PRINCIPII DE PROIECTARE, SECURITATE ŞI ADMINISTRARE A SOLUȚIILOR DE STOCARE ÎN CLOUD

Alin Zamfiroiu zamfiroiu@ici.ro

Carmen Elena Cîrnu carmen.cirnu@ici.ro

Radu Boncea radu@rotld.ro

Carmen Rotună carmen.rotuna@rotld.ro

Monica Anghel monica.anghel@ici.ro

Institutul Național de Cercetare-Dezvoltare în Informatică - ICI Bucureşti

Rezumat: Lucrarea de față explorează noile tehnologii informatice şi noile paradigme apărute în domeniul stocării securizate de date (în orice format s-ar prezenta acestea). Astfel, subiectul dezbătut în cadrul acestei cercetări se referă la principiile de proiectare, securitate şi administrare a soluțiilor de stocare în Cloud.

Cuvinte cheie: Cloud, date personale, stocare, cerințe tehnice, proiectare, securitate.

Abstract: This paper explores the new information technologies and emerging paradigms in secure data storage. Thus, the subject discussed in this research refers to the principles of design, security and storage management solutions in the Cloud .

Keywords: Cloud, personal data, storage, technical requirements, design, security.

1. Stocarea în Cloud

Încă de la începutul acestui secol, schimbările în domeniul TIC au fost mai rapide ca niciodată şi, în plus, a apărut o nouă paradigmă în domeniu: nu numai un calculator şi o persoană, ci un sistem puternic care oferă e-Servicii: aceasta este tehnologia Cloud Computing.

Conceptul de Cloud Computing reprezintă puterea de calcul – stocarea, procesarea şi utilizarea de date pe computer aflate la distanță de utilizator, utilizate prin Internet atunci când este nevoie. Comisia Europeană în documentele sale arată că: "Cloud Computing reprezintă o nouă fază de industrializare a furnizării puterii de calcul în regim de utilitate publică ("utility computing") comparabilă cu industrializarea furnizării de electricitate de către centralele din domeniul energiei" [1], [2].

Chiar dacă Europa nu este suficient de dezvoltată în ceea ce priveşte această tehnologie, au început să apară inițiative în acest domeniu precum inițiativa Cloud for Europe [3] care va oferi o viziune clară asupra cerințelor şi scenariilor de utilizare a tehnologiei de Cloud Computing în sectorul public, industrie şi organismele de standardizare, principalele obiective ale acestei inițiative sunt:

- identificarea obstacolelor pentru utilizarea tehnologiei cloud în sectorul public;
- definirea serviciilor care pot înlătura aceste obstacole;
- utilizarea cercetărilor din industrie pentru găsirea de soluții inovative pentru serviciile de tip cloud.

Principalii termeni utilizați pentru tehnologia Cloud sunt:

- SOA (Service Oriented Architecture) – arhitectura orientată pe servicii. Grupul OASIS şi Open Group au creat definiții formale:

    1. Definiția OASIS pentru Arhitectura Orientată către Servicii: O paradigmă pentru organizarea şi utilizarea capabilităților distribuite care pot fi sub controlul diferitelor domenii de proprietate. Acesta asigură o modalitate uniformă de a oferi, de a descoperi, de a interacționa şi de a folosi capabilități pentru a produce efecte dorite în concordanță cu condiții şi aşteptări măsurabile.

2. Definiția Open Group pentru Arhitectura Orientată către Servicii: Arhitectura SOA este un stil arhitectural care acceptă servicii de orientare. Orientarea-serviciu este un mod de gândire în termeni de servicii, dezvoltare bazată pe servicii şi rezultate ale serviciilor. Un serviciu este o reprezentare logică a unei activități economice repetabile, care are un rezultat specificat (de exemplu verificarea creditului unui client, furnizarea de date meteorologice, consolidarea rapoartelor de foraj).

