

## مروری بر روش‌های تخصیص منابع در رایانش ابری

هادی مویی امام قیسی<sup>۱</sup>، محمد داورپناه جزی<sup>۲</sup>

<sup>۱</sup> دانشجوی کارشناسی ارشد، گروه کامپیوتر، موسسه آموزش عالی صنعتی فولاد، فولادشهر.

<sup>۲</sup> استادیار، گروه کامپیوتر، موسسه آموزش عالی صنعتی فولاد، فولادشهر.

نام و نشانی ایمیل نویسنده مسئول:

هادی مویی امام قیسی

[hadimoei@gmail.com](mailto:hadimoei@gmail.com)

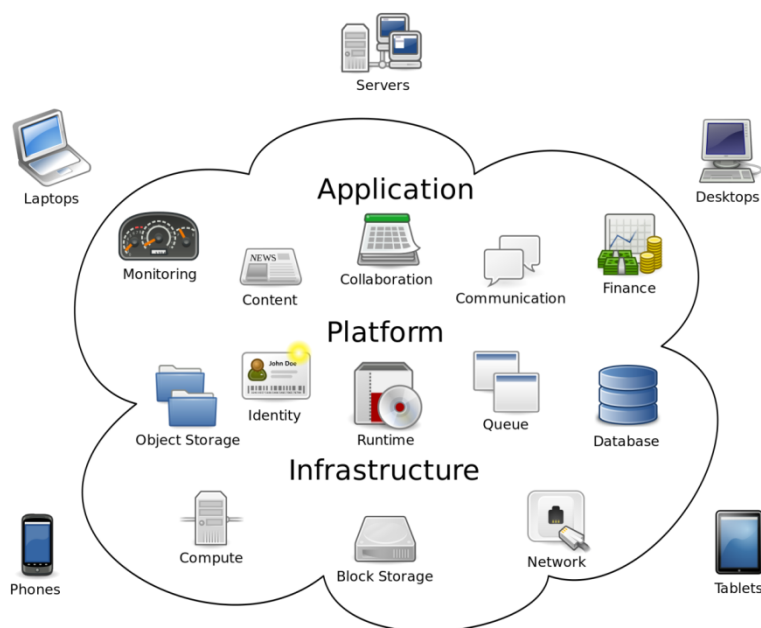
### چکیده

موضوع مدیریت و تخصیص منابع در محیط‌های رایانش ابری، باتوجه به گستردگی مقیاس و همچنین فناوری‌های مدرن پیاده‌سازی شده، به عنوان یک مسئله پیچیده و سخت محسوب می‌شود. مسائلی مانند: ناهمگونی انواع منابع، وابستگی منابع به یکدیگر، پویایی محیط، تنوع حجم کار و همچنین طیف وسیعی از اهداف مدیریتی که ارائه دهندگان خدمات ابر، برای ارائه سرویس در این محیط دارند. بنابراین سالانه تحقیقات وسیعی در محیط‌های دانشگاهی و صنعتی برای حل این مسایل انجام می‌شود و یا در حال انجام هستند. در این مقاله ابتدا توضیح محیط رایانش ابری و موضوعات مرتبط با آن بیان شده است. همچنین باتوجه به موضوع تخصیص منابع، انواع منابع و ذینفعان و همچنین اهداف تخصیص منابع توضیح داده شده است. با بررسی‌های انجام شده چالش‌هایی مانند: عدم وجود یک مدیریت جامع برای منابع در محیط ابر، چگونگی پیش‌بینی فرایند تخصیص منابع، روش‌های بهینه تخصیص منابع به منظور کاهش مصرف انرژی و کاهش زمان دسترسی به منابع و همچنین پیاده‌سازی روش‌های پویا تخصیص منابع در محیط‌های سیار ابری، مشخص شده اند. در نهایت باتوجه به چالش‌های موجود پیشنهادهایی برای بهبود فرایند تخصیص منابع در محیط رایانش ابری ارائه شده است.

**واژگان کلیدی:** رایانش ابری، مدیریت منابع، تخصیص منابع، مجازی‌سازی، مهاجرت ماشین مجازی.

## مقدمه

در دهه‌های اخیر مدل‌های محاسباتی پیشرفت فراوانی داشته‌اند، همچنین تکنیک‌های جدید ارائه خدمات از طریق شبکه اینترنت، برای استفاده عموم کاربران فراهم شده است. برنامه‌ها نیاز کمتری به محاسبات محلی و یا ذخیره‌سازی محلی دارند. با پیشرفت فناوری اطلاعات نیاز به انجام کارهای محاسباتی در همه‌جا و همه‌زمان به وجود آمده است. همچنین نیاز است که افراد بتوانند کارهای محاسباتی سنگین خود را بدون داشتن سخت‌افزارها و نرم‌افزارهای گران، از طریق خدمات قابل دسترسی انجام دهند. رایانش ابری آخرین پاسخ فناوری به این نیازها بوده است. از آنجا که اکنون این فناوری دوران ابتدایی خود را می‌گذراند، هنوز تعریف استاندارد علمی که مورد قبول عام باشد برای آن ارائه نشده است اما بیشتر صاحب‌نظران بر روی قسمت‌هایی از تعریف این پدیده هم‌رای هستند. موسسه ملی فناوری و استانداردها (NIST) رایانش ابری را اینگونه تعریف میکند: "رایانش ابری مدلی است برای فراهم کردن دسترسی آسان بر اساس تقاضای کاربر از طریق شبکه به مجموعه‌ای از منابع رایانشی قابل تغییر و پیکربندی (مثل: شبکه‌ها، سرورها، فضای ذخیره‌سازی، برنامه‌های کاربردی و سرویس‌ها) که این دسترسی بتواند با کمترین نیاز به مدیریت منابع و یا نیاز به دخالت مستقیم فراهم‌کننده سرویس به سرعت فراهم شده یا آزاد گردد" [۱]. رایانش ابری حجم قابل توجهی از منابع که در مراکز داده با مقیاس بزرگ قرار گرفته‌اند را بین برنامه‌های کاربردی، از طریق اینترنت به اشتراک می‌گذارد. رایانش ابری به عنوان یک فناوری محبوب در حال گسترش و پیشرفت است و تمایل آرایه و استفاده از خدمات در محیط‌های ابری در حال افزایش است. در محیط‌های رایانش ابری سرویس‌ها یا بروی زیرساخت‌های استیجاری که فراهم کنندگان ابر در اختیار افراد ثالث (کاربران ابر) قرار می‌دهند، برای استفاده عموم آرایه می‌شوند و یا به عنوان زیرساخت‌های خصوصی برای استفاده سازمان‌های خاص آرایه می‌شوند. مورد اول به معماری ابر عمومی اشاره دارد و مورد دوم معماری ابر خصوصی، همچنین سازمان‌ها می‌توانند منابع خصوصی خود را برای استفاده عموم به اشتراک بگذارند که به آن ابر ترکیبی می‌گویند [۲]. همچنین منابع می‌توانند به صورت غیر متمرکز توسط سازمان‌ها و مراکز داده مختلف در اختیار عموم قرار گیرند. در شکل ۱ معماری مفهومی محیط رایانش ابری نشان داده شده است.



