

Adattárház kialakítása nyílt forráskódú eszközökkel

Ragasits Csaba
Sopron Bank Zrt.
2012

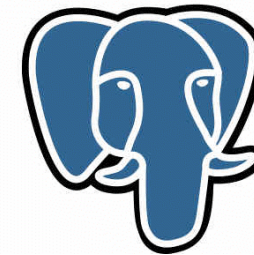
Sopron Bank Zrt.



BRÜLL KALLMUS



Bank és az Open Source



Miért kezdtünk adattárházat építeni?

- Nagyon kevés a lekérdezési lehetőség
 - Szövegfájl kimenet, manuális futtatás
 - Optimalizálás tranzakciókra
- Új lekérdezés csak a szállítón keresztül
 - Lassú átfutás
 - Magas költség
- Helyi nyilvántartások



Fejlesztési folyamatok

- Fejlesztés nulláról, kis lépésekben
- SCRUM
- Mindig csak az adott feladatra koncentrálnunk:
 - Gyorsan tudunk használható megoldást adni
 - Sokszor kell visszatérni az alapokhoz
 - Van amit utólag nehéz berakni



Felhasználói csoportok



- Halandók
 - Előre gyártott, paraméterezhető riportok, lekérdezések
 - Csak a szükséges adatokhoz lehet hozzáférni
- Félistenek
 - Saját adatpiac, munkaterület
 - Lekérdezések készítése
- Szerénységem
 - A fejlesztő



Adatbázis-kezelő



- PostgreSQL

- A Bank preferált adatbázis-kezelője
- OLTP, DW rendszerek szétválasztása

- Replikáció (Master + Slave)
- Terheléselosztás (pgpool)

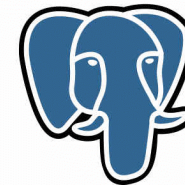
- Monitorozás (Nagios, Munin)

- Kis adatbázis méret

- Csak a szükséges adatok tárolása, nézetek használata
- Horizontális, vertikális particionálás, materializált nézetek

- IO a szűk keresztmetszet

PostgreSQL



Nagios®

ETL

- Saját eszköz 
 - Egy forrásrendszer
 - Soros feldolgozás
 - Primitív hibakezelés
- Pentaho PDI (Kettle)
 - Több forrásrendszer
 - Párhuzamos feldolgozás 
 - Intelligens hibakezelés
- Főbb feldolgozások
 - Forrásrendszerek
 - Historizmus, OLAP
 - Előkészítés
 - Adatszolgáltatás
 - Külső adatok feldolgozása 
 - Felhasználó tevékenységek

Riportok

- Jaspersoft BI Community Edition
 - Fontos, hogy mindenki csak az engedélyezett adatokhoz férhessen hozzá
 - Bevezetéskor megoldott feladatok
 - AD autentikáció
 - AD-ban megadott jogosultságok
 - Forrásrendszerekben megadott jogosultságok
 - Magyar nyelvű kezelőfelület
 - Riportok ütemezett futtatása
 - Verzióváltás 3.7 -> 4.5
 - Font problémák javítása



Egyedi lekérdezések

- MS Access
 - Lassú lekérdezések
 - Riport készítéshez újra kell írni
- Fly Speed SQL Query
 - Natív SQL, optimalizálási lehetőség
 - lekérdezések
 - Könnyű riportkészítés



- Jelenleg az Access teljesen visszaszorult
- Lekérdezési szigetek
 - Egyedi lekérdezés gyűjtemények
 - Párhuzamos lekérdezés készítés eltérő eredményekkel
 - A megoldásra eszköz keresése, vagy saját fejlesztés



Analízis

- Felhasznált eszközök
 - Pentaho Mondrian
 - Jasperserver + jPivot
- Hópehely séma
- Felhasználói fogadtatás
 - jPivot nehezen használható
 - FlySpeed-ben gyorsabb megírni
 - A jövőben talán a Jedox PALO



Jövőbeli lehetőségeink

- Folyamatos a fejlesztés
- Elértük a FOSS lehetőségek határait
- Az analízis megszerettetése
- Az üzleti terület támogatása (Dashboard, KPI)
- A DW beépítése üzleti folyamatokba, előtét rendszer
- Az első adatbányászat (RapidMiner)



Összefoglalás

- Létrehoztunk egy adattárházat
- Ahol lehetett nyílt forráskódú eszközöket használunk
- A fejlesztés nem áll meg
- Az adattárház a Bank mindennapi életének szerves része



Köszönöm a figyelmet!

Ragasits Csaba
ragasitscs@sopronbank.hu
[@ragasits](#)