STUDIA UNIVERSITATIS MOLDAVIAE, 2019, nr.2(122)
Seria "Științe exacte și economice" ISSN 1857-2073 ISSN online 2345-1033 p.23-37

CZU: 004.738.5

UTILIZAREA TEHNOLOGIILOR CLOUD PENTRU SISTEME PARALELE DE CALCUL PERFORMANT

Ionel ANTOHI

Universitatea de Stat din Moldova

Apariția fenomenului, cunoscut sub denumirea de *cloud computing,* reprezintă o schimbare fundamentală în felul în care sunt dezvoltate, livrate, actualizate, întreținute și plătite serviciile IT din cadrul organizațiilor moderne. Serviciile cloud computing devin adevărate incubatoare pentru noile aplicații, menite să răspundă cererii de obținere rapidă de informații și acces facil la acestea. Dinamica pieţelor de desfacere și a proceselor economice obligă managerii IT să devină din ce în ce mai agili în adaptarea aplicațiilor și serviciilor IT&C pentru a răspunde rapid cerințelor din ce în ce mai complexe ale persoanelor din mediul de afaceri.

Cuvinte-cheie: *cloud computing, performanță, sisteme distribuite, sisteme cloud HPC, cluster, grid computing, procese decizionale, securitate și riscuri în cloud.*

USING CLOUD TECHNOLOGIES FOR HPC TYPE CLUSTER

The phenomenon, known as cloud computing, is a fundamental change in how IT services within modern organizations are being developed, delivered, updated, maintained and paid. Cloud computing services become true incubators for new applications designed to meet the demand for quick information and quick access to information. Market dynamics and business processes force IT managers to become more agile in adapting IT & C applications and services to quickly respond to increasingly complex business needs.

Keywords: *cloud computing, performance, distributed systems, HPC cloud systems, cluster, grid computing, decision-making, security and cloud risks.*

Introducere

Cloud computing semnifică convergența a două tendinţe majore ale IT-ului zilelor noastre: *eficienţa IT* – unde puterea calculatoarelor moderne este utilizată mai eficient printr-o scalare înaltă a resurselor de hardware şi software şi *agilitatea de business* – unde tehnologia informațională poate fi folosită ca instrument competitiv pe piaţă prin livrare rapidă, loturi paralele de procesare, utilizarea instrumentelor de inteligenţă a afacerilor, care necesită calcul intensiv și aplicații mobile interactive și care răspund în timp real cerinţelor utilizatorului. Există, probabil, atâtea definiţii câţi comentatori/autori ai subiectului. Definiţia formală a ceea ce înseamnă conceptul de cloud computing este următoarea: *este un model de organizare a serviciilor IT, care presupune că serviciile de prelucrare (hardware şi software) sunt livrate la cerere către clienţi prin intermediul reţelelor de calculatoare, într-o manieră de autoservire, independent de locaţia dispozitivului utilizat.* Resursele necesare pentru a livra un serviciu la un nivel înalt de calitate sunt distribuite, scalabile în mod dinamic, livrate rapid prin virtualizare şi lansate cu minim de interacţiune din partea furnizorului de servicii. Utilizatorii plătesc pentru acest serviciu ca o cheltuială operaţională, fără a necesita investiţii semnificative de capital iniţial. Serviciile de cloud utilizează un sistem de măsurare care alocă resursele de calcul în blocuri corespunzătoare nivelului de prelucrare solicitat.

Cloud computing este un subiect vast, care cuprinde mai multe subiecte diferite. Pentru a descrie în mod adecvat ofertele de cloud computing, trebuie să tratăm în detaliu elemente de infrastructură, arhitecturi orientate spre servicii, reţele sociale, protocoale de comunicare, standarde, interfeţe de programare aplicaţii (API[1]) şi zeci de alte subiecte. Demersul de a scrie o carte atotcuprinzătoare despre acest subiect se va limita permanent doar la noţiunile introductive ale domeniului pentru a putea avea o finalitate.

Cloud computing este conceptul care aduce cele mai mari schimbări în evoluţia dinamică a tehnologiei informaţiei şi comunicaţiei. Zilnic oamenii generează, accesează, prelucrează şi stochează seturi de date noi, beneficiind de mai multă putere de procesare decât oricând înainte. Pentru corporaţii, o consecinţă profundă a acestui consum digital în creştere este necesitatea de a investi permanent sume de bani în echipamente de

[1] API – Application Programming Interface – Interfeţe de programare a aplicaţiilor.

prelucrare şi stocare. Odată cu creşterea puterii de procesare a echipamentelor de calcul, apariţia tehnologiilor de virtualizare a fost un real succes, oferind o nouă dimensiune portabilităţii sistemelor de operare şi aplicaţiilor. Extensibilitatea sistemului informaţional este încă limitată de capacitatea de procesare, memoria RAM şi spaţiul de stocare din cadrul centrului de calcul al companiei. Înfiinţarea sau modernizarea unui centru local de calcul (on-premise) implică investiţii şi activităţi periodice de achiziţie şi configurare a echipamentelor fizice, configurarea instrumentelor de virtualizare, achiziţia şi instalarea sistemelor de operare şi a aplicaţiilor necesare, configurarea mediilor de comunicaţie, inclusiv echipamentele fizice de reţea, configurarea firewall-urilor şi a echipamentelor de stocare de tip Enterprise (SAN, NAS). Elaborarea politicilor de guvernanţă, monitorizare şi întreţinere sau implementarea standardelor de calitate (ISO 9001) sau de securitate (ISO 27001) implică costuri suplimentare pentru companie, responsabilitatea investiţiilor fiind împărţită pe mai multe niveluri ale piramidei organizaţionale şi pe întreg ciclul de viaţă al centrului de calcul.

O practică frecventă în achiziţia de echipamente fizice este acea de supradimensionare a capacităţilor de procesare, memorie RAM şi stocare, justificate de o creştere a nevoilor în timp. În fapt, majoritatea echipamentelor de calcul nu sunt folosite la potenţialul lor complet aproape niciodată, ajungând să fie înlocuite din cauza uzurii morale, lipsei de suport din partea producătorului sau lipsei pieselor de schimb.

Cloud computing oferă o alternativă modernă la centrul de calcul tradiţional. Un furnizor de cloud este singurul responsabil pentru achiziţionarea de echipamente fizice şi de întreţinerea acestora, oferind o gamă largă de servicii şi configuraţii utilizabile după necesarul de prelucrare a fiecărei companii. Astfel, închirierea de servicii din cloud transformă investiţiile de capital în dispozitive fizice şi licenţe în costuri operaţionale, permiţând utilizarea unor fonduri financiare semnificative în scopuri dedicate afacerii de bază. De asemenea, tehnologiile cloud permit accesul punctual la resurse hardware sau software, care în mod normal ar fi prea scumpe pentru a putea fi achiziţionate. Din punct de vedere economic, eficienţa serviciilor cloud este justificată şi prin faptul că acele echipamente sau licenţe sunt plătite doar în momentul în care sunt utilizate.

Prezentul demers științific are ca scop abordarea teoretică şi aplicativă pentru cercetarea platformelor cloud computing, cu scopul de a realiza implementarea aplicaţiilor de comunicaţii pe o astfel de platformă.

1. Caracteristici ale cloud computing-ului

Conceptul de cloud computing a devenit atât de omniprezent în activitatea economică şi socială, încât pare aproape normal să ştim sau să înţelegem ce înseamnă. În fapt, sesizăm că multe dintre principiile şi conceptele care guvernează acest concept devin pe zi ce trece tot mai transparente faţă de utilizatorul final. Caracteristicile esenţiale ale infrastructurilor cloud includ autoservice la cerere, acces în bandă largă la reţea, resurse utilizate în mod partajat, flexibilitate rapidă şi instrumente de comensurare a calităţii serviciilor oferite. Accesul la cloud este permis în mod concurent unui număr mare de consumatori prin intermediul tehnologiilor de virtualizare cu funcţii de *autoscalare* şi provizionare automatizate în funcţie de numărul de cereri de procesare. Din punct de vedere teoretic, cantitatea de resurse de procesare şi stocare de care poate beneficia un utilizator este nelimitată.

1.1. Noțiuni generale

Anii care au urmat lansării TCP/IP şi a Internetului au determinat o cerere constantă de putere de calcul şi necesitatea de a centraliza într-un singur loc datele companiilor, determinând apariţia şi evoluţia a trei tehnologii-cheie care au revoluţionat domeniul reţelelor de calculatoare: clusterele de calculatoare, grid computing-ul şi cloud computing-ul. În accepţiune generală, definirea simplificată a acestor concepte şi tehnologii este:

- clusterele de calculatoare sunt colecţii de PC-uri sau servere interconectate într-o reţea locală de calculatoare caracterizată prin viteză mare de transmitere a datelor şi latenţă redusă;
- grid computing-ul presupune interconectarea unui număr mare de resurse de procesare într-o reţea de tip WAN;
- cloud computing oferă servicii de acces la resurse de prelucrare, stocare sau aplicaţii prin intermediul Internetului. Infrastructura fizică de organizare a centrelor de cloud este transparentă faţă de utilizatorii finali, calea de acces şi administrare fiind reprezentată de browser-ele web, aplicaţii specifice de comunicare securizată sau API-uri specifice.

