

یک مدل ارزیابی عملکرد کارآمد برای خوشه های منابع در محیط ابری با استفاده از زنجیره مارکوف با زمان پیوسته و فرایند پواسن

چکیده

رایانش ابری به ما این امکان را می دهد تا برای پردازش کارها، کارها میان مراکز داده ابری مختلف را با اینترنت به عنوان ستون فقرات به اشتراک بگذاریم. انتخاب مراکز داده جهت پردازش کارها، مبتنی بر پتانسیل مراکز داده، نرخ ورود کارها و بهره برداری یا استفاده بهتر از منابع می باشد. مدلسازی عملکرد و ارزیابی سیستم های ابری، تصویری کامل از احتمال توزیع کارها میان مراکز داده، تعداد متوسط کارها در مراکز داده و زمان کل محیط ابری نشان می دهد. برای ارائه بهترین مدل ارزیابی عملکرد که به محیط ابری مساعد و موفق منتج می گردد، شیوه پیشنهادی نرخ ورود کارها را با استفاده از فرایند پواسن محاسبه می کند که نرخ ورود برای فواصل زمانی نامتناهی محاسبه شده و تعداد متوسط کارها و همچنین بهره برداری از منابع با استفاده از زنجیره مارکوف پیوسته مدلسازی شده اند. برای بهبود زمان بندی کارها با به حداقل رساندن زمان کل (زمان صرف شده) محیط ابری، برنامه های مرتبط با محاسبه نرخ ورود و مدل بهره برداری از منابع شیوه های موجود حل شده اند.

واژگان کلیدی: رایانش ابری، زمان پاسخ، بهره برداری از منابع، ارزیابی عملکرد، فرایند پواسن، زنجیره مارکوف

1. مقدمه

رایانش ابری، به تامین کننده اجازه صدور و انتشار منابع محاسباتی نظیر فناوریهای ذخیره سازی، نرم افزار و پلتفرم ها به عنوان سرویس مشترک برای کاربران به صورت اجاره ای و طبق تقاضا را می دهد. مفهوم مجازی سازی در اساس رایانش ابری نقش حیاتی ایفا نموده و امکان تشکیل گروههایی از منابع محاسباتی از خوشه های سرور و تخصیص یا تخصیص مجدد منابع محاسباتی به اپلیکیشن به شیوه طبق تقاضا را فراهم می آورد [1]. یکی از چالش های اصلی در رایانش ابری، به حداقل رساندن زمان کل محیط ابری، با زمان بندی کارآمد منابع موجود برای کارهای ورودی می باشد. برخی از پارامترها نظیر نرخ ورود کارها، زمان پاسخ و زمان انتظار کار در خوشه های منابع، مورد توجه قرار گرفته و در عین حال، کارها برای منابع موجود در محیط ابری، زمان بندی شده اند. در این مقاله، برخی از مشکلات مشاهده شده با شیوه های موجود در رابطه با نرخ ورود و بهره برداری از دستگاه بررسی و حل شده است.

زمان اجرا یا تکمیل کارها در محیط ابری خاص، زمان کل (صرف شده) آن محیط ابری نامیده شده است [2]. منابع براساس پتانسیلشان خوشه بندی شده اند که این امر منجر به یک زمان بندی کارآمد در محیط ابری می گردد. در شیوه پیشنهادی، منابع با محاسبه پتانسیل هر منبع و با استفاده از الگوریتم $EPRNW^1$ خوشه بندی شده اند [3]. هدف اصلی شیوه پیشنهادی، توسعه یک مدل عملکرد و ارزیابی مناسب برای پارامترهایی نظیر نرخ ورود، نرخ تکمیل و بهره برداری از منابع محیط ابری است. زمان سرویس یا خدمات کارها عمدتاً به نرخ ورود کارها در منابع محاسباتی بستگی دارد. در شیوه پیشنهادی به دو دلیل نرخ ورود با فرایند پواسن محاسبه شده است: یکی اینکه فرایند پواسن از خصوصیات زنجیره مارکوف پیروی می کند که از لحاظ سیستماتیکی قابل مدیریت است و دلیل دوم اینکه نرخ ورود کارها را می توان در فواصل زمانی مستقل محاسبه نمود و هیچ حدی برای فواصل زمانی وجود ندارد [4].

بهره برداری از منابع با تعیین تعداد متوسط کارهای مستقر در خوشه های منابع، محاسبه می شود. از مدل زنجیره مارکوف برای تعیین احتمال تعداد متوسط کارهای مستقر در خوشه های منبع استفاده شده و تضمین شده است که

¹ . equivalence partitioning recurrent node weight

زمان های اجرا در خوشه منبع به صورت نمایی توزیع شده و نرخ ورود پیوسته کارها نیز از توزیع نمایی پیروی می کند. زنجیره مارکوف به دو نوع تقسیم شده است، زنجیره مارکوف با زمان گسسته و زنجیره مارکوف با زمان پیوسته. رویدادها تنها در پایان گام زمانی زنجیره مارکوف با زمان پیوسته، رخ می دهند، در صورتی که در زنجیره مارکوف با زمان پیوسته، رخداد رویدادها به فواصل زمانی مستقل بستگی دارد [5]. در شیوه پیشنهادی از زنجیره مارکوف بازمان پیوسته برای تعیین احتمال بهره برداری از منابع با محاسبه نرخ ورود کارها در فواصل زمانی مستقل با استفاده از فرایند پواسن استفاده شده است. برای تولید یک مدل ارزیابی عملکرد کارآمد، مقاله حاضر، بررسی معایب شیوه های موجود غلبه کرده و راه حل نزدیک به حد بهینه ای حاصل می نماید.

