

# APLICAȚII ÎNTEGRATE PENTRU ÎNȚREPRINDERI

## Tema 4

Evaluarea activității angajaților din cadrul departamentelor de programare și controlul calității într-o companie dezvoltatoare de software

Lansare	Termen de Predare	Pondere
07.01.2014, 23:55	21.01.2014, 23:55	5 / 40

### Obiective

Scopul laboratorului constă în realizarea unui sistem informatic performant și scalabil pentru gestiunea resurselor unei întreprinderi de dimensiuni mari. În acest an universitar, organizația pentru care se va proiecta și implementa aplicația ERP este o companie dezvoltatoare de software.

Se va urmări implementarea graduală a acestui sistem informatic prin intermediul unor teme de casă, pe parcursul a patru etape.

În cadrul acestei etape se va dezvolta o aplicație distribuită care va rula pe mai multe mașini, în condițiile în care volumul de date pe care trebuie să îl gestioneze organizația implică timpi de execuție inacceptabili pentru un utilizator care o interoghează.

După rezolvarea temei de casă, studentul va fi capabil să:

- proiecteze o bază de date distribuită, având o structură dinamică și un număr de înregistrări foarte mare;
- utilizeze paradigma de programare MapReduce pentru a gestiona eficient partițiile unei probleme care pot fi procesate concomitent;
- realizeze o aplicație distribuită ce conține mai multe sarcini, dintre care unele vor fi rulate secvențial, iar altele în paralel.

Cunoștințele necesare<sup>1</sup> pentru rezolvarea temei de casă sunt:

- programarea în limbajul Java;
- manipularea unei bazele de date distribuite HBase (atât prin shell-ul specific cât și prin API);
- paradigma de programare MapReduce.

### Enunț

Întrucât numărul de înregistrări stocat în baza de date relațională determină ca timpul de răspuns pentru operațiile de raportare (ce implică volume mari de date) să nu respecte constrângerea cu privire la operarea în timp real a unei aplicații integrate pentru întreprinderi, la nivelul executiv al organizației dezvoltatoare de software s-a luat decizia migrării spre o soluție distribuită, pentru care a decis achiziționarea unui cluster format din mașini având un cost redus, fără performanțe distribuite.

Se dorește monitorizarea performanțelor angajaților din departamentele de programare și asigurare a calității, prin stabilirea contribuției aduse de aceștia în cadrul fiecărui proiect în care au fost implicați, simplificând atribuțiile responsabilului de echipă și obiectivizând totodată a procesului de evaluare prin utilizarea unor parametri cantitativi concreți<sup>2</sup>.

<sup>1</sup> NU este permisă folosirea altor tehnologii decât a celor specificate în cadrul enunțului.

<sup>2</sup> Fiind vorba de o evaluare strict cantitativă, sigur că aceasta nu se poate constitui în unicul instrument prin care să se aprecieze performanțele unui angajat, însă poate reprezenta un indicator util, mai ales în cadrul organizațiilor de dimensiuni mari unde – în ciuda unei ierarhii foarte bine stabilite – activitatea unui salariat poate să nu fie foarte corect cunoscută factorilor de decizie.

Astfel, vor fi generate două rapoarte sub forma unor fișiere text, în care să se prezinte contribuția fiecărui angajat în cadrul fiecărui proiect, prin contorizarea numărului de defecte (din fiecare tip de severitate) la care și-a adus aportul<sup>3</sup>, respectiv numărul de proiecte și numărul total de defecte în care a fost implicat, de la angajarea sa în organizație.

