

UNAPREĐENJE ENERGETSKE EFIKASNOSTI U CLOUD OKRUŽENJU PRIMENOM KONSOLIDACIJE VIRTUELNIH MAŠINA

Branka Mikavica¹, Aleksandra Kostić-Ljubisavljević², Nevena Simović³

¹Univerzitet u Beogradu - Saobraćajni fakultet, b.mikavica@sf.bg.ac.rs

²Univerzitet u Beogradu - Saobraćajni fakultet, a.kostic@sf.bg.ac.rsž

³A1 Srbija, nevena.simovic@hotmail.com

Rezime: *Pojava Cloud Computing-a uvela je značajne promene u Internet okruženju jer predstavlja paradigmu za obezbeđivanje pristupa skalabilnom i elastičnom skupu deljivih, fizičkih ili virtuelnih resursa sa mogućnošću obezbeđivanja servisa i njihovo administriranje na zahtev. Razvoj servisa i aplikacija u pogledu kvaliteta servisa, troškova, bezbednosti i potrošnje energije zahteva primenu tehnologija koje treba da zadovolje sve strože zahteve. Virtuelizacija je ključna za Cloud Computing, jer je to tehnologija koja stvara inteligentni sloj apstrakcije koji omogućava različitim virtuelnim mašinama da dele isti hardver. U radu je detaljno definisana virtuelizacija i alokacija resursa u cloud okruženju. Takođe, prikazan je značaj smanjenja potrošnje energije. Na kraju, cilj rada je analiza mogućnosti postizanja energetske efikasnosti primenom konsolidacije virtuelnih mašina u cloud okruženju.*

Ključne reči: *virtuelizacija, data centar, migracija, virtuelna mašina, energetska efikasnost*

1. Uvod

U poslednjih nekoliko decenija, društvo je doživelo drastične promene u načinu pristupa informacijama, njihovom skladištenju, prenošenju i obradi. Internet je učinio informacije dostupnim sa bilo kog uređaja koji je povezan na globalnu mrežu. Nove tehnologije su u potpunosti izmenile komunikacione navike korisnika i postale su neophodne u svakodnevnom životu čoveka, a njihov uticaj će u budućnosti postati još veći. Napredak tehnologije i inkorporacija mreža, skladištenja i procesorske snage doveli su do nove ere računarstva, koja se naziva *Cloud Computing* ili opšte poznato kao *Cloud*. Nacionalni institut za standarde i tehnologiju (NIST- *National Institute of Standards and Technology*) definiše *Cloud Computing* kao model za obezbeđivanje odgovarajućeg, potpunog pristupa skupu deljenih, konfigurabilnih resursa (mreža, servera, skladišnih kapaciteta, aplikacija i servisa) na zahtev, pri čemu se obezbeđivanje tih resursa vrši uz minimalno angažovanje i minimalnu interakciju sa provajderom [1]. *Cloud Computing* se

zasniva na virtualizaciji mreža, servera, prostora za skladištenje i servisa i predstavlja skup velikog broja povezanih računara sa pripadajućim procesorima, diskovima, memorijom i drugim hardverskim resursima koji se, zahvaljujući inteligentnom upravljačkom softveru, vide kao jedan virtuelni računar. Sa brzim rastom potražnje za pristupom ovim servisima, potrebe za skladištenjem podataka i protoka velike količine podataka, kompanije proširuju svoje *data centre*.

Data centri se koriste za *host*-ovanje aplikacija koje obično troše ogromne količine resursa, a samim tim i električnu energiju, što dovodi do rasta operativnih troškova i emisije ugljen dioksida. Da bi uslovi Ugovora o nivou servisa, SLA (*Service Level Agreement*) bili zadovoljeni, *data centri* moraju neprekidno da se snabdevaju strujom, što koristi ogromnu količinu energije i posledično povećava troškove ulaganja. Ranije su visoke performanse bile primarni cilj u *data* centrima bez obraćanja pažnje na količinu potrošene energije. Međutim, poslednjih godina ključni izazov je uspostavljanje balansa između performansi sistema i potrošnje energije. Utvrđeno je da se ogromna količina energije troši zbog neaktivnih i preopterećenih servera u *data* centru zbog čega je uloženo mnogo napora da se poboljša energetska efikasnost virtuelizovanih *data* centara. Virtuelizacija je jedna od tehnologija koje omogućavaju smanjenje potrošnje energije *data* centara odvajanjem virtuelnih mašina od osnovnih fizičkih mašina.

