

استفاده از یک الگوریتم بهینه سازی چند هدفه برای تخصیص کارها در سیستم- های مبتنی بر ابر با هدف کاهش انرژی مصرفی

سارا طبقچی میلان* نیما جعفری نویمی پور**

*کارشناسی ارشد، گروه مهندسی کامپیوتر، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران
**دکتری، گروه مهندسی کامپیوتر، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران
تاریخ دریافت: ۱۳۹۸/۱۲/۲۱ تاریخ پذیرش: ۱۳۹۹/۰۷/۳۰
نوع مقاله: پژوهشی

چکیده

امروزه فناوری‌های نوین باعث افزایش تقاضای کسب و کار در محیط وب شده است. لذا افزایش تقاضا منجر به افزایش تنوع و تعداد خدمات و در نتیجه ایجاد مراکز داده رایانش با مقیاس بزرگ شده است که علاوه بر هزینه‌های عملیاتی بالا، مقادیر عظیمی از توان الکتریکی را مصرف می‌کند. از طرفی سیستم‌های خنک کننده ناکافی و ناکارآمد، نه تنها باعث گرم شدن بیش از حد منابع و کاهش عمر کاری دستگاه‌ها می‌شود، بلکه باعث تولید کربن می‌شود که در وضعیت آب و هوا نقش مهمی دارد. از این رو باید با روش‌های مناسب، مصرف انرژی کل این سیستم‌ها را کاهش داد. در این پژوهش، یک روش موثر مدیریت منابع انرژی در مراکز داده ابری مجازی شده ارائه شده است که علاوه بر کاهش مصرف انرژی و هزینه‌های عملیاتی، باعث افزایش کیفیت خدمات نیز شده است. این پژوهش، به ارائه یک استراتژی تخصیص منبع در سیستم‌های ابری با هدف کاهش انرژی و هزینه اجرا پرداخته و کاربرد آن را در محیط رایانش ابری را بررسی می‌نماید. نتایج حاصل از شبیه‌سازی نشان می‌دهد که روش پیشنهادی می‌تواند نسبت به روش‌های NPA، DVFS، ST و MM، میانگین انرژی مصرفی را تا ۰٫۶۲۶ کیلووات ساعت کاهش دهد، همچنین نیاز به مهاجرت و موارد نقض SLA نیز کاهش پیدا نمود.

واژگان کلیدی: رایانش ابری، انرژی مصرفی، تخصیص کار، الگوریتم بهینه‌سازی چند هدفه.

۱- مقدمه

زیرساخت معقول را فراهم می‌نماید [۴] [۵]، [۶]. با ظهور عصر رایانش ابری، تقاضا برای منابع فناوری اطلاعات و خدمات رسانی‌ها برای اتصال به اینترنت به‌طور مداوم در حال افزایش می‌باشد [۵]، [۷]. این فناوری شامل مفهوم رایانش موازی توزیع‌شده^۷ به منظور فراهم نمودن اشتراک‌گذاری منابع، سخت‌افزار، نرم‌افزار و اطلاعات برای

امروزه اینترنت و خدمات وب به اشتراک‌گذاری اطلاعات را تسهیل نموده‌اند [۱-۳]. رایانش ابری^۱ نیز به‌عنوان یک رویکرد جدید و مبتنی بر اینترنت قابلیت دسترسی^۲ بالا، مقیاس‌پذیری مناسب^۳، محاسبات انعطاف‌پذیر^۴، کارایی^۵ بالا، ظرفیت^۶ زیاد و هزینه‌های

⁵ Performance
⁶ Capacity
⁷ Distributed computing

¹ Cloud computing
² Available
³ Scalability
⁴ Flexible

نویسنده مسئول: نیما جعفری نویمی پور jafari@iaut.ac.ir

ادامه این مقاله به ترتیب زیر سازمان یافته است. در فصل دوم، به مطالعه و بررسی کارهای مرتبط پرداخته است. فصل سوم، به ارائه کامل روش پیشنهادی اختصاص یافته است. در فصل چهارم به معرفی محیط شبیه سازی و پارامترهای شبیه سازی و ارائه نتایج حاصل پرداخته شده است و در نهایت فصل پنجم نتیجه گیری و کارهای آتی ارائه شده است.

۲- کارهای مرتبط

در این بخش، برخی از روش‌های مرتبط در زمینه تخصیص منابع در رایانش ابری به همراه مزایا و معایب آن‌ها بررسی و ارائه می‌شوند.

یک روش مبتنی بر نظریه بازی‌ها برای زمانبندی مبتنی بر خدمات ابری با همکاری نیازهای کیفیت خدمات ارائه شده است [۲۰]. ارائه خدمات محاسباتی هزینه‌هایی به همراه خواهد داشت که به تعداد زیادی رقابت وابسته است. هر وظیفه محاسباتی چندین وابستگی و زیر وظایف همگن دارد که در زمان اجرا تاثیر دارند. نظریه بازی‌ها برای یافتن راه حل بهینه در نظر گرفته شده است. روش معرفی شده به این صورت است که در مرحله اول برای دست یافتن به بهینه سازی اولیه روش باینری پیشنهاد شده است که تخصیص دوباره منابع در نظر گرفته نشده است. سپس بر اساس نتایج اولیه مکانیسم تکاملی برای رسیدن به راه حل مطلوب و منصفانه طراحی شده است. در این روش بهره‌وری و کیفیت خدمات افزایش یافته‌اند و پیچیدگی زمانی کاهش یافته است؛ اگر چه هزینه و انرژی مصرفی در نظر گرفته نشده است.

در ادامه یک روش تخصیص منابع انرژی موثر مبنی بر ایمنی کلونال^۱ برای مراکز داده مجازی ارائه شده است [۲۱]. الگوریتم تخصیص منبع موثر، منابع را به کارها تخصیص می‌دهد طوری که انرژی مصرفی مراکز داده و زمان تکمیل اجرای کارها را کمینه می‌کند. در این پژوهش، بهبود تخصیص منابع با استفاده از الگوریتم انتخاب کلونال مبتنی بر بهینه سازی زمان تکمیل کارها و انرژی مصرفی می‌باشد. در روش پیشنهادی ابتدا عملگر کلونال یک نقشه تصادفی آنتی بال می‌باشد. در سیستم ایمنی بیولوژیکی کلونی یک گروه از سلول‌های یکسان می‌باشد و در سیستم ایمنی منظور از

کامپیوترها و یا دیگر دستگاه‌ها می‌باشد [۵، ۸] و حتی کل برنامه‌های کاربردی نرم‌افزاری به کاربران تحت خدمات حین تقاضا تحویل داده می‌شود [۹]. رایانش ابری یک مدل خدمت برای تأمین خدمات فناوری اطلاعات است که اغلب مبنی بر مجازی‌سازی و فناوری‌های رایانش توزیع شده می‌باشد [۱۰]. منابع ابری شامل انواع مختلف منابع از جمله منابع ذخیره‌سازی^۲، پردازش^۳، حافظه^۴، پهنای باند شبکه^۵ و ماشین‌های مجازی^۶ است [۱۱]. نرم‌افزار به‌عنوان خدمت [۱۲، ۱۳]، زیرساخت به‌عنوان خدمت [۱۴، ۱۵] و بستر نرم‌افزاری به‌عنوان خدمت^۹ [۱۶] سه شکل معروف منابع مجازی^{۱۰} در رایانش ابری می‌باشند [۵].

از طرف دیگر، تخصیص منابع یک فرآیند توزیع شده بین برنامه‌ها و کاربران از دیدگاه اقتصادی است که منابع در دسترس را در اختیار برنامه‌های کاربردی ابری قرار می‌دهد که به عنوان مسئله *Np-hard* شناخته شده است [۱۷]. هدف از تخصیص منابع برای ارائه خدمات منحصر به فرد بهینه سازی کیفیت خدمات و بهبود بهره‌وری منابع و انرژی می‌باشد [۱۸]. با توجه به افزایش روزافزون محبوبیت رایانش ابری، اگر انرژی مصرفی در ارائه دهندگان خدمات کنترل نگردد، آنگاه در گام اول هزینه ارائه خدمت افزایش می‌یابد و در پی آن هزینه پرداختی خدمت گیرندگان افزایش خواهد داشت. علاوه بر این، افزایش مصرف انرژی سهم زیادی در افزایش آلودگی محیط زیست خواهد داشت؛ لذا کشف راهکارهای بهره‌وری انرژی بسیار ضروری است. از این رو ضرورت ایجاد می‌کند تا با بهبود الگوریتم‌های تخصیص منبع به همراه شناسایی پارامترهای تأثیرگذار و کاهش مصرف انرژی، سطح رضایت کاربر نیز افزایش داده شود. در نتیجه این امر باعث کاهش هزینه تمام شده خدمات نیز خواهد شد. با توجه به ماهیت *NP-hard* این مساله [۱۹]، در این پژوهش از یک الگوریتم کارآمد به نام بهینه سازی چند هدفه برای تخصیص منابع ابری استفاده شده است که اهداف آن به شرح زیر خواهد بود:

- کاهش مصرف انرژی
- کاهش هزینه کل
- کاهش میزان نیاز به مهاجرت
- افزایش توازن بار

⁷ Software as a Service(SaaS)

⁸ Infrastructure as a Service(IaaS)

⁹ Platform as a Service(PaaS)

¹ Virtualized resources

¹ Clonal immune

¹ makespan

¹ Virtualization

² Storage

³ Processing

⁴ Memory

⁵ Network bandwidth

⁶ Virtual Machine(VM)

بهره‌وری بالای منابع و کاهش هزینه‌ها ارائه شده است [۲۵]. در روش پیشنهادی ابتدا درخواست کاربر بر اساس مهلت زمانی و در دسترس بودن منابع ارسال می‌شود. کیفیت خدمات در روش پیشنهادی شامل پارامترهای همگن و ناهمگن از جمله زمان اجرا، سطح اطمینان مشتری براساس *SLA* می‌باشد. در مرحله بعدی کیفیت خدمات مد نظر متقاضی در منابع موجود بررسی می‌شود و لیستی از منابع ایجاد می‌گردد و مطابق درخواست تامین منابع صورت می‌گیرد و عمل تخصیص صورت می‌پذیرد. در روش ارائه شده، بهره‌وری منابع، کیفیت خدمات و زمان اجرا بهبود یافته است؛ اگرچه هزینه، زمان پاسخ، انرژی مصرفی و زمان تکمیل کارها به خوبی بررسی نشده است.

