

PostgreSQL

Un database libero per la Pubblica Amministrazione Italiana

Gabriele Bartolini

Presidente ITPUG – Italian PostgreSQL Users Group
Vice Presidente PostgreSQL Europe

gabriele.bartolini@2ndquadrant.it
www.itpug.org

Pistoia Linux Day
Pistoia, 25 ottobre 2008



Panoramica sull'intervento

- **Obiettivo:** analisi critica sull'impiego di soluzioni open source
- **Campo di applicazione:** Gestione dei dati e database
- **Contesto sociale:** Pubblica Amministrazione
- **Proposta:** utilizzo e promozione di PostgreSQL



Sommario

- 1) PostgreSQL 8
- 2) Adozione di PostgreSQL nella Pubblica Amministrazione
- 3) PostgreSQL e il Comune di Prato
- 4) Sviluppi riguardanti la comunità di PostgreSQL
- 5) Conclusioni



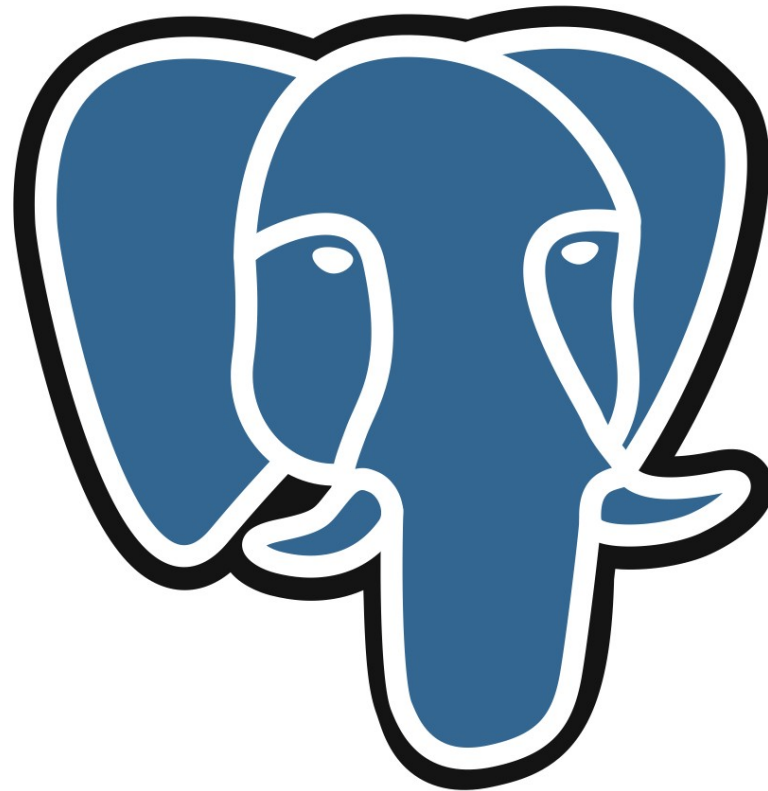
PARTE I

PostgreSQL 8



PostgreSQL o semplicemente Postgres

PostgreSQL



PostgreSQL

- Sistema per la gestione di basi di dati relazionali a oggetti (ORBMS)
- Open Source
- Progetto nato presso l'Università di California (Berkeley):
 - 1986 – 1993: **University Postgres** (Michael Stonebraker)
 - 1994 – 1995: **Postgres95** (Jolly Chen e Andrew Yu)
 - 1996 – oggi: **PostgreSQL**:
 - Fase 1, 1996 – 1998: **eliminazione dei crash**
 - Fase 2, 1998 – 2001: **aderenza allo standard SQL**
 - Fase 3, 2001 – 2004: **robustezza e stabilità**
 - Fase 4, dal 2004 in poi: **funzionalità di classe enterprise**
- Contributo di centinaia di sviluppatori provenienti da tutto il mondo
- Scritto in C (altamente portabile)
- Distribuito secondo la licenza BSD
- Supporta la maggior parte dello standard SQL:2003
- Fornisce un'ampia gamma di estensioni proprie



