



ارائه شده توسط:

سایت ترجمه فا

مرجع جدیدترین مقالات ترجمه شده

از نشریات معتبر

یک الگوریتم زمان بندی پویا محدود شده بر اساس موعد مقرر با صرفه هزینه

برای جریان های کاری علمی در محیط ابری

چکیده

رایانش ابری، که یک الگوی رایانش توزیع شده است، این امکان را فراهم می کند تا منابع IT بر روی اینترنت به کاربران ارائه شود و از مدل پرداخت هزینه به صورت آنی در زمان مصرف، استفاده می کند. برنامه ریزی و زمان بندی برای جریان های کاری چالش بر انگیز ترین مشکل در رایانش ابری می باشد. با وجود این که زمان بندی جریان کاری در سیستم های توزیع شده مانند شبکه ها و یا خوشه بندی ها به صورت گسترده مورد مطالعه قرار گرفته است، اما این راه حل ها برای رایانش ابری قابل استفاده نمی باشند. این موضوع به این علت است که محیط رایانش ابری نسبت به دیگر محیط های توزیع شده از دو منظر متفاوت می باشد: فراهم کردن منابع در زمان نیاز و مدل پرداخت آنی در زمان مصرف. از این رو، برای به دست آوردن مزیت های واقعی محدودیت های جریان کاری بر روی منابع ابری، روش های جدید باید ایجاد شود که بتواند از مزیت های این رایانش ابری استفاده کرده و بتواند چالش های خاص آن را هم رفع کند. در این مطالعه، ما یک الگوریتم اکتشافی محدود شده از نظر موعد مقرر و مقرون به صرفه به صورت پویا را ارائه می کنیم که برای زمان بندی کردن جریان های کاری علمی در رایانش ابری عمومی مورد استفاده قرار می گیرد. هدف تکنیک پیشنهاد شده استفاده کردن از مزیت های ارائه شده توسط رایانش ابری و در عین حال در نظر داشتن عملکرد واقعی ماشین و تغییر در آن و تاخیر های اکتساب آنی برای شناسایی کردن زمان بندی های جریان کاری محدود به موعد مقرر با هزینه ی کمتر، می باشد. ارزیابی عملکردی بر روی بعضی از جریان های کاری شناخته شده نشان می دهد که الگوریتم پیشنهاد شده در مقایسه با جدید ترین الگوریتم های موجود در این زمینه، عملکرد بهتری دارد.

کلمات کلیدی: رایانش ابری، کیفیت خدمات (QoS)، تدارک منابع، زمان بندی، جریان کاری علمی.

دانشمندان در دامنه های تحقیقاتی مختلف مانند فیزیک، بیو انفورماتیک، علوم جغرافیایی و نجوم، برنامه های علمی در مقیاس بزرگ و پیچیده را برای شبیه سازی کردن و تحلیل فعالیت های دنیای واقعی اجرا می کنند. بسیاری از این برنامه های بزرگ مقیاس معمولاً به صورت جریان های کاری ساخته می شوند [۱]. یک جریان کاری، یک برنامه ی کاربردی دانه درشت با تزویج ضعیف است که شامل وظایف محاسباتی می باشد که از طریق وابستگی های داده ها و کنترلی، به هم مرتبط شده اند. جریان های کاری علمی از نظر اندازه ممکن است از چند وظیفه با منابع محدود تا میلیون ها وظیفه نیازمند هزاران ساعت پردازش، ترابایت ها فضای ذخیره و منابع پهنای باند بسیار بالا باشند. این جریان های کاری پیچیده نیازمند محیط های محاسباتی با عملکرد بالا می باشند و معمولاً باید وظایف موجود در این جریان های کاری بین گره های محاسباتی مختلف توزیع شود تا جریان کاری در یک زمان معقول، تمام شود. به صورت متداول، توسعه دهندگان برنامه های علمی از ایستگاه های کاری محلی، ابر کامپیوتر ها، دسته بندی ها و پلتفرم های شبکه ای برای اجرای این جریان های کاری استفاده می کنند. هر کدام از این پلتفرم ها توازن های مختلفی از نظر کارایی، عملکرد و هزینه ارائه می کنند. بسیاری از پروژه های شبکه ای مانند پگاسوس [۲]، آسکالون [۳] و GrADS [۴] سیستم های مدیریت جریان کاری را ارائه کرده اند تا جریان های کاری بر روی شبکه های تعریف، مدیریت و اجرا شود. رایانش ابری اخیراً به عنوان یکی از پلتفرم های مناسب برای اجرای این برنامه های علمی پیچیده ارائه شده است. بسیاری از مطالعه ها [۵ و ۶ و ۷ و ۸ و ۹] به بررسی استفاده از رایانش ابری برای برنامه های علمی پرداخته اند و به این نتیجه رسیده اند که این سیستم های ابری می توانند از نظر عملکرد و هزینه، شرایط مناسبی را ایجاد کنند. اما با ظهور این الگوی محاسباتی جدید، روش های برنامه ریزی زمانی جدید که بتوانند از فواید سیستم های ابری استفاده کنند و چالش های خاص محیط های ابری را هم رفع کنند، در این قسمت مورد نیاز می باشد.