در مراکز داده ابری، ماشین‌های مجازی نیاز به تخصیص در ماشین‌های فیزیکی دارند بطوری که کمترین میزان هدر رفت از منابع را داشته باشند. لذا روشی مبتنی بر گرده افشانی به منظور کاهش مصرف انرژی پیشنهاد شده است [۲۶]. در این روش فریم‌ورکی در نظر گرفته شده است که بوسیله‌ی آن نزدیکترین و بهینه‌ترین راه حل یافت می‌شود که این راه حل بر اساس میزان حافظه و پردازنده منابع خواهد بود. در روش ارائه شده، توازن بار، انرژی مصرفی و بهره‌وری از منابع بهبود یافته؛ اگرچه هزینه، زمان پاسخ و زمان تکمیل کارها به خوبی بررسی نشده است.

طبق مقالات مطالعه شده قبلی همانطور که آشکار است عمده‌ترین مسائل در رایانش ابری انتخاب منبع مناسب برای مشتریان می‌باشد بصورتی که کیفیت خدمات در آن نقض نشود. لذا یک روشی چند هدفه مبتنی بر ژنتیک به منظور تخصیص منابع با به حداقل رساندن هزینه‌ها و افزایش کیفیت خدمات ارائه شده است [۱۹]. بطور کلی فراهم‌کنندگان خدمات ابری تعداد نمونه‌های محاسباتی و پایگاه‌های داده را مجزا می‌نمایند. در روش پیشنهادی منابع مناسب براساس الگوریتم ژنتیک به تناسب تقاضاهای مشتری تخصیص می‌یابند. در روش پیشنهادی زمان پاسخ و کیفیت خدمات بهبود یافته است؛ اگرچه زمان تکمیل اجرای کل کارها، هزینه و انرژی مصرفی در نظر گرفته نشده است.

به طور خلاصه، در هیچ یک از روش‌های ذکر شده، کاهش زمان اجرای کل کارها، انرژی مصرفی و هزینه در محیط‌های ابری با استفاده از تخصیص منابع مورد استفاده قرار نگرفته‌اند؛ لذا روشی که قادر است تخصیص منابع در محیط‌های ابری را با استفاده از الگوریتم بهینه‌سازی چند هدفه انجام دهد در بخش بعدی بررسی و ارائه می‌شود.

کلونی یک گروه از سلولهای یکسان می‌باشد که از یک اجداد مشترک تنها تولید می‌شود. در این روش زمان پاسخ، زمان تکمیل اجرای کارها و انرژی مصرفی کاهش یافته است؛ اما هزینه و توازن بار به خوبی بررسی نشده است.

در ادامه کارها و روش‌های مطالعاتی، سیستمی مبتنی بر فناوری مجازی سازی برای تخصیص منابع مراکز داده ارائه شده و دستیابی به رایانش سبز را با بهینه کردن تعداد سرورها تضمین می‌کند [۲۲]. در این پژوهش، از چولگی برای محاسبه‌ی ناهمواری‌ها در بهره‌وری منابع در یک سرور استفاده شده است و با به حداقل رساندن چولگی می‌توان حجم کارهای مختلفی را ترکیب نمود و بهره‌وری کلی از منابع بهبود می‌یابد. در روش ارائه شده، توازن بار، انرژی مصرفی و بهره‌وری از منابع و تعداد سرورهای استفاده شده بهبود یافته؛ اگرچه هزینه، زمان پاسخ و زمان تکمیل کارها به خوبی بررسی نشده است.

همچنین، در ادامه یک روش مبتنی بر بهینه‌سازی رویکردی تخصیص منبع برای برنامه‌های کاربردی خدمت به عنوان نرم افزار در محیط ابری ارائه شده است [۲۳]. از آنجا که کاربران نیاز به ارسال خدمات به ارائه‌دهندگان ابری دارند، انتخاب ارائه‌دهنده مناسب بسیار مهم است. در روش پیشنهادی ارائه‌دهندگان خدمات به عنوان تامین‌کننده خدمات نرم افزاری عمل می‌کند که سرورها درخواست‌های کاربران خدمت نرم افزاری را با افزایش کیفیت خدمات تضمین می‌کنند. در روش پیشنهادی تخصیص منابع و هزینه انرژی مصرفی بهبود یافته است؛ اگرچه زمان تکمیل اجرای کل کارها، هزینه و انرژی مصرفی در نظر گرفته نشده است.

یک چارچوب مبتنی بر سطح توافقنامه خدمات به منظور تسهیل در تخصیص منابع ارائه شده است که حجم کارها و موقعیت جغرافیایی مراکز داده توزیع شده در نظر گرفته شده است [۲۴]. در این چارچوب، ارائه‌دهنده ابر چندین مراکز داده دارند که در مناطق جغرافیایی گسترش یافته‌اند. در این پژوهش، بر تخصیص منابع پویا در مرکز داده بدون در نظر گرفتن تاخیر در شبکه جهانی تمرکز شده است که تخصیص منابع بر اساس انتخاب مراکز داده در میان مراکز داده توزیع شده می‌باشد. در این روش زمان پاسخ، توازن بار، هزینه بهبود یافته است؛ اگرچه انرژی مصرفی، زمان تکمیل کارها در نظر گرفته نشده است.

تخصیص و تامین منابع بر اساس کیفیت خدمات به شیوه‌ی موثری انجام می‌شود ولی در موارد خاصی تامین منابع بصورت پویا صورت نمی‌گیرد و بسیاری از کارهای انجام شده بر اساس مدیریت منابع همراه با هزینه و سرعت بالایی است. لذا سیاستی مبتنی بر مدیریت موثر در تخصیص منابع با زمان کم بر اساس منابع قابل دسترس و

۳- روش پیشنهادی

در این بخش نسبت به معرفی مساله بهینه سازی چند هدفه و مجموعه پارتو، تشریح الگوریتم بهینه سازی ازدحام ذرات و در نهایت روش پیشنهادی پرداخته خواهد شد.

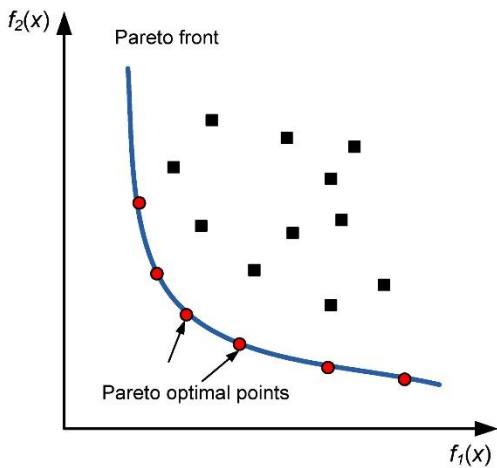
۳-۱. بهینه سازی چند هدفه

بهینه‌سازی با مسائلی سر و کار دارد که هدف آن انتخاب بهترین گزینه از مجموعه‌ای از پاسخ‌ها بر اساس معیارهای خاص می‌باشد. اگر تنها یک معیار در نظر گرفته شود، مساله بهینه‌سازی به مساله بهینه سازی تک هدفه تبدیل می‌شود، که مطالعات گسترده‌ای طی ۵۰ سال گذشته در خصوص آن صورت پذیرفته است. اگر بیش از یک معیار وجود داشته باشد، که می‌بایستی همزمان مد نظر قرار گیرد، مساله بهینه سازی چند هدفه مطرح خواهد شد [۲۷] که این مسائل در طراحی، مدل‌سازی و برنامه‌ریزی سیستم‌های واقعی بسیار پیچیده در حوزه صنعت، حمل و نقل شهری، مالی، مدیریت منابع طبیعی، توزیع انرژی و غیره کاربرد دارد. عمده مسائل تصمیم‌گیری در دنیای واقعی شامل اهداف متعدد هستند و اعمال محدودیت‌های مختلف بر انتخاب راه حل بهینه لازم است که حل مسائل را پیچیده می‌کند. برای حل این مساله راهکارهایی ارائه شده است که مهمترین آنها روشی است که توسط پارتو ارائه شده است [۲۸]. اخیراً الگوریتم بهینه سازی ازدحام ذرات، توجه ویژه‌ای را برای حل مساله بهینه‌سازی چند هدفه به سمت خود جذب کرده است، که نتایج آن در تحقیقات و گزارش‌هایی که به عنوان "ارزیابی بهینه سازی چند هدفه" (EMO) ارائه شده است [۲۹، ۳۰]. صورت کلی مساله بهینه سازی چندهدفه با رابطه (۱) مدل می‌شود.

$$\min F(x) = \{f_1(x), \dots, f_n(x)\} \quad (1)$$

که $f_i(x)$ تابع هدف نام مساله بهینه‌سازی چندهدفه است.