Primele forme de *cluster* [1] apar în perioada anilor 1994 şi introduceau ideea de a unifica puterea de calcul a mai multor calculatoare independente cu scopul de a agrega într-un singur punct puterea de prelucrare. Succesul incipient al acestor modele a fost determinat de costul redus al PC-urilor comparativ cu super-

computerele specializate. Chiar dacă anticipate iniţial, cheltuielile suplimentare legate de conectica de reţea între nodurile clusterului au impus reanalizarea modelului de funcţionare a acestora, majoritatea cercetărilor concentrându-se pe îmbunătăţirea metodelor de eficientizare a comunicaţiei şi disponibilităţii. Caracteristicile prelucrării de tip cluster sunt:

- creşterea puterii de calcul în cadrul organizaţiei;
- asigurarea disponibilităţii anumitor servicii IT prin utilizarea tehnologiilor de load balancing (NLB);
- gestionarea eficientă a momentelor de prelucrare masivă a datelor;
- utilizarea eficientă şi simplitatea gestiunii fizice a resurselor (procesoare, memorie RAM, spaţiu pe disc şi lăţime de bandă de reţea).

Clusterele sunt utilizate în prezent în arhitecturi on-premise în special pentru disponibilitatea ridicată a serviciilor de web şi baze de date, dar pot fi utilizate şi în modele de servicii IaaS[2] în cloud.

Termenul *grid computing* a fost utilizat pentru prima dată în 1998 [2] presupunând agregarea şi partajarea puterii de procesare a mai multor calculatoare într-un format transparent faţă de utilizatorul final şi cu o adresare şi accesibilitate globală. Viziunea care stă la baza conceptului de grid este de a oferi acces la o capacitate teoretic nelimitată de resurse de procesare a informaţiilor şi putere de calcul într-un mod care este la fel de simplu şi asemănător ca accesul la energia electrică.

Grid-ul poate fi implementat şi utilizat atât în mod *privat,* asimilat, dar fără a se confunda, de multe ori cunoscut şi sub denumirea de prelucrare distribuită *(distributed computing),* cât şi în mod *public* cu diferite forme de implementare şi manifestare. În prezent grid-ul public este întâlnit sub mai multe forme: grid-urile de cercetare, sub forma reţelelor *peer-to-peer* (P2P; exemplu: reţelele de sharing tip torrente şi *reţelele de botnet* utilizate în scop distructiv). În anul 2009 am înregistrat cea mai spectaculoasă revenire a grid computing-ului, sub forma reţelelor peer-to-peer, legată de introducerea monedelor electronice şi a instrumentelor de minare a acestora. Rolul pe care îl au modelele de prelucrare grid este exploatat de cercetarea ştiinţifică şi mai puţin în domeniul activităţilor economice. Printre avantajele majore ale grid-ului amintim:

- disponibilitatea unui număr mare de calculatoare conectate la Internet, oferind putere de calcul într-un mod simplu şi transparent, atunci când este nevoie de ea;
- utilizarea eficientă a resurselor de procesare răspândite global;
- limitarea costurilor cu resursele fizice necesare anumitor investigaţii ştiinţifice;
- accesul ca nod de prelucrare în grid este voluntar şi se bazează pe dorinţa oamenilor de a-i ajuta pe alţii.

În viziunea unor autori [3], ar trebui să existe o delimitare în definirea conceptelor de grid ca metodă de prelucrare distribuită utilizată de cercetători şi *volunteer computing* – participare voluntară la procesarea de date. Volunteer computing-ul este mult mai des utilizat în reţelele P2P, dar se bazează pe aceleaşi principii de funcţionare, un nod al reţelei P2P primind sarcini de execuţie şi prelucrare în mod transparent. Nu împărtăşim părerea autorilor conform căreia acest concept face parte din categoria cloud computing-ului, pentru că diferă în modul de implementare, întreţinere şi în calitatea serviciilor pe care le oferă (QoS[3]) şi nu este atractiv pentru companii. Un exemplu sugestiv este acela în care o companie, care are nevoie la finalul unei luni de prelucrarea unui volum imens de date pentru obţinerea unor rapoarte de sinteză, lansează la ora 3 AM o cerere de prelucrare către anumite reţele de volunteer computing în vederea obţinerii rapide a informaţiilor. Neavând niciun contract sau un acord de utilizare a serviciilor (SLA[4]), este foarte probabil ca numărul de noduri de procesare în aria geografică de activitate a firmei să fie foarte mic, ceea ce ar putea provoca întârzieri în obţinerea la timp a rezultatelor dorite. Conform altor surse [4], cloud-ul de tip comunitate, *community cloud,* este în prezent predecesorul de drept al grid computing-ului, dar părerea noastră este că cele două modele de procesare distribuită vor coexista o bună perioadă de timp. Cloud computing-ul apare în anul 2006/2007 ca o alternativă la cerinţele în creştere ale companiilor, în special pentru putere de calcul garantată, sigură, flexibilă şi pentru a răspunde tendinţei de mobilitate a propriilor angajaţi. Specificul serviciilor cloud este acela de a fi livrate prin intermediul Internetului sub forma aplicaţiilor specifice (SaaS[5]), ca platforme de dezvoltare şi prelucrare a propriilor aplicaţii (PaaS[6]) sau ca metodă de emulare a centrelor de calcul din „on-premise" (IaaS).

[2] IaaS – Infrastructure as a service – Infrastructură ca și serviciu.
[3] QoS – Quality of service – Calitatea serviciilor.
[4] SLA – Service-Level Agreement – Acord de utilizare a serviciilor.
[5] SaaS – Software as a service – Software ca serviciu.
[6] PaaS – Platform as a Service – Platformă ca și Serviciu.

În scurt timp de la lansare, marile companii din domeniul IT&C au speculat potenţialul enorm de afaceri al noului domeniu, investind sume considerabile de bani în centre de date specializate, cercetare, sisteme de întreţinere la standarde ridicate, precum şi centre de suport dedicate. Utilizatorii finali sunt interesaţi, deoarece serviciile oferite sunt la un preţ rezonabil şi pot fi accesate de pe orice browser, oferind acces la resursele informaţionale şi de calcul din orice locaţie, facilitând astfel colaborarea şi lucrul de la distanţă. Departamentele IT din cadrul companiilor au devenit interesate de noul model, bazându-se pe o reducere a investiţiilor de capital, eliminarea (cel puţin teoretică) a constrângerilor legate de puterea de prelucrare şi spaţiul de stocare, dezvoltare şi implementare rapidă a aplicaţiilor şi proceselor de afaceri, precum şi pe simplificarea modului de întreţinere a mediilor de reţea complexe.

În sens aproape unanim abordează specialiştii în domeniu conceptul de cloud pornind de la definiţia NIST [5], conform căreia: *cloud computing-ul permite furnizorilor de servicii cloud şi consumatorilor stabilirea unui set iniţial de aşteptări cu privire la managementul, securitatea şi interoperabilitatea, precum şi determinarea valorii juste generate de utilizarea tehnologiei cloud.*

Marea provocare pentru furnizorii de cloud devenea atingerea unui nivel al serviciilor disponibile în termeni de securitate, fiabilitate şi performanţă comparabile cu cele din infrastructurile locale. Pentru măsurarea corectă a fost necesară introducerea unui set de indicatori de calitate şi performanţă la nivel ridicat pentru câştigarea încrederii clienţilor. Astfel, procentele de disponibilitate a serviciilor în cloud sunt mult mai mari decât pot fi asigurate de centrele de calcul locale. Amazon, Google şi Microsoft specifică un timp de peste 99,9% a disponibilităţii*(uptime)* serviciilor în SLA-urile proprii, în anumite condiţii oferind despăgubiri clienţilor dacă procentul de uptime scade sub acest procent. La începutul anilor 2000, odată cu dezvoltarea tehnologiilor de virtualizare, conceptul de consolidare a serverelor fizice vechi a asigurat continuitatea funcţionării unor aplicaţii învechite pe o perioadă considerabilă de timp şi o optimizare a costurilor cu întreţinerea echipamentelor fizice vechi. În schimb, extensibilitatea şi fiabilitatea virtualizării era limitată de resursele fizice ale centrului de calcul, fiind necesare investiţii permanente pentru îmbunătăţirea acestor parametri. De asemenea, modificarea aplicaţiilor, actualizarea sistemelor de operare şi alte operaţiuni de întreţinere introduc timp de inactivitate şi de indisponibilitate în rularea aplicaţiilor. Tehnologiile cloud au preluat succesul virtualizării oferind clienţilor capacităţi de extensibilitate a maşinilor virtuale, teoretic nelimitate, operaţiunile de întreţinere fiind planificate la intervale mari de timp. Specialiştii în proiectarea infrastructurilor de aplicaţii sau maşinilor virtuale în cloud sunt instruiţi să configureze disponibilitatea serviciilor prin alocarea seturilor de resurse în locaţii geografice diferite.

