بسم الله الرحمن الرحيم

دانشکده مهندسی
گروه کامپیوتر

رساله دکتری

ارائه یک معماری شناسایی منابع در سیستم های مشبک داوطلبانه نظیربه نظیر

تکتم غفاریان

استاد راهنما
دکتر حسین دلداری

اساتید مشاور
پروفسور محمدحسین یغمایی
دکتر بهمن جوادی
پروفسور راج کومار بویا

بهمن ماه ۹۱

تعهد نامه

اینجانب تکتم غفاریان دانشجوی دوره دکتری رشته مهندسی کامپیوتر دانشکده مهندسی دانشگاه فردوسی مشهد، نویسنده رساله " ارائه یک معماری شناسایی منابع در سیستم های مشبک نظیر به نظیر" تحت راهنمایی جناب آقای دکتر دلداری متعهد می شوم:

□ تحقیقات در این پایان نامه توسط اینجانب انجام شده و از صحت و اصالت برخوردار است.

□ در استفاده از نتایج پژوهش های محققان دیگر به مرجع مورد استفاده استناد شده است.

□ مطالب مندرج در پایان نامه تاکنون توسط خود و یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.

□ کلیه حقوق معنوی این اثر متعلق به دانشگاه فردوسی مشهد می باشد و مقالات مستخرج با نام

" دانشگاه فردوسی مشهد" و یا " Ferdowsi University of Mashhad" به چاپ خواهد رسید.

□ حقوق معنوی تمام افرادی که در بدست آمدن نتایج اصلی رساله تاثیر گذار بوده اند در مقالات مستخرج از رساله رعایت شده است.

□ در کلیه مراحل انجام این رساله، در مواردی که از موجود زنده (یا بافت های آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.

□ در کلیه مراحل انجام این رساله، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است، اصل رازداری، ضوابط و اصول اخلاق انسانی رعایت شده است.

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزارها و تجهیزات ساخته شده) متعلق به دانشگاه فردوسی مشهد می باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.

استفاده از اطلاعات و نتایج موجود در رساله بدون ذکر مرجع مجاز نمی باشد.

باسمه تعالی

برگه ارزشیابی رساله دکتری

نام و نام خانوادگی دانشجو: **مهندس غفاریان تکتم** شماره دانشجویی: **۸۶۱۶۴۳۰۰۱۶**

گروه آموزشی: **مهندسی کامپیوتر** گرایش **نرم افزار** تاریخ دفاع : **۷/۱۱/۹۱** نام و نام خانوادگی استاد راهنما: **دکتر دلداری**

عنوان رساله : **ارائه یک معماری شناسایی منابع در سیستم های مشبک داوطلبانه نظیر به نظیر**

مشخصات مقاله اصلی مستخرج از رساله (چاپ شده یا پذیرفته شده برای چاپ)

معیارهای ارزشیابی		حداکثر نمره	نمره کسب شده	ملاحظات
کیفیت نگارش	انسجام در تنظیم و تدوین مطالب،حسن نگارش و رعایت دستورالعمل	۳		
	کیفیت تصاویر، اشکال و منحنی های استفاده شده			
کیفیت علمی	بررسی تاریخچه موضوع و بیان سابقه پژوهش در موضوع	۱۱		
	ابتکار ونوآوری			
	ارزش علمی و یا کاربردی			
	استفاده از منابع ومؤاخذ به لحاظ کمی و کیفی (به روز بودن)			
	کیفیت نظرات و پیشنهادات برای ادامه تحقیق			
کیفیت ارائه	تسلط به موضوع و توانایی در پاسخگویی به سئوالات درجلسه دفاع	۳		
	نحوه ارائه (رعایت زمان، تنظیم موضوع، کیفیت ترانس پرنسی و			
مقاله	مقاله مستخرج از رساله که به تأیید استاد راهنما رسیده است	۲		
گزارش ها	تحویل به موقع گزارش ها	۱		
نمره رساله		۲۰		

ردیف	نام و نام خانوادگی	مرتبه علمی	سمت	نام دانشگاه	امضا
۱	دکتر زمانی فر	دانشیار	عضو دفاع خارجی (مدعو)	اصفهان	
	دکتر وحیدیان	استاد	عضو دفاع خارجی (مدعو)	فردوسی	
۲	دکتر نقیب زاده	استاد	عضو دفاع داخلی	فردوسی	
۳	دکتر کاهانی	دانشیار	عضو دفاع داخلی	فردوسی	
۴	دکتر دلداری	دانشیار	استادراهنما	فردوسی	
۵	دکتر یغمایی	استاد	استاد مشاور	فردوسی	
۶	دکتر عربان	استادیار	نماینده تحصیلات تکمیلی	فردوسی	

جلسه دفاع با حضور هیئت داوران تشکیل و رساله با اخذ نمره به عدد ۱۹/۲۵ / حروف نوزده و بیست و پنج صدم با درجه عالی

☐ بدون اصلاحات پذیرفته شد.