هدف کمینه کردن توابع چندگانه $f_i(x)$ بصورت توامان است. همانطور که در شکل ۱ مشاهده می‌شود که مثالی از نمودار پارتو دو هدفه می‌باشد، هدف کمینه کردن دو تابع f_1 و f_2 است. اگر ۲ عضوی از مجموع جواب‌های پارتو باشد، ۱ بر ۲ غلبه دارد و اگر در هیچ یک از دو تابع بدتر از ۲ نباشد، حداقل در یک تابع بهتر از آن می‌باشد.



شکل ۱: مثالی از نمودار پارتو دو هدفه [۳۱]

یک روش کارا برای بهینه سازی چند هدفه، بهترین راه حل را از مجموعه‌ای از راه‌حل‌های (بهترین مجموعه شناسایی شده پارتو) ممکن انتخاب می‌کند، اما هر روشی که برای حل مساله بهینه‌سازی چند هدفه ارائه شده باید سه هدف متضاد زیر را برآورده کند:

- ۱- مجموعه پارتو منتخب باید تا حدی که ممکن است به پارتو درست همگرا باشد، در حالت ایده‌ال، مجموعه پارتو منتخب باید زیر مجموعه‌ای از مجموعه بهینه پارتو باشد.
 - ۲- راه‌حل‌های مجموعه پارتو منتخب باید توزیع یکنواخت داشته باشند و تنوع آنها بر روی مجموعه پارتو حفظ شود تا تصمیم‌گیری واقعی‌تر و تصویر درست‌تری از واقعیت داشته باشد.
- مجموعه پارتو منتخب بایستی تمام طیف مجموعه پارتو را پیمایش کند و اینکار مستلزم آن است که راه‌حلی که بیشترین فاصله را با تابع هدف دارند بررسی شوند. در روش پیشنهادی از بهینه‌سازی پارتو برای دو هدف اصلی پژوهش، کاهش مصرف انرژی و کاهش زمان کل اجرا استفاده شده است، روابط زیر:

$$\text{Obj1: } \min(\text{Energy}) = \sum_{i=1}^m \text{Energy}_i \quad (2)$$

$$\text{Obj2: } \min(\text{TotalTime}) = \sum_{j=1}^n \text{TotalTime}_j \quad (3)$$

که در آن m نشان دهنده تعداد منابع و n نشان دهنده تعداد وظایف است.

۲-۳. الگوریتم بهینه سازی ازدحام ذرات

با توجه به کارایی الگوریتم ازدحام ذرات در حل مسائل تک هدفه، سیرا و کوئلو کوئلو در سال ۲۰۰۶ با ایجاد تغییراتی در ساختار این الگوریتم، الگوریتم چند هدفه بهینه سازی ازدحام ذرات را معرفی کردند. تحقیقات انجام شده، حاکی از عملکرد بالای این الگوریتم در حل مسائل بهینه سازی چند هدفه می باشد. در واقع این الگوریتم تعمیمی است از الگوریتم بهینه‌سازی ازدحام ذرات که برای حل مسائل چند هدفه بکار می‌رود. در الگوریتم بهینه‌سازی ازدحام ذرات یک مفهومی به نام آرشیو یا مخزن نسبت به الگوریتم ازدحام ذرات اضافه شده است، که به تالار مشاهیر نیز معروف است [۳۲]. انتخاب بهترین جواب کلی و بهترین خاطره شخصی برای هر ذره گام مهم و اساسی در الگوریتم بهینه‌سازی ذرات چندهدفه ازدحام ذرات است.

زمانی که ذرات می‌خواهند حرکتی انجام دهند یک عضو از مخزن را به عنوان رهبر انتخاب می‌کنند. این رهبر حتما باید عضو مخزن و همچنین نامغلوب باشد. اعضای مخزن بیانگر جنبه پارتو و شامل ذرات نامغلوب هستند. پس به جای بهترین مکان سراسری ذره^۳ یکی از اعضای مخزن انتخاب می‌شود. از آنجا که در الگوریتم ازدحام ذرات تنها یک هدف وجود دارد، مخزن وجود ندارد و تنها یک ذره است که بهترین است. اما در بهینه‌سازی ازدحام ذرات چند ذره وجود دارد که نامغلوب هستند و در مجموعه جواب جای دارند.

برای مقایسه بهترین مکان محلی ذره^۴ به شکل زیر عمل می‌شود:

(۱) اگر موقعیت جدید، بهترین مکان محلی ذره را مغلوب کند، آنگاه موقعیت جدید جای بهترین مکان محلی ذره را خواهد گرفت. به بیان ریاضی:

$$Pbest_i^{n+1} = X_i^{n+1} \quad (۴)$$

(۲) اگر موقعیت جدید توسط بهترین مکان محلی ذره مغلوب شود، کاری انجام نمی‌گیرد. به بیان ریاضی:

$$Pbest_i^{n+1} = Pbest_i^n \quad (۵)$$

(۳) اگر هیچ کدام یکدیگر را مغلوب نکنند، به صورت تصادفی یکی را به عنوان بهترین مکان محلی ذره در نظر گرفته می‌شود.

ترتیب اجرای الگوریتم پیشنهادی به شرح زیر می باشد:

- (۱) تعیین پارامترهای مورد نیاز برای اجرای الگوریتم چندهدفه ازدحام ذرات: حداکثر تکرار برای اجرای الگوریتم، اندازه جمعیت، مقادیر C_0, C_1, C_2 و میزان اعضای مخزن.
- (۲) ایجاد جمعیت اولیه
- (۳) تعیین بهترین مکان محلی برای هر ذره
- (۴) جداسازی اعضای نامغلوب جمعیت و ذخیره سازی آنها در مخزن.
- (۵) انتخاب لیدر از مخزن توسط هر یک از ذرات و حرکت بر اساس آن (به روزرسانی سرعت و موقعیت)
- (۶) بهترین مکان محلی هر ذره به روزرسانی می‌گردد
- (۷) اعضای نامغلوب جدید به مخزن اضافه می‌گردند.
- (۸) اعضای مغلوب مخزن حذف می‌گردند.
- (۹) اگر شرایط خاتمه تحقق نیافته باشد، از مرحله ۵ تکرار می‌گردد.

۳-۳. مساله تخصیص منابع با رویکرد کاهش انرژی و هزینه اجرا

برای تشریح مساله تخصیص ابتدا پارامترها، شرایط، خصوصیات منابع و کارها بیان می‌شود. برای این کار منبع مجازی با VR نامگذاری شده است و بصورت زیر بیان می‌شود:

$$VR = \{CP_r, SR_r, SW_r, CO_r, ER_r\} \quad (۶)$$

که در آن، CP_r نشان دهنده تعداد منابع محاسباتی، SR_r نشان دهنده فضای حافظه، SW_r نشان دهنده مجموعه نرم افزارهایی که توسط VR پشتیبانی می‌شود، CO_r هزینه VR در واحد زمان و ER_r انرژی مصرفی VR می‌باشد. از آنجایی که ممکن است بسیاری از وظایف درخواستی مشتریان برای انجام به بیش از یک ماشین نیاز داشته باشد، بنابراین هر وظیفه به زیروظایف با اولویت و ترتیب اجرا تجزیه می‌گردد. متناظر با هر VR ، یک زیر وظیفه ST بصورت زیر تعریف می‌گردد:

$$ST = \{CP_t, SR_t, SW_t\} \quad (۷)$$

سه پارامتر ST به ترتیب عبارتند از: تعداد منابع محاسباتی مورد نیاز (CP_t)، مقدار فضای حافظه مورد نیاز (SR_t) و نوع نرم افزار مورد نیاز (SW_t).

بنابراین هر وظیفه که شامل مجموعه زیروظایف است، بصورت زیر تعریف می‌شود:

³ gbest
⁴ pbest

¹ Reyes Sierra and Coello-Coello

² Multi-Objective Particle Swarm Optimization (MOPSO)

VR ، هزینه در واحد زمان (O_r) بصورت یک عدد بیان می‌شود، هزینه کل VR با محاسبه مجموع هزینه‌های مربوط به هر وظیفه به دست می‌آید و به این ترتیب هزینه کل زمانبندی با جمع کردن همه هزینه‌های کل منابع به دست می‌آید. لازم به ذکر می‌باشد این هزینه بستگی به تعداد کل وظایف انجام شده دارد.

• **انرژی کل:** یکی از دغدغه‌های اساسی در تخصیص منابع، کاهش مصرف انرژی توسط منابع جهت صرفه جویی و مسائل زیست محیطی می‌باشد. همانند هزینه کل، انرژی کل نیز از مجموع انرژی مصرفی کل هر یک از VR ها به دست می‌آید و انرژی کل هر VR نیز از مجموع انرژی مصرفی برای هر یک از وظایف (E_r) در VR به دست می‌آید. انرژی مصرفی برابر خواهد بود با توان مصرفی P اندازه گیری شده در زمان T .

$$E = P * T \quad (9)$$

لازم به ذکر می‌باشد که نقض SLA به شرح ذیل محاسبه خواهد شد:

$$Violation_{SLA} = T_{SLAH} \times PerfDegM \quad (10)$$

که $Violation_{sta}$ برابر است با درصد نقض سطح توافق خدمات و T_{SLAH} برابر است با درصدی از زمانی که میزبان فیزیکی ۱۰۰٪ از پردازنده مرکزی را در حالت فعال به کار می‌گیرد. $PerfDegM$ نشان دهنده عملکرد سیستم به هنگام مهاجرت می‌باشد.

