open source. La Fig. 4 mostra le entità dell'attuale database.

La fase successiva ha riguardato il ripristino delle funzionalità web di BANCADATI, in funzione delle modifiche apportate alla base dati. Si è poi proceduto in maniera incrementale alla verifica di tutte le pagine al fine di controllare il completo ripristino delle funzionalità.

Uno sforzo enorme è stato compiuto per evitare la gestione dei 2 server FTP della precedente infrastruttura: GPSGiving e GPSFree. Per questo motivo si è pensato di rendere CLINIC un vero è proprio FTP server con funzionalità aggiuntive di controllo di qualità dei dati acquisiti e di astrazione del file system. I dati sono fisicamente nello stesso archivio ma vengono resi disponibili agli utenti in maniera diversa in funzione dei loro privilegi di accesso. Ciò ha definitivamente permesso di utilizzare un unico FTP server centralizzato configurabile tramite l'applicativo web BANCADATI.

L'ultima fase ha previsto infine il miglioramento dell'interfaccia web di gestione dell'infrastruttura integrata e l'aggiunta di nuove funzionalità che saranno meglio descritte nei prossimi paragrafi.

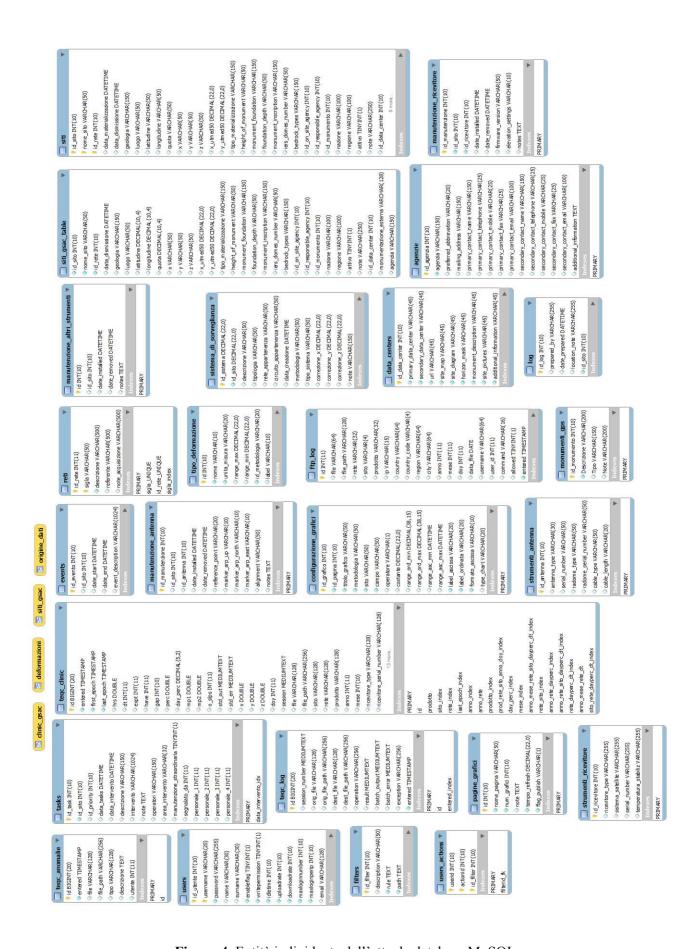


Figura 4. Entità individuate dell'attuale database MySQL.

3.3 Il nuovo CLINIC e l'integrazione con BANCADATI

Rispetto alla precedente architettura, CLINIC è stato integrato all'interno di un'applicazione web che offre funzionalità di FTP server e che utilizza il nuovo BANCADATI. Tale integrazione consente di avere un unico punto di accesso, personalizzabile rispetto alle esigenze della piattaforma, e unifica sia la gestione degli accessi ai file RINEX che la gestione delle informazioni delle reti e dei relativi siti tramite interfaccia web. La Fig. 5 mostra il flusso di lavoro del nuovo CLINIC.

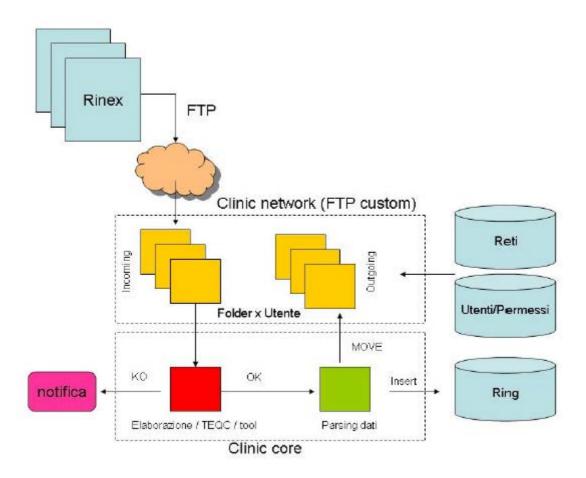


Figura 5. Flusso di lavoro nuovo CLINIC.

Da un punto di vista di configurazione, il nuovo sistema CLINIC, è dotato di funzioni di configurazione, accessibili via web, che oltre a semplificarne l'uso, consentono anche di avere un controllo fine sulle attività di elaborazione dei RINEX, con produzione di report ad hoc.