Ovaj rad je koncipiran na sledeći način. Nakon uvoda, u drugom delu rada opisana je tehnologija virtuelizacije, koncept virtuelnih mašina i način na koji se vrši izbor i konsolidacija virtuelne mašine. U četvrtom delu rada prikazan je uticaj potrošnje energije i značaj energetske efikasnosti. Konsolidacija virtuelnih mašina za postizanje energetske efikasnosti razmatrana je u petom delu. Zaključna razmatranja *data* su na kraju rada.

2. Virtuelizacija *Cloud* resursa

Virtuelizacija je ključna za *Cloud Computing*, jer je to tehnologija koja stvara inteligentni sloj apstrakcije koji sakriva složenost hardvera ili softvera. Omogućuje različitim operativnim sistemima (OS) ili virtuelnim mašinama (VM) da dele isti hardver i olakšava premeštanje OS između različitih fizičkih čvorova, sve dok aplikacije pokreću i koriste tehnike *live* migracije [2]. Virtuelizacija zahteva softver koji je u stanju da pruži izgled hardverske platforme gostujućim OS. To znači da je u stanju da vidi sve interakcije, uključujući i pristup memoriji. OS i korisnici tog OS vide apstraktni uređaj sa skupom hardvera koji nema nikakve veze sa osnovnim hardverom glavnog sistema. [3].

Softverski sloj koji omogućava virtuelizaciju naziva se hipervizor, poznat i kao monitor virtuelne mašine VMM (*Virtual Machine Monitor*). Odgovoran je za stvaranje virtuelnog okruženja na kojem VM rade, kao i za dodeljivanje resursa računarskog sistema jednoj ili više VM. VMM nadzire gostujuće sisteme i osigurava da se resursi dodeljuju po potrebi [2]. Takođe, omogućava da nekoliko OS radi istovremeno na jednoj hardverskoj platformi i nameće izolaciju između ovih sistema, čime se povećava bezbednost. VMM kontroliše kako gostujući OS koristi hardverske resurse i istovremeno omogućava:

1. Više usluga za deljenje iste platforme
2. Premeštanje servera sa jedne platforme na drugu, takozvanu *live* migraciju
3. Modifikaciju sistema uz zadržavanje kompatibilnosti sa originalnim sistemom.

Virtuelizacija apstrahuje osnovne resurse i pojednostavljuje njihovu upotrebu, izoluje korisnike jedne od drugih i podržava replikaciju, što zauzvrat povećava elastičnost sistema. Istovremeno, virtuelizacija je kritičan aspekt *Cloud Computing*-a, podjednako važan za *cloud* korisnike i provajdere, i igra važnu ulogu za:

1. Bezbednost sistema, jer dozvoljava izolaciju servisa koji rade na istom hardveru
2. Performanse i pouzdanost, jer dozvoljava aplikacijama da migriraju sa jedne platforme na drugu
3. Razvoj i upravljanje servisima koje nudi provajder
4. Izolaciju performansi.

Podržavanjem virtuelizacije omogućena je dinamička migracija VM sa jednog *host*-a na drugi *host* u skladu sa zahtevima performansi. *Host*-ovi u stanju mirovanja su podešeni na režim za uštedu energije za efikasno korišćenje energije od strane *data* centra. Alokacija *cloud* resursa se ne koncentriše na dodelu VM radi smanjenja potrošnje energije, iako je cilj alokacije resursa *data* centra da se obezbede visoke performanse uz ispunjavanje zahtevanih SLA [4]. Za upravljanje performansama i efikasnom potrošnjom energije moraju se rešiti tri suštinska pitanja. Prvo, treba otkriti preopterećene *host*-ove, koji ugrožavaju performanse VM i SLA. Drugo, neophodno je identifikovati VM na preopterećenim *host*-ovima i migrirati na efikasne *host*-ove. Treće, dinamička priroda *cloud data* centra je kritična iz perspektive kvaliteta servisa, QoS (*Quality of Service*). Nekoliko VM možda neće dobiti potrebne resurse pod vršnim opterećenjem i neće ispuniti uslove SLA imajući u vidu varijabilnost radnog opterećenja i agresivnu konsolidaciju. Zbog toga se neaktivni *host*-ovi moraju prebaciti u režim za uštedu energije. Sva ova pitanja zahtevaju efikasne politike i algoritme koji mogu da smanje potrošnju energije uz ispunjavanje zahteva SLA. Na osnovu pregleda literature, neki od tih algoritama biće objašnjeni u nastavku rada.