Epoca cloud computing-ului este considerată astăzi o piatră de hotar a tehnologiilor informaţionale, impactul acestora în modul în care se vor derula afacerile viitorului fiind greu de anticipat. Puterea de calcul teoretic nelimitată, accesul de oriunde şi colaborarea la un alt nivel va avea un impact direct asupra eficienţei departamentelor de IT prin schimbarea modului în care îşi vor desfăşura atribuţiile şi asupra activităţii economice în general prin accesul mai rapid la activele informaţionale din cadrul companiei.

1.2. Caracteristicile esenţiale ale tehnologiilor cloud

În opinia unor autori [6], cloud computing nu are o definiţie standard, dar accepţiunea generală simplificată este acea de a oferi servicii computaţionale prin intermediul unei reţele de calculatoare, în special Internetul.

În sens mai larg [7], cloud computing-ul se referă la aplicaţiile şi serviciile care rulează pe sisteme de reţea distribuite, care folosesc tehnicile de virtualizare a resurselor şi care pot fi accesate în mod general prin intermediul Internetului, folosind protocoalele şi serviciile standard de reţea. Cloud-ul asigură transparenţa resurselor fizice şi a configuraţiilor acestora, utilizatorii finali având percepţia că resursele de care dispun sunt teoretic nelimitate.

Alţi autori renunţă să caute/ofere definiţii [8] concrete ale conceptului, preferând să prezinte caracteristicile esenţiale ale cloud-ului:

- scalabilitate[7] masivă;
- abilitatea de a aloca cu uşurinţă resurse;
- o platformă de management al serviciilor.

[7] Menţionăm că în DEX nu există termenul *scalabilitate,* el fiind adoptat în limba română relativ recent prin translatare fonetică a termenului englezesc *scalability* – folosit cu preponderenţă în domeniul tehnologiilor informaţionale. Utilizarea corectă ar fi un derivat al termenului *scalare,* care provine din psihologie şi presupune măsurarea intensităţii opţiunilor, atitudinilor şi cunoştinţelor.

Menţionăm că în opinia generală printre caracteristicile esenţiale este menţionată şi noţiunea de *securitate* cu toate implicaţiile triadei CIA[8] [9].

În demersul nostru de a accede spre zona tehnică, o altă caracteristică de bază este *autoscalarea,* cunoscută şi sub numele de *provizionare,* care asigură flexibilitatea necesară alocării corecte în timp a resurselor de prelucrare. În sensul acestei afirmaţii, provizionarea nu este acelaşi lucru cu configurarea iniţială a mediului de lucru intern (on-premise) şi a celor din cloud pentru accesul la servicii: configurare reţea, implementarea mecanismelor de balansare a cererilor (NLB), configurarea firewall-urilor, configurarea iniţială a imaginilor serverelor. *Autoscalarea* se referă la menţinerea unei alocări dinamice a resurselor de procesare în funcţie de numărul cererilor de procesare în timp real. De exemplu, un cluster din cloud format din două noduri de procesare poate fi configurat să oprească funcţionarea unuia dintre noduri *(decomisionare)* în momentul în care cantitatea de cereri de prelucrare este redusă. În contextul unei creşteri a cererilor, bazându-se pe indicatorii de calitate a latenţei în oferirea răspunsurilor, sistemele de management al cloud-ului, folosind funcţiile de autoscalare pornesc la propriu serverul decomisionat fără a fi necesară intervenţia administratorilor în acest scop. Utilizând acelaşi mecanism de monitorizare a încărcării cu operaţiuni de procesare, *provizionarea* alocă resurse suplimentare de cloud, care iniţial nu au fost prevăzute. Din punct de vedere economic, autoscalarea presupune un mecanism de economisire a resurselor şi, implicit, facturi mai mici; provizionarea presupune plăţi suplimentare pentru putere de procesare în vederea îndeplinirii unei sarcini de lucru punctuale.

Scalabilitatea în alocarea resurselor a apărut înaintea tehnologiilor cloud cu scopul de a construi arhitecturi de prelucrare a datelor care să combată efectele negative ce afectează experienţa de lucru a utilizatorilor [10] în momentul în care încearcă rularea unor procese masive, iar timpul de răspuns este în scădere.

Din punctul de vedere al aplicaţiilor, *scalabilitatea* este măsurată prin numărul de utilizatori care pot accesa şi utiliza în parametri acceptabili aceeaşi resursă. Punctul în care utilizatorii nu se mai pot conecta la acea resursă este cunoscut sub denumirea de *limită de scalabilitate*. În *scalabilitatea pe verticală,* optimizarea procesului şi obţinerea de indicatori superiori de accesibilitate se realizează prin suplimentarea sau adăugarea, sau achiziţia de dispozitive hardware la cele existente: procesor, memorie RAM, spaţiu de stocare, lăţime de bandă de reţea. Altă formă de manifestare a scalabilităţii este *scalabilitatea pe orizontală,* care presupune alocarea de putere de calcul prin adăugarea unor noi noduri (alte calculatoare destinate procesării aceleiaşi cereri). Din punct de vedere fizic, scalabilitatea pe orizontală este mai costisitoare, dar ea a devenit eficientă şi aplicabilă atât în local (on-premise), cât şi în cloud, în contextul utilizării tehnologiilor de virtualizare.

În completarea principalelor caracteristici ale cloud-ului, alţi autori [11] identifică noile oportunităţi de afaceri ale conceptului şi introduc în lista de caracteristici esenţiale conceptele de: utilizare partajată *(multi-tenancy),* plata doar pentru cât utilizezi *(pay-as-you-go)* şi autoservire *(self-service).*

Partajarea resurselor sau *multitenanţa* este diferită de modelele clasice de prelucrare (centrele de calcul locale), care presupun deţinerea unor echipamente specializate (servere) configurate şi izolate prin securizare în mod corespunzător. Aceste resurse fizice pot executa operaţiuni de procesare doar pentru proprietarul lor. În cloud, resursele fizice aparţin de drept unei companii (CSP[9]) care furnizează servicii cloud prin partajarea acestor resurse către clienţii săi. Izolarea accesului se realizează pe mai multe straturi: nivel reţea, nivel maşină virtuală şi nivel aplicaţie, stratul fizic de funcţionare a serviciilor fiind complet transparent faţă de utilizatori. Modelul economic al cloud-ului se bazează pe principiul *plăteşti atât cât utilizezi* (*pay-as-you-go*), referindu-se la puterea de prelucrare şi stocare alocată execuţiei unui proces, precum şi pe durata utilizării acestora. Fiecare contract de furnizare a acestor servicii conţine în mod detaliat preţul pe unitate de procesare, RAM, aplicaţii, capacităţi de stocare, sau sunt oferite pachete predefinite de maşini virtuale, care conţin configuraţiile şi aplicaţiile dorite. Reducerea costurilor pentru pachetele pre-configurate se realizează prin configurarea tehnică a autoscalabilităţii.

Autoservirea (self-service) în termenii tehnologiei informaţiei este un concept elaborat bazat pe capacitatea utilizatorilor din business de a utiliza instrumente IT în îndeplinirea sarcinilor de serviciu. *Knowledge worker* (lucrătorul cu cunoştinţe), în viziunea lui Peter Drucker, este principalul beneficiar al conceptului de *self-service.* Întâlnită destul de rar în companiile din Republica Moldova, autoservirea cu instrumente de procesare este totuşi destul de des întâlnită în procesele de alocare a spaţiului pe discurile companiei sau în instrumentele de *lucru colaborativ* cu portalurile bazate pe tehnologiile SharePoint.

[8] CIA – Confidentiality Integrity Availability – Confidenţialitate Integritate Disponibilitate.
[9] CSP – Cloud Service Provider – Furnizor de servicii cloud.

Din punct de vedere tehnic, self-service presupune utilizarea unor instrumente simple pentru utilizator în vederea construirii unui mediu transparent complex de execuţie a anumitor operaţiuni. Instrumentele de asistenţă a operaţiunilor *(wizard)* şi-au dovedit atractivitatea chiar de la apariţia lor în primele pachete software cu interfaţă grafică. Însă, în partea „nevăzută" a lor aplicaţiile, sistemele de management şi infrastructurile de aplicaţii cloud trebuie să fie capabile să execute operaţiuni, uneori de o complexitate deosebită. În deservirea acestor tipuri de cereri fiecare infrastructură cloud trebuie să conţină un set de instrumente specifice de *automatizare* a operaţiunilor, prin folosirea unor scripturi bazate pe linii de comenzi intuitive şi cu suficient de mulţi parametri, încât să poată fi folosite în cât mai multe cazuri. Aceste comenzi sunt folosite şi de specialişti în scopul provizionării iniţiale a infrastructurilor de aplicaţii şi servicii şi sunt specifice fiecărui furnizor de cloud:

- Microsoft Azure şi Office 365 folosesc pentru automatizare PowerShell, un limbaj de *scripting* şi care permite automatizarea majorităţii tipurilor de procese din cloud;
- Google foloseşte Google Apps Script, care este un derivat pentru cloud al limbajului JavaScript, precum şi SDK pentru API-urile de interconectare la servicii;
- Azure Web Services foloseşte limbajul Java şi Ruby pentru automatizare.

Modelele de autoservire în cloud trebuie să stabilească în mod imperativ un mod detaliat de comensurare a calităţii serviciilor (QoS[10]) pe care le oferă (Yeluri & Castro-Leon, 2014), aşteptările utilizatorilor fiind întotdeauna legate de un timp de răspuns cât mai rapid la cererile lor.

