

FACULTATEA DE AUTOMATICĂ ȘI CALCULATOARE
DEPARTAMENTUL DE AUTOMATICĂ ȘI INFORMATICĂ APLICATĂ
Bd. Vasile Pârvan 2
300223 Timișoara
România
Tel./Fax.: +40 256 40 32 10; +40 256 40 32 14



Rezumatul tezei de doctorat

DISTRIBUTED MAILING SYSTEM (DMS)

elaborată de dl. ing. *Patrik Emanuel Mezo*

I. Actualitatea și oportunitatea tezei de doctorat

Comunicarea reprezintă un element important al societății. De la începutul existenței omenirii, individul uman a încercat să dezvolte noi metode de comunicare. Metodele de interacțiune socială s-au dezvoltat odată cu societatea. În prezent, cel mai utilizat mod de comunicare folosit este reprezentat de poșta electronică (mail sau email). Acest sistem de mail se bazează pe două elemente principale: Client și Server. Termenul de server este desemnat de clădiri specializate (dedicate) care servesc același scop: Server centralizat. Odată cu creșterea numărului de utilizatori ai sistemului de corespondență electronică crește și necesitatea spațiului de stocare. Soluțiile la această problemă constă în comprimare de date și/sau adăugarea de noi sisteme de calcul celor existente. Totodată numărul mare de sisteme de calcul necesită spații dedicate pentru depozitare și administrare. Prin continua evoluție a noilor tehnologii, s-au dezvoltat și alte căi care facilitează comunicarea participând totodată la prosperitatea societății.

Inițial contruit în scopul comunicării, internetul a devenit cel mai vast mediu de stocare pentru informații. Un studiu recent a demonstrat că dimensiunile acestuia s-au dublat în ultimii cinci ani, capacitatea internetului măsurând în prezent sute de exabytes (1 exabyte = 1024 petabytes = 1024² terabytes).

Conceptul sistemului de poșta electronică distribuit a fost realizat pe baza acestor statistici. Spațiul destinat stocării mail-urilor este deținut în totalitate de end-useri și nu de servere dedicate. Maniera structurii arhitecturale este cea de peer-to-peer (P2P), unde fiecare

sistem de calcul este tratat ca Server respectiv Client. Obiectivul acestei teze este de a prezenta o cale prin care fiecare sistem de calcul personal poate contribui la o singură aplicație devenind astfel parte componentă a acestui sistem.

II. Analiza de detaliu, pe capitole a tezei de doctorat

Construcția tezei cuprinde un număr de șapte capitole, conceptele de arhitectură tip P2P și cea de corespondență electronică vor fi prezentate gradual, teza fiind redactată integral în limba engleză. Primul capitol are un scop introductiv cu caracter general privind problemele și soluțiile abordate în această teză. În al doilea capitol, noțiuni generale privind protocoalele utilizate în comunicarea în rețea sunt prezentate. Al treilea capitol constituie primul aport original al autorului prin propunerea unei arhitecturi de tip P2P și prin analiza amănunțită privind statisticile de uptime ale participanților în rețea. Al patrulea capitol dezbate sistemele de poștă electronică, atât pe arhitecturi de tip Client/Server cât și pe cele de tip P2P, autorul subliniind defectele de implementare ale acestora. În capitolul cinci autorul dezvoltă o soluție de intercompatibilitate între sistemele de mailing bazate pe modelul Client/Server și cele pe modelul P2P. În capitolul șase autorul propune trei implementări privind sistemul de corespondență electronică, fiecare implementare evidențându-se distinctiv de celelalte privind securitatea, gradul de distribuție, scalabilitatea și echilibrul de operații. Capitolul șapte prezintă concluziile tezei, oferind o analiză a rezultatelor obținute.