- CDMI (Cloud Data Management Interface) - Interfață Cloud pentru Managementul Datelor; CDMI defineşte interfața funcțională pe care aplicațiile o folosesc pentru a crea, prelua, actualiza şi şterge date din cloud. Prin intermediul acestei interfețe, clientul este capabil să descopere capacitățile cloud de stocare şi să o folosească pentru a gestiona containere şi date. În plus, metadatele pot fi setate pe containere şi date conținute, prin această interfață. Această interfață este, de asemenea, utilizată de aplicații administrative şi aplicații de management pentru a gestiona containere, conturi, informații de securitate acces şi monitorizare/facturare, chiar şi pentru depozitare, care este accesibilă prin alte protocoale. Capacitățile care stau la baza serviciilor de stocare date sunt expuse, astfel încât clienții să poată înțelege oferta [4].
- Criptoperioada - este perioada de timp în care o cheie criptografică este autorizată pentru folosință de către entități legale, sau perioada în care cheile rămân în vigoare pentru un sistem dat; Factorii care pot afecta criptoperioda sunt consecințele expunerii, modul de utilizare al cheii: pentru comunicații sau depozitare, precum şi costurile de revocare şi înlocuire pentru cheie. Cheile care sunt utilizate pentru protecția confidențialității datelor în comunicații pot avea de multe ori criptoperioade mai scurte decât cheile utilizate pentru protecția datelor stocate. Criptoperioada are în general valori mai mari pentru datele stocate, deoarece recriptarea asociată cu schimbarea cheilor poate fi împovărătoare; Criptoperioada definită în mod corespunzător îndeplineşte următoarele caracteristici:
    - limitează cantitatea de informații protejate de o anumită cheie, care este disponibilă pentru criptanaliză;
    - limitează nivelul de expunere, dacă o singură cheie este compromisă;
    - limitează utilizarea unui anumit algoritm particular la durata sa de viață efectiv estimată;
    - limitează timpul disponibil pentru încercările de a pătrunde fizic, procedural şi logic la mecanismele care protejează o cheie de divulgări neautorizate;
    - limitează perioada în care informațiile pot fi compromise prin divulgarea involuntară;
    - limitează timpul disponibil pentru atacuri criptanalitice.

Stocarea în Cloud presupune o serie de principii şi cerințe tehnice în ceea ce priveşte proiectarea, securitatea, implementarea şi modul de administrare al datelor şi documentelor stocate. Realizarea unui sistem de Cloud presupune conştientizarea şi respectarea acestor principii. Nerespectarea lor conduce la realizarea unui Cloud neadecvat şi nefolosibil de către utilizatorii finali.

2. Principii de proiectare a Cloud-ului, bazate pe SOA

Soluția trebuie sa aibă la bază o arhitectură orientată pe servicii, urmând principiile SOA pentru designul soluției:

- interoperabilitate – capacitatea aplicațiilor de a interacționa într-un mod standardizat în vederea realizării unui serviciu IT. Interoperabilitatea, promovată în mod special prin aplicarea consecvent a diferite proiecte în diferite perioade pot fi în mod repetat asamblate împreună într-o varietate de configurații pentru a ajuta la automatizarea proceselor;