Goetsch (2014) expune o ierarhie a caracteristicilor cloud-ului, oferind în acest fel o perspectivă a dependenţei între concepte (*a se vedea* Fig.1).

Într-un mod sec, scurt şi concis, specific standardelor americane, definiţia din NIST [5] a cloud-ului specifică existenţa a cinci caracteristici esenţiale, trei modele de servicii şi patru modele de implementare. În viziunea unor critici, orice ce depăşeşte această delimitare este literatură, fără efecte de implementare rapidă în practica companiilor americane.

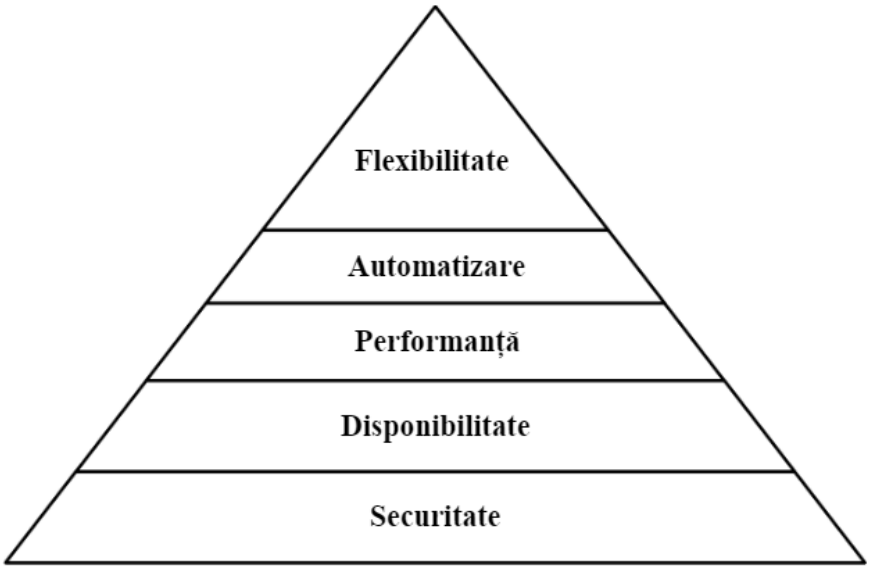


Fig.1. Ierarhia caracteristicilor cloud computing.
Sursa: Prelucrare după (Goetsch, 2014)

Aceste caracteristici sunt:

- Autoservire la cerere;
- Acces pe bază de reţele de calculatoare de bandă largă;
- Partajarea resurselor de prelucrare;
- Flexibilitate rapidă;

[10] QoS – Quality of Services (calitatea serviciilor oferite) – este un concept utilizat în tehnologia informaţiei care semnifică percepţia utilizatorilor despre calitatea unui serviciu oferit prin intermediul reţelelor de calculatoare.

• Servicii comensurabile, care, în sensul definiţiei, se referă la faptul că sistemele cloud trebuie să asigure monitorizarea, controlul şi optimizarea utilizării resurselor prin intermediul unor instrumente *specifice de măsurare, la un nivel de abstractizare specific fiecărui tip de serviciu pus la dispoziţie.*

În domeniul tehnologiilor informaţionale, *abstractizarea* este o tehnică de gestionare a complexităţii sistemelor informatice. Scopul său este de a stabili un anumit nivel de complexitate care poate fi înţeles de un utilizator al sistemului, suprimând detaliile esenţiale care guvernează nivelul de lucru la care utilizatorul are acces.

Cloud computing abstractizează detaliile de implementare a sistemului faţă de utilizatori şi dezvoltatori. Aplicaţiile rulează pe sisteme fizice care nu sunt specificate, datele sunt stocate în locaţii care nu sunt cunoscute, administrarea sistemelor este *externalizată,* iar omniprezenţa utilizatorilor la acele aplicaţii nu este vizibilă decât în zona de cloud închiriată de companie. Unde se află totuşi infrastructura fizică a cloud-urilor? Fiind proiecte de maximă importanţă, suntem convinşi că se află în zone geografice retrase şi cu o securitate fizică a perimetrului bine asigurată. În diferite surse şi prospecte [7] identificăm faptul că furnizorii de cloud au criterii bine stabilite de alegere a zonelor de amplasare a centrelor de calcul pentru cloud:

• Zone în care costul electricităţii este redus. Centrele de date pentru cloud conţin sute de mii de calculatoare, fapt ce determină un consum ridicat de energie electrică;

• Zone care pot oferi surse de energie regenerabilă (zone cu soare, cu vânt constant sau alte forme);

• Zona să dispună de resurse de apă din abundenţă. Apa rece reprezintă o metodă ieftină de răcire a echipamentelor de calcul, transformând-o în apă care poate fi transformată ulterior în aburi la costuri reduse, generând astfel energie suplimentară funcţionării;

• Zona să permită, din punctul de vedere al terenului, implementarea reţelelor de comunicaţii date de mare viteză. Un centru de calcul pentru cloud are nevoie de linii redundante de acces la Internet, pentru a asigura disponibilitatea garantată prin contracte. Această cerinţă implică amplasarea la distanţe cât mai egale a mai multe noduri mari ale reţelei de Internet, regăsite de obicei în marile centre urbane;

• Costul pământului să fie redus şi poziţionat în zone discrete cât mai izolat de zonele populate;

• Furnizorii de cloud plătesc taxe locale pentru proprietăţi (şi nu numai), de aceea vizează locaţiile geografice în care pot obţine reduceri de taxe semnificative;

• Locaţii geografice în care activitatea seismică este cât mai redusă sau fără influenţe climaterice şi meteorologie semnificative.

Toate aceste strategii de amplasare a centrelor cloud oferă furnizorilor o marjă de profit suficient de bună, astfel încât să poată oferi serviciile specifice la preţuri accesibile. În acelaşi timp, sistemele de tip cloud permit utilizarea la maximum a puterii de calcul a serverelor sau a dispozitivelor de prelucrare şi stocare. Calculatoarele pe care le folosim în mod curent pentru activităţile economico-sociale nu ajung decât în cazuri excepţionale să folosească întreaga putere a procesorului sau întreaga memorie RAM şi doar în rare cazuri tot spaţiul disponibil de pe disc. În cloud, toate aceste resurse fizice sunt utilizate permanent în proporţie de 60 până la 80% din capacitate [4]. Această utilizare este posibilă datorită funcţiilor de *virtualizare* implementate în toate modelele de cloud. Spre exemplu, dacă dispunem de un server care are 16 Gb de RAM şi 1024 Gb de spaţiu liber pe disc, putem crea cel puţin şase maşini virtuale cu 2 Gb RAM fiecare şi un spaţiu pe disc de 100 Gb. Restul resurselor sunt păstrate pentru provizionări ulterioare. Prin folosirea funcţiilor de autoscalare putem crea chiar mai multe maşini virtuale pe acelaşi server care să deservească în serie diferite cereri. Majoritatea specialiştilor în domeniul economic şi al tehnologiilor de afaceri sunt oarecum preocupaţi de modul în care funcţionează cloud-ul în partea sa nevăzută, principalele întrebări ce rezultă fiind legate de instrumentele de management al capacităţii cloud-ului de a deservi nevoile de afaceri pe un termen îndelungat de timp la un nivel de calitate specificat în SLA (Service Level Agreement) [13].

Având în vedere varietatea de modele de implementare şi organizare a serviciilor oferite de cloud, este foarte important ca factorii de decizie din cadrul companiilor să fie corect informaţi în legătură cu tipul acestora şi, mai mult, să fie pregătiţi să facă o evaluare corectă a proceselor de afaceri care pot fi deservite de furnizori cloud.

2. HPC Cluster Computing

Cluster computing este nu altceva decât două sau mai multe computere care sunt conectate în rețea pentru a oferi soluții după cum este necesar. Totuși, această idee nu ar trebui să fie confundată cu un model mai general de computere client-server, deoarece ideea din spatele clusterelor este destul de unică.

Un grup de computere se alătură puterilor computaționale ale nodurilor de calcul pentru a oferi o putere computațională mai combinată. Prin urmare, ca și în modelul client-server, mai degrabă decât un simplu client care face cereri de unul sau mai multe servere, cluster computing utilizează mai multe mașini pentru a oferi un mediu de calcul mai puternic, probabil printr-un singur sistem de operare.

În structura sa cea mai simplă, după cum s-a menționat mai sus, clustele HPC sunt destinate să utilizeze calculul paralel pentru a aplica mai multă forță de procesor pentru aranjarea (rezolvarea) unei probleme. Există numeroase cazuri de calcul experimental care utilizează diferite procesoare cu costuri reduse ca parte a paralelului pentru a efectua cantități uriașe de operațiuni. Acest lucru este semnalat ca calcul paralel. Un cluster de înaltă performanță este în mod regulat alcătuit din noduri (numite și lame).

Grupurile HPC vor avea în mod obișnuit un număr mare de calculatoare (numite adesea „noduri") și, în general, majoritatea acestor noduri ar fi configurate identic. Deși din partea exterioară clusterul poate arăta ca un singur sistem, funcționarea internă pentru a face acest lucru poate fi destul de complexă.