Capitolul 1. Acest capitol este unul introductiv, prezentând cadrul actual și oportunitatea implementării temei propuse. Ca element de referință autorul ia în considerare evoluția internetului, prezentând și tendințele modelelor de implementare adoptate. Internetul poate fi clasificat din perspectiva mai multor considerente, cele evidențiate în această teză fiind reprezentate de apartenența sistemelor de calcul la unul din cele două categorii: utilizatori sau furnizori de servicii. Majoritatea serviciilor pe care internetul le oferă sunt implementate prin intermediul modelului Client/Server, serverul desemnând un cadru costisitor, pornind de la numărul sistemelor de calcul și clădirile special amenajate până la personalul instruit în acest scop. Costul necesar administrării acestui tip de serviciu este direct proporțional cu cererea și scalabilitatea implementării respective.

Modelul Client/Server, model adoptat ca standard de majoritatea aplicațiilor care beneficiază de comunicarea în rețea, desemnează utilizatorii de rând ca fiind acea parte a internetului menită să beneficieze de facilitățile oferite de furnizori. Pe de altă parte, modelul bazat pe rețele de tip Peer-to-Peer valorifică, în mod contradictoriu cu modelul Client/Server, sistemele de calcul personale ale utilizatorilor de

rând pentru a furniza servicii. Această implementare revendică un cost minimal pentru administrare, iar arhitectura rețelei este garantată doar de partea de software. Bineînțeles că această implementare are și defectele sale: cadru instabil din punct de vedere al securității, disponibilității și consistenței datelor, datorată faptului că timpul cu care participanții contribuie la rețea nu poate fi anticipat, iar în consecință nu există o standardizare a acestui concept.

Dezvoltarea unui sistem de corespondență electronică pe o rețea de tip P2P necesită multe domenii de cercetare. Autorul specifică în cadrul acestui capitol domeniile cercetate în acest scop: topologia rețelelor P2P, scalabilitatea sistemului de mailing, securitatea datelor, disponibilitatea și consistența datelor, generarea unui cadru stabil prin metode de anticipare a timpului cu care participanții contribuie la rețea și soluții de interoperabilitate cu sistemele actuale de mailing.

Primul capitol se încheie cu prezentarea unei structuri generale a tezei.

Capitolul 2. Acest capitol aprofundează noțiunile privind protocoalele utilizate în comunicarea în rețea. Autorul descrie gradual conceptul de rețea pornind de la definiția acesteia și finalizând cu viziunea sa asupra conceptului Peer-to-Peer. Pornind de la modelul de referință OSI (Open System interconnection Model), autorul descrie implementarea protocolului TCP/IP, unul din protocoalele ce stau la baza proiectării topologiilor de tip P2P. Deasemenea sunt prezentate particularități ale acestui protocol adoptate ca standard din prisma aplicațiilor ajunse la un stadiu de maturitate. Dintre acestea sunt menționate protocoalele utilizate în sistemele tradiționale de corespondență electronică (SMTP, POP, IMAP), management de transfer de fișiere (FTP), asignare de nume adreselor IP (DNS), etc.

Autorul descrie rețeaua de tip P2P ca fiind un nivel adițional la protocolul standard adoptat (în acest caz TCP/IP). Conceptul de Peer-to-Peer denotă o topologie virtuală dezvoltată deasupra celei fizice. Această topologie însă nu este fixă ca cea fizică, ci poate varia în funcție de implementarea adoptată la nivelul aplicației. Autorul subliniază necesitatea dezvoltării unui standard la nivel de protocol privind implementarea rețelelor de tip P2P.

Capitolul 3. Acest capitol debutează cu analiza rețelelor de tip P2P existente. O clasificare amănunțită este prezentată la începutul acestui capitol prin intermediul căruia autorul subliniază defectele și necesitatea de standardizare a acestor implementări. Deși s-au făcut progrese majore în acest domeniu, implementările curente tratând rețeaua de tip P2P la nivel de framework pentru aplicațiile ce urmează a fi implementate, autorul critică fiabilitatea acestor concepte din prisma lipsei de transparență dintre aplicații și aceste soluții arhitecturale tip P2P. În

acest context autorul propune o extensie fiabilă care se pretează la oricare din framework-urile analizate, facilitând astfel crearea cadrului personalizat, în funcție de cerințele aplicației care urmează a fi implementată. Această soluție a fost prezentată în cadrul conferinței CIT2011 (BDI, IEEE), confirmată cu tematică bazată strict pe arhitecturi de tip P2P.

