Soluzioni per la posta elettronica in un Ateneo di medie dimensioni

Roberta Cantaroni

Università di Modena e Reggio Emilia

Abstract. Il servizio di posta elettronica all'interno dell'Università degli studi di Modena e Reggio Emilia è attualmente realizzato con una soluzione ibrida e 2 domini separati. Il dominio @unimore.it, dedicato a personale docente e tecnico-amministrativo, collaboratori esterni, strutture e uffici (3200 mailbox e 8000 indirizzi/alias) è gestito interamente con una soluzione in-house realizzata su infrastruttura VMware e software prevalentemente OpenSource. Il dominio @studenti.unimore.it, dedicato a studenti e dottorandi (40.000 mailbox/indirizzi di studenti attivi e alum), è gestito dal 2008 su piattaforma Google Apps Education.

1. Introduzione

L'Università degli Studi di Modena e Reggio Emilia è un ateneo a reti di sedi (Modena e Reggio Emilia) con 14 Dipartimenti, 19 Centri di Servizio e di Ricerca, 7 Direzioni Operative, circa 1.500 unità di personale e 20.000 studenti. Il sistema di posta elettronica è centralizzato dal 2001 e gestito da due tecnici dedicati. Per la gestione degli indirizzi di posta istituzionali è stata adottata una soluzione ibrida che prevede la gestione di 2 domini separati: @unimore.it, dedicato al personale docente, tecnico-amministrativo, collaboratori esterni, strutture e uffici gestito con una soluzione realizzata interamente in-house e @studenti.unimore.it, dedicato a studenti e dottorandi gestito su piattaforma Google Apps Education. Mailbox e indirizzi seguono il ciclo di vita dell'incarico istituzionale di ciascun utente e sono attivati/disattivati/riattivati in tempo reale in base ai dati presenti nel repository LDAP di Ateneo. Gli indirizzi istituzionali sono iscritti a liste di distribuzione per ruolo/struttura/Dipartimento. Le mailbox devono essere usufruibili 24/24h, da qualunque client di posta interno ed esterno, all'istituzione tramite canali sicuri POPS/IMAPS/HTTPS e devono essere il più possibile libere da spam e virus. L'accesso avviene con le credenziali unificate assegnate all'utente per l'accesso a tutti i servizi Unimore. Nel caso del personale è sentita l'esigenza di recuperare agevolmente messaggi cancellati per errore.

2. Il servizio @unimore.it in-house

Per ragioni di sicurezza, privacy e riservatezza dei dati, il servizio è gestito al momento interamente in-house [1]. I nodi di elaborazione sono raggruppati per la maggior parte su un unico cluster di virtualizzazione VMware che ospita quasi tutti i servizi amministrativi e centralizzati dell'Ateneo (autorizzazione/autenticazione, condivisione files, web, posta, etc) e fornisce infrastruttura di elaborazione ai Dipartimenti, alle strutture dell'Ateneo e a strutture esterne convenzionate.

L'infrastruttura sistemistica, concentrata nella sede di Modena, si basa su uno chassis Blade Server con 16 nodi di elaborazione, collegati con tecnologie di datacenter bridging (Fabric Fiber Channel ed Ethernet Fabric) ad un apparato SAN storage multiprotocollo e 4 storage NAS con una capacità totale di 100 Tb, 1 Tb di RAM e 300GHz di risorse CPU. Il servizio di posta elettronica @unimore.it occupa attualmente in totale circa 7 Tb.

I componenti del servizio sono stati separati logicamente per tipologia e realizzati con software prevalentemente OpenSource scegliendo soluzioni che garantiscano alta affidabilità. I log dei vari componenti sono centralizzati mediante rsyslogd.
Il servizio gestisce attualmente 3200 mailbox e 8000 indirizzi/alias (*fig.1*), conta in media 2.000 accessi distinti POPS/IMAPS al giorno, 400.000 messaggi spediti al mese, 50.000 messaggi/giorno da fuori dominio con oltre il 95% di email di spam individuate e bloccate e più di 400 liste di distribuzione e di discussione. Gli indirizzi nominali sono assegnati nella forma nome.cognome@unimore.it

2.1 MX di dominio

Gli MX di dominio sono realizzati con 4 VM su piattaforma VMware, sistema operativo Debian

Workshop GARR 2014 - Next Network: Costruiamo il futuro della rete - Selected Papers. Roma 2-4 dicembre 2014
ISBN: 978-88-905077-5-5