Ideea este că sarcinile individuale care alcătuiesc o aplicație paralelă ar trebui să funcționeze la fel de bine pe orice nod pe care sunt expediate. Cu toate acestea, unele noduri dintr-un cluster au adesea unele diferențe fizice și logice.

Nod de calcul. Un nod de calcul este locul unde se efectuează tot calculul. Majoritatea nodurilor dintr-un cluster sunt de obicei noduri de calcul. Cu un scop final specific pentru a da un aranjament general, un nod de calcul poate executa una sau mai multe sarcini, ținând cont de sistemul de planificare.

Nodul principal. Clusterele sunt medii complexe, iar administrarea fiecărui segment individual este esențială. Nodul de administrare oferă numeroase capacități, printre care: respectarea statutului nodurilor individuale, emiterea ordinelor de administrare a nodurilor individuale la problemele corecte sau comandarea capacităților de administrare (de exemplu, pornirea / oprirea alimentării).

Nu se poate subestima importanța gestionării clusterului. Este un imperativ în încercarea de a coordona activitățile unui număr mare de sisteme.

Depozitare. În aplicațiile care continuă să ruleze pe un cluster, nodurile de calcul trebuie să aibă acces rapid, fiabil și concurent la un cadru de stocare. Dispozitivele de stocare sunt asociate în mod specific cu nodurile sau asociate cu un nod de stocare care va fi responsabil de facilitarea solicitărilor de stocare.

Scopul general al HPC este de a rula aplicațiile mai rapid sau de a rula probleme care nu pot sau nu se vor executa pe un singur server. Pentru a face acest lucru, trebuie să rulați aplicații paralele pe noduri separate. Deși ați putea utiliza un singur nod și apoi să creați două VM-uri, este important să înțelegeți modul în care aplicațiile rulează pe diferite servere fizice și cum gestionați un sistem de hardware fizic disparat.

Având în vedere acest obiectiv, puteți face unele presupuneri rezonabile cu privire la sistemul HPC. Dacă suntem interesați de calculul paralel folosind mai multe noduri, avem nevoie de cel puțin două sisteme separate (noduri), fiecare cu propriul sistem de operare (OS). Pentru ca lucrurile să funcționeze fără probleme, sistemul de operare pe ambele noduri ar trebui să fie identic. Dacă instalăm un pachet pe nodul 1, atunci acesta trebuie să fie instalat și pe nodul 2. Acest lucru micșorează o sursă de posibile probleme atunci când trebuie să depanați sistemul.

Cel de-al doilea lucru pe care trebuie să-l aibă clusterul este o rețea pentru a conecta nodurile, astfel încât să poată comunica pentru a partaja date, starea soluției la problemă și eventual chiar instrucțiunile care trebuie executate. Rețeaua poate fi teoretic orice ce permite comunicarea între noduri, dar cea mai ușoară soluție este Ethernet.

Depozitarea în fiecare nod poate fi la fel de simplă ca și o cartelă SD pentru menținerea sistemului de operare, a aplicațiilor și a datelor. În plus față de unele stocări de bază și pentru a face lucrurile un pic mai ușor, vom crea un sistem de fișiere partajat de la nodul master la celelalte noduri din cluster.

Cea mai fundamentală arhitectură HPC și software-ul este destul de simplă. Cele mai multe distribuții au instrumentele de bază pentru a face o activitate cluster și pentru a administra instrumentele; cu toate acestea, va trebui să se adauge instrumentele și bibliotecile pentru aplicațiile paralele (de ex., o bibliotecă sau o bibliotecă de compilatoare și orice biblioteci suplimentare necesare aplicației).

Arhitectură. Arhitectura unui cluster este destul de simplă. Sunt prezente câteva servere (noduri) care servesc diferite roluri într-un grup și care sunt conectate printr-un fel de rețea. De obicei, nodurile sunt asemănătoare, dar nu trebuie să fie; cu toate acestea, se recomandă foarte mult ca acestea să fie cât se poate de asemănătoare. Figura 2 este o ilustrare simplă a arhitecturii de bază.

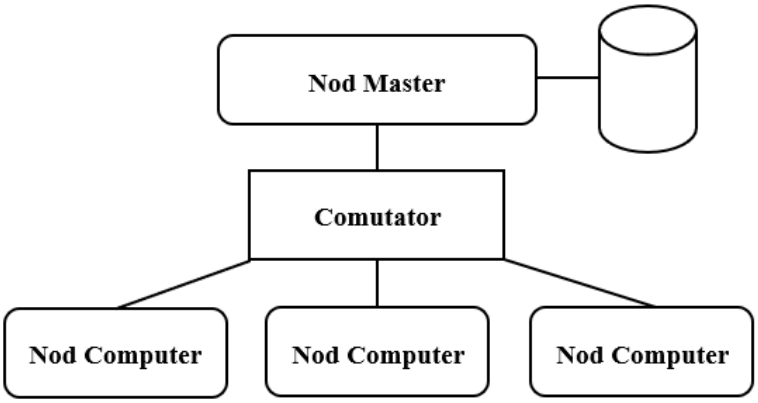


Fig.2. Structura generică a clusterului.

Aproape întotdeauna este prezent un nod care servește rolul unui „nod master" (uneori numit și un „nod cap"). Nodul principal este nodul „controller" sau nodul „management" pentru cluster. Controlează și efectuează gestionarea clusterului și de multe ori este nodul de conectare pentru ca utilizatorii să ruleze aplicații. Pentru grupurile mai mici, nodul principal poate fi folosit atât pentru calcul, cât și pentru management, dar pe măsură ce clusterul crește, nodul principal devine specializat și nu este utilizat pentru calcul.

Alte noduri din cluster completează rolul nodurilor de calcul, care descriu funcția lor. În mod normal, nodurile de calcul nu realizează funcții de gestionare a clusterelor; ele doar calculează. Nodurile computere sunt, de obicei, sisteme care rulează OS-ul minim, ceea ce înseamnă că emulatoarele care nu sunt necesare și pachetele inutile nu sunt instalate și au hardware-ul minim.

Odată cu creșterea clusterului, apar în mod obișnuit și alte roluri, necesitând adăugarea de noduri. De exemplu, serverele de date pot fi adăugate în cluster. Aceste noduri nu rulează aplicații; mai degrabă, stochează și difuzează date către restul clusterului. Nodurile adiționale pot oferi capabilități de vizualizare a datelor în cadrul clusterului (de obicei, vizualizarea la distanță) sau clustere foarte mari ar putea avea nevoie de noduri dedicate monitorizării clusterului sau conectării utilizatorilor la cluster și la rularea aplicațiilor.

Pentru un cluster simplu cu două noduri, pe care îl putem folosi ca introducere la HPC, se desemnează, de obicei, un nod principal și un nod de calcul. Cu toate acestea, pentru că sunt prezente doar două noduri, aplicațiile ar putea funcționa cel mai probabil pe ambele, pentru că se pierd 50% din noduri.

Rețeaua de conectare a nodurilor de cluster ar putea fi orice tehnologie de rețea, dar locul de pornire este cu Ethernet cu fir, care variază de la 100 Mbps la 56 Gbps; cu toate acestea, cele mai frecvent utilizate sunt Fast Ethernet (100 Mbps) sau Ethernet Gigabit (1000 Mbps).

Topologia de rețea utilizată pentru clustere este importantă, deoarece poate avea un efect asupra performanței aplicațiilor. Un layout simplu de rețea are un singur switch cu toate nodurile conectate la acel switch. Această configurație are un singur nivel și este simplă și eficientă, în special atunci când se construiesc sisteme mai mici. Pe masură ce sistemele devin mai mari, putem opta în continuare pe topologia primară, dar probabil va trebui să avem mai multe niveluri de switch-uri.

Pentru sistemele mai mici, comutatoarele Ethernet sunt destul de ieftine, costând doar câțiva dolari pe port. Comutatoarele vor fi mai bune decât un hub Ethernet, deși hub-urile vor limita performanțele, nu vor opri clusterul să funcționeze.

O altă caracteristică-cheie a unei arhitecturi de cluster de bază este un director partajat în cadrul nodurilor. În mod strict acest lucru nu este necesar, dar fără el unele aplicații MPI nu ar fi rulate. Prin urmare, este o idee bună să folosim un sistem de fișiere partajat în cluster. NFS este cel mai ușor de folosit, deoarece serverul și clientul se află în kernel, iar distribuția trebuie să aibă instrumentele necesare pentru configurarea și monitorizarea NFS.

Abordarea clasică NFS la un director partajat este de a exporta un director de la nodul principal la nodurile de calcul. Puteți alege orice director pe care doriți să-l exportați, dar de multe ori utilizatorii doar partajează / acasă de la nodul master, deși uneori vor exporta și un director nou, cum ar fi /shared. Nodurile de calcul montează și directorul partajat ca /home. Deci, dacă ceva în /home este local pentru fiecare nod, acesta nu va fi accesibil.