☑ با اصلاحات پذیرفته شد (دانشجو موظف است تا تاریخ ۹۱/۱۲/۱۰ رساله اصلاح شده خود را که به تأیید استادراهنما رسیده است به گروه آموزشی تحویل دهد).

☐ مردود شناخته شد.

گزارش نماینده تحصیلات تکمیلی: پذیرفته شد

نام و امضای نماینده تحصیلات تکمیلی: دکتر عربان

نام مدیر گروه : دکتر حمید پوررضا تاریخ و امضا

** (لطفاً به توضیحات مندرج در پشت برگه توجه فرمائید)**

Journal Paper

■ Toktam Ghafarian, Hossein Deldari, Bahman Javadi, Mohammad H. Yaghmaee, Rajkumar Buyya, "CycloidGrid: A proximity-aware P2P-based resource discovery architecture in volunteer computing systems," Future Generation Computer Systems (in press), 2012. doi:10.1016/j.future.2012.08.010.

■ Toktam Ghafarian, Hossein Deldari, Bahman Javadi, Rajkumar Buyya, "A proximity-aware load balancing in peer-to-peer based volunteer computing systems," Journal of Supercomputing, vol. 63, no. 1, 2013. doi: 10.1007/s11227-012-0866-7.

Conference Paper

■ Toktam Ghafarian-M., Hossein Deldari, Mohammad H. Yaghmaee, "Proximity-aware resource discovery architecture in peer-to-peer based volunteer computing system," in proceedings of IEEE 11th international conference on Computer and Information Technology (CIT 2011), August 2011, pp.83-90.

■ Toktam Ghafarian-M. , Hossein Deldari, "Task execution availability prediction in the enterprise desktop grid," in proceedings of the ninth IASTED international conference on Parallel and Distributed Computing and Networks (PDCN 2010), February 2010, pp. 16-23.

■ Toktam Ghafarian, Hossein Deldari, Mohammad-R. Akbarzadeh –T., "Multiprocessor scheduling with evolving cellular automata based on ant colony optimization," in proceeding of 14th annual international CSI Computer Conference (CSICC 2009), October 2009, pp. 431-436.

■ Toktam Ghafarian, Mohsen Kahani, "Semantic web service composition based on ant colony optimization method," in proceeding of the first international conference on Networked Digital Technologies (NDT 2009), July 2009, pp. 171-176.

این رساله را ضمن تشکر و قدردانی تقدیم می نمایم به:

پدر و مادر عزیزم، به خاطر همه تلاشهای محبت آمیزی که در دوران مختلف زندگی ام انجام داده اند

همسر مهربانم که در تمام طول تحصیل مشوق و همراه من بوده است

فرزند دلبندم پرنیا که با شکیبایی من را همراهی نموده است

قدردانی

برخود واجب می دانم، از تمام استادان و سروران ارجمند که در دوره دکتری از محضرشان کسب علم نموده ام و این رساله را به سرانجام رساندم، صمیمانه تشکر و قدردانی نمایم به ویژه از استاد بزرگوار و فرهیخته، جناب آقای دکتر دلداری کمال تشکر را دارم که با راهنمایی های بسیار ارزنده خود و حسن خلق، همواره راه گشای من در این رساله بودند. ازجناب آقای دکتر یغمایی سپاسگزارم که با مشاوره های علمی و رهنمودهای ارزنده خود دریچه های تحقیقاتی جدیدی را در این رساله، به روی من گشودند.

در دوره فرصت مطالعاتی در دانشگاه ملبورن استرالیا از مشورت با آقای پروفسور Rajkumar Buyya رئیس Cloud Computing and Distributed Systems Laboratory بهره مند شدم، از ایشان و همکاران ایشان در آن دانشگاه، به خاطر مشاوره های علمی و امکاناتی که برای تحقیق در اختیار بنده قرار دادند سپاسگزارم. همچنین از استاد گرامی جناب آقای دکتر بهمن جوادی از دانشگاه سیدنی غربی، کمال تشکر و قدردانی را دارم که تجربیات واطلاعات ارزشمند خود را بی دریغ در اختیار بنده قرار دادند و مرهون دقت نظر و زحمات ایشانم.

