UNIVERSITATEA POLITEHNICA DIN BUCUREŞTI

Şcoala Doctorală de Electronică, Telecomunicații și Tehnologia Informației

Decizie nr. 778 din 03-12-2021

REZUMAT TEZĂ DE DOCTORAT

Ing. Radu Marius BONCEA

CONTRIBUȚII LA CREȘTEREA SCALABILITĂȚII ȘI FIABILITĂȚII INTERNETULUI OBIECTELOR

CONTRIBUTIONS TO IMPROVE SCALABILITY AND RELIABILITY OF THE INTERNET OF OBJECTS

COMISIA DE DOCTORAT

Prof. dr. ing. Ion MARGHESCU Univ. Politehnica din Bucureşti	Preşedinte
Prof. dr. ing. Ioan BACIVAROV Univ. Politehnica din Bucureşti	Conducător de doctorat
Prof. dr. ing. Mircea POPA Univ. Politehnica din Timişoara	Referent
Prof. dr. ing. Gheorghe ŞERBAN Universitatea din Piteşti	Referent
Prof. dr. ing. Paul ŞCHIOPU Univ. Politehnica din Bucureşti	Referent

BUCUREŞTI 2022

Cuprins

Capitolul 1

Introducere

1.1 Prezentarea domeniului de doctorat

Internetul Obiectelor (IoT) se referă la legătura strânsă dintre lumea digitală și cea fizică [1, 2], evoluând într-un sistem complex care utilizează mai multe tehnologii, de la Internet la comunicațiile wireless, de la sistemele microelectromecanice la sistemele încorporate (*embedded*). De-a lungul timpului, au fost propuse mai multe definiții pentru a descrie ce înseamnă în esență un Internet al Obiectelor, însă, dacă ar fi să diluăm toate aceste definiții, am concluziona că IoT este un sistem complex compus din elemente interconectate precum dispozitivele de calcul, mecanice sau digitale, dispuse ierarhic și care activează servicii digitale, având la bază interacțiuni și procese definite și îmbunătățite continuu.

În 2018, numărul de dispozitive conectate la Internet și utilizate în întreaga lume a depășit cifra de 17 miliarde, din care, numărul de dispozitive IoT este aproximativ 7 miliarde (sunt exluse smartphone-urile, tabletele, laptop-urile sau telefoanele fixe) [3]. Creșterea globală a conexiunilor este determinată, în principal, de dispozitivele IoT - atât pe partea consumatorilor (de exemplu, Smart Home), cât și pe segmentul B2B. Numărul de dispozitive IoT active este de așteptat să crească până la 10 miliarde până în 2020 și 22 miliarde până în 2025. Acest număr de dispozitive IoT include toate conexiunile active și nu ia în considerare dispozitivele care au fost cumpărate în trecut, dar nu mai sunt utilizate.

1.2 Scopul tezei

Deși tehnologiile de bază ale Internetului Obiectelor (IoT) au avansat rapid în ultimii ani, implementările la scară largă a soluțiilor IoT sunt apariții rare în peisajul tehnico-economic. Acest lucru se datorează în principal unor provocări cheie, cum ar fi constrângerile energetice, securitatea, scalabilitatea, interoperabilitatea și comunicațiile.

În acest context, prezenta teză are ca scop identificarea provocărilor deosebite care au impact asupra adopției Internetului Obiectelor și implementării de sisteme IoT de mari dimensiuni, identificarea cerințelor de nivel înalt și a soluțiilor integrabile

conform unei arhitecturi de referință propuse și validarea supozițiilor enunțate prin implementarea și operarea soluțiilor identificate.

Scopul poate fi rezumat în următoarele obiective:

- Identificarea provocărilor majore în implementarea sistemelor IoT complexe. Înțelegem prin sisteme IoT complexe acele sisteme care pot constitui baza suport pentru dezvoltarea de aplicații și servicii deschise și integrabile de către alte sisteme.
- Analizarea stadiului actual în ceea ce privește arhitecturile și modelele de referință, standardele și recomandările.
- Definirea cerințelor de nivel înalt care adresează provocările identificate și a arhitecturii de referință pe baza căreia se pot implementa sisteme IoT.
- Analizarea tehnologiilor și modelelor candidat în rezolvarea problemelor identificate și care pot fi utilizate în dezvoltarea sistemeleor IoT.
- Implementarea arhitecturii de referință ca și studiu de caz și aplicarea modelelor și metodelor propuse.

1.3 Conţinutul tezei

Teza este structurată în 7 capitole și 8 anexe.

Capitolul 1 face o introducere în domeniul tezei de doctorat, unde se urmărește o analiză a tendințelor actuale în ceea ce privește adopția și implementarea sistemelor IoT. Sunt enumerate de asemenea câteva definiții ale Internetului Obiectelor care stau la baza definirii scopului acestei teze.

În **capitolul 2** este prezentat un studiu cu privire la arhitecturile de referință petru sistemele IoT. Sunt analizate modelele considerate mature, cu largă adopție în implementare de către Industria 4.0. Sunt de asemenea identificate provocările majore în adopția IoT precum constrângerea energetică, securitatea și confidențialitatea datelor, comunicațiile, scalabilitatea și interoperabilitatea.

În **capitolul 3** sunt formulate cerințele de nivel înalt, funcționale și nonfuncționale, care adresează în mod specific provocările identificate. Aceste cerințe stau la baza propunerii unei arhitecturi de referință pentru o platformă de monitorizare și operare automată a sistemelor IoT, de tip AIOps, ierarhizată pe 3 nivele: sursele de date, zona de retenție și fuziune a datelor și nivelul analitic.

Capitolul 4 prezintă un studiu aprofundat despre tehnologiile suport în implementarea sistemelor de monitorizare și operaționalizare automată, grupate pe baza domeniului de aplicare: baze de date orientate pe serii de timp, sistemele de configurare și descoperire a serviciilor, sisteme de procesare a volumelor mari de date, sisteme de intermediere a mesajelor și soluții de management a resurselor. În introducerea capitolului se face o comparație între cele două metode de monitorizare a

infrastructurilor IT: metoda tradițională care se bazează pe programarea și codarea evenimentelor și a proceselor de monitorizare; metoda "inteligentă" care inferează evenimente în urma analizării datelor colectate din sistem.
Sunt prezentate de asemenea două metode de selecție optimală a tehnologiilor concurente: metoda analizei maturității folosind modele MADM (Multi Attribute Decision Making), cu un exemplu furnizat pentru bazele de date orientate pe serii de timp; metoda comparativă pe baza performanțelor computaționale și a gradului de utilizare a resurselor, cu un exemplu furnizat pentru bazele de date de tip cheie-valoare.
În încheierea capitolului este prezentat modelul arhitectural bazat pe microservicii și supun atenției oportunitatea integrării tehnologiei Blockchain în arhitectura de referință AIOps.

În **capitolul 5** este prezentat un studiu cu privire la algoritmii de învățare automată, supervizată sau nesupervizată, care pot fi utilizați în rezolvarea unor probleme precum clasificarea, agruparea, regresia sau reducerea dimensionalității. Pentru fiecare algoritm, studiul se concentrează pe descrierea modelului matematic, identificarea avantajelor și dezavantajelor, precum și problemele pe care le rezolvă.

Capitolul 6 prezintă implementarea serviciului AIOps la nivelul infrastructurii de calcul al ICI București și monitorizarea centrului de cloud computing ICIPRO și a registrului român de domenii .ro, RoTLD. Sunt oferite două exemple de servicii monitorizate, serviciul Whois și serviciul DNS, pentru care am propus, ca problemă, identificarea și catalogarea atacurilor cibernetice care au loc la nivelul celor două servicii. În încheierea capitolului sunt prezentate două probleme identificate în urma monitorizării pe baza datelor colectate în timp și sunt propuse soluții.

Capitolul 7 concluzionează teza de doctorat prin enumerarea rezultatelor obținute și a contribuțiilor originale precum articolele științifice publicate în reviste jurnal sau conferințe și proiectele în care am fost implicat. În încheiere sunt prezentate perpectivele de dezvoltare ulterioară.

