

# Lucrare Licenta

5 iunie 2016



# Capitolul 1

## Introducere



# Capitolul 2

## Tehnologii folosite

### 2.1 Java

Java este un limbaj de programare orientat-obiect realizat de către James Gosling în cadrul companiei Sun Microsystems (acum filiala Oracle) și a fost creat să aibă cât mai puține dependențe legate de implementare. A fost creat ca să ajute dezvoltatorii de aplicații, limbajul fiind unul de tip WORA (Write Once Run Anywhere, traducere: Scrie odată rulează peste tot). A fost lansat în anul 1995 iar de atunci a suferit o popularitate foarte mare, majoritatea aplicațiilor distribuite fiind scrise în Java, iar odată cu evoluția tehnologiei, limbajul de programare a devenit utilizat și la scrierea aplicațiilor mobile, agendelor telefonice, etc. Java este utilizat în prezent, având un real succes, și la programarea aplicațiilor destinate intranet-urilor. Astăzi, Java este unul dintre cele mai utilizate limbaje de programare, în special pentru aplicațiile web de tip client-server, având reportat un număr de 9 milioane de dezvoltatori.

Limbajul Java își însușește o mare parte din sintaxa limbajelor C și C++. Diferența dintre aceste limbaje este făcută în cadrul nivelului de jos al programării, Java având mai puține facilități, dar obiectul modelelor este de asemenea simplificat. Un program Java care este corect și care nu suferă modificări poate rula pe orice platformă care are instalată mașina virtuală Java (JVM care înseamnă Java Virtual Machine). Acest lucru este posibil deoarece codul scris în Java este compilat într-un format standard numit cod ce octeți (în engleză byte-code). Codul de octeți diferă față de codul mașină. Codul mașină poate fi executat fără a mai fi prelucrat deoarece este reprezentat de o secvență de instrucțiuni specifice unui anumit tip de procesor și a unei platforme binare de lucru. Codul de octeți reprezintă instrucțiuni asemănătoare codului scris în limbaj de asamblare. Compilatorul generează codul de octeți independent de platforma de lucru. Codul mașină este executat direct de către procesor și va putea rula numai pe platforma pe care a fost creat, pe când codul de octeți va putea rula pe orice platformă care are instalat mediul de programare Java deoarece codul de octeți este interpretat întocmai de acest mediu de programare.

Java Virtual Machine este disponibilă pentru toată lumea începând din anul 2006, atunci când Sun a publicat un articol prin care anunțau că varianta companiei de JVM va fi open-source. Există și alți furnizori de JVM, printre care Oracle, IBM, Bea, FSF.

În funcție de modul de execuție a aplicațiilor, limbajele de programare se împart în două categorii:

- Interpretate: un program numit interpretor citește instrucțiunile linie cu linie și le traduce în instrucțiuni mașină. Această soluție prezintă un avantaj mare datorită simplității și a portabilității. Portabilitatea este dată de faptul că sursa programului este interpretată direct. Viteza de execuție redusă este un dezavantaj evident pentru această categorie. Limbajul de programare Basic este, probabil, cel mai cunoscut limbaj interpretat.

- **Compile:** compilatorul transformă codul sursă al programelor în cod care poate fi executat direct de către procesor. Acest cod se numește cod mașină. Un avantaj îl constituie execuția extrem de rapidă, de cealaltă parte, lipsa portabilității este un dezavantaj important. Un cod de nivel scăzut nu va putea rula decât pe platforma de lucru pe care s-a executat procesul de compilare.

Caracteristicile principale ale limbajului de programare Java au făcut ca acesta să fie, într-un interval relativ scurt, unul dintre cele mai populare și utilizate limbaje de programare. Caracteristicile principale ale limbajului sunt:

- **Simplitate** - supraîncărcarea operatorilor este eliminată, la fel și moștenirea multiplă și toate facilitățile care ar putea duce la scrierea unui cod confuz
- **Ușurința** cu care sunt create aplicațiile complexe (interfața grafică, firele de execuție, baze de date, etc).
- **Robustețe** - Renunțarea la pointeri duce la eliminarea unei surse foarte importante de erori, la fel și gestionarea automată a memoriei și eliminarea scurgerilor de memorie prin introducerea unei proceduri de colectare a obiectelor care nu mai sunt referite, procedura este numită "Garbage Collector".
- **Orientare pe obiecte completă** - programarea procedurală este eliminată complet.
- **Securitate** - limbajul Java este unul foarte sigur deoarece furnizează mecanisme stricte de securitate (codul care verifică secvențele periculoase este verificat dinamic, atunci când se rulează procese la distanță sunt impuse niște reguli foarte stricte).
- **Neutralitate arhitecturală** - arhitectura fizică a mașinii pe care rulează aplicația nu influențează comportamentul acesteia.
- **Portabilitate** - independența față de platforma de lucru reprezintă un avantaj foarte mare pentru limbajul de programare și este o opțiune foarte bună pentru companiile dezvoltatoare de aplicații din cauza reducerii costurilor. Aplicațiile nu mai necesită recompilate și modificate la trecerea de la un sistem de operare pe altul.
- Este un cod compilat și interpretat, aceste două mari caracteristici definesc portabilitatea limbajului de programare.
- **Performanță** - recunoscut pentru faptul că limbajul este mai lent decât cele care generează executabile native pentru o anumită platformă de lucru, performanța ridicată a codului de octeți este asigurată de compilatorul Java, astfel încât viteza mai scăzută nu reprezintă un impediment dezvoltării aplicațiilor complexe ( animații, grafică 3D).