Un raport detaliat conține numele și prenumele angajatului, denumirea proiectului pentru care se afișează informațiile și severitatea defectelor contorizate:

**raport\_detaliat**

NUME<sub>k</sub> Prenume<sub>k</sub>, denumire\_proiect<sub>k</sub>, denumire\_severitate<sub>k</sub> = număr\_defecte

Un raport sintetic conține numele și prenumele angajatului denumirea departamentului în care lucrează acesta, numărul total de proiecte și numărul total de defecte:

**raport\_sintetic**

NUME<sub>k</sub> Prenume<sub>k</sub>, denumire\_departament<sub>k</sub> = număr\_proiecte, număr\_defecte

**Precizări suplimentare**

Se recomandă ca pentru implementarea temei de casă să folosiți mașina virtuală pusă la dispoziție în cadrul laboratorului sau să folosiți aceleași versiuni / proprietăți de configurare ca în cadrul acesteia, întrucât corectarea se va realiza pe o astfel de platformă. Cu alte cuvinte, veți utiliza Hadoop 2.2.0 și HBase 0.96.

Este suficient ca în portarea bazei de date să preluați doar acele tabele necesare pentru rezolvarea cerințelor. Întrucât procesul de populare implică mai multe comenzi pentru o singură înregistrare, se recomandă și dezvoltarea unui program pentru generarea scriptului care realizează această operație.

Nu este necesară includerea unui script pentru crearea unei arhive jar conținând aplicația dezvoltată, întrucât operația este realizată de plugin-ul Hadoop pentru Eclipse, acesta fiind responsabil și de transmiterea sarcinii, permițând totodată consultarea informațiilor din sistemul de fișiere distribuit HDFS.

Nu este permisă utilizarea unui obiect de tip `org.apache.hadoop.hbase.client.Scan` decât ca parametru pentru clasele care realizează sarcini de tip MapReduce, întrucât operațiile realizate de acesta, în contextul unei singure mașini, implică resurse de procesare și memorie care depreciază timpul de execuție al unei aplicații.

În raport vor fi omise informațiile nule, în cazul în care activitatea unui angajat pentru un proiect nu poate fi contorizată conform înregistrărilor reținute în baza de date.

Pentru a se optimiza timpul de execuție, raportul sintetic va fi generat pe baza raportului detaliat, fără a accesa baza de date decât spre a obține departamentul unui angajat.

**Barem de corectare și notare**

Punctaj	Criterii de acordare
1,25 p	reproiectarea bazei de date conform modelului distribuit <ul style="list-style-type: none"><li>definirea tabelor și a structurii statice: 30%</li><li>populare corespunzătoare: 70%</li></ul>
3,50 p	generarea rapoartelor <ul style="list-style-type: none"><li>raport detaliat: 35%</li><li>raport sintetic: 65%</li></ul>
0,25 p	modularizare <ul style="list-style-type: none"><li>structura aplicației și lizibilitatea codului: 50%</li><li>comentarii, README: 50%</li></ul>

<sup>3</sup> În acest sens, va fi necesară reproiectarea bazei de date (prin crearea unei tabele dedicate) astfel încât să se stocheze fiecare contribuție a unui angajat la un defect (nu doar informația despre ultima actualizare asupra defectului respectiv).

**BONUS.** Se pot obține punctaje suplimentare, astfel:

- 0,25p – predarea temei până la 14.01.2014, ora 23:55;
- 0,25p – determinarea, în cadrului raportului detaliat, a numărului de defecte rezolvate<sup>4</sup> (se va afișa raportul defecte\_rezolvate / total\_defecte).

**Condiții de realizare și predare**

Tema va fi realizată individual și va trebui încărcată pe site-ul cursuri.cs.pub.ro sub forma unei arhive de tip .zip (cu denumirea **Grupa34XCX\_NumePrenume\_Tema4.zip**) care să conțină script-ul pentru crearea și popularea tabelor din baza de date, sursele aplicației și un README în care să prezentați definiția bazei de date (tabele, familii de coloane și structura lor), explicând cum ați proiectat sarcinile de tip map și reduce.

Temele vor fi comparate prin aplicații specializate pentru a depista eventualele fraude. În această situație, întreg punctajul pe parcursul semestrului va fi anulat, studenții implicați (atât originalul, cât și copia / copiile) fiind obligați să repete disciplina – cu taxă – în anul universitar următor.

---

<sup>4</sup> Se consideră că un angajat este cel care a rezolvat un defect, dacă ultima modificare realizată asupra acestuia îi aparține.