

بررسی معیارهای ارزیابی کارایی محیط‌های محاسبات ابری

محمد حسین زاده*

دانشجوی کارشناسی ارشد، گروه کامپیوتر، دانشکده مهندسی فناوری اطلاعات و ارتباطات، دانشگاه جامع امام حسین (ع)، تهران، ایران
پست الکترونیکی: mhosseinzadeh@ihu.ac.ir

صادق بجانی

استادیار، گروه کامپیوتر، دانشکده مهندسی فناوری اطلاعات و ارتباطات، دانشگاه جامع امام حسین (ع)، تهران، ایران
پست الکترونیکی: sbejani@ihu.ac.ir

چکیده

گسترش موفق محاسبات ابری در سال‌های اخیر موجب شده است که نیاز به ارزیابی ابرها به یک واقعیت غیرقابل انکار تبدیل شود. اما به دلیل پیچیدگی محیط ابر، ارزیابی عملکرد آن با دشواری‌هایی همراه است. برای غلبه بر این دشواری‌ها باید روش‌های اختصاصی برای ارزیابی عملکرد هر ابر با توجه به قابلیت‌های آن ابر و اهداف ارائه‌دهندگان خدمات آن توسعه داده شود. عدم توجه به کارایی محاسبات ابری موجب پایین بودن بهره ابر خواهد بود که یکی از چالش‌های اصلی این حوزه تلقی می‌شود. در این مقاله روش‌های متنوع و مهم ارزیابی عملکرد ابر را مورد بررسی قرار داده و مشخص کرده‌ایم کدام روش برای چه ابری مناسب است، چه معیارهایی برای ارزیابی استفاده می‌شود و هر روش چه مزایایی دارد تا بدین ترتیب ارائه‌دهندگان ابر بتوانند روش مناسب برای ارزیابی عملکرد ابر خود را انتخاب کنند.

واژگان کلیدی: ارزیابی عملکرد، ابر، کارایی، محاسبات ابری، صف

1. مقدمه

در سال‌های اخیر، محبوبیت و رشد سریع در قدرت پردازش و فناوری ذخیره‌سازی و گسترش سریع اینترنت موجب شده است منابع محاسباتی به منابعی ارزان‌تر، قوی‌تر و در دسترس‌تر از قبل تبدیل شوند [1]. این روند فناوری به‌عنوان محاسبات ابری شناخته می‌شود و برای آن که بتواند پاسخ مناسبی را به گسترش فناوری اطلاعات و ارتباطات بدهد نیازمند پیمودن راه تکاملی مناسبی است [2]. همچنین نیاز به این هست که افراد بتوانند کارهای محاسباتی سنگین خود را بدون داشتن سخت‌افزارها و نرم‌افزارهای گران، از طریق خدمات ابر انجام دهند. رایانش ابری¹ آخرین پاسخ فناوری به این نیازها بوده است. محاسبات ابری بر روی تحویل به‌موقع خدمات، قابلیت اطمینان، تأمین امنیت در ارسال و دریافت داده‌ها، تحمل‌پذیری خطا در ارائه سرویس، پایداری و ایجاد زیرساخت‌های مقیاس‌پذیر برای میزبانی خدمات کاربردی مبتنی بر اینترنت متمرکز شده است. محاسبات ابری در واقع بیانگر نرم‌افزاری است که خدمات را از میان منابع محاسباتی مجازی موجود به کاربران در هر نقطه از جهان ارائه می‌دهد. [3] استفاده از خدمات ابر باعث می‌شود که رابطه در حال توسعه‌ای در میان هر دو بخش دولتی و خصوصی برای خدمت‌رسانی به مردم

* نویسنده مسئول

¹ Cloud computing

به وجود آید. محاسبات ابری از آن جهت برای صاحبان کسب و کار جذاب است که می‌توانند کار خود را با منابع کوچک و کمتر شروع کنند و همزمان با رشد تقاضا از طرف کاربران به ظرفیت منابع خود بیفزایند و برای کاربران نیز از آن جهت جذابیت دارد که به جای پرداخت هزینه‌های گزاف برای در اختیار گرفتن منابع محاسباتی یا نرم‌افزارهای مورد نیاز کافی است هر کدام را به هر اندازه که نیاز دارند به عنوان خدمت از طرف ارائه‌کنندگان تحویل بگیرند و بدین ترتیب صرفه‌جویی بسیار زیادی در زمینه اقتصادی داشته باشند [4]. به دلیل ماهیت خدمت‌گرا بودن محاسبات ابری، ارائه‌دهندگان خدمت نیاز دارند تا آگاهی کاملی از وضعیت سیستم محاسبات ابری خود داشته باشند و به‌طور مداوم بتوانند کارایی آن را ارزیابی کنند تا بتوانند پاسخ مناسب را با کیفیت مناسب به حجم خدمات مورد تقاضای کاربران ارائه دهند. در ارزیابی کارایی محیط محاسبات ابری به دلیل وجود مدل‌های کاربردی و خدماتی مختلف در اجرای سیاست‌های تخصیص منبع و الگوریتم‌های زمان‌بندی، با چالش‌هایی مواجه می‌شویم که از آن میان می‌توان به متفاوت بودن هزینه‌ها، عدم تولید دوباره نتایج آزمون‌ها و عدم اعتماد سرویس‌دهندگان ابر به ارزیابی مبهم و روش آزمون و خطا. همچنین برای آن که از کیفیت خدمت ارائه‌شده توسط ابر اطمینان یابیم و قابلیت اطمینان این خدمات را افزایش دهیم نیازمند ارائه راه‌حلی برای شبیه‌سازی و ارزیابی کارایی^۲ رایانش ابری هستیم [5]. ما در این مقاله چندین روش متداول ارائه‌شده برای ارزیابی کارایی محیط‌های محاسبات ابری را بررسی می‌کنیم و این روش‌ها را از نظر معیارهای مورد توجه هر کدام و طریقه ارزیابی هر روش و نتایج به‌دست‌آمده باهم مقایسه می‌کنیم. سپس بیان می‌کنیم که هر روش ارزیابی کارایی برای کدام نوع محیط‌های محاسبات ابری مناسب است تا بتوان بهترین روش برای ارزیابی کارایی یک محیط محاسبات ابری بخصوص را انتخاب کرد. نتایج بررسی همه این موارد به‌صورت یک جدول در پایان مقاله ارائه‌شده است.

