r t Anno 2009_Numero 113

rapporti tecnici INGV

Progetto e realizzazione del nuovo servizio integrato di posta elettronica della Sede Centrale dell'INGV

113

Istituto Nazionale di Geofisica e Vulcanologia

Anno 2009_Numero 113

rapporti tecnici INGV

PROGETTO E REALIZZAZIONE DEL NUOVO SERVIZIO INTEGRATO DI POSTA ELETTRONICA DELLA SEDE CENTRALE DELL'INGV

Melissa Mendicino, Massimiliano Rossi, Manuela Sbarra, Gianpaolo Sensale, Diego Sorrentino, Francesco Zanolin, Lucio Badiali, Francesca Caprara, Pietro Ficeli

INGV (Istituto Nazionale di Geofisica e Vulcanologia, Centro Nazionale Terremoti – Servizi Informatici e Reti)

113

Indice

Introduzione

Nell'ultimo decennio, con l'aumento delle attività svolte al suo interno ed il conseguente aumento del personale, il nostro Ente si è trovato dinanzi alla necessità di dover gestire una crescita esponenziale delle comunicazioni effettuate attraverso l'uso della posta elettronica nonché un notevole aumento di utenze che ne usufruiscono.
Ad oggi il modo più veloce e sicuro per comunicare, anche per trasferire piccole quantità di dati, è proprio tramite posta elettronica.
L'evoluzione delle tecnologie informatiche hanno permesso di migliorare il sistema precedente con l'introduzione di nuove funzionalità che in precedenza non erano state considerate né valutate per il lavoro.
A tale scopo dopo un lungo periodo di acquisizione delle informazioni, analisi del sistema e tipo di lavoro svolto dal gruppo nella gestione del sistema per le attività straordinarie e di routine simulando anche differenti scenari di situazioni a regime, si è iniziata la progettazione. Il risultato si è ottenuto il 25 di Febbraio 2009 quando è stato messo in funzione il nuovo "Sistema Mail" che in questo documento verrà ampiamente descritto. Il proposito, che pensiamo di aver raggiunto, è di aver superato alcune limitazioni presenti nel precedente sistema per quanto concerne temi che non erano all'ordine del giorno in passato ma comunque sentiti dai colleghi che lo gestivano e che hanno partecipato al progetto ed alla transizione. Si è puntato molto su *l'alta affidabilità*, *la ridondanza dei sistemi* (secondo la letteratura moderna)*, la sicurezza*, *la velocità di ripristino* e per ultima ma non meno importante *l'archiviazione su sistemi* che riducono la possibilità di perdita di dati. Vi è anche da considerare la prepotente velocità con cui l'informatica cambia l'orizzonte delle scelte con sempre nuove possibilità e richiede periodicamente una revisione ed a volte una rivoluzione delle scelte del passato.
Questo documento non ha il compito di descrivere nel dettaglio "come si crea un mail server", in quanto esistono numerosi testi che ne descrivono in modo esaustivo la creazione con differenti tecniche possibili, con pro e contro tutti dal SIR valutate, ma si pone il problema di descrivere al lettore, l'architettura e le metodologie utilizzate per la gestione del nuovo mail server. Si ricorda inoltre quanto vitale sia il sistema di gestione posta nell'invio di dati sugli eventi sismici e quanto la Sede Centrale scambi in tempo reale fasi sismiche e parametri necessari alle localizzazioni. Inoltre presso la Sede Centrale risiede la mail di Presidenza, Direzione Generale e servizi amministrativi.

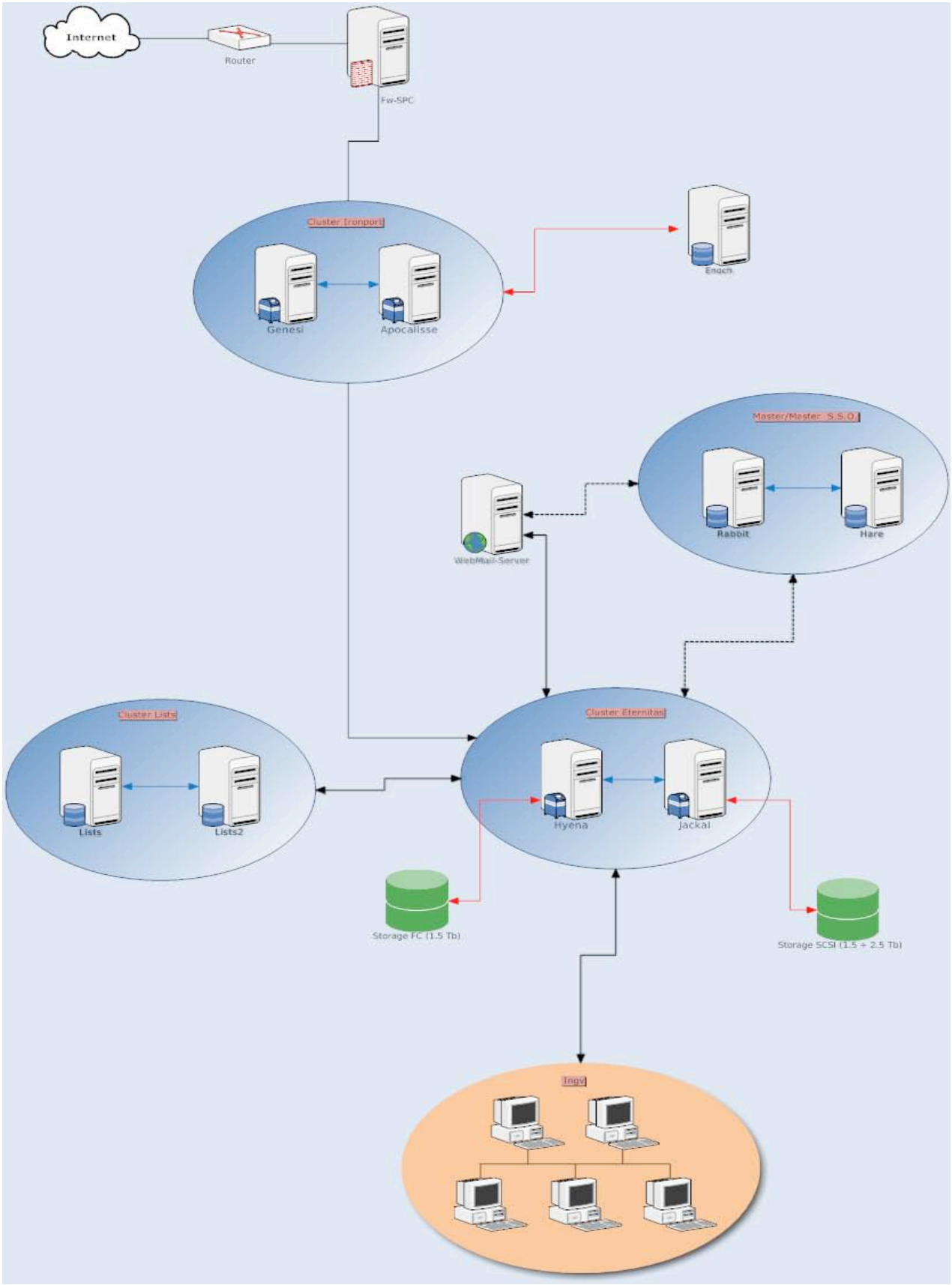

Figura 1. Sistema di gestione posta della Sede Centrale dal febbraio 29009. Firewall e connettività sono in alta affidabilità.

1. Vecchio e nuovo a confronto

In questa parte del documento, vengono messi a confronto alcune delle caratteristiche tecniche che distinguono il precedente sistema di gestione e-mail dal nuovo sistema.

1.1 Scelte Architetturali

Il vecchio sistema, dal momento in cui è stato preso in gestione dal SIR è stato inizialmente modificato per superare alcune problematiche quali la perdita di e-mail, l'elevato livello di spam che ogni utente subiva, problemi strettamente collegati all'archiviazione su raid ormai obsoleti, il blocco del sistema dovuto alla presenza di sistemi anti-spam ed anti-virus che, ad ogni aggiornamento, non riprendevano il normale funzionamento provocando una interruzione del servizio.
Le modifiche orientate al mantenimento del sistema funzionante fino alla completa sostituzione, hanno fatto si che ci fosse un notevole miglioramento del servizio ma si è ritenuto comunque necessario creare un nuovo sistema da zero, che fosse di più facile gestione per tutti.
L'architettura del nuovo sistema doveva garantire alcuni punti fondamentali quali:

- *Scalabilità*: a fronte dell'aumento della richiesta di risorse, fisiologica o eccezionale, da parte di uno o più servizi l'architettura deve permettere l'aumento delle capacità di elaborazione a fronte di modifiche limitate all'architettura stessa.
- *Fault-tollerance*: l'architettura deve permettere al servizio di continuare a funzionare anche in caso di guasto di un qualsiasi componente riducendo al minimo i tempi di ripristino.
- *Tecnologia*: il progetto deve impiegare lo stato dell'arte in campo tecnologico e valutare possibilmente il trend dei prossimi anni.
- *Evoluzione*: l'architettura deve permettere la naturale evoluzione dei sistemi e dei servizi coinvolti. Deve inoltre, per quanto possibile, essere neutra rispetto alla tecnologia dei sistemi che la compongono.

1.2 Caratteristiche del vecchio Mail Server

- Sistema a macchina singola di vecchia generazione non ridondata.
- Spazio di archiviazione di dimensione contenuta e non ridondata.
- Sistema obsoleto e non aggiornabile basato su una distribuzione linux ottenuta dalla riduzione della distribuzione RedHat e non più aggiornabile da tempo.
- Uso di vecchie versioni software non più supportate.
- Uso degli alias per la gestione delle liste utenti.
- Sistema antispam e antivirus gratuito non ottimizzato.
- Uso di una struttura S.S.O (Single Sign On) remotizzata e non amministrata dai sistemisti INGV ad esclusione dell'inserimento degli account con gli ovvi problemi di sicurezza e di rispetto delle leggi.
- Sistema S.S.O contenente le utenze di personale non esistente, errati, non più facenti parte dell'Istituto.
- Accesso alla struttura via ssh con permessi esclusivamente di amministratore senza possibilità di controllare gli accessi e le operazioni dei singoli amministratori.
- Backup inesistenti o manuali.
- Mancanza di applicazioni dedicate alla gestione remota delle informazioni.
- Mancanza di reportistica adeguata al controllo del funzionamento del sistema.
- Mancanza di un sistema automatizzato di allarme in caso di problemi.