شکل ۱ معماری مفهومی رایانش ابری [۱]

## ۲- منابع در رایانش ابری

در حال حاضر رایانش ابری، یک حوزه تجاری است. بنابراین قیمت و کیفیت سرویس از موارد مهم در این حوزه می‌باشد. ارائه‌کننده محاسبات ابری، منابع و نرم‌افزار را به کلاینت‌هایی در شرایط توافق سطح سرویس (SLA<sup>1</sup>) تحویل می‌دهند SLA مقداری بین کاربر و ارائه‌کننده ابر است که شامل منابع مورد نیاز کاربران، محدودیت زمان سرویس و هزینه سرویس می‌باشد که مزایای خیلی

<sup>1</sup> Service Level Agreements

مهمی برای یک سرمایه گذار تجاری دارد. ارائه کننده سرویس ابر خواهان سود از سرویس محاسبات ابری می‌باشد همچنین کاربران نمی‌خواهند پول زیادی بپردازند. کاربران محاسبات ابری، خدمات کیفی خوبی از ارائه کنندگان سرویس دریافت می‌کنند که هزینه خدمات براساس فرایند تخصیص منابع در محیط سرویس خاص می‌باشد. ارائه کننده خدمات ابر باید منابعی را برای کلاینت‌های خاص به یک روش خاص اختصاص دهد. مدل‌های تخصیص منابع مختلفی وجود دارد که در حوزه محاسبات ابری استفاده می‌شود. هر یک از این مدل‌ها از روش خاص و الگوریتم‌هایی برای رسیدن به این هدف استفاده می‌کنند. در محیط‌های ابری مصرف کنندگان درگیر سرمایه گذاری زیاد در زیرساخت‌های فناوری اطلاعات و مسائل پیچیده مربوط به ساخت و نگهداری آنها نمی‌شوند. در این مدل کاربران بدون توجه به اینکه خدمات در کجا میزبانی می‌شوند به خدمات مورد نیاز خود دسترسی دارند. رایانش ابری بر اساس مدل «پرداخت کن - استفاده کن» میزبانی برنامه‌های کاربردی، تجاری و علمی و غیره را بر عهده می‌گیرد. مراکز داده میزبانی کننده برنامه‌های کاربردی، مقدار بسیار زیادی از انرژی را مصرف می‌کنند. نگهداری مراکز داده بزرگ نیازمند مصرف انرژی زیاد می‌باشد. بر آورده شده که هزینه نگهداری مراکز داده از هزینه سرمایه گذاری اصلی آنها بیشتر شده است در این صورت نگهداری مراکز داده با این وضع حتی می‌تواند مضر هم باشد و متاسفانه دیده شده که بدون حد و مرز این کار ادامه می‌یابد. این در حالی است که فناوری اطلاعات کاملاً وارد زندگی مردم شده و باید مقرون به صرفه باشد. از دیدگاه ارائه دهندگان سرویس چگونگی به حداکثر رساندن سود با توجه به هزینه زیاد انرژی نوعی مشکل محسوب می‌گردد. چون سود حاصله برابر است با درآمدهای هزینه، که درآمد خود باید از راه تأمین خدمات به مشتریان حاصل شود. به طور شهودی بالاترین سود و سطح عملکرد عاید آنها می‌شود که بیشترین هزینه انرژی را داده اند از این رو لازم است برای مراکز داده تدبیری اندیشیده شود که در آن سود از هزینه سبقت بگیرد. یک راه برای کاهش هزینه و افزایش سود کاهش مصرف انرژی است. افزایش هزینه انرژی یک تهدید بسیار بزرگ و بالقوه برای افزایش هزینه مالکیت می‌باشد. در نتیجه بازگشت سرمایه را برای ارائه دهندگان ابر کاهش می‌دهد.

## ۲-۱ مدیریت و تخصیص منابع

یکی از مهمترین وظایف فراهم کنندگان ابر مدیریت و تخصیص منابع است. مصرف کنندگان ابر درخواست سرویس را از هر جایی از جهان به ابر می‌فرستند. لازم به ذکر است که بین مصرف کنندگان ابر و کاربران سرویس‌های مستقر شده تفاوت‌هایی وجود دارد. مصرف کنندگان می‌توانند یک شرکت استقرار برنامه‌های وب باشند که جریان‌های کاری متفاوت را مطابق با تعداد کاربرانی که به آن دسترسی دارند ارائه می‌دهند. ارائه دهندگان سرویس‌های ابری باید به مشتریان اطمینان دهند که نیازهای آنها به طور کامل برآورده خواهد شد. تا همین اواخر به دست آوردن کارایی بالا، تنها نگرانی در هنگام تخصیص منابع بود. بدون آنکه هزینه‌های در حال افزایش انرژی در نظر گرفته شود. در محاسبه ابر، تخصیص منابع (RA)، فرایند اختصاص منابع موجود به برنامه‌های کاربردی مورد نیاز بر روی اینترنت است. تخصیص منابع، خدمات را محروم می‌سازد در صورتیکه تخصیص به طور دقیق مدیریت نشود. تأمین کردن منابع، این مسئله را بوسیله امکان مدیریت منابع توسط تأمین کنندگان خدمات برای هر یک از مدل‌ها حل می‌کند.

استراتژی تخصیص منابع (RAS) در مورد تلفیق فعالیت‌های تأمین کننده ابر برای بهره برداری و اختصاص منابع نادر در محدوده محیط ابر برای برآوردن نیازهای برنامه کاربردی ابر می‌باشد. این استراتژی نیازمند نوع و میزان منابع مورد نیاز به وسیله هر برنامه کاربردی به منظور تکمیل کار کاربر می‌باشد. ترتیب و زمان تخصیص منابع هم، برای یک RAS بهینه یک ورودی است. یک RAS بهینه باید از معیارهای زیر به شرح زیر اجتناب کند.

**رقابت منابع:** رقابت منابع هنگامی بوجود می‌آید که دو برنامه کاربردی تلاش می‌کنند تا به منابع یکسان در یک زمان دست پیدا کنند.

**کمبود منابع:** کمبود منابع هنگامی بوجود می‌آید که منابع محدود وجود دارند و درخواست برای منابع بالاست.

**تفکیک منابع:** تفکیک منابع هنگامی بوجود می‌آید که منابع جدا می‌شوند. منابع کافی وجود دارند اما نمی‌توان آن را به برنامه کاربردی مورد نیاز به علت تفکیک به واحدهای کوچک تخصیص داد.

