

# Una webfarm per AZALAI

**Matteo Boschini, Paola Rossi**

*CILEA, Segrate*

## *Abstract*

Il software Azalai, il MetaOPAC realizzato dal CILEA, ha trovato varie applicazioni: dal MAI, il MetaOPAC Azalai Italiano, ai sistemi MetaOPAC specializzati di settore, come MAIA, per l'architettura, e MAL, MetaOPAC della Regione Lombardia per le biblioteche pubbliche. Questi servizi hanno visto crescere l'utenza nel corso di questi ultimi anni, come pure le applicazioni create sono per loro natura in continua crescita in quanto collegano un numero sempre crescente di sistemi OPAC. Si è resa così evidente la crescita costante delle risorse necessarie per questi sistemi. Una soluzione adatta che è stata recentemente messa in opera è la webfarm.

*Keywords:* Telematica, Beni culturali, Webfarm, Biblioteche, Biblioteca virtuale, Catalogo, OPAC.

Il software AZALAI[4,5], realizzato dal CILEA, viene utilizzato ormai da alcuni anni in varie applicazioni. Lo scopo del software è fornire uno strumento per ricerche bibliografiche cumulative.

La funzione svolta, appunto la "metaricerca", è la ricerca di un documento bibliografico svolta dal sistema meta interrogando via web vari sistemi OPAC (online public access catalog) remoti. "AZALAI" è una parola araba che significa "separare per poi ricongiungere", una visione per via metaforica del funzionamento del sistema MetaOPAC.

Il software AZALAI, infatti, invia diverse interrogazioni, costruite nella forma di URL attivate in rete, verso i vari server ("separa"), a fronte dell'interrogazione bibliografica richiesta dall'utente, e restituisce a questi le risposte che via via i server remoti dei sistemi OPAC producono ("raccolge riunendo").

Il servizio MetaOPAC ha visto la nascita di alcune applicazioni di questo software che forniscono servizi in rete, mentre alcune altre stanno nascendo e sono in fase prototipale di allestimento. In questo articolo presentiamo le problematiche relative alle risorse di questo servizio in espansione, illustrando i principali fattori da cui dipende la stima delle risorse necessarie, e la soluzione che si è pensato di adottare: la webfarm.

## **Le applicazioni realizzate con AZALAI**

Dal maggio 1999 è funzionante il sistema MAI - Meta OPAC Azalai Italiano, frutto della collaborazione tra CILEA e AIB (Associazione Italiana Biblioteche) [1]. L'accesso al sistema è disponibile all'URL:

<http://www.aib.it/aib/lis/opac1.htm>

Le caratteristiche generali del MAI sono già state presentate nel Bollettino CILEA n. 68, 1999 [7]. Ulteriori ragguagli sui successivi sviluppi sono stati presentati nel n. 79, 2001 [2]. Il gruppo di lavoro può essere contattato via mail all'indirizzo [aw-opac@aib.it](mailto:aw-opac@aib.it).

L'applicazione MAI, la maggiore delle applicazioni finora realizzate, ha come obiettivo raggiungere il maggior numero di OPAC Italiani.

Accanto a questa realtà si sono create altre possibilità, applicate a settori specialistici diversi[6], che stanno prendendo piede in quanto rendono possibile la ricerca bibliografica soltanto sul materiale riguardante una certa disciplina. Rispetto al sistema nazionale, questi sistemi danno in più la possibilità, offerta dal software, di preselezionare via HTTP il materiale dell'OPAC, cioè possono configurare nell'URL gli opportuni parametri di ricerca. Piuttosto che la ricerca ampia, "a tappeto", che risponde all'esigenza primaria di verificare l'esistenza di un documento in Italia, si effettuano così ricerche in ambito delimitato che risultano con meno rumore, impossibili ad ottenersi a livello generale.

Sono attivi due servizi, applicati nell'ambito dell'architettura e per le biblioteche pubbliche della Regione Lombardia. Altri settori, tra cui quello aerospaziale e della slavistica, sono già in via di allestimento, ed altri ancora si vanno proponendo per realizzare questi progetti entro l'anno.

### **La crescita delle connessioni agli OPAC**

E' il principale fattore di crescita che ha effetto sulle risorse del sistema, assieme alla crescita dell'utenza.