Capitolul 2

Stadiul actual al tehnologiilor IoT

2.1 Arhitecturi de referință IoT

În acest capitol este prezentat un studiu larg elaborat cu scopul identificării provocărilor majore în implementarea sistemelor IoT de mari dimensiuni. În acest sens au fost studiate arhitecturile de referință cele mai utilizate în Industria 4.0, domeniul cu cea mai mare maturitate în aceea ce privește standardizarea proceselor și elementelor constitutive ale unui sistem IoT. Arhitecturile studiate sunt: RAMI - *Reference Architecture Model for Industrie 4.0*, IIRA - *Industrial Internet Reference Architecture*, IDS-RAM 3.0 - *IDS Reference Architecture Model*, BDVA - *Big Data Value Association*, Edgecross, IVI - *Industrial Value Chain*, OpenFog Consortium, Ocean Protocol, X-Road și Fiware. În baza acestui studiu, am propus o arhitectură de referință generală pentru sistemele IoT [4], care poate fi văzută ca un denominator comun al arhitecturilor studiate. Astfel, arhitectura propusă are 4 nivele: nivelul marginal sau *Edge* este nivelul unde se găsesc dispozitivele IoT și unde sunt generate datele în formă brută; nivelul *Gateway* este nivelul unde se face colectarea datelor și trasmiterea lor; *Platformei Cloud* este nivelul unde sunt stocate și analizate în context datele; nivelul de prezentare este nivelul unde datele sunt prezentate sub formă de informații acționabile.

2.2 Provocări tehnologice în implementarea IoT

Cea mai importantă provocare identificată este constrângerea energetică, care afectează nivelul marginal. Metodele de recoltare a energiei și care sunt prezentate în detaliu în acest capitol, pot fi utilizate pentru a genera cel mult energie de ordinul sutelor de $\mu W/cm^2$, energie insuficientă pentru a susține arhitecturi software sau hardware complexe, precum stiva tehnologică IP. Evoluția curentă a tehnologiilor de recoltare a energiei, în lipsa unei singularități tehnologice, conduc spre implementarea unor sisteme minim viabile și fiabile de a efectua măsurători discrete și de a le comunica pe distanțe scurte folosind protocoale ușoare, precum comunicațiile în radiofrecvență. Comunicațiile în sine sunt o provocare, pentru că, așa cum am arătat în teză, nu există tehnologii care să aibe consum energetic mic, rază de acoperire mare, rate de transfer mari și cost redus de implementare și operare. În realitatea economică și tehnologică actuală, se poate observa că, pentru sistemele IoT de mari dimensiuni, implementarea optimală presupune costuri financiare reduse și toleranță

agreată în ceea ce privește reducerea performanțelor (reducerea ratelor de transfer, reducerea razei, consumului energetic sau, cel mai adesea, o combinație a celor trei).

O altă provocare este scalabilitatea sau capacitatea sistemului IoT de a acomoda într-un mod continuu un număr din ce în ce mai mare de elemente, indiferent de nivelul arhitecturii. De exemplu, o creștere a numărului de dispozitive se reflectă într-o creștere a necesarului computațional la nivel de gateway și de platformă cloud. Există câteva instrumente care adresează această provocare, descrise în detaliu în acest capitol, precum procesul de bootstrapping automatizat, controlul sistemului de tip pipeline pentru procesarea Big Data, principiul scalării pe cele 3 axe, integrarea arhitecturii orientate pe microservicii și adoptarea tehnologiilor de stocare a datelor.

Securitatea și confidențialitatea datelor reprezintă de asemenea o constrângere serioasă, mai ales dacă ne referim la scopul sistemelor IoT de a transforma date în informații acționabile, iar coruperea datelor poate avea ca urmări luarea unor decizii incorecte, cu efecte negative[5, 6, 7]. Vulnerabilitățile cheie sunt descrise în teză și includ: identificare, localizarea și urmărirea, profilarea, violarea confidențialității, tranzițiile ciclului de viață, inventarul datelor și conectarea datelor.

Interoperabilitatea este o altă provocare deosebită, fiind asociată cu lipsa standardelor de referință deschise sau având caracter general, agreate la nivel de industrie. Cea mai bună modalitate de a evita această provocare este pregătirea încă de la început a rețelelor IoT pentru interoperabilitate. Peisajul IoT, fiind extrem de fragmentat, se vor aborda trei reguli de bază pentru conectivitatea IoT care facilitează proiectarea rețelei: adoptarea standardelor industriale deschise, se vor adopta tehnologii cu accent pe software, se vor folosi interfețe deschise.

Capitolul 3

Arhitectura de referință

3.1 Cerințele de nivel înalt

În acest capitol este propus un model de arhitectură de referință care adresează provocările identificate în implementarea sistemelor IoT de mari dimensiuni, precum constrângerea energetică, comunicațiile, securitatea și confidențialitatea datelor, scalabilitatea și interoperabilitatea. Astfel, a fost formulat un set de cerințe de nivel înalt, care are scop ghidarea în cadrul procesului de implementare a funcționalității unui ecosistem IoT:

1. **Protocol agnosticitatea** este un principiu conform căruia sistemul nu poate fi "captiv" unui protocol anume, ci, folosind metode și modele specializate precum metoda adaptorului sau metoda intermediarului (*proxy*), sistemul integrează nativ toate protocoalele de colectare a datelor, sau cel puțin protocoalele cele mai utilizate.

2. **Structurare semantică** presupune descrierea elementelor și a serviciilor sistemlui IoT folosind ontologii standardizate, precum SSN (*Semantic Sensor Network*), IoT-Lite, IoT-Stream sau SOSA (*Sensor Observation Sampling Actuator*).

3. **Agregarea, augmentarea și corelarea datelor** este o cerință cu privire la metodele de agregare, sumarizare și filtrare a datelelor colectate de la surse multiple.

4. **Procesarea datelor și învățarea automată** reprezintă procesul de translatare a volumelor mari de date brute colectate din sistem (senzori, date metrice legate de starea sistemului, date augmentate și asociate) în informații acționabile.

5. **Orchestrarea automată și provizionarea** presupune implementarea capabilităților necesare administrării și susținerii proceselor utilizate de serviciile digitale, într-un mod automatizat și sigur, contribuția factorului uman fiind supervizarea cu scopul îmbunătățirii proceselor, identificarea anomaliilor neclasificate și remedierea acolo unde automatizarea eșuează sau este incompletă.

6. **Arhitectură deschisă** facilitează înlocuirea sau extinderea tehnologiilor folosite.

3.2 Model de referință

Modelul referință pentru o platformă AIOps (*Artificial Intelligence Operations*) este dezvoltat ținând cont de setul de cerințe funcționale de nivel înalt și are la bază sursele de date inventariate în cadrul companiei: datele de tip metric generate de dispozitivele IoT și senzori; datele metrice asociate cu performanța sistemului, culese la nivel de sistem de operare; datele metrice asociate cu performața aplicațiilor și serviciilor; date privind suportul tehnic (*ticketing*); surse de inteligență sau date conectate și profilabile având ca sursă *social media*, bloguri și forumuri.

Următorul nivel al arhitecturii este nivelul unde se depozitează datele.Într-o primă fază, datele se salvează într-o bază de date optimizată pentru retenție scurtă (de regulă câteva zile sau ore), precum InfluxDB sau Prometheus. Aceste date sunt folosite pentru analize în timp real.
După trecerea perioadei de retenție, datele sunt *împinse* către baze de date cu retenție lungă, precum OpenTSDB. Aceste date istorice sunt folosite pentru antrenarea algoritmilor de învățare automată și pentru corelarea cu evenimente din trecut. Sunt o permanentă sursă de optimizare a proceselor AIOps și oferă o mai bună înțelegere a sistemelor complexe.

Nivelul al treilea este sistemul de analiză a datelor și nivelul unde datele sunt transformate în informații, folosind algoritmi de învățare automată sau de inteligență artificială, algoritmi specializați în rezolvarea problemelor de clasificare, grupare (*clustering*), regresie și reducere a dimensionalității.

Capitolul 4

Tehnologii suport

4.1 Monitorizarea

Implementarea arhitecturii de referință pentru o soluție de tip AIOps presupune integrarea unui set de tehnologii și soluții specializate. Soluția finală are ca cerință asigurarea proceselor de monitorizare inteligentă, cu scopul integrării operațiunilor IT automatizate, având la bază modele de învățare automată. Platformele AIOps combină tehnologia BigData și modelele de învățare automată, cu scopul de a facilita operarea proceselor de bază într-o infrastructură IT de mari dimensiuni. Platforma AIOps funcționează prin ingestia scalabilă și analiza volumului de date în continuă creștere. Aceste date, se caracterizează printr-o mare varietate și viteză de generare. Platforma permite utilizarea simultană a mai multor surse de date, metode de colectare a datelor și tehnologii analitice și de prezentare.