- abstractizarea serviciilor - presupune ca informațiile publicate într-un contract de servicii să fie limitate la ceea ce este necesar pentru a utiliza în mod eficient serviciile. La nivel fundamental, acest principiu subliniază necesitatea de a ascunde cât mai mult detaliile care stau la baza unui serviciu precum detaliile tehnice ale platformei sau detalii neesențiale despre serviciul în sine;
- autonomia serviciilor - este un principiu de design care presupune furnizarea de servicii independente față de mediile lor de execuție. Aplicând autonomia la nivel de serviciu, limitele serviciilor pot fi deosebite, chiar dacă acestea partajează unele resurse. Compunerea Serviciilor - încurajează proiectarea de servicii care pot fi refolosite în soluții multiple care la rândul lor sunt alcătuite din servicii compuse. Acest principiu este direct responsabil pentru agilitatea promisă de SOA, deoarece promovează compunerea de noi soluții prin reutilizarea serviciilor existente;
- detectarea serviciilor - completarea serviciilor cu metadate, prin care ele pot fi descoperite şi interpretate în mod eficient. Pentru a face un serviciu detectabil este necesar ca informațiile despre serviciu să fie documentate în mod consecvent şi totodată să fie stocate într-un registru pentru a se permite căutarea informațiilor într-un mod eficient;
- servicii cuplate independent - promovează designul independent şi evoluția implementării unui serviciu, în acelaşi timp garantând interoperabilitatea. Acest lucru are ca rezultat contracte de servicii care ar putea evolua fără a afecta consumatorii sau implementarea serviciilor;
- reutilizarea serviciilor - crearea de servicii, care au potențialul de a fi reutilizate la nivel enterprise. Aceste servicii reutilizabile sunt concepute într-un mod, astfel încât soluția lor logică este independentă de orice proces particular sau tehnologie;
- servicii statelessness - minimizarea consumului de resurse prin amânarea gestionării informațiilor de stare atunci când este necesar. Managementul excesiv al informațiilor de stare poate compromite disponibilitatea unui serviciu şi submina potențialul său de scalabilitate;
- orientarea pe servicii şi interoperabilitatea - într-o arhitectură orientată spre servicii, "interoperabilitatea" se referă la capacitatea serviciului de a fi invocat de către orice client potențial al serviciului respectiv. Interoperabilitate semantică depinde de modul în care interfețele unui serviciu sunt descrise şi de modul în care informațiile sunt împărțite cu potențialii clienți ai serviciului. Din punct de vedere sintactic, o arhitectura orientată pe servicii este promițătoare, iar provocarea constă în determinarea numărului de adaptoare de implementare şi determinarea granularității interfețelor de servicii, deoarece nu este întotdeauna cunoscut modul în care sistemele vor utiliza serviciile;
- contract de servicii standardizat - asigură faptul că serviciile care sunt în acelaşi inventar de servicii sunt păstrate în conformitate cu standardele de proiectare pe baza unui contract. Contractul de Servicii Standardizat impune practic considerente specifice ce trebuie luate în calcul atunci când interfața tehnică publică a unui serviciu este în curs de proiectare. Acesta evaluează simultan natura şi cantitatea de conținut care va fi publicat ca parte a contractului oficial al serviciului.

3. Securitatea şi integritatea datelor

- datele aflate la repaus, care nu sunt accesate frecvent trebuie păstrate în mod criptat [5];
- sistemul de criptare trebuie să aibă la bază o soluție dedicată, decuplată de sistemul de stocare, facilitând adaptabilitatea;