2. اهمیت ارزیابی کارایی محیط‌های ابری

به مرور زمان، خدمات مبتنی بر محاسبات ابری در همه جنبه‌های فناوری اطلاعات ظاهر شده و در زندگی روزمره ما نفوذ کرده‌اند. برخلاف روش‌های سنتی که در آن ارائه‌دهندگان خدمات، خودشان از قدرت محاسبات و منابع خود برای ارائه خدمت به کاربران استفاده می‌کردند، امروزه محاسبات ابری خدمات مبتنی بر تقاضای کاربران را در بسترهای شبکه بخصوص اینترنت قرار می‌دهد و این کاربران هستند که بر اساس نیاز خود از منابع ارائه‌دهندگان خدمات استفاده می‌کنند [6,7] مراکز ابر را می‌توان به‌عنوان مجموعه‌ای از گره‌های محاسباتی در نظر گرفت که مسئول اجرای کارهای کاربران هستند. کارهای کاربران طبق یک سری الگوهای از پیش تعیین‌شده به گره‌های محاسباتی می‌رسند، در صف دریافت خدمت قرار می‌گیرند و طبق نیازمندی به گره مناسب تخصیص داده می‌شوند. این به‌طور طبیعی منجر به یک سیستم صف‌بندی با چندین کارساز^۳ (گره) می‌شود. ماهیت صف‌بندی کارها در محیط‌های ابری موجب شده است که ارزیابی عملکرد ابرها تمرکز ویژه‌ای روی مباحث صف‌ها داشته باشد و عملاً ارزیابی کارایی صف بخش مهمی از ارزیابی کلی سیستم ابر را شامل شود. با این حال سیستم‌های مبتنی بر ابر در حال حاضر دارای چالش‌هایی برای مدل‌های صف استاندارد هستند که عبارت‌اند از [8]:

- منابع محاسباتی بسیار زیاد هستند (حداقل چند صد عدد)
- فرآیند ارائه خدمت می‌تواند بسیار متغیر باشد.
- روند ورود کارها می‌تواند تنوع بسیار بالایی داشته باشد.

با وجود فراگیر شدن سیستم‌های ابری و انجام تحقیقات زیاد در مورد این سیستم‌ها، مطالعات بسیار کمی روی ارزیابی عملکرد این سیستم‌ها صورت گرفته است که این امر موجب شده تا بررسی و ارزیابی عملکرد محیط‌های ابری موضوع مناسب و پر چالش برای محققان باشد.

² Performance evaluation

³ server

3. کارهای مرتبط

ما در این مقاله چندین روش ارائه شده برای ارزیابی کارایی محیط‌های ابری را بررسی می‌کنیم. در این بررسی سعی شده است فقط روی مباحث مربوط به صف‌بندی و ارزیابی کارایی ابرها از روی ارزیابی صف کارها، تمرکز نشود بلکه روش‌های نوین و متفاوت ارائه شده که به بررسی و ارزیابی کارایی جنبه‌های دیگری از سیستم‌های ابری می‌پردازند نیز مورد تحلیل قرار گیرند.