4.2 Baze de date orientate pe serii de timp

În această secțiune au fost studiate cele mai utilizate baze de date orientate pe serii de timp, precum Prometheus, InfluxDB, OpenTSDB, TimescaleDB. Au fost comparate performanțele soluțiilor și am propus o metodă de evaluare a maturității tehnologice pe baza unui model de tip MADM (*Multi Attribute Decision Making*) [8], care ține cont de un set larg de atribute calitative și cantitative, precum: interesul pieței cuantificat prin numărul de stele Git, numărul de apariții pe StackOverflow, scorul GoogleTrends; dezvoltarea activă (numărul de iterații pe codul sursă, numărul de contributori la cod); ritmul de dezvoltare (timpul mediu de rezolvare a bug-urilor sau capabilităților noi, ritmul de versionare;documentația și instrumentele. Metoda de evaluare a maturității soluțiilor TSDB reprezintă un instrument util în procesul decizional de selecție a stivei tehnologice, atât din perspectivă tehnică, cât și din perspectiva de afaceri. Metoda poate fi profilată și pentru selecția altor tehnologii sau servicii [9, 10].

4.3 Soluții de configurare și descoperire a serviciilor

Descoperirea serviciilor (*service discovery*) constă în detectarea automată a serviciilor și dispozitivelor într-o rețea. Această capabilitate a devenit critică pentru sistemele complexe și într-un mediu de calcul omniscient (*pervasive computing*). Au fost studiate tehnologiile Etcd, Apache Zookeeper și Consul și au fost comparate performanțele din punctul de vedere al resurselor computaționale: gradul de utilizare a procesorului, memoriei, rețelei, latența și ratele de scriere și citire.. Sunt propuse două modele de implementare, modelul centrat pe server și modelul centrat pe client, fiind prezentate condițiile optime de implementare, precum și cerințele de nivel înalt asociate cu sistemele complexe: stocarea durabilă și consistentă, alegerea liderului și atomicitatea operațiilor.

4.4 Soluții de procesare a datelor

Ca și definiție, procesarea datelor reprezintă un proces automatizat de conversie a datelor, având formate diverse, în informații acționabile. Procesul presupune înregistrarea, analizarea, sortarea, sumarizarea, executarea de calcule, diseminarea și stocarea. Deoarece datele procesate sunt utile atunci când sunt prezentate utilizatorului într-un mod intuitiv și folositor ca informație, sistemele care procesează date sunt cunoscute ca sisteme de informații. Sunt studiate două metode de procesare, procesarea în serie (*batch processing*) și procesarea în flux continuu (*stream processing*) și sunt evaluate tehnologiile integrabile Apache Spark, Apache Storm și Apache Flink în funcție de de metodele de procesare prezentate.

4.5 Soluții de intermediere a mesajelor

Comunicarea și schimbul de informații între aplicații, sisteme și alte servicii se realizează printr-un software, numit broker de mesaje, cunoscut și sub numele de agent de intermediere a mesajelor sau middleware orientat pe mesaje (Message Oriented Middleware – *MOM*). Brokerii de mesaje se bazează pe o componentă numită coadă de mesaje care stochează informația până când aplicațiile se eliberează și o pot procesa. Mesajele sunt stocate în ordinea exactă a intrării și rămân în coadă până la confirmarea primirii. Această funcție se mai găsește și sub denumirea de *FIFO* (First-In, First-Out). Coada de mesaje poate avea și manageri care gestionează interacțiunile dintre cozile multiple, traducerea mesajelor, servicii care furnizează rutarea datelor, dar și persistența și funcționalitățile de gestionare a stării clientului, adică a punctului final. Au fost studiate și analizate performațele tehnologiilor Apache Kafka, Apache ActiveMQ și RabbitMQ.

4.6 Soluții de management a resurselor

Una din cerințele de nivel înalt pentru a asigura scalabilitatea pe orizontala a sistemelor IoT este dată de capacitatea de a proviziona în mod automat noi resurse

computaționale atunci când se impune. Ne putem imagina scenariul în care platforma AIOps, în urma inferenței pe baza datelor metrice colectate și stocate în bazele de date orientate pe serii de timp, inferență care se bazează pe sistemele de procesare a datelor folosind algoritmi de învățare automată specializați, emite alerte cu privire la gradul înalt de utilizare a sistemului, alerte rutate de către brokerii de mesaje către sistemele de management a resurselor care pot adăuga noi resurse în funcție de tipul alertei și parametrii furnizați. Tehnologiile studiate sunt Kubernetes și Apache Mesos.

4.7 Microservicii

Arhitectura orientată pe microservicii este una din abordările importante în arhitectura software, fiind propusă dezvoltarea aplicațiilor ca o suită de servicii orientate pe funcționalități sau domenii distincte, unde fiecare serviciu este executat independent în cadrul propriului proces, comunicația fiind susținută de mecanisme simple, cum ar fi protocolul HTTP. Inclusiv dezvoltarea serviciilor poate fi făcută independent, atât ca poziționare în timp, cât și ca cerințe și procese de lucru. Putem spune că microserviciile sunt un stil arhitectural care structurează o aplicație ca și o colecție de servicii care sunt ușor de întreținut și testat, sunt decuplate, pot fi implementate independent și sunt organizate în jurul capabilităților de business [11].

Un serviciu web este, de regulă, asociat cu o componentă software sau un set de funcționalități, care, în colaborare, realizează o capabilitate. Un serviciu web este unic identificabil și poate fi accesat prin protocoale HTTP (SOAP, Restful). Diferența dintre un serviciu web și un microserviciu web ține de granularitatea componentizării, microserviciile fiind concentrate pe seturi restrânse de funcționalitate. Un concept important al arhitecturii orientate pe microservicii îl reprezintă compunerea serviciilor, care se poate realiza în două moduri: prin „orchestrare" sau „coregrafie".
Orchestrarea presupune controlul activ al tuturor elementelor și interacțiunilor, în timp ce coregrafia presupune stabilirea unui model sau rutină pe care microserviciile o urmează, fără a necesita supraveghere și instrucțiuni.

4.8 Blockchain

Securitatea și confidențialitatea datelor reprezintă a doua cea mai importantă piedică în calea implementării infrastructurilor IoT, prin provocări care includ identificarea, localizarea și urmărirea, profilarea, confidențialitatea, tranzițiile ciclului de viață, inventarul datelor și conectarea datelor. Deși în arhitecturile software ale aplicațiilor web există soluții pentru a adresa aceste provocări, în spațiul M2M unde accesul si procesarea datelor se face în mod automat, aceste soluții nu se pot aplica eficient și uniform, fiind nevoie de ajustări umane care sunt lente și conduc la blocaje. Având în vedere aceste constrângeri, în această secțiune este prezentată ca soluție, implementarea unui contract inteligent (*smart contract*) și înregistrarea lui într-un registru de tip blockchain [12].

Capitolul 5

Inteligența informațională

Definim inteligența informațională ca fiind o metodă tehnică utilizată pentru a transforma volume mari de date complexe în informații relevante și acționabile, cu scopul de a gestiona mai bine riscul și a crește productivitatea și implicit profitabilitatea. Metodele folosite au la bază algoritmi de învățare automată și instrumente de analiză a datelor și sunt utilizate pentru a rezolva un set de probleme clasice, precum: clasificarea sau catalogarea, agruparea sau *clustering*, regresia, reducerea dimensionalității și eliminarea zgomotului din date.

În acest capitol, am realizat un studiu al principalilor algoritmi și modele care pot fi utilizate în problemele enumerate. Algoritmii sunt comparați din perspectiva complexității și performanței și sunt asociați cu domeniile și problemele de aplicare optimală. Algoritmii au fost testați folosind seturi de date colectate de soluția AIOps implementată și au grupați în:

1. algoritmi de învățare automată supervizată: modele liniare (metoda celor mai mici pătrate, regularizarea Tikhonov, modelul Lasso, Multi-task Lasso, Elastic-Net, Least Angle Regression, Orthogonal Matching Pursuit, regresia Bayesiană, regresia Logistică, Gradient Stocastic Descendent, regresia Huber, regresia polinomială), analiza discriminanților liniari pătratici, mașini cu vectori suport, coborârea gradientului stocastic, metoda vecinilor cei mai apropiați, arbori de decizie.

2. învățare automată nesupervizată: modele mixte gaussiene, învățarea multiplă Manifold, asocierea isometrică, Locally Linear Embedding, Modified Locally Linear Embedding, Hessian Eigenmapping, Local Tangent Space Alignment, Multi-dimensional Scaling, t-distributed Stochastic Neighbor Embedding - t-SNE.

Inteligența informațională poate aduce beneficii și rezultate deosebite și în zonele mai puțin vizibile ale arhitecturii de referință propuse. Un exemplu este în zona de securitate cibernetică, unde dorim să identificăm comportamente anormale la nivel de interacțiune operator-sistem[13, 14]. Se mai poate folosi de asemenea în consolidarea cunoștințele echipelor DevOps prin oferirea de instrumente automatizate de documentare și testare[15, 16].