چکیده

سیستم های مشبک داوطلبانه از سیکل های محاسباتی آزاد منابع داوطلب موجـود در اینترنـت بـرای اجرای پروژه های پیچیده علمی استفاده می کنند. بدلیل ناهمگونی منابع در این سیستم ها، ارائه یک معماری شناسایی منابع کارا یکی از فاکتورهای کلیدی در زمینه کارایی آنها محسوب می شود. در این رساله، معماری شناسایی منابع آگـاه از مجـاورت CycloidGrid در سیـستم هـای مـشبک داوطلبانـه پیشنهاد شده است. جستجوی منابع در این معماری با توجه بـه پارامترهـای کیفیـت خـدمت کـاربر، موازنه بار و آگاهی از مجاورت انجام می شود. در این معماری زمانبندی برنامـه هـای مـوازی در قالـب کیسه وظایف و برنامه های جریان کاری بررسی شده است. در معماری پیشنهادی انتخاب منبع جهت اجرای یک برنامه کیسه وظایف طی دو مرحله انجام می شود. در اولین مرحله زیر مجموعه ای ازمنابع بر اساس پارامترهای کیفیت خدمت کاربر و موازنه باردر سیستم انتخاب می شـوند. سـپس در مرحلـه دوم از میان منابع انتخابی در اولین مرحله، منبع نهایی با اولویت بالاتر نسبت به حداقل سربارارتباطی انتخاب می شود. برای محاسبه سربار ارتباطات میان دو ماشین نظیر، یک مدل شـبکه ای مبتنـی بـر تئوری صف پیشنهاد شده است که در آن ترافیک پس زمینه اینترنت در محاسبه سـربار ارتباطـات در نظر گرفته می شود. دو روش جهت پیاده سازی موازنه باردراین معماری پیشنهاد شده است. روش اول یک روش مبتنی بر دانش است که از اطلاعات ماشینهای نظیر مانند طـول صـف بـرای پیـاده سـازی موازنه بار استفاده می کند. درحالیکه روش دوم یک روش تحلیلی مبتنی بر تئوری صف می باشد کـه مبتنی بر مسیریابی درصف های موازی توزیع شده است. زمانبندی برنامه های جریـان کـاری در ایـن معماری با تکیه بر کاهش سربار ارتباطات انجام می شود. در این زمانبنـدی ابتـدا یـک برنامـه جریـان کاری به تعدادی برنامه زیر جریان کاری تقسیم می شود بطوریکه میزان وابـستگی داده میـان برنامـه های زیر جریان کاری حداقل شود، سپس مجموعه ای از منابع نزدیک بهم بـا حـداقل بـار محاسـباتی جهت اجرای نهایی برنامه های زیر جریان کاری انتخاب می شوند.

