

PostgreSQL, un database libero per la Pubblica Amministrazione Italiana

Gabriele Bartolini
Comune di Prato / Italian
PostgreSQL Users Group
Piazza del Comune, 2
59100 Prato PO - Italia
+39 0574 1835214

g.bartolini@comune.prato.it

SOMMARIO

PostgreSQL è unanimemente considerato il **sistema open-source per la gestione di database più avanzato al mondo** (da non confondersi con il più diffuso).

Il titolo gli è valso in seguito alle funzionalità avanzate e al tempo stesso mature che il prodotto ha raggiunto negli anni, fra cui, per citarne solo alcune: viste, schemi, trigger, tablespace e stored procedure.

Deve far riflettere il fatto che oggi **PostgreSQL è maturo e pronto per la maggior parte, se non addirittura la totalità, delle esigenze di un medio comune italiano** (si ricorda che l'85% dei comuni italiani ha un numero di residenti inferiore a 10 mila unità).

Tuttavia, il suo impiego viene spesso ignorato, soprattutto all'interno della pubblica amministrazione. Uno dei possibili motivi è la **mancanza di un supporto professionale** adeguato e diffuso tale da agevolare l'adozione.

Così come per gli altri settori del software libero, un intervento della pubblica amministrazione in pieno spirito di comunità potrebbe inoltre favorire la **generazione di una forma imprenditoriale dinamica** e distrettuale su tutto il territorio italiano.

È inoltre da sottolineare come, contestualmente alla crescita tecnologica del software, anche la comunità globale di PostgreSQL stia organizzandosi in modo serio, trasparente ed aperto: si sono recentemente costituite infatti la prima associazione italiana (**ITPUG – Italian PostgreSQL Users Group**) e la prima associazione europea (**PostgreSQL Europe**) per la promozione di PostgreSQL.

1. INTRODUZIONE

L'adozione diffusa del software libero nella Pubblica Amministrazione italiana non può prescindere da un intervento a livello di infrastruttura software IT. L'obiettivo di questo articolo è quello di promuovere un'analisi critica circa l'impiego di soluzioni open-source per la gestione dei dati nelle realtà medio-piccole della Pubblica Amministrazione in Italia, utilizzando quello che in questo momento è considerato il più avanzato database libero: PostgreSQL.

Una volta descritte le caratteristiche principali di PostgreSQL, l'articolo proverà ad esporre le esigenze di una tipica realtà pubblica istituzionale, evidenziando le barriere che ne hanno frenato l'adozione come RDBMS fino ad ora, proponendo dei

possibili punti di intervento e concludendo con l'esempio del Comune di Prato.

Verrà inoltre fornita una panoramica sulla situazione attuale della comunità di PostgreSQL, sia in Italia che in Europa.

PostgreSQL

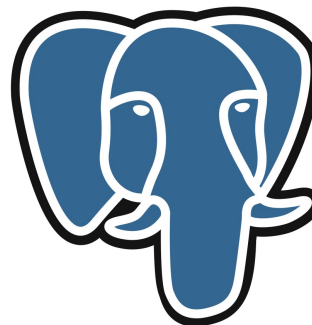


Illustrazione 1: L'elefante, il simbolo di Postgres

2. POSTGRES

PostgreSQL, o semplicemente Postgres, è un sistema open-source molto avanzato per la gestione di basi di dati relazionali a oggetti (ORBMS).

Originariamente sviluppato presso l'Università di California, PostgreSQL è continuamente migliorato grazie al contributo di centinaia di sviluppatori provenienti da tutte le parti del mondo, ed è distribuito secondo la semplice licenza BSD.

PostgreSQL supporta la maggior parte dello standard SQL2003 e fornisce un'ampia gamma di estensioni proprie. Offre la possibilità agli utenti di estendere il sistema tramite la definizione di nuovi tipi di dati, operatori e funzioni.

Oltre a garantire una stabile integrità referenziale e una gestione avanzata delle transazioni, PostgreSQL permette di definire trigger e regole per il controllo di accesso agli oggetti del database.

2.1 Un database di comunità

PostgreSQL è un database di comunità e in tal senso libero da controllo e ingerenze da parte di aziende o singole organizzazioni monopolizzatrici, grazie alla elevata permissività della licenza BSD.

PostgreSQL e il suo codice sorgente non potranno mai essere soggetti ad acquisizioni commerciali, e risulteranno essere sempre un importante patrimonio di conoscenza per la collettività.

2.2 Alcune ragioni per usare Postgres

Le soluzioni proprietarie di sistemi di database sono spesso onerose (non solo per la Pubblica Amministrazione) con offerte e licenze di utilizzo complesse.

