

محاسبات جنگلی: محاسبات سطح بالا و رای ابرها، گریدها و خوشه‌های محاسباتی

اسماعیل آسیابی^۱ و محسن شریفی^۲

^۱دانشگاه علم و صنعت ایران، دانشکده مهندسی کامپیوتر، e_asyabi@comp.iust.ac.ir

^۲دانشگاه علم و صنعت ایران، دانشکده مهندسی کامپیوتر، msharifi@iust.ac.ir

چکیده - امروزه تعداد و تنوع کاربردهایی که نیازمند منابع محاسباتی فوق‌العاده بالا هستند رشد روزافزونی داشته است. این کاربردها غالباً در حوزه‌های علمی، مهندسی و مالی هستند. پاسخگویی مناسب به این نیازهای محاسباتی روبه‌رشد یکی از چالش‌های پیش‌روی دانشمندان است. به‌کارگیری همزمان چندین بستر محاسباتی راهی متداول برای دستیابی به قدرت محاسباتی بیشتر است. یکی از این راه‌کارها، استفاده از سیستم‌های محاسبات جنگلی است. در یک سیستم محاسبات جنگلی چندین بستر محاسباتی سطح بالا مانند گریدها، خوشه و ابرکامپیوتر به طور همزمان برای اجرای یک کاربرد به کار گرفته می‌شوند. دلایل مختلفی برای استفاده از سیستم‌های محاسبات جنگلی وجود دارد. ممکن است نیاز محاسباتی یک کاربرد بیشتر از قدرت محاسباتی فراهم آمده توسط یک سیستم محاسباتی باشد و برای پاسخگویی به این نیاز محاسباتی چاره‌ایی جز استفاده از سیستم محاسبات جنگلی نباشد. به علاوه بخش‌های مختلف یک کاربرد ممکن است نیازهای محاسباتی متفاوت داشته باشند به‌طوری که هیچ سیستمی به تنهایی پاسخگوی همه این نیازها نباشد. از سوی دیگر، محدودیت در تعداد و قدرت منابع محاسباتی و نیازهای روزافزون محاسباتی، چالش دیگری است که توزیع این محاسبات در بسترهای محاسباتی متفاوت می‌تواند راه‌حلی برای رفع این چالش باشد.

کلیدواژه‌ها - محاسبات سطح بالا، محاسبات جنگلی، محاسبات ابری، گریدها، خوشه.

کمیود منابع محاسباتی لزوم استفاده از بستر محاسبات جنگلی را به عنوان یک راه‌حل نمایان می‌سازد.

۱. مقدمه

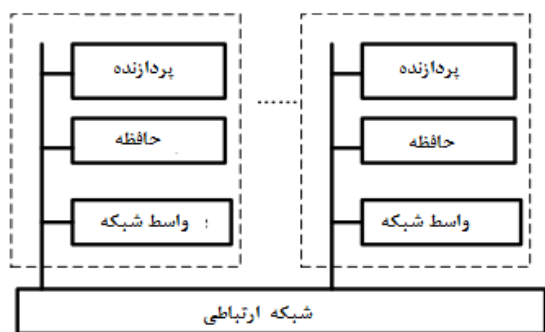
دلایل مختلفی برای استفاده از محیط محاسبات جنگلی وجود دارد. ممکن است نیاز محاسباتی کاربردی بیشتر از قدرت محاسباتی هر سیستم موجود باشد برای پاسخگویی به این نیاز محاسباتی چاره‌ایی جز استفاده از سیستم محاسبات جنگلی نخواهیم داشت. به علاوه بخش‌های مختلف یک کاربرد ممکن است نیازهای محاسباتی متفاوت داشته باشند. طوری که هیچ سیستم تکی به تنهایی پاسخگوی همه این نیازها نیست. از طرف دیگر محدودیت در تعداد و قدرت منابع محاسباتی و نیاز-های روزافزون محاسباتی، چالشی است که توزیع این محاسبات در بسترهای محاسباتی متفاوت می‌تواند راه حلی برای رفع این چالش باشد [2].

در این مقاله بسترهای محاسبات سطح بالا بررسی، چالش‌های آن‌ها شمرده و در نهایت مقایسه می‌شوند. در ادامه در بخش ۲ ابرکامپیوترها بررسی می‌شوند. محاسبات خوشه‌ایی در بخش ۳ توضیح داده می‌شود. در بخش ۴ محاسبات گریدها و

نیاز به قدرت محاسباتی بیشتر یکی از چالش‌هایی است، که علوم مختلفی همانند فیزیک، شیمی، علوم مهندسی و بسیاری دیگر از علم‌ها و کاربردها به آن مواجه هستند [1]. این کاربردها مقدار بسیار زیادی داده و سیکل‌های محاسباتی نیاز دارند.