Oracle a furnizat pentru Java 4 platforme:

- **Java Card** – pentru smartcarduri (carduri cu cip)
- **Java Platform, Micro Edition (Java ME)** – pentru hardware cu resurse limitate gen telefoane mobile
- **Java Platform, Standard Edition (Java SE)**- pentru sisteme gen workstation ( gasim la PC-uri)
- **Java Platform, Enterprise Edition (Java EE)** – pentru sisteme de calcul mari

Mai multe versiuni au apărut de-a lungul timpului, și anume un număr de 9 versiuni. Acestea sunt :

- JDK 1.0 (1996) – versiunea inițială
- JDK 1.1 (1997)
- J2SE 1.2 (1998)
- J2SE 1.3 (2000)
- J2SE 1.4 (2002)
- J2SE 5.0 (2004)
- Java SE 6 (2006)
- Java SE 7 (2012)
- Java SE 8 (2014)

Ultima versiune a limbajului Java (Java SE 8 -2014) este singura versiune gratis suportată de Oracle.

Un IDE (Integrated Development Environment) este, după cum spune numele, un mediu de lucru care permite dezvoltarea aplicațiilor specifice unui limbaj de programare. Pentru Java sunt următoarele : JCreator, Eclipse, NetBeans, BEA Workshop, BlueJ, CodeGuide, Dr Java, IntelliJ IDEA, JBuilder, JDeveloper, KDevelop.

Limbajul de programare Java a fost creat pe baza a cinci obiective majore:

1. Trebuie să fie simplu, orientat pe obiect și familiar
2. Trebuie să fie robust și sigur.
3. Trebuie să fie neutru din punct de vedere al arhitecturii și în același timp, portabil.
4. Trebuie să fie executat cu performanță ridicată.
5. Trebuie să fie interpretat, filetat și dinamic.

Oracle Corporation este proprietarul actual al implementării oficiale a platformei Java Standard Edition, ca urmare a achiziției companiei Sun Microsystems în anul 2010. Implementarea actuală este bazată pe implementarea originală făcută de cei de la Sun. Implementarea Oracle este disponibilă pentru Microsoft Windows (încă merge și pentru Windows XP), Mac OS X, Linux și Solaris.

Implementarea Oracle este distribuită în două moduri:

- The Java Runtime Environment (JRE) - conține părțile din platforma JAVA SE necesare pentru a rula programe Java
- The Java Development Kit (JDK) - este destinat programatorilor și conține programele de dezvoltare precum compilatorul Java, Javadoc, Jar și un debugger.

OpenJDK este altă implementare din familia JAVA SE care este sub licența GNU GPL. Implementarea a început când Sun a lansat codul sursă pentru Java sub licența GPL. Ca și în JAVA SE 7, OpenJDK este referința oficială a implementării limbajului de programare Java.

Scopul limbajului de programare Java este de a face compatibile toate implementările. De-a lungul istoriei, marca înregistrată a limbajului Java pentru utilizare (Sun trademark) a insistat ca toate implementările să fie compatibile. Acest lucru a iscat un conflict legal cu Microsoft după ce Sun a susținut că implementarea realizată de cei de la Microsoft nu avea suport pentru RMI (Remote Method Invocation) sau JNI (Java Native Interface) și au adăugat unele caracteristici specifice doar platformei lor. Sun a dat în judecată Microsoft în anul 1997, iar în anul 2001 au câștigat daune în valoare de 20 de milioane de dolari. Microsoft a fost de asemenea obligată să respecte termenii licenței impuși de Sun. Microsoft nu a acceptat acest lucru și începând din anul 2001 ei nu mai livrează Java împreună cu sistemul de operare Windows.

Platforma independentă Java este esențială către Java EE (Enterprise Edition) și validare mai strictă este necesară pentru a obține o implementare. Acest mediu permite aplicații portabile pe partea de server.

Programele scrise în limbajul Java sunt recunoscute pentru viteza scăzută și consumul mai mare de memorie față de cele scrise în limbajul C++. În orice caz, viteza de execuție a programelor scrise în Java s-a îmbunătățit semnificativ cu introducerea compilării “chiar la timp” în anul 1997/1998 pentru versiunea Java 1.1, adăugarea de noi caracteristici pentru limbaj suportă o mai bună analiză a codului ( clase în clase, clasa StringBuilder, optimizări în mașina virtuală Java). Odată cu versiunea 1.5, performanța a fost îmbunătățită cu adăugarea pachetului `java.util.concurrent`. Această implementare a fost îmbunătățită ulterior cu versiunea 1.6.

Unele platforme oferă suport hardware direct pentru Java, există microcontrollere care rulează Java în hardware și nu în mașina virtuală, și procesoare ARM care oferă suport pentru execuția codului de octeți Java.

Java utilizează un Garbage Collector automat pentru organizarea memoriei în ciclul de viață al unui obiect. Programatorul determină când sunt create obiectele iar Java runtime este responsabil cu recuperarea memoriei odată ce obiectele nu mai sunt utilizate. Odată ce nu rămân referințe la obiecte, memoria neaccesibilă devine disponibilă pentru eliberare automată de către Garbage Collector. Ceva similar unei scurgeri de memorie poate să apară dacă codul implementat de programator stochează o referință la un obiect care nu mai este nevoie de el, de obicei apare acest lucru când obiectele de care nu mai sunt nevoie sunt stocate în containere care încă sunt utilizate. Dacă sunt executate operații asupra unui obiect neexistent, va fi aruncată excepția “Null Pointer Exception”.

Una din ideile din spatele organizării automate a memoriei este aceea că programatorii sunt scutiți de povara de a face manual administrarea memoriei. În unele limbaje de programare, memoria pentru crearea obiectelor este alocată implicit în stack (grămadă), sau explicit alocată și dealocată în memoria heap, în acest ultim caz, programatorului îi revine responsabilitatea de a administra memoria. Dacă programul nu dealocă obiectul, va apărea o scurgere de memorie. Dacă programul încearcă să acceseze sau dealoce o zonă de memorie care a fost deja dealocată, rezultatul nu este definit și este dificil de prezis, iar programul este foarte probabil să devină instabil sau să se distruge. Acest lucru poate fi remediat parțial utilizând pointeri smart, dar aceștia adaugă o complexitate mai mare. De reținut este faptul că Garbage Collector-ul nu previne scurgerile de memorie, acele cazuri în care memoria este referențiată dar nu este folosită.