اگرچه می توانند مدل های محاسبه ی رایانش مبتنی بر کاربرد را ارائه کنند که این امکان را فراهم می کند تا منابع IT بر روی اینترنت تحویل داده شود و سپس از یک مدل پرداخت در زمان مصرف استفاده می کند که در این مدل، کاربر ها هزینه را بر اساس میزان مصرف از منابع پرداخت می کنند [۱۱]. خدمات ابری در اصل به صورت زیر طبقه

بندی می شوند : زیر ساختار ها به عنوان خدمات (IaaS) ، که شامل زیر ساختار های خام بوده و مرتبط با میان افزار ها می باشد ؛ پلتفرم ها به عنوان خدمات (PaaS) که شامل API ها برای توسعه ی برنامه ها بر روی یک پلتفرم انتزاعی می باشد و در نهایت دسته ی نرم افزار به عنوان خدمات (SaaS) که می تواند برای خدمات نرم افزا راه دور، پشتیبانی ایجاد کند. راه حل های مبتنی بر PaaS ها و SaaS ها در حال حاضر به عنوان جایگزین های ممکن برای اجرای جریان های کاری علمی در نظر گرفته نمی شوند. این موضوع به این دلیل است که راه حل های مبتنی بر PaaS شامل سر ریز اجرای برنامه های کاربردی موروثی می باشد در حالی که SaaS در حال حاضر به صورت پراکنده برای کاربرد در دسترس هستند [۷]. سیستم های زیر ساخت ابری در مقابل، می توانند مزیت های مرتبط با هزینه و عملکرد مختلف را برای اجرای برنامه های علمی در مقایسه با محیط های اجرای توزیع شده ی متداول مانند شبکه ها و یا خوشه بندی ها ارائه کنند [۶ و ۱۰]. بعضی از فواید به صورت زیر می باشد.

الف (منابع اقتصادی بی نهایت : سیستم های ابری می توانند توهمی از منابع رایانش بی نهایت را ارائه کنند و این کار با استفاده از مفاهیم مجازی سازی محقق می شود که می توان از آن ها در قاب های زمانی معقول استفاده کرد و هزینه ی استفاده از آن نیز بر اساس میزان مصرف از منابع مشخص می شود.

پلتفرم های ابری ازین رو می توانند یک روش جایگزین برای اجرای برنامه های علمی باشند که در این برنامه ها، منابع دیگر در سازمان های تحقیقاتی قرار ندارد بلکه از یک مرکز داده ی بزرگ در زمان نیاز اجاره می شود. این برون سپاری کردن محاسبه های علمی به پلتفرم های ابری می تواند نه تنها از نظر کاهش بار مالی فراهم کردن منابع، بلکه از نظر کاهش تلاش و هزینه ی عملیاتی نیز به این جریان های کاری کمک کرده و در نتیجه دیگر نیازی به حفظ و به روز رسانی زیر ساختار ها وجود نخواهد داشت.

ب) تدارک منابع درست در زمان نیاز : در شبکه ها و خوشه بندی ها، کاربر مقدار زمان و منابع مورد نیاز برای محاسبه را مشخص کرده و مسئولیت تخصیص منابع نیز متعلق به واحد زمانی بندی کار ها می باشد. ازین رو نیاز برای منابع دسته بندی می شود و بر اساس سیاست های زمان بندی، این منابع در اختیار کاربر ها قرار می گیرد. تخصیص منابع و مشخص کردن وظایف نسبت به منابع، همراه با هم انجام می شود و از کنترل کاربر خارج می باشد. اما در سیستم

های رایانش ابری، کاربرد به صورت مستقیم منابع مورد نیاز برای زمانی بندی کردن محاسبه های خودش را با استفاده از یک برنامه ی زمان بندی تحت کنترل خودش، مشخص می کند. این کار به کاربرد کمک می کند تا سر ریز های زمان بندی را کاهش داده و در نتیجه عملکرد این زمان بندی به صورت محسوس بهبود پیدا می کند.

ج) انعطاف : سیستم های رایانش ابری به کاربر ها این امکان را می دهد تا منابع را در زمان نیاز در اختیار گرفته و سپس آن ها را رها کنند. این موضوع به برنامه های کاربردی این امکان را می دهد تا به راحتی سایز خودشان را زیاد کرده و یا این که مبتنی بر نیاز به منابع، سایز منابع مورد استفاده خودشان را کاهش دهند. جریان های کاری معمولاً مرحله های مختلفی دارند که تعداد منابع مورد نیاز برای هر کدام از این منابع متفاوت می باشد. ازین رو برنامه های کاربردی جریان کاری مبتنی بر سیستم های ابری ، می توانند دقیقاً از همان مقدار منبعی که در لحظه نیاز دارند استفاده کنند و دیگر نیازی نیست که تعداد ثابتی از منابع برای یک برنامه در نظر گرفته شود. این موضوع نه تنها تضمین می کند که میزان کارایی منابع در شرایط قابل قبول است بلکه همچنین منجر به صرفه ی هزینه ی بسیار خوبی برای کاربر می بشود.