1.3 Caratteristiche del nuovo Mail Server

- Macchine di ultima generazione e in topologia ridondata (Cluster).
- Spazio di archiviazione di più grosse dimensioni ridondato, più un sistema di archiviazione periodica su nastro.
- Software di ultima generazione basato su un sistema linux di uso comune con avviso automatico in presenza di aggiornamenti disponibili.
- Accesso alla struttura via ssh con uso di permessi gerarchici.
- Controllo degli accessi e delle operazioni tramite log ed accounting.
- Struttura dedicata per la gestione dell'AntiSpam e dell'AntiVirus ("Ironport").
- Struttura dedicata per la gestione e la centralizzazione degli account (S.S.O) interna all'Istituto con struttura aperta e ridondata.
- Struttura dedicata per la gestione delle mailing list (Mailman).
- Software Webmail con funzionalità di Groupware e gestione progetti (Organizer, Webmail).
- Uso di un'applicazione dedicata per la creazione e modifica delle informazioni degli utenti (MailControlPanel).
- Uso di funzionalità aggiuntive fra cui:
  accesso tramite canale cifrato, *vacancy* e *forward, account a tempo e/o con funzionalità limitate*.
- Backup periodico del sistema, dei dati e dell'archivio e-mail centralizzato.
- Presenza di una infrastruttura automatizzata di allarme (24x7) per la gestione e risoluzione delle eventuali problematiche (Logwatch, Sentinet, Cacti) con avviso tramite sms agli amministratori.

Sintesi del sistema attuale:	
Dimensioni Mail Box Virtuali per singolo account	1Gb
Dimensione massima allegati per singola e-mail	30 Mb
Numero Account Personali	555 Account
Numero Account di Servizio	70 Account
Numero Alias (Account e Mailing List)	656 Alias
Numero Mailing List ufficiali	119 Mailing List

2. Architettura di un Mail Server

2.1 Scelte strategiche

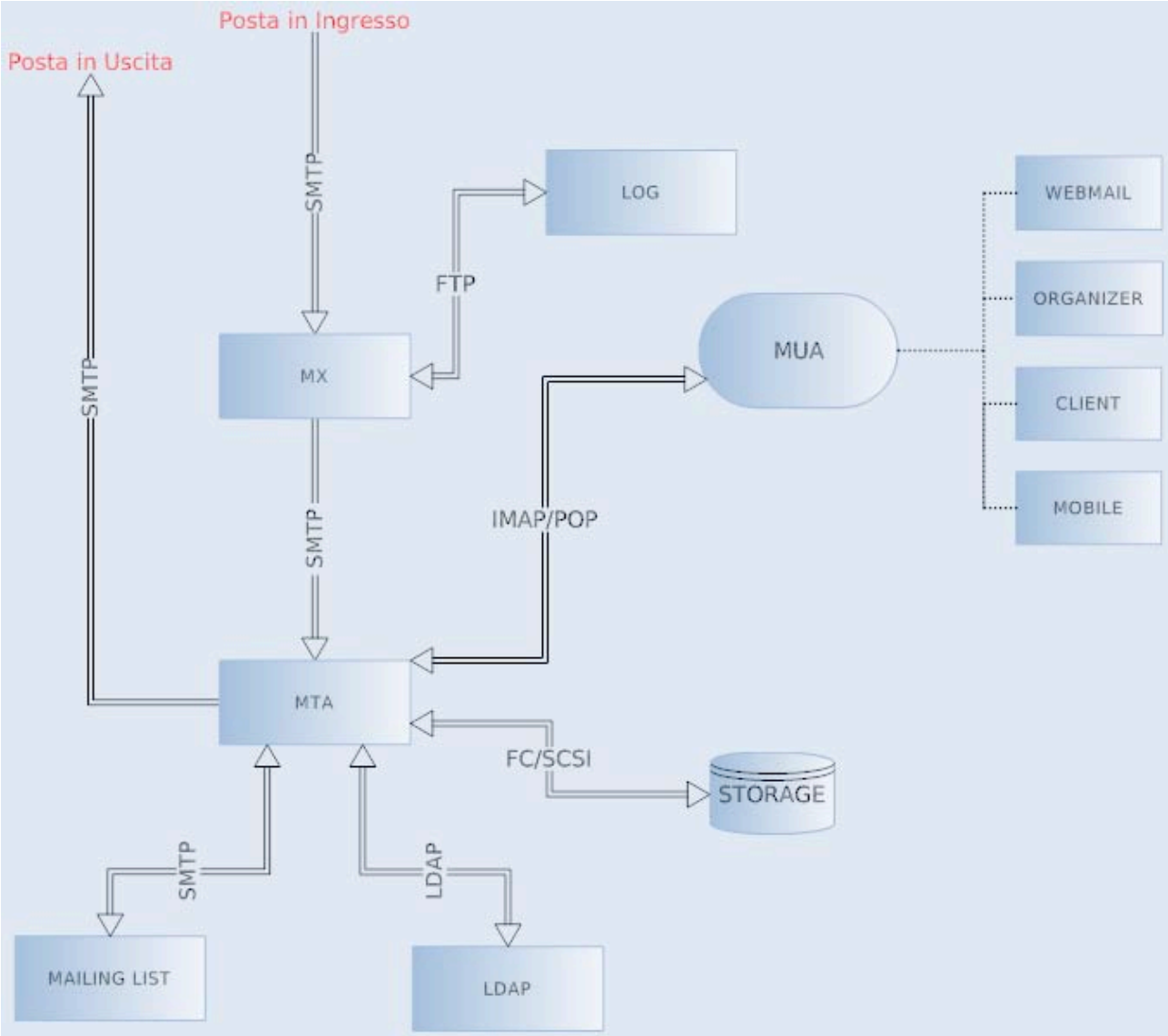

Figura 2. Configurazione Sistema Mail Server.

Per assicurare il soddisfacimento delle qualità sopra espresse è stato deciso di implementare una struttura con servizi distribuiti basati su sistemi ridondati.
Il servizio è stato quindi suddiviso nei seguenti sotto sistemi:

2.1.1 MX: Antispam, Antivirus

Questo servizio si occupa di ricevere le mail provenienti dai mail server esterni.
È basato sull'uso di due appliance hardware, gli Ironport C100, collegati fra loro in cluster, i quali interagiscono tra di loro operando il filtraggio della posta tramite l'uso di filtri di reputazione, proprietari e personalizzati.

Ogni appliance C100 è formato da:

Caratteristiche Hardware	
Chassis/Processore	1 U Rack CPU Single Intel Processor
Storage	Raid: Raid1 Hard Disk: Two 250 Sata drives
Memoria	RAM 4GB
Connettività	2 Ethernet Embedded Intel Gigabit Nics 3 1 DB 9 Serial Port
Caratteristiche Software:	
Protocolli mail	Protocolli in ingress: SMTP, ESMTP, Secure SMTP over TLS. Protocolli in ingress: SMTP, ESMTP, Secure SMTP over TLS.
DNS	Resolver/cache interno; può eseguire le traduzioni op attraverso il dns locale o tramite i DNS server posti in internet.
Interfaces/Configuration	Interfaccia web accessibile tramite HTTP o HTTPS Interfaccia a riga di comando accessibile via SSH o telnet. Wizard di configurazione. Trasferimento file SCP o FTP. Monitoraggio programmato: File di configurazione basati su XML

Su ogni appliance gira un sistema operativo proprietario AsyncOS, sviluppato dall'Ironport appositamente per gestire grossi volumi di mail.
Il sistema di gestione delle mail può essere amministrato sia direttamente a linea di comando che attraverso l'applicazione web proprietaria.
In entrambi i casi è possibile modificare tutta la parte di configurazione sia inserenti al sistema che alle regole di filtraggio.
L'applicazione si occupa inoltre di mantenere un sistema di log grafici, che possono essere esportati, per poi riutilizzarli a piacere.
Inoltre è possibile impostare un invio giornaliero, settimanale o mensile dei report generati dal sistema.

2.1.2 Log: Mail Flow Central

Il Mail Flow Central (MFC) si occupa di mantenere i log delle mail ricevute dagli MX e di archiviarle attraverso un programma sviluppato dall'Ironport specificatamente per l'uso con gli appliance Ironport.
L'applicazione MFC si appoggia ad un DataBase basato su mysql. Acquisisce, attraverso il protocollo FTP, dagli appliance Ironport ogni 30 minuti le intestazioni di ogni mail transitato attraverso gli stessi, archiviandone oltre all'intestazione stessa anche altre informazioni quali data ed ora di arrivo della mail, appliance di transito, ed informazioni sui risultati delle analisi effettuate. Inoltre, viene indicato se la mail e stata accettata o rifiutata.
Tale applicazione è installata su un server DELL PowerEdge 1850 avente le seguenti caratteristiche:

Caratteristiche Hardware:	
Chassis / Processore	1U rack CPU: Single Intel Xeon Processor 3.00 Ghz
Storage	RAID: RAID 1 Hard disk: Two 147GB Scsi drives
Memoria	RAM: 4GB
Connettività	2 Ethernet Embedded Intel Gigabit NICs 1 DB-9 Serial Port
Caratteristiche Software:	
S.O.	Windows 2003 server
Interfaces / Configuration	Interfaccia web accessibile tramite HTTP o HTTPS Interfaccia a riga di comando accessibile via SSH o Telnet Wizard di configurazione Trasferimento File: FTP DataBase: Mysql

2.1.3 MTA: Mail Server

È formata dai server che operano il servizio SMTP principale di spedizione delle mail dall'interno dell'istituto, i servizi IMAP/POP per la visualizzazione della posta da parte dell'utente.
Il sistema è composto da un cluster di due macchine simili ma con alcune differenze.