Procesul de colectare a resturilor poate să aibă loc în orice moment. De obicei, va porni atunci când un program este inactiv. Este garantat că procesul va fi declanșat dacă nu mai este suficientă memorie heap liberă pentru a crea un nou obiect. Acest lucru poate duce la o pauză a procesului de rulare a programului pentru un moment. Administrarea explicită a memoriei nu este posibilă în Java.

Java nu suportă operațiile aritmetice cu pointeri din C sau C++, unde adresele obiectelor și



valorile de tip `unsigned int` (sau `long int`) pot fi utilizate pentru interschimbare. Acest lucru alocă colectorului de deșeuri (garbage collector) să realoce obiectele referențiate și să asigure siguranța și securitatea.

Ca în C++ sau în alte limbaje de programare orientată obiect, variabilele tipurilor primitive de date din Java sunt stocate direct în câmpuri (pentru obiecte) sau în stack (pentru metode) și nu în zona de memorie heap. Această decizie a fost luată de designerii Java din motive de performanță.

Limbajul de programare Java conține diferite tipuri de colectoare de deșeuri (garbage collectors). În mod implicit, HotSpot folosește “parallel scavenge garbage collector”. În orice caz, sunt multe alte colectoare de deșeuri care pot fi utilizate pentru administrarea memoriei heap. Pentru 90% din aplicațiile Java, este suficient utilizarea colectorului de deșeuri “Concurrent Mark-Sweep (CMS)”. Oracle tinde să înlocuiască colectorul de deșeuri CMS cu colectorul “Garbage First Collector (G1)”.

Limbajul de programare Java folosește, în mod nativ, setul de caractere Unicode. Acesta este un standard internațional care a dus la înlocuirea codului ASCII (vechiul set de caractere). În setul Unicode, caracterele sunt reprezentate cu ajutorul a 2 octeți (față de codul ASCII unde se pot reprezenta numai 256 de caractere, setul Unicode suportă reprezentarea a unui număr maxim de 65536 de caractere). Primele 256 de caractere din setul Unicode sunt la fel ca și cele din codul ASCII. Pentru referirea celorlalte caractere se folosește sintaxa `%uxxxx`, unde `xxxx` reprezintă codul caracterului. O caracteristică importantă a setului de date Unicode o reprezintă și faptul că întreg intervalul de reprezentare a simbolurilor este împărțit în subintervale, purtând denumirea de blocuri. Exemple de blocuri: Basic Latin, Gothic, Greek, Arabic, Arrows, Currency, Mathematical, Musical, etc.

Majoritatea cuvintelor rezervate în Java sunt aceleași ca și în limbajul de programare C++. Acestea nu pot reprezenta variabile sau metode, nume de clasă sau interfață. Cuvintele `true`, `false`, `null` nu reprezintă cuvinte cheie, dar nici ele nu pot fi folosite pe post de nume în dezvoltarea de aplicații Java. Cuvintele marcate cu `*` nu sunt folosite, dar sunt rezervate.

<b>abstract</b>	<b>double</b>	<b>int</b>	<b>strictfp</b>
<b>boolean</b>	<b>else</b>	<b>interface</b>	<b>super</b>
<b>break</b>	<b>extends</b>	<b>long</b>	<b>switch</b>
<b>byte</b>	<b>final</b>	<b>native</b>	<b>synchronized</b>
<b>case</b>	<b>finally</b>	<b>new</b>	<b>this</b>
<b>catch</b>	<b>float</b>	<b>package</b>	<b>throw</b>
<b>char</b>	<b>for</b>	<b>private</b>	<b>throws</b>
<b>class</b>	<b>goto*</b>	<b>protected</b>	<b>transient</b>
<b>const*</b>	<b>if</b>	<b>public</b>	<b>try</b>
<b>continue</b>	<b>implements</b>	<b>return</b>	<b>void</b>
<b>default</b>	<b>import</b>	<b>short</b>	<b>volatile</b>
<b>do</b>	<b>instanceof</b>	<b>static</b>	<b>while</b>

## 2.2 MySQL

Un sistem de organizare a unei baze de date relaționale este o implementare necesară în foarte multe domenii, de la chestii tradiționale la afaceri, cercetări, contexte educaționale, la aplicații foarte puternice și complexe ( de exemplu, motoarele de căutare pe internet). În ciuda faptului că o bază de date este foarte importantă pentru a administra și accesa resursele informative, multe organizații s-au aflat în imposibilitatea financiară de a avea un asemenea sistem. De-a lungul timpului, sistemele de baze de date au fost o propunere destul de scumpă, furnizorii cerând sume importante de bani atât pentru partea de software cât și pentru partea de suport. Deoarece un sistem de baze de date necesită echipament hardware substanțial pentru a rula cu o performanță decentă, costul total a devenit și mai mare.

Situația este diferită în zilele noastre, atât din punct de vedere software cât și hardware. Sistemele mici desktop și serverele sunt ieftine și eficiente și este o dorință arzătoare a companiilor dezvoltatoare să scrie sisteme de operare cât mai rapide pentru aceste instrumente. Aceste sisteme de operare sunt disponibile gratuit peste tot pe internet sau uneori la prețul unui CD ieftin. Ele includ și câteva derivate Unix BSD (FreeBSD, NetBSD, OpenBSD) la fel și multe distribuții ale sistemului de operare Linux (Fedora, Debian, Gentoo, Suse și multe altele).

