Archive of SID

www.ElitesJournal.ir

مجله نخبگان علوم و مهندسی

Journal of Science and Engineering Elites

جلد ۴- شماره ۶- سال ۱۳۹۸

ارائه انواع روش های توازن بار مبتنی بر شبکه رایانش ابری

رقیه شفیعی سرخکلائی[۱*]، عبداله کاوه تلاوکی[۲]

۱- کارشناسی ارشد مهندسی کامپیوتر نرم افزار، مسئول فناوری اطلاعات آموزشکده فنی دختران قدسیه ساری

۲- دکترای تخصصی مهندسی برق، معاونت آموزشی و پژوهشی و فناوری آموزشکده فنی دختران قدسیه ساری

[*]confghodsieh.info@gmail.com

ارسال: بهمن ماه ۹۸ پذیرش: بهمن ماه ۹۸

چکیده

هر تغییر و مفهوم جدیدی در دنیای فناوری، مشکلات و پیچیدگی‌های خاص خود را دارد. بهره‌گیری از شبکه رایانش ابری نیز از این قاعده مستثنی نبوده و چالش‌های فراوانی را پیش روی صاحب‌نظران این حوزه قرار داده است. یکی از چالش‌های مهم در زمینه شبکه رایانش ابری، حفظ توازن بار بهینه در محیط ابر است. شبکه رایانش ابری اصطلاحاً شامل مجازی سازی، محاسبات توزیعی، شبکه، نرم افزار و سرویس های وب می باشد. یک ابر از چندین عنصر مثل کلاینت ها، مراکز داده و سرور های توزیع شده تشکیل شده است که در آن مباحث تحمل خطا، دسترسی بالا، مقیاس پذیری، انعطاف پذیری، کاهش سربار برای کاربران، کاهش هزینه ی مالکیت، سرویس دهی به درخواست ها و غیره وجود دارد. تمرکز روی این مسائل برای ایجاد یک الگوریتم توازن بار کارآمد می باشد. این مقاله به مرور روش‌های توازن بار مبتنی بر شبکه رایانش ابری و نقد و بررسی آنها پرداخته است.

کلمات کلیدی: توازن بار، شبکه رایانش ابری، زمانبندی، کیفیت سرویس

۱- مقدمه

اگر کامپیوترهای بزرگ به عنوان نسل اول سیستم‌های محاسباتی در نظر گرفته شود ما با یک سیستم بسیار بزرگ مواجه بودیم که کاربران از طریق یک ترمینال واحد به آن دسترسی پیدا می‌کردند. به مرور این سیستم‌ها کوچک‌تر شدند و با توان پردازشی بیشتری به صورت رایانه‌های شخصی در اختیار همه کاربران قرار گرفتند. سپس این امکان فراهم شد که با اتصال مجموعه‌ای از این سیستم‌های کوچک شبکه‌ای با توان پردازشی بیشتر فراهم شد تا پاسخگوی نیازهای پردازشی بیشتر و سنگین‌تر باشند؛ اما نیازهای پردازشی به شکل فزاینده‌ای در حال افزایش بودند و نیاز به سیستم‌های محاسباتی بزرگ‌تر و قوی‌تر احساس شد [۱]؛ بنابراین تعداد زیادی از این شبکه‌ها به صورت اختصاصی در سرتاسر اینترنت به هم متصل شدند و شبکه محاسبات توری را به وجود آوردند. در این بین مشاهده شد که میلیون‌ها کاربر در اینترنت وجود دارند که در اکثر اوقات از تمام توان رایانه‌ی خود استفاده نمی‌کنند و سیستم محاسباتی دیگری شکل گرفت تا کاربرانی که تمایل دارند، زمان‌های بیکار سیستم خود را برای کارهای محاسباتی دیگران هدیه کنند؛ بنابراین تعداد بسیار زیادی منبع محاسباتی کوچک در شبکه‌ای تحت عنوان محاسبات داوطلبانه به هم پیوستند و توان پردازشی عظیمی را به وجود آوردند [۲]؛ اما هنوز منابع بسیار زیاد دیگری در سازمان‌ها و مراکز اینترنتی وجود داشت که تمام ظرفیت آن‌ها به طور کامل به کار گرفته نشده بود. این منابع نمی‌توانستند در شبکه محاسبات توری به صورت اختصاصی به کار

www.SID.ir

Archive of SID

گرفته شوند زیرا برای آن‌ها وظیفه دیگری تعریف شده بود. در عین حال امکان استفاده از آن‌ها در شبکه داوطلبانه هم وجود نداشت چون فلسفه وجودی آن‌ها کاربردهای تجاری بود. به این ترتیب رویکرد جدیدی شکل گرفت که بتوان با استفاده از فناوری‌های مجازی‌سازی این منابع را به صورت قابل انعطاف و پویا برای کاربردهای مختلف مورد استفاده قرار داد و از تمام ظرفیت آن‌ها به طور مؤثر استفاده کرد. این فناوری محاسبات ابری، لایه زیرساخت نام داشت که امکان استفاده از منابع محاسبات و ذخیره‌سازی را به صورت یک سرویس برحسب نوع نیاز فراهم می‌آورد [۳]. در حقیقت با ایجاد یک لایه انتزاعی بر روی کلیه منابع فیزیکی خود به کمک مجازی‌سازی امکان مدیریت پویای منابع فیزیکی حاصل می‌شود بنابراین محاسبات ابری از دید زیرساخت به گونه‌ای از سیستم‌های توزیع‌شده و موازی اطلاق می‌گردد که مجموعه‌ای از رایانه‌های مجازی را که به یکدیگر متصل هستند شامل می‌شود. این رایانه‌ها به طور پویا عرضه شده و به عنوان یک یا چند منبع محاسباتی یکپارچه بر اساس توافقات سطح سرویس ارائه می‌شوند [۴]. این توافقات در طول انجام مذاکرات سرویس‌دهندگان و مصرف‌کنندگان برقرار می‌گردند. محاسبات ابری سعی دارد ایجاد پویای نسل جدیدی از مراکز داده‌ای را با ارائه کردن سرویس‌ها و خدمات در ماشین‌های مجازی شبکه شده به صورت پویا به گونه‌ای ممکن سازد که ارائه‌دهندگان خدمات کاربردی بتوانند سرویس‌ها و برنامه‌های کاربردی را با انعطاف‌پذیری و سهولت بیشتری ارائه کنند و کاربران نیز بتوانند از هر جایی از دنیا به برنامه‌های کاربردی دسترسی داشته باشند.