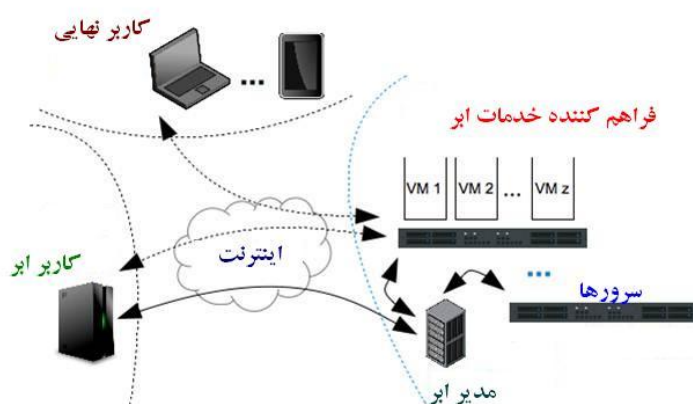
**تأمین بیش از حد:** تأمین بیش از حد هنگامی روی می‌دهد که برنامه کاربردی به منابع اضافی بیشتر از آنچه در خواست شده است دست پیدا میکند.

**تأمین کمتر از حد:** تأمین کمتر از حد منابع هنگامی اتفاق می‌افتد که برنامه کاربردی، منابع بسیار کمتری از آنچه که تقاضا شده است اختصاص می‌یابد.

مصرف انرژی در محیط‌های ابری از دو دیدگاه مورد بررسی قرار می‌گیرد، دیدگاه اول مدیریت استاتیک توان، که بیشتر مربوط به تجهیزات سخت افزاری می‌باشد و دیدگاه دوم مدیریت پویای مصرف انرژی می‌باشد. در این فصل به موضوع مدیریت منابع در رایانش ابری می‌پردازیم. در ابتدا بازیگران و کسانی که در محیط رایانش ابری نقش دارند معرفی می‌کنیم، سپس اهداف تخصیص منابع را بیان می‌کنیم، همچنین انواع منابعی که محیط رایانش ابری می‌تواند دربر داشته باشد و همچنین فناوری‌های موجود آنها را بیان می‌کنیم و در نهایت چارچوب اصلی از منابع در محیط‌های ابری را ارائه می‌دهیم که در بخش‌های بعدی به بررسی روش‌ها و الگوریتم‌های آنها می‌پردازیم.

## ۲-۴ بازیگران و مدل تأمین منابع در ابر

همانگونه که در بخش قبل اشاره شد، منابع در قالب سرویس‌هایی با عنوان‌های SaaS و PaaS و IaaS ارائه می‌گردند. اما ارائه سرویس در قالب روش‌های گفته شده همیشه مشکلات و دشواری‌هایی دارد. در تحقیقات جدید بیشتر بر روی ارائه زیر ساخت به عنوان سرویس تمرکز کرده‌اند. به طور خاص یک زیر ساخت استاندارد باید به گونه‌ای ارائه گردد که در محیط‌های ابری خصوصی و عمومی به راحتی قابل دسترسی باشد. در شکل ۲ قسمت بازیگران و ذینفعان در محیط رایانش ابری مشخص شده است.



شکل ۲: بازیگران و ذینفعان در محیط ابر [۱]

**فراهم کننده ابر<sup>۳</sup>:** فراهم کننده خدمات ابر، مدیریت مجموعه‌هایی از سخت افزارهای مراکز داده، منابع نرم‌افزاری، خدمات ارائه شده به عنوان زیرساخت و به عنوان سکو را در محیط‌های ابر عمومی و خصوصی و همچنین مدیریت تمامی منابع مورد نیاز برای ارائه برنامه کاربردی در قالب سرویس را برعهده دارد. فراهم کننده خدمات ابر مسئولیت تخصیص منابع را مطابق بر توافق سطح سرویس (SLA) بین کاربران نهایی و یا کاربران ابر به منظور دستیابی به اهداف مدیریتی منابع انجام می‌دهد.

**کاربر ابر<sup>۴</sup>:** کاربران ابر بیشتر در محیط ابر عمومی وظیفه میزبانی برنامه‌های کاربردی را برای ارائه به کاربران نهایی برعهده دارد. ابتدا قسمتی از منابع را از فراهم کننده خدمات ابر به صورت استیجاری در اختیار می‌گیرند و سپس بر مبنای توافق سطح سرویس بین کاربران ابر و مشتریان (کاربران نهایی) برنامه‌های کاربردی و سایر سرویس‌ها را به مشتریان ارائه می‌کنند. کاربران ابر بیشتر سرویس‌های خود را طوری ارائه می‌کنند که با کمترین هزینه بیشترین سود را کسب نمایند.

**کاربران نهایی<sup>۵</sup>:** تولید کننده کارهایی (اجرای برنامه‌های کاربردی) که توسط منابع ابر پردازش می‌شوند. کاربر نهایی نقش مستقیم در مدیریت منابع ندارد، اما رفتار او می‌تواند بر روی سیاست‌های تخصیص منابع تاثیرگذار باشد و هم می‌تواند تحت تاثیر این سیاست‌های اعمال شده توسط فراهم کننده خدمات ابر قرار گیرد.

## ۲-۴ انواع منابع

در این قسمت انواع اصلی منابع که به عنوان سرویس در محیط رایانش ابری وجود دارند توضیح داده می‌شوند.

**منابع پردازشی:** منابع پردازشی مجموعه‌ای از ماشین‌های فیزیکی (PM<sup>۶</sup>) هستند که هرکدام شامل یک یا چند پردازنده، حافظه، رابط شبکه، ورودی خروجی محلی هستند. همه‌ی این اجزا به عنوان ظرفیت محاسباتی یک محیط رایانش ابری ارائه می‌گردند. معمولاً

<sup>3</sup> Cloud Provider

<sup>4</sup> Cloud User

<sup>5</sup> End User

<sup>6</sup> Physical Machines

PMها را به صورت مجازی پیاده سازی می‌کنند و ماشین‌های مجازی تولید می‌کنند. این ماشین‌های مجازی از یکدیگر مستقل عمل می‌کنند و هرکدام ممکن است سیستم‌عامل و برنامه‌های کاربردی مختلفی داشته باشند. در حوزه منابع پردازشی بیشتر محققان PMها و VMها را برای افزایش ظرفیت پردازشی و حافظه‌ای محدود می‌کنند.