2.1. Virtuelna mašina i izbor virtuelne mašine

Virtuelna mašina je izolovano okruženje koje predstavlja ceo računar, ali zapravo ima samo pristup delu računarskih resursa. Razlikuju se dve vrste VM: procesna i sistemska VM. Procesna VM je virtuelna platforma koja je stvorena za individualni proces i koja se prekida nakon što se proces završi. Sistemska VM podržava OS zajedno sa mnogim korisničkim procesima [5]. Sistemska virtuelna mašina obezbeđuje kompletan sistem; svaka VM može da pokreće sopstveni OS, koji zauzvrat može da pokreće više aplikacija.

Fizička mašina, PM (*Physical Machine*), može biti domaćin skupa VM. Kada je PM preopterećena, jedna ili više VM moraju da migriraju da bi se smanjilo radno opterećenje PM i promenilo njeno stanje u normalno. Shodno tome, uslovi SLA će biti ispunjeni. Trenutno postoje tri glavna pristupa za izbor virtuelnih mašina [6]:

1. Politika minimalnog vremena migracije koja bira VM kojima će biti potrebno najmanje vremena da se premeste
2. Politika nasumične selekcije koja migrira virtuelne mašine nasumično bez ikakvih ograničenja
3. Politika maksimalne korelacije koja migrira VM koja ima maksimalni koeficijent korelacije u poređenju sa drugim VM na istoj PM.

2.2. Konsolidacija virtuelne mašine

Karakteristike problema konsolidacije VM, u većini slučajeva, se mogu kategorisati kao sledeće [7]:

1. VM se dodeljuju dostupnim fizičkim mašinama koje zadovoljavaju zahteve VM
2. Broj VM se menja tokom vremena u zavisnosti od zahteva korisnika
3. Zahtevi za resurse za VM mogu varirati tokom vremena
4. Potrošnja energije PM se može izračunati tokom rada
5. Migracija se koristi za prenos VM između fizičkih *host*-ova
6. Kršenja SLA (SLAV) opisuju degradacije performansi zbog preopterećenja fizičkih mašina kao i onih uzrokovanih migracijama VM
7. Konsolidacija VM se može izvršiti na mreži ili van mreže.

U zavisnosti od početnog stanja *data* centra postoje dve kategorije VM konsolidacije: *Offline* i *Online* konsolidacija VM [7]. *Offline* ili statička VM konsolidacija ima za cilj da smanji potrošnju energije optimizacijom postojećeg položaja. U tom slučaju, definišu se VM sa specifičnim zahtevima koje treba dodeliti skupu potpuno praznih PM. U statičkoj konsolidaciji, broj VM je unapred poznat. Dakle, ne uzimaju se u obzir novi zahtevi za raspoređivanje VM jer korisnici u *cloud*-u mogu zahtevati bilo koji broj VM prema potrebi. Takve tehnike su korisne u situacijama poput migracije skupa VM sa jednog servera na drugi: na primer, u slučaju reakcije na neplanirane situacije. Međutim, tehnike statičke konsolidacije nameću dodatno opterećenje u *data* centru. Premeštanje postojeće VM korišćenjem migracije zahteva vreme, stvara prekomerne troškove, troši dodatnu količinu propusnog opsega, što potencijalno utiče na odziv u Internet okruženju i može imati negativan uticaj na željeni nivo QoS-a.

