

GARR-B ed il Progetto di Caching Nazionale

Paola Tentoni

CILEA, Segrate

Abstract

Il CILEA è stato formalmente incaricato della realizzazione del servizio nazionale di caching e mirroring. L'incarico avrà durata di due anni e dovrà portare alla creazione di una infrastruttura di cache www collocate sui 4 punti nodali del backbone nazionale GARR e di un analogo gruppo di 4 server dedicati al servizio di mirror FTP.

Dal 1996 il CILEA partecipa, con finanziamenti propri, ad una iniziativa di caching in ambito nazionale GARR, nata spontaneamente a cura di alcuni enti che sentivano la necessità di fornire tale servizio ai propri utenti, soprattutto in un periodo in cui la banda ancora costituiva un problema cruciale e rendeva impossibile o snervante qualunque tipo di accesso a siti esteri.

Con la ristrutturazione della rete, avvenuta in questi ultimi anni, è stata anche prevista la realizzazione di tutta una serie di servizi quali appunto il caching www, le news, il servizio di CERT, il NOC che aveva senso centralizzare e coordinare in maniera diretta.

Di qui l'uscita del primo bando GARR dedicato alla realizzazione dei servizi di caching e news nella primavera 1999, a cui è seguita una richiesta di revisione all'inizio 2000 per giungere alla sua approvazione durante l'estate scorsa.

La realizzazione del progetto

Il progetto approvato prevede, tra le prime cose, l'attivazione in tempi rapidi di 4 server di tipo "appliance" dedicati al caching. Questi verranno collocati in sede Telecom direttamente connessi alla rete dei router di Trasporto e di Concentrazione dei 4 POP di Milano, Bologna, Roma, Napoli.

La scelta di un dispositivo integrato hardware e software per il caching (black-box) è stata una scelta obbligata, motivata dalla sua più elevata affidabilità rispetto alle tradizionali Workstation general purpose, affidabilità hardware innanzitutto, perché costruita secondo criteri di High availability (dischi hot-

spare, doppia alimentazione,...), indispensabili per una apparecchiatura che deve essere posta in una sede non presidiata, quale il POP presso Telecom. Altra caratteristica è la maggiore affidabilità software perché la componente di sistema operativo (microkernel) è ridottissima e specializzata a svolgere unicamente le funzioni necessarie al caching, il file system è proprietario e lo stack TCP/IP è completamente rivisto per supportare fino a 100.000 connessioni TCP/IP simultanee e persistenti. Questo rende il sistema più efficiente, nonché meno esposto agli attacchi di eventuali hacker.

La configurazione prevista **consentirà a tutti i gestori di cache www in ambito GARR l'utilizzo delle cache centrali più vicine come "Parent"**, in modo che esse consentano di aumentare la capacità di caching del sistema nel complesso, cooperando anche al risparmio globale della banda GARR.

In particolare ogni cache locale (presso le università o gli enti di ricerca) dovrà utilizzare nella sua configurazione la cache centrale presente sul POP cui l'ente è connesso. Non si consiglia per ora di utilizzare più di un parent, anche per non generare inutile traffico di controllo sul backbone.

Contestualmente all'attivazione delle cache nazionali i gestori dei vecchi server di primo livello, tra cui il CILEA, dovranno far migrare i server sottostanti sulle cache ufficiali, per non creare più di due livelli nella caching mesh ed indurre quindi ritardi nella consegna delle pagine all'utente finale.

Il servizio sarà accompagnato dalla creazione di un sito web ad esso dedicato:

<http://cache.garr.it/>

mediante il quale verranno fornite tutte le informazioni utili per i gestori di servizi di caching di secondo livello, nonché per gli utenti finali.

Sono inoltre previste alcune mailing-list tra cui una dedicata ai gestori che vogliono essere informati sull'andamento del servizio e sulle novità relative (info@cache.garr.it), nonché quella dedicata alla richiesta di partecipazione alla caching mesh di generico aiuto, o di segnalazione di malfunzionamenti (cache-service@garr.it).

Verrà chiesta la cooperazione dei gestori locali alla registrazione con form web dei server da loro curati in modo da essere contattabili direttamente per ogni comunicazione utile riguardo il servizio, e per la raccolta eventuale di ulteriori dati statistici sul servizio.

FTP mirror a livello centrale GARR

Fa parte dello stesso progetto di caching Nazionale anche la realizzazione di un servizio di ftp mirror a livello Nazionale.

Allo scopo sono stati acquisiti dal CILEA 4 server Netfinity 4500R dotati di 72Gbyte di area disco ciascuno per un totale di 280G da dedicare al mirroring.

Il sistema di Mirror ha come obiettivo ridurre i tempi di reperimento per oggetti (es. software public domain) particolarmente interessanti e ricorrenti nella comunità scientifica, in modo da garantire, più di quanto non possa fare un sistema di caching, tempi di accesso contenuti per il recupero dei dati prescelti, e risparmiare banda internazionale. Ovviamente si tratta di un servizio che per sua natura manca di trasparenza, ovvero deve essere pubblicizzato ed indicizzato opportunamente perché possa diffusamente essere utilizzato.

Proprio per questo motivo verrà allestito un servizio web dedicato, che fornisca sia un accesso guidato per argomento alle risorse presenti sul mirror, sia consenta la ricerca sui filename (<http://mirror.garr.it/ftpsearch.html>).

Un altro software in grado di indicizzare siti ftp e di indicare la presenza del pacchetto voluto (a partire dall'URL) su un certo server è quello dell'FTP Mirror Tracker, un sistema basato su database MySQL che è stato oggetto di un progetto Terena e che verrà sperimentato anche sui nostri server. Al momento i 4 server sono in fase di predisposizione presso il CILEA e sono iniziati i primi mirror dei pacchetti public domain più conosciuti e richiesti, tra cui:

AcrobatReader (1615 Mbyte), W3C(400 MByte)

BIND (67 MByte), Apache (299 MByte)

CPAN (777 MByte), Bind (67 MByte)

CTAN (3416 MByte), Exim (5 MByte),

Eudora (149 MByte), FreeBSD (4469 MByte)

Freeamp (347 MByte) , GnuPG (59 MByte)

Gnuftp (1912 MByte) ,GnuPG-WWW (4 MByte)

Kermit (946 MByte) , KDE (2530 MByte)

Kernel.org(7837MByte),Mandrake(3306 MByte)

LPD (232 MByte) , MySQL (987 MByte)

Netscape(2944 MByte),OpenBSD (7689 MByte)

Postgres (273 MByte) , OpenLDAP (42 MByte)

OpenSSL (54 MByte) , Samba (240 MByte)

PGP (433 MByte) , Sendmail (41 MByte)

RedHat(4320 MByte), Simtel.net(15157 MByte)

Slackware (7373 MByte), Squid (34 MByte).

L'attivazione ufficiale del servizio di mirroring sarà preannunciata dalla convocazione di una serie di seminari, oltre che da comunicazione su una mailing list aperta, cui chiunque vorrà avere informazioni sul servizio potrà sin d'ora iscriversi:

info@mirror.garr.it

Segnalazioni sui malfunzionamenti saranno possibili all'indirizzo:

mirror-service@garr.it

Richieste di nuovi inserimenti sul mirror sono possibili o di qualunque tipo di miglioramento del servizio possono essere indirizzate a:

mirror-service@garr.it

Inoltre il servizio nella sua forma guidata sarà raggiungibile a partire dall'URL:

<http://mirror.garr.it/>