همانطور که مشهود است زمانی که ماشین فیزیکی بیشترین توان خود را به کار می‌گیرد در نتیجه ماشین‌های مجازی کمتری در دسترس خواهد بود و امکان استفاده منابع بیشتری نخواهد بود لذا در نظر گرفتن محدودیت منابع از مهمترین موارد می‌باشد. فرمول ۱۱ نحوه محاسبه T_{SLAH} را نمایش می‌دهد.

$$T_{SLAH} = 1 / N_{Host} * \sum T_{FUi} / T_{Active} \quad (11)$$

N_{Host} برابر با تعداد ماشین‌های فیزیکی می‌باشد. T_{FUi} برابر با زمانی می‌باشد که میزبان فیزیکی I ۱۰۰٪ پردازنده مرکزی را به کار می‌گیرد که ممکن است منجر به نقض SLA شود. $T_{Activei}$ برابر است با مدت زمانی که میزبان i در حالت فعال است.

$$T = \{ST_1, ST_2, \dots, ST_{nT}\} \quad (8)$$

nT ، نشان دهنده تعداد زیر وظایف مربوط به وظیفه T می‌باشد. فرض کنید N وظیفه و به تعداد M دستگاه VR وجود داشته باشد، بنابراین مساله تخصیص وظایف به منابع، عبارت است از تخصیص همه زیروظایف این N وظیفه بر روی M تعداد از VR ها. برای تخصیص چندین شرط اساسی وجود دارد که عبارتست از:

• **محدودیت نرم افزار:** هر وظیفه و زیروظایفش باید به ماشینی که قابلیت اجرا و رفع نیازهای نرم افزاری آن وظیفه را داراست، تخصیص داده شود.

• **محدودیت تعداد منابع:** تعداد منابع مورد نیاز (پردازش و حافظه) وظیفه و زیروظایفش نباید بیشتر از تعداد منابع VR باشد که کار مورد نظر به آن تخصیص داده شود.

• **محدودیت ترتیب و اولویت اجرای زیروظایف:** ترتیب اجرای زیروظایف بایستی بر اساس دنباله روند اجرای آن باشد و تمام زیروظایف دنباله به یک VR برای اجرا تخصیص داده شود.

• **محدودیت زمان:** وظایف بایستی به منابعی تخصیص داده شوند که در یک بازه زمانی تعیین شده اجرا و تمام شوند. عبارتی اجرای همه زیروظایف باید در فاصله زمانی تعیین شده شروع و پایان یابد.

فاکتورهای زیادی توسط ارائه دهندگان خدمت رایانش ابری برای لحاظ کردن، تنظیم زمانبندی و تخصیص منابع، تعیین می‌شود. در این بخش به چند نمونه اشاره می‌کنیم:

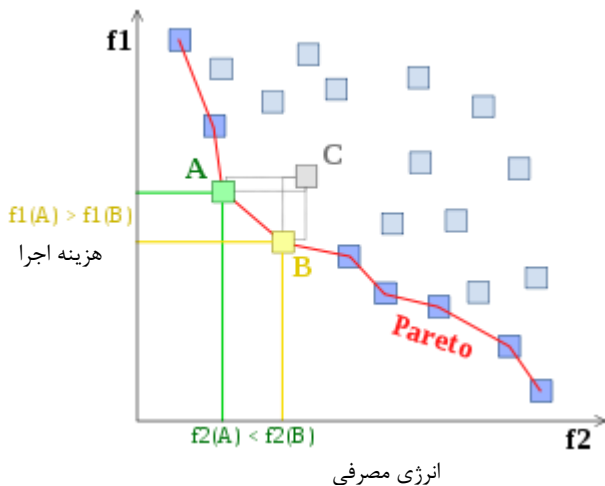
• **زمان کل اجرای وظیفه:** عبارتست از زمان صرف شده برای اجرای کامل یک وظیفه، زمان انتظار، زمان مهاجرت (در صورت نیاز).

• **کیفیت ارائه خدمات:** به منظور پاسخ به درخواست‌های رسیده از مشتریان در یک بازه زمانی مشخص به اندازه کافی VR وجود داشته باشد، در این حالت کیفیت ارائه خدمات نقش اساسی پیدا می‌کند.

• **هزینه کل:** هزینه منابع از مهمترین مواردی است که برای ارائه دهندگان خدمات و مشتریان اهمیت دارد. برای هر

که در آن نشان دهنده سرعت و x_i^t نشان دهنده موقعیت ذره v_i^t در تکرار t ام می‌باشد و pb و gb به ترتیب بهترین موقعیت یا راه حل محلی (فردی) و بهترین راه حل سراسری (کلی) می‌باشد. R و r مقادیر تصادفی در بازه احتمالاتی $[0,1]$ و w و c_1 و c_2 پارامترهای وزنی می‌باشند.

برای انتخاب بهترین موقعیت محلی و سراسری، زمانی که بیش از یک هدف مد نظر باشد، انتخاب به یک چالش اساسی تبدیل می‌شود که برای حل این مساله از مفهوم نظریه پارتو برای مقایسه و انتخاب بهترین زمانبندی یا موقعیت ذره با لحاظ کردن چندین هدف، استفاده می‌شود. فرض کنید دو هدف کاهش انرژی مصرفی و هزینه اجرا دو هدفی باشند که مد نظر هستند، بنابراین نمودار پارتو مربوطه بصورت زیر خواهد بود (شکل ۲).



شکل ۲: نمودار پارتو الگوریتم پیشنهادی

در شکل فوق، مجموعه نقاط قرمز رنگ مشخص شده بر روی روی منحنی مجموعه جبهه پارتو هستند که مخزن پارتو را تشکیل می‌دهند. این نقاط موقعیت ذراتی هستند که بهترین موقعیت محلی را نسبت به سایر موقعیت‌ها بر اساس دو هدف کاهش انرژی و هزینه اجرا دارا می‌باشند. بنابراین برخلاف الگوریتم بهینه‌سازی ازدحام ذرات، در الگوریتم بهینه‌سازی ازدحام ذرات چند هدفه، بیش از یک موقعیت بهینه خواهد بود. در صورتی که از نظریه پارتو استفاده نمی‌شود، می‌بایستی بر روی مقادیر دو هدف نرمال‌سازی انجام می‌شد و از میانگین وزنی برای تعیین یک تابع برآزش استفاده می‌شد. اما از آنجا که تعیین وزن و نرمال‌سازی مقادیر می‌تواند به دقت کار لطمه بزند و انتخاب یک مقدار میانگین می‌تواند دقت پایین‌تری داشته باشد لذا از نظریه پارتو استفاده شده و مجموعه‌ای از موقعیت‌های بهینه ذرات برای تعیین رهبر به کار گرفته می‌شود.

$$PerfDegM = 1/N_{VM} * \sum Deg_j / MIPS \quad (12)$$

N_{VM} برابر با تعداد ماشین‌های مجازی است. Deg_j برابر است با عملکرد ماشین مجازی با ۱۰٪ سرعت اجرای دستورالعمل. $MIPS$ برابر است با مجموع سرعت اجرای دستورالعمل‌ها.

و در آخر تابع هدف به شرح ذیل خواهد بود [36].

$$ESV = E * Violation_{SLA} * cost \quad (13)$$

مکانیسم تخصیص منابع بر اساس الگوریتم بهینه‌سازی ازدحام ذرات چند هدفه در الگوریتم بهینه‌سازی ازدحام ذرات، هر ذره به عنوان یک راه حل تخصیص و به شرح زیر، در نظر گرفته می‌شود.

کد گذاری ذرات: برای کدگذاری ذرات از یک بردار دو بعدی بصورت زوج مرتب استفاده می‌شود، هر زوج مرتب به ترتیب شامل شناسه ماشین مجازی و شناسه منبع می‌باشد. برای مثال، یک مساله با شش ماشین مجازی برای تخصیص به دو منبع را در نظر بگیرید. اگر اولین تخصیص این باشد که ماشین مجازی اول به منبع دوم تخصیص یابد، بصورت زوج مرتب (۱ و ۲) نشان داده می‌شود و دومین تخصیص، تخصیص دومین ماشین مجازی به اولین منبع باشد، آن نیز بصورت (۲ و ۱) نشان داده خواهد شد. در کل کد بندی ذره بصورت زیر خواهد بود:

$$\{(2,6) \text{ و } (2,5) \text{ و } (1,4) \text{ و } (1,3) \text{ و } (1,2) \text{ و } (2,1)\}$$

بهینه‌سازی فرمول‌های حرکتی ذرات: فرمول‌های ارائه شده در بخش قبل برای الگوریتم بهینه‌سازی ذرات، برای طبیعت طراحی شده، اما مقادیر مورد نظر در این پژوهش برای کدگذاری ذرات بصورت اعداد صحیح خاص که اندیس هستند بیان شده است، لذا نیاز به تغییر در فرمول‌های مذکور احساس می‌شود. بصورت ساده، برای اینکار پس از محاسبات، مقادیر حقیقی به دست آمده با مقادیر صحیح مناسب تغییر داده می‌شود. برای اینکه در این تغییر برخی از منابع به دلایل گرد کردن یا روش مشابه از گردونه خارج نشوند از یک روش مقدار دهی چرخشی برای این منظور استفاده می‌شود. ایده اصلی استفاده از الگوریتم بهینه‌سازی ازدحام ذرات، یافتن بهترین راه حل از حرکت کل ذرات با کاهش مصرف انرژی، هزینه اجرا یا زمان کل اجرا می‌باشد. هر ذره، شامل یک موقعیت که نشان دهنده راه حل و یک بردار سرعت است که نشان دهنده جهت و مقدار حرکت ذره می‌باشد. موقعیت و سرعت حرکت هر ذره را با روابط زیر بیان می‌شود [۳۳]:

$$\begin{cases} v_i^{(t+1)} = wv_i^t + c_1r(pb_i^{(t)} - x_i^{(t)}) + c_2R(gb^{(t)} - x_i^{(t)}) \\ x_i^{(t+1)} = x_i^{(t)} + v_i^{(t+1)} \end{cases} \quad (14)$$

۴-۳ مراحل روش پیشنهادی

مراحل روش پیشنهادی به شرح زیر است:

گام (۱): مقدار $t=0$ برای اولین تکرار انتخاب می‌شود. تمامی k ذره در اجتماع S_{set}^t با استفاده از الگوریتم حریصانه مقداردهی اولیه می‌شود.