ابرکامپیوتر، خوشه، گریدها و ابر از جمله بسترهایی هستند که امروزه به منظور محاسبات سطح بالا به کار می‌روند. اما افزایش پیوسته نیاز محاسباتی، به کارگیری همزمان چندین بستر محاسبات سطح بالا را به عنوان راه حلی برای رفع این نیاز مطرح می‌سازد. سیستم توزیعی حاصل محاسبات جنگلی نام دارد.

طبق تعریف به‌کارگیری همزمان چندین بستر محاسباتی همانند ابر، گریدها، خوشه و ابرکامپیوتر، برای اجرای یک کاربرد محاسبات جنگلی نام دارد [1,2]. در واقع انتخاب این بستر برای انجام محاسبات سطح بالا بیشتر از آنکه یک انتخاب آزادانه باشد، یک انتخاب از روی اجبار است. نیاز به قدرت محاسباتی بالاتر و

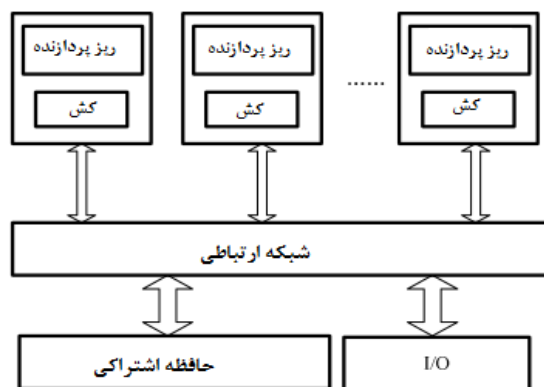


شکل ۲: معماری MPP

به دنبال آن در بخش ۵ محاسبات ابری بررسی می‌گردد. در بخش ۶ سیستم محاسبات جنگلی و چالش‌های پیش‌روی این سیستم توضیح داده می‌شود و در انتها در بخش ۷ جمع‌بندی و نتیجه‌گیری آورده می‌شود.

۲. ابر کامپیوترها

در سال ۱۹۸۰ با پیشرفتی که در حوزه مدارهای مجتمع صورت گرفت، ریزپردازنده‌ها اختراع و قدرت آن‌ها همواره افزایش یافت و معماری‌هایی همانند معماری SMP مبنای معماری بسترهای انجام محاسبات سطح بالا گردیدند. همانطور که در شکل ۳۱. نمایش داده شده است، در معماری SMP چندین ریزپردازنده توسط یک گذرگاه ارتباطی به یک حافظه اشتراکی متصل‌اند. معماری SMP دارای مشکلاتی است. هزینه بالای عملیات ورودی خروجی و همچنین محدودیت در تعداد پردازنده‌ها باعث ایجاد مشکلات مقیاس‌پذیری در این بستر شده است [4].

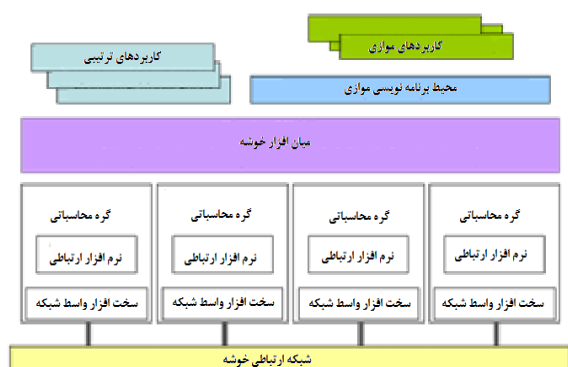


شکل ۱: معماری SMP

معماری دیگری که به منظور محاسبات سطح بالا در سال ۱۹۹۰ به کار گرفته شده است معماری MPP است. معماری MPP در شکل ۲ نمایش داده شده است. در این معماری تعداد زیادی گره محاسباتی که هر کدام دارای پردازشگر، حافظه و واسط شبکه خاص خود است، در بستر یک شبکه ارتباطی به وسیله یک پروتکل ارسال پیام با هم در ارتباط‌اند. امروزه بسیاری از ابر کامپیوترها بر پایه معماری MPP است. اما این بستر دارای چالش‌هایی است، که مهمترین آن هزینه زیاد ساخت چنین سیستمی است [4].