با این وجود، سیستم های ابری مزیت های بسیار زیادی دارند که هنوز در زمینه های خاص چالش هایی را پیش روی خودشان می بینند که باید با سیاست های زمانی بندی مناسب، این چالش ها رفع شود. بعضی از این مشکلات شامل موارد زیر می باشد :

الف) تغییر عملکرد در رایانش ابری : مجازی سازی منابع، ماهیت مشترک زیر ساختار ها و حالت های ناهمگن در منابع فیزیکی در ابر ها منجر به تغییرات عملکردی می شود. چاد و همکارانش [۱۲] نشان دادند که میزان تغییر عملکرد CPU به مقدار ۲۴٪ بر روی رایانش ابری EC2 آمازون وجود دارد. تغییرات عملکردی ممکن است تاثیرات منفی بر روی جریان های کاری محدود به موعد مقرر بشود که بر روی سیستم های ابری زمان بندی شده اند. سیاست های زمان بندی معمولاً مبتنی بر تخمین وظایف بر روی VM های مختلف می باشد. این تخمین بر اساس ظرفیت پردازش VM ها مشخص می شود. در صورتی که این ظرفیت همیشه به صورت بهینه در نظر گرفته شد و زمان اجرای

کار بیشتر زمان ببرد، وظایف با تاخیر اجرایی همراه می شوند. این تاخیر ممکن است تأثیرات آبخاری منفی بر روی وظیفه ی نهایی داشته باشد و در نتیجه برنامه نمی تواند در موعد مقرر کار خودش را انجام دهد.

ب) اکتساب لحظه ای و تاخیر اتمام کار : زمانی که VM مورد اجاره ی یک جریان کاری قرار می گیرد، این VM نیازمند یک زمان شروع کار اولیه می باشد تا بتواند قبل از این که در اختیار کاربر قرار بگیرد، کار خودش را شروع کند. به صورت مشابه، زمانی که یک VM آزاد می شود اندکی زمان نیاز دارد تا بتواند کار خودش را تمام کند. زمان شروع طولانی ممکن است منجر به تاخیر هایی شود که در نتیجه برنامه نتواند در موعد مقرر کار خودش را انجام دهد و این موضوع در زمان بندی ها باید در نظر گرفته شود. اما، در صورتی که زمام اتمام کار VM ها نیز با تاخیر رو به رو شود، این تاخیر بر روی موعد مقرر تعیین شده برای نرم افزار تأثیر ندارد و ازین رو ممکن است هزینه ی بسیار کمی را برای کاربرد ایجاد کند.

ج) منابع زیر ساخت ابری به صورت ناهمگن : زمانی که یک VM از یک ارائه کننده ی خدمات ابری اجاره می شود، کاربر این توانایی را دارد تا ماشین های مختلف با پیکربندی ها و قیمت های مختلف را برای کار های پردازش خودش انتخاب کند. هر راه حل زمان بندی باید تصمیم گیری های تدارک منابع مناسبی را بر اساس توازن عملکرد و هزینه در نظر داشته باشد.

مزیت ها و موضوعات بیان شده در قسمت بالا نشان دهنده ی نیاز برای توسعه ی الگوریتم های زمان بندی مناسب برای محیط های ابری می باشد که می تواند راه حل های موثر از نظر عملکرد و هزینه را برای اجرای جریان های کاری علمی فراهم کند. اجرای جریان های کاری در محیط های ابری شامل دو مرحله ی اصلی می باشد. مرحله ی اول مرحله ی فراهم کردن منابع می باشد که منابع مورد نیاز پردازشی برای اجرای وظایف را شناسایی کرده و آن ها را فراهم می کند. در مرحله ی دوم، یک زمان بندی ایجاد می شود و هر کدام از وظایف مورد نظر بر روی منابع رایانش مناسب، نگاشت می شود. تصمیم های اتخاذ شده در هر کدام از این فاز ها بر اساس محدودیت های تقدم وظایف و نیاز های عملکردی می باشد که توسط کاربر مشخص شده است. بیشتر کار های قبلی در این زمینه، تمرکز خودشان را بر روی برنامه ریزی کردن جریان های کاری بر روی سیستم های توزیع شده قرار داده اند، مانند شبکه ها و خوشه

بندی ها و فقط فاز زمان بندی را بررسی کرده اند. علت این کار این است که محیط های شبکه و خوشه بندی یک مخزن ایستا از منابع را فراهم می کنند که همواره برای اجرای وظایف در دسترس هستند و پیکربندی آن ها از قبل مشخص می شود. اما محیط های ابری، نیازمند هر دو مرحله ی بالا می باشند که همراه با هم اجرا می شوند تا برنامه های اجرایی موثر فراهم شود. یکی دیگر از مشخصه های مهم کار های قبلی برای شبکه ها و خوشه بندی ها این است که بیشتر آن ها تمرکز خودشان را بر روی حداقل کردن زمان اجرای جریان های کاری قرار داده اند. با وجود این که این بهینه سازی برای این محیط ها مناسب می باشد، اما در سیستم های رایانش ابری، یک پارامتر مهم دیگر در کنار زمان اجرا، موضوع هزینه های اقتصادی می باشد. منابع سریع تر معمولا هزینه ی بیشتری را نیاز دارند و ازین رو یک توازن هزینه - قیمت در انتخاب کردن خدمات مناسب در این سیستم ها وجود دارد. ازین رو سیاست های زمان بندی که برای این سیستم های ابری طراحی می شود باید جایگزین های قیمت / زمان مختلف را در نظر داشته باشند تا بتوانند راه حل های موثری را ایجاد کنند که هزینه های غیر ضروری را ایجاد نکنند.