مجازی سازی، تکنولوژی اصلی یک زیرساخت به عنوان یک راه حل سرویس‌دهی می‌باشد. مجازی‌سازی راهی است برای یکی کردن منابع و دریافت کارآیی بیشتر از یک رایانه به وسیله فریفتن آن و ایجاد تصور این که از مجموعه‌ای از رایانه‌های متعدد تشکیل شده است[۵].

ماشین‌های مجازی، متفاوت از یک کامپیوتر فیزیکی نیستند؛ بنابراین در درون یک ماشین مجازی پی توان هر نوع سیستم‌عاملی را به نظم درآورد و در آن هر نوع سرویسی را ایجاد کرد.

اختلاف اصلی بین یک ماشین مجازی و یک کامپیوتر فیزیکی این است که کامپیوتر فیزیکی می‌تواند میزبان یک تعدادی از ماشین‌های مجازی باشد. این تعداد بستگی به منابع کامپیوتر فیزیکی دارد، و منابعی که می‌خواهیم به هر ماشین مجازی اختصاص بدهیم، اما به طور متوسط، شش ماشین مجازی به ازای هر کامپیوتر فیزیکی می‌تواند وجود داشته باشد.

ماشین‌های مجازی راهکاری مناسب جهت اجرای کامل مجازی‌سازی بر اساس نرم‌افزار می‌باشند. در واقع با استفاده از ماشین‌های مجازی، در این نوع مجازی‌سازی، یک سخت‌افزار را با سطح کیفی مناسبی شبیه‌سازی می‌شود.

در سیستم‌های ابری رایج، مسئله توازن بار و زمان‌بندی کارها صرفاً بر اساس الگوریتم‌های نوبت چرخشی [۶] و الگوریتم‌های ایستا [۷] انجام می‌گیرد و از طرفی الگوریتم نوبت چرخشی به‌صورت ایستا عمل می‌کند و فقط روی کارهای واردشده در یک لحظه مشخص کار می‌کند و اطلاعاتی از اجرای کارهای قبلی ندارد و اگر کاری هم بعد از آن زمان که این الگوریتم شروع به کار می‌کند وارد سیستم شود، در نظر گرفته نمی‌شود و باید منتظر بماند تا الگوریتم از حالت اجرا خارج‌شده و دوباره وارد زمان‌بندی شود و کارهای جدید را برای اجرا انتخاب کند به عبارتی در حین اجرا به کارهای قبل و بعد خود توجهی نمی‌نماید؛ و زمان پاسخ کار را طولانی می‌کند اما این دسته الگوریتم‌ها سربار زمان اجرا را کاهش می‌دهند و هزینه اجرا را تخمین می‌زنند که می‌توان این دو را از مزایای این نوع الگوریتم‌ها دانست.

چنین مشکلاتی باعث شده تا نقش الگوریتم‌های توازن بار پویا پررنگ‌تر شود که از جمله الگوریتم‌های به کار رفته در این زمینه، الگوریتم جستجوی لگزی [۸] می‌باشد که زمان‌بندی و اجرای کارهای رسیده به سیستم را با توجه به اطلاعاتی که از کارهای جاری موجود در سیستم دارد، انجام می‌دهد و از طرفی اگر کار جدیدی به سیستم وارد شود هم زمان وارد زمان‌بندی می‌گردد که این عمل باعث می‌شود زمان پاسخ به کاربر کمتر شود و هزینه نیز کاهش پیدا کند و زمان اجرا را حداقل می‌کنند؛ اما هنوز این الگوریتم‌ها به وضعیت قبلی سیستم و سرورها بی‌توجه هستند و در صورت عدم اجرای کار با سرورهای درون هر کلاستر، مهاجرت بین کلاسترها اتفاق می‌افتد.

www.SID.ir

Archive of SID

ساختار این مقاله در ادامه، به شرح زیر است: در بخش ۲، مفاهیم پایه و رویکردهای اصلی توازن بار در شبکه رایانش ابری معرفی می گردند. بخش ۳ به مرور روش های توازن بار مبتنی بر شبکه رایانش ابری پرداخته و بخش ۴، مقاله را جمع‌بندی و نتیجه‌گیری می کند.

۲- مفاهیم پایه و رویکردهای اصلی توازن بار

فناوری اینترنت به‌سرعت رو به رشد است و استفاده از آن نیز گسترده شده است و در این بین محاسبات ابری در حال ظهور و گسترش نیز است. فرض می‌شود این محاسبات به عنوان ابزاری برای پاسخ به نیازهای کاربران ارائه شده است [۹] و کاربران در بستر اینترنت بدون وابستگی مکانی می‌توانند از خدمات آن استفاده کنند. همچنین برای مکان‌هایی که تأمین پویای منابع و استفاده از تکنولوژی مجازی‌سازی اهمیت داشته باشد نیز بکار می‌رود [۱۰]. این مکانیزم نیاز به اینترنت با سرعت بالا دارد و به صورت خوشه‌ای (کلاستری) عمل می‌کند.