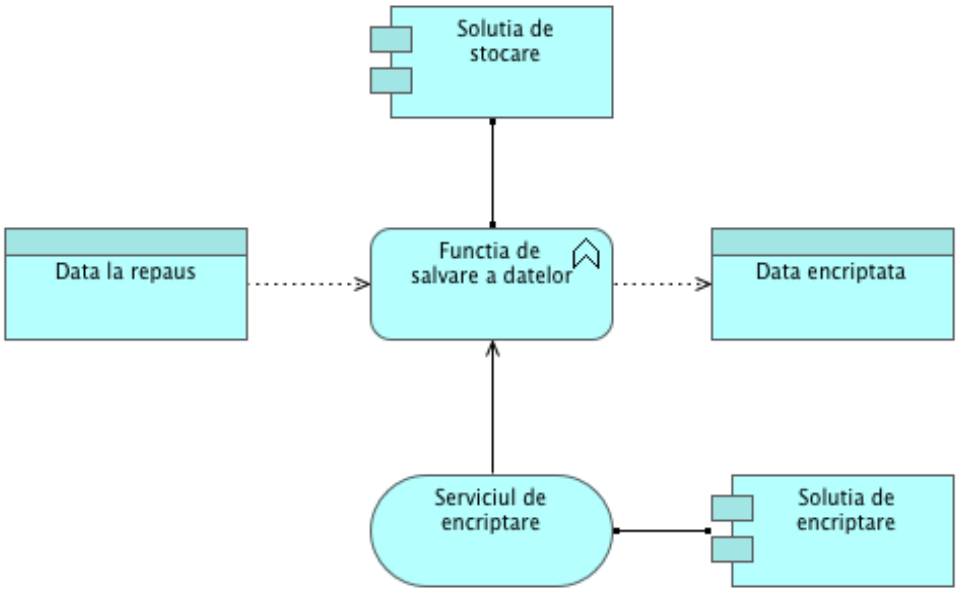

Figura 1. Decuplarea soluției pentru criptare de sistemul de stocare

- criptoperioada şi lungimea cheilor trebuie să ţină cont de recomandările organismelor şi standardelor internaţionale şi europene consacrate în domeniu precum NIST, ANSSI, BSI şi IETF RFC3766 [6];
- accesul la date trebuie să fie controlat folosind un sistem de privilegii pentru utilizatori şi care face distincție clară între proprietarul de date şi persoana mandatată să aibă acces la date şi agentul guvernamental;
- utilizatorii trebuie să fie autentificaţi prin user şi parola pentru a avea acces la sistemul informatic;
- sistemul trebuie să permită autentificarea de tip two-factor prin implementarea HOTP sau TOTP. Autentificarea de tip two-factor (de asemenea, cunoscută sub numele de 2FA) este o autentificare multi-factor ce prevede identificare neambiguă a utilizatorilor prin combinarea a două componente diferite. Aceste componente pot fi ceva ce utilizatorul ştie, ceva ce utilizatorul are sau ceva ce este inseparabil de utilizator. Utilizarea acestui sistem pentru a dovedi identitatea unei persoane se bazează pe premisa că este puțin probabil ca un autor neautorizat să fie în măsură să furnizeze cei doi factori necesari pentru acces. Dacă într-o încercare de autentificare cel puțin una dintre componente lipseşte sau este furnizată incorect, identitatea utilizatorului nu este stabilită cu certitudine şi cererea de acces este respinsă;
- HOTP (HMAC-based One-Time Password algorithm), descris în standardul RFC4226, se bazează pe două lucruri fundamentale: un secret comun şi un factor în mişcare (counter). Acest algoritm este bazat pe evenimente, ceea ce înseamnă că de fiecare dată când este generată o parolă, factorul mişcare va fi incrementat pe bază de evenimente, deci parolele generate ulterior ar trebui să fie diferite de fiecare dată;
- TOTP (Time-based One-Time Password Algorithm-RFC6238) este un algoritm care calculează o parolă unică de la o cheie secretă partajată şi ora curentă folosind o funcție hash criptografică pentru a genera o parolă one-time;