**منابع شبکه‌ای:** منابع پردازشی در مراکز داده درون قفسه‌هایی سازماندهی می‌شوند که معمولاً برای میزبانی از گروه‌های زیادی از برنامه‌های کاربردی که نیاز به منبع دارند آماده هستند. مشکلی که وجود دارد ارتباط با منابع پردازشی است. ماشین‌های فیزیکی به پهنای باند با حجم بالا در حد گیگابیت و بالاتر نیاز دارند تا بتوانند همه درخواست‌های تخصیص منابع که توسط برنامه‌های کاربردی گوناگون ارسال می‌شوند را پاسخ دهند. علاوه بر برنامه‌های کاربردی، برنامه‌هایی که دارای پردازش موازی هستند نیاز به پهنای باند بالا برای انتقال اطلاعات دارند. با وجود این موارد برنامه‌ها باید بدون ایجاد سربار اضافی به مراکز داده متصل شوند و پروتکل شبکه کارایی خودش را حفظ کند. در این جا دو جنبه مهم وجود دارد:

جنبه اول مربوط به توپولوژی طراحی شده برای شبکه است که تاثیر قابل توجهی بر کارایی و تحمل پذیری خطا در محیط ابری دارد. توپولوژی‌های فعلی شبکه در مراکز داده توپولوژی‌های سلسله مراتبی هستند، مانند توپولوژی درختی که در شبکه‌های تلفن همراه اولیه استفاده می‌شد. البته توپولوژی‌های جدیدتری از جمله *fat trees* و *hyper-cubes* و *randomized small-world topologies* نیز ظهور پیدا کرده‌اند. در تمامی این توپولوژی‌ها هدف این است که تعداد پورت ورودی و خروجی شبکه با پهنای باند شبکه به صورت خطی افزایش یابند [۱۰].

جنبه دوم بطور مستقیم با مدیریت منابع گره خورده است، که چگونه زمان تاخیر و پهنای باند شبکه در مراکز داده را با توجه به انواع مختلف الگوهای ترافیکی پیش بینی کنیم. به طور سنتی این مشکل باید در مسایل مربوط به شبکه حل و فصل شود ولی در مراکز بزرگ داده با توجه به لزوم ارتباط و عدم وجود الگوی ترافیکی ثابت باعث افزایش هزینه و همچنین دشواری کار شده است. با توجه به این سعی شده است که یک حرکت رو به جلو به سمت پیاده‌سازی و استقرار سرویس‌های متفاوت و با کیفیت که کارایی آنها از سیاست‌های ترافیکی جدا باشد، انجام شود. این کار نیاز به یک مهندسی سطح بالا برای طراحی الگوهای ترافیکی دارد. گسترش طبیعی این رویکرد به سمت فناوری‌هایی است که منابع شبکه را به صورت مجازی فراهم می‌کنند. ایجاد شبکه‌های مجازی این فرصت را برای پرداختن به طرح‌ها و پروتکل‌های مختلف شبکه فراهم می‌سازد.

**منابع ذخیره سازی:** فراهم‌کنندگان ابر عمومی مانند Amazon اغلب انواع مختلفی از سرویس‌های ذخیره سازی را ارائه می‌کنند. سرویس‌ها شامل دیسک‌های مجازی، سرویس‌های پایگاه داده، و سایر امکانات ذخیره سازی می‌باشند. که در سطوح مختلفی دارای ثبات و قابلیت اطمینان هستند.

یک مسئله دشوار برای ارائه سرویس‌های ذخیره سازی کارایی آنها در محیط‌های پویا است. بطوری که یک سرویس باید در زمان افزایش تعداد کاربران و افزایش حجم کار و یا افزایش حجم داده و یا برعکس در زمان کاهش این معیارها بتواند به صورت بهینه با درخواست‌ها تطبیق پیدا کند. همانگونه که در سیستم‌های سنتی و متمرکز داده‌های ذخیره شده باید با خصوصیات تراکنش‌ها شامل یکپارچگی<sup>۷</sup>، سازگاری<sup>۸</sup>، تجزیه پذیری<sup>۹</sup> و پایایی<sup>۱۰</sup> (ACID) سازگاری داشته باشند، در سیستم‌های محاسبات ابری نیز این خصوصیات ضروری هستند. خوشبختانه بسیاری از برنامه‌های کاربردی اجرا شده در محیط ابری ظرفیت بالایی برای سازگاری دارند، و این به طراحان اجازه می‌دهد که بین سازگاری و کارایی سیستم‌های پیاده سازی شده در محیط ابر سازش برقرار کنند. این کار به وسیله طیف وسیعی از فناوری‌های ذخیره داده پیاده سازی می‌شود که قابلیت پیاده سازی در شرایط عملیاتی و غیر عملیاتی متفاوتی را دارند. نمونه‌هایی از این فناوری‌ها شامل: ذخیره‌سازی متنی، ذخیره‌سازی عمودی، ذخیره‌سازی گراف و همچنین ذخیره‌سازی با استفاده از کلید توزیع شده می‌باشد. به تازگی روش استفاده از کلید توزیع شده<sup>۱۱</sup> برای ذخیره داده توجه زیادی را به خود جلب کرده است. همچنین این روش قابلیت پشتیبانی محیط‌های پویا را دارد. در روش ذخیره سازی با استفاده از کلید توزیع شده، درج، بازیابی و تغییر داده‌ها در محیط‌های محاسبات ابری نیز امکان پذیر است [۱۱].

**منابع انرژی:** هزینه‌های مربوط به انرژی سهم قابل توجهی از هزینه‌های مراکز داده را به خود اختصاص می‌دهند. هزینه‌های انرژی شامل هزینه‌های انرژی مصرفی سیستم‌ها و سخت افزارهای مراکز داده و هم شامل انرژی‌های مربوط به خنک کردن مراکز داده

<sup>7</sup> Atomicity

<sup>8</sup> Consistency

<sup>9</sup> Isolation

<sup>10</sup> Durability

<sup>11</sup> Distributed key-value

می‌شود. همانگونه که گفته شد سهم قابل توجهی از هزینه‌ها مربوط به هزینه‌های انرژی است. به عنوان مثال طبق تحقیقات آقای Hamilton، انرژی مصرفی سیستم‌های موجود در مراکز داده برابر ۱۹ درصد هزینه‌های کلی است و همچنین هزینه‌های مربوط به خنک کردن مراکز داده ۲۳ درصد هزینه‌های کلی را به خود اختصاص می‌دهند [۱۲].

در یک مرکز داده انرژی توسط تجهیزات سرورها، تجهیزات شبکه، تجهیزات توزیع انرژی، خنک کننده‌ها و زیرساخت‌های مربوط به پشتیبانی مصرف می‌شود. مراکز داده معمولاً یک یا چند منبع تامین انرژی برای خود فراهم می‌کنند. به تازگی مراکز داده به سمت تولید انرژی مصرفی خود از طریق انرژی‌های تجدید پذیر روی آورده‌اند. همچنین به منظور کاهش هزینه‌ها تمرکز ویژه‌ای بر روی استفاده از روش‌های کاهش مصرف انرژی دارند.

در این حوزه چهار رویکرد اصلی وجود دارد:

- استقرار و پیاده‌سازی مولفه‌های سخت افزاری که حداقل مصرف انرژی را دارند به منظور بهبود بهره‌وری مصرف انرژی.
- پیاده‌سازی سازی روش‌های بهینه تخصیص منابع به منظور کاهش انرژی.
- پیاده‌سازی و اجرای برنامه‌های کاربردی کارآمد به منظور کاهش مصرف انرژی.
- پیاده‌سازی و استقرار سیستم‌های خنک کننده سازگار با منطقه جغرافیایی و آب‌وهوایی به منظور کاهش انرژی مصرفی.