2. کارهای مرتبط

پژوهشهای زیادی برای مدل ارزیابی عملکرد در مورد زمان بندی کارها روی محیط ابری وجود دارد. برای نیل به کارایی بالا، ارزیابی عملکرد مزارع منابع ابری، محاسبه توزیع تقریبی کارها و تعداد متوسط کارها در سیستم، یک مدل ریاضی پیاده گردید [6]. زمان بندی صف بازخورد چند سطحی توسط مدل زنجیره مارکوف و تحت برخی سیاست های زمان بندی مشترک بررسی گردید تا بدین طریق زمان پردازش CPU تعیین گردد [7]. از تحلیل زنجیره مارکوف برای ارزیابی Drum Buffer Rope و در کنار شیوه های متداول دیگر برای نشان دادن عملکرد بهبودیافته Drum Buffer Rope در محیط فروشگاه Job استفاده شده است [8]. مسئله پارتیشن بندی چند مکانی در ابرهای متحرک (موبایل) بررسی و با استفاده از طرح آفلودچندمکانی انرژی کارآمد، حل و برای دستیابی به مینیوم انرژی مصرفی ابرهای متحرک، با فرایند تصمیم گیری مارکوف مورد تجزیه و تحلیل قرار گرفت [9]. توان فوتوولتائیک گریدها (شبکه ها) با فرایند نیمه مارکوف نمایش داده شد تا بدین طریق توان فوتوولتائیک مطمئنی تولید و واحدهای ذخیره سازی انرژی به شیوه ای تصادفی و همزمان با منابع دیگر، زمان بندی شوند [10]. نتایج شبیه سازی الگوریتم LARAC توجیه می کند که طرح climb offloading برای کانال احتمالی مارکوفی مناسب بود و انرژی مصرفی دستگاههای موبایل را کاهش می دهد [11]. در ابرهای موبایل (متحرک)، با تعیین

خصوصیت فراریت دستگاههای موبایل و پیش بینی حالتها با کمک مدل زنجیره مارکوف ، تحمل پذیری خطا حاصل می گردد [12]. با پیش بینی پهنای باند آتی کلیه ماشین های مجازی با کمک دانش پیشین راجع به حجم کار پهنای باند نیز زمان بندی کارآمدی از ماشین های مجازی در ابر، حاصل می گردد که این امر منجر به تغییرقدرتمندی در مدل مارکوف پنهان در راستای مدیریت سیستم تله هلت یا سلامت از راه دور می گردد [13].

یکی از عوامل مهم برای شبیه سازیها و ارزیابی عملکرد، تولید حجم کار است. یک مولد حجم کار به نام BURSE به خاطر خصوصیات خودتشابهی و انفجاری بودن حجم کار در رایانش ابری مبتنی بر طبیعت و رفتار دو حالت فرایندهای پواسن مدوله شده مارکوف، پیشنهاد گردید [14]. دو پارامتر مهم آفلود یا تخلیه ابری موبایل عبارتند از: مصرف انرژی اکسترم موبایل و زمان پاسخ روی داده با کاردستگاه موبایل. پارامترهایی نظیر جمع وزندار زمان پاسخ انرژی، حاصل ضرب زمان پاسخ انرژی و حاصل ضرب وزندار زمان پاسخ انرژی، بررسی و در استراتژیهای آفلود ابرموبایل استفاده شده اند تا بدین طریق مدل ارزیابی عملکرد به شیوه ای موثر، به نتیجه برسد [15]. زمان بندی کارها و پتانسیل حجم کارهای ناظر کانال پنهان ماشین مجازی متقابل با استفاده از مدل مارکوف پیوسته ارزیابی گردید. از تخصیص کارآمد ماشین مجازی برای به حداقل رساندن نرخ خطا در کانال مذکور استفاده گردید [16].

از مفهوم یادگیری تقویتی برای زمان بندی وظایف و از مفهوم نظریه صف بندی برای بهینه سازی زمان بندی وظایف براساس محدودیت های منابع استفاده گردید که نتیجه این امر نیز افزایش رابطه بین نرخ ورود، بهره برداری از منابع و اندازه ذخیره سازی بود [17]. منابع به شکلی موثر و با پیش بینی مصرف CPU ماشین مجازی و همچنین با توجه به رفتارهای بلندمدت و کوتاه مدت VM با پیاده سازی نمای Hurst تخصیص داده شده اند که یکی از همراهان مدل انتقال مارکوف به حساب می آید [18]. از فرایند تصمیم گیری مارکوفی به عنوان چارچوبی برای نیل به راه حل نزدیک بهینه جهت مسئله آفلود (تخلیه) دینامیکی در رایانش ابری موبایل استفاده گردید [19]. شیوه زمان بندی گره ها از تکنیک پیش بینی زنجیره مارکوف برای تجزیه و تحلیل داده های جریان سازی بزرگ و تضمین پردازش موثر آنها استفاده نمود [20]. روشهای پژوهشهای فوق برای ارزیابی عملکرد خوشه های منابع در رایانش ابری در صورت نرخ ورود بی نهایت و تعداد بزرگ خوشه های منابع کاملاً مناسب نیستند. به طور کلی در چنین

موقعیت هایی، برخی مدل‌های توزیع احتمال نظیر فرایند پواسن برای محاسبه نرخ ورود و زنجیره مارکوف برای برآورد بهره برداری از منابع و همچنین تعداد متوسط کارها در خوشه منبع، ارائه می گردد. هدف پنهان این کار، زمان بندی کارها موثر در محیط ابری را تعیین می نماید. از یک شیوه فراابتکاری بهبودیافته برای بهبود زمان بندی کارها در محیط ابری استفاده گردید [21]. از روش توزیع احتمال برای بهبود زمان بندی کارها در محیط ابری استفاده شده بود [22].