Componente	Hardware	Servizi	Software
MX (Mail eXchanger) di dominio	4 VM	Antivirus/Antispam	Sophos PureMessage
Posta in arrivo	Cluster 2 server fisici + 1 VM	POPS/IMAPS e autenticazione su LDAP	CentOS Cluster Suite, Dovecot, Postfix
Spedizione via SMTP	4 VM	SMTPS con autenticazione su LDAP	Exim4, SpamAssassin, Clamd
Liste di distribuzione per ruolo/struttura	1 VM	Spedizione limitata a indirizzi @unimore.it con controllo d'identità /Aggiornamento giornaliero degli iscritti con LDAP	Sympa
Liste di discussione	1 VM	Amministrazione via web da parte dei proprietari	Mailman
Backup/Restore email	1 VM	Recupero email da interfaccia web	Zimbra
Sistema di monitoraggio	1 VM	Controllo costante dei servizi attivi sui nodi	Munin

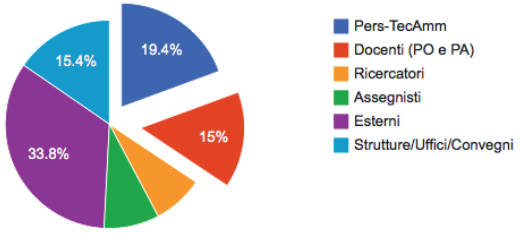

Fig. 1 Distribuzione degli indirizzi @unimore.it

wheezy, 4 Gb RAM, file system di tipo xfs, 80 Gb per software antispam/antivirus e database della quarantena. Il carico è distribuito mediante sistema round robin del DNS con uguale priorità. Gli MX ricevono la posta dall'esterno del dominio e la consegnano al server della posta in arrivo dopo i controlli antivirus e antispam realizzati con software proprietario (Sophos PureMessage [2]) e filtri a livello MTA Postfix (check_recipient_access, reject_non_fqdn_helo_hostname, reject_unknown_sender_domain, reject_non_fqdn_sender, reject_unverified_sender, …).

Ogni VM mantiene una coda locale di email in modo che un eventuale fermo del servizio MTA implichi un ritardo solo nella consegna dei messaggi in quella coda. Per policy le email riconosciute spam (probabilità > 50%) sono bloccate nello spazio di quarantena, condiviso tra i 4 server, per 5 giorni. Ogni utente può autorizzare il mittente richiedendo la consegna in caso di falso positivo, bloccare un mittente "fastidioso" e richiedere la consegna di un digest giornaliero con le intestazioni delle email bloccate.

2.2 Posta in arrivo

Il servizio che gestisce la memorizzazione della posta in arrivo e l'accesso POPS/IMAPS è quello più critico perchè un suo disservizio viene immediatamente avvertito da tutti i client che accedono in quel momento. È stato realizzato su un cluster di tre nodi (2 fisici e 1 VM) con software di clusterizzazione Cluster Suite di CentOS 6, 16 Gb RAM, software OpenSource (*fig.2*). I tre nodi condividono una LUN dedicata su SAN VMware con file system di tipo xfs, 3 Tb di storage per le mailbox (occupazione attuale 55%). Sullo storage è attivo il Tiering automatico (1.63% extreme performance (SSD), 70.02% performance (SAS), 28.35% performance (nSAS). Un solo nodo è attivo in ogni momento ed eroga i servizi di cluster (filesystem, IP a cui corrisponde il record A mail.unimore.it, Mysql, Dovecot, Postfix, httpd). Sul filesystem risiedono i messaggi, il database con mailbox e indirizzi, la coda di posta, gli script per la gestione (creazione/cancellazione/disattivazione) e l'interfaccia web che consente attivazione di vacation/forward e ricerca indirizzi. Ogni mailbox ha una quota personalizzata (default 500 Mb), il formato è Maildir, la dimensione massima dei messaggi (testo + attachment) è di 25 Mb.

L'accesso è realizzato tramite replica locale del sistema LDAP di Ateneo attiva sui 3 nodi. La logica del cluster prevede che i nodi si "sentano" tra-

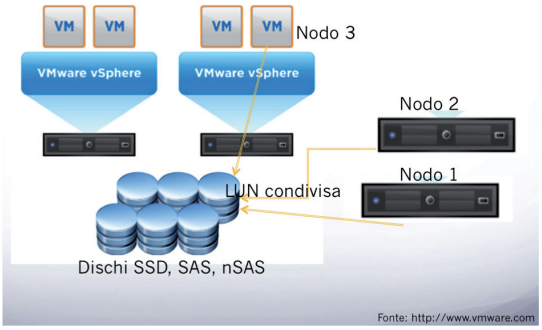

Fig. 2 Architettura in cluster della posta in arrivo

mite dati scritti su quorum disk e che i servizi di cluster migrino su uno degli altri nodi in caso di disservizio su quello attivo. La soluzione adottata è scalabile, consente il riavvio rapido dei servizi in caso di failure su un nodo, consente clone e snapshot periodici sia del nodo realizzato su VM che dello storage e consente la migrazione dei servizi dal nodo attivo ad altro nodo in caso di operazioni di aggiornamento o manutenzione.