Da un punto di vista strettamente tecnico la soluzione realizzata per la componente di gestione del protocollo FTP, integrata direttamente nel modulo CLINIC, è basata sull'utilizzo di Apache FTP server, un server open source consolidato che dispone di numerose funzionalità di base, e che è stato personalizzato per le esigenze della piattaforma. In particolare l'attività principale di sviluppo è stata rivolta alla realizzazione di classi Java personalizzate in grado di intercettare qualsiasi comando FTP tra client e server (CLINIC) e rendere visibile all'utente connesso un file system astratto in funzione dei permessi di accesso ai dati che gli sono stati assegnati tramite la nuova interfaccia web (il nuovo BANCADATI).

I processi elaborativi separati, gestiti tramite scripts dal precedente CLINIC, sono stati invece sostituiti da singoli thread Java, che vengono eseguiti all'interno di un unico processo Java, e che condividono altresì la stessa area di memoria, consentendo una comunicazione immediata. Il processo java utilizzato è lo stesso adoperato per "lanciare" la web application che integra a sua volta anche la componente FTP. Una tale architettura ci consente di poter effettuare dei controlli di esecuzione e configurazione dei thread da una unica console web.

La Fig. 6 mostra la nuova architettura integrata.

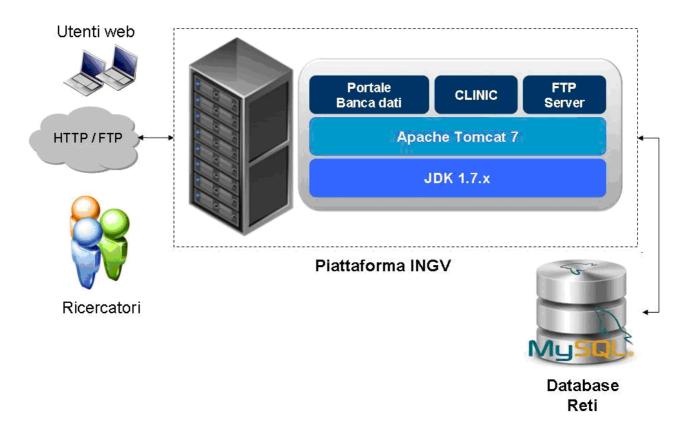


Figura 6. Infrastruttura integrata CLINIC + BANCADATI.

I componenti software di base utilizzati per la realizzazione della nuova architettura sono stati:

- JDK (Java Development Kit) 1.7.x
- Apache Tomcat 7
- Apache FTP Server 1.0.6
- MySQL 5.5.x
- VMWare ESXi 5.1 Essential Kit

Tutti i software utilizzati sono basati su licenza open source tranne che l'hypervisor VMWare Esxi che ospita l'intera infrastruttura software.

Dal punto di vista hardware l'architettura al momento fa uso di un server DELL PowerEdge R710 con doppio processore Xeon quad core, 24 GB di memoria RAM ed un NAS esterno della capienza di circa 2 TB. A breve, inoltre, il NAS sarà sostituito con uno di capienza superiore.

Dal punto di vista software, invece, l'intera infrastruttura può essere schematizzata in maniera semplificata come in figura 7.

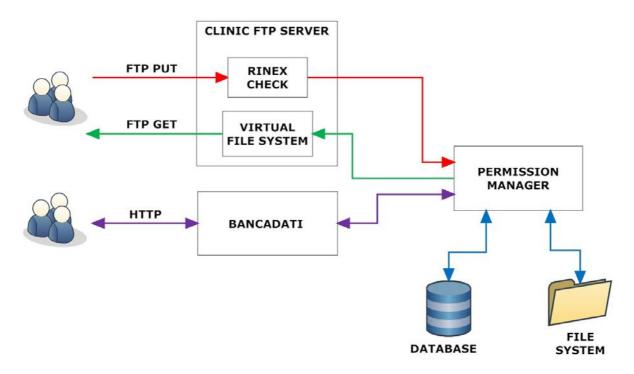


Figura 7. Architettura software sistema di acquisizione e distribuzione dati.

3.4 Il Permission Manager

Il Permission Manager è probabilmente il componente più importante e più critico di tutta l'infrastruttura integrata. È appunto un gestore di permessi scritto in Java; verifica ciò che un utente può fare e su quali dati.

Il permission manager si basa sostanzialmente sui seguenti oggetti ai quali corrispondono entità a se stanti nel database: ACTIONS, FILTERS, USERS.

Le *ACTIONS* rappresentano le azioni basilari che un utente è abilitato o meno ad effettuare. Esse sono ad esempio la sola visualizzazione delle stazioni di una rete, la modifica dei log file, la creazione di una stazione, la sua cancellazione, la possibilità di fare upload FTP su una determinata directory, l'FTP Download per scaricare i dati delle stazioni, la gestione o meno dei task.

I *FILTERS* invece rappresentano delle limitazioni alle singole *ACTIONS*. Entrambe *ACTIONS* e *FILTERS* sono agganciate all'utente al quale si vuole limitare l'uso di quella determinata azione. Per esempio è possibile consentire ad un utente di visualizzare solo determinate stazioni di una determinata rete. In questo caso vi sarà una *ACTION* relativa alla visualizzazione di una specifica rete ed un filtro agganciato a questa *ACTION* che contiene solo i siti abilitati alla visualizzazione. Entrambe, *ACTIONS* e *FILTERS*, sono referenziate all'utente che si intende limitare.

