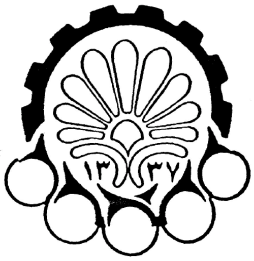

دانشگاه صنعتی امیرکبیر

دانشکدهٔ مهندسی کامپیوتر و فناوری اطلاعات

پایان‌نامه کارشناسی ارشد مهندسی نرم افزار

اجرای جریان کاری از سرویس ها در محیط گرید با توجه به خواص کیفیت سرویس

نگارش

طه قاسمی

استاد راهنما

دکتر محمد رضا رزازی

بهمن ۱۳۸۶

بسمـه تعالـی

دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)
معاونت پژوهشی

فرم اطلاعات پایان نامه
کارشناسی ارشد و دکترا

تاریخ: ..۸۶/۱۱/۲۳.....

پیوست: .....

نام و نام خانوادگی:	طه قاسمی	**دانشجوی آزاد** ■	**بورسیه** □	**معادل** □
شماره دانشجوئی:	۸٤۱۳۱۰٤۳	**دانشکده:** مهندسی کامپیوتر	**رشته تحصیلی:**	مهندسی نرم افزار

نام و نام خانوادگی استاد راهنما: محمد رضا رزازی

عنوان پایان نامه به فارسی: اجرای جریان کاری از سرویس ها در محیط گرید با توجه به خواص کیفیت سرویس

عنوان پایان نامه به انگلیسی: QoS-based Execution of Service Workflows in Grid Environment

نوع پروژه: **کارشناسی ارشد** ■ / **دکترا** □ — **کاربردی** □ **بنیادی** □ **توسعه ای** ■ **نظری** □

تاریخ شروع: ۱۳۸۵/۷/۱ **تاریخ خاتمه:** ۱۳۸۶/۱۰/۲۳ **تعداد واحد:** ۶

سازمان تأمین کننده اعتبار:

واژه های کلیدی به فارسی: کوله پشتی چند بعدی با چند انتخاب، تخصیص جریان های کاری، گرید، بهینه سازی، کیفیت سرویس

واژه های کلیدی به انگلیسی: Multidimensional Multiple-Choice Knapsack, Workflow Assignment, Grid, Optimization, Quality of Service

نظرها و پیشنهادها به منظور بهبود فعالیت های پژوهشی دانشگاه:

استاد راهنما:

دانشجو: توجه بیشتر به کیفیت مقالات تولیدی به جای کمیت آن ها. به عنوان مثال در نظر گرفتن نمره برای تعداد رجوع به مقالات ارائه شده می تواند عامل تشویقی در این زمینه باشد.

امضاء استاد راهنما: **تاریخ: ۸۶/۱۱/۲۳**

نسخه ۱: معاونت پژوهشی

نسخه ۲: کتابخانه و به انضمام دو جلد پایان نامه به منظور تسویه حساب با کتابخانه و مرکز اسناد و مدارک علمی

تقدیم به پدر و مادرم که صمیمانه ترین ایثارها را نثار من کردند..

و تقدیم به خواهرم

تشکر و قدردانی

با سپاس فراوان از دکتر محمد رضا رزازی که راهنمایی این پروژه را بر عهده داشتند. همچنین از آقایان دکتر محمد رضا میبدی و دکتر کوروش عشقی که داوری این پروژه را بر عهده داشتند کمال تشکر را دارم.

چکیده

در محیط گرید سرویس های مختلفی به اشتراک گذاشته شده اند. این سرویس ها باید با یکدیگر همکاری کرده تا عملکرد مورد نیاز کاربر داده شود. این نحوه همکاری سرویس ها توسط جریان کاری از سرویس های مجرد بیان می شود که به صورت اتوماتیک و یا توسط کاربر ایجاد شده است. سرویس مجرد سرویسی است که عملکرد مشخصی را دارا می باشد اما لزوما پیاده سازی آن مشخص نشده است. در محیط گرید پیاده سازی های مختلفی از یک سرویس مجرد وجود دارد که توسط عرضه کننده های سرویس عرضه می شوند. پیاده سازی های یک سرویس مجرد دارای عملکرد یکسان بوده اما خواص کیفیت سرویس متفاوتی را دارا می باشند (همچون قیمت، زمان اجرا، و یا میزان در دسترس بودن).