MAI ha visto crescere fino all'attuale valore di 166 i sistemi OPAC connessi, con un tasso di crescita annuo del 15% circa.

L'attivazione dei prototipi iniziali specializzati è favorita da quanto il nazionale ha già realizzato, ed è possibile avviare prototipi in tempi relativamente brevi. Il numero iniziale delle connessioni può essere quindi relativamente alto, mentre l'aspettativa di crescita *a regime* di queste applicazioni dovrebbe essere confrontabile con la crescita nazionale. L'applicazione MAIA, settore dell'architettura, cioè la prima realizzata, ha ormai raggiunto le 60 connessioni.

### **La crescita dell'utenza**

Nel 2002, rispetto all'anno precedente, si è riscontrato l'aumento del 59% dell'utenza dei servizi MetaOPAC realizzati con AZALAI. La tendenza positiva sembra mantenersi, anche se in misura minore, anche in questi primi mesi del 2003.

### **L'ampliamento del software**

Alcuni moduli di AZALAI che riguardano la produzione delle URL sono soggetti ad ampliamento, proprio in base alle caratteristiche del progetto stesso (perchè nascono nuovi software OPAC in uso in Italia, oppure perchè nuove versioni di alcuni dei software esistenti, già connessi, richiedono l'ampliamento di AZALAI). Questo ampliamento però non incide in modo considerevole nelle risorse necessarie per eseguire le applicazioni MetaOPAC, anzi può essere senz'altro trascurato rispetto agli altri fattori.

Il motore general purpose del software, che presiede alle attivazioni in rete, è rimasto invece pressochè costante rispetto all'architettura iniziale. L'unica eccezione è la connessione al software Aleph500, che richiede maggiori risorse, ma la sperimentazione di questa nuova connessione si è appena conclusa ed il numero attuale delle connessioni è ancora così limitato da non riscontrare effetti sul

servizio. Non siamo quindi in grado di fare previsioni, ma occorre tenere conto che questi OPAC sono piuttosto diffusi, ed il fattore è perciò da considerarsi presente in prospettiva.

### **Le caratteristiche tecniche attuali del servizio**

L'applicazione richiede relativamente molte risorse (un'esecuzione può superare 40 MB di memoria virtuale), in particolare perchè compie sì minime elaborazioni, ma su testo, e per esigenze funzionali necessita di ricordarlo, anche se solo temporaneamente. Utilizza perciò in modo inevitabilmente un po' "disordinato" molta memoria. Alla crescita del numero di sistemi connessi si è evidenziata così una crescita più che lineare delle risorse complessive necessarie, come era appunto prevedibile.

D'altro canto, l'applicazione è intrinsecamente problematica. "Fisiologicamente" la durata di un'esecuzione non può essere ridotta al di sotto di un certo tempo (50 secondi), un tempo che è quasi incompatibile con la durata consigliata di un'esecuzione dello stesso tipo, mentre, se esistessero, limitazioni che fossero introdotte successivamente all'avvio di una esecuzione ovviamente invaliderebbero la funzionalità stessa che si intende fornire (una ricerca parziale invece che globale è assolutamente inaccettabile). Inoltre, la frammentazione in tranche non è gradita all'utenza nè tanto meno risolve il problema dal punto di vista complessivo delle risorse richieste (l'utente è interessato a consultare l'intero esito, non si limita a consultare il primo recupero positivo). Meccanismi di rilascio della memoria non sono risultati compatibili con l'applicazione, perchè sono indiscriminati e hanno perciò effetto sulla corretta esecuzione.

Al crescere delle richieste degli ultimi mesi, si sono inoltre andati evidenziando momenti di rallentamento, soprattutto nei momenti di punta quotidiani, mentre la ripresa, dopo il sovraccarico, sembrava stentare. Come intervento correttivo per questi momenti eccezionali, è stata utile la delimitazione del numero massimo di esecuzioni simultanee che gli utenti possono in un momento richiedere. Questo correttivo è accettabile per un sovraccarico momentaneo, ma le prospettive d'uso del servizio tendono a trasformare il sovraccarico momentaneo in una condizione permanente, se a parità di risorse.