3-1 ارزیابی عملکرد خدمات ابری با بهینه کردن بهره ابر

هدف اصلی ارائه‌دهندگان خدمات ابری بهره‌برداری کامل از منابع و به حداکثر رساندن سود با استفاده از یک برنامه‌ریزی مناسب است. مسئله مهم در محاسبات ابری برنامه‌ریزی خدمات و طریقه برنامه‌ریزی برای تخصیص کارها به‌منظور بالا بردن بهره ابر است. در [9] روشی برای ارزیابی کارایی محیط ابر ارائه شده که از مدل صف‌های اولویت بهره می‌برد تا خدمات اجاره‌ای که توسط ابر فراهم شده‌اند را بتوان ارزیابی کرد. برای این کار زمان خدمت‌دهی، زمان پاسخ و زمان انتظار برای کارهایی که در صف قرار دارند محاسبه می‌شود و بدین ترتیب کارایی سیستم با استفاده از روش صف‌های اولویت محاسبه می‌شود. خدماتی که توسط ابرها ارائه می‌شوند به سه دسته کلی تقسیم می‌شوند: 1) نرم‌افزار به‌عنوان خدمت⁴ (2) بن‌سازه به‌عنوان یک خدمت⁵ و 3) زیرساخت به‌عنوان خدمت⁶. در مدل صف استفاده شده در این روش از نرخ ورود مارکوفی⁷، نرخ عمومی خدمت، m به‌عنوان تعداد کارسازها، نظم صف اولویت و I به‌عنوان اندازه میانگین بهره گرفته شده است. مهم‌ترین مزیت این مدل در طول زمان، این است که ارائه‌دهنده خدمات ابر کارها را به‌گونه‌ای برنامه‌ریزی و صف‌بندی می‌کند تا به بیشترین بهره در استفاده از سیستم ابر خود دست یابد. به دلیل معماری پیچیده ابر، ارزیابی کارایی این سیستم‌ها با استفاده از صف‌های معمولی با دشواری‌هایی همراه است. دلیل این دشواری‌ها عبارت‌اند از:

- محیط ابر به پشتیبانی تعداد زیادی کارساز نیاز دارد درحالی‌که مدل‌های معمولی صف تعداد کارسازهای کمتری را پشتیبانی می‌کنند.
- برای راحتی در مدل‌های صف‌بندی سنتی، زمان خدمت‌رسانی معمولاً به‌صورت توزیع نمایی در نظر گرفته می‌شود درحالی‌که در محیط‌های ابر باید زمان خدمت‌رسانی عمومی کارها در نظر گرفته شود.
- به دلیل ماهیت پویای ابر، بارکاری همیشه یکنواخت نیست و انتظار می‌رود ابر بتواند کیفیت خدمت⁸ را در شرایط پویا و متفاوت نیز تأمین کند.
- کیفیت خدمت نیز باید بر اساس اولویت خدمات در نظر گرفته شود و موازنه‌ای بین اهمیت اولویت و اهمیت کیفیت برقرار شود تا بدین ترتیب بهره ابر نیز به حداکثر خود برسد.

در [9] برای حل مشکلات فوق از روش صف‌های اولویت $M/G/m/m+r$ استفاده شده است که سازوکار ورود به‌صف نیز از نوع به ترتیب ورود (FCFS) انتخاب شده است. در این روش از صف‌بندی وزن‌دار عادلانه استفاده شده تا بتوان به بالاترین بهره رسید. نتایج ارزیابی نشان می‌دهد که این روش بهره بالاتر و کیفیت بهتری را نسبت به صف معمولی با سازوکار به ترتیب ورود به دست می‌دهد. شکل 1 معماری این روش را نشان می‌دهد. مطابق شکل، در این معماری برای هرکدام از سه دسته خدمات ابر یک صف دارای اولویت مجزا در نظر گرفته شده است همچنین برای کارها نیز سه رده مجزا در نظر گرفته شده است که رده یک دارای بالاترین اولویت و رده سه دارای کمترین اولویت است. این سه رده برای هرکدام از صف‌های اول تا سوم در نظر گرفته شده است و هرکدام از صف‌ها دارای زمان خدمت‌دهی و نرخ ورود اختصاصی هستند. در اینجا با توجه به سیاست‌های موردنظر ارائه‌دهنده ابر، خدمات

⁴ Software as a Service

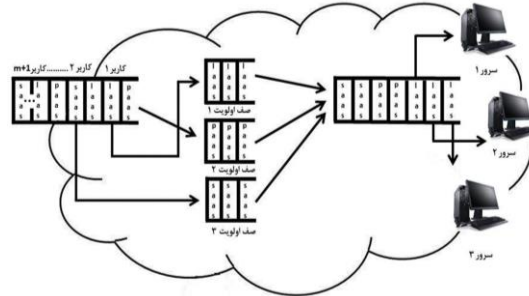
⁵ Platform as a Service

⁶ Infrastructure as a Service

⁷ Markovian arrival rate

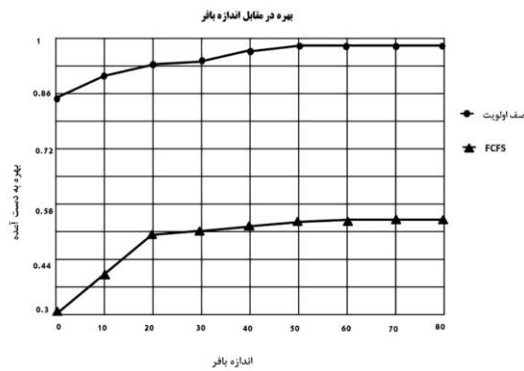
⁸ Quality of service

IaaS دارای بالاترین اولویت و خدمات SaaS دارای پایین‌ترین اولویت است. حال برای صف اول زمان انتظار و متوسط تعداد کارهای منتظر محاسبه می‌شود سپس برای صف دوم نیز همین کار انجام می‌شود با این تفاوت که در اینجا به دلیل پایین‌تر بودن اولویت این صف نسبت به صف قبلی باید زمان انتظار صف قبل نیز به کارهای منتظر در صف دوم اضافه شود. بدین ترتیب برای صف سوم هم دو زمان انتظار صف اول و دوم باید افزوده شود.