محاسبات ابری آینده روشنی دارد اما هنوز مشکلاتی دارد که مرتفع نشده است. توازن بار نیز یکی از این مشکلات است. منظور از توازن بار متعادل‌سازی بار ماشین‌ها و توزیع مساوی حجم کار به‌صورت محلی به تمام ابرها است و از این رو توازن بار به یک موضوع ضروری در محاسبات ابری تبدیل‌شده است و توسط ارائه‌دهندگان سرویس ابر در پلت فرم ابر در نظر گرفته‌شده است همچنین این پلت فرم مکانیزم بین موازنه بار با هزینه کم و منابع نامحدود را برای کاربران بایستی فراهم سازد؛ بنابراین توازن بار توزیع درخواست‌ها بر اساس استقرار نرم‌افزارها در مراکز داده انجام می‌شود.

محاسبات ابری کاملاً به اینترنت وابستگی دارد و تمام برنامه‌ها، فایل‌ها بر روی ابر و در بستر اینترنت قرار دارند. محاسبات ابری مفاهیم موازی و توزیع‌شده را پشتیبانی می‌کند و منابع را بین کاربران به اشتراک می‌گذارد و سخت‌افزار و نرم‌افزار و اطلاعات را برحسب تقاضا به اشتراک می‌گذارد و در مقابل آن کاربران هزینه آن را نیز پرداخت می‌کنند. مشتری نیازی به خریداری سیستم‌عامل یا نرم‌افزار ندارد و با استفاده از اینترنت می‌تواند از منابع و نرم‌افزارهای موردنیاز در زمان مشخصی استفاده کند و تنها هزینه مدت‌زمان استفاده از آن را بپردازد.

استفاده مناسب و بهینه از منابع هم چون حافظه، پردازشگر و زمان‌بندی یا توزیع بار در شبکه ابری یک چالش است. متوازن‌سازی بار فرآیند توزیع بار در بین گره‌های مختلف یک سیستم توزیع شده به منظور بهبود استفاده از منابع و نیز زمان انجام کار می‌باشد به گونه‌ای که از موقعیتی که بسیاری از گره‌ها دچار اضافه بار شده باشند و دیگر گره‌ها بیکار باشند یا کار کمی انجام دهند جلوگیری شود. همچنین زمان‌بندی یا توزیع بار می‌تواند با تقسیم وظایف میان کامپیوترهای مختلف، هم منجر به کاهش هزینه‌ها برای کاربران شود و هم بتواند منابع بیشتری را برای اجاره در اختیار فراهم کنندگان قرار دهد. این عمل رضایت کاربر را حداکثر، زمان پاسخ را حداقل، بهره‌برداری از منابع را افزایش و کارایی سیستم را بالا می‌برد [۱۱].

۲-۱- توازن بار

یکی از مهم‌ترین مؤلفه‌های یک معماری محاسبات ابری مؤلفه توازن بار است. مهم ترین وظیفه این مؤلفه دریافت کارهای کاربران و توزیع آن‌ها بر روی سرورهای متفاوت می‌باشد، به‌گونه‌ای که دسترسی به داده‌ها تا حد امکان محلی شده و محاسبات نیز به‌صورت متعادل بر روی هر سروری قرار گیرد؛ یعنی آنکه در صد به‌کارگیری سرورها تقریباً یکسان باشد.

فرایند توازن بار[1] درواقع یک فرآیند جابه‌جایی و بارگذاری کلی است تا گره‌های منحصربه‌فرد در یک سیستم اشتراکی ، استفاده مؤثر از منابع را داشته باشند و زمان پاسخ کارها را بهبود بخشند؛ و به طور هم زمان تقسیم‌کار را بین گره‌ها انجام می‌دهد [۱۲].

یک الگوریتم توازن بار پویا، ذاتاً شامل حالت قبلی سیستم و یا رفتار سیستم نمی‌شود و به رفتار و حالت فعلی سیستم وابسته است. معیارهای مهمی که برای توسعه‌ی این الگوریتم باید در نظر گرفته شود [۱۳]:

- تخمین بارگذاری

[1] load balancing

www.SID.ir

Archive of SID

- دلیل بارگذاری
- پایداری سیستم‌های متفاوت
- ارتباط بین نودها
- کارایی سیستم
- ذات و اصل کار منتقل‌شده
- انتخاب نودها

۲-۲- اهداف توازن بار

لیونی[2] اهداف توازن بار را به‌صورت زیر بیان می‌دارد [۱۴]:

- افزایش کارایی
- داشتن طرح پشتیبان برای سیستمی که حتی به‌صورت جزئی با شکست مواجه شده است.
- تأمین پایداری سیستم
- سازگار نمودن سیستم با اصلاحات

هدف اصلی از توازن بار سرعت بخشیدن به اجرای برنامه‌های کاربردی با منابع مختلف و کاهش زمان پاسخ است [۱۵]. روش‌های توازن بار در محیط‌های همگن و ناهمگن بکار می‌رود. اساساً دو نوع تکنیک توازن بار وجود دارد:۱-ایستا ۲-پویا

الگوریتم‌های ایستا تنها در زمانی کار می‌کنند که گره‌ها تنوع بار کمی دارند. بنابراین این الگوریتم‌ها برای محیط ابری که دارای تنوع کاری بسیاری است مناسب نیست و الگوریتم‌های پویا سودمندتر عمل می‌کنند. اما برای رسیدن به این هدف باید هزینه اضافی در مورد جمع‌آوری و نگهداری اطلاعات بار پرداخت.

تکنیک‌های پویا برای متعادل کردن کارها در بین منابع ناهمگون بسیار موفق عمل می‌کنند. روش پیشنهادی ما یک روش پویا است که به جز توازن بار، اولویت کارها در صف انتظار ماشین‌های مجازی را نیز محاسبه می‌کند.

در محیط ابری هر زمان که ماشین مجازی بار سنگینی از کارها را دارد می‌تواند این کارها را به ماشین‌های مجازی که بار کمتری دارد انتقال دهد. در این حالت، زمانی که بیش از یک کار از ماشین مجازی پرکار منتقل می‌شود، اگر بیش از یک ماشین مجازی فعال وجود داشته باشد این کارها برحسب اولویت باید منتظر پردازش باقی بمانند. توازن بار معمولاً در سطح مرکز داده انجام می‌شود.