Producția de sisteme de operare gratuite a început odată cu dezvoltarea aplicațiilor gratuite ca și GCC, compilatorul GNU C. Aceste eforturi de a face software-ul disponibil oricui are nevoie de el fac parte din mișcarea Open Source. Proiectele Open Source au produs lucruri importante în industria software. De exemplu, Apache este cel mai utilizat server web pe internet. Perl, Python și Ruby sunt limbaje de scripting foarte utilizate și totodată foarte stabile, iar PHP este un limbaj care a devenit popular din cauza ușurinței cu care se pot scrie pagini web dinamice. Toate aceste lucruri sunt în contrast cu soluțiile aflate în proprietatea cuiva și au un preț foarte ridicat și nici măcar nu există accesul la codul sursă al acestor programe.

Software-ul pentru bazele de date a devenit și el mai accesibil iar sistemele Open Source de administrare a bazei de date sunt disponibile gratuit. Unul dintre acestea este MySQL, un sistem SQL de tip client-server utilizat la administrarea bazei de date relaționale și este originar din Scandinavia. MySQL include un server SQL, programe care să permită clientului să acceseze serverul SQL, programe administrative și interfețe pentru a scrie propriile programe.

Rădăcinile MySQL au prins naștere în anul 1979, când Michael Monty Widenius de la compania suedeză TcX a creat o aplicație pentru baze de date numită UNIREG. În anul 1994, compania TcX a început să caute un sistem de administrare a bazei de date relaționale cu o interfață SQL pentru a putea fi folosită în dezvoltarea aplicațiilor web. Au testat câteva servere comerciale, dar au descoperit doar faptul că erau prea lentu pentru tabelele foarte largi a companiei. De asemenea, au investigat și programul mSQL, dar acesta nu dispunea de multe proprietăți pe care compania TcX le-a marcat ca fiind necesare.

Prin urmare, Monty a început dezvoltarea unui nou sever. Interfața de programare a fost creată explicit să fie similară cu cea folosită de aplicația mSQL deoarece multe aplicații gratuite erau disponibile pentru mSQL, și folosind o interfață similară, toate acele aplicații gratuite puteau fi folosite și pentru MySQL cu un minimum efort de portabilitate,

În anul 1995, David Axmark de la compania Detron HB a început să îi stimuleze pe cei de la TcX să lanseze MySQL pe internet. David a lucrat, de asemenea, și la documentație și la posibilitatea ca MySQL să ruleze cu utilitățile GNU. MySQL 3.11.1 a fost lansat în toată lumea în anul 1996 sub forma unor distribuții binare pentru sistemele de operare Linus și Solaris. Astăzi, MySQL este disponibil în mult mai multe platforme atât sub formă binară cât și sub formă de cod sursă. Compania MySQL AB a fost formată pentru a pune la dispoziție distribuții MySQL atât pentru versiunea Open Source cât și pentru cele cu licență comercială și să ofere suport tehnic, servicii de monitorizare și training. În anul 2008, compania Sun Microsystems a achiziționat compania MySQL AB ( angajamentul pentru versiunea Open Source a continuat să

fie la fel de puternic deoarece planul companiei Sun era de a distribui majoritatea produselor lor sub forma Open Source).

Inițial, MySQL a devenit popular la scară globală pentru viteza și simplitatea sa. MySQL nu a fost lipsit de critici, mai ales din cauza faptului că îi lipsea multe caracteristici printre care tranzacțiile și suport pentru cheia străină. MySQL a continuat să se dezvolte adăugând nu doar acele caracteristici lipsă cât și altele, printre care amintim replicațiile, subinterogările, procedurile stocate, vederi și trigerr-e (declanșatoare). Aceste capabilități au dus MySQL pe tărâmul aplicațiilor enterprise (aplicații care sunt destinate mai degrabă companiilor și nu utilizatorilor individuali). Ca și rezultat, oamenii care odată considerau aplicațiile de tip "big iron" (calculatoare foarte mari, foarte rapide dar și foarte scumpe) potrivite pentru implementarea unui sistem de baze de date, acum acordă foarte mare atenție MySQL.

MySQL este portabil și rulează pe sistemele de operare comerciale (cum ar fi Mac OS X, Windows, HP-UX) și pe dispozitive hardware până la rangul de sus, reprezentat de serverele companiilor foarte mari. Mai mult decât atât, performanța sa concurează cu performanța oricărui sistem de baze de date și poate să găzduiască baze de date care conțin miliarde de rânduri. În domeniul afacerilor, MySQL continuă să își mărească prezența deoarece companiile descoperă că MySQL poate fi capabil să le gestioneze baza de date la o zecime de preț față de cât plăteau de obicei pentru licențele comerciale și suport.

MySQL face parte din imaginea care se dezvoltă în spatele nostru încontinuu, și aceasta este: sisteme de operare gratis care rulează pe componente hardware foarte puternice dar ieftine, punând puterea substanțială de procesare și capabilitățile în mâinile unor utilizatori individuali și ale unor afaceri, mai mult ca oricând, pe o varietate foarte largă de sisteme. Această coborâre a barierelor economice în domeniul calculatoarelor a făcut ca soluțiile foarte rapide de baze de date să ajungă la mai mulți oameni și la mai multe organizații decât oricând altcândva. Organizațiile care doar visau să aibă un sistem de administrare a unei baze de date relaționale, pot face acest lucru la un cost foarte scăzut. Acest lucru este adevărat și pentru utilizatori individuali.