3. روش پیشنهادی

محیط ابری محتوی سرورهای زیادی است که n تعداد از کارها را برای خوشه های منابع موجود، زمان بندی می کنند. هر خوشه محتوی ترتیبی میان کارها است که از آن برای زمان بندی استفاده نموده و با استفاده از الگوریتم EPRNW بهترین خوشه منبع را انتخاب نموده و عملیاتی بر روی کارهایش اجرا می نماید [1]. در این الگوریتم، پتانسیل خوشه های منبع براساس سه گام محاسبه شده است. گام اول تخصیص وزن به گره ها و ایجاد یک هویت برای آن گره ها است. گام دوم، برقراری اتصال بین گره ها در محیط برای یک ارتباط موثر میان آنها است. گام سوم طبق روش الگوریتم، جدا نمودن شاه منابع از منابع دیگر است. زمان کل یا صرف شده، یک پارامتر مهم در محیط ابری به حساب می آید که کارها را زمان بندی می نماید. زمان کل به صورت زمان تکمیل کل کارها برای یک محیط ابری خاص تعریف شده و عمدتاً بر زمان ورود کار، زمان پاسخ کارها و همچنین بهره برداری از دستگاه تمرکز می کند. زمان کل محیط ابری به صورت زمان بین ورود کارها به خوشه منبع و خروج از آن مشتمل بر کارهای ورودی به همان خوشه منبع و همچنین خوشه منبع متفاوت محاسبه شده است. کلیه خوشه های منابع از طریق رسانه شبکه به هم متصل شده اند تا بدین طریق از تسهیلات همسایگی میان خوشه ها اطمینان حاصل گردد. سرورها از مرتب سازی متوالی (ترتیبی) برای خدمت کردن به کارها استفاده می کنند. هر خوشه منبع در محیط ابری، کارها را در داخل یا خارج می پذیرد. در شیوه های موجود، انتخاب خوشه های منبع مبتنی بر شیوه احتمالی است که منجر به یک راه حل ناکارآمد می گردد. برای بدست آوردن راه حل نزدیک بهینه، بهترین خوشه های منبع با محاسبه

پتانسیل منبع و با استفاده از الگوریتم معرفی شده، انتخاب می شوند. کارهای روی خوشه منبع A به خوشه B با پتانسیل خوشه CP_{AB} حمل شده یا از خوشه منبع خارج می شوند به شرطی که پتانسیل خوشه منبع کارهایی با پتانسیل $CP_{A,out} = 1 - \sum_B CP_{AB}$ را ارضا نماید.

نرخ ورود کل به خوشه منبع A، جمع نرخ کارهای ورودی از خارج به خوشه منبع A و نرخ کارهای ورودی از داخل محیط ابری به خوشه منبع A می باشد.

$$\text{ورود کل } (\lambda_A) = \text{ورود خارجی } (e_A) + \sum_B \lambda_B CP_{BA}$$

در همان زمان، برای تعیین λ_A کل، کلیه پارامترهای معادلات حل شده است. همچنین معادله بالا را می توان به صورت زیر نوشت

$$\lambda_A (1 - CP_{AA}) = e_A + \sum_{B \neq A} \lambda_B CP_{BA}$$

هر دو معادله مساوی هستند، اما انتقال کار از خوشه منبع A مجدداً به خوشه منبع A حذف گردد.

3.1 فرایند ورود در هر خوشه منبع

محاسبه λ_A به فرایند پواسن بستگی دارد، زیرا فرایند ورود، یک فرایند پواسن است که به توزیع کارهای مختلف در میان خوشه های منبع کمک می کند. به طور مثال، 'Z' یک متغیر تصادفی است که ورود تعدادی کار به خوشه منبع در فاصله زمانی خاص را نشان می دهد. بعضاً هیچ کاری وارد خوشه های منبع نمی شود. در این رابطه، احتمال ورود کارها و همچنین احتمال عدم ورود کارها، با کمک توزیع احتمال تعیین شده است. در یک خوشه منبع خاص، دو موقعیت بروز می کند. یکی ورود کارهای زیاد به خوشه منبع و موقعیت دیگر عدم ورود کارها به خوشه منبع در فاصله زمانی خاص است. خوشه های منبع حتی در زمان عدم ورود کارها در حالت بیکار قرار نخواهند داشت. در آن زمان، کارهای کارهای منابع دیگر با خوشه های منبع تهی به اشتراک گذاشته شده و بدین ترتیب کارها

پردازش می شوند. فواصل زمانی ورود کارها متفاوت بوده و در میان خوشه های منبعی که منجر به فرایند پواسن می گردند، مستقل می باشند. از فرایند پواسن برای تعیین ارزش یا مقدار مورد انتظار (E) 'Z' در فاصله زمانی پیوسته با برآورد تعداد متوسط کارهای ورودی (λ_j) به خوشه منبع استفاده شده است. مفهوم توزیع دو جمله ای نشان می دهد ارزش مورد انتظار (E) متغیر تصادفی، برابر است با تعداد آزمایشاتی که در آنها متغیر تصادفی پیدا شده و احتمال موفقیت آزمایشات. توزیع دوجمله ای مطابق مسئله خوشه منبع پیشنهادی با تعیین تعداد آزمایشات به صورت پردازش در فواصل زمانی مستقل و با تعیین احتمال موفقیت به صورت وضعیت ورود و عدم ورود کارها به خوشه منبع، اصلاح شده است.

$$E(Z) = \lambda_j = Q_t \times P_{AJ,NJ}$$

که λ_j - ورود کارها، Q- تعداد آزمایشات (فواصل زمانی مستقل)، AJ- کارهای ورودی و NJ- کارهای غیر ورودی.

به طور مثال، برآورد تعداد کارهای ورودی در هر ساعت منجر به تعدادی آزمایش در هر ساعت (60 دقیقه = 1 ساعت، بنابراین، 60 آزمایش) با احتمال موفقیت در هر ساعت می گردد.

$$\lambda_{j/h} = 60_{\text{min/h}} \times \lambda / 60_{j/\text{min}}$$

$$P(Z = K) = \left[\frac{60}{K} \right] \left[\frac{\lambda}{60} \right]^K \left[1 - \frac{\lambda}{60} \right]^{60-K}$$

(60 - k → no job arrival)

آزمایشات را برای یک ثانیه نیز می توان انجام داد زیرا احتمال ورود بیش از یک کار در دقیقه وجود دارد. بنابراین، برآورد باید با رابطه زیر انجام شود

$$P(Z = K) = \left[\frac{3600}{K} \right] \left[\frac{\lambda}{3600} \right]^K \left[1 - \frac{\lambda}{3600} \right]^{3600-K}$$

(3600 → 3600s/min)