شکل 1: معماری روش استفاده از صف‌های اولویت در ابر [9]

بعد از شبیه‌سازی و اجرای دستورالعمل فوق نتیجه کار در شکل 2 نشان داده شده است که مشاهده می‌شود استفاده از صف‌های اولویت موجب افزایش بهره سیستم نسبت به حالت استفاده از صف‌های معمولی شده است.



شکل 2: مقایسه استفاده از صف‌های اولویت و عدم استفاده از آن‌ها [9]

2-3 ارزیابی عملکرد خدمات ابر با در نظر گرفتن بازیابی خطا

برای این که قابلیت اطمینان خدمات ابر بالا رود نیاز به استفاده از بازیابی خطا داریم [10]. همچنین استفاده از قابلیت بازیابی خطا موجب تأثیر بر روی عملکرد خدمات یک ابر خواهد شد. در [3] این موضوع مورد بررسی قرار گرفته است. کیفیت عملکرد ابر در این روش توسط زمان پاسخ خدمت سنجیده شده است. در این روش ابتدا زمان پاسخ خدمات مدل‌سازی شده است که از رابطه 1 به دست می‌آید:

$$T_{SRT} = T_{SUB} + T_W + T_S + T_E + T_R \quad (1)$$

در این رابطه T_{SUB} نشانگر زمان رسیدن درخواست مشتری به ابر و T_W زمان انتظار، T_S زمان سرویس، T_E کل زمان اجرای سرویس و T_R زمان بازگشت جواب به کاربر است. البته در این روش از زمان‌های T_{SUB} و T_R صرف‌نظر شده و رابطه 2 در نظر گرفته شده است:

$$T_{SRT} = T_W + T_S + T_E \quad (2)$$

قابلیت بازیابی خطا تأثیری روی زمان انتظار و زمان خدمت ندارد ولی در T_E مؤثر است به طوری که اگر از قابلیت بازیابی خطا استفاده کنیم مدت زمان T_E فرق خواهد کرد. فرض می‌کنیم میزان خطا برای گره j برابر λ_j باشد. سپس مجموع احتمال خطا برای گره j در مدت زمان t تا 0 از رابطه 3 به دست می‌آید:

$$\Pr\{N_j(t) = n\} = \frac{(\lambda_j t)^n}{n!} \exp(-\lambda_j t), n = 0, 1, \dots \quad (3)$$

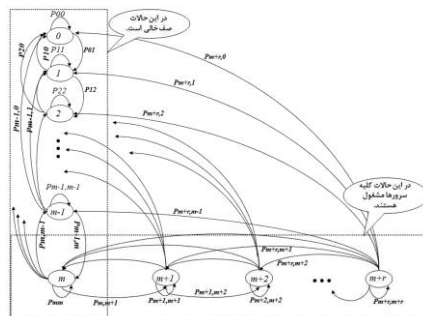
همچنین زمان بازیابی خطا نیز از رابطه 4 به دست می‌آید:

$$TR_j(t) = \sum_{k=1}^{N_j(t)} TR_j^{(k)} \quad (4)$$

با دخالت دادن فاکتورهای فوق در T_E زمان پاسخ خدمات به دست می‌آید. بدین ترتیب می‌توانیم عملکرد ابرهای با قابلیت بازیابی خطا را ارزیابی کنیم.

3-3 ارزیابی کارایی رایانش ابری با استفاده از زنجیره مارکوف

در [11] یک مدل جدید برای ارزیابی کارایی سیستم‌های رایانش ابری ارائه شده است. این مدل از یک سو بر روی تعیین تعداد کارسازهای رایانش ابری و اندازه میانگیرهای ورودی آن‌ها توجه کرده و از سوی دیگر به فاکتورهای کارایی سیستم مانند تعداد کارهای سیستم و کارهایی که نیاز به خدمت‌دهی فوری دارند توجه می‌کند. در این روش برای محاسبه کارایی از روش مارکوف و صف‌های $M/G/m/m+r$ استفاده شده است که زمان رسیدن کارها دارای توزیع نمایی (M) هستند. درحالی‌که زمان خدمت‌دهی، متغیرهای تصادفی مستقل با توزیع یکسان هستند که توزیع عمومی (G) برابر با مقدار متوسط μ دارند. در این مدل از تعداد m کارساز استفاده شده که دارای فرآیند FCFS هستند که فرآیند ورود توسط مدل‌های مارکوف مدل‌سازی شده است. بنابراین با توجه به وجود m کارساز و اندازه میانگیر r ، ظرفیت کل سیستم $m+r$ می‌شود. صف‌های $M/G/m/m+r$ ممکن است به‌عنوان یک مدل نیمه مارکوف در نظر گرفته شود [12] که می‌توانند توسط تکنیک‌های زنجیره‌های مارکوف تجزیه و تحلیل شوند. در این روش زنجیره مارکوف طبق مدل شکل 3 طراحی شده است که هر یک از حالات با شماره‌گذاری از صفر تا $m+r$ نشان داده شده‌اند و احتمالات گذر از یک حالت به حالت دیگر نیز مشخص شده است. در این مدل، حالات به دو دسته تقسیم شده‌اند، دسته اول که از صفر تا $m-1$ شماره‌گذاری شده‌اند در حالت خدمت‌گیری و دسته دیگر که از m تا $m+r$ مشخص شده‌اند، منتظر خالی شدن صف یک کارساز برای قرار گرفتن در صف خدمت‌گیری هستند. برای وضوح بیشتر، تنها برخی از انتقال‌ها به طور کامل کشیده شده‌اند و بقیه به صورت ناتمام نشان داده شده‌اند. این مدل نشان می‌دهد هر یک از حالات سیستم در حال انتظار، با چه احتمالی ممکن است به حالت خدمت‌گیری منتقل شود. با تحلیل این مدل می‌توان کارایی سیستم را اندازه‌گیری کرد.