Multe sisteme de administrare a bazei de date sunt disponibile gratuit sau la un cost foarte mic, printre care MySQL, PostgreSQL sau SQLite. Atunci când comparăm MySQL cu alte sisteme de acest tip trebuie să ne gândim la cele mai importante lucruri: performanță, suport, caracteristici disponibile, condiționare și restricționare a licenței și preț. Luând toate acestea în considerare, MySQL are foarte multe lucruri atrăgătoare de oferit:

- Viteza. MySQL este foarte rapid. Dezvoltatorii acestui sistem susțin că este cel mai rapid sistem de administrare a bazei de date relaționale.
- Ușor de utilizat. MySQL este foarte performant dar relativ simplu și este mult mai puțin complex de setat și administrat decât sistemele mai mari.
- Suport pentru interogări. MySQL înțelege limbajul SQL (Structured Query Language), limbajul standard pentru sistemele de baze de date moderne.
- Capabilitate. Serverul MySQL este multi-thread (mai mult fire de execuție) așa că mai mulți clienți pot să se conecteze la el în același timp. Fiecare client poate să folosească mai multe baze de date în același timp. Accesarea serverului se poate realiza interactiv folosind diferite interfețe care permit adăugarea de interogări și vizualizarea de operații: clienți din linia de comandă, browser-e web sau aplicații grafice. În completare, interfețe de programare sunt disponibile pentru multe limbaje, de exemplu: C, Perl, Java, PHP, Python și Ruby. Putem să accesăm MySQL și cu ajutorul aplicațiilor care suportă ODBC și .NET (protocoale dezvoltate de Microsoft). Acest lucru ne dă posibilitatea de a scrie pachetele noastre pentru client în vederea accesării informațiilor din baza de date.

- Conectivitate și securitate. MySQL este complet de tip rețea iar bazele de date pot fi accesate de oriunde există conexiune la internet, deci poți să împărtășești datele cu oricine, oriunde. Dar, în același timp, MySQL controlează accesul la informație așa încât o persoană care nu are dreptul de a vizualiza conținutul bazei de date nu va putea face acest lucru. Ca să ofere securitate și mai mare, MySQL suportă conexiuni criptate folosind protocolul SSL (Secure Sockets Layer).
- Portabilitate. MySQL rulează pe o varietate foarte mare de sisteme de operare de tip Unix și Linux, la fel și pe alte sisteme cum ar fi Windows și NetWare. MySQL rulează pe echipamente hardware de la servere foarte mari până la dispozitive care încap în palmă.
- Spațiu redus. MySQL este distribuit pe o dimensiune redusă, mai ales comparativ cu uriașele discuri de stocare ale altor sisteme de baze de date.
- Disponibilitate și cost. MySQL este un proiect disponibil de tip Open Source sub mai multe licențe. În primul rând, este disponibil sub licența celor de la GNU General Public License (GPL). Acest lucru însemnând că MySQL este disponibil fără niciun cost pentru majoritatea operațiilor necesare. În al doilea rând, organizațiile care preferă aranjamentele formale sau nu vor să fie legate de condițiile de licențiere a celor de la GPL, licențele comerciale sunt disponibile.
- Distribuție deschisă și codul sursă. MySQL este foarte ușor de obținut. Trebuie doar să folosești browserul web. Dacă nu înțelegem cum funcționează o caracteristică sau vrem să facem un audit de securitate, putem să obținem codul sursă și să îl examinăm. Dacă utilizatorul care inspectează codul sursă găsește ceva în neregulă, este recomandat să raporteze această eroare pentru a putea fi corectată.

Distribuția MySQL include următoarele unelte:

- Un server SQL. Acesta este motorul care pornește MySQL și oferă acces la baza de date.
- Programul pentru client și pentru alte utilizări. Acestea includ un program interactiv pentru client care permite acestuia să introducă direct interogările și să vizualizeze rezultatele. De asemenea sunt disponibile mai multe programe administrative care te ajută să rulezi MySQL împreună cu site-ul tău: un program permite administratorului să monitorizeze și să controleze serverul, alte programe permit acestuia să importe date, să își salveze datele la un interval de timp (proces numit back-up), să verifice tabelele de probleme și multe alte chestii.
- O librărie pentru client care să îi permită să scrie propriile programe. Se pot scrie programe client în limbajul C deoarece librăria este creată în C. Librăria poate fi corelată cu alte limbaje de programare precum Perl, PHP, Ruby pentru a oferi serviciile de bază pentru interfețele MySQL.

Adițional software-ului disponibil împreună cu MySQL, MySQL este utilizat de mulți oameni talentați și capabili cărora le place să scrie software pentru a-și mări productivitatea și care sunt disponibili să împărtășească acest program. Rezultatul este unul benefic, accesul la o varietate largă de unele care fac MySQL să fie mai ușor de utilizat sau de integrat în diferite arii, cum ar fi dezvoltarea de site-uri web.

## 2.3 Hibernate

Majoritatea proiectelor de dezvoltare software implică cel puțin o bază de date relațională. Baza majorității aplicațiilor comerciale este reprezentată de informație ordonată la scară mare, de exemplu: cataloage, liste de clienți, detalii de contracte, texte publicate sau design arhitectural.

Odată cu apariția World Wide Web ( rețea largă internațională), cererea pentru baze de date a crescut considerabil. Chiar dacă poate nu știu, clienții unei librării online, de exemplu, sau a unui ziar online necesită folosirea unei baze de date pentru stocarea lor. Undeva în adâncul aplicației, o bază de date este interogată și un răspuns al acesteia este oferit.

Hibernate este o bibliotecă care simplifică utilizarea bazelor de date relaționale în aplicațiile realizate cu ajutorul limbajului de programare Java. Acest lucru este posibil reprezentând datele relaționale ca fiind simple obiecte, accesate în urma unei sesiuni. Hibernate prezintă două tipuri de interfețe programabile: o interfață nativă Hibernate și The Java EE-standard Java Persistence API.

Sunt situații în care utilizarea Hibernate este foarte utilă, dar sunt și situații pentru care abordarea tradițională reprezentată prin accesul direct via The Java Database Connectivity (JDBC) Api este recomandată. Totuși, Hibernate rămâne favorită pentru prima opțiune deoarece nu exclude nici folosirea simultană a mai multor abordări, deși trebuie avută foarte mare grijă atunci când datele sunt modificate cu ajutorul a două API-uri diferite.