### ۳ روش‌ها و چارچوب مفهومی تخصیص منابع

در این قسمت روش‌های اصلی تخصیص منابع در محیط‌های ابری به طور خلاصه بیان می‌شود و در ادامه به طور مفصل توضیح داده می‌شود. روش‌ها و چارچوب مفهومی و همچنین عملیات و تعاملات بین آنها در شکل ۳ نشان داده شده‌است. عناصر عملیاتی باید به گونه‌ای با یکدیگر هماهنگ شوند که توانایی ارائه یک راه حل جامع در راستای کنترل و نظارت بر انواع منابع و اهداف مدیریتی تخصیص منابع را داشته باشند.

در شکل ۳ نمای عملیاتی و وظایف فراهم کننده ابر و کاربر ابر به عنوان ارائه دهنده زیرساخت به عنوان سرویس نشان داده شده است. ارائه دهنده خدمات ابر مسئول نظارت بر استفاده از منابع پردازشی، شبکه‌ای، ذخیره‌سازی و انرژی می‌باشد و برای این کار از فرایندهای محلی و سراسری استفاده می‌کند. کاربر ابر وظیفه نظارت و کنترل واحدهای مختلف استقرار یافته به عنوان زیرساخت‌های مجازی که از ارائه دهنده خدمات ابر به صورت اجاره در اختیار دارد را برعهده دارد. همچنین میزان حجم کار که مشتریان (کاربران نهایی) از برنامه‌های کاربردی دریافت می‌کنند را کنترل می‌کند. هر دو ارائه دهنده خدمات ابر و کاربر ابر می‌توانند قیمت‌های اجاره زیر ساخت را با توجه به منافع و حجم کار استفاده شده تغییر دهند. مسئولیت کاربر ابر بیشتر مربوط به مسائل مربوط به مدیریت زیرساخت‌های ارائه شده است، اما با انعطاف پذیری بیشتر در زمان درخواست منبع و یا آزاد کردن منبع تخصیص داده شده است. در نهایت مشتری (کاربر نهایی) مسئولیت محدودی در مدیریت منابع دارد. مشتری‌ها با توجه به قیمت‌ها و حجم کار ارائه شده و در بعضی موارد با توجه به چگونگی موقعیت زمانی و مکانی کار ارائه شده، درخواست خود را ارائه می‌دهند.

چارچوب ارائه شده منابع را از دیدگاه IaaS نشان می‌دهد. با این حال این چارچوب از دیدگاه PaaS و SaaS نیز قابل اجرا است، که مانند عناصر عملیاتی زیرساخت محسوب می‌شوند ولی ارائه کننده خدمات ابر مسئولیت کمتری در این موارد دارد.

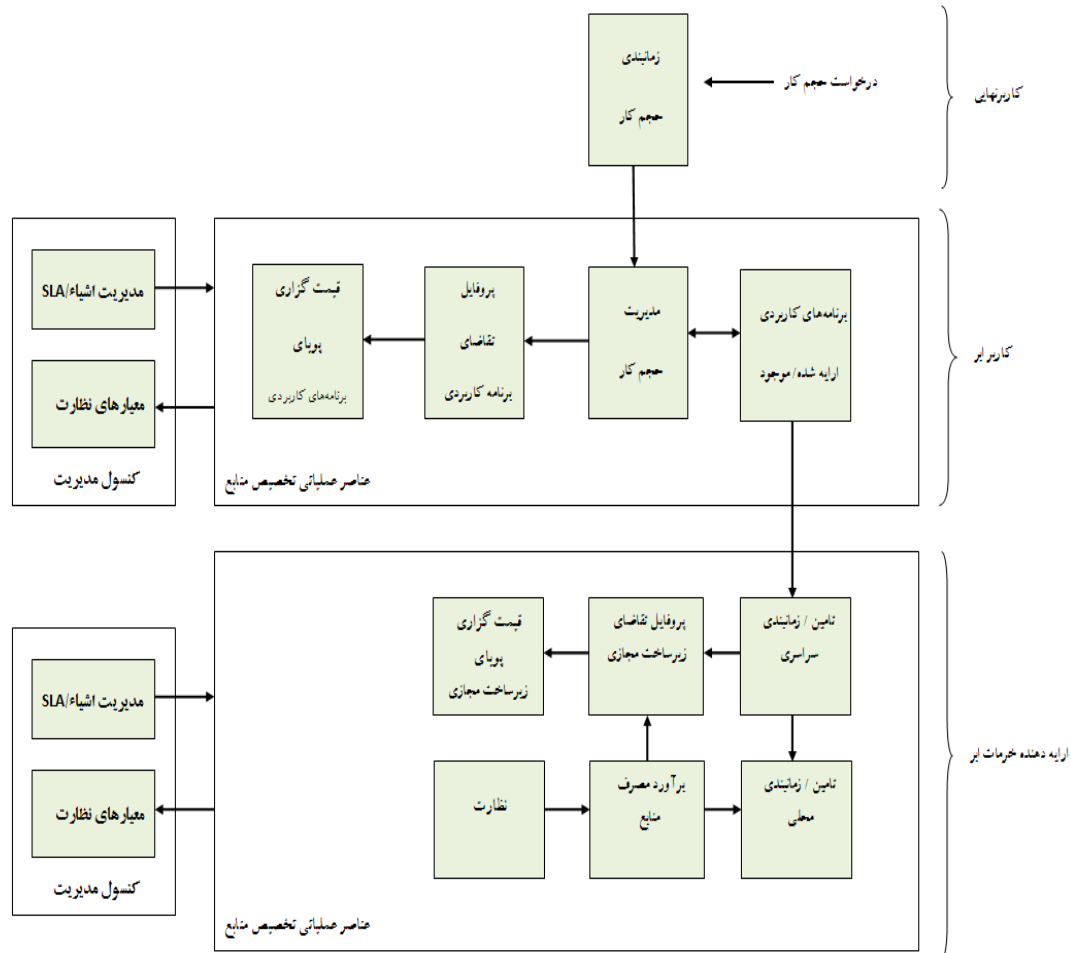
در PaaS، کاربر ابر نقش اصلی در ارائه پلتفرم یا یک برنامه کاربردی را دارد. در مورد SaaS، کاربر ابر معمولاً از همان سازماندهی که توسط ارائه دهنده خدمات ابر ارائه می‌شود، استفاده می‌کند. این سازماندهی باید قابلیت تخصیص منابع را داشته باشد. در ادامه به توضیح مختصر هر کدام از عناصر می‌پردازیم.