شکل 3: طراحی زنجیره مارکوف با استفاده از صف $M/G/m/m+r$ [11]

با استفاده از مدل شرح داده شده در شکل 3 و با داشتن رابطه توزیع تجمعی (CDF) $A(x) = \text{prob}[A < x]$ و همچنین تابع چگالی احتمال (Pdf) $a(x) = \lambda e^{-\lambda x}$ می توان با تبدیل لاپلاس میانگین زمان ورود (A^*) را به دست آورد. به همین ترتیب مدت زمان خدمت (B) نیز قابل محاسبه است همچنین در این روش با استفاده از B، زمان خدمت باقی مانده کارها نیز محاسبه می شود. همگن و ارگودیک بودن زنجیره مارکوف در این روش، باعث شده که یک حل حالت پایدار از سیستم به دست آید. بدین ترتیب چون سیستم به حالت پایدار رسیده است می توان احتمال خروج هر کار را نیز با رابطه های معین به دست آورد. سپس با به دست آوردن ماتریس حالت گذار سیستم، معادلات جریان برای آن نوشته شده و احتمال هر یک از حالت ها محاسبه می شود و بدین ترتیب معیارهای مربوط به کارایی سیستم ابر نیز به دست می آید. در پایان [11] نیز روش ارائه شده شبیه سازی و آزمایش شده و صحت عملکرد آن برای صف هایی با تعداد خدمت دهنده و تعداد کارهای مختلف به دست آمده است. با توجه به ماهیت محیط ابر، در این روش مدت زمان عمومی خدمت برای درخواست ها و همچنین تعداد زیاد کارسازها در نظر گرفته شده اند که باعث شده است تا این مدل از نظر مقیاس و تنوع زمان خدمت دهی، انعطاف پذیر باشد. همچنین این روش اثبات کرده است که مراکز ابری که از کارسازهای ناهمگن استفاده می کنند در مقایسه با مراکزی که از کارسازهای همگن استفاده می کنند، مدت انتظار بیشتری را به کاربران تحمیل می کنند.

3-4 ارزیابی عملکرد رایانش ابری برای خدمات محاسباتی با تعداد زیاد کار

در [13] به بررسی استفاده از رایانش ابری برای پردازش کارهای حجیم علمی پرداخته شده است و طریقه ارزیابی عملکرد هنگامی که با حجم بسیار بالایی از کارها مواجه هستیم، شرح داده شده است. در این حالت کاربران بسیار زیادی از برنامه هایی مبتنی بر ابر استفاده می کنند تا پردازش های خود را انجام دهند و بدین ترتیب حجم عظیمی از کارها بین ابر و کاربران مبادله خواهد شد. این امر موجب می شود که ارزیابی کارایی ابر اهمیت ویژه ای پیدا کند تا با آگاهی از وضعیت خدمت رسانی ابر از اعمال بار بیش از حد توان به آن جلوگیری کنیم و همچنین با ارزیابی کارایی سعی در بالا بردن بهره ابر داشته باشیم. در [13] برای آن که بتوان به تجارب واقعی دست یافت، کارایی چندین خدمت دهنده بزرگ همچون آمازون ارزیابی شده و فاکتورهای کارایی آن ها استخراج شده است. نتایج نیز نشان می دهد که باید بهبودهایی در عملکرد ابرهای امروزی صورت گیرد تا بتوان عملکرد بهینه برای کارهای با حجم بالا حاصل شود و برای رسیدن به این هدف مهم ترین چیزی که باید در نظر گرفته شود میزان اختلاف عرضه و تقاضا در ابر است. نکته مهمی که در نظر گرفته شده است زیاد بودن تعداد کارهای کاربران است نه این که حجم هر کار زیاد باشد، یعنی تعداد تبادلات کارها بین ابر و کاربران بسیار زیاد است که به این حالت MTC⁹ گفته می شود [14].

اولین چیزی که در این حالت باید در نظر داشت شناسایی کاربرانی هستند که به صورت MTC هستند که برای این کار روش هایی ارائه شده است [15,16] یعنی باید ابر بدانند چه موقع ممکن است با افزایش کاربران MTC مواجه شود و تعداد این کاربران چند تا است و حجم کاری اعمال شده از این کاربران را پیش بینی کند. بعد از این مرحله نوبت به ارزیابی عملکرد ابر می رسد.