Acest nou concept promovează ideea ca fiecare implementare arhitecturală tip P2P să beneficieze doar de resursele necesare selectate din cadrul numărului total de participanți, ca de exemplu putere de calcul, lățime de bandă, uptime, etc. Prin caracteristicile astfel selectate, noul concept facilitează crearea unor grupări ierarhice (module ierarhice) controlate și adresate prin intermediul unor fișiere de control din prisma aplicației care a inițiat crearea acestora. Prin această propunere autorul generează suportul necesar dezvoltării mai multor aplicații de tip P2P simultan, facilitând astfel un cadru comun ce permite standardizarea implementării la nivel de protocoale folosite, topologii, etc.

Capitolul continuă cu analiza timpului cu care fiecare participant contribuie în cadrul aplicațiilor de tip P2P. Prin intermediul acestor investigații, autorul poate estima și încadra sistemul de mailing într-o anumită categorie de aplicații, putând astfel anticipa media timpului de uptime a fiecărui participant. Deasemenea prin această investigație autorul propune un algoritm de anticipare a resursei de uptime prin care se poate face replicarea datelor în cadrul sistemului de mailing. Noul algoritm propus poate estima timpul într-o marjă de interval bine definită, această marjă fiind minimală în raport cu alți algoritmi dezvoltați până în prezent.

Capitolul 4. În acest capitol autorul prezintă sistemul tradițional de corespondență electronică cât și alte implementări curente bazate pe modelul P2P. Aportul autorului în cadrul acestui capitol este unul investigativ, el subliniind caracteristicile și defectele acestor implementări.

Sistemul tradițional de corespondență a ajuns la un stadiu matur, prin îmbunătățirea constantă a mai multor mecanisme, cum ar fi protocolul standard utilizat, topologii de rețea folosite în manipularea sarcinilor distribuite la nivel de cluster, necesitățile de resurse ale sistemelor și spații de depozitare. Deși s-au făcut îmbunătățiri, există de asemenea scenarii care depășesc actualul model tradițional de corespondență: probleme de accesibilitate atunci când gateway-ul se află în spatele unui link care a fost întrerupt sau suprasolicitat, problemele de stocare și procesare din cauza atașamentelor multiple ale mail-urilor. Nevoia constantă de scalare a acestor servicii în funcție de numărul mare a cerințelor survenite din partea utilizatorilor, presupune costuri mai ridicate în ceea ce privește clădirile dedicate și personal de

specialitate calificat pentru gestionarea acestora, topologii pentru manipularea sarcinilor distribuite, resurse de calcul, etc.

Sistemele de corespondență bazate pe arhitecturi Peer-to-Peer au fost dezvoltate ca un răspuns la problemele ridicate de către serviciile poștale tradiționale. Soluții recente au fost dezvoltate pe platforme P2P tip hibrid sau framework. Soluțiile dezvoltate pe platforme hibride prezintă o performanță mai bună în situații de uptime scăzut, utilizatorii conectându-se la rețea doar pentru a revendica beneficiile acelei aplicații. Această performanță se datorează faptului că soluțiile hibride s-au folosit de la început de sistemele de calcul care prezentau proprietăți ale resurselor mai ridicate decât celelalte noduri din rețea. Sistemele de corespondență dezvoltate prin adoptarea platformelor P2P tip framework prezintă o complexitate avansată în termeni de operații de implementare, folosire de lățime de bandă, implementare și dezvoltare la nivel de arhitectură de corespondență (topologie).

Autorul încheie capitolul subliniind toate aceste defecte, pregătind în același timp cititorul pentru soluțiile de mailing propuse în această teză.

Capitolul 5. Înainte de a-și prezenta modelele proprii de implementare ale sistemelor de corespondență, autorul propune o modalitate de intercompatibilitate între sistemele tradiționale și cele bazate pe rețele tip P2P. Contribuția autorului a fost valorificată în cadrul conferinței SIITME 2011 (BDI, IEEE).