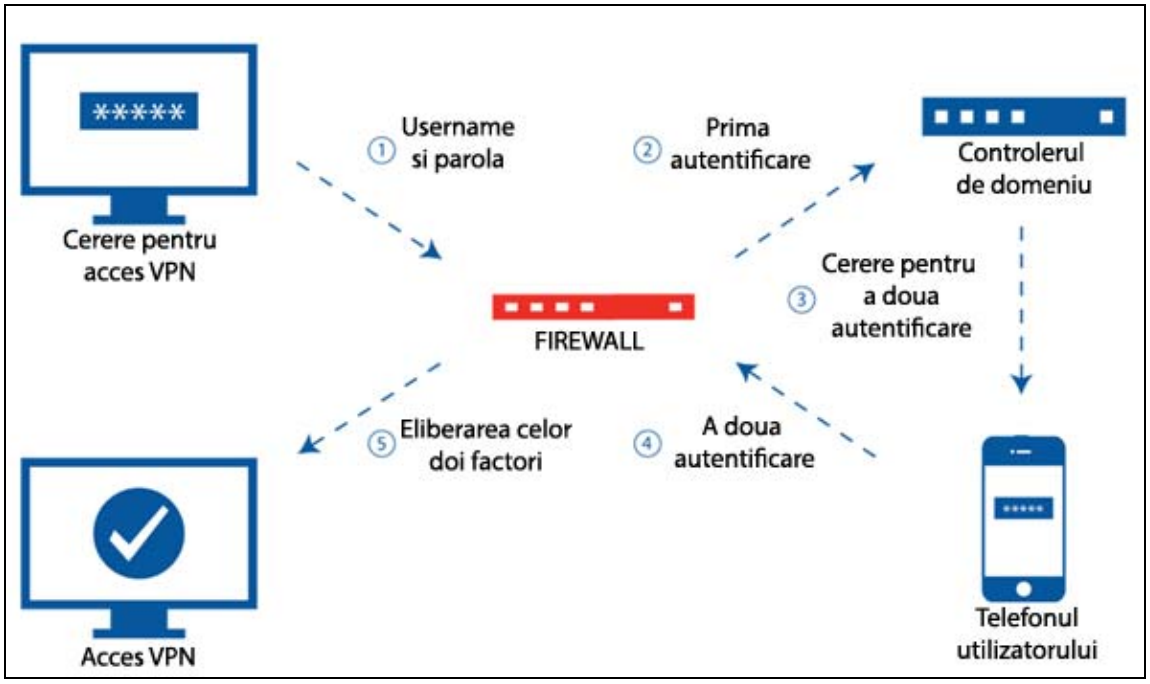

Figura 2. Autentificare de tip two-factor

- mesajele la nivel de aplicație în arhitectura SOA trebuie să fie scanate în vederea detecției informațiilor malformate – message screening, deoarece un atacator poate trimite conținut dăunător serviciului, care conduce la un comportament nedorit. Conform [7] soluția constă în verificarea tuturor datelor de intrare şi a mesajelor primite de către serviciul web.

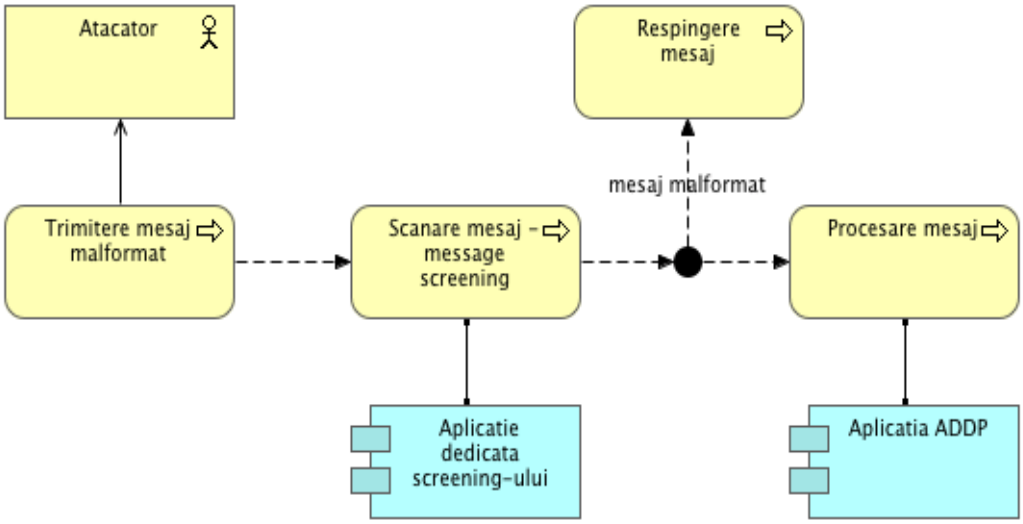

Figura 3. Message screening

În cazul în care mesajul este malformat sau are conținut dăunător este respins, altfel acesta intră în etapa de procesare şi este transmis aplicației.