Dacă Hibernate este o soluție, atunci care este problema? Un răspuns la această întrebare comună ar fi faptul că atunci când facem exact aceleași lucruri folosind JDBC, avem nevoie de o cantitate considerabilă de cod și de observarea atentă a diferitelor reguli (de exemplu, administrarea conexiunilor) ca să ne asigurăm că aplicația noastră nu produce o scurgere de resurse.

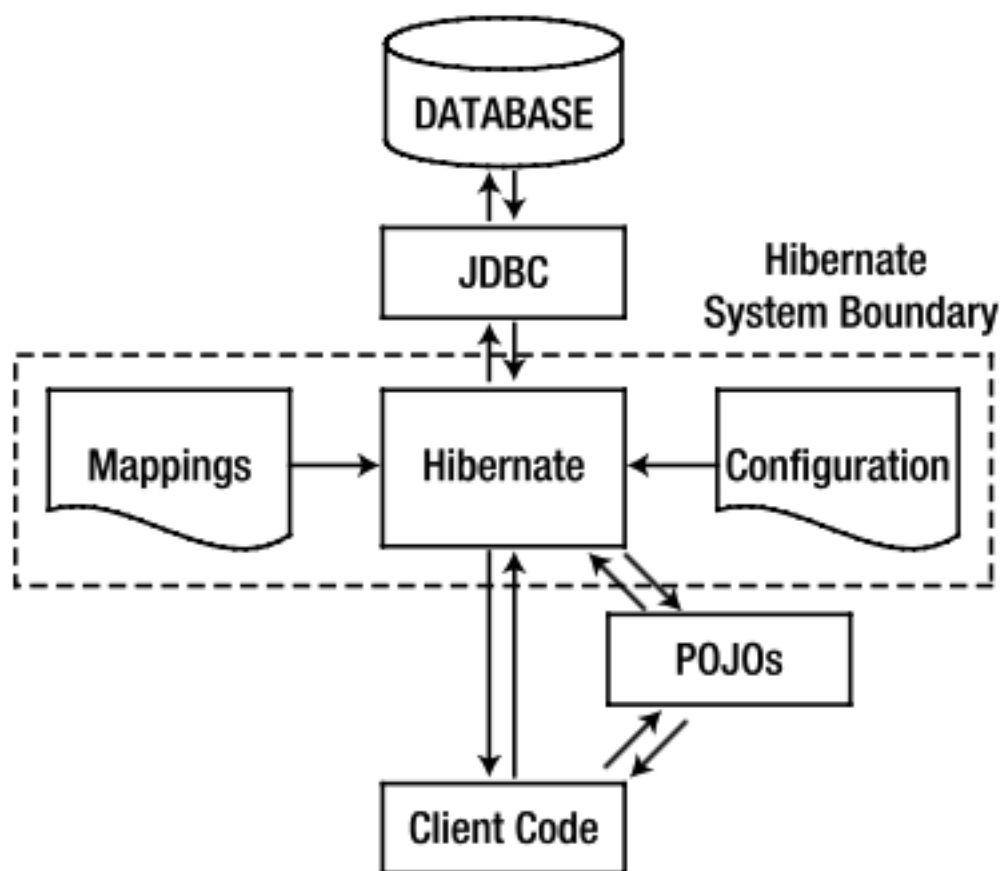
Într-o lume ideală ar fi neimportant să luăm orice obiect Java și să îl persistăm (continuăm) într-o bază de date. Nicio introducere de cod nou nu va fi necesară pentru a realiza această caracteristică, nu vor apărea nici penalități de performanță iar rezultatul va fi portabil în totalitate. În această lume ideală probabil că vom realiza operația respectivă în următorul mod:

```
POJO pojo = new POJO();  
ORMSolution magic = ORMSolution.getInstance();  
magic.save(pojo);
```

Nu vor apărea surprize neplăcute, nu va fi nevoie de muncă în plus pentru a corela clasele cu tabelele bazei de date iar performanța nu va fi afectată.

Hibernate este foarte aproape de această lume ideală, cel puțin atunci când este comparată cu celelalte alternative, dar este nevoie de crearea unor fișiere de configurare și trebuie luate în calcul și problemele de performanță și sincronizare. Hibernate, în orice caz, își atinge scopul fundamental: permiterea stocării de obiecte de tip POJO (Plain Old Java Object) în baza de date.

Termenul comun pentru persistența directă a obiectelor Java este ”obiect/ mapare relațională”, care înseamnă maparea obiectelor Java direct la entitățile relaționale ale unei baze de date. Următoarea imagine arată modul în care Hibernate se mulează într-o aplicație, de la codul clientului și până la baza de date:



Comparativ cu JDBC, Hibernate furnizează o administrare a resurselor mai curată, acest lucru însemnând faptul că nu mai trebuie să ne îngrijorăm de conexiunile actuale ale bazei de date sau de a introduce blocuri gigantice de cod de tipul "try/catch/finally". Totuși, erorile nu vor dispărea, doar că nu va mai trebui să ne ocupăm noi de ele. În orice caz, aceste condiții sunt excepționale (adică ne ocupăm doar de excepțiile de care trebuie să ne ocupăm, nu de fiecare excepție care ar necesita tratament).

Hibernate manipulează maparea obiectelor la tabelul bazei de date incluzând construirea schemei bazei de date dacă o configurăm să facă acest lucru. Nu necesită construirea unui tabel pentru fiecare tip de obiect, putem să mapăm foarte ușor un obiect la mai multe tabele. Și, de asemenea, Hibernate administrează și relațiile dintre obiecte, de exemplu: dacă am adăugat o listă de adrese unei persoane, putem foarte ușor să avem adresele stocate într-un al doilea tabel, construit și organizat de Hibernate.

În plus față de maparea obiectelor la tabelul unei baze de date, Hibernate poate să gestioneze mapările unor tipuri noi la baza de date.