مساله ای که در اینجا تعریف می شود ایجاد مکانیزمی است که با گرفتن یک جریان کاری از سرویس های مجرد آن را در محیط گرید به اجرا در آورد. برای این منظور باید جریان کاری سرویس های مجرد به جریان کاری از سرویس های واقعی نگاشت یابد. در این نگاشت کاربر اهداف و محدودیت های خود را، در مورد کیفیت نهایی جریان کاری، داراست که باید رعایت شوند. در نتیجه این مساله به یک مساله بهینه سازی تبدیل می شود. هدف از این پروژه مدل سازی این مساله و سپس حل آن می باشد. برای این منظور از یک تبدیل تقریبی برای کاهش مساله به مساله کوله پشتی چند بعدی با چند انتخاب استفاده شده است. به منظور حل مساله کوله پشتی چند بعدی با چند انتخاب، مفهوم هسته برای آن توسعه یافته است. هسته زیر فضایی از فضای راه حل است که در آن راه حل بهینه قرار گرفته است. بر مبنای هسته، الگوریتم هایی برای حل دقیق و تقریبی مساله ارائه شده اند. در نهایت معماری سیستم برای به کار گیری این الگوریتم در محیط گرید بر مبنای میان افزار گلوباس مورد بحث قرار گرفته است.

کلمات کلیدی

کوله پشتی چند بعدی با چند انتخاب (Multidimensional Multiple Choice Knapsack)، تخصیص جریان های کاری (Workflow assignment)، گرید (Grid)، بهینه سازی (Optimization)، کیفیت سرویس (Quality of Service).

فهرست مطالب

۱ مقدمه ۱

۱-۱ انگیزه ۱

۱-۲ تعریف پروژه ۱

۱-۳ کارهای مرتبط ۴

۱-۴ ساختار پایان نامه ۶

۲ گرید، سرویس ها، و مدیریت جریان کاری ۷

۲-۱ تعاریف اولیه ۷

۲-۲ گرید ۸

۲-۲-۱ مفهوم گرید ۸

۲-۲-۲ تعریف گرید ۱۳

۲-۲-۳ نسل ها و دسته بندی گریدها ۱۶

۲-۳ جریان های کاری ۱۸

۲-۳-۱ مفاهیم و تعاریف ۱۸

۲-۳-۲ اجرا و مدیریت جریان ها کاری ۲۰

۲-۴ مساله نگاشت ۲۲

۲-۴-۱ نگاشت وظائف در گرید ۲۲

۲-۴-۲ مساله تخصیص جریان کاری از سرویس ها در محیط گرید ۲۴

۳ مساله کوله پشتی و انواع آن ۳۰

۳-۱ مساله کوله پشتی ۳۰

۳-۲ انواع مساله کوله پشتی ۳۳

۳-۲-۱ کوله پشتی با چند انتخاب ۳۳

۳-۲-۲ کوله پشتی چند بعدی ۳۷

۳-۲-۳ کوله پشتی چند بعدی با چند انتخاب ۳۷

۳-۳ هسته ۳۹

۴ حل مساله تخصیص ۴۵

۴-۱ مدل سازی مساله ۴۵

۴-۱-۱ مدل محیط اجرایی ۴۵

۴-۱-۲ مدل جریان کاری ۴۵

۴-۱-۳ مدل درخواست کاربر ۴۶

۴-۱-۴ مدل مساله ۴۷
۴-۱-۵ تبدیل مساله به مساله کوله پشتی (چند بعدی) با چند انتخاب ۴۷
۴-۲ نگاهی دوباره به هسته ۴۸
۴-۳ کوله پشتی چند بعدی با چند انتخاب ۵۰
۴-۳-۱ توسعه هسته برای MMKP ۵۰
۴-۳-۲ حل تقریبی MMKP ۵۷
۴-۳-۳ حل دقیق MMKP ۵۸
۴-۴ کوله پشتی با چند انتخاب ۶۴
۴-۴-۱ هسته ۶۴
۴-۴-۲ الگوریتم شاخه و قید ۶۶
۴-۵ افزودن کیفیت سرویس رابط ها ۶۸
۵ نتایج ۷۱
۵-۱ کوله پشتی چند بعدی با چند انتخاب ۷۱
۵-۱-۱ حل دقیق ۷۲
۵-۱-۲ حل تقریبی ۸۳
۵-۲ کوله پشتی با چند انتخاب ۸۷
۶ اجرای جریان کاری در گرید ۸۹
۶-۱ جمع آوری و به روز رسانی اطلاعات ۸۹
۶-۲ توزیع بار اجرا ۹۱
۶-۳ مدیریت قرار دادهای سطح سرویس ۹۳
۶-۴ سلسله مراتب تخصیص-زمان بندی ۹۴
۶-۵ کاربر محور در مقابل سیستم محور ۹۵
۶-۶ درخواست کاربران ۹۶
۶-۷ نظارت زمان اجرا و نگاشت مجدد ۹۷
۷ جمع بندی ۱۰۰
۷-۱ خلاصه ۱۰۰
۷-۲ مشکلات و کارهای آینده ۱۰۱
مراجع ۱۰۳
واژه نامه انگلیسی به فارسی ۱۱۰
واژه نامه فارسی به انگلیسی ۱۱۱
ضمیمه ۱۱۲