فهرست مطالب

فهرست مطالب..... I

فهرست جداول..... IV

فهرست شکل ها..... V

۱ معرفي ۱

۱-۱ سیستم های مشبک رومیزی..... ۲

۱-۲ دسته بندی سیستم های مشبک رومیزی..... ۴

۱-۳ برنامه های کاربردی متداول در سیستم های مشبک رومیزی..... ۶

۱-۴ پرسش های تحقیق..... ۸

۱-۵ اهداف رساله..... ۹

۱-۶ روش انجام رساله..... ۱۲

۱-۷ مشارکت علمی..... ۱۴

۱-۶ ساختار رساله..... ۱۵

۲ مروري بر پيشينه تحقيق ۱۷

۲-۱ مروری بر سیستم های مشبک رومیزی متمرکز..... ۱۷

۲-۲ مروری بر سیستم های مشبک رومیزی نظیر به نظیر..... ۲۱

۲-۲-۱ الگوریتم های جستجوی منابع در اجرای برنامه های کیسه وظایف..... ۲۲

۲-۲-۲ الگوریتم های جستجوی منابع در اجرای برنامه های جریان کاری..... ۴۶

۲-۳ نتیجه گیری..... ۵۲

۳ معرفي معماري شناسايي منابع آگاه از مجاورت CycloidGrid ۵۳

۳-۱ منابع و مدل برنامه کاربردی.....۵۳

۳-۲ معماری CycloidGrid.....۵۴

۳-۲-۱ مدیریت ورود و خروج گره ها در CycloidGrid ۵۸

۳-۲-۲ فرایند انتخابات.....۶۱

۳-۳ مدلی مبتنی بر تئوری صف برای محاسبه سربارارتباطات.....۶۱

۳-۴ زمانبندی برنامه های کیسه وظایف درCycloidGrid ۶۶

۳-۵ مدل تحلیلی مبتنی بر تئوری صف برای پیاده سازی موازنه باردر CycloidGrid.....۷۱

۳-۵-۱ مدل تحلیلی مبتنی بر تئوری صف.....۷۱

۳-۵-۲ پیاده سازی موازنه بار مبتنی بر مدل تحلیلی پیشنهادی.....۷۹

۳-۵-۲-۱ تعیین نرخ ارسال بهینه ورودی به هریک از خوشه های منطقی....۷۹

۳-۵-۲-۲ الگوریتم های توزیع درخواست ورودی میان خوشه های منطقی...۸۰

۳-۵-۳ زمانبندی برنامه های کیسه وظایف بر اساس مدل تحلیلی پیشنهادی.....۸۵

۳-۶ زمانبندی آگاه از مجاورت برنامه های جریان کاری در معماری CycloidGrid ۸۷

۳-۶-۱ سیستم زمانبندی آگاه از مجاورت برنامه های جریان کاری.....۸۹

۳-۷ نتیجه گیری.....۹۷

۴ ارزيابي معماري شناسايي منابع آگاه از مجاورت CycloidGrid ۹۸

۴-۱ شبیه ساز CycloidGrid ۹۸

۴-۲ ارزیابی کارایی CycloidGrid در اجرای برنامه های کیسه وظایف ۱۰۳

۴-۲-۱ مدل تولید بار محاسباتی.....۱۰۳

۴-۲-۲ پارامترهای ارزیابی.....۱۰۶

۴-۲-۳ الگوریتم های پایه جهت مقایسه با معماری شناسایی منابع پیشنهادی....۱۰۷

۴-۲-۴ نتایج حاصل از آزمایش ها.....۱۰۸

۴-۳ ارزیابی کارایی مدل تحلیلی مبتنی بر تئوری صف برای پیاده سازی موازنه بار.....۱۱۸