Attualmente, PostgreSQL riesce a coprire circa l'80% delle funzionalità e delle richieste di performance offerte dai principali concorrenti proprietari, ad un costo di licenza nullo, privo di *vendor lock-in* e con conseguente riduzione del *costo totale di proprietà*¹.

A partire dalla versione 8, PostgreSQL ha ampiamente migliorato la sua scalabilità e le sue performance, tanto da risultare adatto a data warehouse, data mining e a sistemi di supporto alle decisioni in genere.

Mette a disposizione in modo nativo e maturo funzionalità come viste, schemi, trigger, tablespace, stored procedure (attualmente in 12 linguaggi di programmazione), interfacce di connessione (in particolare ODBC e JDBC), UNICODE, *two-phase commit*, alta disponibilità.

È disponibile per tutti i principali sistemi operativi, fra cui GNU/Linux, Windows, Solaris, FreeBSD, Mac OSX, ecc.

Grazie al suo approccio *object-oriented* e alla sua *estendibilità*, Postgres offre soluzioni valide anche in campi come i sistemi informatici territoriali (GIS), fondamentali per la Pubblica Amministrazione.

Infine, la sua stabilità, la sua sicurezza e l'efficiente supporto da parte della comunità sono ampiamente riconosciuti nel settore RDBMS.

2.3 PostgreSQL 8.3

La 8.3 è l'ultima *major release* stabile di Postgres, rilasciata il 4 febbraio 2008. Come dichiarano le note di rilascio di PostgreSQL 8.3.0 [1], sono state migliorate le prestazioni grazie all'implementazione di funzionalità come *Heap Only Tuples (HOT)*² e checkpoint distribuiti. Altre novità importanti riguardano:

- supporto per SQL/XML³;
- ricerca *full text*⁴;
- supporto all'autenticazione GSSAPI e SSPI;
- adozione di nuovi tipi di dati, in particolare UUID, ENUM ed array di tipi composti.

3.ADOZIONE NELLA PUBBLICA AMMINISTRAZIONE

L'adozione del software libero nella Pubblica Amministrazione può consolidarsi partendo dal livello infrastrutturale. La gestione dei dati è uno dei punti chiave di questo fenomeno, in quanto rappresenta il fulcro di tutte le attività e le applicazioni che vi ruotano intorno. L'unità informativa deve essere sempre disponibile e integrabile.

PostgreSQL è uno strumento che può svolgere egregiamente questa funzione.

Tenendo inoltre conto che la realtà della Pubblica Amministrazione italiana risulta essere molto frammentata (vista la presenza di moltissimi piccoli comuni e di pochi

grandi comuni⁵), le esigenze della maggior parte delle singole PA si riducono fortemente.

Assume pertanto una fondamentale importanza un'azione *scatenante* a livello di amministrazione centrale, che possa promuovere e favorire l'impiego di soluzioni open-source per la gestione dei dati. Per soluzioni RDBMS libere si intendono anche altri prodotti "*concorrenti*" a PostgreSQL, come MySQL e Firebird, per citarne alcuni.

3.1 Requisiti per data storage di un RDBMS

In genere, le esigenze in termini data storage di un RDBMS, sia libero che commerciale, di una PA italiana di medie dimensioni, possono essere riepilogate in:

- stabilità, affidabilità e sicurezza dei dati;
- elevate prestazioni;
- accesso via rete;
- supporto SQL;
- alta disponibilità;
- memorizzazione di grandi quantità di dati;
- integrazione con altre fonti dati;
- interfacciamento tramite standard universali di connessione;
- supporto per codifiche di caratteri non latine;
- supporto per la memorizzazione standard OGC⁶ di dati geografici e analisi geografiche evolute;
- disponibilità di applicazioni verticali (anche proprietarie) in grado di interfacciarsi senza alcun problema con i dati;
- supporto e assistenza professionale 24h/24h;
- diffusione ed utilizzo del RDBMS presso realtà simili o aziende di un certo spessore.

3.2 Lacune e punti di intervento

I due punti deboli dell'*indotto PostgreSQL* sono sicuramente:

- la limitata disponibilità di applicazioni verticali che lo supportano;
- la limitata presenza di aziende in grado di fornire supporto e assistenza professionale.

Probabilmente esiste una correlazione fra i due *fenomeni*.

Le applicazioni generalmente sviluppate per le procedure riguardanti una pubblica amministrazione locale, spesso si rivolgono a database proprietari; in alcuni casi e comunque con una minore frequenza anche ad altri database del mondo open-source.

Una delle asserzioni plausibili è il fatto che le Pubbliche Amministrazioni richiedono che una particolare applicazione risieda su un particolare RDBMS che è in dotazione presso la loro infrastruttura informatica.