۳. محاسبات خوشه‌ای

خوشه‌ها ره‌یافت متمایزی برای رسیدن به قدرت محاسباتی بالاتر با هزینه کمتر در مقایسه با ابر کامپیوترها هستند. طبق تعریف یک خوشه مجموعه‌ای از کامپیوترهای غالباً همگن است، که به صورت موازی یا توزیعی بر بستر یک شبکه پرسرعت برای حل یک مسئله با همدیگر همکاری می‌کنند. معماری یک خوشه در شکل ۳ نمایش داده شده است.



شکل ۳: معماری یک خوشه

معمولاً خوشه‌ها در سه دسته خوشه برای دسترس‌پذیری بالاتر، خوشه برای توازن بار و خوشه برای محاسبات طبقه‌بندی می‌شوند. در خوشه محاسباتی قدرت محاسباتی بالاتر از طریق همکاری گره‌های خوشه بدست می‌آید. کاری که برای اجرا به یک خوشه سپرده می‌شود، به کارهای کوچکتر شکسته می‌شود و هر کار کوچک توسط یکی از گره‌های خوشه به صورت موازی با سایر کارها انجام می‌شود. در واقع گره‌های موجود در خوشه امکانات خود را تحت یک کامپیوتر مجازی تک به اشتراک می‌گذارند.

۵. محاسبات ابری

محاسبات ابری بستری جدید برای محاسبات سطح بالا است. این الگو زیرساخت محاسبات را به شبکه انتقال می‌دهد و در نتیجه هزینه‌های نگهداری و مدیریت منابع سخت‌افزاری و نرم‌افزاری محاسبات سطح بالا را کاهش می‌دهد و کاربران را قادر می‌سازد از هر جای دنیا سرویس این بستر محاسباتی را مورد استفاده قرار دهند.

طبق تعریف محاسبات ابری مدلی برای دسترسی آسان و بر حسب تقاضا به مجموعه‌ای از منابع محاسباتی اشتراکی قابل تنظیم، تحت بستر شبکه است. منابع در کمترین زمان با کمترین هزینه و تعامل با فراهم کننده سرویس آماده می‌شوند [3].

محیط محاسبات ابری سرویس‌های خود را در ۳ سطح ارائه می‌دهد:

• زیرساخت به عنوان سرویس

در این لایه سرویس‌هایی مانند پردازش، فضای ذخیره سازی، شبکه و سایر منابع محاسباتی به عنوان سرویس ارائه می‌شود. بعنوان مثال، Amazone EC2 ابری است که سرویس زیرساخت ارائه می‌دهد [7].

• بستر به عنوان سرویس

در این سطح همه سرویس‌های که برای توسعه یک کاربرد لازم است، ارائه می‌شود. این سرویس‌ها شامل سرویس‌هایی برای ساخت، به کارگیری، تست، نگهداری و سایر سرویس‌های لازم است. بنابراین توسعه دهنده فقط بر روی عملکردهای مهم سرویس خود تمرکز می‌نماید.

• نرم‌افزار به عنوان سرویس

در این نوع سرویس کاربران زیادی می‌توانند از نرم‌افزاری که فراهم آورنده سرویس میزبان آن است استفاده نمایند. کم شدن هزینه‌های خرید، نگهداری، به روز نمودن نرم-افزار و خرید سخت‌افزار مورد نیاز از مزایای این سرویس است.

۶. محاسبات جنگلی

محاسبات جنگلی یک بستر محاسباتی توزیع شده است که منابع

گره‌های موجود در خوشه از طریق یک تکنولوژی ارسال پیام مانند MPI (واسط مبادله پیام) یا PVM (ماشین موازی مجازی) با هم در ارتباط‌اند. کاربردهایی که مناسب اجرا در محیط محاسبات خوشه‌ای است، باید به صورت مجموعه‌ای از وظایف باشد که می‌توانند بر روی گره‌های مختلف خوشه به اجرا درآیند. ضمن آنکه هزینه ارتباطات نیز حداقل ممکن باشد [4].

توسعه کاربردهای مناسب محیط محاسبات خوشه‌ای از جمله چالش‌های این محیط است. بستر محاسبات سطح بالا باید در مقابل افزایش درخواست‌های محاسباتی مقیاس‌پذیر باشد و بتواند بدون خللی در کارایی سیستم پاسخگوی این درخواست‌ها باشد. مقیاس‌پذیری از دیگر چالش‌های محیط محاسبات خوشه-ای است، که لزوم استفاده از بستر دیگری را برای غلبه بر این گونه مشکلات روشن می‌سازد [6].

۴. محاسبات گریدی

طبق تعریف گرید یک سیستم توزیعی و موازی است. که قابلیت اشتراک‌گذاری، انتخاب و اجماع منابع توزیع شده در گستره جغرافیا را به صورت پویا و در زمان اجرا بر اساس ویژگی‌های منابع، مانند در دسترس بودن، قدرت محاسباتی، هزینه و کیفیت سرویس مورد نیاز کاربران، فراهم می‌آورد [5].