آزمایشات را برای یک میلی ثانیه نیز می توان انجام داد، زیرا احتمال ورود بیش از یک کار در دقیقه وجود دارد. بنابراین، برآورد باید با رابطه زیر انجام شود

$$P(Z = K) = \left[\frac{3,600,000}{K} \right] \left[\frac{\lambda}{3,600,000} \right]^K \times \left[1 - \frac{\lambda}{3,600,000} \right]^{3,600,000-K}$$

(3,600,000 → 3,600,000 ms/h)

تعداد نامحدودی از آزمایشات در فواصل زمانی پیوسته انجام شده است، زیرا هیچ حدی برای نرخ ورود، وجود ندارد. در زنجیره مارکوف پیوسته، نمی توان حدی برای ورود کارها لحاظ نمود. اگر قصد چنین کاری را دارید، باید به فرایند مارکوف با زمان گسسته تبدیل شود. این مسئله منجر به بیان مجدد معادله فوق به صورت زیر می گردد:

$$P(Z = K) = \lim_{Q \rightarrow \infty} \left[\frac{Q}{K} \right] \left[\frac{\lambda}{Q} \right]^K \left[1 - \frac{\lambda}{Q} \right]^{Q-K}$$

$$\lim_{Q \rightarrow \infty} \frac{Q!}{(Q-K)!K!} \left[\frac{\lambda^K}{Q^K} \right] \left[1 - \frac{\lambda}{Q} \right]^Q \left[1 - \frac{\lambda}{Q} \right]^{-K}$$

$$\lim_{Q \rightarrow \infty} \left[\frac{Q(Q-1)(Q-2)\dots(Q-K+1)}{Q^K} \right] \left[\frac{\lambda^K}{Q^K} \right]$$

$$\left[1 - \frac{\lambda}{Q} \right]^Q \left[1 - \frac{\lambda}{Q} \right]^{-K}$$

$$\lim_{Z \rightarrow a} B(Z) C(Z) = \lim_{Z \rightarrow a} B(Z) \lim_{Z \rightarrow a} C(Z)$$

$$\lim_{Q \rightarrow \infty} \frac{Q^K + \dots + \lambda^K}{Q^K} \lim_{Q \rightarrow \infty} \left[1 - \frac{\lambda}{Q} \right]^Q \left[1 - \frac{\lambda}{Q} \right]^{-K}$$

$$= 1 \times \frac{\lambda^K}{K!} e^{-\lambda} \times 1$$

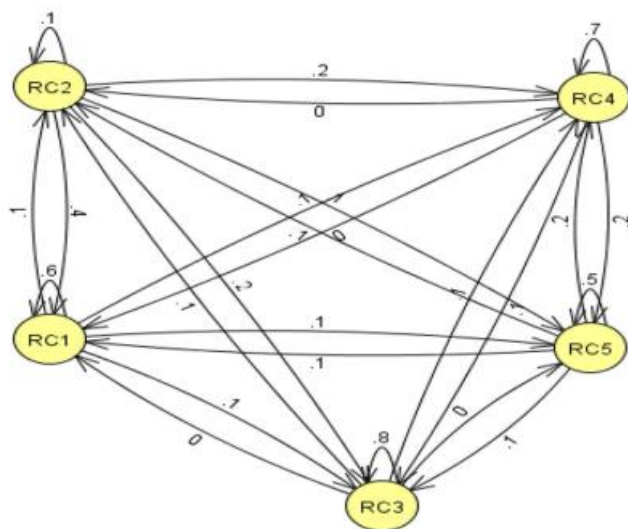
$$P(Z = K) = \lim_{Q \rightarrow \infty} \left[\frac{Q}{k} \right] \left[\frac{\lambda}{Q} \right]^K \left[1 - \frac{\lambda}{Q} \right]^{Qk} = \frac{\lambda^K}{K!} e^{-\lambda}$$

طبق معادله فوق، فرایند پواسن نرخ ورود کامل را با یافتن تعداد زیادی آزمایش به صورت فرایند پیوسته نشان داده و گرانولارپته یا دانه بندی مورد نیاز بدست آمده است.

3.2 توزیع کارها میان خوشه های منبع

زنجیره مارکوف ابزاری عالی برای تعیین توزیع کارها میان خوشه های منبع به حساب می آید به شرطی که زمان های خدمات و زمان های بین ورود کارها از توزیع نمایی پیروی کنند. بیاپید پنج خوشه منبع RC1, RC2, RC3

RC4 و RC5 را برای سیستم پیشنهادی در نظر بگیریم، که توسط زنجیره مارکوف مدلسازی شده اند و بدین طریق احتمال توزیع کارها را تعیین نماییم. مدل زنجیره مارکوف پیوسته، یک مدل ارزیابی برای مدیریت تعداد زیادی از کارها با محاسبه زمان ورود در زمان نامتناهی با استفاده از فرایند پواسن ایجاد می کند (2 در 2). در مقایسه با فرایند زنجیره مارکوف با زمان گسسته، فرایند زنجیره مارکوف با زمان پیوسته، در فواصل زمانی نامتناهی نیز عملکرد خوبی به معرض نمایش می گذارد، زیرا در زنجیره مارکوف با زمان گسسته، حد زمانی را باید تعیین نمود. شکل، نمودار انتقال از خوشه منبع مدلسازی شده با زنجیره مارکوف را نشان می دهد که توزیع کارها میان خوشه های منبع را به تصویر کشیده و احتمال تعداد متوسط کارها در هر خوشه منبع را مشخص می نماید. کارها از داخل یا خارج وارد RC1 می شوند، 60 درصد از کارها در RC1 مستقر بوده و 40 درصد باقیمانده در میان خوشه های منبع دیگر نظیر RC2، RC3، RC4 و RC5 پراکنده شده اند. کارهای مستقر در خوشه منبع براساس محاسبه پتانسیل منبع با استفاده از الگوریتم EPRNW بدست آمده اند. 10 درصد از کارها در RC2 مستقر هستند که از داخل یا خارج وارد RC2 شده اند و 90 درصد باقیمانده از کارها در میان خوشه های منبع پراکنده شده اند. برای خوشه های منبع باقیمانده، از همین خصوصیت مارکوف استفاده شده است. ماتریس های احتمال انتقال نمایش روشن توزیع احتمال کارها و تعداد متوسط کارها در خوشه منبع را نشان می دهند.