میان افزار گلوباس ۱۱۲
معماری گلوباس ۱۱۳
بخش ها ۱۱۶
خلاصه ۱۳۲

۱ مقدمه

۱-۱ انگیزه

ایده گرید به منظور استفاده بهینه از منابع توزیع یافته و امکان به اشتراک گذاری آن ها به صورت کنترل شده مطرح شد. با تکامل این ایده بحث استفاده از سرویس ها در محیط گرید مطرح شد. در این مدل سرویس ها توسط فراهم کنندگان سرویس به اشتراک گذاشته می شوند و کاربران در هر جایی از گرید می توانند از این سرویس ها استفاده کنند. با افزایش تعداد این سرویس ها می توان انتظار داشت که بر روی گرید تعداد زیادی از سرویس هایی که از لحاظ عملکرد یکسانند وجود داشته باشند که کاربر مجبور به انتخاب از بین آن ها باشد. این انتخاب بر اساس مولفه های کیفیت سرویس (جنبه های غیر عملکردی سرویس) همچون زمان خدمت و قیمت آن می باشد.

با تخصصی تر شدن سرویس ها اینکه یک سرویس بتواند تمامی نیازهای یک کاربر را بر آورده کند کمرنگ می شود. در نتیجه کاربر باید مجموعه از سرویس ها را به کار گیرد تا در مجموع عملکرد مورد نیازش به او داده شود. برای این منظور کاربر باید سرویس های متناسب با نیاز و با کیفیت مد نظر خود را پیدا کرده و با آن ها به تعامل بپردازد. این عمل باعث پیچیده تر شدن استفاده از گرید می شود که با اهداف آن در تناقض است. از طرفی با مطرح شدن برنامه های کامل تر و پیچیده تر بر روی گرید از سوی سازمان ها و پروژه های علمی نیازمند اجرای کنترل شده این فرایند می باشیم. این بحث در حوزه مهندسی نرم افزار پیش از این در مورد مدیریت جریان های کاری مطرح بوده است. جریان کاری شامل ترتیبی از اجرای کارهاست که به صورت اتوماتیک انجام می شود. نرم افزارهای مدیریت جریان کاری سعی می کنند که یک یا چندین جریان کاری را به صورت سیستماتیک اجرا و کنترل کنند. ایده این نرم افزارها برای محیط سرویس گرید نیز قابل استفاده بود و لذا چندین پروژه در این راستا تعریف شدند. در این صورت کاربر می تواند جریان سرویس های مورد نیاز خود را تعریف کرده و این نرم افزارها اقدام به مدیریت جریان کاری در محیط گرید بنمایند.

در این پروژه سعی می کنیم با تکامل سیستم های مدیریت جریان کاری در گرید همچون در نظر گرفت کیفیت سرویس و انتخاب بهینه سرویس ها گامی در جهت استفاده بهینه تر از گرید برداریم.

۱-۲ تعریف پروژه

همانطور که گفته شد در محیط سرویس گرید سرویس های متعددی وجود دارند. می توان این سرویس ها را به دسته هایی افراز کرد که در هر دسته سرویس هایی با عملکرد یکسان وجود دارد. به

مجموعه این سرویس ها، "سرویس های از نظر عملکردی یکسان" می گوئیم. این سرویس ها دارای یک واسط مشترک می باشند که تحت عنوان سرویس مجرد آن را می شناسیم. به عبارت دیگر هر دسته مجموعه سرویس هایی هستند که سرویس مجرد را پیاده سازی کرده و دارای عملکرد یکسان می باشند. هر سرویس در مجموعه سرویس های با عملکرد یکسان توسط یک فراهم کننده سرویس عرضه می شود در نتیجه دارای خواص کیفیت سرویس متفاوت می باشد. می توان این خواص کیفیت سرویس را با برداری برای هر سرویس نمایش داد. این بردار دارای ابعاد مختلفی همچون زمان اجرا، قیمت، شهرت فراهم کننده سرویس، و قابلیت دسترسی می باشد. باید توجه نمود که در برخی ابعاد بیشینه بودن مقدار مطلوب است همچون شهرت و در برخی دیگر کمینه بودن آن. به منظور بیان جریان کاری سرویس ها نیازمند مدلی می باشیم. این مدل بر حسب ساختارهای مورد نیاز در یک جریان کاری می تواند ساده یا پیچیده باشد. به عنوان مثال می توان از یک گراف جهت دار (DAG) برای این منظور استفاده نمود که در آن ترتیب کارها قابل بیان است اما مواردی همچون تصمیم گیری ها در جریان کاری قابل مدل سازی نیست. علاوه بر مدل DAG می توان از استانداردهای موجود در این زمینه نیز استفاده نمود که این کار با تبدیل مدل ها می تواند مورد پشتیبانی قرار گیرد.