3. Sisteme de Cloud pentru calcul performant orientat spre clase aplicaţii

Virtualizarea resurselor reprezintă nucleul oricărei arhitecturi cloud computing, permiţând utilizarea unei interfeţe abstracte logic pentru accesarea resurselor fizice (servere, reţele, medii de stocare). Printre metodele de simulare a interfeţei către obiectele fizice sunt:

a) multiplexarea – crearea mai multor obiecte virtuale dintr-o singură instanţă a unui obiect fizic, de exemplu un procesor este multiplexat pentru a prelucra mai multe procese înlănţuite *(thread-uri);*

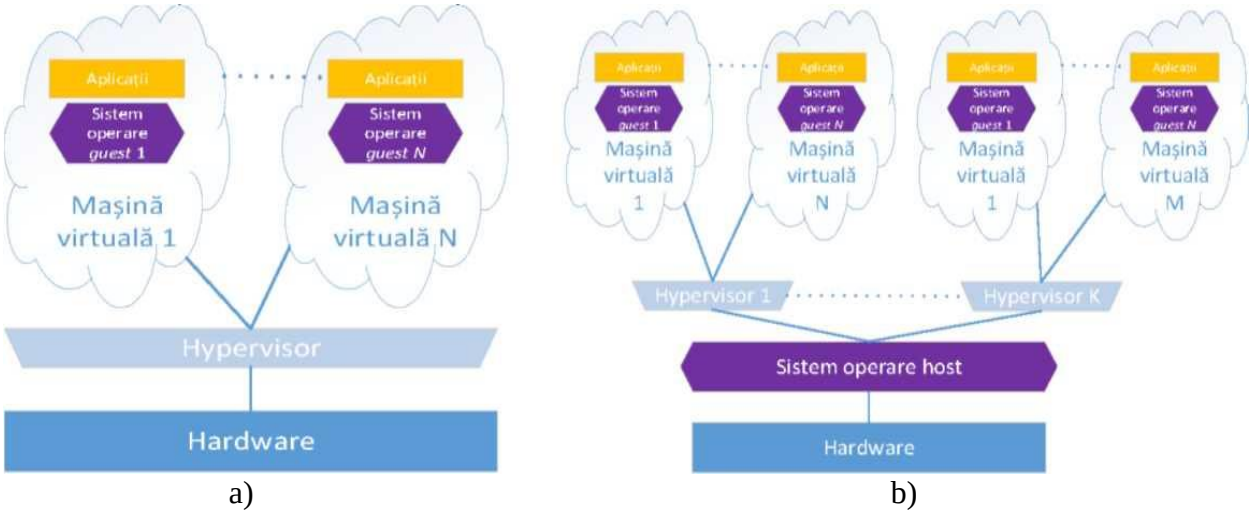
b) emularea – construirea unui obiect virtual dintr-un obiect fizic de alt tip, de exemplu un hard disk fizic poate emula memorie *RAM* (prin intermediul unui fişier sau partiţii interschimbabile – de tip *swap);*

c) agregarea – crearea unui singur obiect virtual din mai multe obiecte fizice, de exemplu un număr de hard disk-uri pot forma un disk agregat *RAID;*

d) multiplexare combinată cu emulare – de exemplu, protocolul TCP emulează un canal de comunicaţii sigur şi multiplexează transferul de date între canalul de comunicaţii fizic şi procesor.

Conceptul de virtualizare se aplică pentru a grupa resurse fizice sub formă omogenă şi de a le gestiona ca un tot unitar, oferind aceste resurse în funcţie de cereri. La baza acestei abordări a fost definit conceptul de *maşină virtuală,* înlocuind maşina reală fizică. Este definit conceptul de *hypervisor* (denumit şi *VMM – Virtual Machine Monitor),* care gestionează resursele de calcul pentru maşinile virtuale, simplificând procedurile de alocare şi monitorizare a resurselor fizice, de exemplu starea unei maşini virtuale poate fi salvată şi mutată pe un alt server. Astfel, un *hypervisor* este o entitate hardware, software sau *middleware* care creează şi rulează maşini virtuale. Un calculator fizic pe care rulează un *hypervisor* este denumit calculator gazdă (denumit în engleză *host),* iar fiecare calculator virtualizat care rulează pe calculatorul gazdă este denumit maşină virtuală vizitator (denumit în engleză *guest).* Pentru fiecare maşină virtuală este nevoie să fie oferit suport pentru virtualizarea componentelor hardware (procesor, memorie, stocare, dispozitive I/O, comunicaţii) şi partajarea acestora între diferitele aplicaţii.

Clasificarea tehnologiilor de tip *hypervisor* se face în următoarele categorii, în funcţie de nivelul la care rulează pe hardware, după cum este prezentat şi în Figura 3: (a) tip 1 – nativ (denumit şi *bare metal),* unde modulul de tip *hypervisor* rulează direct pe sistemul hardware al calculatorului gazdă şi gestionează resursele pentru maşinile virtuale, (b) tip 2 – găzduit (denumit şi *hosted),* unde modulul de tip *hypervisor* rulează pe un sistem de operare convenţional.



a) b)

Fig.3. Comparaţie între tehnologiile de tip *hypervisor:*
(a) nivel 1 – nativ şi (b) nivel 2 – găzduit *(hosted)* [13].

Pentru un modul de tip *hypervisor* de nivel 1 exemple tipice sunt *Citrix XenServer* [14] sau *VMWare ESX* [15], iar pentru un *hypervisor* de nivel 2 cele mai des utilizate sunt *VMWare Workstation* şi *VirtualBox* [16]. Există de asemenea şi sisteme de tip *hypervisor,* pentru care este mai dificil de aplicat această clasificare, de exemplu pentru *KVM (Kernel-based Virtual Machine)* [17], care permite programului kernel Linux să funcţioneze ca un modul de tip *hypervisor* de nivel 1. Totuşi, Linux este un sistem de operare în sine şi, conform

clasificării anterioare, orice modul de tip *hypervisor,* ce rulează peste acest sistem, este considerat de nivelul 2. De asemenea, *Microsoft Hyper-V* [18] a fost considerat un sistem de tip *hypervisor* de nivelul 2, deşi versiunea actuală din *Windows Server 2008* şi, mai recent, *Windows Server 2012* sunt încărcate înaintea rulării sistemului de operare. Se impune astfel definirea unei a treia categorii de sisteme de virtualizare *hibride,* în care modulul de tip *hypervisor* partajează resursele cu un sistem de operare, ca în Figura 4.

Sunt prezentate modalităţile de virtualizare şi analizate modurile în care sistemele de tip *hypervisor* controlează utilizarea resurselor hardware de către sistemele de tip invitat, pentru a nu a afecta alte maşini virtuale:

– virtualizarea sistemelor de operare,
– virtualizarea platformelor,
– virtualizarea stocării,
– virtualizarea reţelei,
– virtualizarea aplicaţiilor.

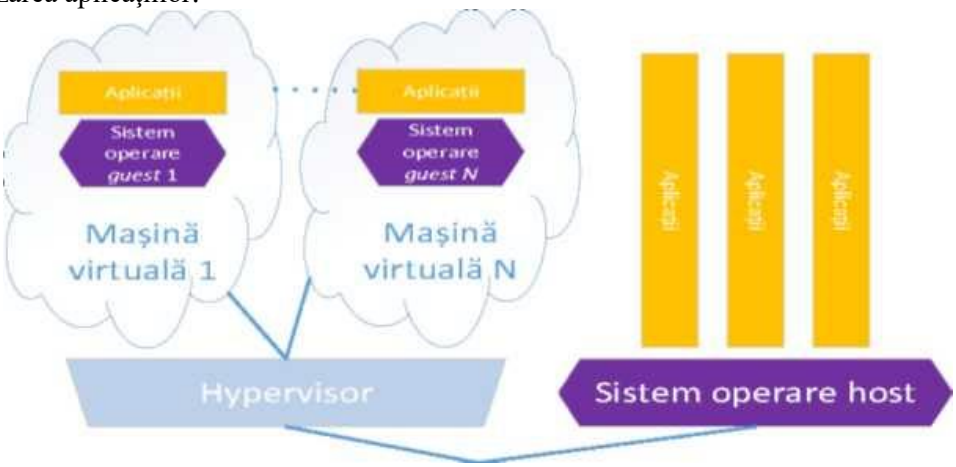


Fig.4. Exemplu de modul de tip *hypervisor* hibrid [13].

De asemenea, sunt analizate funcţiile unui modul de tip *hypervisor* de a partaja platforma fizică pentru mai multe aplicaţii sau servicii, şi posibilitatea migrării unei maşini de la o platformă la cealaltă.

4. Securitate şi riscuri în cloud

Lumea ştiinţifică şi tehnică este deosebit de preocupată de aspectul securităţii şi confidenţialităţii datelor în cloud. Faptul că datele companiei sau ale persoanelor sunt rezidente pe calculatoarele altor companii naşte suspiciuni privind accesul la acestea. Nu în ultimul rând, comunităţile de infractori cibernetici sunt foarte interesate de a studia şi exploata riscurile directe şi indirecte ale implementărilor de cloud, în speranţa colectării unui volum de date şi informaţii care le pot aduce beneficii sau venituri directe şi indirecte [19]. În percepţia generală a utilizatorilor de cloud, cele mai importante concepte legate de securitatea în cloud sunt [20]: criptarea datelor, autentificarea şi identificarea, protecţia antivirus, modul de configurare a firewall-urilor şi disponibilitatea serviciilor.