Il primo nodo è formato da un server HP DL360G5:

Caratteristiche Hardware:	
Chassis / Processore	1U rack 2 Intel(R) Xeon(R) 5160 a 3.00GHz Dual core
Storage	RAID: RAID 1 Hard disk: HP Smart Array Controller with 73 GB SAS 15000 in RAID 1
Memoria	RAM: 32GB in 8 banchi da 4Gb
Connettività	Broadcom NetXtreme II BCM5708 Gigabit Ethernet 1 DB-9 Serial Port Fiber HBA: Qlogic QLA2340
Dischi Esterni	Unità DAS sterna DOTHILL SANNET II 200 da 1,5 TB FiberChannel in RAID 5

mentre il secondo nodo è formato da un server HP DL360G5:

Caratteristiche Hardware:	
Chassis / Processore	1U rack 2 Intel(R) Xeon(R) 5160 a 3.00GHz Dual core
Storage	RAID: RAID 1 Hard disk: HP Smart Array Controller with 73 GB SAS 15000 in RAID 1
Memoria	RAM: 32GB in 8 banchi da 4Gb

Connettività	Broadcom NetXtreme II BCM5708 Gigabit Ethernet 1 DB-9 Serial Port Fiber HBA: Qlogic QLA2340
Dischi Esterni	Unità DAS sterna PLASMON RAIDTEC CS3120 da 4 TB in RAID 5

La differenza dei sistemi riguarda principalmente il sistema disco esterno, uno SCSI e l'altro FiberChannel di produttori diversi, per risolvere i problemi illustrati in [TOS0403-60] e ridurre le possibilità di perdita di dati a causa di guasti ai sistemi disco e da eventuali errate implementazioni di protocollo.
Il sistema è basato su una distribuzione DEBIAN 5.0 Lenny (Stable) installata con il minimo possibile di software a corredo, protetto tramite un firewall iptables e l'impostazione delle configurazioni di sicurezza utilizzate per tutte le macchine gestite dai Servizi Informatici e Reti necessarie per ridurre al minimo la superficie di attacco.
Per aumentare la velocità di ripristino e la resistenza ad eventuali guasti sono stati utilizzate diverse tecniche:

- è stato configurato il servizio di *heartbeat* che si occupa di controllare lo stato delle macchine spostando l'indirizzo ip ufficiale del server, e aprire le porte dei servizi, tra il nodo primario e quello secondario in caso di problemi. Su entrambi i nodi girano sempre tutti i servizi, SMTP/POP/IMAP, attivi in modo da ridurre le tempistiche di migrazione e permettere in caso di sovraccarico del sistema attivo anche lo smistamento manuale sul nodo di backup di parte degli utenti;

- per aumentare la sicurezza dei dati si è configurato un metadisco DRBD, simile a un raid 1 via IP, utilizzando le due unità disco esterne di modo che ogni nodo possa avere una copia perfettamente sincronizzata dell'archivio delle mail; il sistema DRBD opera anche un insieme di tecniche e analisi per prevenire la perdita di dati dovuti ad eventuali guasti non bloccanti di una delle unità disco;

- è stata prevista una procedura di backup periodica delle mail, che consente in caso di guasto distruttivo di ridurre al minimo la perdita relative;

- l'uso di un sistema MX separato permette di conservare fino a 3 giorni le mail ricevute dall'esterno anche se il sistema interno, magari per manutenzione, non è disponibile.

Per i servizi di posta invece sono stati utilizzati alcuni dei principali programmi Open-Source e largamente usati da molti ISP:

- "*Postfix*" progetto sviluppato da componenti IBM è stato progettato per fornire tutte le funzionalità di "Sendmail" con un altissimo grado di sicurezza e un sistema molto flessibile di "plug-in" per l'aggiunta di filtri e sistemi di processamento delle mail;

- "*Courier*" sviluppato dalla "Double Precision Inc." è un sistema integrato di mail server e groupware che da solo può fornire tutto il necessario per un mail server. Nel passaggio dal vecchio al nuovo sistema si è comunque deciso di utilizzare il "Courier" solo per la parte POP/IMAP e lasciare SMTP/Webmail/Groupware e Mailing List a software più specializzati e performanti;
- "*SASL*" è un componente ponte derivante dal progetto "Cyrus" per l'integrazione dei software sopra esposti con i più diversi sistemi di autenticazione, nel nostro caso il sistema scelto è stato il protocollo LDAP;

- "*Gnarwl*" è un programma per la gestione delle funzionalità di vacation/autoreply che l'utente può configurare tramite l'interfaccia webmail.

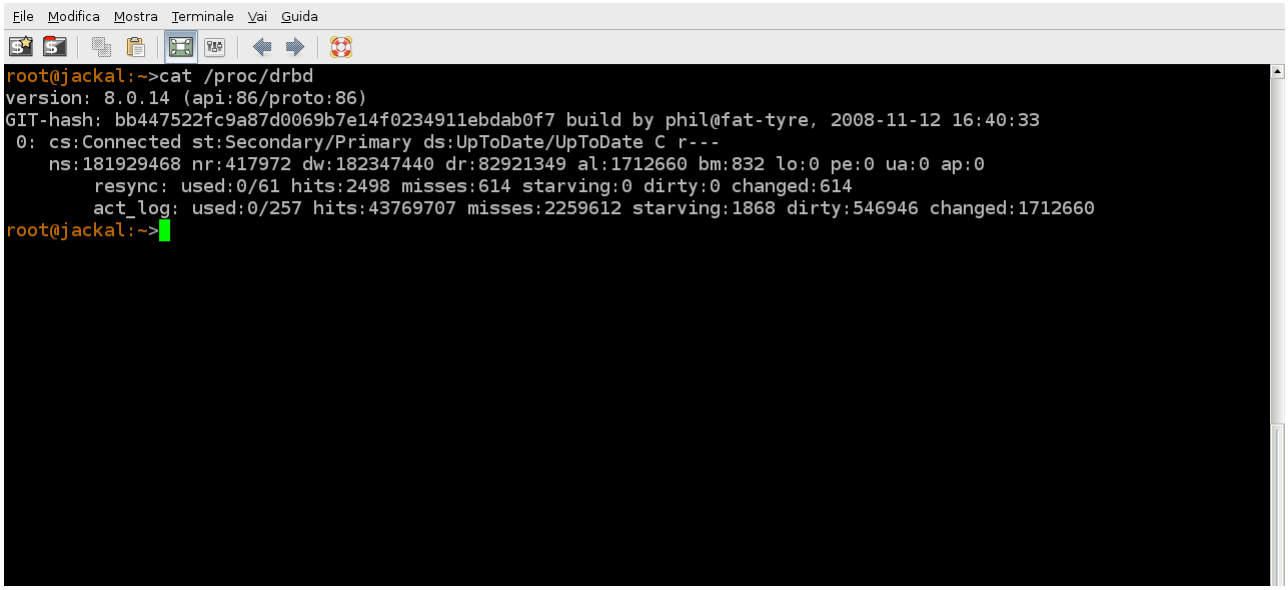

Figura 3.1 Gestione DRBD sul nodo Jackal.

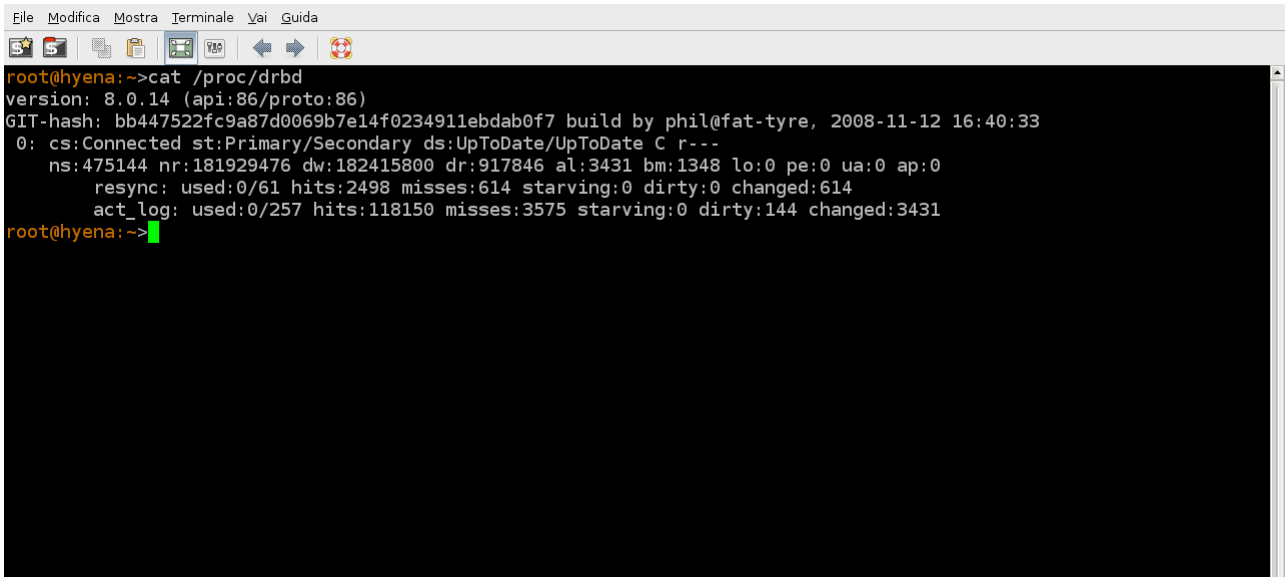

Figura 3.2 Gestione DRBD sul nodo Hyena.

2.1.4 LDAP

Due server si occupano di fornire l'autenticazione degli utenti e le informazioni ad essi associate necessari ai vari servizi. Durante il passaggio dal vecchio sistema di posta si è provveduto a migrare tutti gli account precedenti, comprese le password, tramite l'attivazione di un servizio kerberos interno e l'attivazione di un insieme di controlli di sicurezza sull'inserimento delle nuove password in modo da garantire una certa resistenza agli attacchi a forza bruta.

Figura 4. Applicazione per la gestione degli account.