4. Prezervarea şi retenția datelor

- integritatea datelor trebuie asigurată în cazul replicării sau migrării [8];
- integritatea datelor poate fi verificată folosind funcții hash;
- sistemul trebuie să permită inventarierea formatelor datelor şi monitorizarea acestora în eventualitatea deprecierii lor. Soluția ar putea implementa un sistem de tip watchdog timer,

unde toate formatele trebuie confirmate periodic iar formatele neconfirmate sunt monitorizate;

- sistemul trebuie să implementeze funcțiile pentru prezervarea datelor în caz de litigiu (Litigation Hold), prevedere necesară pentru păstrarea datelor cu caracter personal; această cerință garantează că datele vor fi disponibile în cazul unui litigiu între părțile implicate;
- sistemul trebuie să ofere posibilitatea ştergerii permanente a datelor [9], doar în cazul în care se doreşte acest lucru; în sens contrar realizându-se doar o ştergere fictivă;

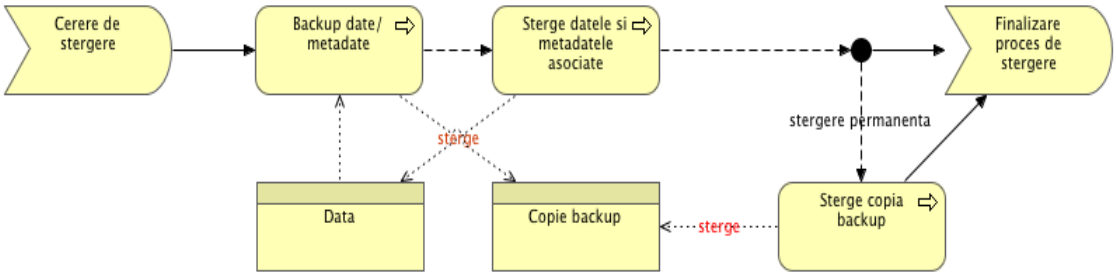

Figura 4. Procesul de ştergere permanentă a datelor

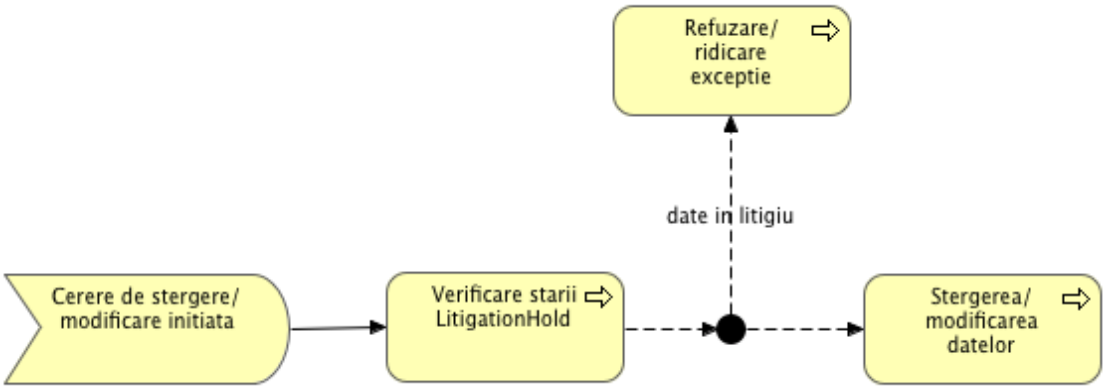

Figura 5. Verificarea stării de litigiu în procesul de ştergere a datelor

- sistemul de arhivare trebuie să fie conform cu normele europene şi internaționale în domeniu, precum:
  - MoReq 2 (Model Requirements for the Management of Electronic Records - Modelul cerințelor de gestionare a documentelor electronice), a fost publicat în 2008 şi este destinat utilizării în întreaga Uniune Europeană. Acesta nu are statut oficial de standard însă la nivel european este considerat un standard european [10];
  - SOX (Sarbanes-Oxley) este un act adoptat de Congresul SUA în 2002 pentru protejarea publicului de erorile contabile sau practicile frauduloase.
- datele pot avea asociate metadate. Metadatele se supun aceloraşi reguli ca şi datele.
- metadatele trebuie să aibă la bază vocabulare controlate precum:
  - Dublin Core este un set de termeni folosiți pentru descrierea resurselor web (video, audio, imagini) şi a resurselor fizice precum cărți, CD-uri [11];
  - ISA Core Vocabularies este un vocabular cu un set minim de clase şi proprietăți pentru descrierea prin nume şi adresa a anumitor locuri;

- W3C Open Annotation Data Model este folosit pentru crearea de asocieri prin adnotări între resurse conexe, utilizând o metodologie conform cu arhitectura World Wide Web. O adnotare este un set de resurse legate între ele [12].