در [13] برای ارزیابی عملکرد از روش محک زنی¹⁰ سیستم استفاده شده است. در [17] نشان داده شده است که محک زنی عملکرد اجزای سیستم های مختلف با طیف گسترده ای از معیارهای خرد و هسته های برنامه های کاربردی، می تواند یک برآورد مرتبه اول از عملکرد سیستم را فراهم کند. البته روش استفاده شده در [13] کاملاً مشابه [17] نیست بلکه به جای آن، مشخصات اجزای ابر اضافه شده، چند معیار محک زنی ارزیابی جامع و مستقل انتخاب شده و همچنین روی معیارهای خاص سیستم های با مقیاس بزرگ تمرکز شده است. محک هایی که برای ارزیابی عملکرد مشخصات ابر استفاده شده به طور کامل در [13] بیان شده است. ابزار تحلیل عملکردی که در این روش استفاده شده است گسترشی از روش GrenchMark [18] است که در [19] بیان شده است. این گسترش با هدف پشتیبانی از سیستم های توزیع شده با مقیاس بزرگ توسعه داده شده است که برای آزمایش کارایی سیستم های ابر

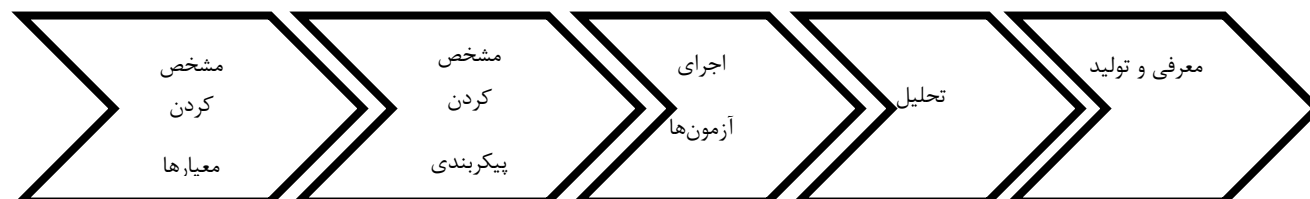
⁹ Many Task Computing

¹⁰ Benchmark

قابل استفاده است. این چارچوب قادر به تولید و ارائه حجم کار واقعی و مجازی برای سیستم‌های توری¹¹، ابر و سایر سیستم‌های مقیاس بزرگ است. در پایان نیز عملکرد تجربی چهار ابر بزرگ و متداول بخصوص EC2 آمازون مورد بررسی قرار گرفته است.

3-5 ارزیابی عملکرد ارائه‌دهندگان محاسبات ابری

در [20] یک روش برای ارزیابی عملکرد پیکربندی رایانش ابری ارائه شده است. در حالت عادی ارائه‌دهندگان خدمت بر روی بُن‌سازهٔ سطوح خدمتی که ارائه می‌کنند توجه دارند ولی در [20] به تک‌تک خدمات به جای بُن‌سازهٔ کلی، توجه شده است. در ارزیابی کارایی محیط‌های ابری ما نیازمند یک روش هستیم که به امکانات خاص پیکربندی و گزینه‌های اندازه‌گیری توجه کند. در [20] برای ارزیابی عملکرد ابر از روش ASTAR استفاده شده است. [21,22,23] ASTAR یک روش ارزیابی کارایی برای محیط‌های خدمت‌گرا (SOA) با معماری‌های شفاف است. رویکرد جعبه سیاه¹² روش ASTAR (ارزیابی عملکرد توسط پایش عملکرد خارجی کارسازهای وب انجام می‌گیرد) باعث می‌شود تا این روش یک گزینه خوب برای ارزیابی عملکرد محیط‌های ابری باشد. روش ارائه شده در [20] با گسترش امکانات و تغییر در پیکربندی روش ASTAR، آن را برای ارزیابی محیط‌های ابری مهیا کرده و از آن در این جهت استفاده شده است. برای این کار یک سری معیارهای مناسب که مربوط به ارزیابی ابرها هستند شناسایی شده و در نظر گرفته می‌شوند سپس این معیارها در پیکربندی ASTAR دخالت داده شده و پیکربندی جدید تنظیم می‌شود. در این پیکربندی باید به انواع خدماتی که ابر ارائه می‌دهد توجه شود و معیارها متناسب با هر خدمت در پیکربندی اضافه شود. همانند روش ASTAR روش مطرح شده در [20] نیز جزء روش‌های تجربی علوم کامپیوتر طبقه‌بندی می‌شود. در روش‌های تجربی با استفاده از آزمایش سعی در ارزیابی یک نظریه داریم که یا نظریه تأیید خواهد شد و یا پدیده‌های جدیدی برای نظریه به دست خواهد آمد. این روش نیازمند ابزاری است تا بتواند کارایی را ارزیابی کند که این ابزار توسط [24] و [25] بسط داده شده‌اند. سپس روش مدنظر برای ارزیابی عملکرد خدمت‌دهنده آمازون یعنی EC2 مورد استفاده قرار گرفته است و نتایج حاصل ارائه شده است. شکل 4 مراحل آماده‌سازی این روش برای ارزیابی محیط‌های ابر را نشان می‌دهد.