Uno sforzo enorme è stato realizzato in particolare per le *ACTIONS FTP_UPLOAD* e *FTP_DOWNLOAD*. Queste *ACTIONS* anche se configurate tramite l'interfaccia web del nuovo BANCADATI sono elaborate dal nuovo CLINIC che in sostanza è un FTP Server adattato a partire da Apache FTP Server. In questo caso nei filtri si sono adoperate delle vere e proprie espressioni regolari (regexp) che hanno consentito un elevato livello di flessibilità nell'attribuzione dei permessi a ciascun utente.

3.5 L'astrazione del file system

L'FTP server modificato (CLINIC) intercetta ogni comando FTP di accesso a file e cartelle e ne verifica l'autorizzazione col Permission Manager prima di negare o accettare la richiesta dell'utente. Le regexp sono così potenti che ci consentono ad esempio di creare delle vere e proprie astrazioni del file system FTP in fase di navigazione da parte dell'utente.

È possibile, ad esempio, configurare *ACTIONS* e *FILTERS* affinché un utente FTP si trovi dinanzi un file system astratto nel quale riesce a visualizzare i dati di sole determinate reti e per ciascuna di esse solo di determinate stazioni GPS e addirittura all'interno di un certo intervallo temporale per ciascuna stazione GPS

nonostante tutti i file siano fisicamente nello stesso archivio. Tutte le richieste di GET non consentite dal Permission Manager saranno negate dall'FTP Server all'utente finale. Si riporta un esempio di filtro in Fig. 8.



Figura 8. Configurazione di un filtro FTP tramite interfaccia web.

In questo caso, l'utente cui sarà agganciato il seguente filtro, denominato FTP_DOWNLOAD_SITI_OPEN sarà in grado di accedere tramite FTP ai soli dati RINEX a 30 secondi della rete RING e per le stazioni GPS elencate e per ciascuna di esse potrà prelevare solo i dati di un determinato intervallo temporale. Per esempio MOCO[2006259-*] significa che è possibile prelevare della stazione MOCO tutti i RINEX a partire dal giorno giuliano 259 dell'anno 2006 fino ad oggi. I restanti dati non saranno neppure visibili all'utente a cui sarà agganciato questo filtro.

PARTE SECONDA: L'applicativo web

4. L'applicativo web integrato CLINIC-BANCADATI

Di seguito verranno brevemente esplorate le nuove pagine dell'applicativo web BANCADATI e le sue nuove funzionalità.

4.1 Home page utente

L'accesso all'applicativo web avviene mediante autenticazione. Ad accesso effettuato l'utente del sistema viene reindirizzato automaticamente alla propria home. La home elenca le reti GPS (RING + reti convenzionate) che l'utente è abilitato a visualizzare. Nell'esempio riportato in Fig. 9 l'utente "admin" è abilitato alla visualizzazione delle reti RING, ITALPOS e PUGLIANET.

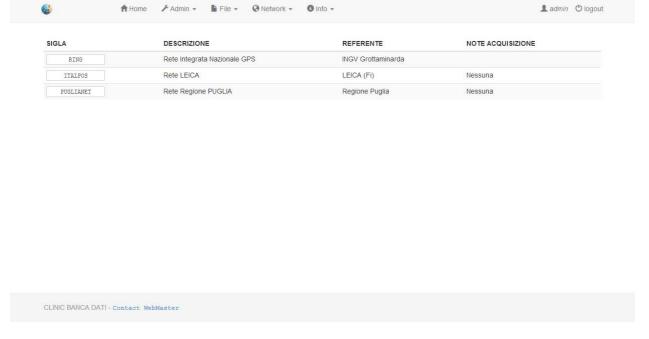


Figura 9. Home utente.

Le reti e le loro stazioni sono inoltre visualizzabili anche su mappa tramite l'apposito tab. La Fig. 10 ne è un esempio.

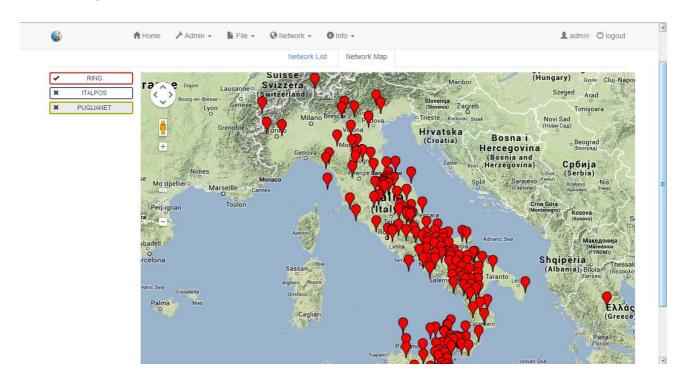


Figura 10. Dislocazione su mappa delle stazioni GPS di ciascuna rete.

Nella parte superiore della pagina sarà sempre presente il menù principale che consentirà ogni volta di tornare alla *home page*, accedere alle impostazioni del sistema (menu *ADMIN*), visualizzare le elaborazioni dei file RINEX (menu *FILE*) o modificare le impostazioni delle varie reti (menu *NETWORK*). I menù o parte di essi saranno visibili all'utente solo se questi dispone dei relativi permessi. I vari sotto-menù saranno maggiormente approfonditi nei paragrafi che seguiranno.

4.2 Home-> Le stazioni di una rete GPS

Cliccando dalla home sulla sigla di una rete GPS vengono mostrate le stazioni che l'utente connesso è abilitato a visualizzare per quella rete. Le modalità di visualizzazione delle stazioni sono quattro: grid, map, list e task. La modalità di default è quella grid.