در واقع در محاسبات گریدی کامپیوترهای که ممکن است در دامنه‌های مدیریتی متفاوت باشند، برای دسترسی به یک هدف مشترک و یا حل یک مسئله به کار گرفته می‌شوند. این شبکه از منابع ممکن است موقتی باشد و خیلی سریع حذف شود. استراتژی گرید تامین منابع محاسباتی همانند صنایع همگانی مثلاً شبکه برق است. کافی است در این بستر کاربر به شبکه متصل و منابع محاسباتی را بدون اطلاع از چگونگی و محل فراهم‌آوری این منابع به کار گیرد.

بستر محاسبات گریدی نیز دارای چالش‌هایی است. در بستر گرید منابع در مالکیت و مدیریت سازمان‌های مختلف قرار دارند [4]. در نتیجه ممکن است، هر زمان یک گره از شبکه گرید حذف و یا اضافه شود. مدیریت این وضعیت در کنار مشکلات مربوط به زمانبندی انجام کارها و رفع مشکلات مربوط به ناهمگونی زیاد در این بستر از جمله چالش‌های گرید است [6]. به علاوه ارتباطات ضعیف گره‌ها و پردازش توزیع شده کارها موجب ایجاد پیچیدگی برنامه‌نویسی کاربردها در این بستر می‌شود.

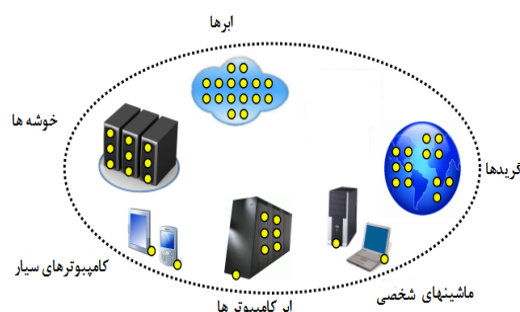
رزرو شود و یا منبع محاسبات ابری از طریق کارت اعتباری مورد دسترس قرار گیرد.

چالش دیگری که محیط محاسبات جنگلی با آن روبه‌رو است، مشکلات مربوط به ارتباطات گره‌های موجود در سیستم است. اصولاً برخی منابع موجود در سیستم محاسبات جنگلی همانند خوشه‌ها و ابرکامپیوترها برای ارتباط با دنیای بیرون طراحی نشده‌اند [2]. حل مشکل ارتباطی این گونه گره‌های محاسباتی موجود در سیستم و مشکلات ارتباطی دیگر مثل خرابی و محدودیت پهنای باند، از جمله مشکلات دیگری است که سیستم محاسبات جنگلی باید در جستجوی راه‌حلی برای آن‌ها باشد [1].

سیستم محاسبات جنگلی می‌باید محاسبات را مستقل از نوع منابع و میان‌افزارها به صورت شفاف توزیع نماید. سیستم مذکور دارای شبکه ارتباطی مطمئن و تحمل‌پذیر در مقابل انواع خطا است. روشن است که برای دستیابی به یک سیستم محاسبات جنگلی ایده‌آل باید چالش‌های فراوانی رفع شود. عملاً Ibis تلاشی برای پیاده‌سازی اهداف ذکر شده است.

یک چهارچوب نرم‌افزاری Ibis است، که در تلاش است چالش‌های سیستم محاسبات جنگلی را رفع نماید [8]. این سیستم تعدادی از نیازهای پایه‌ای سیستم محاسبات جنگلی را از طریق این چهارچوب نرم‌افزاری برطرف می‌سازد. Ibis فقط تعدادی از نیازهای پایه‌ای سیستم محاسبات جنگلی را برآورده و تا رسیدن به سیستم محاسبات جنگلی شفاف که چالش‌های ذکر شده را برطرف نماید، فاصله زیادی وجود دارد. در ادامه ویژگی‌های سیستم محاسبات جنگلی ایده‌آل با سایر سیستم‌های محاسبات سطح بالا در جدول ۱ مقایسه شده است.

مختلف محاسباتی را به صورت همزمان جهت اجرای یک کاربرد به کار می‌گیرد [1]. همانطور که در شکل ۴ نشان داده شده است، این بستر شامل منابع مختلف محاسباتی همچون گریدها، خوشه، ابر و یا حتی کامپیوترهای شخصی است. یک سیستم محاسبات جنگلی که می‌تواند بسترهای متفاوت محاسباتی را به صورت شفاف و کارا به کار گیرد با چالش‌های فراوانی روبه‌رو است.