Capitolul 6

Implementarea experimentală

În acest capitol este descrisă implementarea unei platforme de tip AIOps, folosind arhitectura de referință și tehnologiile descrise în capitolele anterioare. Scopul inițial al acestei platforme a fost monitorizarea infrastructurii IT a registrului român de domenii .ro, ROTLD. Implementarea platformei a început în 2015 și, chiar de la început, s-a avut în vedere o dezvoltare iterativă continuă, cu interferențe minime la nivelul infrastructurii sau aplicațiilor existente. Scopul a fost extins un an mai târziu prin includerea monitorizării centrului de date și a serviciilor ICI PRO, în mod special, serviciile de tip IaaS. Sistemul astfel rezultat poate fi descris tehnologic ca având o infrastructură convergentă (hiperconvergentă începând cu 2020), virtualizată atât la nivel de rețea cât și la nivel de provizionare și administrarea a mașinilor de calcul virtualizate, folosind hipervizori VMWare și având în componență aproximativ 200 de mașini fizice de calcul și peste 1000 de mașini virtuale.

În ceea ce privește serviciile și aplicațiile monitorizate și specifice activății registrului de domenii .ro și serviciile de tip cloud computing, sistemul are în componență servicii DNS la nivel de domeniul înalt (*Top Level Domain*), serviciul Whois, DAS (*Domain Availability Service*), EPP (*Extensible Provisioning Protocoll*), interfețe programatice pentru registrari, servicii de date deschise și aplicații ale clienților (baze de date, aplicații web). Datele astfel colectate sunt furnizate platfomei analitice unde sunt procesate și evaluate în contextul detecției de anomalii și a evenimentelor. Informațiile obținute sunt furnizate mai departe nivelelor de decizie, precum SOC (*Security Operation Center*) - nivelul care adresează politica de securitate, nivelul DevOps - nivelul de control și remediere a incidentelor și nivelul de orchestrare, unde controlul și remedierea sunt automatizate.

Alegerea tehnologiilor suport folosite pentru implementarea platformei de tip AIOps s-a făcut pe baza unei analize a maturității și a performanțelor tehnice, astfel:

1. Analiza maturității se bazează pe un model multi-atribut [8], în care sunt evaluate proprietăți cantitative și calitative ale soluțiilor propuse.

2. Analiza performaţelor are la bază testarea funcțiilor principale prin simularea unui mediu de execuție în condiții de stres, urmărindu-se indicatori precum utilizarea resurselor și capacitatea de procesare.

Sistemul care se dorește monitorizat este implementat folosind ca stil arhitectural modelul *middleware*. Middleware este o componentă care agreghează funcțiile sistemului într-un set de noi funcții de nivel înalt, astfel încât aplicațiile client implementează strict integrarea cu middleware, fără a fi nevoie de integrarea cu fiecare element al sistemeului. De exemplu, în cazul nostru middleware oferă acces la funcția de înregistrare a unui nume de domeniu .ro. Aplicația client va trimite toate datele necesare înregistrării unui nume de domeniu la middleware, iar middleware va executa logica acestrui proces: se va conecta la baza de date, va face interogările și va înregistra domeniul cu datele asociate, va emite factură dacă este cazul, va înregistra evenimentul în sistemul de loguri, etc.
La nivelul de middleware a fost implementată interfața de publicare a datelor de tip metric asociate cu aplicațiile și care este folosită de Prometheus în vederea colectării datelor.

Există însă și situații în care, din motive de performanță sau de securitate, aplicații și serviciile sunt implementate direct pe modelul client-server, cum este cazul serviciilor Whois și DNS. Serviciul Whois este un serviciu care acceptă criterii de căutare (ex. numele unui domeniu .ro), execută căutări în baza de date și oferă informații privind disponibilitatea domeniului sau informații despre domeniu. Serviciul DNS oferă clienților posibilitea înregistrării de atribute specifice numelor de domenii, folosind protocoale asociate cu modificări dinamice sau oferă informații despre domenii pe baza unor criterii de căutare.

Serviciul Whois a fost implementat pe baza specificațiilor RFC3912 și este folosit pentru furnizarea de informații cu privire la numele de domenii. Un proces tipic de interacțiune cu serviciul Whois, presupune trimiterea de către o aplicație client a unui criteriu de căutare, precum numele de domeniu, iar serviciul va interoga baza de date și va furniza informațiile conform criteriului. Rotld are 4 servere asociate aplicației whois care rulează în orice moment și alte 4 servere în așteptare (*hot standby*) care sunt pornite automat atunci când numărul de interogări crește sau capacitatea celor 4 servere inițiale este depășită.
Pentru orice cerere ajunsă la aplicațiile Whois se efectuează verificarea existenței unui răspuns asociat în baza de date de cache. Dacă există răspunsul, aplicația va trimite acest răspuns către client. Dacă nu există răspunsul sau răspunsul a fost generat cu prea mult timp în urmă, aplicația Whois va trimite cererea către un load balancer al bazei de date cu înregistrările de domenii. Ca și aplicațiile Whois, există mai multe instanțe ale bazei de date.

Se impun câteva precizări. Există două tipuri de răspuns posibile oferite de serviciul Whois: în cazul în care domeniul există, serviciul returnează informații despre acesta, dacă domeniul nu este înregistrat, serviciul va returna o eroare specifică. Prin implementarea unui contor conform specificațiilor Prometheus, vom colecta numărul de răspunsuri pentru fiecare categorie de răspuns.

Ne propunem, pentru acest serviciu să identificăm, pe baza datelor metrice colectate din sistem și a răspunsurilor oferite, starea sistemului din perspectiva securității, prin stare înțelegând una din situațiile:

– Sistemul se află sub atac de tip *bruteforce*, caracterizat printr-un număr mare de cereri cu multe răspunsuri de tip eroare (interogarea de domenii inexistente).

– Atac de tip "țintit", caracterizat printr-un număr mare de cereri pentru domenii existente.

– Atac de tip "combinat", caracterizat printr-un număr mare de cereri atât pentru domenii existente cât și inexistente. De regulă acest tip de atac constă în generarea de nume de domenii pe bază de dicționar.

– Sistemul se află în starea normală dacă nu se regăsește in niciuna din stările anterioare.

În tabelul 6.1 avem un exemplu de date colectate necesare identificării stării sistemului Whois, în care, pe lângă numărul de răspunsuri de tip *Found* și *Not Found*, am colectat și gradul de încărcare al procesorului. Prin procesor înțelegem aici gradul de încărcare la niveul întregului sistem prin agregarea metricelor asociate cu fiecare procesor de pe fiecare server component al sistemului. Corelarea datelor din tabel poate fi vizualizată în figura 6.1, unde putem observa următoarele:

– Atacul de tip *bruteforce* se corelează cu un număr mare de răspunsuri *Not Found* și un grad de procesare relativ scăzut. Gradul de procesare scăzut se explică prin faptul că pentru răspunsurile de tip *Not Found* nu se execută interogări adiționale de obținere a datelor despre domenii.

– Atacul de tip *targetted* se corelează cu un număr mare de răspunsuri *Found* și grad înalt de procesare, în timp ce atacul combinat putem spune că are un număr mare de răspunsuri *Found* și *Not Found* și un grad de procesare mediu-înalt.

– Starea normală constă într-un număr mic de cereri și, prin urmare, un grad mic de procesare.

– În setul de date furnizat observăm două date catalogate greșit (outliers), punctul verde ar fi trebuit să fie galben, iar unul din punctele galbene ar trebui să fie verde.

Se pot găsi și alte corelații. De exemplu, pe baza data metrice colectate se poate identifica un atac de tip *slowloris*. Un atac slowloris presupune deschiderea a foarte multe conexiuni către server. fără a trimite însă date sau fără a formula cererea, epuizând bazinul de conexiuni pe care serverul le poate avea. Există desigur mecanisme de protecție, precum implementarea pe parte de server a unui mecanism de închidere automată a conexiunilor care așteaptă date mai mult de t secunde. Însă un atacator poate stabili de pe un simplu calculator sute de mii de astfel de conexiuni într-un interval foarte scurt. Iar dacă atacul devine și distribuit, cu mii de mașini care stabilesc zeci de mii de conexiuni, sistemul poate deveni indisponibil pe o perioada lungă de timp.