Online ili dinamička konsolidacija, suprotno statičkoj, zahteva odluke o smeštanju koje se mogu izvesti u dve različite faze: inicijalno postavljanje i pokretanje *runtime*-a. Inicijalno postavljanje se vrši u trenutku primene, i to na sledeći način: VM instance se dodeljuju fizičkim *host*-ovima koji su zasnivaju na njihovim zahtevima i ugovoru SLA između korisnika i provajdera. Faza *runtime* je proces mapiranja VM koja je već pokrenuta na drugoj PM. *Live* migracija je neophodna za konsolidaciju VM tako što omogućava prenos pokrenute VM jednog *host*-a na drugi. Međutim, troškovi migracije se ne mogu zanemariti, a nivoi servisa pokrenutih aplikacija će verovatno biti negativno pogođeni tokom same migracije VM, što će dovesti do negativnog uticaja na SLA. Zbog toga je važno minimizirati broj *live* migracija.

Zbog ovih tehnika virtuelizacije potrošnja energije se povećava. Ako se uzme u obzir *cloud*, to bi imalo petu najveću potrošnju električne energije na svetu. Tačnije, potrošnja energije linearno zavisi od broja pokrenutih fizičkih mašina. Čak i ako su serveri u praznom hodu, i dalje troše više od polovine svoje vršne snage. Jedna od opcija za smanjenje potrošnje energije jeste da se smanji broj uključenih servera konsolidacijom VM u minimalni broj *host*-ova i prebaci na režim niske snage. Ova strategija, koja se naziva konsolidacija VM, međutim, može prouzrokovati neke probleme koji mogu uticati na obezbeđeni nivo QoS. Zapravo, konsolidacija VM se postiže migracijom VM. Kapacitet odredišnog *host*-a trebalo bi da bude dovoljan da zadovolji potrebe resursa svih

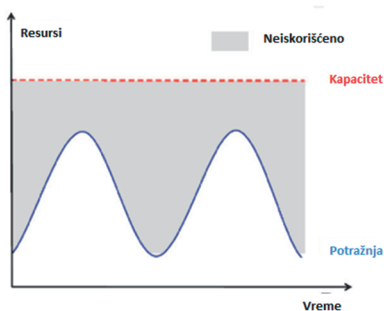
VM koji se pokreću na njemu. U suprotnom, *host* će biti preopterećen, što rezultuje degradacijom performansi svojih VM i kršenjem SLA.

3. Potrošnja energije i energetska efikasnost *Cloud* sistema

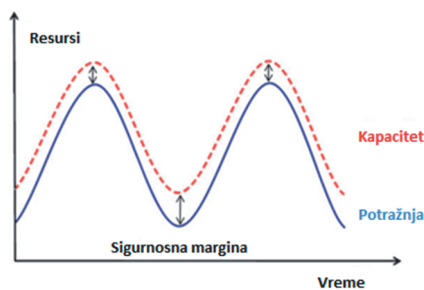
Sve veći obim podataka koji se svakodnevno obrađuje, skladišti i kome se pristupa putem savremene Internet infrastrukture dovodi do toga da potražnja za energijom u *data* centrima raste sve brže i brže. Iz tog razloga, energetske orijentisani *data* centri se istražuju kako bi se smanjio njihov ekološki i ekonomski uticaj. Da bi se *data* centri mogli smatrati *cloud* resursima, važno je da oni obezbeđuju skalabilnost i elastičnost, pristup bilo kada i bilo gde; obezbeđuju servise na zahtev uz mogućnost merenja i mogućnost korišćenja od strane više korisnika.

Potrošnja energije u *data* centrima, kao i njen ekonomski i ekološki uticaj mogu se razmatrati kroz analizu koncepta energetske proporcionalnih sistema. Ovo je veoma važan koncept jer je strategija za upravljanje resursima *cloud*-a koncentrisana na podskup servera i prebacivanje ostalih servera u režim pripravnosti kad god je to moguće. Ova strategija ima za cilj smanjenje potrošnje energije i troškova pružanja računarskih i skladišnih resursa.