Un database di comunità e “internazionale”

- PostgreSQL è un database di comunità
- Libero da controllo e ingerenze da parte di:
 - Aziende
 - Singole organizzazioni monopolizzatrici
- Non soggetto ad acquisizioni commerciali
- Patrimonio di conoscenza per la collettività
- Come la maggior parte dei progetti open-source di successo:
 - È sensibile agli standard internazionali (SQL:2003)
 - È sensibile all'interscambio dei dati (supporto XML)
 - È sensibile all'internazionalizzazione e alla localizzazione:
 - Supporto UTF-8 per la codifica di caratteri Unicode
 - Supporto per National Language (NLS)



Alcune ragioni per usare PostgreSQL (1/2)

- Costo di licenza nullo
- Assenza di *vendor lock-in*
- A partire dalla versione 8:
 - migliore scalabilità
 - migliori performance
 - adatto per sistemi di supporto alle decisioni (data mining e warehouse)
- Funzionalità native mature e stabili:
 - viste
 - schemi
 - stored procedure (attualmente in 12 linguaggi di programmazione)
 - trigger
 - tablespace
 - interfacce di connessione (in particolare ODBC e JDBC)
 - supporto UNICODE
 - two-phase commit
 - alta disponibilità



Alcune ragioni per usare PostgreSQL (2/2)

- Multi-piattaforma:
 - GNU/Linux
 - Microsoft Windows
 - Unix
 - Solaris
 - FreeBSD
 - Mac OS X
 - ...
- Supporto per la memorizzazione standard OGC di dati geografici e analisi geografiche evolute (**PostGIS**)
- Stabilità
- Sicurezza
- Efficiente supporto da parte della comunità
- Esistono diverse soluzioni per la replica e l'alta disponibilità (HA)



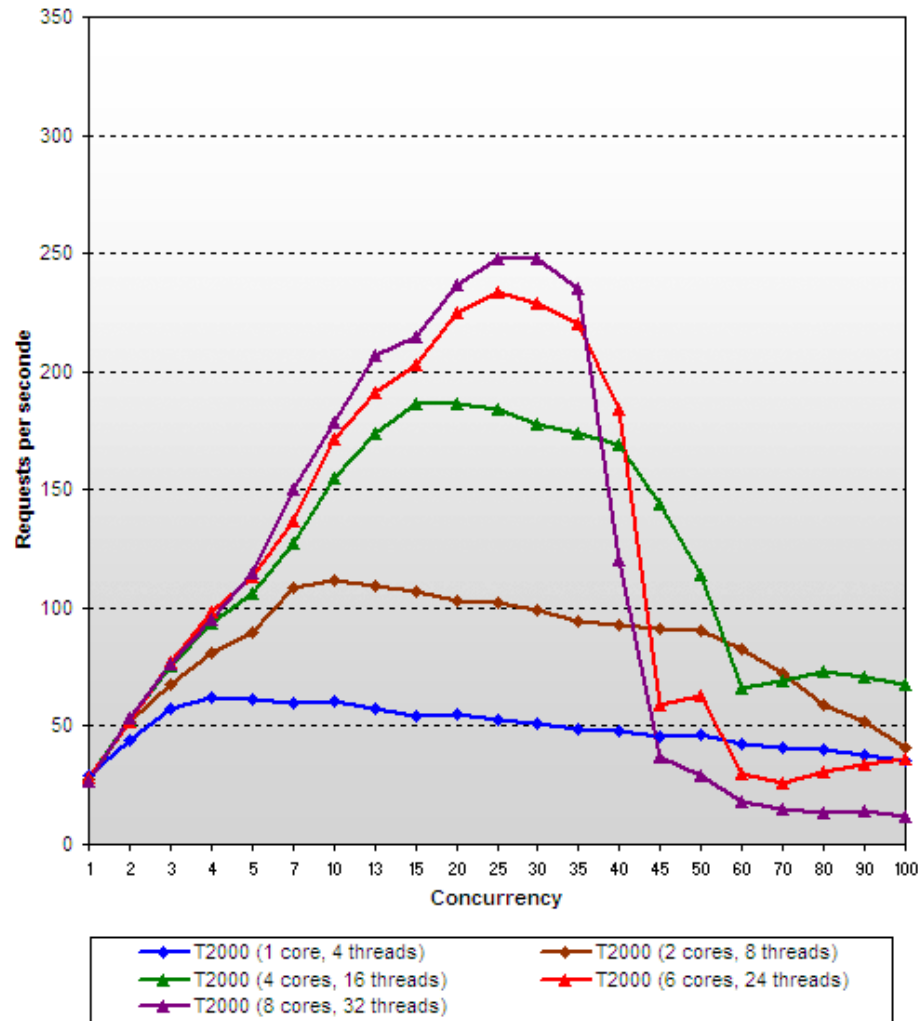
Lo sapevate che ...