این کار یک الگوریتم اکتشافی محدود به موعد مقرر با حداقل سازی هزینه به صورت پویا را ارائه می کند که برای زمان بندی کردن برنامه های علمی جریان کاری در محیط های رایانش ابری عمومی مورد استفاده قرار می گیرد. هدف تکنیک پیشنهاد شده استفاده از مزیت های ارائه شد در محیط های ابری و در عین حال در نظر داشتن تغییرات عملکرد VM ها و تاخیر اکتساب لحظه ای می باشد تا بتواند یک زمان بندی دقیق را برای جریان های کاری علمی ارائه کند که بتواند در موعد مقرر مشخص شده با هزینه ی کمتر، وظایف را انجام دهد.

قسمت های بعدی این مقاله به صورت زیر سازماندهی شده است. بخش ۲ نشان دهنده ی برنامه های کاربردی جریان کاری علمی و معماری مورد نظر برای اجرای جریان کاری بر روی منابع منعطف می باشد. تعریف مسئله نیز در بخش ۳ ارائه شده است و بعد از آن الگوریتم های زمان بندی پیشنهاد شده و توصیف آن ها، در بخش ۴ ارائه شده است. عملکرد های مبتنی بر آزمایش ها و ارزیابی آن ها نیز در بخش ۵ ارائه شده است. در بخش ۶ نیز کار های مربوطه و در بخش ۷ نیز جمع بندی ارائه شده است.

۲. معماری مد نظر برای اجرای جریان کاری در سیستم رایانش ابری

این بخش، مدل برنامه ی کاربردی، مدل منابع ابری و معماری کلی برای قالب کاری رایانش اجرای جریان کاری در سیستم ابری را نشان می دهد که در این مطالعه مورد استفاده قرار گرفته است.

۲,۱ مدل برنامه ی کاربردی

یک برنامه ی جریان کاری $W = (T, E)$ به عنوان یک گراف جهت دار غیر چرخه ای (DAG) تعریف می شود

که $T = \{t_1, t_2, \dots, t_n\}$ نشان دهنده ی مجموعه ای از راس ها می باشد که هر کدام نشان دهنده ی وظایف

مد نظر و E هم نشان دهنده ی یال های جهت دار می باشد که داده ها یا وابستگی های کنترلی بین وظایف را نشان

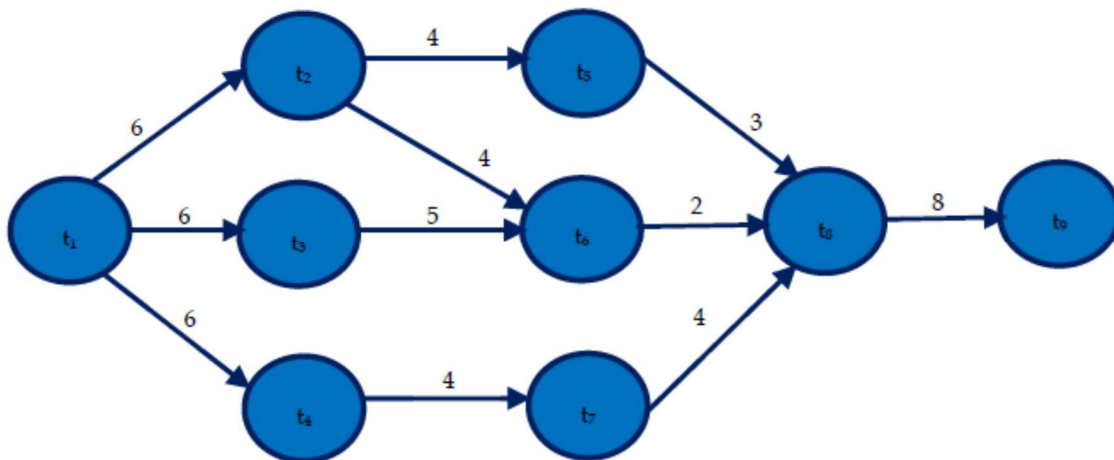
می دهند. وابستگی به صورت e_{ij} یک محدودیت تقدم از حالت (t_i, t_j) می باشد که $t_i, t_j \in T$ و $t_i \neq t_j$

. این موضوع نشان می دهد که وظیفه ی t_j (وظیفه ی فرزند) تنها بعد از اتمام وظیفه ی t_i (وظیفه ی سرپرست)

می تواند شروع شود و وابستگی های مرتبط با داده ها نیز به t_i منتقل می شود. از این رو، یک وظیفه ی فرزند تنها

زمانی می تواند وظیفه ی خودش را شروع کند که تمام گره های سرپرست وظایف خودشان را انجام داده باشند و

وابستگی های مورد نیاز چه از نظر کنترل و چه از نظر داده ها، فراهم شده باشد.