4.2.1 GridView

Nella modalità Grid-View le stazioni di una rete vengono disposte a griglia così come mostrato dalla Fig. 11. A ciascun riquadro identificante la singola stazione GPS è associata una ulteriore griglia che ne descrive la presenza o meno di file RINEX acquisiti regolarmente negli ultimi 7 giorni. Il colore rosso identifica un giorno in cui il dato GPS RINEX non è stato acquisito, il colore verde, invece, il caso contrario.



Figura 11. Visualizzazione a griglia delle stazioni GPS di una rete.

4.2.2 MapView

Nella visualizzazione a mappa (Fig. 12) le stazioni GPS sono rappresentate da un cerchio che può assumere la colorazione verde, arancio o rosso a seconda che negli ultimi 2 giorni vi siano:

- i dati di entrambi i giorni (verde)
- almeno un giorno di dati (arancio)
- nessun dato (rosso)

Cliccando sulla singola stazione un *tooltip* visualizzerà lo stato dell'acquisizione degli ultimi 7 giorni della stazione così come avviene nella modalità di visualizzazione GridView. Infine cliccando sulla sigla della stazione si accede alle informazioni di dettaglio della singola stazione.

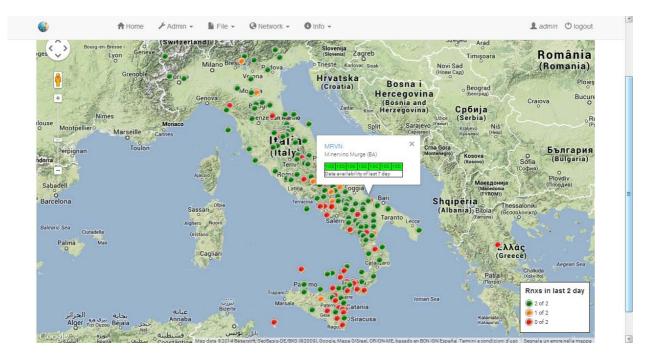


Figura 12. Visualizzazione su mappa delle stazioni GPS di una rete.

4.2.3 ListView

Nella modalità di visualizzazione ListView (Fig. 13) le stazioni sono elencate in tabella in ordine alfabetico insieme alle strumentazioni di cui si compongono (tipo ricevitore, tipo antenna, versione firmware, seriali...). È possibile inoltre applicare filtri a ciascun campo della tabella in modo da estrarre ad esempio solo le stazioni che hanno un determinato ricevitore GPS. La tabella filtrata o non può essere esportata in formato CVS tramite il tasto.

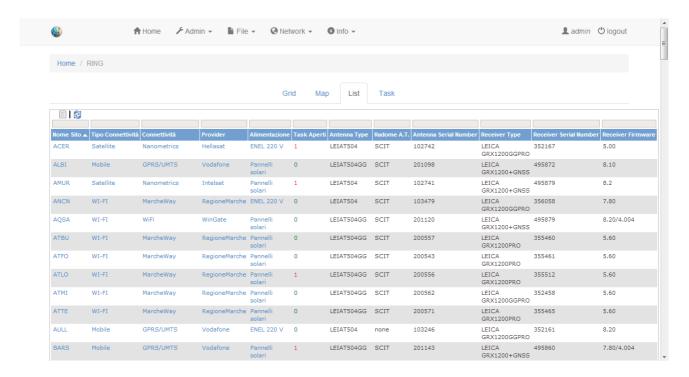


Figura 13. Visualizzazione a lista delle stazioni GPS di una rete.

4.2.4 TaskView

L'ultima modalità di visualizzazione è la TaskView (Fig. 14). In questa modalità è possibile visualizzare l'elenco dei task aperti per ciascuna stazione GPS o i task relativi ad un particolare anno. La tabella è esportabile in formato CSV tramite l'apposito tasto .

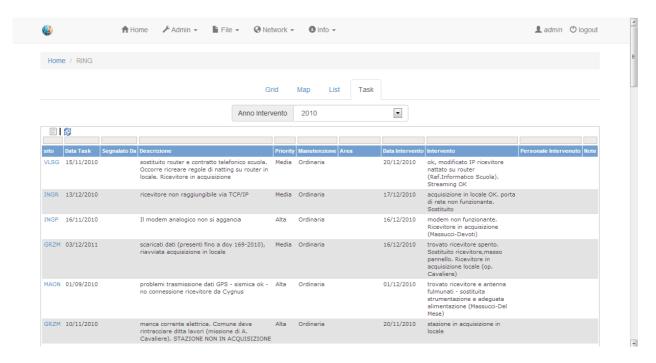


Figura 14. Visualizzazione dei task delle stazioni GPS di una determinata rete.

4.3 Informazioni di dettaglio stazione GPS

Da ciascuna modalità di visualizzazione è possibile cliccare sulla sigla stazione per visualizzarne le sue informazioni di dettaglio.

4.3.1 Statistics Stazione

Quando ci si riferisce ad una particolare stazione GPS la prima informazione che viene visualizzata è la sua disponibilità di dati degli ultimi 4 anni come mostrato in Fig. 15. Le osservazioni GPS sono archiviate in file RINEX giornalieri di 24 ore. Il parametro PERC ne stima il livello di completezza elaborando il numero di osservazioni giornaliere registrate rispetto a quelle attese nell'arco delle 24 ore. Nella visualizzazione di default la colorazione dei quadranti dal rosso al verde identifica la percentuale di osservazioni giornaliere di quel mese/anno con un livello di completezza giornaliera del dato superiore all'80%. È possibile dinamicamente interrogare il database affinché mostri la percentuale di dati mensile che ha superato una soglia di completezza personalizzabile. Come pure è possibile analizzare un intervallo temporale superiore ai 4 anni di default. La Fig. 15 mostra lo Statistics di dettaglio di una stazione GPS.