در مدل جریان کاری، روابط بین سرویس های مجرد بیان می شود و نه سرویس های واقعی. در نتیجه کاربر (یا سیستمی که جریان کاری را به صورت اتوماتیک ایجاد می کند) نیازمند این نیست که بداند در محیط واقعی چه سرویس هایی، سرویس مورد نظرش را پیاده نموده اند. در محیط واقعی جریان کاری باید با سرویس های واقعی اجرا شود. با توجه به اینکه هر سرویس واقعی دارای خواص کیفیتی خود است در مجموع کل جریان کاری دارای خواص کیفیت سرویس خاص خود می باشد که می توان آن را نیز به صورت یک بردار تعریف نمود. به عنوان مثال هزینه کل جریان کاری و یا مدت زمان اجرای آن. این دیدگاه از این نظر نیز قابل تفسیر است که خود جریان کاری در واقع یک سرویس مرکب می باشد که متشکل از سرویس های ساده و یا حتی مرکب دیگر است. در نتیجه این سرویس نیز همانند سایر سرویس ها دارای خواص کیفیت سرویس مشخصی می باشد. در نتیجه نیازمند روشی هستیم که کیفیت سرویس کل جریان کاری را با توجه به کیفیت سرویس اجزاء آن بدست آورد.

از طرف دیگر هر کاربر نیازمند ویژگی های کیفیت سرویس متفاوتی می باشد. این نکته با توجه به اینکه معمولا کیفیت سرویس بیشتر در ازاء قیمت بالاتر است مهمتر می نماید. در نتیجه کاربران مختلف اهداف متفاوتی در زمینه کیفیت سرویس دارا می باشند. به عنوان مثال کاربری ممکن است بخواهد جریان کاری او هر چه سریعتر تمام شود اما میزان هزینه اش برای او اهمیتی ندارد. از طرف دیگر کاربر دیگر متمایل است که جریان کاری با کمترین هزینه انجام شود. علاوه بر این اهداف کاربر ممکن است دارای محدودیت هایی نیز باشد که باید بتواند آن ها را به طریقی بیان نماید. به عنوان مثال ممکن است کاربر بخواهد که جریان کاری با کمترین هزینه اما حداکثر در دو ساعت انجام شود. این محدودیت ها می تواند محدودیت روی بردار کیفیت سرویس کل جریان و یا روی بردار کیفیت سرویس تک تک اجزاء آن

باشد. در حالت کلی این مساله را می توان با یک مساله بهینه سازی چند هدفه و با چند محدودیت مدل نمود. در اینجا هر هدف معادل هدف کاربر در یکی از ابعاد کیفیت سرویس در نظر گرفته می شود و هر محدودیت محدودیت کاربر در زمینه یکی از ابعاد کیفیت سرویس است.

هدف از این پروژه توسعه سیستمی است که بتواند با گرفتن جریان کاری بین سرویس های مجرد و همچنین اهداف و محدودیت های کاربر در زمینه کیفیت سرویس جریان کاری نهایی اقدام به اجرای جریان کاری کند. برای این منظور نیازمند روشی هستیم که با در نظر گرفتن سرویس های واقعی با عملکرد یکسان با سرویس مجرد ذکر شده در جریان کاری، به گونه ای مجموعه ای از این سرویس ها را انتخاب کند و در جریان کاری به کار برد که هدف کاربر تامین شده و محدودیت های آن ارضاء شود. به عبارت دیگر نگاشت مناسبی از سرویس های مجرد در جریان کاری به سرویس های واقعی را بیابد. البته با توجه به اینکه مساله در حالت کلی NP-Hard می باشد باید الگوریتم های کارا برای حل آن را توسعه داد.

به منظور یافتن نگاشت مناسب جریان کاری در محیط واقعی سیستم باید مسائل موجود در محیط گرید را در نظر بگیرد. به علت عدم تجانس و توزیع یافتگی در گرید برای این منظور نیازمند معماری متناسب با چنین محیطی می باشیم. در این پروژه نکات مربوط به معماری چنین سیستمی بر مبنای میان افزار گلوباس نیز عرضه خواهد شد.

به طور خلاصه پروژه دارای این مشخصات است:

هدف:

- توسعه سیستمی برای اجرای جریان های کاری از سرویس های مجرد در محیط گرید. این عمل با در نظر گرفتن خواص کیفیت سرویس، سرویس ها انجام خواهد شد.

فرض ها:

- جریان کاری سرویس های مجرد توسط کاربر بیان شده است.
- سرویس های دارای عملکرد یکسان با یک سرویس مجرد به همراه ویژگی های کیفیت سرویس آن ها مشخصند.
- معماری بر مبنای میان افزار گلوباس خواهد بود.

خروجی ها:

- الگوریتم هایی برای انتخاب بهینه و نیمه بهینه سرویس های واقعی برای یک جریان کاری از سرویس های مجرد.
- معماری یک سیستم برای نگاشت جریان کاری در گرید.