شکل ۱ جریان کاری نمونه

یک موعد مقرر D در این قسمت به عنوان محدودیت زمانی مشخص شده توسط کاربر تعریف می شود که برای اجرای جریان کاری در نظر گرفته می شود. یک جریان کاری نمونه در شکل ۱ نشان داده شده است. هر کدام از این گروه ها نشان دهنده ی یک وظیفه و هر کدام از یال ها نیز نشان دهنده ی انتقال داده بین وظایف می باشد.

۲,۲ مدل منابع ابری

مدل ابری، که در این کار در نظر گرفته شده است شامل منابع مجازی می باشد که توسط زیر ساخت های خدمات ابری ارائه می شود. این خدمات شامل خدمات محاسباتی مانند سیستم های محاسبات منعطف آمازون (EC2) ، و سیستم های ذخیره سازی مانند سیستم های منعطف ذخیره سازی آمازون می باشد که برای ذخیره سازی و یا ورود و خروج داده ها به صورت محلی مورد استفاده قرار می گیرد. تمام خدمات محاسبات و ذخیره سازی در این قسمت در یک مرکز داده و یا یک منطقه قرار دارند در نتیجه پهنای باند میانگین بین خدمات محاسبات تا حد دقیقی، برابر می باشد. علاوه بر این، خدمات محاسبات و رایانش نیز در این قسمت به صورت انواع ماشین های مجازی (VM ها) ارائه می شود. این ماشین های مجازی دارای انواع پیکربندی مختلف برای CPU ها، سائز حافظه و دیگر پیکربندی می باشند که با قیمت های مختلف در دسترس کاربران قرار دارد. فرض ما در این قسمت این است که هیچ محدودیتی بر روی نمونه های VM که مورد استفاده یک برنامه قرار می گیرد وجود ندارد. همچنین، زمانی که یک VM مورد استفاده قرار می گیرد، این VM نیازمند یک زمان شروع کار اولیه می باشد تا قبل از این که در دسترس کاربر قرار بگیرد، کار خودش را شروع بکند. به صورت مشابه زمانی که یک VM آزاد می شود نیازمند اندک زمانی است که به خوبی خودش را خاموش کند. مدل قیمت گذاری مبتنی بر هزینه در زمان مصرف می باشد که مشابه با سیستم های رایانش ابری تجاری فعلی می باشد و کاربر ها بر اساس بازه های زمانی که از یک VM استفاده می کنند، باید هزینه ی آن را پرداخت کنند ، حتی در صورتی که VM ها به صورت کامل در بازه های زمانی مورد استفاده قرار نگیرند. این بازه های زمانی توسط فراهم کننده ی سیستم های ابری مشخص می شود. به عنوان مثال یک سیستم خدمات دهنده مانند آمازون ، هزینه ی کاربران را بر اساس بازه های زمانی یک ساعته مشخص می کند. ازین رو، حتی در صورتی که یک VM تنها برای چندین دقیقه مورد استفاده قرار بگیرد، کاربر باید هزینه ی یک ساعت کامل را پرداخت کند.

اما، بعضی از خدمات دهنده ها مانند پلتفرم های رایانش ابری گوگل [۲۴] ، اخیرا بازه های زمانی کوتاه را ارائه کرده اند و مدل های قیمت گذاری شان بر اساس دقیقه مشخص می شود و کمترین دوره ی پرداخت هزینه شان، به صورت بازه های ده دقیقه ای می باشد. به دلیل این که انتقال داده ها به صورتی داخلی در بیشتر محیط های ابری رایگان است، هزینه ی انتقال داده ها صفر می باشد. علاوه بر این، با وجود این که خدمات دهنده های واقعی برای خدمات ذخیره سازی مورد استفاده برای ورودی یا خروجی فایل ها مبتنی بر حجم اختصاص یافته هزینه را مشخص می کنند، این هزینه ها در این مدل در نظر گرفته نمی شود زیرا این هزینه ها ارتباطی با الگوریتم های زمان بندی ندارند. در مدل های منابع رایانش ابری، نوع ماشین های مجازی (VM_v) بر اساس یک دو جمله ای $\{(ET_{t_i})_v, C_v\}$ مشخص می شود که نشان دهنده ی زمان پردازش تخمین زده شده برای هر وظیفه ی t_i و هزینه بر اساس هر بازه ی زمانی می باشد. در این قسمت فرض ما این است که مان تخمین زده شده برای وظایف بر روی انواع مختلف از ماشین های مجازی را می توان با استفاده از تکنیک های تخمین عملکرد موجود، به دست آورد (یعنی ، مدل های تحلیلی ، مدل سازی های تجربی و داده های پیشینه ای). هزینه ی اجرای یک وظیفه T_i بر روی یک نوع از ماشین مجازی به صورت

محاسبه می شود. علاوه بر این، زمان انتقال داده ها $TT(e_{ij})$ بین وظایف VM_v ، به صورت