Securitatea sistemelor informaţionale este un domeniu vast şi poate fi evaluată pe baza principiului celei mai slabe componente [21]. Pentru a securiza fiecare componentă a sistemului informaţional, compania trebuie să parcurgă o serie de etape specifice, care includ, printre altele, inventarierea activelor, stabilirea unui nivel de risc pe fiecare activ pe baza vulnerabilităţilor specifice şi realizarea unui plan de tratare a riscurilor cu scopul diminuării lor. Unii autori afirmă că, pe lângă un management comprehensiv al riscurilor, companiile trebuie să adopte şi să adapteze permanent cele mai noi standarde şi coduri de bună practică în domeniul securităţii informaţionale. Rhoton [22] oferă exemple concrete de calcul al riscului în adopţia şi operarea tehnologiilor cloud, incluzând aspecte generale cu privire la pierderea de informaţii, indisponibilitatea serviciilor, costul neutilizării anumitor resurse, dar şi imposibilitatea adaptării la cerinţele dinamice ale pieţelor. Autorul propune un sistem de evaluare a riscurilor calculat pe baza impactului, probabilităţii şi numărului de produse sau servicii influenţate în cazul producerii riscului anticipat. Înţelegerea corectă a modelelor de implementare şi de servicii oferite de cloud şi a interdependenţelor dintre acestea constituie primul pas în determinarea nivelului de risc acceptabil pentru adopţia tehnologiilor cloud. *Încrederea* în cloud poate fi echivalentă cu încrederea care se

investeşte într-o aplicaţie web clasică [4], prin intermediul căreia se pot manipula date şi se pot lansa diferite comenzi şi operaţiuni. În cloud dispare teoretic un nivel din arhitectura de securitate a unui sistem informaţional: *nivelul fizic*. Datele companiei nu mai sunt pe un server aflat în spatele mai multor uşi închise. Zittrain [23] precizează că o persoană poate avea acces la informaţiile din cloud doar prin intermediul unei parole.

În seria sa de mituri despre cloud, Smith [24] specifică faptul că în cazul în care *neîncrederea* în disponibilitatea serviciilor cloud şi în confidenţialitatea datelor, în contextul complexităţii şi lipsei de transparenţă, se poate alege calea implementărilor de cloud hibrid. Latura ironică a acestei idei are la bază dualitatea dintre *complexitate* şi *control*. Arhitecturile de reţea foarte complexe sunt mai greu de controlat atât de cei care le implementează, cât şi de cei care încearcă să le înţeleagă pentru a le exploata vulnerabilităţile. Lipsa *transparenţei* poate fi considerată un alt factor de risc în implementarea soluţiilor cloud, pentru că nu asigură un nivel corespunzător de înţelegere a interdependenţelor dintre componentele unei implementări complexe [11].

Neîncrederea este uneori directă cu nevoia de *intimitate* a datelor pe care o au anumite companii în legătură cu datele pe care le stochează şi prelucrează în cloud. Chiar dacă în cele mai multe materiale intimitatea este tratată în categoria *confidenţialităţii,* termenul „intimitate" este mult mai potrivit în contextul manipulării datelor cu caracter personal. Scepticii cloud-ului sunt de părere că marii furnizori de cloud sau alte servicii din web-ul social folosesc în scopuri de afaceri datele personale ale angajaţilor şi ale oamenilor în general. În fapt, această teamă este suficient de simplă de susţinut cu probe concrete. O simplă căutare pe oricare motor de căutare vă va oferi publicitate contextuală în oricare alt site. Motivul este unul profesionist de canalizare a mesajelor publicitare pe nevoile reale ale unui individ. Dar, cum ar putea face furnizorii acest lucru fără stocarea unei cantităţi de date, fie şi infimă, care poate determina un profil temporar sau pe termen îndelungat al unui utilizator al Internetului?

Pentru a determina vulnerabilităţile unui sistem cloud, trebuie să ţinem seama de o serie de factori diferenţiatori faţă de implementările locale:

- ✓ Intern *versus* Extern – cu referire la locaţia de stocare a datelor;
- ✓ Deschis *versus* Proprietar – cu referire la forma de proprietate a echipamentelor de prelucrare şi stocare şi modului de acces la resursele respective;
- ✓ Externalizat *versus* internalizat – cu referire la modul în care se asigură suportul pentru serviciile puse la dispoziţie;
- ✓ În perimetrul de securitate *versus* în afara perimetrului – cu referire la modelul clasic de securizare a infrastructurii de reţea locale prin intermediul dispozitivelor de tip firewall.

Standardele americane de securitate în domeniul cloud-ului definesc mai multe clase de controale specifice pentru determinarea nivelului de vulnerabilitate a unui serviciu în general:

- Managementul configuraţiilor;
- Achiziţia de sisteme şi servicii;
- Protecţia sistemelor şi comunicaţiilor;
- Integrarea sistemelor şi informaţiilor.

Fiecare clasă de controale are propriile sale categorii de riscuri şi vulnerabilităţi specifice aplicabile la nivel general sistemelor informaţionale, inclusiv arhitecturilor cloud.

Vulnerabilităţi ale clienţilor de conectare în cloud. Având în vedere că metodele de conectare la serviciile cloud includ browser-e şi aplicaţii specifice, clienţii trebuie să se asigure că acestea sunt protejate şi securizate. Una dintre cele mai comune probleme legate de browser-e este aceea a instalării *add-on*-urilor de tip *adware,* care pot colecta datele de conectare ale utilizatorilor, în felul acesta expunând întreaga companie la riscuri de acces neautorizat la conţinutul informaţional stocat în cloud. Salvarea parolelor în browser-e sau aplicaţii poate constitui un alt risc în cazul pierderii sau furtului dispozitivului de acces. Pentru diminuarea riscurilor companiile folosesc metode specifice de securizare a browser-elor şi aplicaţiilor. Dintre acestea cele mai întâlnite sunt: localizarea geografică a unui dispozitiv, ştergerea de la distanţă a conţinutului stocat, blocarea dispozitivului sau implementarea mecanismelor de dublă autentificare: nume de utilizator + parolă şi un cod temporar generat de o aplicaţie specifică.

Criptarea datelor şi comunicaţiilor. Fiecare serviciu de cloud transferă datele prin intermediul canalelor securizate: HTTPS, SSH, RDP, FTPS, dar revine în sarcina clientului să se asigure în momentul conectării că accesează site-ul corect pentru autentificare. *Phishing*-ul a ajuns o tehnică destul de elaborată în redirecţionarea

utilizatorilor către site-uri care emulează cât mai fidel site-urile originale, reuşind astfel să intre în posesia datelor de autentificare ale utilizatorilor. Majoritatea browser-elor moderne au mecanisme de protecţie a phishing-ului prin validarea certificatelor de criptare a comunicaţiei dintre client şi serviciul cloud. Un alt risc legat de certificatele digitale şi criptarea comunicaţiilor şi datelor este reprezentat de *scurgerea de informaţii* din cadrul companiei şi *repudierea*[11] conţinutului informaţional. Mecanismele de tipul managementul drepturilor digitale (DRM[12]) şi mecanismele privind prevenirea scurgerii informaţiilor (DLP[13]) sunt din ce în ce mai utilizate atât în interiorul companiilor, cât şi în cloud, furnizorii punând la dispoziţie servicii DRM integrate cu mecanismele de autentificare şi identificare a utilizatorilor. Dintre cele mai frecvente măsuri implementate în DRM amintim: semnarea conţinutului informaţional, criptarea conţinutului informaţional, protecţia asupra anumitor operaţiuni comune, precum copierea, redistribuirea, şi mecanisme complete de expirare a conţinutului informaţional după o anumită perioadă de timp. DRM asigură în acelaşi timp mecanisme concrete de protecţie de tip FYEO,[14] asigurându-se astfel că numai destinatarul unui mesaj este capabil să îl deschidă şi citească.

Protecţia perimetrului fizic este o cerinţă concretă de securitate pentru locaţia clientului care foloseşte serviciile de cloud. În acelaşi timp, este un criteriu de selecţie a furnizorului de cloud, care trebuie să ofere posibilitatea de acces la aceleaşi date dintr-o locaţie secundară sau mai multe. Mecanismele de diminuare a riscurilor împotriva atacurilor fizice pun în prim-plan strategiile de realizare şi gestionare a cópiilor de siguranţă a datelor.

Riscuri legate de performanţă şi funcţionalităţi. Clienţii de cloud trebuie să se asigure în momentul în care aleg un serviciu cloud că au testat suficient de bine performanţele serviciului şi că funcţionalităţile sunt corespunzătoare specificului proceselor de afaceri ale companiei. Majoritatea furnizorilor de cloud pun la dispoziţia clienţilor versiuni de evaluare a unor funcţionalităţi de cloud, pentru efectuarea implementărilor pilot şi definirea planurilor şi strategiilor de implementare, migrare şi crearea documentaţiilor suport. De asemenea, este foarte importantă cunoaşterea tuturor funcţionalităţilor oferite de anumite servicii. De multe ori, firmele aleg anumite planuri de servicii care oferă funcţionalităţi limitate, dar la un preţ mai redus, făcând dificilă ulterior dezvoltarea unor funcţionalităţi personalizate. Alteori, clienţii au cunoştinţe detaliate despre anumite aplicaţii şi servicii implementate local, aleg o soluţie similară din cloud şi după o perioadă de timp constată că le lipsesc anumite componente-cheie de dezvoltare, personalizare şi administrare a serviciilor. Serviciile şi aplicaţiile cloud nu întotdeauna au implementate toate funcţionalităţile, sau anumite funcţionalităţi nu pot fi personalizate sau administrate asemănător cu cele din implementările locale, dând naştere nemulţumirilor unor categorii de utilizatori.