گام (۲): برای هر ذره محاسبات مربوطه با استفاده از روابط (بخش قبل) و توضیحات بند (۲) این بخش انجام و K ذره جدید تولید می‌شود. مجموع ذرات جدید و قدیم به تعداد $2k$ اجتماع حافظه موقتی T_{set} را تشکیل می‌دهد.

گام (۳): با استفاده از نظریه پارتو ذرات بهینه یا همان موقعیت‌های محلی بهینه پیدا می‌شود و با PO_{set} نامگذاری می‌گردد و تعداد اعضای مجموعه بهینه با N_{po} نشان داده می‌شود. اگر $N_{po} < K$ باشد گام (۴) و در غیر اینصورت گام (۵) اجرا می‌شود.

گام (۴): به تعداد $K - N_{po}$ ذره از بین ذراتی که بهینه پارتو نیستند برای اضافه کردن به مجموعه PO_{set} به خاطر تکمیل کردن تعداد مجموعه و رساندن آن به K تعداد انتخاب می‌شود تا مجموعه اجتماع ذرات در مرحله بعد S_{set}^{t+1} را تشکیل دهند.

گام (۵): به صورت تصادفی K تعداد از مجموعه انتخابی بهینه پارتو PO_{set} برای ایجاد مجموعه اجتماع ذرات مرحله بعد S_{set}^{t+1} انتخاب می‌شود.

گام (۶): شرط خاتمه الگوریتم بررسی می‌شود که شرایط به دست آمده مطلوب است یا تعداد تکرارها پایان یافته یا خیر. لازم به ذکر است که منظور از شرایط خاتمه شرایط تخصیصی می‌باشند که در بخش ۳-۳ به آن پرداخته شده است که شامل محدودیت نرم افزار، محدودیت تعداد منابع، محدودیت ترتیب و اولویت اجرای زیر وظایف است. (رجوع شود به بخش ۳-۳) در صورت تحقق شرط خاتمه گام (۸) و در غیر اینصورت گام (۷) اجرا می‌شود.

گام (۷): مقدار $t=t+1$ به روز شده و گام (۲) اجرا می‌شود.

گام (۸): الگوریتم پایان یافته و به مجموعه بهینه پارتو PO_{set} برای تصمیم گیری و انتخاب راه حل بهینه سراسری مراجعه می‌شود.

۴-۴ نتایج

در این بخش محیط شبیه سازی، پارامترهای شبیه سازی و نتایج حاصل بررسی می‌شود.

۴-۱-۱- محیط شبیه سازی

از آنجا که امکان آزمایش عملی با توجه به گستردگی محیط های ابری و در دسترس نبودن آن میسر نیست، برای آزمایش و اعتبارسنجی الگوریتم پیشنهادی، از زبان برنامه‌نویسی جاوا در محیط نت بینز آی دی ای ۸٫۱ و نسخه ۳٫۰٫۱ بسته شبیه سازی کلودسیم [۳۴]، کتابخانه الگوریتم های فرااکتشافی و ابزار کمی محیط جاوا با عنوان MOEA framework استفاده شده است. کامپیوتر مورد استفاده با مشخصات پردازنده اینتل با حافظه ۴ گیگابایت و تحت سیستم عامل ویندوز ۷ انتخاب گردید.

۴-۲-۴ پارامترهای شبیه سازی

برای شبیه سازی از یک سناریو با اعمال مقادیر مناسب برای منبع تغذیه و انرژی مصرفی استفاده شده است [۲۹] که مقادیر آن در جدول ۱ آمده است.

جدول ۱: پارامترها و مقادیر اولیه شبیه سازی

HP ProLiant ML110 G4 (1 x [Xeon 3040 1860 MHz, 2 cores], 4GB) HP ProLiant ML110 G5 (1 x [Xeon 3075 2660 MHz, 2 cores], 4GB)	ماشین فیزیکی
سرعت اجرای دستور العمل: {۱۸۶۰ و {۲۶۶۰ تعداد پردازنده ها: ۲ رم: ۴۰۹۶ پهنای باند: ۱ گیگا بایت بر ثانیه حافظه: ۱ گیگا بایت	
سرعت اجرای دستورالعمل: {۲۵۰۰ و {۲۰۰۰ و ۱۰۰۰ و ۵۰۰ تعداد پردازنده ها: ۱ رم: {۸۷۰ و ۱۷۴۰ و ۱۷۴۰ و ۶۱۳ پهنای باند: ۱۰ مگابیت بر ثانیه	ماشین مجازی
طول: ۲۱۶۰ مگا بایت تعداد پردازنده ها: ۱	وظایف

۴-۳-۴ نتایج بدست آمده

لحظه آماده به کار شدن می‌گردد. میانگین زمان کل اجرا برای این سناریو برابر است با ۰,۰۰۵۳۳ ثانیه، که در مقایسه با سایر سناریوها تحلیل خواهد شد. انحراف معیار زمان کل اجرای آن ۰,۰۰۵۱۱ ثانیه است که مقدار زیادی است و نشان از پراکندگی زیاد در مقادیر زمان اجرا است و حاکی از عدم توازن بار بین منابع است.

در سناریوی دوم (اعمال سیاست پیشنهادی مبتنی بر الگوریتم ازدحام ذرات چند هدفه در کلاس VMselectionpolicy)؛ کلاس انتخاب ماشین مجازی از سیاست الگوریتم ارائه شده مبتنی بر الگوریتم ازدحام ذرات چند هدفه استفاده گردید تا ماشین‌های مجازی مناسب از بین ماشین‌های مجازی در دسترس انتخاب شوند. در کلاس تخصیص ماشین مجازی از سیاست انتخاب مبتنی بر آستانه و به تبعیت از سیاست اتخاذی در کلاس انتخاب ماشین مجازی استفاده گردید. در واقع در این روش، الگوریتم ارائه شده در کلاس VM selection policy اعمال گردید که در جدول ۳ قابل مشاهده است.

مقادیر ستون ۲ از جدول ۳، در مقایسه با مقادیر به دست آمده از سناریوی اول در ستون ۱ جدول ۳ تفاوت محسوسی دارد. در سناریوی دوم، میزان انرژی مصرفی حدود 50% کاهش پیدا کرده است، تعداد مهاجرت‌ها از ۵۷۰۰ مورد به ۵۶ مورد کاهش یافته که دلیل بر تخصیص بهینه است. موارد نقض توافقیانه به 0.00143% کاهش پیدا کرده است و درصد موارد نقض به دلیل مهاجرت صفر شده است. موارد خاموش شدن‌های میزبان‌ها نیز به ۴۶ مورد کاهش پیدا کرده که در مقایسه با مقدار آن در سناریوی اول، ۶۳۴ مورد بسیار پایین است. میانگین زمان اجرای کل از ۰,۰۰۵۳۳ به ۰,۰۰۱۵۷ سناریوی نخست و اطلاق زمان می‌باشد. انحراف معیار پایین سناریوی دوم نسبت به سناریوی اول کمتر است، و حاکی از توازن بار بیشتر است.

در سناریوی سوم (اعمال سیاست پیشنهادی مبتنی بر الگوریتم ازدحام ذرات چند هدفه در کلاس VMallocationpolicy) از سیاست انتخاب ماشین مجازی با کمترین مقدار بهره‌وری استفاده شد و در کلاس تخصیص ماشین مجازی از سیاست الگوریتم ارائه شده مبتنی بر الگوریتم ازدحام ذرات چند هدفه استفاده گردید تا بهترین تخصیص بر اساس الگوریتم پیشنهادی صورت پذیرد (جدول ۳).

در اولین ارزیابی، نتایج شبیه سازی برای چهار حالت مختلف ارائه می‌گردد. این چهار حالت شامل اتخاذ سیاست‌های مختلف در دو کلاس منشعب از کلاس‌های مادر ماشین مجازی allocation policy برای اعمال سیاست تخصیص و ماشین مجازی selection policy برای اعمال سیاست انتخاب می‌باشد که در جدول ۲، حالت‌های مختلف اعمال شده در شبیه سازی آورده شده است.

جدول ۲: حالت‌ها و سناریوهای مختلف شبیه سازی

سناریو	ماشین مجازی selection policy	ماشین مجازی allocation policy
۱	Minimum Utilization(mu)	Threshold Allocation (thr)
۲	Minimum Energy(me) with MOPSO	Threshold Allocation (thr)
۳	Minimum Utilization(mu)	MOPSO Allocation (mopso)
۴	Minimum Energy(me) with MOPSO	MOPSO Allocation (mopso)

در سناریوی اول (حالت استاندارد)؛ کلاس انتخاب ماشین مجازی، از سیاست انتخاب ماشین مجازی با کمترین مقدار بهره‌وری استفاده شده است، و در کلاس تخصیص ماشین مجازی از سیاست آستانه بهره‌وری با مقدار آستانه ۰,۸ استفاده گردید. در واقع این روش حالت نرمال شبیه ساز کلود سیم بوده و سیاست خاصی در کلاس تخصیص اعمال نشده بلکه از سیاست انتخاب مبتنی بر انتخاب بیکارترین ماشین‌های مجازی با یک آستانه مشخص استفاده شده است (جدول ۳).