توازن بار استاتیک بدین مفهوم است که در زمان کامپایل منابع موردنیاز جهت حفظ توازن بار بررسی می‌شود. اجرای این الگوریتم ساده اما سیستم دچار سربار می‌شود. این الگوریتم‌ها تنها درزمانی که تنوع کارهای سیستم کم است می‌تواند مفید واقع شوند پس نمی‌تواند نقش مؤثری در محیط ابری داشته باشد. الگوریتم‌های توازن بار پویا کارهای مختلف را با توجه به زمان اجرا بین گره‌ها توزیع می‌کند [۱۶].

۳- روش‌های توازن بار مبتنی بر شبکه رایانش ابری

در این بخش، الگوریتم‌های رایج در توازن بار مورد بررسی قرار می‌گیرد:

- الگوریتم خوشه‌بندی فعال

بر اساس گروه‌بندی گره‌های شبیه به هم کار می‌کنند، فرایندها شامل: یک گره فرآیند را آغاز می‌کند و از بین گره‌های همسایه معتبر خود یک گره متفاوت نسبت به قبلی به عنوان گره واسط انتخاب می‌کند. گره واسط پس از آن یک ارتباط بین

[2] liveny

www.SID.ir

Archive of SID

یک همسایه هم نوع گره اولیه آن شکل می‌دهد. سپس گره واسط ارتباط گره اولیه با خودش را جدا می‌کند. مجموعه فرآیندهای بالا مرتب تکرار می‌شوند. در این الگوریتم، کارایی سیستم با افزایش منابع افزایش می‌یابد در نتیجه توان عملیاتی با استفاده از این منابع کارآمد بالا می‌رود. این الگوریتم، با افزایش تنوع[3] سیستم، تنزل پیدا می‌کند[۱۷].

- الگوریتم کاوش زنبورعسل

حالتی از توازن بار در وب سرورها افزایش یا کاهش تقاضا می‌باشد، سرویس‌های تخصیصی به صورت پویا تقاضاهای کاربر را تنظیم می‌کنند. این سرورها تحت عنوان سرورهای مجازی گروه‌بندی‌شده‌اند و هر سرور مجازی صف‌های سرویس مجازی خود را دارد. هر سرور با پردازش یک درخواست و تقاضا از صف، مقدار منافع و مزایا موردنیاز را محاسبه می‌کند، مشابه آنچه زنبورهای عسل در حرکات موزون از خود نشان می‌دهند یک اندازه‌گیری از این مزایا مقدار زمان مصرفی پردازنده برای پردازش یک درخواست می‌باشد. قسمت رقص زنبورعسل مشابه تابلو آگهی در اینجا می‌باشد. همچنین این صفحه برای تبلیغ مزایا (منافع موردنیاز) در کل کلونی استفاده می‌شود. هرکدام از این سرورها نقش یک کاوشگر و یا یک دیده بان (پیشاهنگی) رامی گیرد. سرور پس از پردازش یک درخواست می‌تواند مزایا و منافع را به تابلوهای آگهی با احتمال rp ارسال کند. یک سرور می‌تواند یک صف از سرورهای مجازی را با احتمال xp انتخاب کند xp، نشان‌دهنده رفتار کاوشگر و یا می‌تواند آگهی‌ها را بررسی کند (دیدن رقص) و سرویس‌دهی کند که بنابراین رفتار پیشاهنگی را نشان می‌دهد. یک سرور برای سرویس‌دهی یک درخواست، پس از محاسبه منافع و مقایسه‌ی آن با کل منافع کلونی، مقدار xp را برای آن در نظر می‌گیرد. اگر این منافع بالابود سرور در همان سرور مجازی جاری باقی می‌ماند و یک آگهی با احتمال rp برای آن ارسال می‌شود، اگر منافع کم باشد سرور نقش پیشاهنگی و یا کاوشگری را می‌گیرد. این روش به توازن بار عمومی از طریق فعالیت‌های سرورهای محلی دست پیدا می‌کند. با افزایش تنوع سیستم، کارایی سیستم افزایش می‌یابد؛ اما توان عملیاتی با افزایش اندازه سیستم، افزایش نمی‌یابد. برای شرایطی مناسب است که در آن جمعیت متنوعی از انواع سرویس موردنیاز است[۱۸].

- گرایش نمونه‌گیری تصادفی

در اینجا یک گراف مجازی ساخته‌شده است که در آن هر گره (سرور به‌عنوان گره) یک اتصال برای نمایش بار سرور دارد. در گراف، هر سرور به‌صورت یک گره با هر درجه‌ای به منابع آزاد سرور هدایت می‌شود. هرگاه یک گره یک کار را انجام دهد، آن از لبه‌های ورودی حذف می‌شود که نشان‌دهنده‌ی کاهش دسترسی به منبع آزاد می‌شود. بعد از تکمیل یک کار، گره یک لبه ورودی ایجاد می‌کند که نشان‌دهنده‌ی افزایش دسترسی به منبع آزاد می‌باشد. افزودن و حذف فرآیند به‌صورت نمونه‌گیری تصادفی انجام می‌شود. در ابتدای مسیر در هر گره به‌صورت تصادفی یک همسایه انتخاب می‌شود. آخرین گره، برای تخصیص بار انتخاب‌شده است .همچنین می‌توان روش دیگری برای انتخاب گره جهت تخصیص بار استفاده شود و انتخاب گره بر اساس معیارهای خاص باشد. با این‌حال انتخاب و تخصیص بار می‌تواند برای گره‌های کم بار باشد. در نهایت آنچه به دست می‌آید یک گراف جهت‌دار می‌باشد این روش، بار را در عرض تمام گره‌های سیستم متعادل می‌کند. کارایی سیستم با جمعیت مشابه و بالای منابع افزایش می‌یابد، در نتیجه توان عملیاتی با افزایش منابع سیستم، افزایش می‌یابد طرح توازن بار در اینجا کاملاً غیر متمرکز می‌باشد؛ لذا بکار گیری آن در سیستم‌های شبکه‌های بزرگ مثل ابر مناسب می‌باشد[۱۹].