شکل 4: مراحل آماده‌سازی روش ارزیابی عملکرد [20]

4. نتیجه‌گیری

ارزیابی عملکرد محیط‌های ابر یکی از جنبه‌های مهم محاسبات ابری است که موضوع مورد علاقه ارائه‌دهندگان و مشتریان خدمات ابر است. ارزیابی عملکرد ابر موجب آگاهی ارائه‌دهندگان خدمت از وضعیت ابر شده و در نتیجه به بهبود خدمات ابر کمک می‌کند. با توجه به اهمیت ویژه این بحث در محاسبات ابری نیاز به روش‌های متعددی برای این کار است اما به دلیل پیچیدگی محیط‌های ابری و دخالت معیارهای زیاد در عملکرد آن، ارزیابی عملکرد با دشواری‌هایی همراه است و به همین دلیل نیز روش‌های انگشت‌شماری در دنیا برای ارزیابی عملکرد محیط‌های ابری ارائه شده است. ما در این مقاله روش‌های ارزیابی عملکرد ابر را بررسی کردیم. در این بررسی سعی کردیم روش‌های متنوع را با توجه به تنوع ابرها تحلیل کنیم. نتایج تحلیل نشان می‌دهد به دلیل متنوع بودن اهداف توسعه ابرها و همچنین تفاوت در معماری آن‌ها، هر روش ارزیابی برای ابر مشخصی مناسب است و انتخاب روش، بستگی به هدفی دارد که ارائه‌دهندگان ابر دنبال رسیدن به آن هستند. همچنین هر روش نیاز به یک سری ابزار و معیارهای

¹¹ Grid

¹² Black box

مشخص دارد که باید با توجه به دسترسی ارائه‌دهندگان ابر به این ابزار و معیارها از روش مناسب استفاده شود. نتیجه همه این بررسی‌ها به‌طور کامل در جدول شماره 1 ارائه شده است.

جدول 1: مقایسه روش های مختلف ارزیابی عملکرد ابر

نام روش	مشخصات کلی	روش استفاده شده برای ارزیابی	معیارهای مورد استفاده برای ارزیابی	مناسب برای محیط‌های ابری
ارزیابی عملکرد خدمات ابری با بهینه کردن بهره ابر	1- به حداکثر رساندن بهره ابر 2- سازوکار ورود کارها FCFS	صف‌های اولویت	1- زمان خدمت 2- زمان پاسخ 3- زمان انتظار	قابل اولویت‌بندی کردن خدمات ابر، مدنظر بودن بالا بردن بهره ابر
ارزیابی عملکرد خدمات ابر با در نظر گرفتن بازیابی خطا	1- دخالت دادن قابلیت بازیابی خطا 2- مدل‌سازی ریاضی سیستم	مدل‌سازی ریاضی با استفاده از فاکتورهای مختلف	1- زمان پاسخ 2- زمان رسیدن درخواست مشتری 3- زمان انتظار 4- زمان خدمت 5- کل زمان اجرای خدمت 6- احتمال خطا برای گره	مدنظر داشتن قابلیت بازیابی خطا، داشتن قابلیت مدل‌سازی ریاضی سیستم، سیستم‌های ابری با پیچیدگی کم
ارزیابی کارایی رایانش ابری با استفاده از زنجیره مارکوف	1- تعیین تعداد کارسازهای رایانش ابری و اندازه میانگیرهای ورودی آنها 2- یک حل حالت پایدار از سیستم	صف معمولی $M/G/m/m+r$ ، مدل مارکوف	1- تابع توزیع تجمعی زمان رسیدن کارها 2- تابع چگالی احتمال ورود کارها در یک زمان مشخص 3- ماتریس گذر حالت	برای سیستم‌های منعطف از نظر مقیاس و تنوع زمان خدمت
ارزیابی عملکرد رایانش ابری برای خدمات محاسباتی پر تعداد	1- تعداد کارها بسیار زیاد 2- توجه به میزان اختلاف عرضه و تقاضا در ابر	شناسایی کاربرانی که حجم کاری زیادی به ابر اعمال می‌کنند، استفاده از روش محک‌زنی، ابزار تحلیل عملکردی GrenchMark	1- تعداد تبادلات کارها بین ابر و کاربران 2- استفاده از محک‌زنی و معیارهای مخصوص سیستم‌های با مقیاس بزرگ [13]	پردازش کارهای حجیم علمی، قابلیت شخصی‌سازی برای ابرهای مختلف
ارزیابی عملکرد ارائه‌دهندگان محاسبات ابری	1- ارزیابی عملکرد پیکربندی رایانش ابری 2- توجه به تک‌تک خدمات به‌جای بُن‌سازه کلی 3- جزء روش‌های تجربی	روش ASTAR	1- معیارهای روش ASTAR [21] به‌علاوه برخی معیارهای مختص محیط‌های ابری	ارزیابی عملکرد پیکربندی رایانش ابری، اهمیت داشتن تک‌تک خدمات به‌جای بُن‌سازه کلی