## La soluzione “webfarm”

Si è deciso di adottare una tecnologia di tipo “webfarm”, ormai consolidata ed ampiamente usata dalla comunità Internet. La webfarm sembra indicare una via ottimale in questa situazione di forte pressione della richiesta. Essa mantiene ad una crescita lineare la disponibilità delle risorse necessarie quando crescono le richieste: così permette di affrontare la prospettiva della crescita di questo servizio.

La webfarm è costituita da un insieme di server che ricevono alternativamente le varie richieste attraverso un meccanismo di smistamento basata su “DNS Round Robin” [3].

La webfarm realizzata si appoggia a due server, differenti come configurazione (hardware e software). Il meccanismo di smistamento delle richieste verso i vari server rispecchia, attualmente, la proporzione tra le diverse prestazioni dei due sistemi della farm<sup>1</sup>.

Inoltre, sono state effettuate prove di estensione della farm ad un numero maggiore di server, tra l'altro con altri sistemi operativi, che hanno dato esito positivo. Questa è un'ulteriore conferma della flessibilità, facile estendibilità e scalabilità offerte dalla soluzione tecnologica.

Per contro, oltre ai vantaggi già menzionati, la webfarm comporta un aggravio per la maggiore attenzione nell'aggiornamento dei sistemi, cioè l'allineamento tra i server nella farm. Nel nostro caso specifico, l'aggiornamento del software e dei dati di connessione (dati prodotti altrove ma che sono *locali* al software AZALAI), ha un ritmo non del tutto trascurabile.

D'altro canto le realizzazioni che utilizzano lo stesso software AZALAI come abbiamo visto sono numerose ma sono, soprattutto, “diverse”. Dal punto di vista del software, lo condividono, mentre i dati sono diversi per le diverse applicazioni nazionale e specializzate, e oggi l'utilizzo per l'applicazione nazionale è molto maggiore rispetto alle specializzate. Le funzionalità MAI di ricerca, la ricerca globale, selettiva e ricerca per regioni, inoltre, possono essere gestite come indipendenti.

La realizzazione attuale della webfarm ha visto oggi una equidistribuzione del carico delle metaricerche richieste, ma in futuro lo sviluppo dei vari servizi potrebbe prevedere una distribuzione in parte diversa sui server della webfarm, che concorresse a ridurre le attività di gestione dell'aggiornamento o meglio

incontrasse le diverse esigenze di risorse delle varie funzionalità.

## Bibliografia

- [1] A. De Robbio, R. Ridi, F. Valenziano, “AIB99”, XLV Congresso nazionale AIB, Roma, 16-19 maggio 1999, URL: <http://www.aib.it/aib/congr/co99metaopac.htm>
- [2] A. De Robbio, P. Rossi, “MAI - MetaOPAC Azalai”, Bollettino del CILEA, n. 79, 2001, URL: <http://cdl.cilea.it/cdlnew/NRiviste/dett.asp?ID=cils12001b79p19-26&H=22o38>
- [3] R. S. Engelschall, “Load Balancing Your Web Site”, Web Techniques Magazine, n. 3, 1998, URL: <http://freebsd.ntu.edu.tw/doc/webfarm>
- [4] L. Guglielmi, F. Valenziano, “Azalai : il MetaOPAC del CILEA”, Bollettino del CILEA, n. 57, 1997, URL: <http://www.cilea.it/bollettino/b57/p13-14.pdf>
- [5] P. L. Miglioli, P. Rossi, “Azalai: un motore per la ricerca bibliografica”, NIR-IT: "Didattica, società, cultura, mercato: le nuove frontiere di Internet", Milano, 13-15 gennaio 1998, URL: <http://www.cilea.it/GARR-NIR/Nir-it98/atti/miglioli.pdf>
- [6] P. Rossi, “MAI: MetaOPAC Azalai Italiano e metaOPAC specializzati”, AIDA informazioni, n. 2-3, 2002, URL: <http://www.aidainformazioni.it/2002/232002.html>
- [7] P. Rossi, F. Valenti, “MAI-MetaOPAC Azalai Italiano”, Bollettino del CILEA, n. 68, 1999, URL: [http://www.cilea.it/bollettino/b68/Rossi\\_Valenti6\\_10.pdf](http://www.cilea.it/bollettino/b68/Rossi_Valenti6_10.pdf)

<sup>1</sup> Weighted DNS Round Robin