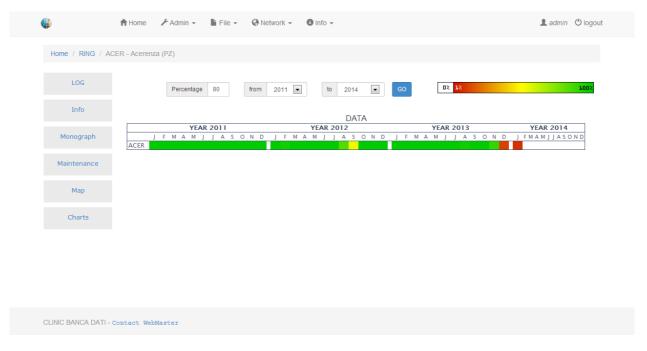


Figura 15. Statistics relativo ad una particolare stazione GPS.

Cliccando sul quadrante che identifica un particolare mese vengono visualizzate in maniera dettagliate tutte le osservazioni giornaliere di quel mese e la loro percentuale di completezza (Fig. 16).

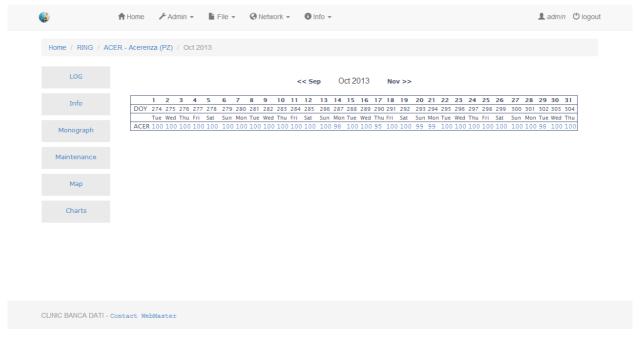


Figura 16. Percentuale completezza dati di uno specifico mese.

Cliccando sulla singola percentuale è possibile inoltre scaricare il dato RINEX giornaliero o visualizzarne il risultato del controllo di qualità (l'output del software TEQC) come mostrato in Fig. 17.

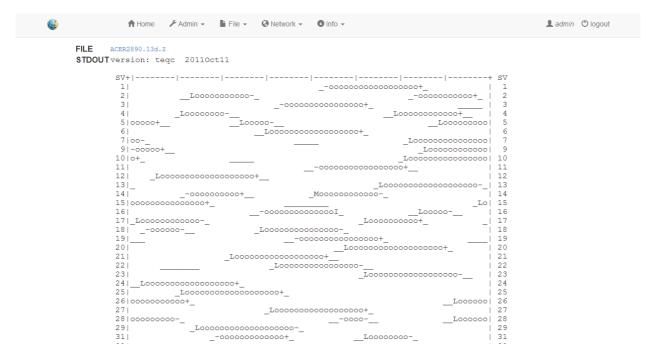


Figura 17. Output di controllo qualità eseguito dal tool TEQC.

4.3.2 LOG FILE di Stazione

I LOG file delle stazioni GPS [IGS Site Log Standard] rappresentano la carta di identità delle stesse e sono al centro del processo di elaborazione di CLINIC. Forniscono informazioni dettagliate riguardo il tipo di ricevitore, antenna, seriali delle stazioni e tengono traccia nel tempo di eventuali sostituzioni dovute a guasti o riparazioni. Da questa pagina (Fig. 18) l'utente può scaricare il LOG file della stazione o compilarlo in ogni sua sezione ammesso che ne abbia l'autorizzazione.

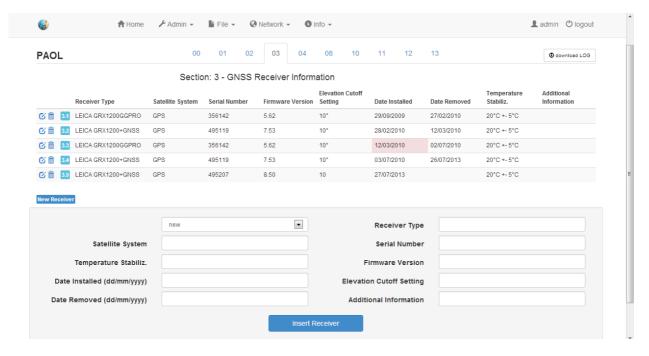


Figura 18. Pagina di editing del LOG di una stazione.

4.3.3 Informazioni aggiuntive di stazione

La scheda info (Fig. 19) consente di memorizzare per la singola stazione GPS informazioni riguardanti la connettività del sito, il provider di telecomunicazione, il tipo di alimentazione ed il referente di quel sito. È possibile inoltre archiviare un album fotografico della stazione ed in particolare le foto di sito e monumento che verranno automaticamente utilizzate dall'applicativo per la composizione automatica del PDF relativo alla monografia del sito.