در جدول ۳، میزان کل مصرف انرژی به کیلووات ساعت را نشان داده شده است که برابر با ۳۰,۵۹ می‌باشد، تعداد مهاجرت‌های ماشین‌های مجازی ۵۷۰۰ می‌باشد که رقمی بالاست و حاکی از تخصیص اولیه غیر کارآمد است. موارد نقض سطح توافقیانه خدمت نسبتاً زیاد است و موارد نقض به دلیل مهاجرت 0.45% می‌باشد که موجب عدم رضایت مشتریان خواهد شد. دلیل این نقض، ناشی از تخصیص‌های نامناسب و مهاجرت‌های زیاد می‌باشد. تعداد خاموش شدن‌های میزبان‌ها ۶۳۴ می‌باشد، که نسبتاً زیاد است. خاموش شدن‌های بی مورد و زیاد، باعث افزایش زمان آماده به کار شدن آنها و در نتیجه افزایش زمان کل اجرا و همچنین افزایش انرژی مصرفی اولیه در

¹ Minimum Utilization

استفاده از یک الگوریتم بهینه سازی چند هدفه برای تخصیص کارها در سیستم‌های مبتنی بر ابر با هدف کاهش انرژی مصرفی

9- Mean time before a host shutdown(sec)	1443.59	2278.74	1443.59	2278.74
10- StDev time before a host shutdown(sec)	1319.58	8257/02	1319.58	8257.02
11- Mean time before a VM migration	20.14	19.90	20.14	19.09
12- StDev time before a VM migration	8.01	7.86	8.01	7.86
13- Execution time - VM selection mean	0.00019	0.00024	0.00019	0.0019
14- Execution time - VM selection stDev	0.00039	0.00062	0.00084	0.00040
15- Execution time - host selection mean	0.00033	0.00038	0.00026	0.00033
16- Execution time - host selection stDev	0.00048	0.00075	0.00046	0.00048
17- Execution time - VM reallocation mean	0.00230	0.0003	0.00196	0.0003
18- Execution time - VM reallocation stDev	0.00303	0.00018	0.00246	0.00018
19- Execution time - total mean	0.00533	0.00157	0.00484	0.00187
20- Execution time - total stDev	0.00511	0.00276	0.00514	0.00370

نتایج سناریوی سوم که در ستون ۳ از جدول ۳ آمده است، نشان می‌دهد که در این پیاده‌سازی نتایج مشابه سناریوی اول است، به غیر از میانگین زمان کل اجرا که از ۰,۰۰۵۳۳ به ۰,۰۰۴۸۴ رسیده است و اندکی کاهش را نشان می‌دهد. اما انحراف معیار از ۰,۰۰۵۱۱ به ۰,۰۰۵۱۴ افزایش یافته است که نشان از توزان بار کمتر نسبت به سناریوی اول است.

در سناریوی چهارم (اعمال سیاست الگوریتم پیشنهادی در هر دو کلاس VM selection policy و VM allocation policy)، در هر دو کلاس انتخاب ماشین مجازی و تخصیص ماشین مجازی از سیاست الگوریتم پیشنهادی استفاده گردید (جدول ۳).

نتایج جدول ۳ ستون ۴ نشان می‌دهد که عملکرد سناریوی چهارم در کاهش مصرف انرژی، میزان نیاز به مهاجرت، درصد موارد نقض توافقنامه و خاموش شدن‌های میزبان مشابه سناریوی دوم است اما میانگین زمان کل اجرای آن که در سناریوی دوم ۰,۰۰۱۵۷ بود به ۰,۰۰۱۸۷ افزایش یافته است و انحراف معیار زمان کل اجرا نیز از ۰,۰۰۲۷۶ به ۰,۰۰۳۷۰ افزایش یافته است که حاکی از ضعف سناریوی چهارم نسبت به سناریوی دوم است. سناریوی چهارم در مقایسه با سناریوهای اول و سوم بهتر عمل کرده است و در رده دوم قرار دارد.

جدول ۴: نتایج سناریوهای شبیه سازی شده

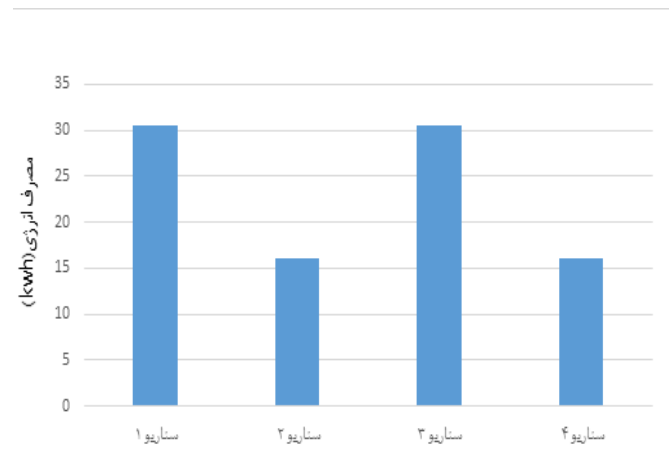
الف- مقادیر یکسان شبیه سازی				
Number of hosts: 50				
Number of VMs: 50				
Total simulation time: 86400.00 sec				
ب- نتایج سناریوهای اجرایی				
سناریو ۴	سناریو ۳	سناریو ۲	سناریو ۱	پارامترها
16.09	30.59	16.09	30.59	1-Energy consumption(kwh)
56	5700	56	5700	2- Number of VM migrations
0.00143	0.10356	0.00143	0.10356	3- SLA(%)
0.00	0.45	0.00	0.45	4- SLA perf degradation due to migration(%)
73.72	22.8	73.72	22.8	5- SLA time per active host(⊗%)
33.07	7.74	33.07	7.74	6- Overall SLA violation(%)
36.27	18.83	36.27	18.83	7- Average SLA violation(%)
46	634	46	634	8- Number of host shutdowns

۴-۳-۱- ارزیابی کارایی روش پیشنهادی در کاهش مصرف انرژی

از آنجا که یکی از اهداف اصلی روش پیشنهادی کاهش مصرف انرژی در ابر می‌باشد، در این بخش نقش الگوریتم پیشنهادی در کاهش مصرف انرژی مورد ارزیابی قرار می‌گیرد. برای این منظور نتایج حاصل از اعمال شرایط چهار حالت و سناریوی مختلف که در بخش قبل توضیح داده شد در شکل ۳ آورده شده است.

همانطور که در شکل مشخص است میزان کاهش مصرف انرژی در سناریوی دوم و چهارم که الگوریتم پیشنهادی در کلاس انتخاب ماشین مجازی اعمال شده است به طور کاملاً محسوسی کاهش پیدا کرده است. نتایج این نمودار نشان می‌دهد که در سناریوی چهارم که در مقایسه با سناریوی دوم، در کلاس تخصیص نیز الگوریتم پیشنهادی بصورت توأمان اعمال شده است، بهینگی صورت نپذیرفته

همانطور که در شکل ۴ مشاهده می‌شود، سناریوی دوم و چهارم تعداد مهاجرت کمتری را داشته‌اند و شرایط مطلوب‌تری دارند. تعداد مهاجرت کمتر حاکی از تخصیص اولیه بهتر است، زمانیکه تخصیص‌های اولیه منابع بهینه باشد. تطابق قدرت منابع با ماشین‌های مجازی باعث عدم نیاز به مهاجرت شده و در نتیجه باعث افزایش کارایی سیستم می‌گردد. این عمل هم در کاهش زمان اجرا و هم در کاهش مصرف انرژی به دلیل عدم اتلاف انرژی و در نتیجه منجر به تامین خواسته‌های هر دو طرف خدمت دهنده و خدمت گیرنده خواهد شد.



شکل ۳: نمودار کاهش مصرف انرژی برای سناریو ۴

۴-۳-۳-۳-۴- ارزیابی کارایی روش پیشنهادی در کاهش زمان کل اجرا و توازن بار

از پارامترهای مهم دیگر برای ارزیابی الگوریتم پیشنهادی، پارامتر زمان اجرای کل کارها است که نقش بسیار مهمی در رایانش ابری دارد، بدیهی است که کاهش زمان اجرا هم جزو شرایط کیفیت ارائه خدمات و درخواست‌های مشتریان است، هم باعث کاهش انرژی مصرفی و علاوه بر این باعث سود اجاره دهندگان خواهد بود. در این بخش از ارزیابی زمان کل اجرای چهار سناریوی پیاده‌سازی الگوریتم پیشنهادی مبتنی بر الگوریتم بهینه‌سازی ازدحام ذرات چند هدفه را با همدیگر در شکل ۵ آورده‌ایم.



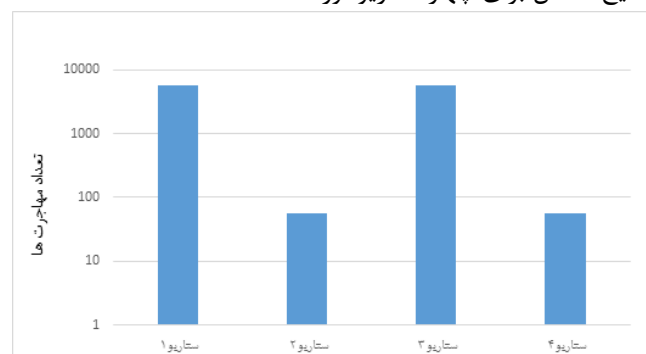
شکل ۵: نمودار مقایسه زمان اجرای کل ۴ سناریو پیاده‌سازی

همانطور که در شکل مشاهده می‌شود سناریوی دوم و چهارم وضعیت بهتری نسبت به دو سناریوی دیگر دارند، در بین این دو سناریو نیز سناریوی دوم زمان اجرای کل پائین‌تری دارد، و حاکی از عملکرد خوب آن است. در این نمودار همچنین مشخص است که میزان انحراف معیار سناریوی دوم نیز کمتر از بقیه است و نشان دهنده آن است که توازن بار نیز در پیاده‌سازی مذکور رعایت شده است.