U idealnom slučaju, energija koju troši sistem u stanju mirovanja treba da bude približno jednaka nuli i linearno treba da raste sa opterećenjem sistema. Po definiciji, idealan sistem proporcionalan energiji uvek radi na 100% efikasnosti [8]. U stvarnosti, koriste više od polovine snage u praznom hodu nego što koriste pri punom opterećenju. Teoretska dnevna varijacija radnog opterećenja za tipičan *data* centar prikazana je na Slici 1. Na grafiku se može primetiti da, iako opterećenje/potražnja prati pseudosinusoidalni trend, potrošena snaga ostaje skoro konstantna tokom visokog i niskog perioda opterećenja. Ovo je posledica uticaja računarskih resursa koji se uvek održavaju i rade i kada se nedovoljno koriste ili se uopšte ne koriste, pa se tako troši velika količina energije tokom perioda niskog opterećenja. Osnovna ideja jeste da *data* centri postanu „svesni” energije uvođenjem elastičnosti čime se prilagođava promena kapaciteta *data* centra tako da prati trenutnu potražnju/opterećenje, kao što je prikazano na Slici 2. To se može postići dinamičkim upravljanjem na nivou raspoređivanja resursa što podrazumeva da se vrši raspoređivanje zadataka na dostupne servere, a neiskorišćeni se stavljaju u režim mirovanja [9].



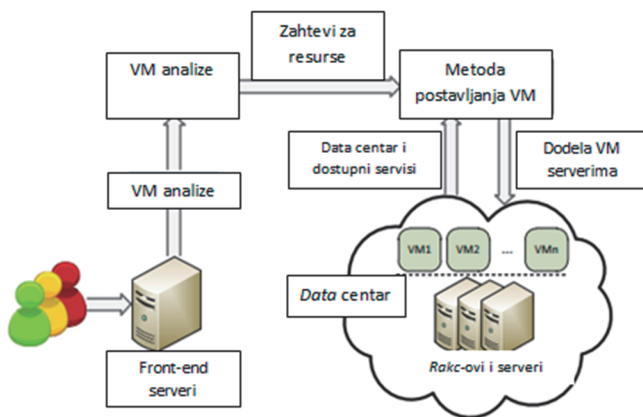
Slika 1. Opterećenje i energetska potrošnja u *data* centrima



Slika 2. Koncept elastičnog obezbeđenja resursa u data centrima

4. Konsolidacija VM za unapređenje energetske efikasnosti

Algoritmi koji se primenjuju za energetske efikasno postavljanje VM koriste prednosti strukture *data* centra i karakteristika VM, uključujući računarske sposobnosti svake VM u smislu MIPS-a (*Mega Instructions Per Second*) i potrebne memorije. Operativni tok okvira je prikazan na Slici 3. i sastoji se od dva osnovna modula: (a) analiza VM i (b) metoda postavljanja VM. Analiza VM određuje zahtevani MIPS i memoriju svake VM. Metoda postavljanja VM uzima dva parametra kao ulaz: (a) zahteve za resursima izračunate analizom VM i (b) trenutni status (uključeno/isključeno) opreme *data* centra i raspoloživih resursa svakog uključenog servera. Metoda postavljanja VM pokušava da pronađe minimalni broj servera i *rack*-ova potrebnih za ispunjavanje zahteva VM za resurse i pokušava da upravlja napajanjem na minimalnom broju *switch*-eva i uređaja za hlađenje. Ostatak resursa koje čine *switch*-evi za neaktivne mreže, rashladni sistemi i serveri, mogu se isključiti da bi se postigla veća energetska efikasnost [10].



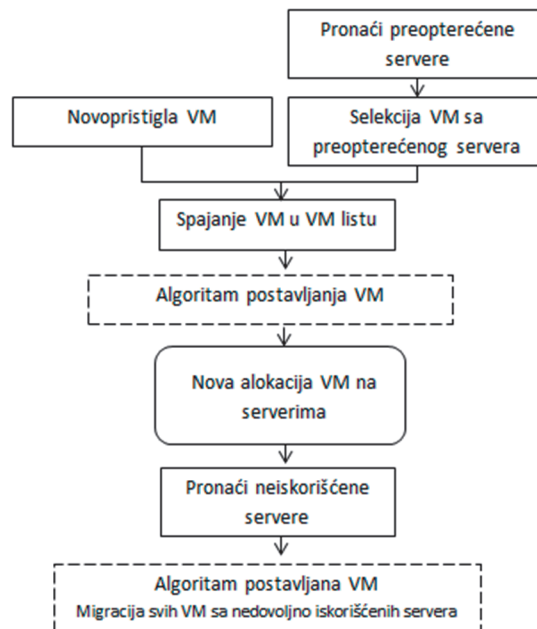
Slika 3. Pregled okvira za energetske efikasnost u virtuelizovanom data centru [9]

U nastavku je opisan problem konsolidacije VM na mreži, gde u unapred određenim vremenskim intervalima, odnosno, fazama, mogu nailaziti novi zahtevi za uspostavljanje VM, promene u korišćenju CPU (*Central Processing Unit*) postojećih VM

i povlačenje nekih drugih VM. Takođe, razmatra se stanje konsolidacije VM uzimajući u obzir strukturne karakteristike *data* centra gde je *cloud* implementiran.