ماتریس احتمال انتقال

$$P = \begin{matrix} & \text{RC1} & \text{RC2} & \text{RC3} & \text{RC4} & \text{RC5} \\ \text{RC1} & \begin{pmatrix} 0.6 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.4 & 0.1 & 0.2 & 0.2 & 0.1 \\ 0 & 0.1 & 0.8 & 0.1 & 0 \\ 0 & 0 & 0.1 & 0.7 & 0.2 \\ 0.1 & 0.1 & 0.1 & 0.2 & 0.5 \end{pmatrix} \end{matrix}$$

ماتریس توزیع حالت اولیه برای خوشه های منبع. در اینجا از ماتریس توزیع حالت برای تعیین درصد یا احتمال بهره برداری محیط ابری از دستگاه یا منبع در توزیع کارها استفاده شده است.

RC1 دیگر

$$SDC1_0 = [0.3 \ 0.7]$$

ماتریس توزیع حالت اولیه برای RC1 می گوید که RC1 محتوی 30 درصد از کارها است و 70 درصد باقیمانده از کارها در میان خوشه های منبع دیگر پراکنده شده اند.

$$SDC2_0 = [0.2 \ 0.8]$$

ماتریس توزیع حالت اولیه برای RC2 می گوید که RC2 محتوی 20 درصد از کارها است و 80 درصد باقیمانده از کارها در میان خوشه های منبع دیگر پراکنده شده اند.

RC3 دیگر

$$SDC3_0 = [0.4 \ 0.6]$$

ماتریس توزیع حالت اولیه برای RC3 می گوید که RC3 محتوی 40 درصد از کارها است و 60 درصد باقیمانده از کارها در میان خوشه های منبع دیگر پراکنده شده اند.

RC4 دیگر

$$SDC4_0 = [0.6 \ 0.4]$$

ماتریس توزیع حالت اولیه برای RC4 می گوید که RC4 محتوی 60 درصد از کارها است و 40 درصد باقیمانده از کارها در میان خوشه های منبع دیگر پراکنده شده اند.

RC5 دیگر

$$SDC5_0 = [0.3 \ 0.7]$$

ماتریس توزیع حالت اولیه برای RC5 می گوید که 30 درصد از کارها در RC5 مستقر هستند و 70 درصد باقیمانده از کارها در میان خوشه های منبع دیگر پراکنده شده اند. با استفاده از این ماتریس های توزیع حالت، احتمال توزیع کارها میان خوشه های منبع در فاصله زمانی مستقل محاسبه می گردد.

احتمال کارهای توزیع شده در RC1

$$\begin{aligned} &= (0.3)(0.6) + (0.7)(0.5) \\ &= 0.18 + 0.35 = 0.53 \end{aligned}$$

بنابراین، ماتریس انتقال حالت اولیه برای RC1 در فاصله زمانی مستقل $SDC1_1 = [0.53 \ 0.47]$

می باشد. این رابطه نشان می دهد که 53 درصد از کارها از کارهای ورودی در RC1 مستقر بوده و 37 درصد از کارها در میان خوشه های منبع دیگر پراکنده خواهند شد.

احتمال کارهای توزیع شده در RC2

$$\begin{aligned} &= (0.2)(0.4) + (0.8)(0.7) \\ &= 0.8 + 0.56 = 0.64 \end{aligned}$$

بنابراین، ماتریس انتقال حالت اولیه برای RC1 در فاصله زمانی مستقل $SDC1_2 = [0.64 \ 0.36]$

است. این رابطه نشان می دهد که 64 درصد از کارهای ورودی در RC2 مستقر بوده و 36 درصد از کارها در میان خوشه های منبع دیگر پراکنده خواهند شد.

احتمال کارهای توزیع شده در RC3

$$\begin{aligned} &= (0.4)(0) + (0.6)(0.11) \\ &= 0 + 0.66 = 0.66 \end{aligned}$$

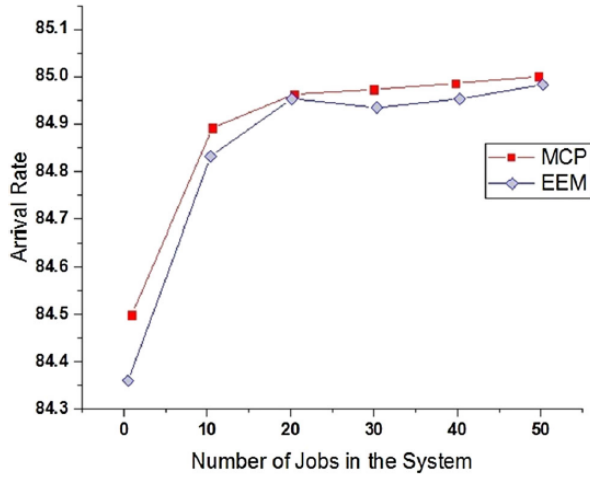
بنابراین، ماتریس انتقال حالت اولیه برای RC1 در فاصله زمانی مستقل $SDC1_3 = [0.66 \ 0.34]$ است. این رابطه نشان می دهد که 66 درصد از کارهای ورودی در RC2 مستقر است و 34 درصد از کارها در میان خوشه های منبع دیگر پراکنده خواهد شد.