Found Hits	Not Found Hits	CPU	Label	Label Name
1.3	0.19	1.5	0	NORMAL
1.6	0.9	1.8	0	NORMAL
0.90	0.23	1.2	0	NORMAL
1.1	0.19	1.3	0	NORMAL
0.83	0.23	1.1	0	NORMAL
1.12	0.26	1.2	0	NORMAL
0.56	0.22	0.89	0	NORMAL
2.31	0.98	1.82	0	NORMAL
7.52	1.22	7.42	0	NORMAL
16.12	1.9	18.2	1	TARGETTED
14.32	6.42	16.32	1	TARGETTED
12.37	1.08	11.28	1	TARGETTED
9.12	0.98	11.23	1	TARGETTED
1.23	0.24	1.23	1	TARGETTED
2.05	0.14	2.17	1	TARGETTED
3.9	0.38	4.1	1	TARGETTED
4.3	1.2	4.9	1	TARGETTED
5.32	0.76	5.6	1	TARGETTED
0.53	11.21	3.8	2	BRUTEFORCE
1.3	12.65	4.5	2	BRUTEFORCE
0.48	18.32	5.45	2	BRUTEFORCE
0.65	12.14	3.83	2	BRUTEFORCE
0.49	11.40	3.57	2	BRUTEFORCE
0.98	4.7	4.23	2	BRUTEFORCE
12.21	11.59	13.99	3	COMBINED
11.20	13.19	12.95	3	COMBINED
9.83	12.97	11.74	3	COMBINED
11.42	11.46	12.67	3	COMBINED
11.20	16.32	13.02	3	COMBINED
12.35	11.21	12.98	3	COMBINED

Table 6.1: Exemplu de set de date folosit pentru învățare automată în cazul monitorizării stării serviciului Whois.

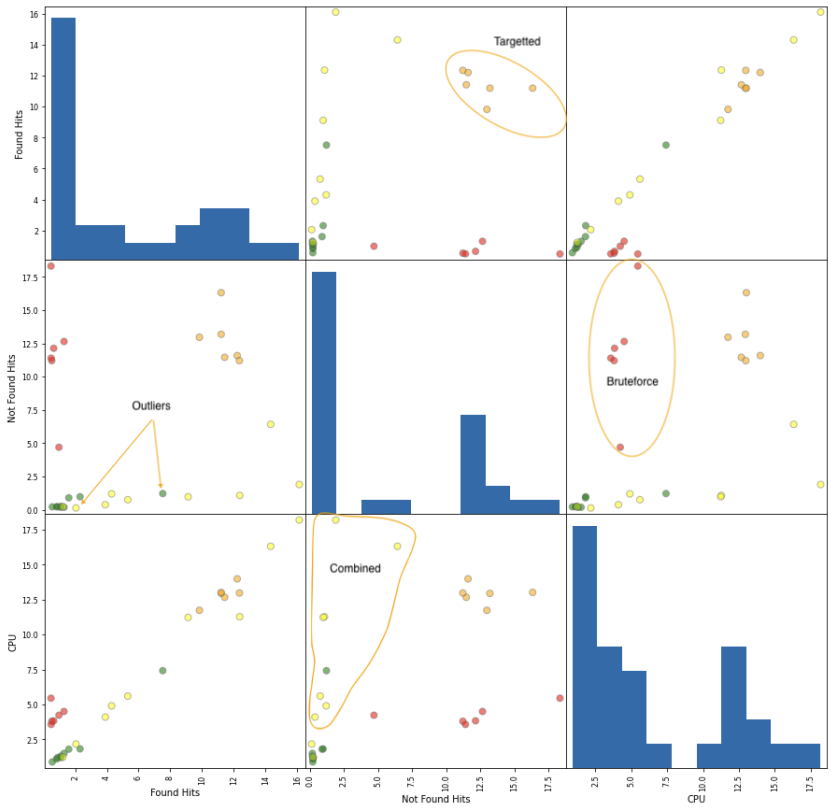

Figure 6.1: Corelarea răspunsurilor și a gradului de procesare pentru determinarea stării sistemului

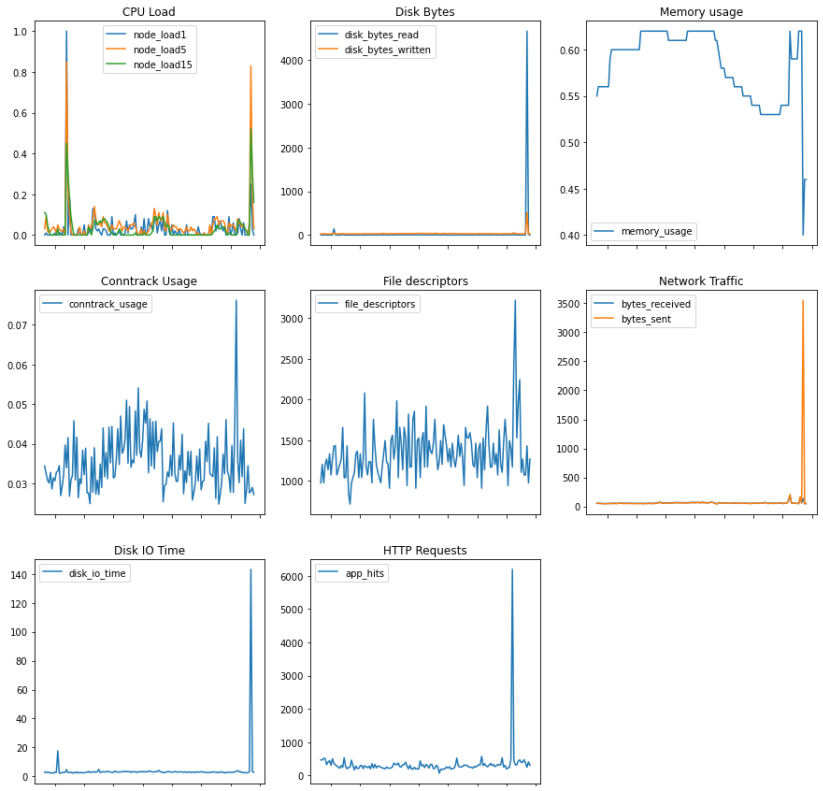

Figure 6.2: Diferite date metrice asociate cu rularea serviciului Whois

Figure 6.3: Serii de timp asociate cu monitorizarea serviciului Whois, vizualizate în Grafana și stocate în Prometheus și OpenTSDB

Capitolul 7

Concluzii

7.1 Rezultate obținute

În prezent platforma AIOps implementată are un grad de maturitate suficient pentru a permite dezvoltarea de servicii comerciale de monitorizare. Astfel, soluția a fost implementată, testată și lansată în execuție pentru Ministerul Afacerilor Externe, pentru monitorizarea platformei cariera.mae.ro în faza inițială.

De asemenea, în 2015 am avut șansa de a câștiga și conduce în calitate de director de proiect, un proiect în valoare de aproximativ 40 de milioane de lei, în cadrul ***"Programului Operațional Competitivitatea 2014- 2020 Axa prioritară 1"***, proiect intitulat ***"Crearea de laboratoare privind cercetarea datelor de mari dimensiuni în vederea dezvoltării unor produse inovative și a unor aplicații în domeniul internetul viitorului"***, cod de proiect ID_34_462, SMIS 2014+, beneficiar fiind compania Anagrama SRL. Unul din obiectivele proiectului era de a demonstra capacitățile infrastructurii dezvoltate prin implementarea și definirea ca produse sau servicii, a patru aplicații de tip *smart-city*: Smart City Map, Buy Local, City Drop și Jobs Nearby, aplicații de tip *mobile* Am avut astfel oportunitatea de a implementa arhitectura descrisă în această teză cu scopul de a culege date de pe dispozitivele mobile în vederea analizării și îmbunătățirii serviciilor și aplicațiilor.

7.2 Contribuții originale

Ca și scop general al prezentei teze, s-a avut în vedere identificarea provocărilor și piedicilor în calea implementării de sisteme IoT de mari dimensiuni, identificarea cerințelor de nivel înalt care adresează aceste provocări, dezvoltarea unei arhitecturi de referință pentru implementarea unui sistem IoT, dezvoltarea unei arhitecturi de referință a unui sistem de monitorizare a ecosistemului IoT și demonstrarea acestor concepte și modele în condițiile unui mediu industrial. Demonstrarea și validarea modelelor descrise au putut fi efectuate în cadrul unor proiecte de cercetare pe care le-am câștigat în competițiile naționale, pe care le-am enumerat în secțiunea 7.3.2. A "Proiecte CDI câștigate și conduse în calitate de director de proiect".
Contribuțiile mele originale sunt în direcția realizării scopului propus, astfel:

1. Am realizat un studiu cu privire la stadiul actual privind adopția tehnologiilor IoT și a standardelor și modelelor de referință pentru implementearea de sistemele IoT complexe și multistratificate. Am pus accentul în mod deosebit pe studierea arhitecturilor de referință dezvoltate de mari actori industriali cu dublu scop: generalizarea unui model unic de referință și identificarea zonelor care prezintă provocări deosebite. [C1]