- Skype utilizza PostgreSQL come database di backend per memorizzare i profili dei propri utenti
- Attualmente il database contiene circa **400 milioni** di profili
- Ma è scalabile fino a **1 miliardo** di profili
- Skype sviluppa soluzioni open-source per PostgreSQL riguardanti:
 - Alta disponibilità (HA)
 - Scalabilità
 - Clustering
 - Partitioning
 - ...

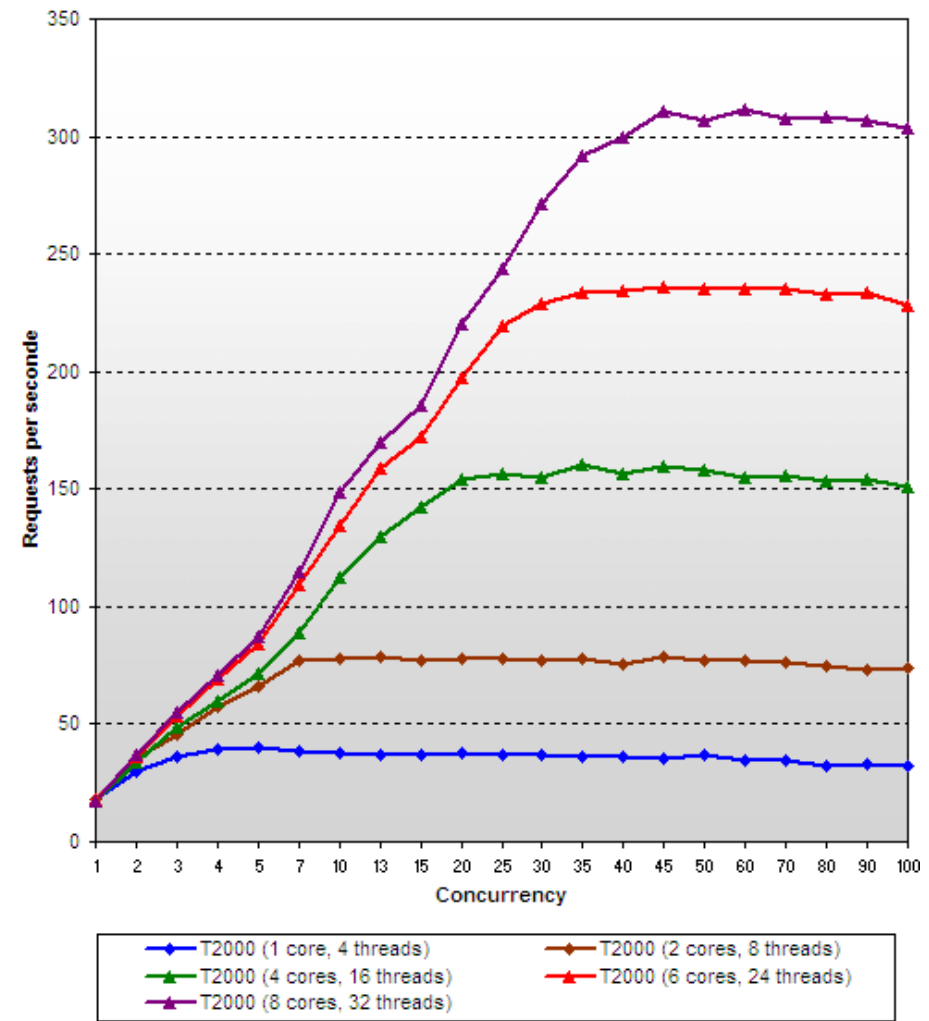


MySQL vs PostgreSQL: scalabilità (26/10/06)