- روش کارتن

برای کنترل استفاده از توازن بار و نرخ توزیع‌شده‌ی محدود[4]، مکانیسم کارتن را برای ابر، پیشنهاد کرده است. توازن بار برای توزیع برابر کارها به سرورهای متفاوت استفاده می‌شود، بنابراین هزینه‌های مربوطه می‌تواند کاهش یابد و نرخ توزیع‌شده‌ی محدود برای حصول اطمینان از اینکه منابع به روشی توزیع شده‌اند که تخصیص عادلانه‌ی منابع حفظ می‌شود، استفاده می‌شود.

[3] Diversity
[4] DRL

www.SID.ir

Archive of SID

با سربار محاسباتی و ارتباطی خیلی کم، این الگوریتم ساده و برای پیاده‌سازی آسان است. محیط استفاده از این الگوریتم، چارچوب متحد برای کنترل ابر است. از میان معیارهای توازن بار، معیار سربار و بهره‌برداری از منابع در نظر گرفته‌شده است[۲۰].

- روش رویداد محور

یک الگوریتم توازن بار رویداد محور را برای بازی‌های آنلاین چند بازی‌کننده حجیم زمان واقعی[5] پیشنهاد کرده است. این الگوریتم پس از دریافت رویدادهای ظرفیتی به‌عنوان ورودی، اجزای خودش را درزمینه ی منابع و حالت کلی نشست بازی[6]، تجزیه‌وتحلیل می‌کند. درنتیجه فعالیت‌های توازن بار، نشست بازی را تولید می‌کند. این روش قادراست که مقیاس یک نشست بازی را روی چندین منبع بر اساس بار متغیر کاربر کاهش یا افزایش دهد اما نقض کیفیت سرویس گاه و بی گاه دارد. در این الگوریتم از میان معیارهای توازن بار، معیار بهره برداری از منابع در نظر گرفته شده است[۲۱].

- توازن بار مبتنی بر سرور برای سرویس‌های توزیع‌شده‌ی اینترنتی

یک سیاست توازن بار مبتنی بر سرویس جدید برای سرورهای وب که در سرتاسر جهان توزیع شده‌اند پیشنهاد شده است. این سیاست به کاهش زمان سرویس با استفاده از محدود کردن تعداد تغییر مسیرهای یک درخواست به نزدیک‌ترین سرور راه دور بدون سربار کردن آن‌ها، کمک می‌کند. یک میان افزار برای پیاده‌سازی این پروتکل توصیف شده است، همچنین از یک روش اکتشافی برای کمک به سرورهای وب به منظور تحمل بار استفاده می‌کند[۲۲].

- منطق فازی

یک الگوریتم توازن بار مبتنی بر نوبت گردشی را در محیط‌های ماشین مجازی در شبکه ابری، به‌منظور دستیابی به زمان پاسخ و زمان پردازش بهتر، طراحی کرده است. الگوریتم توازن بار قبل از اینکه سرورهای پردازشی از راه برسند، انجام‌شده است. کار بر اساس پارامترهای مختلفی مانند سرعت پردازنده، بار تخصیص داده‌شده به ماشین مجازی و غیره، زمان‌بندی‌شده است. این الگوریتم اطلاعات داخل هر ماشین مجازی و تعداد درخواست‌های جاری تخصیص داده شده به ماشین مجازی سیستم را نگهداری می‌کند. زمانی که یک کار برای تخصیص می‌آید این الگوریتم ماشینی با حداقل بار را مشخص می‌کند و اگر بیشتر از یک ماشین با این ویژگی وجود داشته باشد، این الگوریتم ماشین اول را مشخص می‌کند. تلاش بسیاری شده که تکنیک توازن بار جدید را بر اساس منطق فازی پیاده‌سازی کنند. از آنجایی‌که منطق فازی مانند زبان طبیعی است می‌توانند مسائل خود را تدوین و فرموله نمایند. در این معماری فازی ساز فرآیند فازی سازی را اجرا می‌کند که دو نوع از داده‌های ورودی مانند سرعت پردازنده و بار تخصیص داده شده به ماشین مجازی را تبدیل می‌کند و یک خروجی مانند بار موازنه‌شده را که در استنتاج سیستم موردنیاز است. این طرح سرعت پردازنده و بار ماشین مجازی را به عنوان دو پارامتر ورودی، ب منظور توازن بار بهتر در ابر با استفاده از منطق فازی در نظر می‌گیرد. این پارامترها به‌عنوان ورودی به فازی ساز داده می‌شوند که برای اندازه‌گیری بار متعادل به‌عنوان خروجی استفاده می‌شود. دو پارامتری که به‌عنوان سرعت پردازنده و بار تخصیصی ماشین مجازی نامیده می‌شوند با هم برای ارزیابی بار متعادل شده روی مراکز داده‌ی محیط‌های شبکه ابری از طریق منطق فازی، استفاده می‌شوند. نتایج به دست‌آمده با ارزیابی کارایی می‌تواند با کاهش زمان پردازش و نیز بهبود زمان پاسخ به توازن بار دست یابد که منجر به استفاده‌ی حداکثر از منابع می‌شود. سرعت پردازنده و بار اختصاص‌یافته به ماشین مجازی برای توازن بار در شبکه ابری از طریق منطق فازی اعمال می‌شوند[۲۳].