5. Provizionarea facilă a serviciilor

Provizionarea serviciilor cloud se bazează pe crearea de conturi într-un mediu cloud iar autorizarea şi setările utilizatorului sunt configurate pentru servicii şi aplicaţii localizate la distanţă şi livrate prin Internet.

Acest proces se poate desfăşura în mai multe moduri diferite:

- provizionare în avans - furnizorul pregăteşte resursele corespunzătoare înainte de începerea serviciului;
- provizionare dinamică - furnizorul alocă mai multe resurse dacă sunt necesare şi le îndepărtează, dacă nu sunt;
- auto-provizionarea - clientul solicită resurse direct de la furnizorul de cloud prin intermediul unui formular web.

6. Administrarea facilă a soluției

Un sistem de management pentru cloud combină software şi tehnologii într-un design de gestionare a mediilor cloud. La un nivel minim, un sistem de management în cloud trebuie să aibă capacități pentru a:

- gestiona un fond de resurse de calcul eterogene;
- asigura accesul utilizatorilor finali;
- optimiza volumul de lucru;
- monitoriza securitatea;
- gestiona alocarea resurselor;
- gestiona sistemul de monitorizare.

Data Management Interface Cloud – CDMI defineşte o interfață funcțională pe care aplicațiile o vor folosi pentru a crea, prelua, actualiza şi şterge date din cloud. Prin intermediul acestei interfețe, clientul va fi capabil să descopere capacitățile cloud de stocare şi să o folosească pentru a gestiona containerele şi datele existente. În plus, metadatele pot fi setate pe containere şi date conținute, prin această interfață.

Majoritatea ofertelor de stocare cloud existente astăzi sunt în măsură să implementeze o interfață fie direct folosind un adaptor pentru integrarea cu interfața proprie, sau în paralel cu aceasta.

Prin urmare, această interfață poate fi utilizată de aplicații administrative şi de management pentru a gestiona containere, domenii, acces de securitate, precum şi pentru informare monitorizare / facturare.