¹ Anche conosciuto come *total cost of ownership (TCO)*

² HOT è una funzionalità in grado di migliorare sensibilmente le performance in caso di frequenti aggiornamenti di tabelle

³ Il supporto rispetta lo standard ANSI ed include anche l'export in XML

⁴ Postgres implementa in modo nativo uno strumento di ricerca avanzata nel testo, *Tsearch2*, in grado di supportare lingue e dizionari diversi

⁵ I comuni con più di 100.000 abitanti sono 43, secondo l'ultima fonte ISTAT disponibile.

⁶ *Open Geospatial Consortium*

A parità di funzionalità tecniche richieste⁷, si evince chiaramente che la discriminante decisionale è affidata ai servizi professionali che ruotano intorno a un particolare RDBMS.

E spesso, una scelta che può essere vista anche *opportunistica* e *comoda* è quella di appoggiarsi a soluzioni che riducano le responsabilità e garantiscano *tranquillità*, supporto e assistenza, indipendentemente dal loro costo di licenza e di proprietà totale.

In questo scenario, PostgreSQL paga pesantemente il fatto di essere interamente un **progetto di comunità**, senza una azienda che ne detiene il codice, che si occupa del marketing e che fornisce assistenza professionale. Paradossalmente, la libertà a 360 gradi è il grosso limite all'espansione di Postgres.

In una realtà in cui il bacino di utenza PostgreSQL non è esteso, il principio di *economicità* per le aziende che sviluppano software verticale viene notevolmente a mancare.

Lo scenario potrebbe cambiare se fosse la Pubblica Amministrazione Italiana a promuoverne l'utilizzo. Un primo passo potrebbe essere la promozione dello sviluppo collaborativo di una applicazione verticale open-source che sia basata su Postgres: un esempio su tutti è quello dell'**Anagrafe Unica Libera**.

3.3 Alcune considerazioni

La promozione di PostgreSQL come soluzione per il data storage nelle pubbliche amministrazioni italiane, favorirebbe soluzioni di riuso del software oltre a opportunità di sviluppo collaborativo di software.

Inoltre, esso potrebbe rappresentare un ulteriore incentivo per le imprese italiane a investire nel software libero e, conseguentemente, a investire nella conoscenza e nell'innovazione.

In un contesto fortemente basato sulla piccola e media impresa, la pubblica amministrazione potrebbe favorire la crescita di piccoli distretti locali ICT specializzati nel software libero, dove aziende dinamiche diventano soggetti attivi del sistema FLOSS. Nello specifico caso è auspicabile la crescita di aziende in grado di fornire supporto e assistenza professionale Postgres, che possano fornire la *tranquillità richiesta*, precedentemente accennata.

L'utilizzo di PostgreSQL inoltre porta con sé tutti i vantaggi di libertà e di **diffusione del sapere** tipici del software libero verso i quali la pubblica amministrazione non può rimanere insensibile. Inoltre, il supporto allo standard SQL2003 permette agli utilizzatori di acquisire conoscenze trasferibili anche ad altri RDBMS⁸.

4. POSTGRESQL E IL COMUNE DI PRATO

Il Comune di Prato utilizza PostgreSQL come database per alcuni dei servizi offerti al cittadino nell'ambito della Rete Civica Po-Net [4], una delle principali nel contesto italiano.

⁷ Il divario a livello tecnologico fra RDBMS commerciali e liberi fino a qualche anno fa era molto ampio. Database come MySQL, per certi aspetti, e soprattutto PostgreSQL dalla versione 8, hanno contribuito a ridurre sensibilmente l'ampiezza. Nella comunità di PostgreSQL è diffusa l'idea che circa l'80-90% delle funzionalità richieste ad un RDBMS proprietario siano condivise da Postgres e che altre caratteristiche avanzate siano presenti solamente in quest'ultimo [3].

⁸ Uno degli obiettivi della comunità di Postgres di ITPUG è quello di favorire l'impiego negli istituti superiori e nelle università italiane come strumento didattico per l'apprendimento di sistemi di gestione di database relazionali.

I risultati sono eccellenti, sia in termini di prestazioni che di affidabilità. Fra i servizi implementati con PostgreSQL, si citano:

- newsletter cittadine [5];
- ordinanze su trasporti e viabilità [6];
- archivio delle sintesi e degli avvisi di concorsi [7];
- archivio del tempo libero ed eventi culturali del territorio pratese (programmazione cinematografica, concerti, mostre, ecc.) [8];
- versione on-line dell'opera "I segni del territorio" [9];
- statistiche dei siti web della rete civica Po-Net [10], all'interno del progetto ht://Miner [11].

Il risultato sicuramente più significativo per le prestazioni di Postgres è rappresentato dal database di ht://Miner, che memorizza le informazioni sulle richieste effettuate dall'utenza Internet ai siti web della rete civica Po-Net. Per dare un'idea del volume del traffico, si consideri che la rete civica ha ottenuto nel 2007 oltre 50 milioni di richieste e 10 milioni di visite (opportunitamente filtrate).