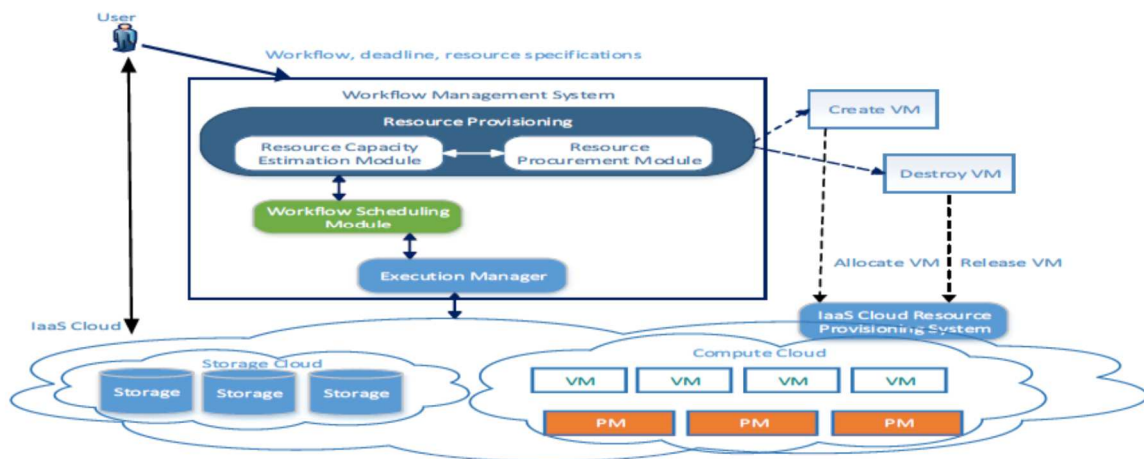
$$\left[\frac{(ET_{t_i})_v}{\text{time interval}} \right] \times C_v.$$

زمان بندی شده بر روی ماشین های مجازی مختلف به صورت $\frac{a_{t_i}}{\beta}$ محاسبه می شود که d_{t_i} نشان دهنده ی سایز فایل داده های خروجی است که باید از t_i به t_j منتقل شود و β هم نشان دهنده ی پهنای باند میانگین در مرکز داده ی ابری می باشد . تنها استثنا در این قسمت زمانی است که یک وظیفه ی سرپرست t_i و وظیفه ی فرزند t_j بر روی یک VM زمان بندی شده باشد که در این حال زمان انتقال داده ها به صورت $TT(e_{ij})$ ، صفر خواهد بود.

۲.۳ مدل پلتفرم رایانشی

شکل ۲ در این قسمت نشان دهنده ی مدل پلت فرم رایانش برای اجرای جریان کاری بر روی منابع ابری می باشد. پلتفرم های رایانش که در این مطالعه مورد استفاده قرار می گیرند، مشابه با نمونه هایی هستند که در مقاله ی [۲۰] مورد استفاده قرار گرفته است. به دلیل این که سیستم های رایانش ابری از کاربر ها می خواهند که مقدار منابع برای

اجرای برنامه ها را مشخص کنند (یعنی با مشخص کردن نوع منبع و مدت زمان اجازه) دو مرحله ی مهم در زمان برنامه ریزی اجرای جریان کاری در سیستم های رایانش ابری وجود دارد. در مرحله ی اول تدارک منابع ، شناسایی و تدارک منابع محاسباتی مناسب برای اجرای وظایف جریان کاری انجام می شود. مرحله ی دوم نیز زمان بندی اجرای وظایف را مشخص کرده و وظایف را بر روی این زمان بندی ها نگاشت می کند.



شکل ۲ مدل های پلتفرم رایانش برای اجرای جریان های کاری علمی بر روی منابع رایانش ابری

یک کاربر جریان کاری را همراه با الزام های مختلف QoS ، مانند محدودیت های موعده مقرر و مشخصات منابع را به سیستم مدیریت جریان کاری ارائه می کند (WMS). محدودیت های موعده مقرر نشان دهنده ی محدودیت های زمانی و مشخصات منابع نشان دهنده ی الزام منابع (رایانش، حافظه، I/O) در برنامه های کاربردی می باشد. با در نظر داشتن این ورودی ها، WMS می تواند به صورت خودکار منابع مورد نیاز را شناسایی کرده و آن ها را فراهم کند، وظایف را بر روی منابع تدارک دیده نگاشت کرده و اجرای جریان کاری را مدیریت کند. برای کاهش دادن هزینه ی اجرای این برنامه ها، WMS منابع را کسب کرده و زمانی که مورد نیاز هستند آن ها را مورد استفاده قرار می دهد و فوراً بعد از استفاده، آن ها را آزاد می کند. این WMS شامل سه ماژول اصلی می باشد: ماژول تدارک منابع، ماژول زمان بندی جریان کاری و مدیریت اجرای کار. ماژول تدارک منابع شامل دو زیر ماژول می باشد: تخمین ظرفیت منابع و تهیه ی منابع. ماژول تخمین منابع، ساختار جریان کاری را تحلیل می کند تا بتواند منابع مورد نیاز را تعیین کند. ماژول تهیه ی منابع نیز با سیستم تدارک منابع در طرف زیر ساختار های رایانش ابری مذاکره می کند تا بتواند

مقادیر منابع مشخص شده را به دست بیاورد. مازول زمان بندی جریان کاری نیز به صورت هماهنگ با قسمت مدیریت اجرایی، نداشت بین منابع فراهم شده و وظایف در جریان کاری را مشخص می کند. وظایف زمان بندی شده توسط مدیریت اجرایی، اجرا می شوند. تفاوت اصلی بین این مدل رایانشی و دیگر مدل های متداول، این است که تخصیص منابع در این مدل بر اساس جریان کاری مشخص می شود و سایز مجموعه ی منابع ممکن است در زمان اجرا، تغییر کند.