Pentru a rezolva problema generată de incompatibilitatea dintre cele două abordări la nivel arhitectural, autorul a propus o interfață menită să ruleze ca un serviciu în spatele sistemului de operare. Mai multe probleme sunt ridicate de o astfel de interfață, una fiind faptul că sistemele de corespondență tip P2P trebuie să se ocupe de protocoalele interne pentru un anumit format standard. O altă problemă constă în modul în care se face referire la participanții din afara rețelei. Abordarea autorului presupune separarea protocolului standard RFC de cel de comunicare internă, facilitând interoperabilitatea și în situația în care unul din protocoale este modificat.

Capitolul 6. Acest capitol cuprinde trei propuneri distincte de sisteme electronice de corespondență, fiecare implementare desemnând una din contribuțiile autorului de-a lungul celor trei ani de cercetare : DMS (SIITME2011 – BDI, IEEE), HMAIL (SACI2012 – BDI, IEEE), DMAIL (ICIE2012 - ISI).

DMS. Constituie un sistem de corespondență electronică dezvoltat pe o platformă P2P tip hibrid. Această implementare beneficiază doar de resursele special selectate pentru a susține backbone-ul topologiei : super node – SN (sisteme de calcul cu o medie a resurselor care îndeplinește cerințele arhitecturii), ordinary node - ON (celelalte noduri

din rețea). Topologia virtuală este implementată la nivel de inele interconectate, fiecare inel îndeplinind în același timp proprietatea de comunicate. Fiecare comunitate este adresată la nivel geografic, evitând astfel suprasolicitarea lățimii de bandă pentru care rețelele tip P2P sunt foarte bine cunoscute. Aportul autorului este cel de a implementa operațiile aferente sistemului de corespondență electronică într-o manieră echilibrată pe unul din nivelele menționate : ON, SN, comunități interconectate. Sistemele de calcul reprezentate de ON sunt folosite pentru replicarea conținutului de e-mail. Replicarea e-mail-urilor are loc numai între noduri de același tip, SN - ul desemnând lista de replicare conform căreia aceasta poate avea loc la inițiativa unui participant pentru a efectua operația de trimitere (send new mail).

Deși această implementare este dezvoltată pe o arhitectură hibridă, unde fiecare operație la nivel de rețea este transmisă prin mesaje tip broadcast, autorul limitează bariera de propagare a acestora la nivel geografic. În această manieră, efectele secundare pot fi diminuate, arhitectura sistemului de corespondență beneficiind doar de nodurile apropiate geografic.

HMAIL. Această implementare beneficiază de facilitățile arhitecturii P2P dezvoltate în acest sens în capitolul trei. Distribuția participanților este făcută la nivel geografic, aplicația de corespondență controlând topologia prin intermediul unor fișiere de control. Modulele ierarhice sunt structurate pe două nivele, ultimul nivel facilitând stocarea propriu-zisă a e-mail-urilor (Inbox și Outbox). Unul din aportul autorului îl constituie nivelul adițional de ierarhie care facilitează un element de securitate adițional la sistemul de corespondență propus. Deasemenea prin intermediul aplicației sunt filtrate nodurile participante la rețea în scopul implementării operațiilor necesare pentru transmiterea, stocarea și recepționarea mail-urilor.

O a doua contribuție a autorului reiese din metoda de abordare a intercompatibilității cu sistemele tradiționale de corespondență. Sistemul astfel propus poate fi clasificat ca fiind standalone sau hibrid. În cazul standalone, acesta poate intercepta email-uri din sistemul tradițional cât și trimite în același format. În modul hibrid, acesta poate fi considerat ca fiind parte componentă a sistemului tradițional, mail-urile fiind trimise mai întâi spre partea tip P2P, iar dacă acesta este suprasolicitată, a doua alternativă o reprezintă destinația bazată pe modelul Client/Server. A doua modalitate reprezintă o soluție fiabilă pentru sistemele tradiționale, aceasta diminuând drastic costurile necesare pentru a scala acest sistem.