- [1] M. Armbrust, A. Fox et al. "A view of cloud computing." *Communications of the ACM*, 53(4), pp. 50-58, 2010.
- [2] D. Puthal, B. Sahoo, S. Mishra and S. Swain. "Cloud Computing Features, Issues and Challenges: A Big Picture" *International Conference on Computational Intelligence & Networks*, IEEE, pp. 116-123, 2015.
- [3] Bo Yang, Feng Tan, Yuan D and Suchang. "Performance Evaluation of Cloud Service Considering Fault Recovery" *First International Conference, CloudCom*, Springer Berlin Heidelberg, LNCS 5931, pp. 571-576, 2009.
- [4] Sasikala, P. "Research challenges and potential green technological applications in cloud computing." *International Journal of Cloud Computing*, Online ISSN: 2043-9997, 2(1), pp. 1-19, 2013.
- [5] C. Vecchiola, S. Pandey and R. Buyya, "High Performance Cloud Computing A View of Scientific Applications", *10th International Symposium On Pervasive Systems*, IEEE, pp. 4-16, 2009.
- [6] Xu, X. "From cloud computing to cloud manufacturing", *Robotics and computer-integrated manufacturing*, 28(1), pp. 75-86, 2012.
- [7] Dikaiakos, M. D., Katsaros, D., Mehra, P., Pallis, G., and Vakali, "A. Cloud computing: Distributed internet computing for IT and scientific research", *Internet Computing*, IEEE, 13(5), pp. 10-13, 2009.
- [8] T. Atmaca, T. Begin, A. Brandwajn and H. Castel. "Performance Evaluation of Cloud Computing Centers with General Arrivals and Service", *IEEE Transactions on Parallel and Distributed Systems*, 2015.
- [9] M. Jaiganesh, B. Ramadoss, A. Vincent and S. Mercy, "Performance Evaluation of Cloud Services with Profit Optimization", *Procedia Computer Science*, vol. 54, pp. 24-30, 2015.
- [10] Armbrust, M., Fox, A., Griffith, R., et al, "Above the clouds: A Berkeley view of cloud computing" *Technical Report No. UCB/EECS-2009-28*, University of California at Berkeley, 2009.
- [11] H. Khazaei, J. Mistic, and V. Mistic, "Performance Analysis of Cloud Computing Centers Using M/G/m/m+r Queuing Systems", *IEEE Transactions on Parallel and Distributed Systems*, 23(5), pp. 936-943, 2012.
- [12] G. Grimmett and D. Stirzaker, "Probability and Random Processes", third ed, Oxford University Press, July 2010.
- [13] A. Losup, S. Ostermann and H.j. Epema, "Performance Analysis of Cloud Computing Services for Many-Tasks Scientific Computing", *IEEE Transactions on Parallel and Distributed Systems*, 22(6), pp. 931-945, 2011.
- [14] I. Raicu, Z. Zhang, M. Wilde, I.T. Foster, P.H. Beckman, K. Iskra, and B. Clifford, "Toward Loosely Coupled Programming on Petascale Systems", *Proceedings of the 2008 ACM/IEEE conference on Supercomputing*, pp. 22, 2008.
- [15] A. Iosup, H. Li, M. Jan, S. Anoep, C. Dumitrescu, L. Wolters, and D. Epema, "The Grid Workloads Archive," *Future Generation Computer Systems*, 24(7), pp. 672-686, 2008.
- [16] The Parallel Workloads Archive Team, "The Parallel Workloads Archive Logs," <http://www.cs.huji.ac.il/labs/parallel/workload/logs.html>, Jan. 2009.
- [17] R.H. Saavedra, A.J. Smith, "Analysis of Benchmark Characteristics and Benchmark Performance Prediction," *ACM Trans. Computer Systems*, 14(4), pp. 344-384, 1996.
- [18] A. Iosup and D.H.J. Epema, "GrenchMark: A Framework for Analyzing, Testing, and Comparing Grids," *Sixth IEEE International Symposium on Cluster Computing and the Grid*, pp. 313-320, 2006.
- [19] N. Yigitbasi, A. Iosup, S. Ostermann, and D. Epema, "C-Meter: A Framework for Performance Analysis of Computing Clouds," *Proceedings of the 2009 9th IEEE/ACM International Symposium on Cluster Computing and the Grid*, pp. 472-477, 2009.
- [20] V. Stantchev, "Performance Evaluation of Cloud Computing Offerings", *Third International Conference on Advanced Engineering Computing and Applications in Sciences*, IEEE, pp. 187-192, 2009
- [21] V. Stantchev and C. Schröpfer, "Techniques for service level enforcement in web-services based systems," *Proceedings of 10th International Conference on Information Integration and Web-based Applications and Services*. New York, NY, USA: ACM, pp. 7-14, 2008.
- [22] H. Krallmann, C. Schröpfer, V. Stantchev, and P. Offermann, "Enabling autonomous self-optimization in service-oriented systems," *Proceedings of 8th International Workshop on Autonomous Systems – Self Organisation, Management and Control*, Springer, pp. 127-134, 2008.
- [23] V. Stantchev and C. Schröpfer, "Negotiating and enforcing qos and slas in grid and cloud computing," *Advances in Grid and Pervasive Computing*, ser. *Lecture Notes in Computer Science*, vol. 5529. Springer, pp. 25-35, 2009.
- [24] G. Alonso, F. Casati, H. Kuno, and V. Machiraju, "Web services: concepts, architectures and applications." Springer Berlin Heidelberg, 2004.

[25] Microsoft, Comparing Web Service Performance: WS Test 1.1 Benchmark Results for .NET 2.0, .NET 1.1, Sun One/ JWS DP 1.5 and IBM WebSphere 6.0. <http://www.theserverside.net/tt/articles/content/NET2Benchmarks>, 2006.