۴-۳-۱ مدل تولید بار محاسباتی.....۱۱۸

۴-۳-۲ نتایج حاصل از آزمایش ها.....۱۲۱

۴-۴ ارزیابی کارایی سیستم زمانبندی جریان کاری پیشنهادی.....۱۳۱

۴-۴-۱ برنامه های جریان کاری محک.....۱۳۲

۴-۴-۲ مدل تولید بارمحاسباتی.....۱۴۰

۴-۴-۳ الگوریتم های پایه جهت مقایسه با سیستن زمانبندی جریان کاری.....۱۴۲

۴-۴-۴ نتایج حاصل از آزمایش ها.....۱۴۳

۴-۵ نتیجه گیری.....۱۵۲

۵ خلاصه ، نتیجه گیري و راهکارهاي آینده ۱۵۴

۵-۱ نتیجه گیری.....۱۵۵

۵-۲ راهکارهای آینده.....۱۵۹

واژه نامه انگلیسي به فارسي ۱۶۲

مراجع ۱۶۸

پیوست ها ۱۷۹

پیوست الف سیستم های نظیر به نظیر ۱۷۹

پیوست ب محاسبه مدل کارایی برای برنامه های کیسه وظایف در معماریCycloidGrid ۱۸۹

فهرست جداول

۱-۲ مقایسه سیستم های متفاوت جستجوی منابع..... ۴۰

۱-۳ معرفی نمادهای استفاده شده در مدل موازنه بار مبتنی بر تئوری صف..... ۷۴

۲-۳ مقایسه سه الگوریتم توزیع بار محاسباتی پیشنهادی..... ۸۵

۱-۴ پارامترهای ورودی مدل تولید بار محاسباتی..... ۱۰۴

۲-۴ پارامترهای ورودی مدل تولید بار محاسباتی..... ۱۱۸

۳-۴ پروفایل زمان اجرای Montage [Juv12] ۱۳۴

۴-۴ پروفایل زمان اجرای CyberShake [Juv12]..... ۱۳۶

۵-۴ پروفایل زمان اجرایEpigenomics [Juv12]..... ۱۳۸

۶-۴ پروفایل زمان اجرایSipht [Juv12] ۱۴۰

۷-۴ پارامترهای ورودی مدل تولید بار محاسباتی..... ۱۴۱

فهرست شکل ها

۱-۲ معماری سرویس دهنده BOINC [And09].....۱۸

۲-۲ معماری سیستم Entropia [Chi03] ۱۹

۳-۲ معماری سیستم OurGrid [Ang10].....۳۵

۴-۲ معماری سیستم ShareGrid [Ang10].....۳۶

۱-۳ ساختار یک CycloidGrid یازده بعدی.....۵۵

۲-۳ درخت تصمیم گیری CycloidGrid جهت دسته بندی منابع.....۵۷

۳-۳ درصد تعداد گره های گزارش دهنده فعال به تعداد منابع سیستم.....۵۹

۴-۳ مدل شبکه ای مبتنی بر تئوری صف برای محاسبه تاخیر ارتباطی میان دو گره نظیر.....۶۳

۵-۳ زمانبندی برنامه های کیسه وظایف درCycloidGrid.....۶۸

۶-۳ مدل مبتنی بر تئوری صف جهت پیاده سازی موازنه بار درCycloidGrid.....۷۳

۷-۳ سیستم زمانبندی جریان کاری آگاه از مجاورت CycloidGrid.....۹۱

۸-۳ مثالی از الگوریتم بخش بندی FM برای یک گراف نمونه (الف) بخش بندی اولیه (ب) خروجی الگوریتم FM.....۹۲

۹-۳ معماری جستجوی منابع در زمانبندی یک برنامه جریان کاری.....۹۵

۱-۴ چیدمان تصادفی (شکل سمت چپ) و چیدمان توزیع دم پهن (شکل سمت راست).....۱۰۱

۲-۴ پراکندگی اندازه برنامه های کیسه وظایف تولید شده در طول یکی از آزمایشات.....۱۰۵

۳-۴ متوسط زمان پاسخ به ازای الگوریتمهای متفاوت با ۱۰۰۰ ماشین نظیر و ترافیک سنگین پس زمینه اینترنت. (الف) نتایج حاصل با تغییر پارامتر α و تغییر فاصله زمانی بین دو درخواست

متوالی (ب) نتایج حاصل با تغییر متوسط زمان اجرای یک وظیفه درBoT.....۱۰۹

۴-۴ متوسط زمان پاسخ به ازای الگوریتمهای متفاوت با ۱۰۰۰ ماشین نظیر و ترافیک متوسط پس زمینه اینترنت. (الف) نتایج حاصل با تغییر پارامتر α و تغییر فاصله زمانی بین دو درخواست متوالی (ب) نتایج حاصل با تغییر متوسط زمان اجرای یک وظیفه درBoT.....۱۱۰

۵-۴ متوسط زمان پاسخ به ازای الگوریتمهای متفاوت با ۳۰۰۰ ماشین نظیر و ترافیک سنگین پس زمینه اینترنت. (الف) نتایج حاصل با تغییر پارامتر α و تغییر فاصله زمانی بین دو درخواست متوالی (ب) نتایج حاصل با تغییر متوسط زمان اجرای یک وظیفه درBoT.....۱۱۲

۶-۴ متوسط زمان پاسخ به ازای الگوریتمهای متفاوت با ۳۰۰۰ ماشین نظیر و ترافیک متوسط پس زمینه اینترنت. (الف) نتایج حاصل با تغییر پارامتر α و تغییر فاصله زمانی بین دو درخواست متوالی (ب) نتایج حاصل با تغییر متوسط زمان اجرای یک وظیفه درBoT.....۱۱۳

۷-۴ متوسط زمان پاسخ به ازای الگوریتمهای متفاوت با ۳۰۰۰ ماشین نظیر و ترافیک متوسط پس زمینه اینترنت و سیستم پویا. (الف) نتایج حاصل با تغییر پارامتر α و تغییر فاصله زمانی بین دو درخواست متوالی (ب) نتایج حاصل با تغییر متوسط زمان اجرای یک وظیفه درBoT.....۱۱۴

۸-۴ متوسط زمان پاسخ به ازای الگوریتمهای متفاوت با ۳۰۰۰ ماشین نظیر و ترافیک متوسط پس زمینه اینترنت به ازای تغییرمتوسط فاصله زمانی ورود-خروج ماشین های نظیراز سیستم ۱۱۶