و نشان دهنده این است که اعمال الگوریتم پیشنهادی در کلاس انتخاب که زودتر از کلاس تخصیص در کلود اجرا می‌شود و یک لایه قبل تر است نتایج خوبی ارائه می‌دهد و اعمال آن در کلاس تخصیص تغییری در بهینه‌گی ایجاد نمی‌کند. عدم بهینه‌گی کاهش مصرف انرژی در سناریوی سوم که الگوریتم پیشنهادی در آن در کلاس تخصیص اعمال شده است نیز دلیل دیگری بر عملکرد صحیح الگوریتم پیشنهادی در کلاس انتخاب است. کاهش تقریباً ۵۰٪ در مصرف انرژی، توسط الگوریتم پیشنهادی نسبت به حالت استاندارد تخصیص منابع که بر اساس ترتیب لیست عمل می‌نماید، حاکی از قدرت الگوریتم بهینه‌سازی ازدحام ذرات چند هدفه، در انتخاب بهینه منابع ماشین مجازی برای تخصیص است.

۴-۳-۲- ارزیابی کارایی روش پیشنهادی در کاهش نیاز به مهاجرت

شاخص دیگری که برای ارزیابی الگوریتم پیشنهادی مورد نظر قرار گرفت تعداد مهاجرت‌های مورد نیاز ماشین‌های مجازی است، بدیهی است هر چقدر که مقدار این مهاجرت‌ها نسبت به نتایج حاصله کمتر باشد بهتر است و حاکی از تخصیص‌های بهینه می‌باشد. در شکل ۴ نتایج حاصل برای چهار سناریو آورده شده است.



شکل ۴: تعداد مهاجرت‌های ماشین مجازی در سناریو ۴

SLA perf degradation due to migration: 0.00%
SLA time per active host: 74.95%
Overall SLA violation: 28.00%
Average SLA violation: 30.91%
Number of host shutdowns: 79
Mean time before a host shutdown: 2174.04 sec
StDev time before a host shutdown: 9968.25 sec
Mean time before a VM migration: 2.05 sec
StDev time before a VM migration: 0.00 sec
Execution time - VM selection mean: 0.00054 sec
Execution time - VM selection stDev: 0.00067 sec
Execution time - host selection mean: 0.00080 sec
Execution time - host selection stDev: 0.00064 sec
Execution time - VM reallocation mean: 0.00005 sec
Execution time - VM reallocation stDev: 0.00023 sec
Execution time - total mean: 0.00513 sec
Execution time - total stDev: 0.00995 sec

با توجه به نتایج جدول ۸، میزان کل انرژی مصرفی روش ارائه شده برابر است با ۶۲,۶۹ کیلو وات ساعت، تعداد ماشین‌های مجازی مهاجرت داده شده برابر است با ۱۸۶ ماشین مجازی، موارد نقض توافقنامه کیفیت ارائه خدمات ۰,۰۰۰۰۷ درصد است که مقدار بسیار پایینی است و حاکی از کارایی الگوریتم پیشنهادی در تضمین کیفیت ارائه خدمات است. با توجه به پایین بودن میزان نیاز به مهاجرت ماشین‌ها، موارد نقض توافقنامه به دلیل مهاجرت صفر شده است. تعداد خاموش شدن‌های میزبان‌ها ۷۹ مورد است، در نهایت میانگین زمان کل اجرای کارها ۰,۰۰۵۱۳ ثانیه و انحراف معیار آن ۰,۰۰۹۹۵ ثانیه است. در ادامه این نتایج با نتایج سایر روش‌های پیشین با شرایط شبیه‌سازی یکسان مورد بررسی قرار خواهد گرفت. در [۳۳]، که مروری بر کارهای گذشته در این حوزه داشته است، مقایسه الگوریتم‌های ان پی آی، دی وی اف اس، اس تی و MM را در شرایط یکسانی که در جدول (۳) آمده است، صورت پذیرفت. در جدول (۴-۵) نتایج نهایی این آزمایش‌ها در مقایسه با روش پیشنهادی ارائه شده است. با توجه به جدول (۳) میزان انرژی مصرفی توسط الگوریتم پیشنهادی برابر است با ۶۲,۶۹ کیلو وات

۴-۳-۴ ارزیابی کارایی روش پیشنهادی در مقایسه با روش‌های

پیشین

در ارزیابی دیگری به عنوان ارزیابی نهایی، کارایی الگوریتم پیشنهادی در مقایسه با الگوریتم‌های قبلی مورد بررسی قرار گرفت، برای این منظور ابتدا شرایط و تنظیمات یکسان بصورت جدول ۴ اعمال گردید [۳۵]. در این مرحله از ارزیابی سناریوی دوم که بهترین سناریو و پیاده‌سازی از الگوریتم پیشنهادی بود اجرا گردید.

جدول ۴- تنظیمات شبیه سازی یکسان برای مقایسه با الگوریتم‌های

پیشین

مقدار دهی اولیه و تنظیمات شبیه‌سازی برای مقایسه روش

پیشنهادی با روش‌های دیگر

```

NUMBER_OF_VMS = 290;
NUMBER_OF_HOSTS = 100;
VM_TYPES = 4;
VM_MIPS = { 250, 500, 750, 1000 };
VM_PES = { 1, 1, 1, 1 };
VM_RAM = { 128, 128, 128, 128 };
VM_BW = 100000; // 100 Mbit/s
VM_SIZE = 1000; // 1 GB
HOST_TYPES = 3;
HOST_MIPS = { 1000, 2000, 3000 };
HOST_PES = { 1, 1, 1 };
HOST_RAM = { 8192, 8192, 8192 };
HOST_BW = 1000000; // 1 Gbit/s
HOST_STORAGE = 1000000000; // 1 TB
HOST_POWER =
PowerModelSpecPowerHpProLiantM110G4Xeon3040()
    
```

در این مرحله از شبیه‌سازی، تعداد ماشین‌های مجازی: ۲۹۰، تعداد میزبان: ۱۰۰، انواع ماشین مجازی: ۴ نوع با مشخصات پردازشی، تعداد هسته، حافظه رم، پهنای باند و اندازه حافظه جانبی مندرج در جدول و سه نوع میزبان با مشخصات ذکر شده در جدول ۴ ایجاد گردید. خروجی شبیه‌سازی در جدول ۵ آورده شده است.

جدول ۵- خروجی شبیه‌سازی با شرایط مقایسه‌ای

نتایج شبیه سازی

Experiment name: thr_me_0.8
Number of hosts: 100
Number of VMs: 290
Total simulation time: 86400.00 sec
Energy consumption: 62.69 kWh
Number of VM migrations: 186
SLA: 0.00007%

³ ST

¹ NPA

² DVFS

۵- نتیجه گیری و کارهای آتی

نتایج به دست آمده از شبیه‌سازی نشان داد که استفاده از الگوریتم فرااکتشافی بهینه‌سازی ازدحام ذرات برای تخصیص منابع می‌تواند به طور قابل توجهی میزان صرفه‌جویی در مصرف انرژی را بهبود بخشد. اهداف تعریف شده در این پژوهش کاهش مصرف انرژی و کاهش زمان اجرای کل بود که به آنها دست یافته شد. کاهش مدت زمان اجرا به ۰,۰۰۵۱۳ ثانیه و کاهش میانگین انرژی مصرفی به ۰,۶۲۶ کیلووات ساعت در مقایسه با بهترین روش قبلی که ۱,۱۴۲ کیلووات ساعت بود، همچنین کاهش میزان نیاز به مهاجرت و توازن بار از دستاوردهای دیگر این پژوهش بود. طی این تحقیق مشخص شد که از قابلیت‌های ویژه الگوریتم‌های فرااکتشافی به ویژه الگوریتم بهینه‌سازی ازدحام ذرات و بهینه‌سازی چندهدفه می‌توان به نتایج بهتری بدست آورد. تعریف توابع جدید در بهینه‌سازی چندهدفه، برای دستیابی به حصول سایر شرایط و اهداف مورد نیاز می‌تواند موثر باشد.

به عنوان کارهای پژوهشی آینده استفاده از روش پیشنهادی برای زمانبندی وظایف در ماشین‌های مجازی، اعمال پارامترهای سودمند دیگر برای تعیین تابع برازندگی، استفاده از روش ارایه شده برای سایر اهداف سیستم‌های ابری، همانند کاهش هزینه مشتری و افزایش سود ارایه دهندگان خدمت، استفاده از روش‌های دیگر فرا ابتکاری و روش‌های ترکیبی مشابه جهت تخصیص منابع، آزمایش روش پیشنهادی بر روی محیط اتحادیه ابر برای صرفه‌جویی بیشتر در مصرف انرژی و نتایج بهتر می‌توانند در جهت توسعه روش ارائه شده در این پژوهش انجام شوند.

مراجع

1. Jafari Navimipour, N. and B. Zareie, *A model for assessing the impact of e-learning systems on employees' satisfaction*. Computers in Human Behavior, 2015. **53**: p. 475-485.
2. Chiregi, M. and N.J. Navimipour, *A new method for trust and reputation evaluation in the cloud environments using the recommendations of opinion leaders' entities and removing the effect of troll entities*. Computers in Human Behavior, 2016. **60**: p. 280-292.
3. Souri, A. and N. Jafari Navimipour, *Behavioral modeling and formal verification of a resource discovery approach in Grid computing*. Expert Systems with Applications, 2014. **41**(8): p. 3831-3849.
4. Ding, Y., et al., *Energy efficient scheduling of virtual machines in cloud with deadline constraint*. Future Generation Computer Systems, 2015. **50**: p. 62-74.