REFERENCES

- [1] Deelman, E., Gannon, D., Shields, M., & Taylor, I. (2009). Workflows and e-Science: An overview of workflow system features and capabilities. *Future Generation Computer Systems*, 25(5), 528-540.
- [2] Deelman, E., Singh, G., Su, M. H., Blythe, J., Gil, Y., Kesselman, C., & Katz, D. S. (2005). Pegasus: A framework for mapping complex scientific workflows onto distributed systems. *Scientific Programming*, 13(3), 219-237.
- [3] Wicczorek, M., Prodan, R., & Fahringer, T. (2005). Scheduling of scientific workflows in the ASKALON grid environment. *ACM SIGMOD Record*, 34(3), 56-62.
- [4] Berman, F., Casanova, H., Chien, A., Cooper, K., Dail, H., Dasgupta, A., & Yarkhan, A. (2005). New grid scheduling and rescheduling methods in the GrADS project. *International Journal of Parallel Programming*, 33(2-3), 209-229.
- [5] Hoffa, C., Mehta, G., Freeman, T., Deelman, E., Keahey, K., Berriman, B., & Good, J. (2008, December). On the use of cloud computing for scientific workflows. In *eScience, 2008. eScience'08. IEEE Fourth International Conference on* (pp. 640-645). IEEE.
- [6] Juve, G., Deelman, E., Vahi, K., Mehta, G., Berriman, B., Berman, B. P., & Maechling, P. (2009, December). Scientific workflow applications on Amazon EC2. In *E-Science Workshops, 2009 5th IEEE International Conference on* (pp. 59-66). IEEE.
- [7] Ostermann, S., Iosup, A., Yigitbasi, N., Prodan, R., Fahringer, T., & Epema, D. (2010). A performance analysis of EC2 cloud computing services for scientific computing. In *Cloud Computing* (pp. 115-131). Springer Berlin Heidelberg.
- [8] Deelman, E., Singh, G., Livny, M., Berriman, B., & Good, J. (2008, November). The cost of doing science on the cloud: the montage example. In *Proceedings of the 2008 ACM/IEEE conference on Supercomputing* (p. 50). IEEE Press.
- [9] Juve, G., Deelman, E., Berriman, G. B., Berman, B. P., & Maechling, P. (2012). An evaluation of the cost and performance of scientific workflows on amazon ec2. *Journal of Grid Computing*, 10(1), 5-21.
- [10] Juve, G., & Deelman, E. (2011). Scientific workflows in the cloud. In *Grids, Clouds and Virtualization* (pp. 71-91). Springer London.
- [11] Mell, P., & Grance, T. (2011). The NIST definition of cloud computing
- [12] Schad, J., Dittrich, J., & Quiané-Ruiz, J. A. (2010). Runtime measurements in the cloud: observing, analyzing, and reducing variance. *Proceedings of the VLDB Endowment*, 3(1-2), 460-471.
- [13] Amazon elastic compute cloud (Amazon EC2). [Online] Available: <http://aws.amazon.com/ec2/>
- [14] Amazon elastic block store (Amazon EBS). [Online] Available: <http://aws.amazon.com/ebs/>
- [15] Yu, J., Buyya, R., & Tham, C. K. (2005, July). Cost-based scheduling of scientific workflow applications on utility grids. In *e-Science and Grid Computing, 2005. First International Conference on* (pp. 8-pp). IEEE.
- [16] Nudd, G. R., Kerbyson, D. J., Papaefstathiou, E., Perry, S. C., Harper, J. S., & Wilcox, D. V. (2000). PACE—A toolset for the performance prediction of parallel and distributed systems. *International Journal of High Performance Computing Applications*, 14(3), 228-251.
- [17] Abrishami, S., Naghibzadeh, M., & Epema, D. H. (2012). Cost-driven scheduling of grid workflows using partial critical paths. *Parallel and Distributed Systems, IEEE Transactions on*, 23(8), 1400-1414.
- [18] Jang, S., Wu, X., Taylor, V., Mehta, G., Vahi, K., & Deelman, E. (2004). Using performance prediction to allocate grid resources. *Texas A&M University, College Station, TX. GriPhyN Technical Report*, 25.
- [19] Smith, W., Foster, L., & Taylor, V. (1998, January). Predicting application run times using historical information. In *Job Scheduling Strategies for Parallel Processing* (pp. 122-142). Springer Berlin Heidelberg
- [20] Byun, E. K., Kee, Y. S., Kim, J. S., & Maeng, S. (2011). Cost optimized provisioning of elastic resources for application workflows. *Future Generation Computer Systems*, 27(8), 1011-1026
- [21] Juve, G., Chervenak, A., Deelman, E., Bharathi, S., Mehta, G., & Vahi, K. (2013). Characterizing and profiling scientific workflows. *Future Generation Computer Systems*, 29(3), 682-692.
- [22] Bharathi, S., Chervenak, A., Deelman, E., Mehta, G., Su, M. H., & Vahi, K. (2008, November). Characterization of scientific workflows. In *Workflows in Support of Large-Scale Science, 2008. WORKS 2008. Third Workshop on* (pp. 1-10). IEEE.