A differenza di altri servizi che hanno necessitato di una struttura ad alta affidabilità tramite programmi appositi, il servizio di autenticazione utilizza le funzionalità di replica già presenti nei protocolli LDAP e Kerberos per creare la ridondanza su due sistemi; i due server hanno inoltre un sistema di backup crittografato su dvd contenente la copia incrementale giornaliera dei database e i relativi master settimanali per ogni mese. Le copie di backup vengono periodicamente estratte dai server è conservate in una cassaforte di sicurezza posta in un luogo distante dalla sala macchine.
Le macchine che forniscono il servizio sono dei server "DELL PowerEdge 1850 con 2 processori XEON 3.0Ghz dual-core e dischi scsi da 139GB in RAID 1" con S.O. DEBIAN 5.0 Lenny (Stable), software "OpenLDAP" e "MIT-Kerberos".

2.1.5 Mailing List

Il software di gestione delle Mailing List è "Mailman" uno dei software più utilizzati, in quanto fornisce la possibilità di controllare le Mailing List sia tramite comandi via mail, via ssh che tramite interfaccia web; il software fornisce
inoltre funzionalità molto interessanti per quanto riguarda le Mailing List stesse consentendone una gestione ottimale e molto più flessibile di quanto era possibile fare con il sistema basato su alias.

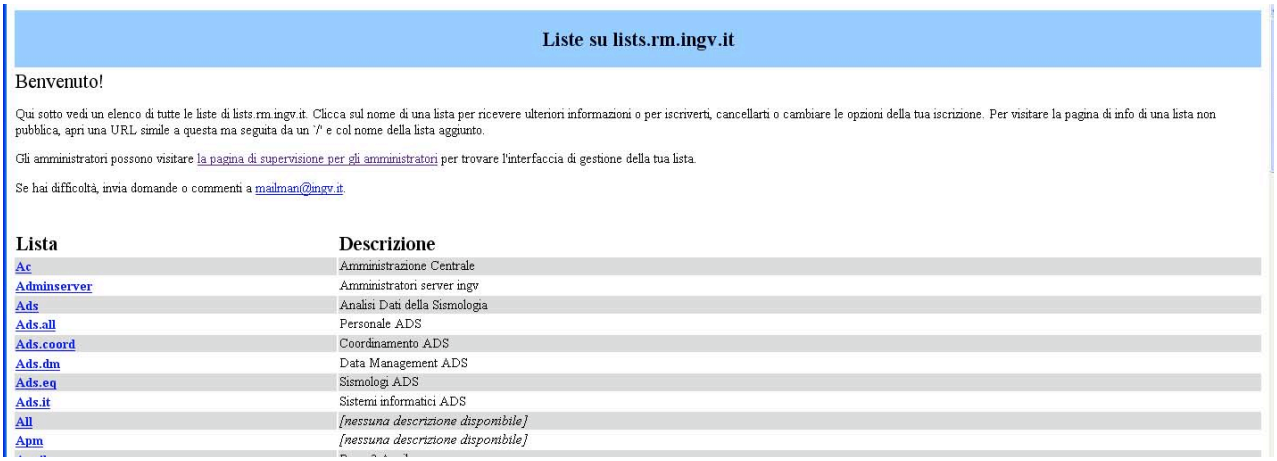

Figura 5. Applicazione per la gestione delle mailing list.

La macchina che eroga, al momento, il servizio è un “DELL Poweredge 1950 con 2 Processori Intel(R) Xeon(R) E5310 1.60GHz, 4 GB di RAM, hard disk da 146GB SCSI in RAID1” ed è in procinto di essere ridondata con una macchina gemella.
Il software installato è un S.O. DEBIAN 5.0 Lenny (Stable) con il software “Mailman” che fornisce il Mailing List manager, “Apache2” che si occupa della pubblicazione dell'interfaccia di amministrazione; “Postfix” demone SMTP utilizzato per lo scambio delle mail tra il Mailing List server e il server di posta “Eternitas”.

2.1.6 MUA: Webmail, Organizer, Client e Mobile

Per l'accesso alla casella di posta da parte degli utenti, oltre ai classici accessi tramite programmi di posta, è stato previsto anche l'attivazione di due sistemi di accesso via web.
Il primo sistema è basato sul framework horde/imp ed è ospitato su un immagine Debian posta all'interno della webfarm dell'istituto. Già nel vecchio sistema era stato adottato un’analoga applicazione per accesso via web ma quello attualmente utilizzato risulta più immediato ed inoltre contiene funzionalità non presenti nella versione precedente.

Figura 6.1 User Agent: Webmail Login.

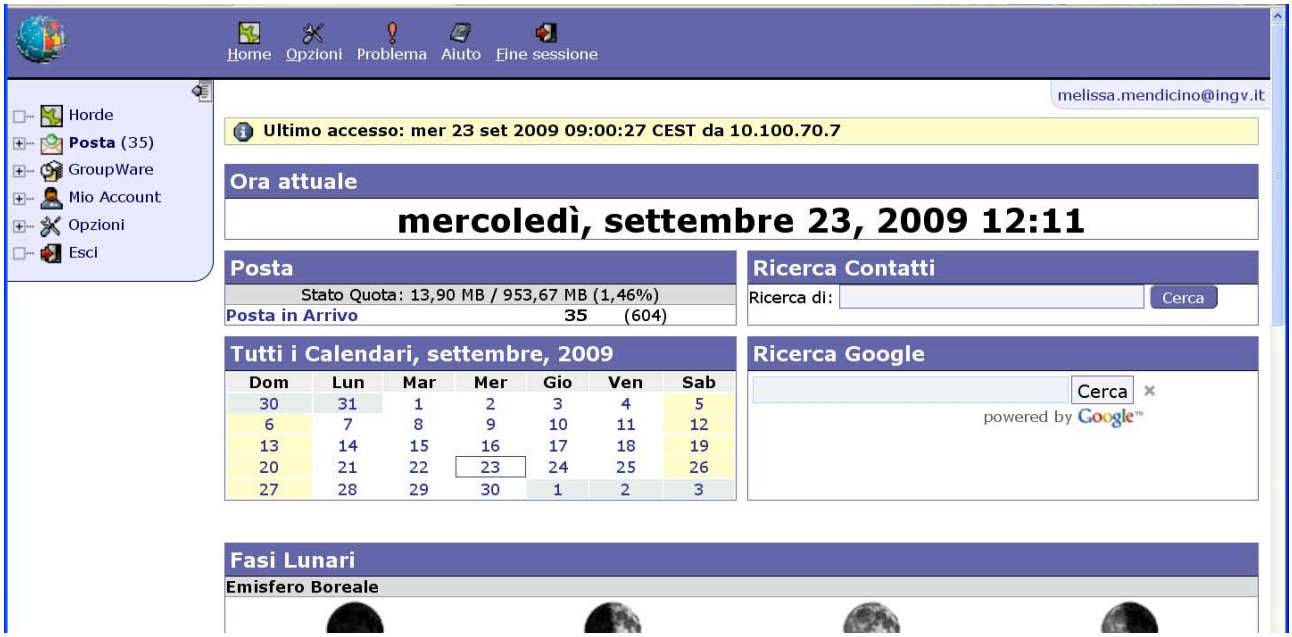

Figura 6.2 User Agent: Webmail Schermata iniziale.

Il secondo sistema è un componente del software enterprise “E-groupware” utilizzato in Istituto che oltre a permettere l’accesso alla propria casella di posta via web, offre anche un servizio di gestione contatti, progetti (GANTT), “to-do list” e molto altre utilità.
Questo software è eseguito sullo stesso server che ospita anche il gestore delle Mailing List.
Di seguito vengono riportate alcune delle immagini più rappresentative del sistema.

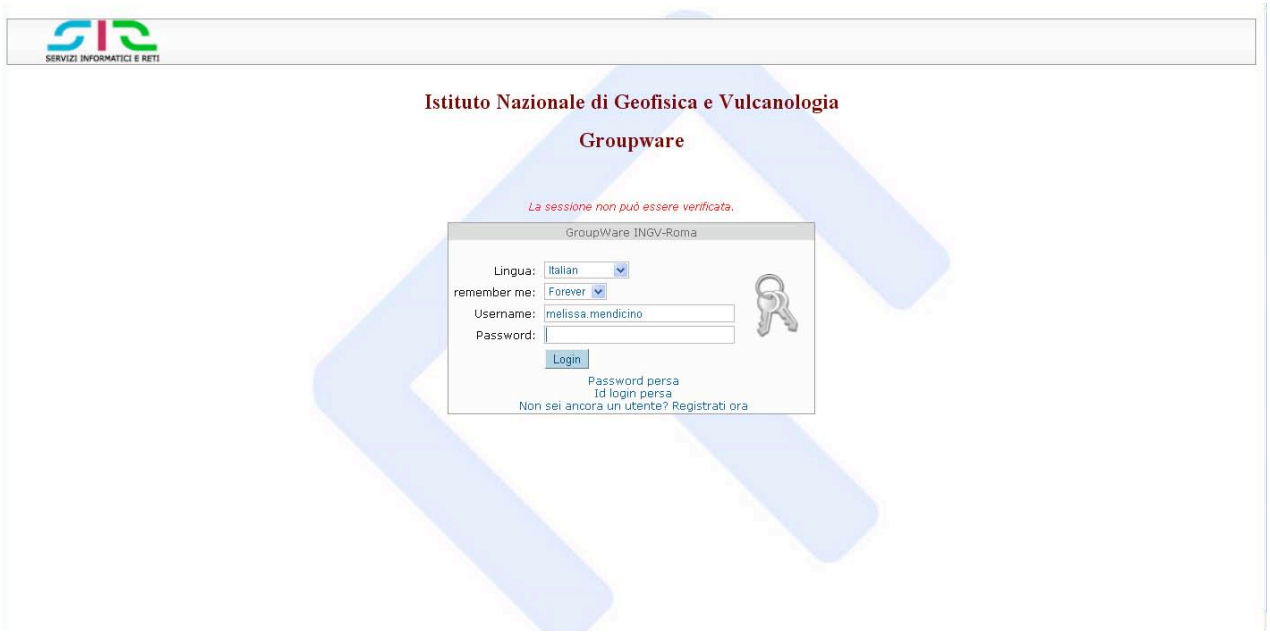

Figura 6.3 User Agent: Organizer Login.