2. Am propus un model de referință generalizat pentru implementarea sistemelor IoT care poate fi văzut ca un denominator comun al modelelor studiate, model care prezintă 4 nivele: nivelul marginal care conține majoritatea dispozitivelor IoT, nivelul intermediar sau *gateway* unde sunt agregate și filtrate datele, nivelul platformei cloud unde sunt stocate, procesate și analizate datele și nivelul de prezentare.[C2, B3]

3. Am realizat un studiu detaliat al provocărilor deosebite în implementarea sistemelor IoT. Cea mai importantă provocare identificată, constrângerea energetică, are un impact major asupra arhitecturii. Metodele curente de recoltare și stocare a energiei nu permit implementarea de protocoale complexe de comunicații și nici metode de monitorizare a stării la nivelul dispozitivelor IoT. Acesta este motivul pentru care am efectuat un studiu comparativ al protocoalelor de comunicații din perspectiva consumului de energie versus cost,arie de acoperire și ratele de transfer, folosind metoda scorului Z. [A3,A4]

4. Am realizat un studiu cu privire la securitatea și confidențialitatea datelor, în care am identificat și ordonat, după impact și relevanță, problemele de securitate asociate cu sistemele IoT, atât din perspectiva consumatorului de servicii IoT, cât și a furnizorului.[C2,A3]

5. Am formulat cerințele de nivel înalt, de tip funcțional, pentru implementarea unui sistem IoT care adresează provocările identificate, precum agnosticitatea protocoalelor de comunicații, structurare semantică, agregarea și augmentarea datelor, procesare și învățare automată, orchestrare și provizionare automată și necesitatea unei arhitecturi deschise. Pe baza acestor cerințe am propus o arhitectură de referință pentru implementarea unei platforme AIOps care are ca scop monitorizarea automată a sistemelor IoT în vederea creșterii scalabilității, disponibilității și fiabilității sistemului IoT. [A1, B3]

6. Am identificat și testat tehnologiile care pot fi utilizate în implementarea platformei AIOps, cu accent pe acele tehnologii care facilitează tranziția către infrastructuri hiperconvergente. [C4, B9]

7. Am dezvoltat și propus un model de analiză a maturității soluțiilor care poate fi folosit în procesul de decizie strategică la nivel organizațional. Algoritmul se bazează pe modelele de decizie multi-atribut MADM (Multi Attribute Decision Making) și l-am utilizat în alegerea soluției optime în ceea ce privește bazele de date orientate pe serii de timp. Modelul poate fi însă generalizat și folosit pentru orice alte tehnologii IT.[A2,B1]

8. Am coordonat activitățile de cercetare și dezvoltare în cadrul unui pachet de lucru al proiectului European Cloud for Europe C4E, activități care au avut ca obictiv dezvoltarea specificațiilor tehnice (arhitectură de referință și cerințe)

necesare implementării unei soluții de stocare în cloud securizată și bazată pe reguli "legislative", *"Secure Legislation Aware Storage"*. [B2,B4]

9. Pe baza modelului de analiză a maturității soluțiilor, am contribuit la dezvoltarea a două modele MADM de selecție a furnizorilor de servicii cloud.[A10,A11]

10. Am realizat o analiză comparativă a performanțelor soluțiilor de tip bază de date cheie-valoare, prin efectuarea de teste de stres și măsurând performanța în diferite scenarii. Metoda poate fi aplicată și pentru alte soluții și poate constitui criteriu cantitativ de evaluare pentru metoda analizei maturității.[C4]

11. Am testat oportunitatea integrării tehnologiei Blockchain ca metodă de provizionare și autentificare automată a serviciilor IoT, folosind tranzacțiile Blockchain pentru înregistrarea dispozitivelor și serviciilor IoT și utilizând *smart contracts* pentru validarea accesului la servicii și dispozititive. Au fost utilizate mașinile virtuale Ethereum și Arwen WASM (Elrond).[B9]

12. Am contribuit la dezvoltarea de metode de automatizare a proceselor de testare a codului aplicațiilor implementate continuu, conform metodologiei de lucru *agile* și într-un mediu DevOps. [B6]

13. Am realizat un studiu al principalilor algoritmi de învățare automată care pot fi utilizați în rezolvarea problemelor asociate cu agruparea, clasificarea, regresia, reducerea dimensionalității și eliminarea zgomotului, cu scopul comparării avantajelor, dezavantajelor, complexității și domeniilor de aplicare optimă. [C3, A5]

14. Am adaptat modelul de referință pentru platforma AIOps cu scopul dezvoltării unui sistem distribuit de scanare a vulnerabilităților cibernetice la nivel național prin colectarea de date de identificare a tehnologiilor folosite de site-urile web din România (*"Fingerprint web server"*). Soluția a fost implementată la nivelul registrului de domenii .ro ca și concept. [B5]

15. Am implementat, testat și operat platforma AIOps pe baza arhitecturii de referință descrisă, în condiții și mediu industrial, platforma fiind operațională la registrul de domenii .ro (RoTLD). Soluția dezvoltată a fost profilată în scop comercial, fiind în prezent utilizată de anumite instituții ale statului român.[C4]

16. Folosind datele colectate de platforma AIOps de la cele două centre de date, ICIPRO și RoTLD, pe parcursul a aproximativ 3 ani de operare continuă, am contribuit la dezvoltarea de noi metode și modele pentru consolidarea centrelor de date în vederea optimizării consumului de energie [A9].

17. Am investigat posibilitatea detecției de comportamente anormale din partea utilizatorilor unei aplicații monitorizate AIOps, prin înregistrarea interacțiunilor utilzator-interfață web în platforma AIOps. Dacă subiectul interacțiunii poate fi vizualizat ca un nod într-un graf, iar înteracțiunea în sine ca o muchie, atunci secvența de interacțiuni într-o perioadă de timp poate fi modelat ca un graf orientat.[A6]

18. În cadrul proiectului "Crearea de laboratoare privind cercetarea datelor de mari dimensiuni în vederea dezvoltării unor produse inovative și a unor aplicații în domeniul internetul viitorului", în calitate de director de proiect și director științific, am contribuit la implementarea infrastructurii de calcul Big-Data, folosind metode și metodologii descrise în prezenta teză, precum selecția tehnologiilor pe baza analizei maturității și implementarea arhitecturii de referință AIOps.[B1]

7.3 Lista lucrărilor originale

7.3.1 Articole științifice

A. Articole științifice în publicații indexate ISI

A1 Boncea, R., A. Zamfiroiu, and I. Bacivarov, *New method for monitoring microservices in a federated and distributed architecture*, Proceedings of the IE 2018 International Conference, 17-20 May, Iasi, Romania, 2018.

A2 I Petre, R Boncea, CZ Radulescu, A Zamfiroiu, I Sandu, *A Time-Series Database Analysis Based on a Multi-attribute Maturity Model*, Studies in Informatics and Control, ISSN 1220-1766, vol. 28(2), pp. 177-188, 2019. WOS:000473284800006

A3 Alin Zamfiroiu, Ionut Petre, and Radu Boncea. 2019. *Cloud Computing Vulnerabilities Analysis*. In Proceedings of the 2019 4th International Conference on Cloud Computing and Internet of Things (CCIOT 2019). Association for Computing Machinery, New York, NY, USA, 48–53. DOI:https://doi.org/10.1145/3361821.3361830. WOS:000526710900008

A4 Carmen Elena CÎRNU, Carmen Ionela ROTUNĂ, Adrian Victor VEVERA, Radu BONCEA, *Measures to Mitigate Cybersecurity Risks and Vulnerabilities in Service-Oriented Architecture*, Studies in Informatics and Control, 08 2018, ISSN: 1220-1766 eISSN: 1841. WOS:000447079500011

A5 Boncea, R., A. Zamfiroiu, and E. Mitan, *Proposing algorithm to improve student evaluation process*, 10th Annual International Conference on Education and New Learning Technologies, Palma de Mallorca (Spain), 2nd - 4th of July, 2018. WOS:000531474300023

A6 A. Zamfiroiu and R. Boncea, *Modelling the users' profiles based on their behaviour in social applications*, 2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Iasi, Romania, 2018, pp. 1-6, doi: 10.1109/ECAI.2018.8678990. WOS:000467734100060

A7 Petre, I., Cohal, A.M., Boncea, R., *e-Participation Platform for Facilitating Citizens Involvement in Smart City Initiatives*, Revista Română de Informatică şi Automatică, Volume: 28 Issue: 2 Pages: 5-14, 2018. WOS:000455837200001

A8 Zamfiroiu, A., Boncea, R., Petre, I., *Quality of mobile applications based on their development*, Romanian journal of information technology and automatic control, Volume: 28 Issue: 1 Pages: 35-46, 2018. WOS:000455836300003