- [23] Abrishami, S., Naghibzadeh, M., & Epema, D. H. (2013). Deadline-constrained workflow scheduling algorithms for Infrastructure as a Service Clouds. *Future Generation Computer Systems*, 29(1), 158-169.
- [24] Google Cloud Platform. [Online] Available: <https://cloud.google.com/compute/>
- [25] Palankar, M. R., Iammitchi, A., Ripeanu, M., & Garfinkel, S. (2008, June). Amazon S3 for science grids: a viable solution?. In *Proceedings of the 2008 international workshop on Data-aware distributed computing* (pp. 55-64). ACM
- [26] Rodriguez M. A. & Buyya R. (2014). Deadline based resource provisioning and scheduling algorithm for scientific workflows on clouds. *IEEE Transactions on Cloud Computing*, 2(2), 222-235
- [27] Mao, M., & Humphrey, M. (2011, November). Auto-scaling to minimize cost and meet application deadlines in cloud workflows. In *Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis* (p. 49). ACM
- [28] Johnson, D. S., & Garey, M. (1979). Computers and Intractability: A guide to the theory of NP-completeness. *Freeman&Co, San Francisco*.
- [29] Amazon EC2 Pricing [Online] Available: <https://aws.amazon.com/ec2/pricing/>
- [30] Malawski, M., Juve, G., Deelman, E., & Nabrzyski, J. (2012, November). Cost-and deadline-constrained provisioning for scientific workflow ensembles in iaas clouds. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis* (p. 22). IEEE Computer Society Press.
- [31] Yu, J., & Buyya, R. (2006). Scheduling scientific workflow applications with deadline and budget constraints using genetic algorithms. *Scientific Programming*, 14(3), 217-230.
- [32] Chen, W. N., & Zhang, J. (2009). An ant colony optimization approach to a grid workflow scheduling problem with various QoS requirements. *Systems Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 39(1), 29-43.
- [33] Workflow Generator [Online] Available: <https://confluence.pegasus.isi.edu/display/pegasus/WorkflowGenerator>
- [34] Yuan, Y., Li, X., Wang, Q., & Zhu, X. (2009). Deadline division-based heuristic for cost optimization in workflow scheduling. *Information Sciences*, 179(15), 2562-2575.
- [35] Sakellariou, R., Zhao, H., Tsiakkouri, E., & Dikaiakos, M. D. (2007). Scheduling workflows with budget constraints. In *Integrated Research in GRID Computing* (pp. 189-202). Springer US.
- [36] Prodan, R., & Wicczorek, M. (2010). Bi-criteria scheduling of scientific grid workflows. *Automation Science and Engineering, IEEE Transaction on*, 7(2), 364-376.
- [37] Duan, R., Prodan, R., & Fahringer, T. (2007, November). Performance and cost optimization for multiple large-scale grid workflow applications. In *Supercomputing, 2007. SC'07. Proceedings of the 2007 ACM/IEEE Conference on* (pp. 1-12). IEEE.
- [38] Afzal, A., Darlington, J., & McGough, A. (2006, September). Qos-constrained stochastic workflow scheduling in enterprise and scientific grids. In *Proceedings of the 7th IEEE/ACM International Conference on Grid Computing* (pp. 1-8). IEEE Computer Society.
- [39] Saifullah, A., Ferry, D., Lu, C., & Gill, C. (2012). Real-time scheduling of parallel tasks under a general dag model.
- [40] Poola, D., Garg, S. K., Buyya, R., Yang, Y., & Ramamohanarao, K. (2014, May). Robust scheduling of scientific workflows with deadline and budget constraints in clouds. In *Advanced Information Networking and Applications (AINA), 2014 IEEE 28th International Conference on* (pp. 858-865). IEEE.
- [41] Szabo, C., Sheng, Q. Z., Kroeger, T., Zhang, Y., & Yu, J. (2014). Science in the cloud: allocation and execution of data-intensive scientific workflows. *Journal of Grid Computing*, 12(2), 245-264.
- [42] Pandey, S., Wu, L., Guru, S. M., & Buyya, R. (2010, April). A particle swarm optimization-based heuristic for scheduling workflow applications in cloud computing environments. In *Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on* (pp. 400-407). IEEE.
- [43] Mao, M., & Humphrey, M. (2012, June). A performance study on the vm startup time in the cloud. In *Cloud Computing (CLOUD), 2012 IEEE 5th International Conference on* (pp. 423-430). IEEE.



این مقاله، از سری مقالات ترجمه شده رایگان سایت ترجمه فا میباشد که با فرمت PDF در اختیار شما عزیزان قرار گرفته است. در صورت تمایل میتوانید با کلیک بر روی دکمه های زیر از سایر مقالات نیز استفاده نمایید:

لیست مقالات ترجمه شده ✓

لیست مقالات ترجمه شده رایگان ✓

لیست جدیدترین مقالات انگلیسی ISI ✓

سایت ترجمه فا ؛ مرجع جدیدترین مقالات ترجمه شده از نشریات معتبر خارجی