شکل ۴: سیستم محاسبات جنگلی

سیستم محاسبات جنگلی بسیار ناهمگون است. منابع موجود در این سیستم خصوصیات پایه‌ای متمایزی دارند. خصوصیتی مانند معماری پردازشگر، میزان حافظه و کارایی در منابع مختلف موجود در سیستم متفاوت است و همین امر باعث پیچیدگی زیادی در به‌کارگیری آن‌ها به عنوان یک تک سیستم، برای اجرای یک کاربرد شده است. سیستم ممکن است برای غلبه بر این ناهمگونی بخش‌های متفاوت یک کاربرد را برای منابع محاسباتی متفاوت دوباره‌نویسی و یا ترجمه دوباره نماید که عملاً دارای هزینه زیادی خواهد بود.

در سیستم محاسبات جنگلی واحد مدیریت مرکزی وجود ندارد. منابع موجود در سیستم راه‌های دسترسی متفاوتی دارند. مثلاً یک منبع محاسباتی گریدها باید قبل از به‌کارگیری

جدول ۱: مقایسه بسترهای محاسبات سطح بالا

خوشه	گریدها	ابر	جنگل	
-	-	✓	✓	قابلیت اطمینان
✓	-	✗	-	امنیت
✗	-	✓	✓	واسط کاربری آسان
-	-	✓	✓	مجازی سازی
✓	✓	-	-	استاندارد سازی شده
✗	✗	✓	-	مدل تجاری
✗	✓	✓	-	SOA
✗	✓	✓	✓	چند کاربری
✗	✓	✓	✓	ناهمگونی
✗	-	✓	✓	مقیاس پذیری
✗	✗	✓	✓	هزینه کمتر

۷. جمع‌بندی و نتیجه‌گیری

در این مقاله بسترهای محاسبات سطح بالا بررسی و چالش‌های آنها شماره شده شدند. سیستم محاسبات جنگلی که سایر بسترهای محاسبات سطح بالا را در به طور همزمان به کار می‌گیرد، معرفی و چالش‌های آن بررسی گردیدند. بسیاری از چالش‌هایی که سیستم‌های محاسبات جنگلی با آن روبه‌رو هستند در گرید نیز مطرح بوده‌اند. سیستم Ibis یک نمونه عملیاتی از سیستم محاسبات جنگلی است که فناوری‌های گرید را برای ساخت سیستم محاسبات جنگلی به کار گرفته است. این سیستم فاصله زیادی با یک سیستم محاسبات جنگلی کارا و شفاف دارد و فقط قدم‌های اول از یک راه طولانی است که سعی در برآوردن نیازهای پایه و اساسی سیستم محاسبات جنگلی دارد.

مراجع

- [1] F. Seinstra, J. Maassen, R. V. van Nieuwpoort, N. Drost, T. Van Kessel, B. van Werkhoven, J. Urbani, C. Jacobs, T. Kielmann, and H. E. Bal, "Jungle Computing: Distributed Supercomputing Beyond Clusters, Grids, and Clouds", Springer, 2011.
- [2] NielsDrost, Jason Maassen, Maarten A.J. van Meersbergen and Henri E. Bal, "High-Performance Distributed Multi-Model / Multi-Kernel Simulations: A Case-Study in Jungle Computing", 21st International Heterogeneity in Computing Workshop, China, 2012.
- [3] SherifSakr, Anna Liu, Daniel M. Batista and Mohammad Alomari, "A Survey of Large Scale Data Management Approaches in Cloud Environments", IEEE Journal of Communications Surveys & Tutorials, 2011.
- [4] Li Bo Zhou, Zhenliu Wang and Xiangfeng, "A Survey of HPC Development", IEEE International Conference on Computer Science and Electronics Engineering (ICCSEE), 23-25 March 2012.
- [5] M.Chetty and R. Buyya, "Weaving Computational Grids: How Analogous Are They with Electrical Grids?", Computing in Science and Engineering (CiSE), 2002.
- [6] Naidila Sadashiv and S. M Dilip Kumar, "Cluster, Grid and Cloud Computing: A Detailed Comparison", IEEE 6th International Conference on Computer Science & Education (ICCSE 2011), Singapore, 2011.
- [7] Amazon Web Services LLC, "Amazon Elastic Compute Cloud (Amazon EC2)", <http://aws.amazon.com/ec2/>, Last Visit: October 2012.
- [8] Ibis, "Ibis: Computing in the Jungle", <http://www.cs.vu.nl/ibis/index.html>, Last Visit: October 2012.