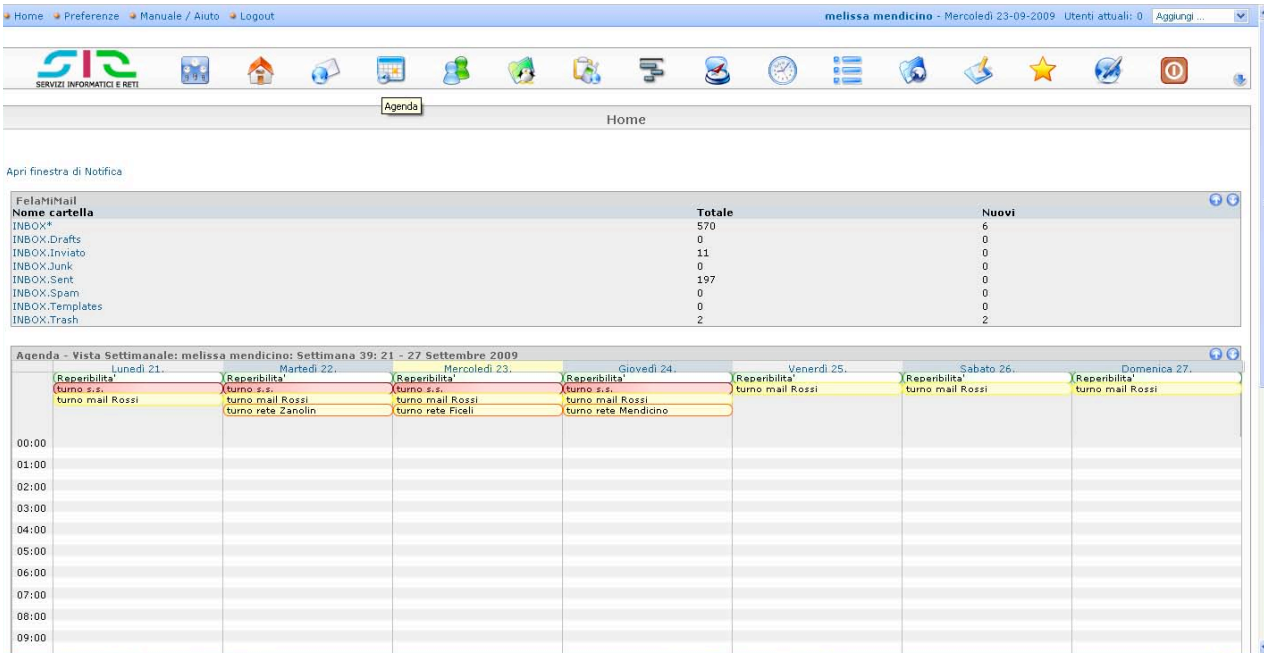

Figura 6.4 User Agent: Organizer Schermata iniziale.

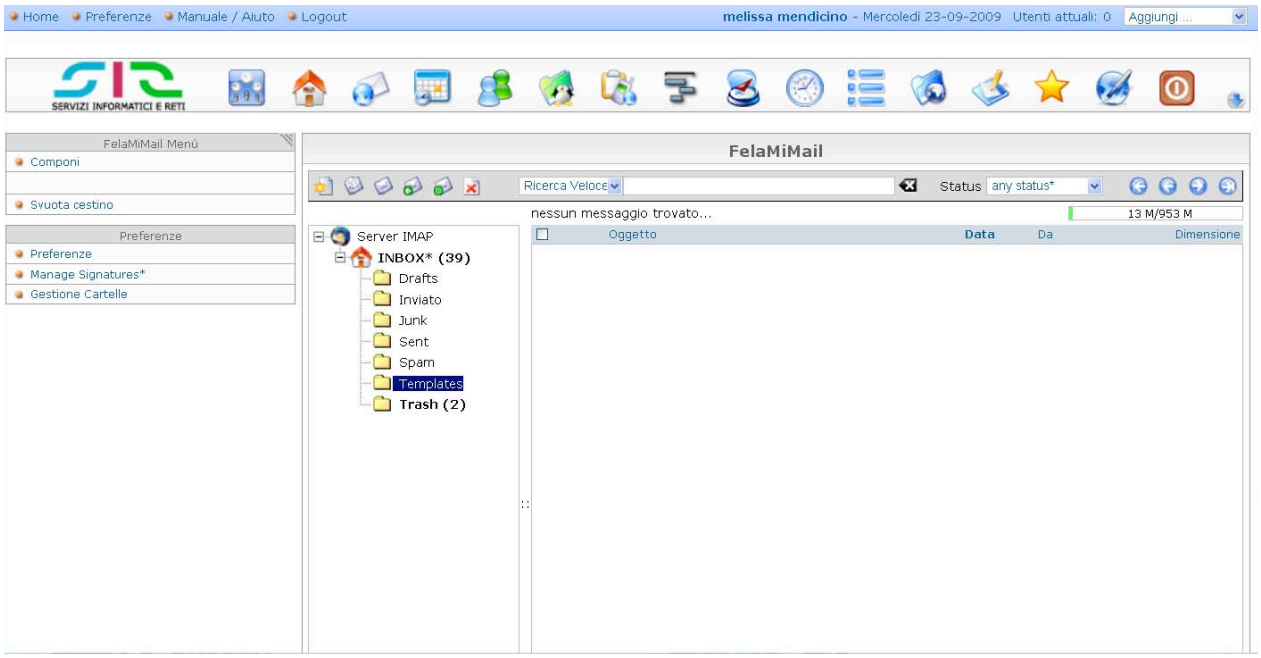

Figura 6.5 User Agent: Organizer FelaMiMail.

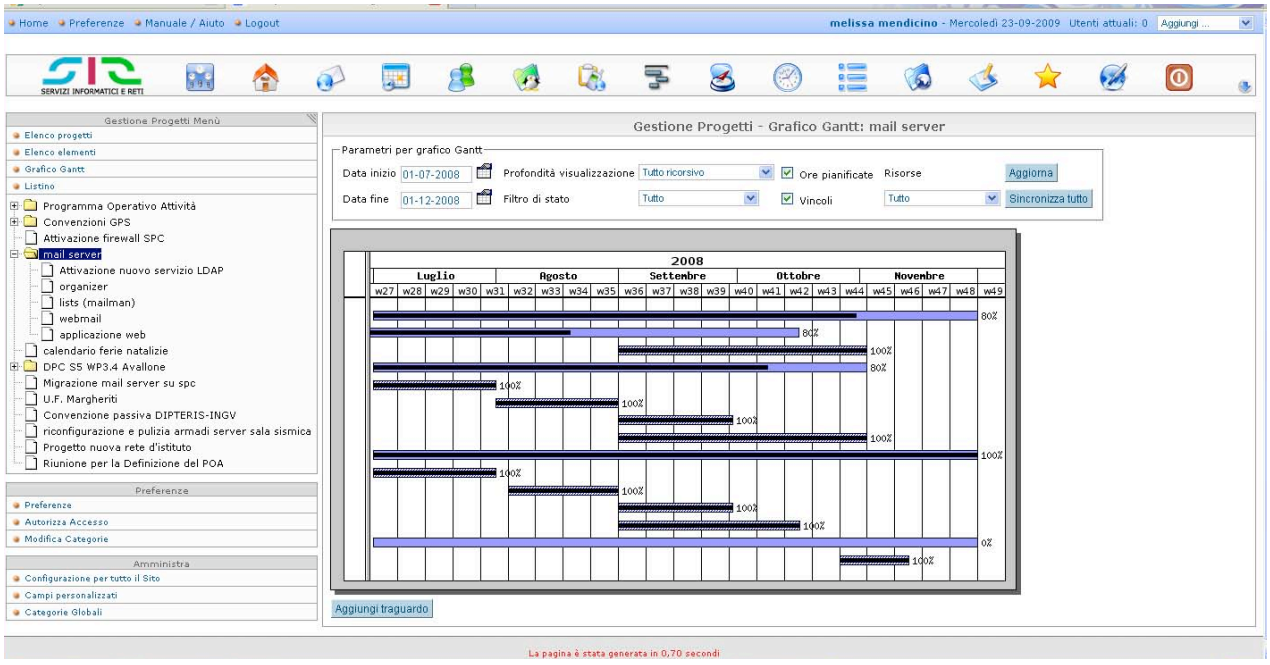

Figura 6.6 User Agent: Organizer Gantt.

2.2 MailControlPanel

Come accennato nel paragrafo 1.3, è stata sviluppata ad hoc un'applicazione dedicata alle operazione di creazione e gestione degli Account e delle Mailing List.
Tale applicazione chiamata *"MailControlPanel"* nasce dall'esigenza degli amministratori, di semplificare e fare convergere in un'unica applicazione tutte le attività collegate alla gestione di un servizio e-mail. Pertanto sarà possibile prendere visione dello stato del sistema, di ogni account e/o mailing list, di modificarne e/o cancellarne le informazioni in modo veloce e senza creare disagi all'operatore il quale non dovrà più interagire direttamente con i singoli server della catena, ma potrà usufruire dell'interfaccia preposta.
Attualmente tale applicazione è in fase di test.