۱-۳ کارهای مرتبط

بحث زمان بندی جریان های کاری پیش از این در حوزه سیستم های موازی مطرح بوده است. [۱] بررسی جامعی در این زمینه داشته است. با این وجود در محیط گرید مساله دارای ابعاد بیشتری می باشد و محیط نامتجانس است. در محیط گرید، همانند سایر محیط ها، بسیاری از زمان بندهای موجود جریان های کاری، عمل زمان بندی را بدون در نظر گرفتن محدودیت ها انجام می دهند [۲-۸]. در [۹] روشی برای زمان بندی جریان های کاری با در نظر گرفتن محدودیت هزینه با استفاده از الگوریتم های ژنتیکی عرضه شده است. در [۱۰] نیز یک روش ابتکاری برای زمانبدی جریان های کاری با فرض وجود محدودیت در هزینه ها ارائه شده است. [۱۱، ۱۲] سعی نموده اند مساله نگاشت جریان های کاری را در حالتیکه هدف کمینه سازی هزینه با محدودیت زمان است حل کنند. این کار با توزیع زمان بر روی جریان کاری انجام شده است. برای این منظور از ترکیبی از روش های BFS و DFS و آنالیز مسیر بحرانی استفاده شده است سپس بر روی هر شاخه با توجه به اینکه محدودیت زمان آن مشخص شده است با استفاده از Markov Decision Process اقدام به انتخاب منابع به طوریکه هزینه را کمینه کنند کرده است. در [۱۳] دو روش برای زمان بندی جریان کاری یکبار با هدف بهینه کردن زمان با فرض محدودیت بر روی هزینه و بار دیگر با هدف بهینه کردن هزینه با فرض محدودیت بر روی زمان ارائه شده است. به منظور حل مساله از الگوریتم های ژنتیک استفاده شده است. ما در این پروژه دید را فراتر برده و اولا علاوه بر زمان و هزینه سایر ابعاد کیفیت سرویس را نیز در نظر می گیریم ثانیا کاربر آزاد است که نیاز خود را به صورت بهینه سازی در یک یا چند بعد با فرض موجود بودن محدودیت در یک یا چند بعد دیگر بیان کند. یکی دیگر از تفاوت های این پروژه با کارهای انجام شده آن است که به واسطه سرویس گرا بودن ما بر روی مساله تخصیص به جای مساله زمان بندی تاکید می کنیم.

در خارج از محیط گرید کارهای بیشتری در زمینه بحث کیفیت سرویس و جریان های کاری انجام شده است. در [۱۴] سعی شده است که روشی برای محاسبه کیفیت سرویس جریان کاری بر مبنای کیفیت سرویس مولفه های آن ارائه شود. در این روش جریان کاری شامل ساختار شرطی به مجموعه ای از شاخه های قطعی تبدیل می شود. همچنین ساختار حلقه به ساختار معمولی کاهش می یابد. سپس در هر کدام از این جریان های کاری ساده شده برای هر کدام از ابعاد، کیفیت سرویس همچون زمان، قیمت و قابلیت دسترسی محاسبه می شود. بعد زمان و قیمت خاصیت جمعی داشته اما قابلیت دسترسی ضربی است. البته در برخی از روش ها با گرفتن log فرض می شود آن نیز جمعی است. برای محاسبه بعد زمان از روش مسیر بحرانی استفاده شده است. Cardoso نیز در تز دکتری خود روشی را برای محاسبه کیفیت سرویس جریان کاری از روی کیفیت سرویس مولفه های آن ارائه داده است. در این روش با توجه به احتمال شرط ها و یا احتمال تعداد پیمایش حلقه اقدام به محاسبه کیفیت سرویس در ابعاد مختلف به صورت احتمالی می کند. در روش ارائه شده برای محاسبه هر بعد از کیفیت سرویس، فرمول خاصی داده

شده است. در [۱۵] روشی برای محاسبه کیفیت سرویس یک مولفه با توجه به کاربر ارائه شده است. در این روش کیفیت سرویس برای هر کاربر متفاوت بوده و ابعاد مختلف کیفیت تبدیل به یک عدد می شود. برای این منظور از تکنیک وزن دهی استفاده شده است. در [۱۶، ۱۷] از الگوریتم های ژنتیک برای حل مساله تخصیص استفاده شده است. در روش های ارائه شده شیوه برخورد با محدودیت ها و همچنین یک هدفه کردن مساله مورد بحث قرار نگرفته اند. در این روش کیفیت سرویس می تواند دارای هر تعداد بعدی باشد و روشی بهبود یافته، بر مبنای روش Cardoso، برای تشخیص کیفیت سرویس جریان از روی کیفیت سرویس مولفه ها در حالتیکه جریان دارای حلقه می باشد ارائه شده است. روش باز کردن حلقه ارائه شده دارای هزینه سنگین بوده و در این زمینه توازنی بین بهینگی راه حل با زمان محاسبه مورد بحث قرار گرفته است. در [۱۸] مساله با دو روش ابتکاری حل شده است. [۱۴] مساله را در دو حالت بهینه سازی محلی و بهینه سازی سراسری با استفاده از برنامه ریزی اعداد صحیح حل نموده است. در حالت بهینه سازی محلی هر سرویس به طور مجزا به صورت بهینه انتخاب می شود و بر روی یک سرویس می توان محدودیت تعریف نمود. محدودیت ها در حالت سراسری به صورت خطی است. در حالتیکه جریان کاری شامل چندین مسیر اجراست مسیر کاری پرکاربردتر در گذشته به عنوان ملاک تصمیم قرار می گیرد. نویسندگان عقیده دارند که روش ارائه شده در محیط پویا زمان قابل قبولی ندارد و باید روش هایی برای ترکیب روش سراسری با روش های محلی ارائه داد. در روش ارائه شده همچنین بحث غیر مطمئن بودن مدت زمان اجرای یک سرویس با فرض توزیع نرمال برای آن مورد پشتیبانی قرار گرفته است. روش کاهش یک جریان کاری مرکب شامل حلقه و شرط برای جریان های کاری پیچیده موثر نیست.