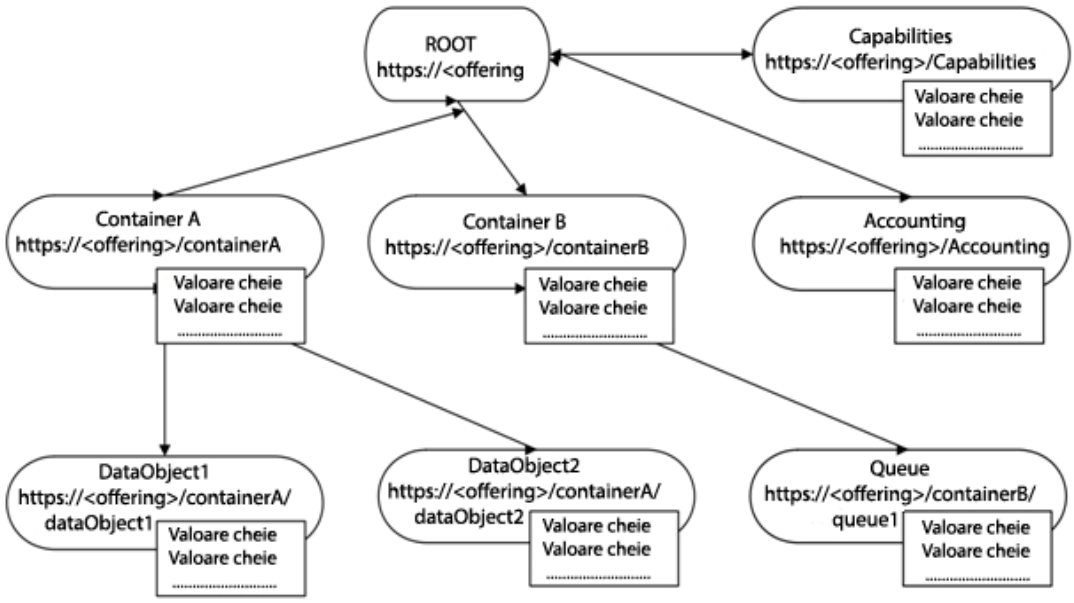

Figura 6. Modelul de Interfata CDMI

Capacitățile care stau la baza serviciilor de stocare date sunt expuse, vizibile sau publice astfel încât clienții să poată înțelege oferta de servicii. Specificația CDMI folosește principii REST pentru designul interfeței.

7. Concluzii

Tehnologia Cloud Computing este privită ca fiind la stadiul de adoptare timpurie în Europa. În timp ce majoritatea guvernelor sunt dornice să adopte soluții Cloud, crearea unor politici durabile, a unor strategii şi planuri de implementare a fost dificilă în multe cazuri. Fiind o tranziție majoră în cadrul operațiunilor şi responsabilităților, această tranziție s-ar putea confrunta cu reticențe. Cu toate acestea, există exemple care pot fi urmate şi îmbunătățite, stimulente din partea guvernelor pentru a susține această tranziție şi exemple de bune practici.

Credibilitatea şi securitatea, precum şi aspectele legale şi de reglementare trebuie să fie coordonate perfect între țările cu mare putere de decizie în acest domeniu. O soluție sau un cadru de lucru nu se va armoniza în rândul tuturor organizațiilor, aşa cum nici un model de Cloud nu se va potrivi cu toate profilurile de risc.

Atât la nivel european, dar şi la nivel național, interoperabilitatea şi Cloud computing-ul reprezintă subiecte de actualitate, iar dezvoltarea soluțiilor destinate îmbunătățirii serviciilor publice au la bază aceste două mari teme.

Articolul îşi propune să efectueze o introducere generalizată a principiilor de stocare a datelor în Cloud cu scopul de a face cunoscută această metodă la nivelul național al sistemelor publice şi de a evidenția beneficiile unei stocări securizate, accentuând în speță nivelul de interoperabilitate oferit de această soluție.

BIBLIOGRAFIE

1. Valorificarea potențialului cloud computing-ului în Europa, http://europa.eu/rapid/press-release_MEMO-12-713_ro.htm, 27 septembrie 2012
2. Agenda digitală: o nouă strategie de impulsionare a productivității întreprinderilor şi administrațiilor europene prin intermediul cloud computingului, http://europa.eu/rapid/press-release_IP-12-1025_ro.htm, 27 septembrie 2012

3. Cloud for Europe, www.cloudforeurope.eu
4. Cloud Data Management Interface (CDMI) v1.1.1, http://www.snia.org/cdmi, 19 martie 2015
5. S.O. Kuyoro, F. Ibikunle, O. Awodele, Cloud Computing Security Issues and Challenges, International Journal of Computer Networks(IJCN), vol. 3, nr. 5, 2011.
6. BlueKrypt - Cryptographic Key Length Recommendation, www.keylength.com, 26 februarie 2015
7. Hogg, Smith, Chong, Hollander, Kozaczynski, Brader, Delgado, Taylor, Wall, Slater, Imran, Cibraro, Cunningham, Message Screenning, http://soapatterns.org/design_patterns/message_screening
8. K. VIjay Kumar, N Chandra Sekhar Reddy, B. Srinivas Reddy, Preserving Data Privacy, Security Models and Cryptographic Algorithms in Cloud Computing, International Journal of Computer Engineering and Applications, vol. VII, nr. 1, iulie 2014.
9. Litigation Hold (Preservation Orders or Hold Orders) Definition, http://searchstorage.techtarget.com/definition/litigation-hold
10. MoReq2, http://moreq2.eu/
11. Metadata Innovation Dublin Core, http://dublincore.org/
12. **SANDERSON, R.; CICCARESE, P.; VAN DE SOMPEL, H.:** Open Annotation Data Model .