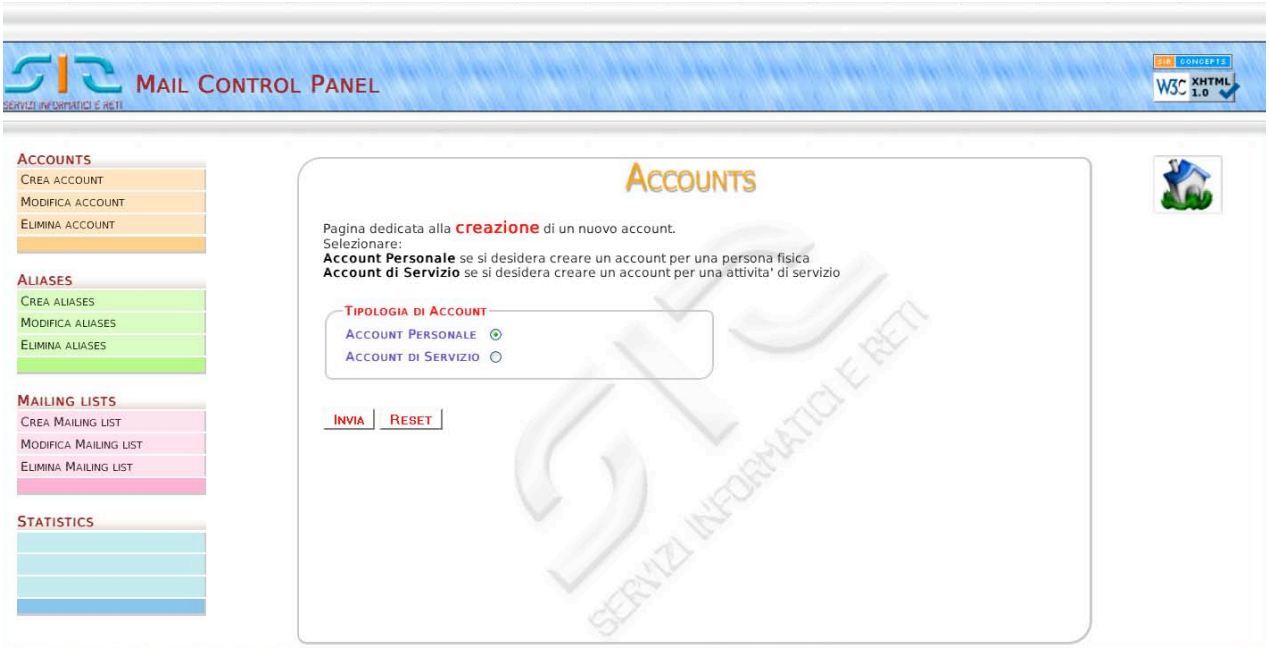

Figura 7.1 MailControlPanel: Creazione di un account.

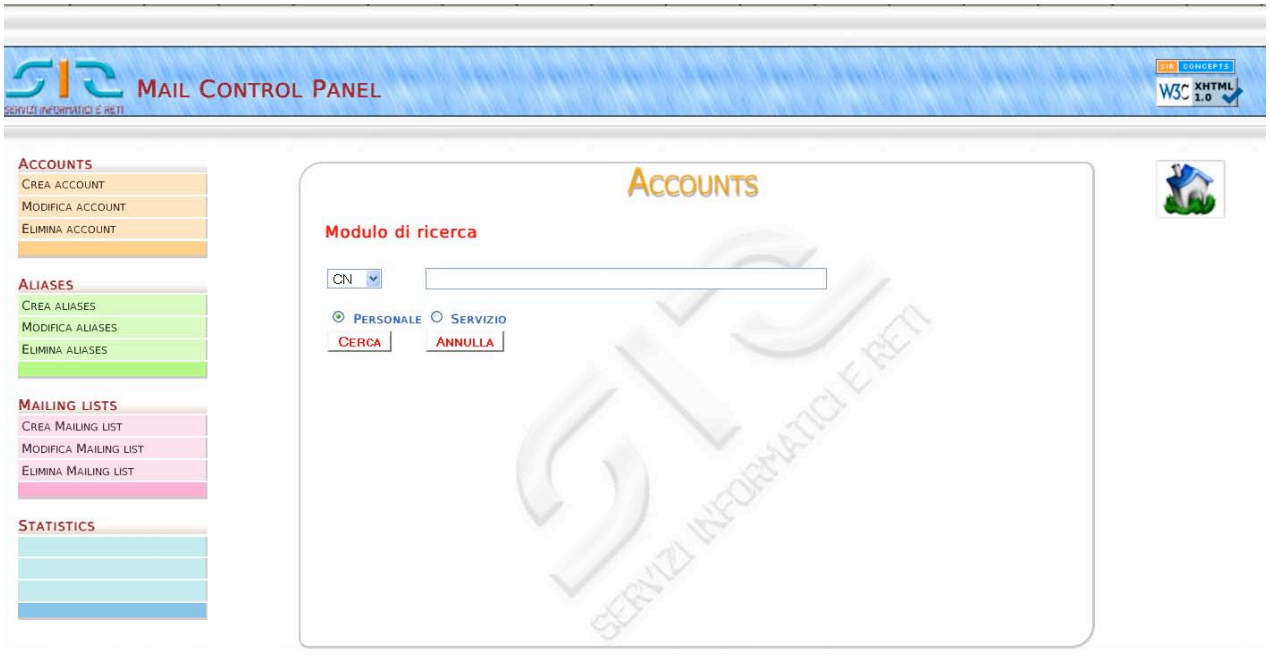

Figura 7.2 MailControlPanel:modifica e/o cancellazione di un account.

2.3 Percorso di una e-mail

Di seguito cercheremo di descrivere come una e-mail viene gestita dal sistema.
Si ponga il caso che un generico utente A che utilizza un indirizzo di posta elettronica esterno debba inviare una e-mail all'indirizzo di posta elettronica di un utente B appartenente al nostro sistema di posta.
L'utente A invia all'indirizzo dell'utente B (generalmente indicato come B@ingv.it) una e-mail che viene consegnata dal server MTA di A al nostro cluster MX (che per semplicità definiremo "Ironport") da una prima analisi esso riconosce che il dominio di gestione è quello dell'Ente ed inizia il processamento dell'e-mail.
Durante l'analisi il sistema verifica se si tratta di una mail conforme alle regole, su di esso impostate, oppure se si tratta di SPAM o mail contenente Virus. Se la mail risulta conforme, viene passata attraverso il protocollo SMTP al cluster MTA (che per semplicità definiremo "Eternitas"), altrimenti viene rigettata dal sistema il quale ne mantiene in archivio una copia per circa 2 settimane e traccia nei log per circa 3 mesi.
Una volta ricevuta da "Eternitas", il sistema invia una richiesta al sistema LDAP (che per semplicità definiremo "Ldap") il quale si occupa di verificare l'esistenza di un account o di una Mailing List, associata all'indirizzo e-mail.
Nel caso esista una corrispondenza, nel DataBase, con un account l'e-mail viene presa in consegna da "Eternitas" il quale la deposita nella maildir dell'utente associato all'account in questione.
Se invece la corrispondenza è con una Mailing List "Eternitas" attraverso, il protocollo SMTP, invia la mail al sistema Mailing List (che per semplicità definiremo "Lists").
"Lists" una volta individuata la lista degli indirizzi che fanno capo all'indirizzo di B, associa una copia della mail ad ogni singolo indirizzo appartenente alla lista e le invia, sempre attraverso il protocollo SMTP, ad "Eternitas". A questo punto "Eternitas" prende in consegna le singole e-mail e le deposita nella maildir corrispondente o le spedisce agli indirizzi esterni in caso ci siano nella mailing list indirizzi non appartenente al mail server.
Nel caso in cui non si verificano nessuno dei precedenti ipotesi la mail viene rifiutata.
L'utente B a secondo della tipologia di MUA che utilizza potrà prendere visione o scaricare la mail che l'utente A gli ha inviato ed eventualmente rispondere alla stessa o scriverne una nuova sempre dietro autenticazione.

Si osservi che all'inizio della descrizione è stato sottolineato il fatto che l'utente A non invia la mail attraverso un account INGV, ma da un indirizzo esterno.
Nel caso in cui entrambi gli utenti utilizzino un indirizzo INGV, l'operazione risulta semplificata, in quanto la mail non viene inizialmente gestita dall' "Ironport" ma viene direttamente gestita da "Eternitas".

3. Problematiche riscontrate

Come per ogni installazione di nuovi sistemi, anche nel nostro caso ci siamo trovati di fronte ad alcune problematiche non direttamente imputabili al sistema, ma strettamente collegate alla necessità di mantenere sempre attivo il servizio in quanto di vitale importanza per la sala di monitoraggio sismico e per il controllo di molti dei servizi offerti, ma anche per la quasi completa assenza di un'organizzazione dei dati aggiornata e conforme all'evoluzione dell'Ente.
Elemento molto importante sia nella scelta delle tecniche che dei programmi impiegati per la creazione del nuovo sistema, è stato infatti riuscire a mantenere una conformità di dati e di strumentazione al fine di permettere un graduale e poco invadente passaggio dal vecchio al nuovo sistema di gestione, che è si più complicato ma allo stesso tempo molto più ricco di funzionalità aggiuntive (vedi paragrafo 1.2).
Di particolarmente importanza è stato il periodo di raccolta dati ed analisi del sistema, che hanno fatto si che il disagio creato agli utenti nella fase di passaggio fosse notevolmente inferiore rispetto a quella che generalmente viene preventivato in casi simili al nostro.
Al fine di permettere agli utenti di mantenere la posta archiviata nel sistema fino al momento del passaggio è stato necessario mantenere attivo per un periodo di 6-8 mesi, il vecchio sistema. Tale sistema può essere raggiunto esclusivamente dalla rete interna dell'istituto e permette solamente di leggere e/o scaricare in locale le e-mail archiviate.
Tale problematica è nata dal fatto che sostituendo alcuni elementi essenziali nell'architettura del servizio non è stato possibile trasferire l'archivio mail rispettando le regole di gestione legate ad ogni singola utenza.
Al termine di questo periodo i sistemi ed l'hardware verranno definitivamente abbandonati ed eventualmente riutilizzati per altre finalità.
Altro vincolo, è stato l'adeguamento dell'indirizzo e-mail al nuovo regolamento [Regolamento Informatico] che prevede l'uso esclusivo del solo indirizzo e-mail formato da nome.cognome@ingv.it nel caso di account personali, maildiservizio@ingv.it nel caso di mail di servizio e mailinglist@ingv.it nel caso di mailing list istituzionali.
Tale adeguamento non potrà essere completato a breve termine in quanto molti degli indirizzi precedentemente utilizzati per identificare gli utenti del nostro istituto, sono stati distribuiti per molti anni ed a molti contatti esterni. Pertanto si è deciso di effettuare un graduale passaggio, che ha avuto inizio con l'introduzione dell'uso dell'account esteso nome.cognome al posto del precedente account basato su parole chiave a volte molto fantasiose. Passo successivo sarà la definitiva eliminazione del vecchio account per l'autenticazione al sistema, seguito da una graduale ma definitiva eliminazione di tutti gli alias, associati all'account nome.cognome, che non hanno senso di esistere se non per motivi strettamente collegati al lavoro svolto presso il nostro Ente. Tale adeguamento pertanto subisce una notevole estensione temporale, che potrebbe aggirarsi anche intorno ai 12 mesi, in quanto di vitale importanza per la comunicazione con l'esterno.