[۷، ۸] بررسی نسبتا کاملی در مورد سیستم های مدیریت جریان کاری در گرید داشته اند. بر اساس این بررسی ها مشخص شده است که بحث پشتیبانی QoS و همچنین کنترل خطا از بخش هایی است که هنوز مورد توجه قرار نگرفته اند. [۴، ۱۹-۲۲] نمونه هایی از سیستم مدیریت جریان کاری در محیط گرید می باشند. در این سیستم ها از قابلیت های گرید همچون بستر توزیع یافته استفاده مناسب نشده است. [۲۳، ۲۴] به معرفی استاندارد مدیریت جریان کاری می پردازند. [۲۵] نیر استاندارد بیان نمودن جریان های کاری از سرویس ها در محیط وب را معرفی نموده است. در [۲۶] برنامه ریزی مجدد اجرای جریان کاری در محیط عملی بواسطه تغییر شرایط کیفی مطرح گشته است. در [۲۷] نیز سیستمی برای نظاره بر اجرای جریان های کاری به صورت توزیع یافته در گرید توسعه داده شده است. در [۲۸] سیستمی کلی برای جمع آوری اطلاعات و وقایع از سطح گرید مطرح گشته است.

۱-۴ ساختار پایان نامه

ساختار این پایان نامه بدین صورت است. در فصل ۲، مفاهیم پروژه توضیح داده شده اند. برای این منظور دو بحث پایه ای یعنی گرید و جریان های کاری بررسی شده اند. نقش سرویس ها نیز در این دو حوزه توضیح داده شده است. پس از این تعاریف به بررسی مسائل نگاشت بر روی گرید پرداخته و در نهایت مساله مطرح شده در این پایان نامه به صورت مفهومی بررسی می شود. این مساله نگاشت جریان کاری از سرویس ها در محیط گرید با توجه به خواص کیفیت سرویس می باشد. با توجه به آن که حل مساله بر مبنای تکنیک های توسعه داده شده در مسائل کوله پشتی است، فصل ۳ به معرفی این مسائل و راه حل های آن ها پرداخته است. توجه ویژه ای به مفهوم هسته برای حل این مسائل شده است. در فصل ۴ مساله به صورت فرمال مساله مدل شده و راه حلی پیشنهادی برای حل آن آن ارائه می شود. مدل مساله بر مبنای سه زیر مدل مدل محیط اجرایی، مدل جریان کاری، و مدل نیاز کاربر می باشد. پس از مدل سازی نشان داده می شود که مساله با استفاده از تبدیل مناسب به مساله کوله پشتی چند بعدی با چند انتخاب کاهش می یابد. لذا در این فصل الگوریتم هایی برای حل دقیق و تقریبی این مساله توسعه داده شده اند. هر دو الگوریتم مبتنی بر تکنیک هسته می باشند. در فصل ۵ نتایج ارزیابی عملی الگوریتم های ارائه شده برای حل مساله گزارش شده است. همچنین این نتایج با نتایج الگوریتم های موجود مورد مقایسه قرار گرفته است. در نهایت در فصل ۶ نحوه به کار بردن الگوریتم ارائه شده در محیط واقعی توضیح داده شده است. برای این منظور چالش ها مطرح شده و نیازمندی ها و راه حل های مربوطه در سطح معماری و بر مبنای میان افزار گوباس ارائه می شوند.

۲ گرید، سرویس ها، و مدیریت جریان کاری

در این بخش به معرفی زمینه پروژه می پردازیم. برای این منظور ابتدا به معرفی گرید و معرفی اصطلاحات موجود در این حوزه می پردازیم. پس از آن با حوزه مدیریت جریان کاری آشنا می شویم. در هر دو حوزه بحث سرویس گرا شدن پیش کشیده شده است. در آخر بخش نیز به مساله نگاشت جریان کاری از سرویس ها در محیط گرید خواهیم پرداخت که هدف اصلی از این پروژه می باشد.

۲-۱ تعاریف اولیه

در زیر مفاهیم پایه مورد استفاده در طول گزارش تعریف شده اند.