*Expunerea la atacurile de inginerie socială***.** Având în vedere că mecanismele de protecţie a cloud-urilor îngreunează sau reduc aproape la minimum atacurile clasice de tip *brute-force, password guessing, denial-of-services*, infractorii cibernetici caută alte metode prin care să aibă acces la conţinutul informaţional al companiilor. Ingineria socială este un subiect sensibil pentru majoritatea companiilor pentru că, cel puţin în ultimii ani, este cea mai eficientă măsură de acces direct la datele companiei. Bazându-ne pe studii şi experimente concrete [20], am constatat că la nivelul multor companii din România nu există mecanisme de protecţie şi/sau conştientizare a angajaţilor privind aceste tipuri de riscuri.

„IT-ul tradiţional" se bazează pe securitatea unui firewall, dar esenţa cloud-ului presupune transferul şi accesul la date de oriunde, folosind într-adevăr canale securizate, dar mobilitatea expune angajaţii la interacţiuni cu diferite persoane care le pot convinge prin diferite metode să divulge informaţii utile pentru construirea unui atac de tipul ingineriei sociale.

Concluzii

Caracteristicile fundamentale determină o serie de avantaje de care pot beneficia utilizatorii tehnologiilor cloud. Abilitatea tehnică a furnizorilor de cloud de exploatare la un nivel ridicat a echipamentelor de calcul, precum şi mediul concurenţial destul de agresiv, determină un nivel redus al preţurilor pentru care sunt

[11] ISO/IEC 13888-1:2009(en) – Information technology – Security techniques – Non-repudiation. Part 1: General. https://www.iso.org/obp/ui/#iso:std:iso-iec:13888:-1:ed-3:v1:en
[12] DRM – Digital Rights Management.
[13] DLP – Data Loss Prevention.
[12] FYEO – For Your Eyes Only – Exclusiv doar pentru ochii tăi.

comercializate serviciile cloud. Astfel, clienţii beneficiază de *preţuri mici* comparativ cu costul total de apartenenţă (TCO[15]) al puterii de calcul similare pe care ar trebui să o implementeze local (on-premise).

Datorită pachetelor predefinite de maşini virtuale şi software preinstalate, utilizatorii vor beneficia de un *acces mai uşor* la serviciile informaţionale de care au nevoie pentru desfăşurarea activităţilor. De exemplu, pentru a instala un server de baze de date on-premise este nevoie de un server dedicat pe care trebuie instalat un sistem de operare, configurate servicii specifice de stocare, instalare de aplicaţii prerechizite, instalarea aplicaţiilor pentru bazele de date şi configurarea şi securizarea corectă a acestora. Toate aceste operaţiuni necesită licenţe, timp şi oameni specializaţi. În cloud se poate alege un pachet predefinit de server care să aibă preinstalat sistemul de baze de date dorit. În felul acesta, beneficiarul nu va plăti licenţe sau timpul necesar instalării şi configurării, ci doar timpul de utilizare a respectivei maşini virtuale, fiindu-i în acest fel mult mai uşor să aibă acces la serviciul de baze de date dorit. *Fiabilitatea* şi *calitatea serviciilor oferite* sunt alte beneficii ale tehnologiilor cloud. Infrastructurile de reţea şi maşinile virtuale pot fi configurate să asigure un nivel de balansare în deservirea cererilor în mod dinamic, asigurând astfel *disponibilitatea* ridicată a serviciilor la un cost redus şi un nivel tehnic mult mai performant decât instrumentele care ar putea fi configurate on-premise.

O reducere considerabilă a costurilor poate fi desprinsă şi din *externalizarea* serviciilor de administrare şi întreţinere a infrastructurilor hardware şi de reţea. Chiar dacă este asemănător cu *outsourcing*-ul, beneficiarul de cloud nu trebuie să încheie contracte de întreţinere şi de suport separate cu alţi furnizori specializaţi. Această metodă de administrare simplifică şi modul în care se realizează operaţiunile de întreţinere şi update, responsabilitatea pentru acestea revenind exclusiv furnizorului de cloud. Uneori *upgrade*-ul sistemelor de operare se face la costuri de licenţiere ridicate şi implică o serie de teste prealabile de funcţionalitate viitoare a aplicaţiilor implementare. În cazuri excepţionale, operaţiunile de întreţinere şi upgrade din on-premise induc o întrerupere pe o perioadă destul de mare a livrării serviciilor către beneficiari.

Referințe:

1. ANDERSON, T.E., CULLER, D.E., PATTERSON, D.A. A case for NOW (networks of workstations). In: *Micro, IEEE*, 1995, p.54-64.
2. FOSTER, I., KESSELMAN, C. *The Grid 2: Blueprint for a New Computing Infrastructure.* San Francisco, USA: Elsevier, 2003.
3. KONDO, D., JAVADI, B., MALECOT, P., CAPPELLO, F., ANDERSON, D. P. Cost-benefit analysis of cloud computing versus desktop grids. In: *Parallel & Distributed Processing*, 2009. IPDPS, 1-12.
4. YELURI, R., CASTRO-LEON, E. *Building the Infrastructure for Cloud Security. A solutions View.* New York, USA: Apress Open, 2014.
5. MELL, P., GRANCE, T. *The NIST Definition of Cloud Computing.* 2011 / Preluat de pe National Institute of Standards and Technology: http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf
6. GROSSMAN, R.L. The case for cloud computing. In: *IT professional*, 2009, no11(2), p.23-27.
7. SOSINSKY, B. *Cloud Computing Bible.* Indianapolis, USA: Wiley Publishing, Inc., 2011.
8. SULLIVAN, D. *The Definitive Guide To Cloud Computing*, 2010. http://www.realtimepublishers.com/: Realtime Publishers.
9. OPREA, D. *Protecția și securitatea informațiilor.* Iași: Polirom, 2007.
10. WILDER, B. *Cloud Architecture Patterns.* Gravenstein Highway North, Sebastopol, CA: O'Reilly Media, Inc., 2012.
11. MATHER, T., KUMARASWAMY, S., LATIF, S. *Cloud Security and Privacy.* Gravenstein Highway North, Sebastopol: O'Reilly Media, Inc., 2009.
12. GOETSCH, K. *eCommerce in the Cloud.* Gravenstein Highway North, Sebastopol: O'Reilly Media, Inc., 2014.
13. SUCIU, G. *Contribuții la teoria și implementarea aplicațiilor de comunicații pe platformă Cloud Computing.* București, 2013.
14. SABHARWAL, N., WALI, P. *Cloud Capacity Management.* New York, USA: Apress, 2013.
15. Citrix Systems, Inc, „Citrix XenServer - Efficient Server Virtualization Software," [Interactiv]. Available: http://www.citrix.com/xenserver.
16. VMware, Inc, „vSphere ESX and ESXi Info Center", [Interactiv]. Available: http://www.vmware.com/products/esxi-and-esx/overview.
17. Oracle Corp., „Oracle VM VirtualBox", [Interactiv]. Available: https://www.virtualbox.org/.

[15] TCO – Total Cost of Ownership – Costul total de apartenenţă, care include costurile de achiziţie şi întreţinere a unui echipament de calcul.

18. KVM, „Kernel Based Virtual Machine", [Interactiv]. Available: http://www.linux-kvm.org.
19. Microsoft Corp., „Microsoft Hyper-V Server 2012 – Virtualization", [Interactiv]. Available: http://www.microsoft.com/hyper-v-server.
20. CHARIF, B., & AWAD, A.-I. Business and Government Organizations' Adoption of Cloud Computing. In: *Intelligent Data Engineering and Automated Learning–IDEAL*, 2014, p.492-501.
21. GREAVU-ȘERBAN, V. Prezentare ISO27001: sistemul de management al securităţii informaţionale. În: *Progrese în teoria deciziilor economice în condiții de risc și incertitudine,* 2010, vol.XI, p.96-104. Iasi: Tehnopress.
22. RHOTON, J. *Cloud Computing Explained: Implementation Handbook for Enterprises.* Milton Keynes, UK: Recursive Press, 2009.
23. ZITTRAIN, J. *The Future of the Internet – And How to Stop It.* New Haven, UK: Yale University Press, 2008.
24. SMITH, D.M. *The Top 10 Cloud Myths.,* 2014. Preluat de pe Gartner: http://www.gartner.com/doc/2860422

Date despre autor:
Ionel ANTOHI, doctorand, Școala doctorală *Matematică și Știința Informației*, Universitatea de Stat din Moldova.
E-mail: antohi.ionel@gmail.com
ORCID: 0000-0002-3776-451X

Prezentat la 04.04.2019