- مدل پیام گرا[7]

[5] MMOG
[6] Game Session
[7] Message oriented model

www.SID.ir

Archive of SID

خوشه‌ها فرصت استفاده از برنامه‌های کاربردی توزیع‌شده توسط کامپیوترهای مختلف روی شبکه‌ها را فراهم می‌کنند. این موضوع مرتبط با خوشه‌های مطرح‌شده در عملکرد شبکه می‌باشد، اگر کل بار در شبکه توزیعی توسط یک کامپیوتر توزیع شود، باعث کندی شبکه می‌شود .برای جلوگیری از این وضعیت، مدیریت منابع می‌تواند به صورت معیارهای نرم‌افزاری برای توزیع ترافیک بین ایستگاه‌ها استفاده می‌شود؛ به طوری‌که کارایی شبکه در حد بالایی حفظ شود. وب سرویس‌ها عمدتاً در برنامه‌های پیام فوری آنلاین استفاده می‌شود، این فناوری برای ارتباطات بلادرنگ بین طرف‌های مختلف می‌باشد با این‌حال وجود و یا دسترس بودن برنامه کاربردی مهم می‌باشد. مدلی ارائه شده است که از XMPP برای توازن بار استفاده می‌کند. کلاینت‌هایXMPP اطلاعات حاضر را به سرور حاضر XMPP می‌فرستند و جریان‌های XML جزئیات اطلاعات حاضر مشتریان تولید شده توسط این سرورها را در بردارد. با استفاده از یک موازنه کننده بار در بالای یک سرور XMPP اجازه می‌دهد درخواست‌های ورودی توسط سرویس‌های عمومی، اولویت‌بندی و به کار گرفته شوند در این الگوریتم از میان معیارهای توازن بار، معیار زمان پاسخ و کارایی در نظر گرفته‌شده است[۲۴].

- الگوریتمMin-Min

این الگوریتم با مجموعه‌ای از وظایف تخصیص داده نشده، شروع می‌شود. اول از همه، زمان اتمام حداقل برای تمام وظایف یافت می‌شود. سپس در میان این زمان‌های حداقل، کمترین مقدار انتخاب می‌شود که حداقل زمان میان تمام وظایف روی هر منبع موجود است. سپس مطابق با زمان حداقل، وظیفه روی ماشین مربوطه زمان‌بندی می‌شود. سپس زمان اجرا برای همه‌ی وظایف دیگر روی ماشین با اضافه کردن زمان اجرای وظیفه‌ی تخصیص داده شده به زمان اجرای دیگر وظایف روی ماشین به روزرسانی می‌شود و وظیفه تخصیص داده شده از لیست وظایفی که به ماشین تخصیص داده می‌شود، حذف می‌شود. دوباره همین روال تا زمانی که همه‌ی وظایف به منابع تخصیص داده شوند، دنبال می‌شود؛ اما این روش یک اشکال اصلی دارد که می‌تواند منجر به گرسنگی شود. در این الگوریتم از میان معیارهای توازن بار، معیار بهره‌برداری از منابع، سربار، توان عملیاتی، زمان پاسخ و کارایی در نظر گرفته‌شده است [۲۵].

- الگوریتمMin-Max

تقریباً همان الگوریتم Min-Min است به جز موارد زیر: بعد از یافتن زمان‌های اجرای حداقل، بیشترین مقدار انتخاب می‌شود که زمان حداکثر میان تمام وظایف روی هر منبع موجود است. سپس مطابق بازمان حداکثر، وظیفه روی ماشین مربوطه زمان‌بندی می‌شود. سپس زمان اجرا برای همه‌ی وظایف دیگر روی ماشین با اضافه کردن زمان اجرای وظیفه‌ی تخصیص داده شده به زمان اجرای دیگر وظایف روی ماشین به‌روزرسانی می‌شود و وظیفه تخصیص داده‌شده از لیست وظایفی که به ماشین تخصیص داده می‌شود، حذف می‌شود. در این الگوریتم از میان معیارهای توازن بار، معیار بهره برداری از منابع، سربار، توان عملیاتی، زمان پاسخ و کارایی در نظر گرفته‌شده است[۲۵].

- الگوریتم‌های توازن بار دومرحله‌ایOLB + LBMM

یک الگوریتم زمان‌بندی دو مرحله‌ای پیشنهادشده است که الگوریتم‌های زمان‌بندی (LBMM +OLB) فرصت توازن بار و توازن بار (Min-Min) را برای استفاده از اجرای بهتر و حفظ توازن بار سیستم، ترکیب می‌کند. الگوریتم زمان‌بندی OLB هر گره را برای رسیدن به هدف توازن بار، در حالت کار نگه می‌دارد و الگوریتم زمان‌بندی LBMM برای کاهش زمان اجرای هر وظیفه روی گره استفاده می‌شود، درنتیجه زمان اجرای کل را کاهش می‌دهد. محیط استفاده از این الگوریتم شبکه‌های شبکه ابری سه سطحی است و در آن، معیار کارایی و بهره‌برداری از منابع در نظر گرفته‌شده است. این الگوریتم ترکیبی، در استفاده‌ی مؤثر از منابع کمک می‌کند و بهره‌وری را افزایش می‌دهد نتایج بهتری را نسبت به کاوش زنبورعسل، نمونه‌برداری تصادفی و خوشه‌بندی فعال ارائه می‌دهد[۲۶].

www.SID.ir

Archive of SID

۴- بحث و ارزیابی

به منظور مقایسه و ارزیابی روش های مذکور در بخش ۳، از چندین معیار استفاده می شود که به شرح زیر بیان می گردد:

۱- سربار محاسباتی: کلیه هزینه هایی که طی فرآیند اجرای الگوریتم ایجاد می شود به طوری که نمی توان به صورت مستقیم یا مشخص به هزینه های اصلی الگوریتم نسبت داد، سربار محاسباتی می گویند.

۲- توان عملیاتی: به متوسط نرخ رسیدن به پاسخ صحیح توسط الگوریتم مربوطه، توان عملیاتی گفته می شود.

۳- کارایی: کارایی یک الگوریتم، با توجه به میزان انرژی مصرفی آن سنجیده می شود. هرچقدر میزان پاسخ صحیح نسبت به انرژی مصرفی بیشتر باشد، کارایی نیز بیشتر است.