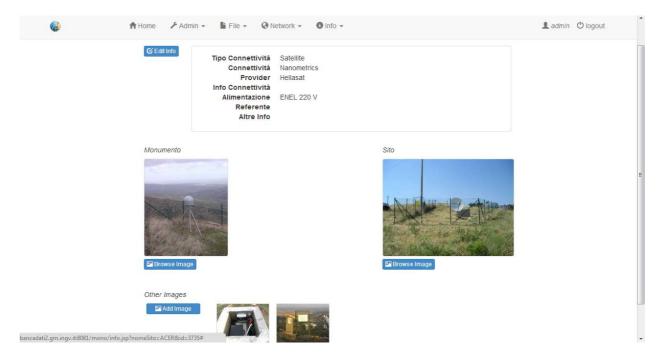


Figura 19. Pagina di editing delle informazioni aggiuntive di stazione.

4.3.4 Monografia di stazione

Il tasto monograph consente di realizzare "on the fly" e scaricare la monografia del sito in formato PDF. Di seguito un esempio di monografia (Fig. 20).



Figura 20. Monografia in formato PDF di una stazione GPS.

4.3.5 Manutenzione stazione

Grande importanza è stata data alla gestione degli interventi di manutenzione che vengono effettuati sui siti remoti. La pagina Maintenance (Fig. 21) mostra per ciascun sito gli interventi da effettuarsi (task aperti) o gli interventi effettuati (task chiusi).

I task sono ordinati per data di chiusura; ad ogni task viene assegnata una priorità ed è possibile tener traccia anche fotograficamente degli interventi effettuati per la risoluzione di un determinato task.

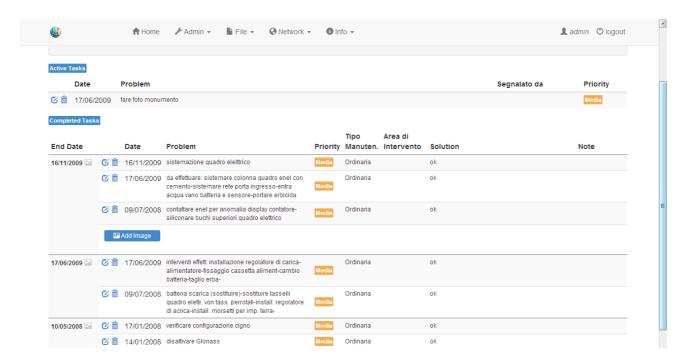


Figura 21. Pagina di gestione task di una stazione GPS.

4.3.6 Mappa dislocazione stazione

Il tasto MAP consente di visualizzare su google-maps la collocazione geografica della stazione GPS.

4.3.7 Grafici di stazione

Abbiamo dato molta importanza alla rappresentazione grafica (Fig. 22) di alcuni parametri di qualità della stazione GPS. In questo modo riusciamo a stimarne la qualità del dato e a confrontarlo con le altre stazioni della stessa rete. I grafici sono realizzati in tempo reale sui seguenti parametri di qualità: MP1, MP2, SLPS, HRS, GAP e PERC. Nella parte superiore dello schermo vi è la rappresentazione grafica del singolo parametro di stazione in funzione del giorno giuliano di acquisizione. Nel caso in cui presso la stazione è stato effettuato un cambio ricevitore o antenna questo sarà visibile con una barra verticale tratteggiata.

Nella parte sottostate invece lo stesso parametro viene mediato negli ultimi 45 giorni e confrontato con lo stesso parametro delle altre stazioni della stessa rete. La barra verticale ne rappresenta lo scarto quadratico medio.



Figura 22. Grafici sui parametri MP1, MP2, PERC, SLPS, GAP e HRS di una stazione GPS.

4.4 Pannello Admin

Il menù Admin è per ovvie ragioni riservato agli amministratori del sistema. Da qui è infatti possibile aggiungere utenti, modificarli e assegnare loro permessi di accesso al portale web e al recupero dati tramite FTP. È inoltre possibile impostare i filtri di cui si è discusso nel paragrafo 3.4.

4.4.1 Admin --> TeqcLog

Un importante strumento di analisi delle elaborazioni dei dati RINEX effettuate dal sistema è visualizzabile dalle pagine raggiungibili tramite il menù file. In particolare *TeqcLogs* consente all'amministratore di sistema di visualizzare tutte le operazioni alle quali i file RINEX vengono sottoposti fino al momento in cui, superati i controlli di qualità, vengono posti in archivio pronti per essere distribuiti. Eventuali errori prodotti da TEQC durante il processamento di questi file possono essere visualizzati sia nella pagina *TeqcLogs* sia nella pagina *TeqcAnomalie*. Inoltre è possibile, in qualsiasi momento, filtrare i LOG files per verificare ad esempio i passi di elaborazione a cui è stato sottoposto un determinato RINEX; o ancora filtrare tutti i RINEX che non hanno superato determinati controlli di qualità. La Fig. 23 mostra la pagina *TeqcLogs*.

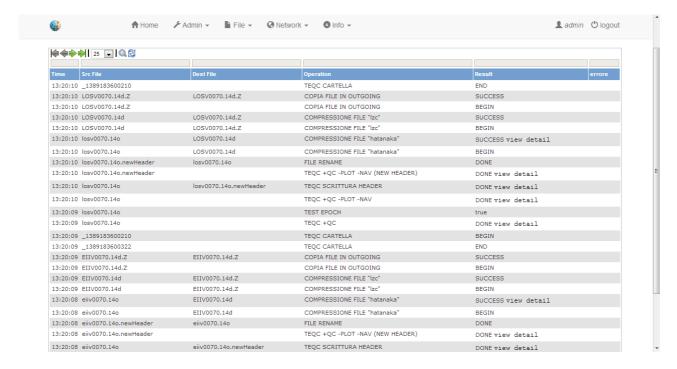


Figura 23. Pagina TeqcLogs.