2.3 Spedizione via SMTP

Il servizio SMTP permette la spedizione con autenticazione TLS e le stesse credenziali della posta in arrivo. Consegna al server della posta in arrivo nel caso di destinatari @unimore.it ed è abilitato alla consegna all'esterno di Unimore. È realizzato con 4 VM su piattaforma VMware, con sistema operativo Debian wheezy, 2 Gb RAM, 10 Gb storage, file system di tipo xfs. Il nome è unico (smtp.unimore.it) e ad esso corrispondono 4 record A nel DNS, il carico è distribuito mediante sistema round robin del DNS.

Le email spedite sono controllate da antivirus e antispam realizzati con software OpenSource (Clamd, SpamAssassin) e filtri a livello MTA (Exim4) in grado di bloccare l'invio con credenziali compromesse e/o limitare il numero di email spedite all'ora per username. Ogni VM mantiene una coda locale di email, un eventuale fermo del servizio MTA implica un ritardo solo nella consegna dei messaggi in quella coda.

2.4 Liste di distribuzione e di discussione

Le liste di distribuzione per struttura e ruolo sono circa 250 e sono gestite con il software OpenSource Sympa [3]. La spedizione è limitata agli indirizzi @unimore.it mediante controllo di identità. L'elenco degli iscritti è ottenuto con filtri sugli indirizzi degli utenti attivi in LDAP. La consegna dei messaggi alle liste numerose è effettuata utilizzando la modalità hard link di Dovecot (comando dovecot-lda con opzione –p [4]) ottenendo in questo modo un notevole risparmio di spazio disco dato che l'invio ad una lista di N indirizzi equivale alla memorizzazione su disco del singolo messaggio. Sono inoltre attive circa 150 liste di discussione gestite con il software OpenSource Mailman [5]. Gli amministratori delle liste possono accedere via interfaccia web alla configurazione delle liste, alle iscrizioni e alla moderazione dei messaggi.

2.5 Backup e restore

Per gestire il disaster recovery, ma anche per consentire il rapido recupero di messaggi cancellati per errore dagli utenti, sono attivi diversi sistemi paralleli di backup:

- snapshot periodici schedulati sulla SAN (1/gg con retention di 7gg);
- sincronizzazione settimanale del contenuto dello spazio mailbox su storage parallelo mediante rsync e ripristino su richiesta di messaggi o folder;
- copia in tempo reale su altro server di posta (realizzata mediante l'opzione recipient_bcc di Postfix) dei messaggi ricevuti sugli indirizzi istituzionali.

Il server di posta è basato su Zimbra Collaboration OpenSource Edition installato su una VM con Centos 6, 4 Gb RAM, 800 Gb storage e consente agli utenti di accedere via webmail al recupero di messaggi ricevuti e cancellati per errore negli ultimi 3 mesi.

2.6 Sistema di monitoraggio

Su ogni server è installato il software OpenSource Munin, componente "node", che consente di monitorare non solo il carico in termini di cpu, memoria e spazio disco, ma anche la coda di email e di avvisare con warning e alert in caso di superamento delle soglie impostate. La configurazione dei servizi e delle soglie è gestita con Munin, componente "master", installato su una VM dedicata. Il controllo e il riavvio automatico dei processi critici è gestito mediante il software OpenSource Monit.

3. Il servizio @studenti.unimore.it su Google Apps Education

Ogni studente, all'atto dell'immatricolazione, riceve un indirizzo nel formato <ID>@studenti.unimore.it, dove ID è l'identificativo assegnato dalle Segreterie Studenti. I dottorandi ricevono anche l'alias nel formato nome.cognome@unimore.it che possono impostare nel campo From.

L'indirizzo rimane attivo per 3 anni dopo il conseguimento del titolo (stato di alumn).