Pornirea Hibernate durează puțin mai mult decât codul direct JDBC. Oricum, timpul de inițializare a sistemului nu este o măsură foarte semnificativă. Majoritatea aplicațiilor au perioade foarte lungi de rulare iar procentul de timp petrecut în procesul de inițializare a librăriei Hibernate este nesemnificativ comparativ cu performanța actuală a aplicațiilor. Avantajele Hibernate în întreținerea și administrarea obiectelor mai mult decât compensează timpul utilizat în configurare și inițializare. Ca de obicei, cel mai bun mod de a evolua performanțele unei aplicații este de a testa și a analiza aplicația propriu-zisă.

Orice obiect Java capabil de a fi persistat la o bază de date este un candidat pentru a fi persistat cu ajutorul Hibernate. Prin urmare, Hibernate este o înlocuire naturală pentru soluțiile ad-hoc (ca și JDBC de exemplu) sau un motor de persistare pentru o aplicație care nu are deja

încorporată o bază de date. În plus, prin alegerea Hibernate, utilizatorul nu este legat de nicio decizie (schimbarea designului, de exemplu) pentru obiectele aflate în aplicație, incluzând care bază de date să fie folosită pentru Hibernate.

După cum era de așteptat, Hibernate are nevoie de cineva să îi spună care tabel corespunde cărui obiect. În limbajul Hibernate, acest proces este numit mapare. Mapările pot fi implementate prin adnotări Java sau cu ajutorul unui fișier de mapare XML. Adnotările sunt preferate de majoritatea dezvoltatorilor deoarece oferă o imagine clară a structurii chiar la nivelul de cod. Hibernate dispune și de o abordare de tipul "configurare prin excepție" pentru adnotări: dacă suntem mulțumiți cu valorile pe care Hibernate ni le prezintă, nu mai trebuie să le precizăm explicit cu ajutorul adnotărilor. Pentru instanțe, Hibernate folosește numele clasei POJO ca și valoarea implicită a tabelului bazei de date pentru care se realizează maparea. De fapt, dacă accesul la baza de date se realizează doar prin intermediul Hibernate, nu trebuie să știm numele tabelului. Acest lucru este posibil deoarece interogările Hibernate se realizează pe tipul obiectului și nu pe numele tabelului. Hibernate construiește automat interogarea în așa fel încât este folosit numele corect al tabelului, chiar dacă ulterior schimbăm numele tabelului.

Integrarea Hibernate în aplicația Java se realizează relativ ușor. Designer-ii Hibernate au evitat câteva dintre cele mai comune capcane și probleme care existau cu soluțiile persistenței Java și au creat o arhitectură curată și puternică în același timp. Practic, acest lucru înseamnă că nu e nevoie să rulăm Hibernate în interiorul unui container sau schelet Java EE. Rularea Hibernate necesită doar instalarea versiunii Java 6 sau versiunile următoare.

La început, adăugarea Hibernate în proiectul Java poate să intimideze utilizatorul deoarece distribuția include un set foarte mare de librării. Pentru a avea prima aplicație Hibernate, trebuie setate referințele bazei de date și configurările Hibernate, care pot include maparea obiectelor la baza de date. De asemenea, trebuie creată și clasa de tip POJO împreună cu toate adnotările folosite în procesul de mapare. După ce au fost făcute toate acestea de mai sus, trebuie scrisă partea logică din spatele aplicației care utilizează Hibernate în vederea obținerii unor rezultate. Odată ce procesul de integrare Hibernate într-o aplicație Java este înțeles, chestiile de bază se aplică pentru orice proiect care folosește Hibernate.

Una dintre caracteristicile cheie a Hibernate este faptul că Hibernate nu interferează cu aplicația Java mai mult decât este nevoie, acest lucru a fost posibil datorită designer-ilor care au creat Hibernate.

Pentru a configura și integra Hibernate este necesar parcurgerea următorilor pași:

1. Identificarea claselor de tip POJO care trebuie reprezentate cu ajutorul bazei de date.
2. Identificarea proprietăților clasei POJO care trebuie persistate (mapate)
3. Adnotarea fiecărei proprietăți pentru a mapa obiectele Java la numele coloanelor bazei de date
4. Crearea schemei bazei de date folosind unealta pentru exportarea schemei, folosind o bază de date actuală sau crearea propriei scheme a bazei de date.
5. Adăugarea librăriilor Hibernate în aplicația Java.
6. Crearea unui fișier XML de configurare Hibernate care indică spre baza de date și spre clasele mapate acesteia.
7. În aplicația Java, trebuie creat un obiect de tip Hibernate Configuration care face referire la fișierul de configurare XML.

8. Tot în aplicația Java, trebuie creat un obiect de tipul Hibernate SessionFactory din obiectul Configuration.
9. Restabilește obiectele de tip HibernateSession din sesiunea FactorySession și scrie datele pentru accesul logic al aplicației ( crează, șterge, actualizează,etc).

Utilizatorul poate chema Hibernate direct din aplicația Java sau poate accesa Hibernate prin intermediul unei alte arhitecturi. Putem chema Hibernate prin intermediul unei aplicații de tip Swing, unui servlet, portlet, unei pagini JSP sau prin intermediul oricărei aplicații Java care are acces la o bază de date. De obicei, vom utiliza Hibernate pentru a crea un înveliș pentru accesul de date al aplicației sau pentru a înlocui un înveliș existent.

Hibernate suportă Java Management Extensions (JMX), J2EE Connector Architecture (JCA) și Java Naming and Directory Interface (JNDI). Folosind JMX, putem configura Hibernate în timp ce rulează. Hibernate poate fi lansat ca un conector de tip JCA și putem utiliza JNDI pentru a obține sesiunea curentă Hibernate. Adicional, Hibernate folosește driverele standard JDBC pentru a accesa baza de date relațională. Hibernate nu înlocuiește JDBC ca fiind un strat de conectivitate la baza de date. Hibernate stă cu un nivel mai sus de JDBC.