4. نتایج و بحث

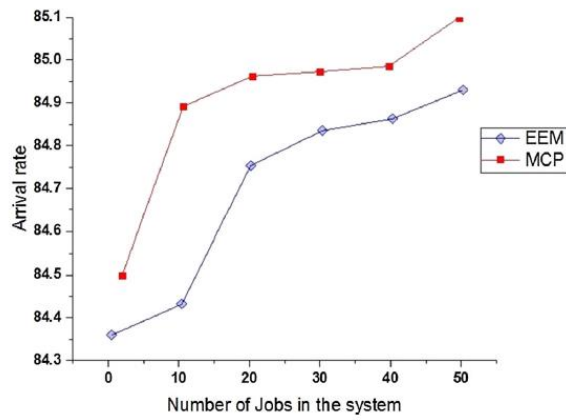
گراف های آزمایشی زیر ارزیابی عملکرد مدل برآورد موجود و زنجیره مارکوف پیشنهادی در کنار مدل پواسن را به تصویر می کشند. در سیستم پیشنهادی، از تعداد متفاوتی از خوشه های منبع (RC=50,100) برای اجرای وظیفه در محیط ابری استفاده شده است. اشکال 1 و 2 محاسبه نرخ ورود کارها در 50 و 100 خوشه منبع با مدل برآورد شده موجود (EEM) و مدل پواسن زنجیره مارکوف (MCP) را نشان می دهند. مدل MCP در مقایسه با EEM، تعداد بیشتری از کارهای ورودی را تعیین و زمان بندی می نماید. درصد احتمال نرخ ورود مقادیر نمایی محاسبه و جدول مقایسه ثابت می کند که MCP تعداد بیشتری از ورودها را حتی در زمان نمایی تعیین و درصد احتمال تعیین نرخ ورود کارها در MCP افزایش یافته است.

EEM				MCP			
Percentage of arrival rates	Expected exponential percentage	Lower percentage	Upper percentage	Percentage of arrival rates	Expected exponential percentage	Lower percentage	Upper percentage
84.36	14.28571	0.38242	1.89471	84.4994	13.08771	1.58563	7.85604
84.4319	28.57143	0.76872	3.80865	84.8917	28.56718	3.17125	15.71209
84.7533	42.85714	1.15898	5.74222	84.9627	47.51252	4.75688	23.56813
84.8346	57.14286	1.55329	7.69583	84.9727	71.93731	6.3425	31.42418
84.8627	71.42857	1.95173	9.6699	84.9857	106.36212	7.92813	39.28022
84.93	85.71429	2.35438	11.66486	85.1	165.21173	9.51376	47.13626

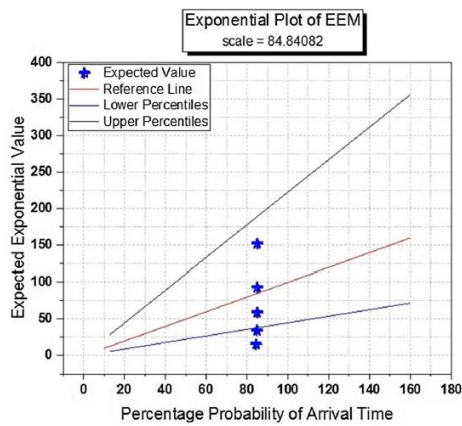
اشکال 3 و 4 محاسبه درصد احتمال نرخ ورود کارها در 50 و 100 خوشه منبع با EEM و MCP را نشان می دهند.