۹-۴ متوسط پیامهای مبادله شده در سیستم به ازای یک برنامه کیسه وظایف.....۱۱۷

۱۰-۴ متوسط زمان پاسخ به ازای سیاست های توزیع متفاوت با ۱۰۰۰ ماشین نظیر در محیط ایستا. (الف وپ) نتایج حاصل با تغییر پارامتر α با تغییر فاصله زمانی بین دو درخواست متوالی (ب وت) نتایج حاصل با تغییر متوسط زمان اجرای یک وظیفه درBoT.....۱۲۳و۱۲۲

۴-۱۱ متوسط زمان پاسخ به ازای سیاست های توزیع متفاوت با ۱۰۰۰ ماشین نظیر در محیط پویا. (الف) نتایج حاصل با تغییر پارامتر α با تغییر فاصله زمانی بین دو درخواست متوالی (ب) نتایج حاصل با تغییر متوسط زمان اجرای یک وظیفه درBoT..... ۱۲۴

۴-۱۲ متوسط زمان پاسخ به ازای سیاست های توزیع متفاوت با ۳۰۰۰ ماشین نظیر در محیط ایستا. (الف) نتایج حاصل با تغییر پارامتر α با تغییر فاصله زمانی بین دو درخواست متوالی (ب) نتایج حاصل با تغییر متوسط زمان اجرای یک وظیفه درBoT..... ۱۲۶

۴-۱۳ متوسط زمان پاسخ به ازای سیاست های توزیع متفاوت با ۳۰۰۰ ماشین نظیر در محیط پویا. (الف) نتایج حاصل با تغییر پارامتر α با تغییر فاصله زمانی بین دو درخواست متوالی (ب) نتایج حاصل با تغییر متوسط زمان اجرای یک وظیفه درBoT..... ۱۲۷

۴-۱۴ متوسط زمان پاسخ به ازای سیاست های توزیع متفاوت با ۱۰۰۰ ماشین نظیر به ازای تغییرمتوسط فاصله زمانی ورود-خروج ماشین های نظیراز سیستم..... ۱۲۸

۴-۱۵ درصد درخواست هایی که ضرب الاجل آنها رعایت می شود به ازای افزایش درصد درخواست های با فوریت بالا (الف) و افزایش نرخ ورود درخواست ها با تغییر پارامتر(ب).... ۱۳۰

۴-۱۶ برنامه جریان کاری Montage [Bha08] ۱۳۴

۴-۱۷ برنامه جریان کاری CyberShake [Bha08]..... ۱۳۶

۴-۱۸ برنامه جریان کاریEpigenomics [Bha08]..... ۱۳۸

۴-۱۹ برنامه جریان کاری Sipht [Bha08]..... ۱۴۰

۴-۲۰ متوسط زمان پاسخ به ازای سیاست های متفاوت برای برنامه جریان کاری Montage با تغییر پارامتر αدر تغییر فاصله زمانی بین دو درخواست متوالی (الف) ۱۰۰۰ ماشین نظیر (ب)۳۰۰۰ ماشین نظیر..... ۱۴۵

۴-۲۱ متوسط زمان پاسخ به ازای سیاست های متفاوت برای برنامه جریان کاری CyberShake با تغییر پارامتر αدر تغییر فاصله زمانی بین دو درخواست متوالی (الف) ۱۰۰۰ ماشین نظیر (ب)۳۰۰۰ ماشین نظیر..... ۱۴۶

۴-۲۲ متوسط زمان پاسخ به ازای سیاست های متفاوت برای برنامه جریان کاری Epigenomics با تغییر پارامتر αدر تغییر فاصله زمانی بین دو درخواست متوالی (الف) ۱۰۰۰ ماشین نظیر (ب)۳۰۰۰ ماشین نظیر..... ۱۴۸

۴-۲۳ متوسط زمان پاسخ به ازای سیاست های متفاوت برای برنامه جریان کاری Sipht با تغییر پارامتر αدر تغییر فاصله زمانی بین دو درخواست متوالی (الف) ۱۰۰۰ ماشین نظیر (ب)۳۰۰۰ ماشین نظیر..... ۱۴۹

۴-۲۴ متوسط زمان پاسخ به ازای سیاست های متفاوت و ۱۰۰۰ ماشین نظیر با تغییرمتوسط فاصله زمانی ورود-خروج ماشین های نظیر از سیستم (الف) Montage workflow (ب) CyberShake workflow ۱۵۱

۴-۲۵ متوسط زمان پاسخ به ازای سیاست های متفاوت و ۱۰۰۰ ماشین نظیر با تغییرمتوسط فاصله زمانی ورود-خروج ماشین های نظیر از سیستم (الف) Epigenomics workflow (ب) Sipht workflow ۱۵۲