ساعت و تعداد میزبان برابر است با ۱۰۰ بنابراین میانگین مصرف انرژی برای هر میزبان برابر است با ۰,۶۲۶ کیلو وات ساعت و در مقایسه با روش‌های قبلی کمترین میزان مصرف انرژی را دارد و از بهترین نتایج به دست آمده در روش‌های ان پی آ، دی وی اف اس، اس تی با دو سناریوی پیاده‌سازی و ام ام با سه سناریوی پیاده‌سازی به طور محسوسی بهینه است.

جدول ۶- نتایج نهایی مقایسه روش پیشنهادی با روش‌های قبلی

میانگین SLA	مهاجرت ماشین مجازی	SLA violation	انرژی (کیلو وات ساعت)	الگوریتم
-	-	-	9.151	ان پی آ
-	-	-	4.402	دی وی اف اس
81.8%	35226	5.4%	2.034	اس تی ۵۰٪
87.9%	34519	7.4%	1.622	اس تی ۶۰٪
56.9%	3360	1.1%	1.476	ام ام ۳۰٪-۷۰٪
65.7%	3241	2.8%	1.266	ام ام ۴۰٪-۸۰٪
76.1%	3121	6.7%	1.142	ام ام ۵۰٪-۹۰٪
30.91%	186	28%	0.626	روش پیشنهادی

ستون سوم نشان دهنده درصد نقض شرایط توافقنامه کیفیت ارائه خدمات سراسری است که الگوریتم پیشنهادی ضعیف تر از بقیه است. در ستون مهاجرت ماشین مجازی مشاهده می‌شود که تعداد مهاجرت‌ها در روش پیشنهادی به طور قابل مقایسه‌ای کاهش یافته است و این حاکی از تخصیص بهینه می‌باشد که نتایج‌اش در کاهش مصرف انرژی نیز ظهور پیدا کرده است. ستون آخر این ارزیابی مربوط است به میانگین توافقنامه سطح خدمت که اعداد مربوط به این ستون نشان دهنده درصد موارد رعایت توافقنامه کیفیت ارایه خدمات است، به عبارتی اعداد این ستون درصد تخصیص‌های نامناسب را به لحاظ سرعت اجرای دستور العمل ماشین فیزیکی نسبت به سرعت اجرای دستور العمل درخواستی نشان می‌دهد، هر چقدر این مقدار کم باشد حاکی از بهینگی روش مذکور است. مشاهده می‌شود که روش پیشنهادی کمترین مقدار (۳۰,۹۱٪) را در بین روش‌های ارائه شده دارد.

- Clouds. European Journal of Operational Research, 2014. **236**(1): p. 326-339.
17. Yousafzai, A., et al., *Cloud resource allocation schemes: review, taxonomy, and opportunities*. Knowledge and Information Systems, 2017. **50**(2): p. 347-381.
 18. Hameed, A., et al., *A survey and taxonomy on energy efficient resource allocation techniques for cloud computing systems*. Computing, 2016. **98**(7): p. 751-774.
 19. Devarasetty, P. and S. Reddy, *Genetic algorithm for quality of service based resource allocation in cloud computing*. Evolutionary Intelligence, 2019.
 20. Wei, G., et al., *A game-theoretic method of fair resource allocation for cloud computing services*. The Journal of Supercomputing, 2010. **54**(2): p. 252-269.
 21. Shu, W., W. Wang, and Y. Wang, *A novel energy-efficient resource allocation algorithm based on immune clonal optimization for green cloud computing*. EURASIP Journal on Wireless Communications and Networking, 2014. **2014**(1): p. 64.
 22. Xiao, Z., W. Song, and Q. Chen, *Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment*. IEEE Transactions on Parallel and Distributed Systems, 2013. **24**(6): p. 1107-1117.
 23. Li, C., Y.C. Liu, and X. Yan, *Optimization-based resource allocation for software as a service application in cloud computing*. Journal of Scheduling, 2017. **20**(1): p. 103-113.
 24. Son, S., G. Jung, and S.C. Jun, *An SLA-based cloud computing that facilitates resource allocation in the distributed data centers of a cloud provider*. The Journal of Supercomputing, 2013. **64**(2): p. 606-637.
 25. Jala, J. and K.R.H. Rao. *QoS-Based Technique for Dynamic Resource Allocation in Cloud Services*. 2019. Singapore: Springer Singapore.
 26. Usman, M.J., et al., *Energy-efficient Virtual Machine Allocation Technique Using Flower Pollination Algorithm in Cloud Datacenter: A Panacea to Green Computing*. Journal of Bionic Engineering, 2019. **16**(2): p. 354-366.
 27. Lau, H.Y.K. and Y. Zhao, *Integrated scheduling of handling equipment at automated container terminals*. International Journal of Production Economics, 2008. **112**(2): p. 665-682.
 28. Holland, J.H., *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. 1992: MIT press.
 29. Coello, C.A.C., *Evolutionary multi-objective optimization*. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2011. **1**(5): p. 444-447.
 30. Abraham, A. and L. Jain, *Evolutionary multi-objective optimization*. Evolutionary Multi-objective Optimization, 2005: p. 1-6.
 5. Jafari Navimipour, N., et al., *Expert Cloud: A Cloud-based framework to share the knowledge and skills of human resources*. Computers in Human Behavior, 2015. **46**: p. 57-74.
 6. Xiao, P., Z.-G. Hu, and Y.-P. Zhang, *An Energy-Aware Heuristic Scheduling for Data-Intensive Workflows in Virtualized Datacenters*. Journal of Computer Science and Technology, 2013. **28**(6): p. 948-961.
 7. Song, J., et al., *Study on energy-consumption regularities of cloud computing systems by a novel evaluation model*. Computing, 2013. **95**(4): p. 269-287.
 8. L.D, D.B. and P. Venkata Krishna, *Honey bee behavior inspired load balancing of tasks in cloud computing environments*. Applied Soft Computing, 2013. **13**(5): p. 2292-2303.
 9. Alami Milani, B. and N. Jafari Navimipour, *A comprehensive review of the data replication techniques in the cloud environments: Major trends and future directions*. Journal of Network and Computer Applications, 2016. **64**: p. 229-238.
 10. Lombardi, F. and R. Di Pietro, *Secure virtualization for cloud computing*. Journal of Network and Computer Applications, 2011. **34**(4): p. 1113-1122.
 11. Jafari Navimipour, N., et al., *Behavioral modeling and automated verification of a Cloud-based framework to share the knowledge and skills of human resources*. Computers in Industry, 2015. **68**: p. 65-77.
 12. Almorsy, M., J. Grundy, and A.S. Ibrahim, *Adaptable, model-driven security engineering for SaaS cloud-based applications*. Automated Software Engineering, 2014. **21**(2): p. 187-224.
 13. Wu, L., S. Kumar Garg, and R. Buyya, *SLA-based admission control for a Software-as-a-Service provider in Cloud computing environments*. Journal of Computer and System Sciences, 2012. **78**(5): p. 1280-1299.
 14. Ramamoorthy, S. and S. Rajalakshmi, *A Preventive Method for Host Level Security in Cloud Infrastructure*, in *Proceedings of the 3rd International Symposium on Big Data and Cloud Computing Challenges (ISBCC - 16')*, V. Vijayakumar and V. Neelanarayanan, Editors. 2016, Springer International Publishing: Cham. p. 3-12.
 15. Dam, S., et al., *An Ant Colony Based Load Balancing Strategy in Cloud Computing*, in *Advanced Computing, Networking and Informatics- Volume 2: Wireless Networks and Security Proceedings of the Second International Conference on Advanced Computing, Networking and Informatics (ICACNI-2014)*, M. Kumar Kundu, et al., Editors. 2014, Springer International Publishing: Cham. p. 403-413.
 16. Anselmi, J., D. Ardagna, and M. Passacantando, *Generalized Nash equilibria for SaaS/PaaS*

31. Li, H., K. Deb, and Q. Zhang, *Variable-length Pareto Optimization via Decomposition-based Evolutionary Multiobjective Algorithm*. IEEE Transactions on Evolutionary Computation, 2019.
32. Sheikholeslami, F. and N.J. Navimipour, *Service allocation in the cloud environments using multi-objective particle swarm optimization algorithm based on crowding distance*. Swarm and Evolutionary Computation, 2017. **35**: p. 53-64.
33. Clerc, M., *Standard particle swarm optimisation*. 2012.
34. Calheiros, R.N., et al., *CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms*. Software: Practice and Experience, 2011. **41**(1): p. 23-50.
35. Beloglazov, A., J. Abawajy, and R. Buyya, *Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing*. Future Generation Computer Systems, 2012. **28**(5): p. 755-768.
36. Mandal, R., Mondal, M.K., Banerjee, S. et al. An approach toward design and development of an energy-aware VM selection policy with improved SLA violation in the domain of green cloud computing. *J Supercomput* (2020). <https://doi.org/10.1007/s11227-020-03165-6>

Using a multi-objective optimization algorithm to allocate tasks in cloud-based systems to reduce energy consumption

Due to the increasing demand for business in the web environment, the variety and number of services have increased. As a result, large-scale computing data centers have been created that increase costs and electrical power. Inadequate and inefficient cooling systems not only overheat the resources and reduce the devices' service life but also produce carbon, which plays an important role in weather conditions.

Therefore, it is necessary to reduce the total energy consumption of these systems with appropriate methods. In this research, an effective method of energy resource management in virtualized cloud data centers is presented. In addition to reducing energy consumption and operating costs, it has also increased the quality of services. This study presents a resource allocation strategy in cloud systems to reduce energy and implementation costs and examine its application in the cloud computing environment. The simulation results showed that the proposed method could reduce the average energy consumption up to 0.626 kWh compared to NPA, DVFS, ST, and MM methods. Also, the need for migration and SLA violations are decreased.

Keywords: Cloud Computing, Energy Consumption, Job Allocation, Multi-Objective Optimization Algorithm.