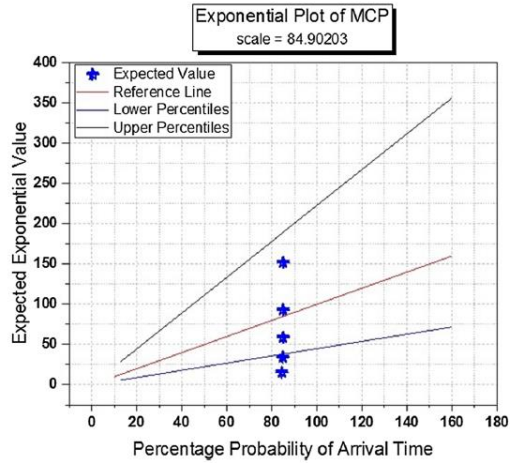
شکل 1 RC=50



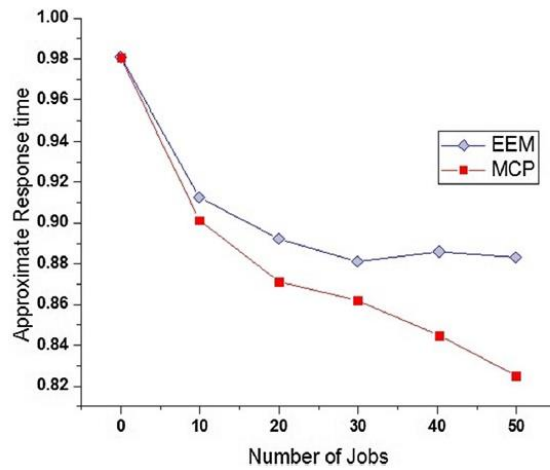
شکل 2 RC=100



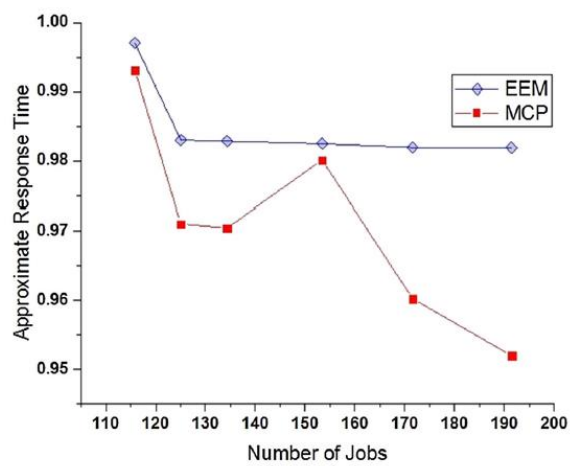
شکل 3. درصد احتمال نرخ ورود در RC=50



شکل 4. احتمال نرخ ورود در $RC=100$



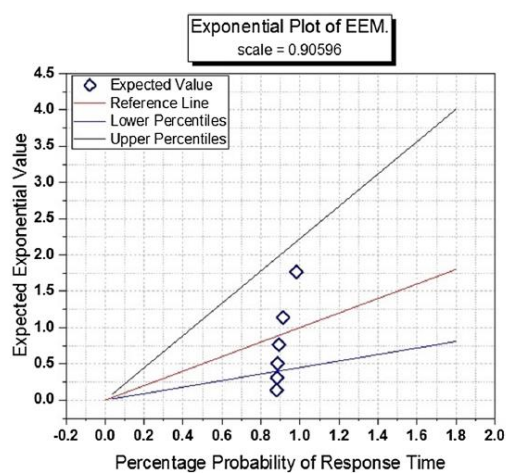
شکل 5. محاسبه زمان پاسخ در $50RC$



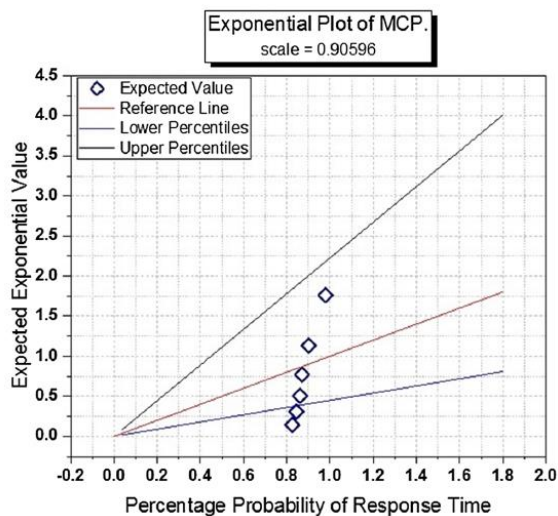
شکل 6. محاسبه زمان پاسخ در $100RC$

اشکال 5 و 6 محاسبه زمان پاسخ کارها در 50 و 100 خوشه منبع با EEM و MCP را نشان می دهند. مدل MCP زمان پاسخ کارها را به حداقل رسانده، تعداد متوسط کارهای موجود در محیط ابری را تعیین و بهره برداری بهتر از منابع را تضمین می نماید. با به حداقل رساندن زمان پاسخ، زمان کل یا صرف شده محیط ابری کاهش یافته و در عین حال، کارها در میان خوشه های منبع در محیط ابری، زمان بندی می شوند.

اشکال 7 و 8 محاسبه درصد احتمال زمان پاسخ کارها در 50 و 100 خوشه منبع با EEM و MCP را نشان می دهند.



شکل 7. درصد احتمال زمان پاسخ در $RC=50$

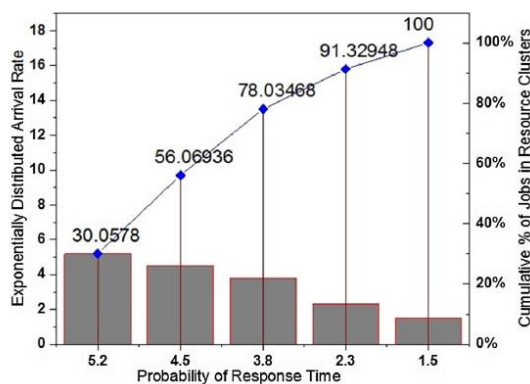


شکل 8. درصد احتمال زمان پاسخ در $RC=50$

درصد احتمال زمان پاسخ برای مقادیر نمایی محاسبه و جدول مقایسه ثابت می کند که MCP زمان پاسخ کارها را حتی در زمان نمایی به حداقل رسانده و درصد احتمال کارهای نمایی در MCP در مقایسه با EEM افزایش یافته است.

EEM				MCP			
Response time	Expected exponential percentage	Lower percentage	Upper percentage	Percentage of arrival rates	Expected exponential percentage	Lower percentage	Upper percentage
0.88106	10.9375	0.00409	0.02027	0.95199	10.9886	0.91256	0.96234
0.88319	26.5625	0.00822	0.04074	0.96024	26.8765	0.91876	0.96404
0.88593	42.1875	0.0124	0.06142	0.97044	42.6328	0.92266	0.96515
0.8922	57.8125	0.01662	0.08232	0.97103	58.4231	0.92556	0.96602
0.91251	73.4375	0.02088	0.10344	0.98013	74.7653	0.92791	0.96674
0.9809	89.0625	0.02518	0.12478	0.99312	90.1852	0.92989	0.96737

در شکل 9 یک گره چند هدفه نمایش داده شده است. این گراف احتمال زمان پاسخ خوشه های منبع، درصد تجمعی کارها در خوشه های منبع را نشان می دهد که بهره برداری از منابع محیط ابری و نرخ ورود توزیع شده به صورت نمایی را تضمین می کند. با کاهش زمان پاسخ، زمان خدمات و زمان تکمیل، زمان کل محیط ابری نیز کاهش می یابد. زمان کل به زمان صرف شده از سوی الگوریتم یا روش برای تکمیل کلیه کارهای محیط ابری اتلاق می گردد. اختلاف بین زمان خدمات و زمان تکمیل حاکی از آن است که زمان خدمات، حفره درونی زمان تکمیل می باشد. زمان خدمات به زمان صرف شده برای تکمیل کار خاصی از محیط اتلاق می گردد. زمان تکمیل شامل زمان پاسخ و همچنین زمان خدمات می شود.



شکل 9. تعیین نرخ ورود، زمان پاسخ و بهره برداری از منابع

5. نتیجه گیری

درست مثل اعطا کنندگان و مصرف کنندگان ابری، ارزیابی عملکرد منابع نقش حیاتی ایفا می کند. برای ارائه یک مدل ارزیابی عملکرد کارآمد برای خوشه های منبع محیط ابری، یک مدل ریاضی در مورد زنجیره مارکوف با زمان پیوسته جهت تعیین بهره برداری از منابع و زمان خدمات، در کنار (ترکیب) فرایند پواسن در مورد نرخ ورود کارها پیشنهاد شده است. مدل ریاضی پیشنهادی در تعیین نرخ ورود نمایی، زمان پاسخ و بهره برداری از منابع برای انواع متفاوتی از خوشه های منبع در محیط ابری، عملکرد درستی به معرض اجرا می گذارد. در این راستا تعدادی آزمایشات ریاضی جهت ارزیابی سیستم پیشنهادی انجام شده است. یکی از مدل‌های ارزیابی عملکرد کارآمد که راه حل نزدیک به حد بهینه ای برای محیط ابری تولید می کند با تعیین نرخ ورود کارها در زمان نمایی، احتمال زمان پاسخ خوشه های منبع و درصد تجمعی کارها در خوشه های منبع بدست آمده است. زمان ل یا صرف شده محیط ابری خاص با کمک مدل پیشنهادی به حداقل رسانده شده که این امر موجب تعیین درست نرخ ورود و بهره برداری از منابع می شود. در آینده با مقایسه زنجیره مارکوف با زمان گسسته با زنجیره مارکوف با زمان پیوسته می توان مدل پیشنهادی را تعمیم داد و احتمال عملکرد خوشه های منبع و تبدیل زنجیره مارکوف با زمان گسسته به زنجیره مارکوف با زمان پیوسته و بالعکس را بیشتر ارزیابی نمود. هدف از تبدیل، تعیین ترتیب ورود کارها در فواصل زمانی گسسته و فراهم نمودن حدی برای آزمایشات است که حتی برای محیط های کوچک کافی خواهد بود.