فصل ۱

معرفي

سیستم های مشبک رومیزی[1] یا سیستم های محاسبات داوطلبانه[2] که در آنها ماشین های داوطلب منابع محاسباتی و ذخیره سازی آزاد خود را دراختیار سیستم قرار می دهند بعنوان یک سکوی اجرایی با کارایی بسیار بالا برای اجرای پروژه های علمی بزرگ مورد توجه است [Kon09]. سیستم های محاسبات داوطلبانه نظیر به نظیر[3] محیطی غیر متمرکز، خودسازمانده[4] و مقیاس پذیر[5] را برای اجرای برنامه های کاربردی بزرگ ارائه می دهند بطوریکه در آنها منابع بطور گسترده در مناطق مختلف جغرافیایی توزیع شده اند. دو گروه از برنامه های کاربردی در قالب برنامه های کاربردی موازی کیسه وظایف[6] و برنامه های جریان کاری[7] در این سیستم ها متداول می باشند. یکی از چالشهای اصلی در این محیط اشتراک منابع توزیع شده ، غیر متمرکز و بزرگ طراحی یک معماری شناسایی منابع با کارایی بالا می باشد. موازنه بار بعنوان یکی از فاکتورهای کارایی یک معماری شناسایی منابع مطرح است. برای پیاده سازی موازنه بار در سطح سیستم دو روش مبتنی بر دانش[8] منابع موجود در سیستم و روش تحلیلی عاری از دانش[9] وجود دارد. از دیگر پارامترهای قابل بررسی در یک معماری شناسایی منابع کارا می توان به پارامترهای کیفیت خدمت کاربر[10] و آگاهی از مجاورت[11] اشاره کرد.

---

[1] Desktop Grids
[2] Volunteer computing systems
[3] Peer-to-peer based volunteer computing systems
[4] Self-organized
[5] Scalable
[6] Bag-of-Tasks(BoT)
[7] Workflow
[8] Knowledge-based
[9] Knowledge-free
[10] Quality of Service(QoS)
[11] Proximty-aware

در ادامه این فصل ابتدا در زمینه سیستم های مشبک رومیزی و دسته بندی آنها بحث خواهد شد. درادامه برنامه های کاربردی متداول در این سیستم ها معرفی خواهند شد. سپس مسئله و پرسش های تحقیق ،اهداف تحقیق و در نهایت روش انجام تحقیق و ساختار رساله بیان خواهد شد.

۱ - ۱ سیستم های مشبك رومیزي

هدف از سیستم های مشبک رومیزی (سیستم های محاسبات داوطلبانه) آنست که از کامپیوترهای رومیزی بیکار در سطح اینترنت (منابع داوطلب) برای اجرای پروژه های علمی و برنامه های کاربردی با حجم بالای محاسبات استفاده نمایند [And04,Chi03, Fos03, Cap04,Cho07].

از جمله پروژه های معروف در این سیستم ها می توان به SETI@home [And02] ، Folding@home [Beb09] ، EDGeS [Fed08] ، DEGISCO [DEG] ، EDGI [EDG] و Climate@home [Gui03] اشاره کرد. این پروژه ها در زمینه های مختلف علمی نظیر ریاضیات، رمزنگاری، فیزیک، نجوم،مطالعه آب و هوا و پزشکی بوده و نیاز به قدرت محاسباتی بالا در طول یک مدت زمان نسبتا طولانی دارند. سیستم های مشبک رومیزی با سیستم های مشبک[1] تفاوتهایی دارند که از آن جمله می توان به موارد زیر اشاره کرد[Cho07]:

- منابع سیستم های مشبک رومیزی اغلب کامپیوترهای شخصی هستند (نظیر کامپیوترهای رومیزی) درحالیکه منابع سیستم های مشبک می توانند از سوپرکامپیوترها ، خوشه[2]، تجهیزات آزمایشگاهی، پایگاه های داده و ... نیزمی باشند.
- منابع سیستم های محاسبات داوطلبانه بر خلاف سیستم های مشبک بسیار فرار[3] ، غیر خصوصی[4] و ناهمگون هستند. همچنین آنها غیر قابل اعتمادتر و مستعدتر به خطا نسبت به سیستم های مشبک می باشند.

---

[1] Grids
[2] Cluster
[3] Volatile
[4] non-dedicated

- در سیستم های مشبک رومیزی منابع بطور خصوصی توسط افراد مدیریت می شوند در حالیکه در سیستم های مشبک توسط مدیران حرفه ای مدیریت می شوند.