4. Caso studio reale

Al fine di meglio descrivere i "test sul campo" che dimostrano i punti di forza del sistema è stato scelto come caso studio il terremoto dell'aquilano che ha notevolmente stressato il nuovo sistema.
Di seguito verrà riportata integralmente parte della relazione [Prot. N.0005584], redatta dal SIR, inviata dal resp. SIR Dott. Lucio Badiali ai direttori dell'INGV ed al Collegio di Struttura del CNT, e protocollata nel periodo successivo agli eventi catastrofici che hanno colpito l'Aquila nell'Aprile dell'anno corrente.

4.1 Distribuzione e-mail delle ultime tre settimane

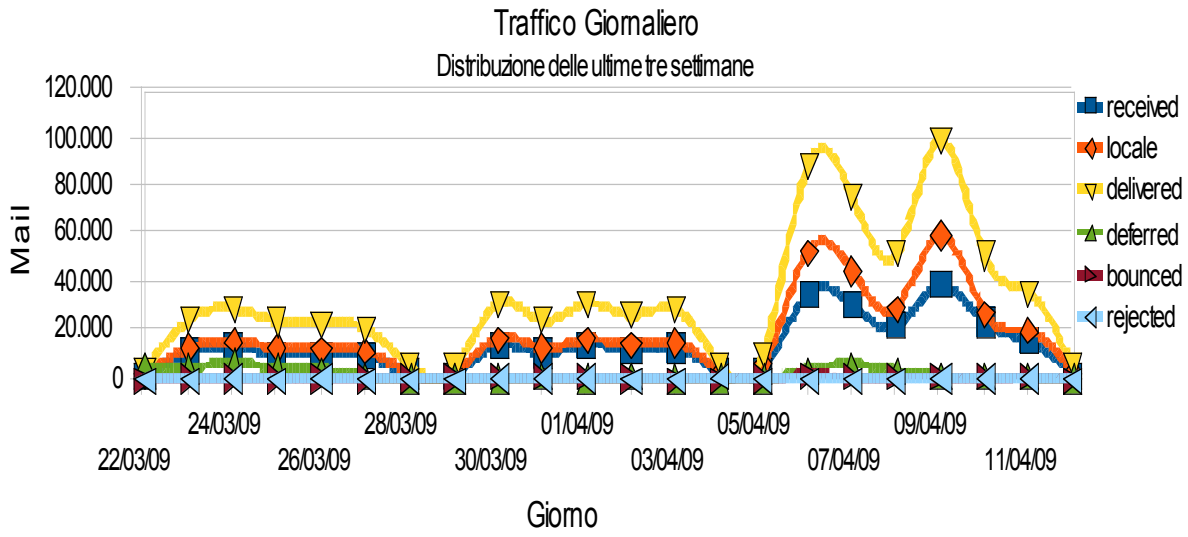

Figura 8.1 Evoluzione temporale del traffico di mail.

Il sistema è in funzione dal 26 febbraio. È possibile individuare un forte incremento di e-mail, concentrato nel periodo tra il 05 aprile ed il 12 aprile c.a. con un incremento rispetto alle due settimane antecedenti, di circa il 300 % (vedi grafici di seguito allegati).

Prendendo in esame la distribuzione oraria delle e-mail durante il primo giorno di allarme sismico (lunedì 6 aprile) e confrontandolo con il lunedì della settimana precedente è possibile notare l'evidente incremento di comunicazioni via e-mail, concentrato negli orari in cui si sono verificate le scosse più violente del sisma.
È interessante notare come il carico nelle singole ore di maggior traffico, cioè nelle ore immediatamente successive alle scosse di maggior magnitudo, gestito dal nuovo sistema sia stato molto superiore a quello che era stato sufficiente a mandare in crisi il precedente sistema di posta il quale accusava problemi durante un carico standard (c'è da notare che era stato progettato in modo non facilmente scalabile, con distribuzioni non del tutto *open* e non pensando ad una Sede quale sarebbe potuta diventare in seguito nel tempo).
Durante tale periodo di crisi, le mail inviate sono generalmente state indirizzate in contemporanea verso più indirizzi e-mail (mailing list: ingvall), con allegati di grosse dimensioni. A differenza della vecchia struttura che entrava in sofferenza e tendeva a bloccarsi non appena anche un solo utente destinatario della mail era in overquota, questo problema non si ri-presenta nel nuovo sistema grazie all'uso di vere Mailing List.
È possibile inoltre osservare, per un ulteriore raffronto, che nella settimana del sisma, la quantità di dati consegnati in locale è stata pari a 33 Gb, ovvero di dimensioni comparabili al completo archivio e-mail presente, al momento della chiusura, sul vecchio sistema. In conclusione, un ultimo cenno va fatto al supporto fornito dal sistema " Antispam ed Antivirus" che ha garantito il flusso di "e-mail legittime" inviate dall'esterno; cosa non solo utile ma necessaria per una gestione efficiente delle comunicazioni elettroniche.

Infatti, dal grafico in figura 4 si può notare che a fronte di una media di circa 33mila e-mail ricevute settimanalmente, di cui circa 7-8mila rigettate in quanto spam, nella settimana di crisi (da 5 al 12 aprile) rispetto al carico ricevuto di circa 66mila e-mail circa 12mila sono state rigettate dal sistema " Antispam ed Antivirus". Virus e spam contribuiscono purtroppo ad occupare e diminuire la banda utile disponibile.

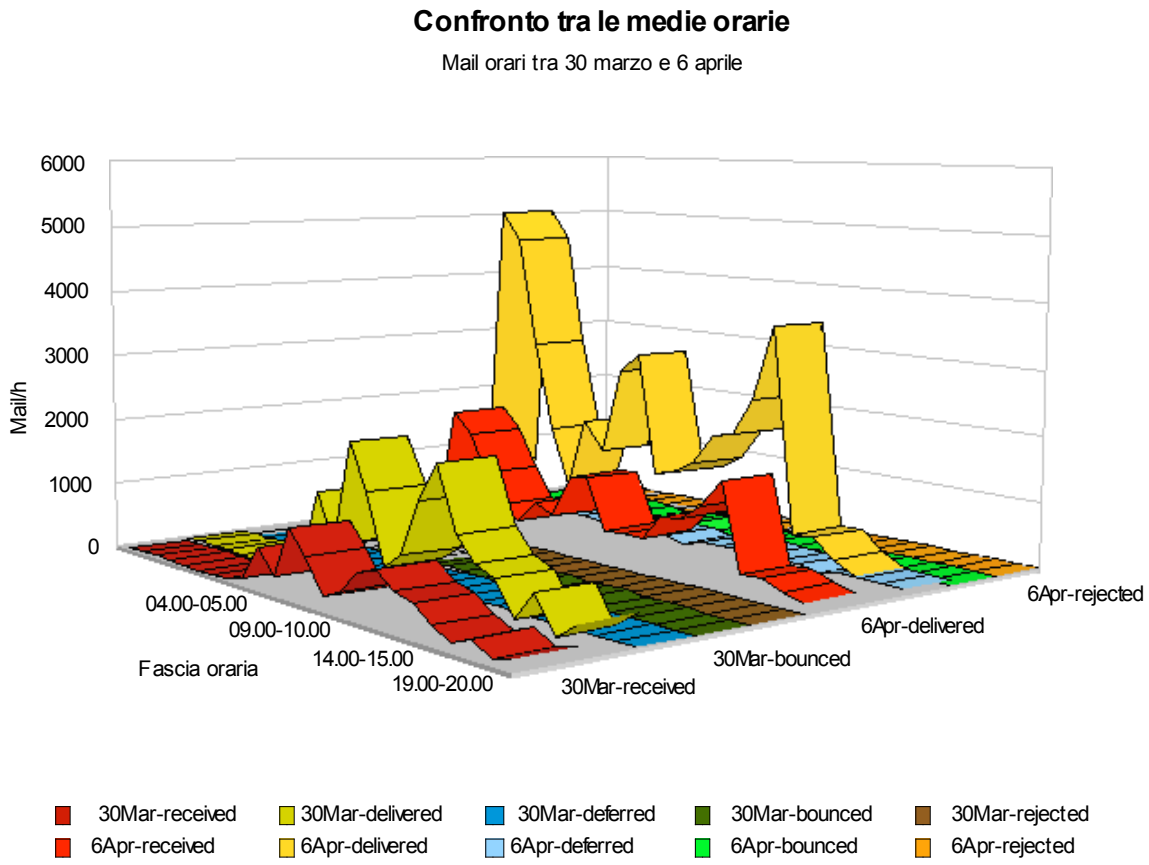

Figura 8.2 Confronto tra le medie orari.