**زمانبندی سراسری منابع مجازی:** ارائه دهنده خدمات ابر باید بطور مستقیم<sup>۱۲</sup> VIهایی که به کاربران ابر اجاره داده است و برنامه‌های کاربردی که در ابر میزبانی می‌شوند را پشتیبانی کند. زمانبندی سراسری مربوط به نظارت و کنترل منابع مجازی و فیزیکی سیستم در راستای اهداف تخصیص منابع در سیستم می‌باشد. این موضوع شامل کنترل و پذیرش درخواست‌های کاربران ابر برای پیاده سازی VI و قراردادن کارهای اولیه مربوط به VIها روی زیرساخت‌های فیزیکی منطبق با فرایند تخصیص زیر ساخت در کل زمان استفاده از سیستم می‌باشد.

**پروفایل تقاضای منبع:** یک تخصیص منبع موثر، نیاز دارد که بین واکنش پذیری (تخصیص منابع با توجه به تغییر تقاضاها) و قابلیت پیش‌بینی (تخصیص منابع با توجه به پیش‌بینی تقاضا)، با استفاده از معیارهای اندازه‌گیری، تعادل برقرار کند. در چارچوب ارائه شده پروفایل تقاضا به عنوان ورودی و مبنای قیمت‌گذاری VIها محسوب می‌شود. ورودی‌های این قسمت معمولاً از قسمت زمانبندی سراسری

<sup>۱۲</sup> زیرساخت‌های مجازی

دریافت می‌شود. همچنین پیش‌بینی برآورد استفاده از منبع نیز به عنوان ورودی این دریافت می‌شود، که این نیز می‌تواند مبنای محاسبه قیمت و حجم کار استفاده شده باشد.



شکل ۳: چارچوب مفهومی مدیریت منابع در محیط رایانش ابری

**برآورد استفاده از منبع:** تصمیماتی که برای مدیریت منابع توسط ارایه دهنده ابر و کاربر ابر گرفته می‌شود، نیازمند برآورد اولیه از وضعیت منابع فیزیکی و مجازی برای ارایه به برنامه‌های کاربردی هستند. برآورد وضعیت منابع پردازشی، ذخیره‌سازی، شبکه‌ای و انرژی توسط این قسمت انجام می‌شود. همچنین این قسمت ورودی‌هایی را برای قسمت‌های نظارت و زمانبندی منابع فراهم می‌کند.

**قیمت‌گذاری منابع و افزایش سوددهی:** از دید تجاری، سرویس ارایه شده در محیط محاسبات ابری معمولاً با میزان منابع ارایه شده یا محاسبات انجام شده برای برنامه‌های کاربردی اندازه‌گیری می‌شود. این اندازه‌گیری می‌تواند بر مبنای زمان استفاده و یا سطح استفاده شده از منابع باشد. در ارایه زیرساخت با عنوان IaaS، معمولاً بر اساس زمان استفاده از زیرساخت و ماشین‌های مجازی، قیمت‌گذاری انجام می‌شود. برای محاسبه قیمت سرویس با عنوان PaaS، معمولاً از ترکیبی شامل زمان استفاده از سرویس، پهنای باند مورد استفاده، حجم داده‌های استفاده شده و همچنین تعداد تراکنش‌های انجام شده، استفاده می‌کنند. قیمت سرویس ارایه شده در ابر معمولاً از یک قیمت‌گذاری ثابت که از طرف ارایه دهنده خدمات ابر ارایه می‌گردد، مشخص می‌شود. با این حال، اخیراً سعی بر این شده که قیمت‌گذاری خدمات ابر به صورت پویا انجام شود. به این صورت که، زمانی که مراکز داده حجم کمتری از تقاضا را دارند قیمت سرویس‌ها کاهش یابد و استفاده کنندگان از سرویس برای استفاده بیشتر از سرویس‌ها تشویق شوند. همچنین کاربران ابر به دنبال کاهش هزینه‌های اجاره زیرساخت، و کاهش هزینه‌های کلی و در نتیجه افزایش سود بدست آمده هستند. با فرض اینکه روش قیمت‌گذاری پویا متداول شود، قسمت قیمت‌گذاری زیرساخت و برنامه‌های کاربردی ارایه شده در چارچوب، نیازمند پروفایل تقاضا به عنوان ورودی می‌باشد، که طبق آن حداکثر تاثیر خود را در سیستم داشته باشد.



**زمانبندی محلی منابع مجازی:** مدیریت منابع محلی و زمانبندی کارها برای استفاده از منابع فیزیکی انجام می‌شود و چگونگی قابلیت دسترسی به منابع را مشخص می‌کند. منابع پردازشی مجازی، در واقع یک امکان مجازی شده از ماشین‌های فیزیکی برای به اشتراک گذاری پردازشگرها، ورودی خروجی‌های محلی و رابط‌های شبکه بین ماشین‌های مجازی و برنامه‌های میزبانی شده در ابر می‌باشد. با توجه به قابلیت‌های مجازی سازی، قسمت تامین و زمانبندی محلی منابع، می‌تواند بصورت پویا منابع را بین ماشین‌های مجازی به اشتراک قرار دهد. همچنین در سیستم‌های ذخیره‌سازی توزیع شده، امکان دسترسی به ظرفیت ذخیره‌سازی، کنترل ورودی شبکه و ارسال منابع از مسیرهای مختلف و همچنین انرژی مصرفی تجهیزات را مدیریت و کنترل می‌کند. در تمامی موارد بیان‌شده، عملیات اجرا شده توسط قسمت زمانبندی و تامین منابع باید در راستای اهداف مدیریتی تخصیص منابع باشد.

**تامین و اولویت‌بندی برنامه‌های کاربردی:** کاربران ابر کسانی هستند که برنامه‌های کاربردی را بر روی VI‌های اجاره شده از فراهم کننده خدمات ابر قرار می‌دهند، و بطور معمول نیاز دارند کسانی که از واحدهای نرم‌افزاری استفاده می‌کنند را کنترل نمایند. در حالی که در ارایه سرویس به صورت IaaS و PaaS، ارایه کننده خدمات ابر بطور مستقیم این مسئولیت را به عهده دارد. برای برنامه‌های کاربردی که از چندین واحد نرم‌افزاری تشکیل شده‌اند یا گروهی از نرم‌افزارهای وابسته به هم، باید مکان اولیه در VI برای آنها مشخص شود. بطور خاص برای سیستم‌هایی که به محاسبات متمرکز نیاز دارند مانند پیاده‌سازی MapReduce، واحدهای نرم‌افزاری برای استفاده معینی از منابع مجازی باید پیکربندی شوند. علاوه بر این تغییر تقاضاهای برنامه‌های کاربردی می‌تواند باعث ایجاد تغییر در قرار دادن و پیکربندی واحدهای نرم‌افزاری به صورت پویا شود.