Dal 2008 il dominio studenti.unimore.it è gestito su piattaforma Google Apps Education. I servizi attivati sono Posta, Calendar, Drive e Chat. L'interfaccia web non presenta banner pubblicitari [6]. L'autenticazione via webmail avviene me-

diante Shibboleth con le credenziali centralizzate Unimore. Per mantenere la tracciabilità dei log, le email ricevute e spedite passano dai server MX e SMTP di Unimore. I limiti attuali sono quelli impostati da Google: spazio "illimitato", dimensione massima dei messaggi (testo + attachment) 25 Mb. Per ogni studente, i dati inviati ai server di Google sono soltanto username, nome, cognome e una password secondaria utilizzata dallo studente per accedere via client di posta. Si richiede che questa password sia mantenuta diversa da quella centralizzata Unimore.

Il servizio gestisce attualmente circa 40.000 mailbox/indirizzi attivi (studenti + alumni) con una media di 15.000 accessi distinti al mese e 500.000 email spedite/mese. Sono attivi 17 gruppi Google sincronizzati giornalmente con LDAP (tutti gli studenti attivi, gli alum, i dottorandi, gli studenti suddivisi per Dipartimento) a cui corrispondono 17 liste di distribuzione utilizzate da personale e uffici per le comunicazioni con gli studenti previo controllo di identità. Le procedure di aggiornamento dei gruppi sono state realizzate con script Ruby e Google Admin SDK.

4. Ciclo di vita di mailbox e indirizzi

Un utente ha diritto all'accesso ai servizi Unimore per tutto il periodo di durata dell'incarico a lui assegnato. Un applicativo denominato "correlatore" (*fig. 3*) si preoccupa di sincronizzare il sistema LDAP di Ateneo con i dati provenienti da 3 distinte banche dati: quella del personale, quella degli studenti e quella dei collaboratori esterni.

Per la gestione del ciclo di vita di mailbox e indirizzi è utilizzato il sistema di messaggistica Apache ActiveMQ. I sistemi POSTA ed LDAP sono entrambi producer e consumer di messaggi. Nel caso degli indirizzi @unimore.it:

- ogni variazione su LDAP (utenti aggiunti, utenti cancellati, account che cambiano utente di riferimento) invia un messaggio ad una coda a cui segue l'azione corrispondente nel sistema POSTA @unimore.it (creazione/riattivazione/disattivazione di mailbox e indirizzi) realizzata mediante script Ruby;
- ogni nuovo indirizzo attivato è segnalato con un messaggio su una coda letta da LDAP che si preoccupa di riempire il campo unimoreMail dell'entry utente.

Nel caso degli indirizzi @studenti.unimore.it:

- ogni variazione su LDAP (studenti aggiunti, studenti cessati, studenti che cambiano organizational unit (ou) cioè che diventano alumn o passano da studente a dottorando o da registered a studente) genera un messaggio ad una coda a cui segue l'azione corrispondente nel sistema POSTA @studenti.unimore.it (creazione/riattivazione/disattivazione) realizzata mediante script Ruby e Google Admin SDK;
- ogni nuovo indirizzo attivato è segnalato con un messaggio su una coda letta da LDAP che si preoccupa di riempire il campo unimoreMail dell'entry studente.

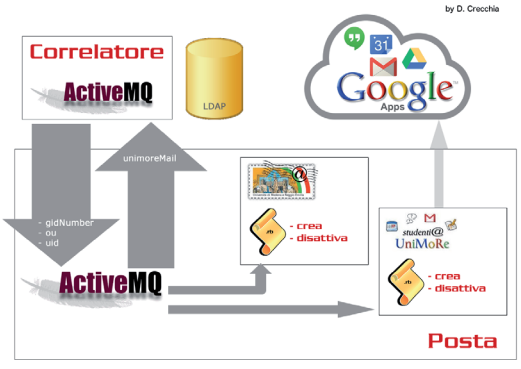

Fig. 3 Gestione del ciclo di vita di mailbox e indirizzi

Riferimenti bibliografici

[[1] Portale dei servizi di posta elettronica @unimore.it http://posta.unimore.it
[2] http://www.sophos.com
[3] Portale liste di distribuzione Unimore http://circolari.unimore.it
[4] http://wiki.dovecot.org/LDA
[5] Portale liste di discussione Unimore http://liste.mail.unimo.it
[6] Portale dei servizi di posta elettronica @studenti.unimore.it http://start.studenti.unimore.it

Roberta Cantaroni
roberta.cantaroni@unimore.it

Laureata in Matematica nel 1987. Lavora dal 1990 come personale tecnico presso i servizi informatici (rete, fonia e sistemi) dell'Università degli studi di Modena e Reggio Emilia. Gestisce il servizio di posta elettronica e l'infrastruttura per il rilascio di PEC e firma digitale.