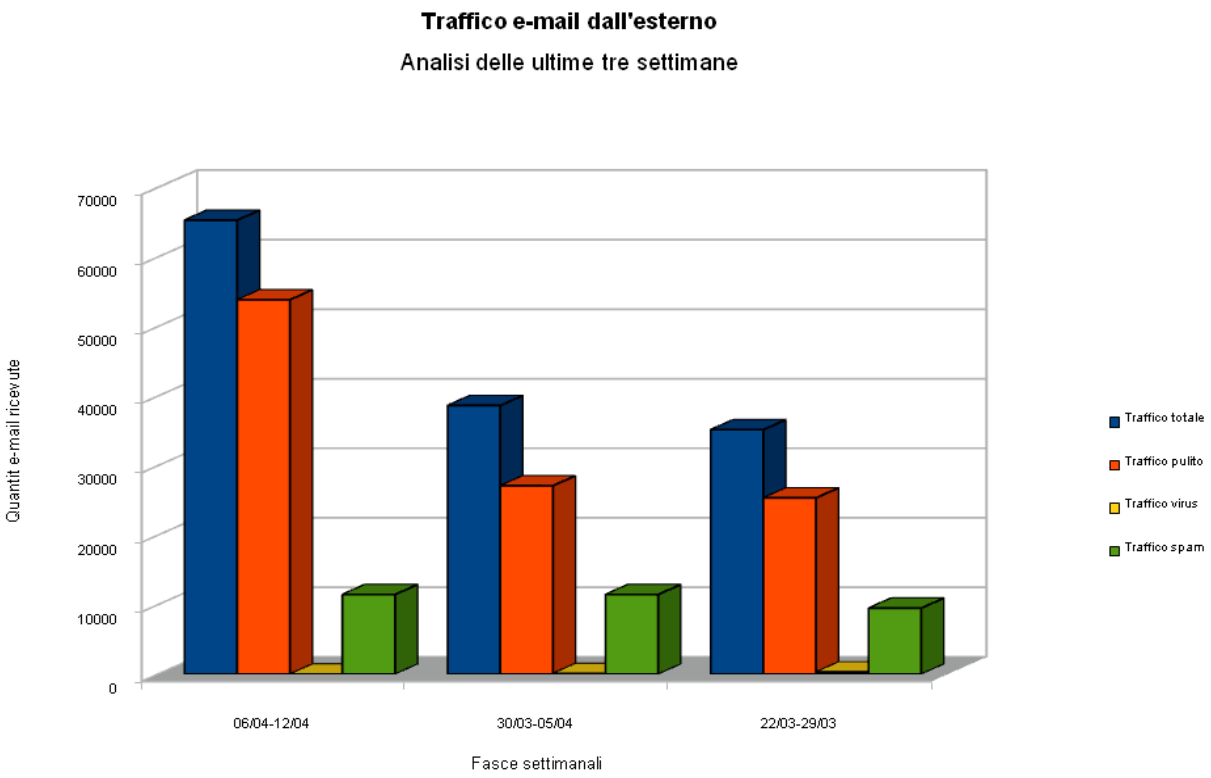

Figura 8.3 Analisi traffico comparato.

Glossario

A.

Account.
Con il termine *Account* si definiscono le credenziali che un amministratore di una risorsa, assegna ad un utilizzatore per consentirgli un accesso riservato.

B.

Browser.
Browser consiste in programma che permette di visualizzare le pagine informative presenti su internet. Fra i più comuni browser ricordiamo Mozilla Firefox ed Internet Explorer.

E.

E-Mail.
Una E-Mail o Mail abbreviazione di Electronic Mail è un messaggio elettronico costituito da un Header e da un corpo, quest'ultimo contiene il testo del messaggio stesso;

G.

Gestore di Mailing List
Un *gestore di Mailing List* è un sistema automatizzato che si occupa di gestire una serie di liste di indirizzi. Ogni e-mail inviata all'indirizzo della mailing list, viene processata dal sistema ed inviata individualmente ad ogni membro iscritto alla lista. Fra le varie mailing list ricordiamo MailMan, Couriermlm, Majordomo, Ezmlm,etc.

I.

IMAP
Un *Internet Message Access Protocol* è un protocollo che si occupa di fornire all'utente l'accesso autorizzato alla propria mailbox, residente su un mail server, senza che l'utente si preoccupi di dover scaricare in locale le mail ricevute.
Attualmente il protocollo è arrivato alla versione 4: IMAP4.
Tale protocollo viene definito per la versione 4 nell' [RFC3501], mentre per la versione 3 nell'[RFC1203].

L.

LDAP
Un *Lightweight Directory Access Protocol* è un protocollo standard si occupa di mantenere le informazioni necessarie all'autenticazione dell'utente verso il server ed alla conservazione delle configurazioni ad esso associate.
LDAPv2 è definito nell'[RFC1777] mentre la versione successiva LDAPv3 è definita nell'[RFC2251].

M.

MailBox
Una *Mailbox* è un file Unix contenente tutte le mail, di un singolo utente, concatenate fra di loro e distinguibili attraverso un Marker.

Maildir
Una *Maildir* è una cartella contenente un file per ogni singola mail ricevuta dall'utente ad esso associato; ogni file viene opportunamente nominato.

Mail Server

Un *Mail Server* è un sistema complesso che permette il trasferimento di E-Mail da un utente sorgente ad un o più utenti destinatari, sia essi sulla stessa rete che distribuiti su internet.

MDA

Un *Mail Delivery Agent* riceve la posta dal *Transport Agent* ad esso associato, la consegna in locale alla *Mailbox / Maildir* del destinatario. Destinatari di un'e-mail possono essere singoli utenti, mailing list, nonché processi/programmi che a secondo della loro natura possono riutilizzarli in vari modi, ad esempio per l'analisi dei LOG.
Fra i vari Delivery Agent ricordiamo procmail e maildrop.

MTA

Un *Mail Transport Agent* si occupa di ricevere la posta da un *User Agent* ed attraverso l'analisi dell'indirizzo destinatario consegnarla all' *MTA* destinatario.
I Transport Agent utilizzano il protocollo SMTP (Simple Mail Transport Protocol) definito nell'[RFC821] oppure il protocollo ESMTP (Extended SMTP) definito negli [RFC1869], [RFC1870], [RFC1891] e [RFC1895].
Esistono molti Transport Agent ma quelli più importanti sono: Postfix, SendMail, Qmail, Exim.

MUA

Un *Mail User Agent* consiste nel client di posta che si occupa della ricezione, trasmissione e organizzazione della posta elettronica di ogni singolo utente.
Esistono varie tipologia di MUA, fra le applicazioni client installati sul sistema operativo dell'utente ricordiamo Mozilla Thunderbird, Eudora Mail, Microsoft Outlook, Evolution, etc.
Frai gli MUA esistono delle applicazioni web che permettono di accedere direttamente attraverso il Browser web utilizzabili su qualsiasi piattaforma senza installare programmi sui computer degli utenti.
Fra di essi ricordiamo Webmail, Squirrelmail, E-groupware, etc.

P.

POP

Un *Post Office Protocol* è il protocollo che si occupa di fornire all'utente l'accesso alla propria casella di posta elettronica da cui visualizzare i messaggi ricevuti oppure scriverne di nuovi, eventualmente scaricandoli direttamente dalla propria mailbox su di un client locale o lasciandole sulla stessa. Gli utenti che utilizzano tale protocollo hanno la possibilità di mantenere una doppia copia (sul server e sul client) delle mail ricevute oppure scaricarle completamente e occuparsi personalmente del loro backup.
Attualmente il protocollo è arrivato alla versione 3: POP3.
Tale protocollo viene definito negli [RFC1734], [RFC1939], [RFC1957] e [RFC3206].

S.

SMTP

Un *Simple Mail Transport Protocol* è il protocollo che si occupa di trasportare, attraverso il protocollo di livello Transport TCP (come i protocolli POP ed IMAP), i messaggi di posta elettronica da un client ad un altro. Fra i vari protocolli SMTP ricordiamo quelli maggiormente utilizzati che sono: Postfix, Courier, Sendmail e Dovecot.
Attualmente tale protocollo viene definito negli [RFC2821] e [RFC2822].

Bibliografia

"Linux, manuale per l'amministratore di sistema, Prima ed.", Evi Nemeth, Arth Snyder, Trent R. Hein. Mondadori, 2008.

"Email Server in Linux", Magnus Back, Patrick Ben Koetter, Ralf Hilderbrandt, Alisteir McDonald, David Rusenko, Carl Taylor. Mc Graw Hill, 2007.

[TOS0403-60]: WEIHANG JIANG, CHONGFENG HU, and YUANYUAN ZHOU, University of Illinois at Urbana Champaign E ARKADY KANEVSKY , Network Appliance Inc. ,"Are Disks the Dominant Contributor for Storage Failures? A Comprehensive Study of Storage Subsystem Failure Characteristics", ACM Transactions on Storage, Vol. 4, No. 3, Article 7, Nov. 2008.

[Regolamento Informatico] "Regolamento Informatico dell'Istituto Nazionale di Geofisica e Vulcanologia, sede di Roma ed sue collegate", Lucio Badiali, Alessandro Piccio, Diego Sorrentino e Francesco Zanolin, 30 Nov. 2006.

[Prot. N.0005584], "La rete telematica della Sede Centrale dell'Istituto Nazionale di Geofisica e Vulcanologia ed i suoi principali servizi informatici durante le prime fasi dell'evento aquilano dell'aprile 2009", 24 Aprile 2009.

[RFC1203]: http://www.ietf.org/rfc/rfc1203
[RFC1734]: http://www.ietf.org/rfc/rfc1734
[RFC1777]: http://www.ietf.org/rfc/rfc1777
[RFC1869]: http://www.ietf.org/rfc/rfc1869
[RFC1870]: http://www.ietf.org/rfc/rfc1870
[RFC1891]: http://www.ietf.org/rfc/rfc1891
[RFC1895]: http://www.ietf.org/rfc/rfc1895
[RFC1939]: http://www.ietf.org/rfc/rfc1939
[RFC1957]: http://www.ietf.org/rfc/rfc1957
[RFC2251]: http://www.ietf.org/rfc/rfc2251
[RFC2821]: http://www.ietf.org/rfc/rfc2821
[RFC2822]: http://www.ietf.org/rfc/rfc2822
[RFC3206]: http://www.ietf.org/rfc/rfc3206
[RFC3251]: http://www.ietf.org/rfc/rfc3251
[RFC821]: http://www.ietf.org/rfc/rfc821
Postfix: http://www.postfix.org
Courier-Imap: http://www.courier-mta.org/imap
Mailman: http://www.gnu.org/software/mailman/
E-groupware: http://www.egroupware.org/
Web-mail: http://www.horde.org

Coordinamento editoriale e impaginazione
Centro Editoriale Nazionale | INGV

Progetto grafico e redazionale
Laboratorio Grafica e Immagini | INGV Roma

Via di Vigna Murata, 605
00143 Roma
Tel. +39 06518601 Fax +39 065041181

http://www.ingv.it

Istituto Nazionale di Geofisica e Vulcanologia