Problem se može opisati na sledeći način. Na početku svake faze, analizira se lista novopristiglih VM, trenutno iskorišćenje CPU od strane postojećih VM i VM isključenih u prethodnoj fazi, struktura *cloud data* centra uključujući lokaciju servera u *rack*-ovima i mrežu koja povezuje servere i *rack*-ove, kao i mehanizam koji se koristi za njihovo hlađenje. Na osnovu toga se VM dodeljuju fizičkim mašinama, u nekim slučajevima koristeći mehanizam *live* migracije, tako da je ukupna potrošnja energije u *cloud data* centru svedena na minimum.

Ukupni dijagram optimizacije je prikazan na Slici 4. Ovde se koristi heuristika kako bi se smanjila složenost problema i kako bi se obezbedila odgovarajuća rešenja za postavljanje VM za *online* upotrebu. Prvo se posmatraju samo novopristigle VM i VM na preopterećenim serverima, i nakon što se postave na fizičke mašine, premeštaju se VM sa nedovoljno iskorišćenih fizičkih mašina da bi se isključile. Kao što pokazuje gornji deo Slike 4, u svakoj fazi se biraju odgovarajuće VM sa preopterećenih servera i spajaju se sa novopristiglim VM kao ulaz za blok algoritma postavljanja VM u prvom koraku. Nakon postavljanja ovih VM na servere, sve VM sa nedovoljno iskorišćenih servera se migriraju na odgovarajuće servere koristeći algoritme za postavljanje VM, kao što je prikazano u donjem delu Slike 4. Ova činjenica dovodi do prebacivanja servera u režim mirovanja kako bi se smanjila ukupna potrošnja energije *data* centra. Shodno tome, primećuje se da se algoritam za postavljanje VM pokreće dva puta da bi se završilo postavljanje po fazi.



Slika 4. Algoritam za postavljanje VM [10]

Za migraciju VM, razmatraju se dva praga koji sadrže donju i gornju granicu iskorišćenja CPU za servere, tako da se iskorišćenost CPU održava između ovih pragova

[10]. Ako iskorišćenost CPU padne ispod donjeg praga, server se naziva nedovoljno iskorišćenim i sve VM moraju da se migriraju sa ovog *host*-a. Shodno tome, *host* se prebacuje u režim mirovanja kako bi se eliminisala potrošnja energije. Ako iskorišćenost CPU premašuje gornji prag, server se naziva preopterećenim i neke VM moraju da migriraju sa *host*-a da bi se smanjilo opterećenje. Ovaj poslednji potez omogućava da se privremeni skokovi u iskorišćenju CPU-a za VM mogu prilagoditi bez kršenja SLA. Za izbor najbolje VM za migraciju sa preopterećenog *host*-a, uzimaju se u obzir dva uslova: prvo, biraju se VM koji imaju iskorišćenost veću od razlike između iskorišćenja *host*-a i gornjeg praga; drugo, ako VM treba da se migrira sa *host*-a, bira se VM koja ostvaruje najveću iskorišćenost (iako je ispod gornjeg praga) na *host* serveru. Algoritam se nastavlja sve dok iskorišćenost *host*-a ne postane manja od gornjeg praga iskorišćenja. Ova politika vodi do minimalnog broja migracija VM kako bi se smanjili troškovi migracije [10].

5. Neki predloženi algoritmi za energetska efikasnu alokaciju resursa

Algoritmi za energetska efikasnu alokaciju resursa mogu se klasifikovati na: HVRAA (*Heuristics Virtual Resource Allocation Algorithm*), MP-VRSA (*Multidimensional Power-aware based Virtual Resources Scheduling Algorithm*) i VROA (*Virtual Resource Optimization algorithm*) [8].