۴- زمان پاسخ: متوسط زمانی که یک الگوریتم برای رسیدن به پاسخ صحیحی مصرف می کند، زمان پاسخ می گویند.

با توجه به تعاریف فوق، جدول زیر به مقایسه نتایج حاصل از الگوریتم های پیشنهادی می پردازد.

جدول ۱- مقایسه روش‌های ارائه‌شده در مقاله

نویسندگان	الگوریتم پیشنهادی	نتایج
نها و همکاران [۱۷]	الگوریتم خوشه‌بندی فعال (Active Clustering)	سربار محاسباتی ندارد-توان عملیاتی بالا-کارایی بالا-زمان پاسخ بالا
ژونلای و همکاران [۱۸]	الگوریتم کلونی مصنوعی زنبورعسل (ABC)	سربار محاسباتی دارد-توان عملیاتی بالا-کارایی بالا-زمان پاسخ کم
ناروکی و همکاران [۱۹]	الگوریتم نمونه گیری تصادفی	سربار محاسباتی دارد-توان عملیاتی بالا-کارایی بالا-زمان پاسخ کم
اسووشکی و همکاران [۲۰]	روش کارتن	سربار محاسباتی دارد-توان عملیاتی بالا-کارایی بالا-زمان پاسخ بالا
پریانکا و همکاران [۲۱]	الگوریتم رویداد محور	سربار محاسباتی ندارد-توان عملیاتی بالا-کارایی بالا-زمان پاسخ بالا
کومار و همکاران [۲۲]	سیاست توازن بار مرکزی برای ماشین‌های مجازی	سربار محاسباتی ندارد-توان عملیاتی بالا-کارایی بالا-زمان پاسخ پایین
صوبهرا و همکاران [۲۳]	استفاده از منطق فازی در توازن بار	سربار محاسباتی ندارد-توان عملیاتی پایین-کارایی بالا-زمان پاسخ پایین
صوفیان و همکاران [۲۴]	الگوریتم پیام گرا	سربار محاسباتی دارد-توان عملیاتی بالا-کارایی بالا-زمان پاسخ کم
کودهارای و همکاران [۲۵]	الگوریتم min min	سربار محاسباتی ندارد-توان عملیاتی بالا-کارایی بالا-زمان پاسخ بالا
کودهارای و همکاران [۲۵]	الگوریتم max min	سربار محاسباتی دارد-توان عملیاتی بالا-کارایی بالا-زمان پاسخ کم
میشارا و همکاران [۲۶]	الگوریتم توازن بار دومرحله‌ای OLB + LBMM	سربار محاسباتی دارد-توان عملیاتی بالا-کارایی بالا-زمان پاسخ کم

شبکه‌های ابری به منظور ارائه خدمات مؤثر به کاربران ایجاد گردید. مهم ترین رکن در این شبکه‌ها، بازدهی بالای شبکه و رضایت مشتریان آن‌ها از لحاظ زمان و هزینه می‌باشد. هدف از توازن بار، درگیر کردن تمامی سرورهای موجود در شبکه برای اجرای نیازهای کاربران می‌باشد. در نتیجه با استفاده از روش‌های توازن بار می‌توان هزینه شبکه و سرعت پاسخگویی به نیازهای کاربران را افزایش داد.

الگوریتم های بهینه سازی و جستجوی تصادفی و تکاملی روش های نوین و کارامدی هستند که به ویژه برای یافتن جواب های بهینه سراسری مسائل به کار می روند. ویژگی تصادفی بودن این الگوریتم ها مانع از گیر افتادن در نقاط بهینه موضعی می شوند. در مسائل بهینه سازی عملی مانند طراحی های مهندسی، مدیریت سازمان ها و سیستم های اقتصادی معمولاً توجه اصلی معطوف به بدست آوردن جواب های بهینه سراسری است. بسیاری از این الگوریتم ها الهام گرفته شده از سیستم های زیستی هستند و پاسخ مناسبی در توازن بار از خود ارائه می دهند.

www.SID.ir

Archive of SID

می‌توان نتیجه گرفت که استفاده از الگوریتم‌های تکاملی می‌توان توازن بار را بهبود بخشد. با مقایسه روش های پیشنهادی، الگوریتم تکاملی زنبور، مراتب نتایج بهتر و متعادل تری نسبت به سایر روش ها از خود نشان داده است.

۵. نتیجه‌گیری

در این مقاله به بررسی روش‌های ارائه شده برای توازن بار در شبکه رایانش ابری پرداخته شده است. هر تغییر و مفهوم جدیدی در دنیای فناوری، مشکلات و پیچیدگی‌های خاص خود را دارد. بهره‌گیری از شبکه ابری نیز از این قاعده مستثنی نبوده و چالش‌های فراوانی را پیش روی صاحب‌نظران این حوزه قرار داده است که از آن جمله می‌توان به مواردی نظیر: توازن بار، امنیت، قابلیت اطمینان، مالکیت، پشتیبان گیری از داده ها، قابلیت حمل داده ها، پشتیبانی از چندین پلتفرم اشاره کرد. یکی از چالش‌های مهم در زمینه شبکه ابری، حفظ توازن بار بهینه در محیط ابر است. شبکه ابری اصطلاحاً شامل مجازی سازی، محاسبات توزیعی، شبکه، نرم افزار و سرویس های وب می باشد. یک ابر از چندین عنصر مثل کلاینت ها، مراکز داده و سرور های توزیع شده تشکیل شده است که در آن مباحث تحمل خطا، دسترسی بالا، مقیاس پذیری، انعطاف پذیری، کاهش سربار برای کاربران، کاهش هزینه ی مالکیت، سرویس دهی به درخواست ها و غیره وجود دارد. تمرکز روی این مسائل برای ایجاد یک الگوریتم توازن بار کارآمد می باشد. توازن بار از سنگین شدن بار روی ماشین های مجازی جلوگیری می کند در واقع اگر حجم کار یک منبع در محیط محاسباتی بیش از حد باشد باید حجم کار به یک منبع دیگر منتقل شود و از وضعیتی که برخی از گره ها بیش از حد بارگذاری می شوند در حالی که دیگران کم بار می شوند اجتناب می کند و این باعث کاهش زمان پاسخ دهی، استفاده بهینه از منابع، رضایت مشتری و ... می شود.