Tweakers.net Database Simulatie - MySQL 5.0.20a schaalgedrag



Tweakers.net Database Simulatie - PostgreSQL 8.2 schaalgedrag



Fonte: <http://tweakers.net/reviews/649/7/database-test-sun-ultrasparc-t1-vs-punt-amd-opteron-pagina-7.html>



PARTE II

Adozione di PostgreSQL nella Pubblica Amministrazione



Adozione di FLOSS nella PA

- Deve partire dal livello infrastrutturale
- La gestione dei dati è un punto chiave
- Il dato è un valore fondamentale per la PA:
 - unità informativa
 - sempre disponibile
 - integrabile
- Pubblica Amministrazione locale italiana molto frammentata
- **Esigenze ridotte per la maggior parte delle PA**
- Necessità di un'azione scatenante a livello di amministrazione centrale che favorisca l'impiego di soluzioni libere per la gestione dei dati:
 - PostgreSQL
 - ma anche MySQL, Firebird, ecc.



PA: requisiti per data storage di un RDBMS

- **stabilità, affidabilità e sicurezza dei dati**
- **elevate prestazioni**
- **accesso via rete e 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 OGC 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 x 7g*
- *diffusione ed utilizzo del RDBMS presso realtà simili o aziende di un certo spessore*



Lacune di PostgreSQL e punti d'intervento

- Punti deboli dell'indotto PostgreSQL:
 - 1) Limitata disponibilità di applicazioni verticali che lo supportano
 - 2) Limitata presenza di aziende di supporto e assistenza professionale
- Correlazione fra i due fenomeni
- Molto spesso:
 - le applicazioni sviluppate per procedure di una PA locale si rivolgono a database proprietari
 - le PA richiedono applicazioni per uno degli RDBMS in loro dotazione
- **A parità di funzionalità tecniche richieste:**
 - la discriminante decisionale è affidata ai servizi professionali che ruotano intorno a un particolare RDBMS
 - una scelta comune, anche opportunistica, è appoggiarsi a soluzioni che, indipendentemente dal loro costo di licenza e di proprietà totale:
 - riducano le responsabilità
 - garantiscano tranquillità, supporto e assistenza



Punti di intervento

- Progetto di comunità, nessuna una azienda:
 - Sviluppi una mirata e efficace strategia di marketing
 - Fornisca assistenza professionale
- Bacino di utenza PostgreSQL non è esteso
- Ridotta economicità per le software house tale da giustificare un investimento in sviluppo di applicazioni su PostgreSQL
- Le cose potrebbero cambiare se la Pubblica Amministrazione Italiana:
 - Promuovesse lo sviluppo collaborativo di applicazioni verticali open-source basate su PostgreSQL: **Anagrafe Unica Libera**
 - Promuovesse realmente il riuso del software libero
 - Promuovesse PostgreSQL come soluzione preferibile per il data storage



Investire nel FLOSS

- Occorre incentivare le imprese italiane (piccole e medie) a:
 - Investire nel software libero, ovvero
 - Investire in conoscenza competitiva, ovvero
 - Investire in innovazione e in qualità
 - Auto-organizzarsi in distretti locali ICT specializzati