A9 Delia Mihaela Radulescu, Constanta Zoie Radulescu, Gheorghe Lazaroiu, Radu Boncea, *Binary programming models for the server consolidation problem in data centers*, The 18 International Multidisciplinary Scientific GeoConference SGEM 2018, 30 June - 9 July 2018, Albena, Bulgaria, The accepted article will be published in the Conference Proceedings (ISSN 1314-2704) and will be submitted for evaluating and indexing by ISI Web of Knowledge, Web of Science, Thomson Reuters

A10 C. Z. Rădulescu, I. C. Rădulescu, R. Boncea and E. Mitan, *A group decision approach based on rough multi-attribute methods for Cloud Services Provider selection*, 2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Iasi, Romania, 2018, pp. 1-6, doi: 10.1109/ECAI.2018.8678966. WOS:000467734100036

A11 Constanţa Zoie RADULESCU, Marius RADULESCU, Radu BONCEA, Ionut PETRE, Ionut-Eugen SANDU, Mihail DUMITRACHE, *A Multicriteria Framework for Cloud Service Providers Selection Based on the Matter Element Extension Method*, Studies in Informatics and Control, ISSN 1220-1766, vol. 30(1), pp. 77-87, 2021. https://doi.org/10.24846/v30i1y202107

A12 Boncea, R., Petre, I., Vevera, V., Gheorghiţă, A. *Machine Learning Based Methods Used for Improving Scholar Performance*. In The International Scientific Conference eLearning and Software for Education 2019 (Vol. 2, pp. 471-478). Carol I National Defence University. WOS:000473324400065.

A13 BANCIU, D., PETRE, I., BONCEA, R. *Information and Documentation through New Technologies in E-Learning Process*. In The International Scientific Conference eLearning and Software for Education 2019 (Vol. 2, pp. 465-470). Carol I National Defence University. WOS:000473324400064.

A14 Alin Zamfiroiu, Radu Boncea. *Using Decision Tree and Machine Learning to recognize users by their behaviour*. Proceedings of the 16th international conference on informatics in economy (IE 2017). WOS: 000418463600015.

B. Articole științifice în publicații indexate BDI

B1 Radu BONCEA, Ionuț PETRE , Dragoș-Marian SMADA, Alin ZAMFIROIU, *A Maturity Analysis of Big Data Technologies*, Informatica Economică vol. 21, no. 1/2017, DOI 10.12948, ISSN 14531305

B2 Alin Zamfiroiu, Carmen Elena Cîrnu, Radu Boncea, Carmen Rotună, Monica Anghel, *Principii de proiectare, securitate și administrare a soluțiilor de stocare în cloud*, Revista Română de Informatică și Automatică, vol. 25, nr. 2, 2015

B3 Radu BONCEA, Ioan BACIVAROV *System Architecture for Monitoring the Reliability of IoT*, PROCEEDINGS of the 15th International Conference on Quality and Dependability, pp. 143-149, 2016

B4 Radu BONCEA, Carmen Elena CÎRNU, *Cloud for Europe Project: New Solutions for Addressing Cloud Security Issues*, PROCEEDINGS of the 15th International Conference on Quality and Dependability, PROCEEDINGS of the 15th International Conference on Quality and Dependability, pp.156-160

B5 Eugenie STĂICUȚ, Radu BONCEA, Carmen ROTUNĂ, *A Reliable Architecture for a Massive and Continuous Scanner of Web Vulnerabilities in Internet*, PROCEEDINGS of the 15th International Conference on Quality and Dependability, pp. 176-180

B6 Dragoș SMADA, Carmen ROTUNĂ, Radu BONCEA, Ionuț PETRE. *Automated Code Testing System for Bug Prevention in Web-based User Interfaces*, Revista de Informatica Economică, vol. 22, no. 3/2018

B7 Alin ZAMFIROIU, Radu BONCEA, Ionuț PETRE, *Quality of mobile applications based on their development*, Romanian Journal of Information Technology and Automatic Control, ISSN 1220-1758, vol. 28(1), pp. 35-46, 2018

B8 Badea, V.E., A. Zamfiroiu, and R. Boncea, *Big Data in the Aerospace Industry*, Revista Română de Informatică şi Automatică, vol. 22, no. 1, pp. 17-24, 2018

B9 Radu BONCEA, Ionut PETRE, Victor VEVERA, *Building Trust Among Things in Omniscient Internet Using Blockchain Technology*, Romanian Cyber Security Journal, Spring 2019, No. 1, vol. 1.

B10 Gabriel Neagu, Ionut Petre, Radu Boncea, Mihail Dumitrache, *Building a business model for service offer integrator in case of cloud-iot based monitoring*, Conference: 18th International Conference on INFORMATICS in ECONOMY. Education, Research and Business Technologies, May 2019, DOI: 10.12948/ie2019.02.05.

C. Rapoarte științifice în cadrul programului doctoral

C1 Raportul științific nr.1 "*Analiză State of Art a Internetului Obiectelor*"

C2 Raportul științific nr.2 "*Arhitecturi de referinta pentru IoT*"

C3 Raportul științific nr.3 "*Aplicarea algoritmilor de inteligenta artificiala in retelele wireless de senzori pentru detectarea erorilor si anomaliilor de sistem*"

C4 Raportul științific nr.4 "*Integrarea învățării automate în arhitecturile de monitorizare a infrastructurilor de date*"

C5 Raportul științific nr.5 "*Arhitectură de referință pentru soluții AIOps*"

7.3.2 Proiecte de cercetare, dezvoltare și inovare

A. Proiecte CDI câștigate și conduse în calitate de director de proiect

D1 **PN 19370401 "Soluții noi pentru probleme complexe din domenii actuale de cercetare TIC bazate pe modelare și optimizare"**, Programul Nucleu 2019-2022. Proiectul se concentrează pe elaborarea unei colecții

de soluţii noi pentru probleme complexe din domenii actuale de cercetare TIC, soluții bazate pe modelare şi optimizare. Prin soluții noi se înțeleg noi metode şi modele, metodologii, algoritmi şi software, ca rezultate ale modelării unor probleme complexe. Problemele complexe abordate de proiect se referă la: (a) creșterea fiabilității și securității sistemelor complexe, (b) achiziția și analiza datelor cantitative prin soluții integrate, (c) îmbunătățirea eficienței energetice în centre de date, (d) selecția optimală a furnizorilor de servicii cloud, de produse și energie, (e) probleme complexe de teoria grafurilor NP hard și (f) securitatea datelor în sisteme software pentru dispozitive mobile.

D2 **PN18190101 "Noi cercetări în modelarea și optimizarea sistemelor complexe cu aplicații în industrie, mediul de afaceri și cloud computing"**. Proiectul a propus și realizat soluții noi pentru sisteme complexe cu aplicabilitate în industrie și afaceri. Au fost elaborate metode, modele și instrumente software de optimizare utilizabile în context industrial, metode de decizie multi-criteriale, algoritmi și software cu aplicații în cloud computing și business precum și algoritmi din teoria grafurilor pentru probleme NP-hard cu aplicații reale.

D3 **"Crearea de laboratoare privind cercetarea datelor de mari dimensiuni în vederea dezvoltării unor produse inovative și a unor aplicații în domeniul internetul viitorului"**, cod de proiect ID_34_462, SMIS 2014+, Programului Operațional Competitivitatea 2014- 2020 Axa prioritară 1. Proiectul are ca obiective dezvoltarea de laboratoare BigData cu scopul dezvoltării de produse și aplicații inovative în domeniul internetului viitorului și dezvoltarea a patru aplicații comerciale de tip *smart-city* care dovedesc capacitatea laboratoarelor.

D4 **"Studiu privind sisteme adaptive de recunoaștere în stadii incipiente a atacurilor cibernetice asupra resurselor statale"** în cadrul Planului Sectorial al MCSI 2018-2020. Proiectul a realizat o arhitectură de referință pentru un sistem de detecție a atacurilor cibernetice folosind modele și algoritmi de învățare automată și analiza datelor de tip metric colectate din sistem.

B. Alte proiecte

E1 Proiectul european **"EuroCC - Centre Naționale de Competență în HPC (High Performance Computing)"** își propune stabilirea, conectarea și operarea unui număr de 33 de centre naționale de competențe în HPC cu scopul de a oferi acces la tehnologii, cunoștințe, expertiză și competențe HPC, aliniate la nevoi specifice naționale și în funcție de nivelul de maturitate al HPC al fiecărui stat.

E2 Proiectul european **"Cloud for Europe – C4E"** adresează obiectivele Parteneriatului European Cloud (Cloud European Partnership) şi ajută partenerii să adopte o strategie europeană bine definită privind tehnologia Cloud Computing pentru sectorul public. Obiectivele proiectului sunt: identificarea obstacolelor de utilizare a tehnologiei Cloud Computing în

sectorul public;definirea serviciilor care vor permite depășirea acestor obstacole;facilitarea cercetării din industrie pentru a identifica soluţii inovatoare pentru servicii de Cloud.