۱. منبع

هر موجودیتی که می تواند مورد استفاده و اشتراک قرار گیرد به عنوان منبع شناخته می شود. منابع را می توان به موارد کلی زیر تقسیم کرد:

- منابع محاسباتی[1] همانند پردازشگر مرکزی، مجموعه کامپیوتر های یک کلاستر، و یا یک سوپر کامپیوتر. در برخی تعاریف کل کامپیوتر به عنوان یک منبع محاسباتی محسوب می شود.
- منابع ذخیره سازی: مانند هارد دیسک ها و سایر سیستم های ذخیره سازی، و سرور های مخصوص ذخیره سازی داده ها.
- منابع شبکه ای: از قبیل خطوط ارتباطی و ظرفیت آن ها.

در برخی نگاه ها منبع را علاوه بر موارد بالا شامل نرم افزارها، داده ها، سرویس ها، و سخت افزارهای مختلف مانند پرینتر به شمار آورده می شود. پس به عبارت دیگر می توان منبع را نماینده یک موجودیت منطقی یا فیزیکی دانست. در واقع در این حالت دسته بندی دیگری از منابع را نیز می توان در نظر گرفت: منابع سخت افزاری (فیزیکی) و منابع نرم افزاری (منطقی). در بحث پیش رو بنابر تعریفی که انجام شد، منبع هر عنصری است که می تواند مورد استفاده قرار گیرد، و لذا شامل تمامی موارد بالا در نظر گرفته می شود.

۲. سرویس

موجودیتی بر روی شبکه است که خدمت مشخصی را عرضه می کند. سرویس را می توان مجموعه ای دانست از پروتکل ارتباطی با آن و رفتاری که از سرویس در اثر تعامل با آن از طریق این پروتکل سر می زند. لزومی ندارد که یک سرویس همواره در دسترس باشد، می تواند دارای حق دسترسی مخصوص

[1] Computational Resources

به خود باشد، می تواند نیاز به پرداخت هزینه برای استفاده از خدمات آن باشد، و حتی خود به صورت توزیع یافته در حال اجرا باشد. سرویس باید واسط ارتباطی با خود را به طور مشخص تعریف کرده باشد تا دیگران بتوانند از آن استفاده نمایند. نمونه هایی از سرویس ها می توان به سرویس انتقال فایل، سرویس محاسبه یک تابع، سرویس حسابداری، و .. اشاره کرد.

۳. کیفیت سرویس

مشخص کننده سطح کیفیت سرویس ارائه شده یا مورد درخواست می باشد. کیفیت سرویس می تواند شامل زمان اجرا، هزینه اجرا، قابلیت اطمینان، میزان دسترسی (یا در دسترس بودن)، بازدهی، تحمل پذیری خطا، امنیت، و شهرت عرضه کننده باشد. این موارد باید بتوانند به صورت کمی بیان شوند. میزان کیفیت سرویس ممکن است به صورت حداقل میزان قابل قبول، متوسط و یا تغییرات قابل تحمل نیز باشند که ما در این پروژه از آن ها استفاده نمی کنیم.

۴. استاندارد باز

استانداردی که در دسترس عموم قرار داده شده است. اشخاص مختلف می توانند به راحتی از آن استفاده کرده و یا در فرایند توسعه آن نقش داشته باشند.

۲-۲ گرید

۲-۲-۱ مفهوم گرید

برای اینکه مفاهیم بهتر در جای خود قرار گیرند خوب است به این پرسش پاسخ دهیم که چه ارتباطی بین گرید، اینترنت، و وب وجود دارد؟ آیا گرید جایگزینی برای اینترنت می باشد؟ آیا گرید بر روی وب اجرا می شود؟

بهتر است در این زمینه ابتدا مقایسه ای بین وب و گرید داشته باشیم. وب را می توان به عنوان یکی از سرویس های روی اینترنت شناخت. بر روی اینترنت سرویس های دیگری همچون پست الکترونیکی و یا FTP اجرا می شوند که هر کدام خدمات متفاوتی را عرضه می کنند. سرویس وب به منظور اشتراک اطلاعات بر روی اینترنت به کار می رود. از نظر کاربر وب این امر که اطلاعات دریافتی او از چه مکان فیزیکی دریافت شده اهمیتی ندارد. همچنین دیدی که کاربر از وب دارد مستقل از زمان و یا مکانی است که از طریق آن به اینترنت متصل شده و از خدمات وب بهره گرفته است. تمام آنچه کاربر مشاهده می کند فضایی است مجازی که دارای معماری اطلاعات بخصوصی می باشد و او می تواند از در این فضا اطلاعات مورد نیاز خود را دریافت کند.

سرویس وب در عین حال دارای محدودیت های خود می باشد. با استفاده از سرویس وب نمی توان در سطح فرایند ها (محاسبات) به امر اشتراک پرداخت و یا نمی توان سیستم های ذخیره سازی[1] را مورد استفاده مشترک قرار داد. اما این قابلیت ها را می توان از یک گرید انتظار داشت. به عبارت بهتر گرید سرویسی است بر روی بستری همچون اینترنت که امکان اشتراک انواع منابع در سطوح مختلف و استفاده از سرویس های متنوع را فراهم می آورد. در اینجا منظور از منبع هر نوع قابلیتی است که سیستم های محاسباتی و ذخیره سازی می توانند داشته باشند. نمونه هایی از این منابع شامل قدرت محاسباتی، ظرفیت ذخیره سازی، ظرفیت ارتباطی، و یا منابع جانبی همچون چاپگر می باشد.