4.4.2 Admin --> TeqcReprocess

Un'importante funzionalità del sistema è quella di riprocessare in maniera automatica i file dati di una determinata stazione GPS in un intervallo temporale ben definito. Spesso accade, ad esempio, che interventi di manutenzione effettuati sulle stazioni remote non vengono riportati nella BANCADATI in maniera celere oppure l'operatore che ha effettuato l'intervento omette di comunicare eventuali operazioni importanti come il cambio firmware o sostituzione antenna. In questi casi i dati GPS continueranno ad essere acquisiti con informazioni dell'header errate. Per ovviare a questo inconveniente, se ci si accorge di una modifica strumentale presso una stazione, è possibile in maniera del tutto automatica chiedere al sistema di riprocessare i dati GPS in un determinato intervallo di tempo in modo che i realitivi header possano essere aggiornati con le corrette informazioni strumentali. La Fig. 24 mostra la pagina *TeqcReprocess*.

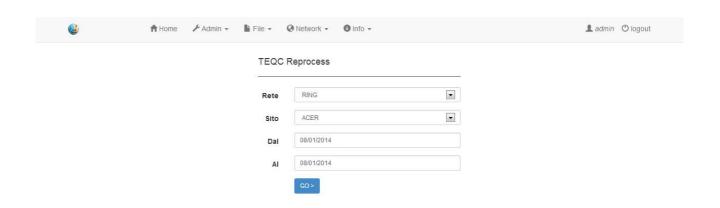


Figura 24. Pagina TeqcReprocess.

4.5 Statistics

La pagina Statistics presente nel sistema consente di verificare per ciascuna stazione della rete selezionata la percentuale di osservazioni giornaliere pervenute in un determinato mese/anno. È la stessa pagina che viene visualizzata nel dettaglio di ciascuna stazione ma in questo caso le informazioni riguardano l'intera rete. È possibile anche in questo caso selezionare l'intervallo temporale di riferimento oppure interrogare il database per le sole osservazioni che hanno superato una percentuale di completezza giornaliera dei dati superiore ad un determinato valore. La Fig. 25 mostra lo Statistics della rete RING.



Figura 25. Pagina Statistics dell'intera rete RING.

5. GSAC: integrazione archivio RING con le reti convenzionate UNAVCO

GSAC Web Service è un pacchetto software sviluppato da UNAVCO e U.S. National Science Foundation insieme ai partner CDDIS e SOPAC. Esso fornisce accesso web agli archivi geodetici dislocati in tutto il mondo. Lo scopo principale di GSAC è quello di fornire un servizio web completo, moderno e consistente che permetta agli utenti remoti di esplorare i siti GPS delle varie reti mondiali, accedere alle informazioni riguardanti le loro strumentazioni e scaricare i dati acquisiti dai diversi archivi. Tutti i datacenter che acquisiscono e distribuiscono pubblicamente dati GPS possono utilizzare GSAC per contribuire alla condivisione dei dati GPS con gli scenziati di tutto il mondo. Usando GSAC, un datacenter può fornire dei servizi web standard agli utenti affinché possano interrogarlo e scaricare i dati dai loro archivi.

Un archivio GSAC può essere interrogato attraverso le pagine web del datacenter presso cui è installato, attraverso un client software fornito con GSAC stesso o direttamente da una linea di comando Linux. Una singola richiesta a GSAC può trovare e scaricare centinaia di dati GPS insieme alle informazioni strumentali delle stazioni. La Fig. 26 mostra il funzionamento di GSAC.

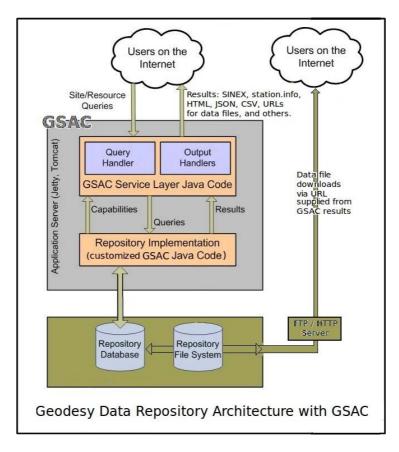


Figura 26. Schema del software GSAC.

GSAC è stato integrato nell'architettura di acquisizione e distribuzione dati GPS; per consentirgli di accedere alle informazioni delle stazioni presenti nel nostro database sono state create delle opportune viste nel formato predisposto dagli sviluppatori di GSAC. La Fig. 27 mostra il web service installato presso il datacenter della RING a Grottaminarda.



Figura 27. Servizio Web Gsac della RING.

6. Conclusioni

Il sistema di acquisizione e distribuzione dati GPS, presentato in questo rapporto tecnico, costituisce un'importante infrastruttura di ricerca dell'INGV; nato inizialmente allo scopo di acquisire e distribuire i soli dati GPS della RING, è oggi in grado di fornire lo stesso servizio anche alle reti GPS di Enti o privati già presenti sul territorio nazionale e convenzionate con l'INGV. Le nuove funzionalità di BANCADATI consentiranno ai gestori delle singole reti di tener aggiornate, in maniera autonoma, le informazioni strumentali delle proprie stazioni di monitoraggio.