E3 Proiectul nucleu **"Cercetări privind politici și soluții avansate de securizare a infrastructurilor critice împotriva atacurilor cibernetice"** se concentrează pe zona sistemelor de control industrial, pe identificarea vulnerabilităților structurii organizatorice, a sistemelor și rețelelor de control industrial, precum și a procedurilor de securitate aplicate pot fi exploatate involuntar sau în mod voit de către unul sau mai mulți atacatori generând astfel evenimente cibernetice cu puternic impact asupra afacerii, infrastructurii critice și a siguranței naționale.

7.4 Perspective de dezvoltare ulterioară

Una din provocările majore identificate în implementarea sistemelor IoT de mari dimensiuni, este securitatea cibernetică, în special securitatea datelor. O soluție pentru securitatea datelor prin metode de autorizare al accesului și auditare, este integrarea tehnologiei Blockchain, tehnologie testată în condiții de laborator prin definirea unui nomenclator al dispozitivelor și serviciilor IoT, unde înregistrarea unui element are ca efect crearea unei tranzacții în Blockchain care conține ca informație (*payload*) identificatorul elementului. De asemenea au fost testate diferite metode de autentificare și autorizare al accesului la dispozitive și servicii folosind *smart contracts* și mașinile virtuale Ethereum și Arwen WASM. Însă integrarea Blockchain în vederea operării în condiții industriale nu a fost testată.
Astfel o perspectivă de dezvoltare este testarea în condiții apropiate de realitate a integrării Blockchain, prin verificarea performanțelor și a costurilor asociate, astfel:

- vor fi comparate costurile legate de tranzacții pentru diferite soluții Blockchain;
- vor fi comparate performațele asociate cu timpii medii de acces la ledger, latențele;
- vor fi studiate noi modele de compresie a informației stocate în ledger pentru a reduce costurile;
- se va studia oportunitatea dezvoltării de subchain-uri dedicate și operate independent similar arhitecturii IoTex.
- se vor studia cazuri și exemple de aplicații descentralizate (DApps) în domeniul IoT.

O altă direcție de dezvoltare este simularea, atât la nivel de rețea sau protocoale de comunicații, cât și la nivel de date generate, a sistemelor IoT prin implementarea unor instrumente asemănătoare *ns3* (Network Simulator 3). Implementarea unui astfel de simulator este în sine un proiect de anvergură care presupune dezvoltarea unui set de instrumente, fiecare instrument fiind specializat. În urma unui studiu de piață sumar pe care l-am făcut, am ajuns la concluzia că există simulatoare specializate pe anumite capabilități (NetSim, Proteous, CupCarbon, Bevywise, NS3, Cooja, MANET), dar nu există un mediu de simulare care integrează simulatoare pe toate nivelele arhitecturale.

Sistemele încorporate avansate implementate la nivelul marginal sau de gateway (*advanced embedded systems for edge computing*) reprezintă de asemenea un domeniu de interes, în condițiile în care actori industriali precum NVidia au dezvoltat soluții hardware și software care pot rula modele complexe de rețele neurale în domenii precum procesarea video sau de imagini, cu un consum redus de energie. Un exemplu este portofoliul de produse Nvidia Jetson, care, integrate la nivel de gateway în arhitectura de referință IoT descrisă în această teză, pot schimba substanțial fluxul datelor între platforma cloud și gateway. Ne putem imagina astfel un scenariu în care există o cameră video amplasată într-o intersecție și ne propunem să numărăm mașinile care trec prin intersecție la fiecare t minute. Prin integrarea unei soluții precum NVidia Jetson Nano ca și gateway amplasat lângă cameră, procesul de numerotare poate fi executat chiar de Jetson care are implemementate modele instanțiabile de recunoaștere a obiectelor, inclusiv mașini. Evităm astfel trimiterea fluxului video către platforma cloud în vederea procesării, în schimb vom trimite strict starea contorului, economisind resurse semnificative.

În calitate de șef de departament al departamentului "Ingineria software și a sistemelor complexe" din cadrul ICI București, mi-am propus înființarea unui laborator de inovație în domeniul Internetului Obiectelor în cadrul institutului. În acest sens, am creat acorduri de principiu cu părți interesate, precum firme și companii care au ca obiect de activitate cercetarea și inovarea, universități și autorități locale interesate de implementarea tehnologiilor smart city. Scopul laboratorului este de a oferi baza materială și de cunoaștere tinerilor cercetători din ICI sau studenților care fac practică în ICI, prin implicarea lor în proiecte de cercetare, dezvoltare și inovare cu participarea părților interesate. Un alt obiectiv al laboratorului este de a face spin-off la soluțiile care ating nivelul de maturitate tehnologică cel puțin demonstrabilă în condiții simulate sau de laborator și de a asigura tranziția lor către produse și servicii valabile comercial.

Bibliography

[1] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Comput. Netw.*, vol. 54, pp. 2787–2805, Oct. 2010.

[2] B. Sterling, *Shaping Things (Mediaworks Pamphlets)*. The MIT Press, Sept. 2005.

[3] K. L. Lueth, "State of the iot 2018: Number of iot devices now at 7b – market accelerating," tech. rep., 2018.

[4] R. Boncea and I. Bacivarov, "A system architecture for monitoring the reliability of iot," in *Proceedings of the 15th International Conference on Quality and Dependability*, pp. 143–150, SOCIETATEA ROMÂNĂ PENTRU ASIGURAREA CALITĂȚII, 2016.

[5] R. Boncea and C. E. Cîrnu, "Cloud for europe project: New solutions for addressing cloud security issues," in *Proceedings of the 15th International Conference on Quality and Dependability*, pp. 156–160, SOCIETATEA ROMÂNĂ PENTRU ASIGURAREA CALITĂȚII, 2016.

[6] A. Zamfiroiu, I. Petre, and R. Boncea, "Cloud computing vulnerabilities analysis," in *Proceedings of the 2019 4th International Conference on Cloud Computing and Internet of Things*, CCIOT 2019, (New York, NY, USA), p. 48–53, Association for Computing Machinery, 2019.

[7] C. CÎRNU, C. I. ROTUNĂ, A. V. VEVERA, and R. BONCEA, "Measures to mitigate cybersecurity risks and vulnerabilities in service-oriented architecture," *Studies in Informatics and Control*, vol. 27, no. 3, pp. 359–368, 2018.

[8] I. PETRE, R. BONCEA, C. Z. RADULESCU, A. ZAMFIROIU, and I. SANDU, "A time-series database analysis based on a multi- attribute maturity model," *Studies in Informatics and Control*, vol. 28, no. 2, pp. 177–188, 2019.

[9] C. Z. Rădulescu, I. C. Rădulescu, R. Boncea, and E. Mitan, "A group decision approach based on rough multi-attribute methods for cloud services provider selection," *2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, pp. 1–6, 2018.

[10] I. Z. RADULESCU, M. RADULESCU, R. BONCEA, I. PETRE, I.-E. SANDU, and M. DUMITRACHE, "A multicriteria framework for cloud service providers selection based on the matter element extension method," *Studies in Informatics and Control*, vol. 30, no. 1, pp. 77–87, 2021.

[11] R. Boncea, A. Zamfiroiu, and I. Bacivarov, "New method for monitoring microservices in a federated and distributed architecture," in *Proceedings of the 17th International Conference on Informatics in Economy*, pp. 13–18, Bucharest University of Economic, 2018.

[12] R. BONCEA, I. PETRE, and V. VEVERA, "Building trust among things in omniscient internet using blockchain technology," *Romanian Cyber Security Journal*, vol. 1, no. 1, pp. 17–24, 2019.

[13] A. Zamfiroiu and R. Boncea, "Modelling the users' profiles based on their behaviour in social applications," *2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, pp. 1–6, 2018.

[14] A. Zamfiroiu and R. Boncea, "Using decision tree and machine learning to recognize users by their behaviour," in *PROCEEDINGS OF THE 16TH INTERNATIONAL CONFERENCE ON INFORMATICS IN ECONOMY (IE 2017)*, pp. 90–95, Bucharest University of Economic Studies, 2017.

[15] R. Boncea, A. Zamfiroiu, and E. Mitan, "Proposing algorithm to improve student evaluation process," in *EDULEARN18 Proceedings*, 10th International Conference on Education and New Learning Technologies, pp. 5799–5805, IATED, 2-4 July, 2018 2018.

[16] R. Boncea, V. Vevera, I. Petre, and A. Gheorghita, "Machine learning based methods used for improving scholar performance," in *The International Scientific Conference eLearning and Software for Education*, vol. 2, pp. 471–478, Carol I National Defence University, 2019.