5.1. HVRAA algoritam

Glavni cilj HVRAA algoritma je da dodeli sve VM koje zahtevaju korisnici, minimalnom broju PM. Osnovna ideja HVRAA je sledeća: izaberu se sve VM koje mogu da odgovaraju *host*-u i na kraju, ako ne postoji nijedan koji odgovara trenutnom *host*-u, onda se pokreće novi *host* dok sve VM ne budu dodeljene *host*-ovima. U retkim situacijama, u kojima postoji mnogo slobodnih *host*-ova dostupnih za dodatne VM, a odnos PM i VM je visok, primenjuje se algoritam optimizacije za migraciju VM sa nedovoljno iskorišćenih *host*-ova na druge *host*-ove i isključivanje *host*-ova, čime se smanjuje potrošnja energije. U tu svrhu koristi se niži prag tako da VM na *host*-u ispod donjeg praga migriraju na druge *host*-ove, a zatim se *host* gasi.

5.2. MP-VRSA algoritam

Osnovni cilj MP-VRSA algoritma je dalje smanjenje potrošnje energije identifikacijom i otkrivanjem preopterećenih *host*-ova. MP-VRSA se sastoji od četiri koraka:

1. Otkrivanje preopterećenih *host*-ova
2. Odabir vm koje treba da migriraju sa preopterećenih *host*-ova
3. Izbor novih *host*-ova za vm koje će biti migrirane
4. Primena operacije migracije za sve preopterećene *host*-ove.

Prvi korak u MP-VRSA je otkrivanje preopterećenih *host*-ova gde se koristi strategija otkrivanja preopterećenih *host*-ova da bi se utvrdilo da li VM treba migrirati. Koraci u nastavku su neophodni za otkrivanje preopterećenih *host*-ova i postavljanje praga za iskorišćenost kapaciteta. Ako iskorišćenost CPU *host*-a premašuje prag, takav

host se može otkriti i neke VM treba migrirati. Nakon prvog koraka u MP-VRSA se primenjuje strategija izbora VM maksimalne korelacije. Ideja strategije izbora VM maksimalne korelacije je sledeća: što je veća korelacija između opterećenja VM koje rade na *host*-u, veća je verovatnoća preopterećenja *host*-a. Dakle, prema ovoj ideji, biraju se VM za migraciju koje imaju najveću korelaciju u iskorišćenosti CPU-a sa drugim VM. Treći korak u MP-VRSA je strategija postavljanja VM, gde je glavni zadatak ove strategije da izabere odgovarajući *host* za VM koje su migrirane. Međutim, kada se VM ponovo lociraju na druge *host*-ove, to će povećati iskorišćenost CPU-a. Dakle, strategija povećanja minimalne snage (MPIS - *Minimum Power Increasing Strategy*) je dizajnirana da brzo postavi VM u *host*-ove i smanji potrošnju energije.

5.3. VROA algoritam

Algoritam za optimizaciju virtuelnih resursa (VROA) migrira VM sa *host*-ova sa najmanjim korišćenjem resursa na druge PM i prebacuje originalni *host* u režim mirovanja. Zbog toga može smanjiti potrošnju energije *data* centara. VROA se sastoji od tri glavna koraka: nakon što je korak planiranja virtuelnih resursa završen, VROA će izabrati PM sa najnižim iskorišćenjem CPU-a i pokušati da migrira VM na druge *host*-ove. Zatim će sistem postaviti *host* sa najmanjim iskorišćenjem CPU u režim mirovanja nakon uspešne migracije VM na druge *host*-ove. Ako bilo koja od VM na *host*-u sa najmanjim iskorišćenjem CPU ne može da se migrira, onda *host* ostaje aktivan i sve migracije VM se otkazuju [8].

6. Zaključak

Virtuelizacija, kao jedna od najznačajnijih karakteristika *Cloud Computing*-a, dovela je do transformacije konvencionalnih *data* centara u fleksibilnu *cloud* infrastrukturu. Upravljanje potrošnjom energije primenom konsolidacije VM je kritična tehnika kada se razmatra kvalitet servisa.