In ambito internazionale, inoltre, l'infrastruttura ben si candida all'integrazione con le altre reti di monitoraggio terrestri europee nell'ambito dello European Plate Observing System (EPOS).

EPOS è l'infrastruttura integrata di ricerca delle scienze che studiano la terra solida approvata dall'ESFRI (European Strategy Forum on Research Infrastructures) ed inclusa nella sua roadmap dal Dicembre 2008. EPOS costituisce un piano di integrazione a lungo termine delle varie infrastrutture di ricerca nazionali.

Lo scopo di EPOS è promuovere e rendere possibili approcci innovativi per una migliore comprensione dei processi fisici che governano i terremoti, le eruzioni vulcaniche, gli tsunami e la dinamica della superficie terrestre. L'integrazione delle infrastrutture di ricerca nazionali esistenti incrementerà l'accesso e l'uso dei dati multidisciplinari registrati dalle rete di monitoraggio terrestri, acquisiti in esperimenti di laboratorio e/o prodotti da simulazioni computazionali.

Un'importante upgrade al sistema ci consentirebbe di entare in EPOS garantendo un maggior livello di disponibilità del servizio e una strategia di disaster recovery in caso di guasti.

In tal senso, da un'analisi preliminare, l'infrastruttura descritta potrebbe essere distribuita non solo nella sede Irpinia dell'INGV (dove attualmente risiede il datacenter) ma anche nella sede INGV di Roma. La replica di questi servizi potrebbe essere quella raffigurata in Fig. 28.

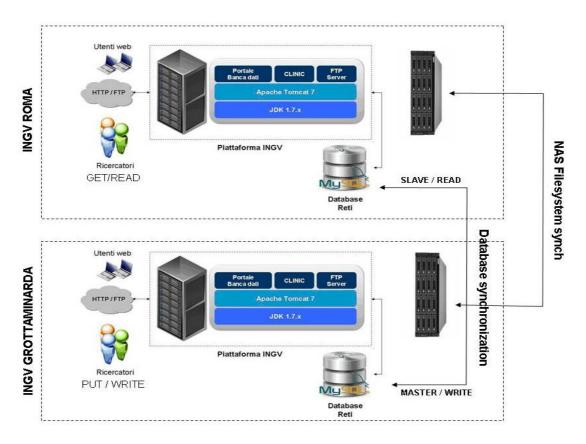


Figura 28. Replica servizi infrastruttura RING tra le sedi di Roma e Grottaminarda dell'INGV.

Bibliografia

- Avallone A., Selvaggi G., D'Anastasio E., D'Agostino N., Pietrantonio G., Riguzzi F., Serpelloni E., Anzidei M., Casula G., Cecere G., D'Ambrosio C., De Martino P., Devoti R., Falco L., Mattia M., Rossi M., Obrizzo F., Tammaro U., Zarrilli L. (2010). The *RING network: improvements to a GPS velocity field in the central Mediterranean*. Annals of Geophysics, 53, 2, APRIL 2010; doi: 10.4401/ag-4549.
- Cardinale V., Castagnozzi A., D'Ambrosio C., Falco L., Memmolo A., Minichiello F. (2010). Wi-Fi Mesh Network: integrazione dell'infrastruttura telematica della rete sismica e geodetica nazionale. Rapporti Tecnici INGV, n. 141, 2010.
- Cecere G. (2007). La piattaforma tecnologica di gestione dati e informazioni della Rete Integrata Nazionale GPS (RING). Rapporti Tecnici INGV, n. 52.
- Falco L. (2006). Realizzazione rete di acquisizione dati Nanometrics e segmento PDMZ (Partial DeMilitarized Zone) della rete telematica della sede di Grottaminarda dell'Istituto Nazionale di Geofisica Vulcanologia. Rapporti Tecnici INGV, n. 35.
- Falco L. (2008). Implementazione e gestione di una rete di monitoraggio GPS e sismica mediante tecnologie GPRS/EDGE/UMTS/HSDPA. Rapporti Tecnici INGV, n. 69.
- IGS Site Log Standard ftp://igs.org/pub/station/general/blank.log, ftp://igs.org/pub/station/general/sitelog_instr.txt
- Oracle. Java EE 7 Documentation API http://docs.oracle.com/javaee/7/api/
- Oracle. Database 10 XE Documentation http://www.oracle.com/pls/xe102/homepage
- Oracle. MySql 5.6 Documentation http://dev.mysql.com/doc/refman/5.6/en/index.html
- RINEX Format Standard ftp://igscb.jpl.nasa.gov/pub/data/format/RINEX210.txt
- The Apache Software Foundation. Apache Ftp Server Documentation http://mina.apache.org/ftpserver-project/documentation.html
- The Apache Software Foundation. Apache Tomcat 7 Documentation http://tomcat.apache.org/tomcat-7.0-doc/
- UNAVCO GSAC WEBServices http://facility.unavco.org/data/gsacws/gsacws.html
- Yuki Hatanaka. *A Compression Format and Tools for GNSS Observation Data* http://www.gsi.go.jp/common/000045517.pdf

Coordinamento editoriale e impaginazione

Centro Editoriale Nazionale | INGV

Progetto grafico e redazionale

Daniela Riposati | Laboratorio Grafica e Immagini | INGV

© 2014 INGV Istituto Nazionale di Geofisica e Vulcanologia Via di Vigna Murata, 605 00143 Roma Tel. +39 06518601 Fax +39 065041181