در واقع هدف سیستم های مشبک رومیزی بدست آوردن بازده بالا[1] می باشد (تاکید بر روی مقدار کاری است که کامپیوترهای رومیزی می توانند در طی یک بازه زمانی انجام دهند) در حالیکه در سیستم های مشبک اساسا تاکید بر روی کارایی[2] بالا باشد (بالاترین سرعتی که در آن مجموعه ای از وظایف قابل اجرا هستند).

شناسایی منابع یکی از چالشهای مطرح در سیستم های مشبک خصوصا سیستم های مشبک رومیزی می باشد. هدف از معماری شناسایی منابع تخصیص وظیفه های یک برنامه کاربردی به مناسب ترین منابع در سیستم می باشد.

بحث زمانبندی در سیستم های مشبک بسیار متفاوت از سیستم های مشبک رومیزی می باشد. دلیل این امر آنست که سیستم های محاسبات داوطلبانه به لحاظ نوع منبع، قابلیت اطمینان، نوع برنامه کاربردی و نظائر آن با آنها متفاوتند [And04,Chi03, Fos03, Cap04,Cho07, Kon04]. بطور کلی در این سیستم ها اولا بیشتر تمرکز بر روی پروسه تخصیص وظایف به مناسب ترین منابع در سیستم است [And05,Cap04, Chi03,Cho06,Kon04] . جستجوی منبع ممکن است بصورت متمرکز یا بصورت کاملا توزیع شده انجام شود. ثانیا برخلاف سیستم های مشبک در این سیستم ها نیاز به زمانبند محلی نمی باشد زیرا واحد زمانبندی بر خلاف سیستم های مشبک که یک سایت می باشد یک کامپیوتر رومیزی است. در نهایت زمانبندی در سیستم های مشبک رومیزی از نوع فرصت طلبانه[3] است. در سیستم های محاسبات داوطلبانه انتظار می رود که بحث مختار بودن منابع داوطلب حفظ شود (به این معنا که منابع داوطلب می توانند آزادانه در محاسبات شرکت کنند یا سیستم را ترک نمایند) بنابراین بحث زمانبندی در این سیستم ها باید بتواند در حد امکان و به سرعت از منابع

---

[1] High throughput
[2] High performance
[3] Opportunistic

داوطلب در طول مدت زمانی که در دسترس هستند و یا بیکار می باشند حداکثر استفاده را نماید[Cho06, And04].

۲-۱ دسته بندي سيستم هاي مشبك روميزي

بطور کلی سیستم های مشبک رومیزی را می توان بر حسب روش سازماندهی منابع به دو گروه سیستم های متمرکز[۱] و نظیر به نظیر[۲] تقسیم نمود [Cho07].

یک سیستم مشبک رومیزی متمرکز شامل مشتری ، سرویس دهنده و تعدادی منابع داوطلب[۳] است. سرویس دهنده شامل اطلاعات کلیه منابع داوطلب موجود در سیستم بوده و وظیفه جستجوی منبع، زمانبندی[۴] کارهای محاسباتی و تخصیص آنها به منابع داوطلب را بر عهده دارد. از جمله سیستم های مشبک رومیزی متمرکز می توان به BOINC [And04, And06] ، [Chi03] Entropiaو XtremeWeb [Cap04] اشاره کرد.

سیستم های مشبک رومیزی نظیر به نظیر شامل مشتری و منابع داوطلب هستند. برخلاف سیستم های متمرکز ماشین سرویس دهنده ای وجود ندارد و هریک از منابع داوطلب می توانند نقش سرویس دهنده را ایفا کنند و شامل بخشی از اطلاعات سایر منابع داوطلب موجود درسیستم باشند. علاوه بر این وظیفه هر یک می توانند بعنوان زمانبند وظیفه زمانبندی کارهای محاسباتی را نیز بر عهده گیرند. در این سیستم ها ماشین ها خودشان را بر اساس یک شبکه پوششی محاسباتی[۵] (CON) که یک شبکه منطقی بین منابع است سازماندهی می کنند [Ste05]. ساختن CON به نوعی گروهبندی منابع است و بوسیله ماشین های داوطلب به شیوه ای توزیع شده ایجاد می شود. در این

---

[1] Centralized
[2] Peer-to-peer
[3] Volunteer
[4] Scheduling
[5] Computational overlay network (CON)