U ovom radu je dat pregled energetske efikasnosti algoritama za konsolidaciju VM u *cloud data* centrima. Cilj je da se smanji potrošnja energije tako da bude u skladu sa zahtevima SLA. Glavna ideja je da se minimizira potrošnja energije s obzirom na iskorišćenost *host*-ova. Zbog brzog rasta kapaciteta *data* centara, smanjenje potrošnje energije postalo je glavna briga u odnosu na njenu cenu i uticaj na životnu sredinu. Većina pristupa fokusirana je na korišćenje *live* migracija da bi se rasteretili nedovoljno iskorišćeni *host*-ovi i na taj način prešli u stanje niske potrošnje, rešavajući na taj način problem energetske efikasnosti. Postoji niz algoritama za konsolidaciju VM koji imaju za cilj smanjenje potrošnje energije *data* centara uzimajući u obzir strukturne karakteristike, kao što su *rack*-ovi i mrežna topologija *data* centra koji leži u osnovi *cloud*-a. Na osnovu pregleda literature, pokazano je da je energetska efikasnost u *cloud* okruženju sve aktuelnija tema. U budućnosti se očekuje da će joj biti posvećena još veća pažnja za rešavanje izazova u smislu balansiranja između potrošnje energije i performansi sistema.

Zahvalnica

Rad je finansiran od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije.

Literatura

- [1] The NIST Definition of Cloud Computing - Recommendations of the National Institute of Standards and Technology, 2011.
- [2] R. Yadav, O. Kaiwartya, W. Zhang, P.Singh, "Adaptive Energy-Aware Algorithms for Minimizing Energy Consumption and SLA Violation in Cloud Computing", *IEEE Access*, 6, 2018, pp. 55923-55936.
- [3] M.F. Bari, "Data Center Network Virtualization: A Survey," *IEEE Communications Surveys & Tutorials*, 15(2), 2013, pp. 909-928.
- [4] R. Messier, *Collaboration with Cloud Computing: Security, Social Media, and Unified Communications*, Syngress, 2014.
- [5] D. Marinescu, *Cloud Computing: Theory and Practice*, Elsevier Science, 2013.
- [6] Y. Saadi, S. El Kafhali, "Energy-Efficient Strategy for Virtual Machine Consolidation in Cloud Environment," *Soft Computing*, 24, 2020, pp. 14845-14859.
- [7] N. Hamdi i W. Chainbi "A Survey on Energy Aware VM Consolidation Strategies," *Sustainable Computing: Informatics and Systems*, 23, 2019, pp. 80–87.
- [8] M. Deiab, D. El-Menshawy, S. El-Abd, A. Mostafa, "Energy Efficiency in Cloud Computing", *International Journal of Machine Learning and Computing*, 9(1), 2019, pp. 98-102.
- [9] S. Ricciardi, F. Palmieri, J. Torres-Vinals, B. Di Martino, G. Santos-Boada, J. Sole-Pareta, "Green Data Center Infrastructures in the Cloud Computing Era," in *Handbook of Green Information and Communication Systems*, Elsevier
- [10] A. Pahlavan, M. Goudarzi, "Structure-Aware Online Virtual Machine Consolidation for Datacenter Energy Improvement in Cloud Computing", *Computers & Electrical Engineering*, 42, 2015, pp. 74-89.

Abstract: *The development of Cloud Computing introduced significant changes in the Internet environment since it represents the paradigm for enabling network access to a scalable and elastic pool of shareable physical or virtual resources with self-service provisioning and administration on-demand. The development of Cloud Computing in terms of quality of service, costs, security and energy consumption requires technology that can satisfy strict requirements. Virtualization is crucial in Cloud Computing, representing an enabling technology that creates an intelligent abstraction layer which allows different virtual machines to share the same hardware. This paper describes the virtualization and allocation of cloud resources in detail. The importance of reduced energy consumption is presented, as well. Finally, the aim of the paper is to analyze the possibility of achieving energy efficiency by consolidating virtual machines in the cloud environment.*

Keywords: *virtualization, data center, migration, virtual machine, energy efficiency*

IMPROVING ENERGY EFFICIENCY IN THE CLOUD ENVIRONMENT BY THE VIRTUAL MACHINES CONSOLIDATION

Branka Mikavica, Aleksandra Kostić-Ljubisavljević, Nevena Simović