Attualmente, il database di ht://Miner contiene 333.317.098 record memorizzati in 236 tabelle con 934 indici a partire dal 1 gennaio 2007; lo spazio fisico complessivamente occupato è 143GB.

Tuttavia, l'utilizzo di PostgreSQL all'interno del Comune potrebbe essere maggiore. Uno degli obiettivi a medio-lungo termine è di utilizzare Postgres anche per servizi maggiormente istituzionali – e non solo per quelli *web oriented*.

5. LA COMUNITÀ ITALIANA E EUROPEA

PostgreSQL, soprattutto dalla versione 8, è uscito dal guscio protettivo di progetto universitario ed è entrato prepotentemente nella fascia *enterprise* di RDBMS.

La comunità di PostgreSQL ha finalmente preso atto della necessità di un'attività adeguata di promozione del progetto (per i motivi esposti in precedenza, non esiste una azienda che cura il marketing di Postgres).

In seguito al successo della prima edizione del PostgreSQL Day Italiano⁹, nel novembre del 2007 la comunità nazionale di PostgreSQL si è organizzata in un'associazione *no-profit* denominata **ITPUG – Italian PostgreSQL Users Group** [12]. Il suo obiettivo principale è la diffusione di Postgres sul territorio italiano.

L'associazione, in continuo e stretto contatto con la comunità internazionale, è stata uno dei promotori più attivi per la costituzione di **PostgreSQL Europe**, una associazione *no-profit* europea per la promozione di Postgres. L'associazione PostgreSQL Europe è stata fondata in occasione del FOSDEM 2008 a Bruxelles¹⁰ da rappresentanti delle maggiori nazionalità europee (Italia, Francia, Germania, Inghilterra, Svezia, Olanda, ecc.).

Entrambe le associazioni stanno attivamente organizzando il primo **PostgreSQL Day Europeo**, che si terrà in Italia nei giorni 17 e 18 ottobre 2008 contestualmente alla seconda edizione del PostgreSQL Day Italiano.

Alla luce degli eventi esposti, si evince chiaramente come anche la comunità stia rispondendo con i fatti alla maturità raggiunta dal software.

⁹ Il PGDay 2007 si è tenuto a Prato presso il Monash Prato University Centre, nei giorni 6 e 7 luglio 2007

¹⁰ Bruxelles, 23 e 24 febbraio 2008

6.CONCLUSIONI

PostgreSQL 8.3 è sicuramente un prodotto eccellente nel settore. Un prodotto maturo e adatto a soddisfare la maggior parte delle richieste di data storage di una tipica pubblica amministrazione.

Viste le funzionalità avanzate che esso supporta, il suo spirito aperto e gli sviluppi relativi alla sua comunità, PostgreSQL richiede maggiore attenzione da parte della pubblica amministrazione e degli addetti ai lavori.

L'auspicio è che istituzioni, aziende e comunità possano trovare il modo di collaborare in sinergia e di ottenere una crescita globale ed economica del sistema.

7.BIBLIOGRAFIA

- [1] The PostgreSQL Global Development Group. Note di rilascio della versione 8.3.0, traduzione in italiano ad opera di Stefano Reksten.
<http://www.postgresql.org/about/press/presskit83.html.it>
- [2] The Open Geospatial Consortium. Sito web:
<http://www.opengeospatial.org/>
- [3] Fujitsu Supported PostgreSQL, The Business Case For PostgreSQL.
http://postgresql.fastware.com/community/business_case_for_postgresql
- [4] Enti vari. Progetto Rete Civica Po-Net. <http://www.po-net.prato.it/>
- [5] Comune di Prato, Servizio Informativo tramite newsletter e SMS “PuntoPerPunto”.
<http://punterpunto.comune.prato.it/>
- [6] Comune di Prato, Ordinanze su Trasporti e Viabilità.
<http://www.comune.prato.it/servizicomunali/ordinanze/trasporti/>
- [7] Comune di Prato, Archivio dei concorsi.
<http://www.comune.prato.it/concorsi/>
- [8] Comune di Prato, Cartellone eventi.
<http://www.comune.prato.it/appuntamenti/>
- [9] Comune di Prato e Marco Piccardi. I segni del territorio.
<http://segnidelterritorio.comune.prato.it/>
- [10] Rete Civica Po-Net. Statistiche degli accessi alla rete civica Po-Net. <http://statistiche.po-net.prato.it/>
- [11] Comune di Prato. Progetto ht://Miner.
<http://www.htminer.org/>
- [12] Italian PostgreSQL Users Group. Sito web.
<http://www.itpug.org/>