Adicional standardului JAVA APIs, majoritatea aplicațiilor Java de tip web integrează Hibernate. API-urile simple și curate Hibernate fac integrarea Hibernate ușor de suportat într-un fel sau altul. Scheletul Spring distribuie un excelent mod de a integra Hibernate, incluzând suport generic pentru obiectele persistente, un set generic de excepții și o administrare a tranzacțiilor.

Neținând cont de mediul în care este integrat Hibernate, anumite cerințe rămân constante. Trebuie definite detaliile configurației care se aplică. Aceste detalii sunt reprezentate, în pasul următor, de un obiect de tip Configuration. De la acest obiect se crează un singur obiect de tipul SessionFactory. De la acest obiect, obiectele de tip Session sunt instanțiate, obiecte prin care aplicația accesează reprezentarea Hibernate a bazei de date.

## 2.4 Apache PDFBox

Librăria Apache PDFBox este o librărie pentru limbajul de programare Java de tip Open Source care lucrează cu documentele de tip PDF. Această librărie permite crearea de documente noi de tip PDF, manipularea documentelor existente cât și abilitatea de a extrage conținutul documentelor. Apache PDFBox include, de asemenea, importante utilități pentru linia de comandă. Apache PDFBox este publicat sub licența Apache License v2.0

Apache PDFBox conține următoarele componente:

- PDFBox: partea de bază
- FontBox: se ocupă cu informațiile despre fonturi
- XmpBox: se ocupă de metadatele de tip XMP
- Prefligh (componentă opțională): verifică fișierele PDF în conformitate cu standardul PDF%-1a

Dezvoltarea PDFBox a început în anul 2002 în cadrul companiei SourceForge de către Ben Litchfield care a vrut să extragă textul din fișierele PDF pentru librăria Lucene. A devenit parte a proiectului Apache Incubator în anul 2008 iar în anul 2009 a ajuns unul dintre proiectele de top ale companiei. În anul 2015, ApachePDFBox a fost numit partener a organizației Open Source Partner Organization de către PDF Association.

Una dintre caracteristicile importante a librăriei Apache PDFBox este aceea de a extrage rapid și precis text dintr-o varietate de documente de tip PDF. Această funcționalitate este încapsulată în org.apache.pdfbox.util.PDFTextStripper și poate fi executat foarte ușor din linia de comandă.



Lucene este o librărie de căutare a textului de tip Open Source făcând parte din proiectul Apache Jakarta Project. Pentru ca Lucene să poată indexa un document PDF, trebuie prima dată să fie convertit sub formă de text. PDFBox dispune de o abordare simplă pentru a adăuga documentele PDF într-un index Lucene.

```
Document luceneDocument = LucenePDFDocument.getDocument( ... );
```

După ce s-a creat obiectul de tip Lucene Document, poate fi adăugat la indexul Lucene în același mod ca și când ar fi fost creat dintr-un fișier text sau HTML. LucenePDFDocument extrage automat o varietate de câmpuri de metadate din fișierul PDF care sunt adăugate indexului Lucene, javadoc afișând detaliile în acele câmpuri. Această abordare este foarte simplă și este suficientă pentru majoritatea utilizatorilor, în caz contrar se pot utiliza tehnici mai avansate de extragere a textului.

Unele aplicații conțin cerințe mai complexe de extragere a textului, nici linia de comandă sau LucenePDFDocument nu pot îndeplinescă acele cerințe. Este posibil ca utilizatorul să utilizeze sau să extindă clasa PDFTextStripper pentru a satisface o parte din cerințele respective.

Există destule moduri de a limita textul care este extras în timpul procesului de extragere. Cel mai simplu mod este de a specifica intervalul de pagini din care să se extragă textul.

Extragerea textului în limbile ale căror conținut merge de la dreapta la stânga (Arabă sau Ebraică, de exemplu) poate să rezulte în text care ordinea literelor este inversată. PDFBox poate să normalizeze și să inverseze acest text cu ajutorul fișierului de tip jar ICU4J. Trebuie de asemenea activată sortarea folosind `org.apache.pdfbox.util.PDFTextStripper` sau `org.apache.pdfbox.ExtractText` pentru a asigura un text de calitate.

Există mai multe librării de procesare a fișierelor PDF, printre care amintim Adobe PDF Library, iText, JPedal, PDFTron Systems, etc. Cea mai populară librărie este iText. Există o întreagă dezbateră pe tema "Care librărie este mai bună dintre PDFBox sau iText?", de aceea vom face o comparație între cele două:

- O diferență majoră o constituie faptul că PDFBox procesează textul simbol cu simbol pe când iText procesează textul ca fiind un simplu string. Acest lucru reduce semnificativ numărul de resurse folosite de librăria iText.
- Arhitectura orientată pe evenimente a librăriei iText atunci când parcurge un text înseamnă o povară mai mică pentru resursele iText față de resursele PDFBox.
- PDFBox stochează informațiile, care nu sunt necesare pentru extragerea textului, pentru o perioadă mai lungă de timp. Acest lucru este posibil datorită utilizarea mai multor resurse.
- Modul în care librăriile încarcă documentul este, de asemenea, o diferență. PDFBox oferă multiple moduri de supraîncărcare a funcției `PDFDocument.loads` dar și a funcției `PDFDocument.loadNonSeq`.
- iText dispune de o strategie mai simplă și mai avansată a parcurgerii unui text. Cea mai simplă strategie presupune ca textul să apară în ordinea citirii pe când cea mai avansată presupune sortarea textului. Librăria PDFBox sortează mereu textul.
- În ciuda dezavantajelor clare, PDFBox este recomandată atunci când manipularea fișierelor de tip PDF este minimă, iText având mult mai multe funcții și îngreunează puțin conținutul aplicației. Acest lucru m-a determinat să aleg și eu librăria PDFBox pentru a manipula textul din fișierele de tip PDF.