DMAIL. Constituie o implementare de sistem de corespondență dezvoltată pe o topologie tip P2P susținută de un framework studiat în această teză. Sistemul propus utilizează de toate facilitățile prezentate anterior în sistemele abordate. Aportul autorului în cazul acestui sistem îl constituie compatibilitatea cu echipamentele mobile din ziua de astăzi. Autorul propune o implementare arhitecturală prin intermediul căreia un utilizator poate desemna sistemul său de calcul ca fiind echipamentul

primar cu care el se conectează la rețeaua tip P2P. Odată desemnat sistemul de calcul primar, acesta poate fi folosit ca un gateway extern pentru celelalte echipamente cu care autorul se conectează la internet : tabletă, telefon mobil, etc.

Capitolul 7. Capitolul final al tezei este dedicat analizei gradului în care obiectivele enunțate în primul capitol au fost atinse, precum și evidențierii contribuțiilor originale ale autorului. Enumerarea acestora demonstrează un efort semnificativ al domnului inginer Patrik-Emanuel Mezo în direcția punerii la punct a unor noi concepte și arhitecturi, contribuțiile originale fiind prezente în capitolele trei, cinci și șase ale tezei, care sunt și cele de consistență.

Teza se încheie cu sumarizarea, în ansamblu, a stadiului în care se găsește proiectul *Sistem Distribuit de Mailing* în momentul de față, din punct de vedere arhitectural, urmată de prezentarea unor posibile direcții de dezvoltare.

III. Concluzii finale

Teza de doctorat elaborată de dl. ing. Patrik Emanuel Mezo reprezintă o lucrare științifică ale cărei contribuții au făcut obiectul unor prezentări la conferințe de prestigiu :

- **P. E. Mezo**, M. Vladutiu and L. Prodan, “Design of a Hierarchical based DHT Overlay P2P routing Algorithm”, 11th IEEE International Conference on Computer and Information Technology, Paphos, Cyprus, Aug. 2011, pp. 415 – 420, ISBN: 978-1-4577-0383-6 (BDI, IEEE rank).
- **P. E. Mezo**, M. Vladutiu and L. Prodan, “Interoperability solution between Peer-to-Peer and Client-Server based mailing systems”, 2011 IEEE 17th International Symposium for Design and Technology in Electronic Packaging (SIITME), Timisoara, Romania, Oct. 2011, pp. 45 – 48, ISBN: 978-1-4577-1276-0. (BDI, IEEE rank).
- **P. E. Mezo**, M. Vladutiu and L. Prodan, “Distributed Mailing System (DMS)”, 2011 IEEE 17th International Symposium for Design and Technology in Electronic Packaging (SIITME), Timisoara, Romania, Oct. 2011, pp. 349 – 354, ISBN: 978-1-4577-1277-7. (BDI, IEEE rank).
- **P. E. Mezo**, M. Vladutiu and L. Prodan, “HMail: A hybrid mailing system based on the collaboration between traditional and Peer-to-Peer mailing architectures”, 2012 IEEE 7th International Symposium on Applied Intelligence and Informatics (SACI), Timisoara, Romania, May. 2012, pp. 255 – 260, ISBN: 978-1-4673-1014-7. (BDI, IEEE, Australian Research Council list class C rank).
- **P. E. Mezo**, M. Vladutiu, L. Prodan and F. Opritoiu, “DMail: Distributed mailing system based on the collaboration between traditional and Peer-to-Peer mailing

architectures”, 2012 International Conference on Information Engineering, Lecture Notes In Information Technology, Vol. 25, Singapore, Singapore, Jun. 27-28, pp. 128 -135, ISBN: 978-1-61275-024-8. (Ei Compendex, Cambridge Scientific Abstracts, Google Scholar, IEE, ISI rank).

Valoarea lucrării rezultă din complexitatea problematicii abordate, precum și din contribuțiile originale ale autorului, atât de factură conceptuală, cât și de factură arhitecturală. Elaborarea tezei a necesitat un volum important de muncă, acompaniat de o foarte bună informare în domeniu. Autorul tezei demonstrează atingerea obiectivelor specifice unui doctorat: extragerea și sinteza de cunoștințe noi în contexte noi, dezvoltarea și implementarea unor strategii și soluții inovative, precum și capacitatea de evaluare a rezultatelor obținute.