La coscienza del dipartimento IT di una PA

- Qualora una PA decida di NON adottare un database *open-source* come PostgreSQL, dobbiamo per forza di cose essere coscienti che:
 - Nella quasi totalità dei casi
 - **NON SI TRATTA DI REQUISITI TECNICI**
- Se la scelta è puramente tecnica, è lecito **chiedere le motivazioni** che hanno portato alla esclusione di un prodotto come PostgreSQL, nella fattispecie
- L'era del “*non lo sapevo*”, oppure “*il db libero non è idoneo*” è finita
- Occorre sensibilizzare il personale della PA al movimento e alla cultura open-source, soprattutto in questo periodo, dando impulso e spinta a una **generazione di giovani informatici coscienti e consapevoli**



PARTE III

PostgreSQL e il Comune di Prato



PostgreSQL nel Comune di Prato

- Impiegato dal 2000/2001
- Database per alcuni dei servizi principali della Rete Civica Po-Net:
 - newsletter cittadine (PuntoPerPunto)
 - ordinanze su trasporti e viabilità
 - archivio delle sintesi e degli avvisi di concorsi
 - archivio del tempo libero ed eventi culturali del territorio pratese
 - versione on-line dell'opera “I segni del territorio”
 - statistiche dei siti web della rete civica Po-Net - progetto [ht://Miner](http://Miner)
- Eccellenti risultati:
 - Ottime prestazioni
 - Alta affidabilità
- Tuttavia PostgreSQL potrebbe ambire a qualcosa di più critico (*“speranza personale”*: anagrafe?)



Alcuni numeri significativi

- Il database più complesso gestito con Postgres nel Comune è sicuramente il data warehouse con le statistiche di accesso al web
- Parte del progetto ht://Miner
- Raggiungibile su: statistiche.po-net.prato.it
- La rete civica ha ottenuto nel 2007:
 - oltre 50 milioni di richieste
 - 10 milioni di visitatori unici
- Attualmente ht://Miner di Po-Net contiene:
 - 425.298.160 record
 - 309 tabelle (in partizionamento verticale)
 - 1.218 indici
 - dati dal 1 gennaio 2007 – 16 settembre 2008
 - spazio fisico: 182 GB



PARTE IV

Sviluppi riguardanti la comunità di PostgreSQL e conclusioni



La comunità ITPUG

- La comunità di PostgreSQL ha finalmente preso atto della necessità di un'attività adeguata di promozione del progetto
- Dopo il primo PostgreSQL Day Italiano la comunità nazionale di PostgreSQL si è organizzata in un'associazione no-profit denominata ITPUG – Italian PostgreSQL Users Group
- Obiettivo principale è la diffusione di Postgres sul territorio italiano
- Promotore fra i più attivi per la costituzione di PostgreSQL Europe
- PostgreSQL Europe è stata fondata in occasione del FOSDEM 2008 a Bruxelles da rappresentanti delle maggiori nazionalità europee (Italia, Francia, Germania, Inghilterra, Svezia, Olanda, ecc.)
- Partecipazione aperta a tutti
- Informazioni e iscrizioni: www.itpug.org



PGDay Italiano, Prato, 6-7 luglio 2007



PGDay 2008



PGDay 2008

- ITPUG ha organizzato il **primo PostgreSQL Day Europeo**, che si è tenuto:
 - a Prato, presso il Monash University Prato Centre
 - nei giorni venerdì 17 e sabato 18 ottobre 2008
- Sinergia con:
 - PostgreSQL Europe
 - PostgreSQL Global Development Group
- 171 partecipanti nei due giorni



Conclusioni

- PostgreSQL 8.3 è un eccellente software di database:
 - Libero
 - Funzionalità avanzate e di qualità
 - Stabile e maturo
 - Adatto per il data storage di una tipica pubblica amministrazione
- PostgreSQL 8.4 uscirà nel 2009
- PostgreSQL merita maggiore attenzione da parte:
 - delle istituzioni
 - degli addetti ai lavori
- E' necessaria la collaborazione sinergica fra:
 - Istituzioni (promozione del software libero)
 - Aziende (supporto e sviluppo)
 - Comunità (promozione e sviluppo)



Domande?



Grazie



Licenza Creative Commons

Attribuzione
Non commerciale
Condividi allo stesso modo
2.5 Italia

<http://creativecommons.org/licenses/by-nc-sa/2.5/it/>