۶– منابع

1. J. Broberg., S. Venugopal and R. Buyya,(2016) ,"Market-oriented Grid and utility computing: The state-of-the-art and future directions", Journal of Grid Computing,Elsevier, pp:255,276.

2. Wood T,(2017) ."Black-box and Gray-box Strategies for Virtual Machine Migration", Proceedings of the 4th Int'l Conference on Networked Systems Design & Implementation, IEEE.

3. B. Yagoubi., Y. Slimani,(2015) "Task load balancing strategy for grid computing", Journal of Computer Science, pp:186,194.

4. V. Venkatesa Kumar., R. Revathi and Newlin Rajkumar,(2014) ,"An Assessment On Various Laod Balancing Techniques In Cloud Computing" ,IJAICT, 1(8).

5. Jasmin James and Dr. Bhupendra Verma,(2016) ,"Efficient VM load balancing algorithm for a cloud computing environment", pp:1258,1368.

6. Mamta Khanchi and Sanjay Tyagi, (2016), "An Efficient Algorithm For Load Blancing In Cloud Computing" , International Journal Of Engineering Sciences & Research Technology , IJESRT.

7. Md. Shahjahan Kabir, Kh. Mohaimenul Kabir and Dr. Rabiul Islam,(2015) ,"Process Of Load Balancing In Cloud Computing Using Genetic Algorithm" , Electrical & Computer Engineering: An International Journal (ECIJ), pp:4,.

8. Auday Al-Dulaimy., Ahmed Zekri., Wassim Itani and Rached Zantout, (2015), "Towards Solving the Problem of Virtual Machine Placement in Cloud Computing: A Job Classification Approach" , Journal of Computer Sciences , Science Publications, Elsevier.

9. christos stergiou, kostas e.psannis, (2018), "secure integration of iot and cloud computing", elsevier , future generation computer systems, 78(3), pp: 964-975.

10. Mumtaz M.Ali AL-Mukhtar and Asraa Abdulrazak Ali Mardan ,(2014) ,"Performance Evaluation of Private Clouds Eucalyptus versus CloudStack" , (IJACSA) International Journal of Advanced Computer Science and Applications, 5(5).

www.SID.ir

Arcl

11. ping li, jin li, (2018), "privacy-preserving outsourced classification in cloud computing",springer, 21(1), pp: 277–286.

12. jinli, yinghuizhang, (2018), "secure attribute-based data sharing for resource-limited users in cloud computing", elsevier, computers & security, volume 72, pp 1-12.

13. blessonvarghese, rajkumarbuyya, (2018), "next generation cloud computing: new trends and research directions", elsevier, future generation computer systems,79(3), pp 849-861.

14. Sreenivas Velagapudi, M.Prathap and Kemal Mohammed,(2014) ,"Load Balancing Techniques: Major Challenge in Cloud Computing – A Systematic Review", IEEE, International Conference on Electronics and Communication Systems (ICECS), IEEE.

15. Prithpal Mohini Singh and Shaveta Angurala,(2015) ,"Refined VM-Assign Load Balancing in Cloud Computing", International Journal of Advanced Research in Computer and Communication Engineering, 4(10).

16. Shridhar. G. Donamal,(2014) ,"Optimal Load-Balancing in Cloud Computing by efficient utilization of virtual resources", IEEE.

17. Neha Roy and Rishabh Jain,(2015) ,"Cloud Computing: Architector And Concept Of Virtualization", International Journal of Science, Technology & Management,4(2).

18. huanlai xing, fuhong song, lianshan yan, wei pan, (2019), "a modified artificial bee colony algorithm for load balancing in network-coding-based multicast", springer, soft computing, 23(15), pp: 6287–6305.

19. Nawrocki, P. and W. Reszelewski,(2017) ,"Resource usage optimization in Mobile Cloud Computing"., ppl 1-12.

20. Srushti Patel., Hiren Patel and Nimisha Patel,(2016) ,"Dynamic Load Balancing Techniques for Improving Performance in Cloud Computing", International Journal of Computer Applications, pp:138,3.

21. Priyanka Singh., Palak Baaga and Saurabh Gupta,(2016) ,"Assorted Load Balancing Algorithms in Cloud Computing: A Survey", International Journal of Computer Applications, Elsevier, pp:143,147.

22. Rakesh Kumar and Neha Gupta,(2014) ,"Open Source Solution for Cloud Computing Platform Using OpenStack ",Conference Paper , IEEE.

23. Subhra Priyadarshini Biswal, Satya Prakash Sahoo, (2018),"Fuzzy Logic Based Cost and Energy Efficient Load Balancing in Cloud Computing Environment", IEEE, 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS).

24. soufiane maguerra, azedine boulmakoul, (2019), "load balancing of distributed actors in an asynchronous message processing boundary", smc '19 proceedings of the new challenges in data sciences: acts of the second conference of the moroccan classification society.

25. Ankita Choudharya., Shilpa Ranab and K.J. Matahaic,(2015), "A Critical Analysis of Energy Efficient Virtual Machine Placement Techniques and its Optimization in a Cloud Computing Environment", International Conference on Information Security & Privacy (ICISP2015), IEEE, pp:11,22.

26. Nitin Kumar Mishra and Nishchol Mishra,(2016) ,"CELBT: An Algorithm for Efficient Cost based Load Balancing in Cloud Environment", International Journal of Computer Applications, 134(1).

www.SID.ir