انجام یک فرایند موفق مشروط به تخمین درست از تقاضاهای بعدی می‌باشد. این قابلیت توسط قسمت تامین و اولویت‌بندی برنامه‌های کاربردی انجام می‌شود. در بعضی موارد کاربران ابر، قابلیت‌های واحدهای نرم‌افزاری موجود در VI را به کاربران نهایی پیشنهاد می‌کنند، که این کار نیازمند تصمیم‌گیری در مورد این است که، آیا منابع کافی برای واحدهای نرم‌افزاری در دسترس می‌باشد یا خیر؟ و در نهایت برای پاسخ به تغییرات تقاضاهای برنامه‌های کاربردی و پشتیبانی از VI‌های استیجاری نیاز به اولویت‌بندی پویای منابع است.

**مدیریت حجم کار:** وقتی که واحد‌هایی از نرم‌افزار بر روی VI قرار داده شده است، کاربر ابر باید روی حجم کار درخواست شده برای اجرا توسط کاربر نهایی (مشتری)، کنترل‌هایی را اعمال کند. در ابتدا باید تصمیم بگیرد که آیا حجم کار درخواست شده در راستای اهداف مدیریتی کاربر ابر می‌تواند محاسبه شود یا خیر؟ اهداف مدیریتی از قبیل وضعیت قابلیت دسترسی و کارایی برای حجم کار درخواست شده باید به درستی برآورد شود. هنگامی که حجم کار قابل قبول است که با کاری که یک واحد نرم‌افزاری می‌تواند مدیریت و اجرا کند، مربوط باشد و همچنین در راستای اهداف مدیریتی کاربر ابر نیز باشد.

**سیستم‌های مدیریت محیط ابر:** سیستم‌های مدیریتی توسط ارایه دهنده ابر و بعضی مواقع توسط کاربر ابر استفاده می‌شوند. این سیستم‌ها برای ارایه گزارش، و یا استفاده از رابط کاربری برای اجرای کنترل و فرمان صادر شده توسط مدیریت انسانی هستند. از این قسمت برای تعریف معیارهای مدیریتی و همچنین دریافت بازخوردهای دریافتی از عملیات تخصیص منابع در سیستم استفاده می‌شود. بنابراین از معیارهایی برای برآورد میزان بهره‌برداری از منابع و پروفایل‌های تقاضاها، همراه با اطلاعات سیستم و همچنین اطلاعات نظارتی سیستم استفاده می‌کنند. این قسمت بطور خاص بر اجرای موارد توافق شده در SLA نظارت می‌کند. برای ارایه دهنده خدمات ابر SLA با مشتریان که کاربران ابر هستند نظارت می‌شود، در حالی که کاربران ابر علاوه بر SLA توافق شده با ارایه دهنده خدمات ابر، باید به SLA توافق شده با کاربران نهایی (مشتریان) نیز پایبند باشند.

#### ۴- نتایج کارهای انجام شده

با توجه به گستردگی موضوع، افراد مختلف نیز از دیدگاه‌های مختلف و با اهداف مختلفی به مسایل مربوط به تخصیص منابع در محیط‌های ابری پرداخته‌اند. در جدول ۱ سعی شده که به طور خلاصه یک جمع بندی کلی از روش‌های ارایه شده به همراه مزایا و معایب هر کدام و همچنین هدف از ارایه هر روش را بیان کنیم.



جدول ۱: بررسی روش‌های تخصیص منابع در محیط رایانش ابری

موضوع	هدف	مزایا و معایب
روش‌های برآورد زمان اجرا	<ul style="list-style-type: none"> <li>• تخصیص بهینه ماشین مجازی، مهاجرت زنده</li> <li>• کاهش انرژی</li> </ul>	<ul style="list-style-type: none"> <li>○ کاهش تعداد ماشین‌های مجازی.</li> <li>○ کاهش سربار شبکه.</li> <li>○ در صورتی که برآورد دقیق نباشد باعث شکست خواهد شد</li> </ul>
روش‌های مبتنی بر عامل	<ul style="list-style-type: none"> <li>• تخصیص منابع پویا</li> </ul>	<ul style="list-style-type: none"> <li>○ تخصیص پویا بر اساس نیاز.</li> <li>○ در منابع توزیع شده کاربرد ندارد</li> </ul>
تخصیص منابع با استفاده از زمانبندی محلی و سراسری	<ul style="list-style-type: none"> <li>• تخصیص بهینه</li> <li>• کاهش مصرف انرژی</li> <li>• کاهش ترافیک شبکه</li> </ul>	<ul style="list-style-type: none"> <li>○ مشکل زمانبندی در محیط‌های ناهمگون</li> </ul>
روش‌های مبتنی بر حجم کار	<ul style="list-style-type: none"> <li>• پیش بینی حجم کار</li> <li>• کاهش مصرف انرژی</li> <li>• کاهش زمان پاسخ</li> </ul>	<ul style="list-style-type: none"> <li>○ مشکل برآورد حجم کار اولیه</li> </ul>
روش‌های هوشمند	<ul style="list-style-type: none"> <li>• تخصیص بهینه در محیط‌های پویا</li> </ul>	<ul style="list-style-type: none"> <li>○ سازگار با محیط‌های پویا</li> <li>○ پیچیدگی اجرا</li> </ul>

## ۵ جمع بندی و نتیجه‌گیری

بر اساس تجزیه و تحلیل و بررسی‌های انجام شده منافع زیادی در تخصیص منابع وجود دارند در حالیکه از رایانش ابری صرف نظر از اندازه سازمان و بازارهای تجاری استفاده می‌شود. اما محدودیت‌هایی هم وجود دارد، زیرا رایانش ابری یک تکنولوژی در حال شکل‌گیری می‌باشد. در رایانش ابری منابع به صورت سرویس‌هایی در اختیار کاربران قرار می‌گیرند که این تخصیص منابع یا سرویس‌دهی مزایا و محدودیت‌هایی دارند که در جدول ۲ به صورت کلی بیان شده‌اند.