References

1. Zhang, Q., Cheng, L., Boutaba, R.: Cloud computing: state-of-the-art and research challenges. *J. Internet Serv. Appl.* 10(1), 7–18. ISSN 1869-0238 (2010)
2. Kowsigan, M., Balasubramanie, P.: Scheduling of jobs in cloud environment using soft computing techniques. *Int. J. Appl. Eng. Res.* 10(38). ISSN 0973-4562 (2015)
3. Kowsigan, M., Balasubramanie, P.: An improved job scheduling in cloud environment using auto associative memory network. *Asian J. Res. Soc. Sci. Humanit.* 6(12), 390–410. ISSN 2249-7315 (2016)
4. Karlin, S., Taylor, H.M.: *A First Course in Stochastic Processes*, 2nd edn, pp. 221–228. Academic Press, Cambridge (1975)
5. Harchol-Balter, M.: *Performance Modeling and Design of Computer Systems-Queuing Theory in Action*, pp. 106–118. Cambridge University Press, Cambridge (2013)
6. Khazaei, H., Mistic, J., Mistic, V.B.: Performance analysis of cloud computing centers using M/G/m/m+r queuing systems. *IEEE Trans. Parallel Distrib. Syst.* 23(5), 936–943 (2012)
7. Jain, S., Jain, S.: Analysis of multi level feedback queue scheduling using markov chain model with data model approach. *Int. J. Adv. Netw. Appl.* 07(06), 2915–2924. ISSN 0975-0290 (2016)
8. Rabbani, M., Tanhaie, F.: A Markov chain analysis of the effectiveness of drum-buffer-rope material flow management in job shop environment. *Int. J. Ind. Eng. Comput.* 6(4), 457–468. ISSN 1923-2934 (2015)
9. Terefe, M.B., Lee, H., Heo, N., Fox, Geoffrey C., Sangyoon, Oh: Energy-efficient multisite offloading policy using Markov decision process for mobile cloud computing. *Pervasive Mob. Comput.* 27, 75–89 (2016)
10. Barnes, A.K., Balda, J.C., Escobar-Mejía, A.: A Semi-Markov model for control of energy storage in utility grids and microgrids with PV generation. *IEEE Trans. Sustain. Energy* 6(2), 546–556. ISSN 1949-3037 (2015)
11. Zhang, W., Wen, Y., Wu, D.O.: Energy-efficient scheduling policy for collaborative execution in mobile cloud computing. In: *Proceedings IEEE INFOCOM*. ISSN: 0743-166X (2013)
12. Park, J.S., Yu, H.C., Chung, K.S., Lee, E.Y.: Markov chain based monitoring service for fault tolerance in mobile cloud computing. In: *IEEE Workshops of International Conference on Advanced Information Networking and Applications* (2011)
13. Wang, J., Qiu, M., Guo, B.: High reliable real-time bandwidth scheduling for virtual machines with hidden Markov predicting in telehealth platform. *Future Gener. Comput. Syst.* 49, 68–76 (2015)
14. Yin, J., Lu, X., Zhao, X., Chen, H., Liu, X.: BURSE: a bursty and self-similar workload generator for cloud computing. *IEEE Trans. Parallel Distrib. Syst.* 26(3). ISSN 1045-9219 (2015)
15. Wu, H., Knottenbelt, W., Wolter, K.: Analysis of the energy response time tradeoff for mobile cloud offloading using combined metrics. In: *IEEE 27th International Teletraffic Congress* (2015)
16. Zhang, R., Su, X., Wang, J., Wang, C., Liu, W., Lau, R.W.H.: On mitigating the risk of cross-VM covert channels in a public cloud. *IEEE Trans. Parallel Distrib. Syst.* 26(8), 1. ISSN 1045-9219 (2015)
17. Peng, Z., Cu, D., Zuo, J., Li, Q., Bo, Xu, Lin, Weiwei: Random task scheduling scheme based on reinforcement learning in cloud computing. *Cluster Comput.* 18(4), 1595–1607 (2015)
18. Lu, C.-T., Chang, C.-W., Li, J.-S.: VM scaling based on Hurst exponent and Markov transition with empirical cloud data. *J. Syst. Softw.* 9, 199–207 (2015)
19. Hyytiä, E., Spyropoulos, T., Ott, J.: Offload (only) the right jobs: robust offloading using the Markov decision processes. In: *IEEE 16th International Symposium on a World of Wireless, Mobile and Multimedia Networks* (2015)
20. Zhang, Q., Chen, Z., Yang, L.T.: A nodes scheduling model based on Markov chain prediction for big streaming data analysis. *Int. J. Commun. Syst.* 28(9), 1610–1619. ISSN 1099-1131 (2014)
21. Kowsigan, M., Rajkumar, S., Seenivasan, P., Vikram Kumar, C.: An enhanced job scheduling in cloud environment using improved metaheuristic approach. *Int. J. Eng. Res. Technol.* 6(2), 184–188 (2017)
22. Kowsigan, M., Kalicharan, S., Karthik, P., Manikandan, A., Manikandan, R.: An enhanced job scheduling in cloud environment using probability distribution. *Int. J. Eng. Technol.* 9(2), 1374–1381 (2017)