جدول ۲: مزایا و معایب کلی تخصیص منابع در رایانش ابری

مزایا	محدودیت‌ها
<ul style="list-style-type: none"> <li>• بزرگترین فایده تخصیص منابع این است که کاربر نه نرم‌افزار و نه سخت افزار را برای دسترسی به کاربردها نصب می‌کند، برنامه کاربردی را توسعه می‌دهد و مجری برنامه در اینترنت است.</li> <li>• فایده اصلی بعدی این است که هیچ محدودیت مکانی و محیطی وجود ندارد ما می‌توانیم به برنامه‌های کاربردی و داده‌هایمان در هر جایی در دنیا و بروی هر سیستمی دسترسی پیدا کنیم.</li> <li>• نیازی نیست کاربر، سیستم‌های سخت‌افزاری و نرم‌افزاری را توسعه دهد.</li> <li>• تامین کنندگان ابر در منابعشان بروی اینترنت در طول کمبود منابع شریک هستند.</li> </ul>	<ul style="list-style-type: none"> <li>• به دلیل اینکه کاربران، منابع را از سرورهای دور برای هدف خودشان کرایه می‌کنند آنها کنترلی بر منابعشان ندارند.</li> <li>• مسئله مهاجرت هنگامی روی می‌دهد که کاربران می‌خواهند برای ذخیره‌سازی بهتر داده‌هایشان به یک تامین‌کننده دیگر روی آورند. انتقال داده‌های زیاد از یک تامین‌کننده به تامین‌کننده دیگر آسان نیست.</li> <li>• ابزارهای پیرامونی مانند پرینترها یا اسکنرها نمی‌توانند با ابر کار کنند. بسیاری از آنها نیازمند نرم‌افزاری هستند که به صورت محلی نصب می‌شوند. تجهیزات جانبی شبکه بندی شده مشکلات کمتری دارند.</li> <li>• دانش بیشتر و عمیق‌تری برای تخصیص و مدیریت منابع در ابر مورد نیاز است، زیرا تمامی اطلاعات در مورد کارکردن ابر عمدتاً وابسته به تامین‌کنندگان خدمات ابر هستند.</li> </ul>

علاوه بر مزایا و محدودیت‌های گفته شده، چالش‌هایی در فرایند تخصیص منابع وجود دارد که می‌تواند مبنای تحقیقات آینده باشند. بعضی از چالش‌های موجود، مسائلی هستند که به طور ذاتی پیچیده و مشکل هستند، که برای برطرف کردن این نوع مسائل نیاز به کشف روش‌های جدید می‌باشند و بعضی از چالش‌ها نیز براساس محدودیت‌های موجود یا مسائل مدیریتی و طراحی به وجود می‌آیند و دلیل آن عدم وجود یک استاندارد معین برای محیط‌های رایانش ابری می‌باشد. طبق بررسی‌ها و مقایسه‌های انجام شده، یکی از چالش‌های موجود در فرایند تخصیص منابع که نیاز به توجه کافی و تدوین به یک روش استاندارد و مشخص دارد، دستیابی به روش‌های پیش‌بینی عملکرد برنامه‌های کاربردی و منابع درخواستی است. پیش‌بینی و فهمیدن عملکرد برنامه‌های کاربردی مانند زمان دقیق اجرا، زمان درخواست منبع و زمان آزادسازی منبع به عنوان چالشی که در زمان تدوین SLA باید مشخص شود. برای این منظور نیاز به یک مدل‌سازی درست و زمانبندی دقیق از ارایه درخواست‌ها داریم، که این کار به دلیل پویایی و پیچیدگی مسئله کار مشکلی است. علاوه بر این هیچ کنترل ثابتی برای پیش‌بینی عملکرد برنامه‌های کاربردی در محیط ابر وجود ندارد. به طور کلی پیش‌بینی عملکرد در بیشتر سیستم‌های کامپیوتری به عنوان چالش محسوب می‌شود. در محیط‌های ابری نیز با توجه به وجود یک لایه بیشتر به عنوان لایه مجازی، کار پیش‌بینی را سخت‌تر می‌کند. در نتیجه پیش‌بینی درخواست منابع مجازی سخت‌تر از پیش‌بینی درخواست منابع فیزیکی می‌باشد. همچنین زمان اجرای برنامه‌های کاربردی و زمان استفاده از منابع با توجه به اینکه در هر اجرا از ماشین‌های مجازی متفاوتی استفاده می‌کنند، کار را سخت‌تر می‌کند. بنابراین روش پیشنهادی برای این منظور، استفاده از الگوریتم‌های داده‌کاوی برای دسته‌بندی درخواست‌های دریافتی می‌باشد. با توجه به ماهیت الگوریتم‌های دسته‌بندی و همچنین علم داده‌کاوی می‌توان از عملیات‌های تخصیص منابع انجام شده قبلی در محیط ابری و همچنین بررسی شرایط موجود هر درخواست، یک دسته‌بندی از درخواست‌های دریافتی ارایه کرد و طبق آن منابع مورد نیاز هر دسته از درخواست‌ها را فراهم نمود. همچنین این کار نیاز به مجموعه داده از تاریخچه درخواست‌ها دارد. بنابراین پس از مشخص کردن دسته‌های مختلف از درخواست‌ها، در زمان دریافت یک درخواست و مشخص شده دسته و الگوی درخواست بنابراین منابعی که مربوط به این دسته از درخواست‌ها می‌باشد ارایه شده و همچنین زمان آزادسازی و نوع استفاده از منابع توسط درخواست، برای انجام کارهای مدیریتی نیز مشخص می‌شود.

## منابع و مراجع

- [1] Wikipedia .[Online].Available: <http://fa.wikipedia.org>
- [2] NIST.[Online].Available: <http://www.nist.gov/itl/cloud>
- [3] R Siva Theja Maguluri , Lei Ying, "Heavy traffic optimal resource allocation algorithms for cloud computing clusters," journal homepage: [www.elsevier.com/locate/peva](http://www.elsevier.com/locate/peva), 2014.
- [4] R Madhumathi, A. S. Balagopalan, "Dynamic Resource Allocation in Cloud Using Bin-Packing Technique," 2015 International Conference on Advanced Computing and Communication Systems, 2015.
- [5] G Qian Zhao, "Research on Resource Allocation for Cloud Computing Platform based upon Service Requirement," Advanced Science and Technology Letters, vol. 78, 2014.
- [6] T Mansoor Alicherry, "Net work Aware Resource Allocation in Distributed Clouds," Proceedings IEEE INFOCOM, 2012.
- [7] J Lu Dai, "An optimal Resource Allocation Algorithm in Cloud Computing Environment," Applied Mechanics and Materials, vol. 733, 2015.
- [8] S Linlin Wu, Steve Versteeg and Rajkumar Buyya, "SLA-based Resource Provisioning for Software-as-a-Service Applications in Cloud Computing Environments," IEEE TRANSACTIONS ON SERVICES COMPUTING, 2013.
- [9] S Gopal Kirshna Shyam, "Resource Allocation in Cloud Computing Using Agents," in 2015 IEEE International Advance Computing Conference (IACC), 2015.
- [10] Al-Fares, M., Loukissas, A., Vahdat, A," A scalable, commodity data center network architecture In": Proceedings of the ACM SIGCOMM 2008 Conference on Data communication (SIGCOMM 2008), 2008
- [11] Robinson, I., Webber, J., Eifrem, E" Graph Databases", 1st edn. O'Reilly, Media, 2013
- [12] Hamilton, J.R." Cooperative expendable micro-slice servers (CEMS): low cost, low power servers for internet-scale services", Proceedings of 4th Biennial Conference on Innovative Data Systems, 2009
- [13] Dilip Kumar, and Bibhudatta Sahoo"Energy Efficient Heuristic Resource Allocation for Cloud Computing",Computer Science & Engineering, National Institute of Technology,2014