کاربری که از خدمات گرید استفاده می کند انتظار دارد که مستقل از مکان و زمان عملیات بتواند از خدمات گرید بهره بگیرد. خدمات گرید شامل استفاده از منابع مشترک و سرویس های موجود می باشد. از دید کاربر گرید کامپیوتری مجازی[2] را برای او فراهم می آورد که کاربر می تواند از منابع آن کامپیوتر بدون در نظر گرفتن چگونگی و محل قرار گیری آن منابع به سهولت از آن ها استفاده کند. در این صورت وظیفه گرید را می توان هماهنگ کردن و در دسترس کردن منابع و در اختیار گذاشتن سرویس های مختلف دانست. با دقت در این مفهوم از گرید می توان ارتباط آن با وب را دریافت. در واقع وب را می توان گریدی ناکامل پنداشت که دسترسی به منابع را فراهم می آورد اما به خوبی از آن ها بهره نمی گیرد. اشتراکی که در سطح وب عرضه می شود اشتراکی محدود در سطح داده بوده و سایر قابلیت ها مورد توجه قرار نمی گیرند. البته با عرضه شدن بحث سرویس های وب و همچنین سرویس گرا شدن گرید ها این دو حوزه به یکدیگر بسیار نزدیک تر شده اند.

نکته دیگری که در بالا به آن اشاره شد این است که گرید بر روی بستری همچون اینترنت می تواند اجرا شود. از نظر تئوری گرید را می توان بر روی هر بستری که قابلیت دسترسی به منابع را فراهم کند به کار گرفت. اما باید در نظر داشت که گرید می خواهد همچون وب به صورت گسترده در اختیار کاربران باشد و لذا باید از بسترهای موجود که مورد استفاده همگان است نهایت بهره را بگیرد. لذا از نظر عملی می توان اینترنت را بستر خوبی برای گرید دانست. هرچند که ممکن است اینترنت فعلی دارای مشکلات خاصی باشد که باید به سمت اصلاح و تکامل پیش رود. پس در این صورت گرید را می توان سرویسی بر روی اینترنت دانست که امر اشتراک و بهره گیری از منابع را ممکن می سازد. در اینجا می توان جواب این سوال را که آیا گرید جایگزینی برای اینترنت است را نیز داد که گرید وسیله ای برای بهره گیری بهتر از اینترنت است نه جایگزینی برای آن.

از بحث انجام شده اهمیت مفهوم دیگری به نام The Grid مشخص می شود. در واقع همانطور که یک اینترتت و وب جهانی در اختیار داریم که در صورت نیاز به آن ها مراجعه می کنیم می توان یک

[1] Storage

[2] Virtual Computer

گرید جهانی را نیز تصور کرد که کاربران در صورت نیاز به بهره گرفتن از خدمات آن و استفاده از منابع مشترک به آن متصل می شوند و از آن استفاده می کنند. مسلما رسیدن به چنین مفهومی نیازمند تکامل و به بلوغ رسیدن تکنولوژی های مرتبط با گرید و به خصوص توسعه استانداردهای باز، فراگیر، و مورد قبول همگان همچون استانداردهای موجود بر روی اینترنت می باشد.

با وجود آن که از اهداف مهم گرید ارائه خدمات در مقیاس وسیع است، اما می توان گریدی را متصور بود که در مقیاس های کوچکتری به کار گرفته می شود به عنوان مثال در درون یک سازمان. به این نوع گرید ها می توان اطلاق گرید محلی[۴] نمود. این نوع گرید ها بر روی مجموعه محدودی از منابع با سیاست های تعریف شده درون سازمانی و جامعه کاربری شناخته شده مورد استفاده قرار می گیرند. در این نوع گرید ها ممکن است بحث هایی چون استفاده از سیکل های آزاد پردازنده های مربوط به کامپیوتر های متصل شده مورد توجه بیشتری قرار گیرد. همچنین بواسطه وجود ارتباط با سرعت های بالاتر امکان حل مسائل با دانه بندی ریز تر و مسائلی که نیاز به ارتباطات بیشتری دارند وجود دارد. این امور بر خلاف گرید های مقایس وسیع است که معمولا در پی آنند که از سرور های اختصاصی برای امر قدرت محاسباتی خود استفاده کنند و مسائل را در سطح کلان تری می بینند.

با توجه به آن که نام گرید از شبکه برق اقتباس شده است می توانیم با مقایسه ای بین گرید با شبکه برق به برخی از اهداف و مفاهیم اصلی مربوط به گرید برسیم. البته ذکر این نکته ضروری است که اگرچه شباهت هایی بین این دو موضوع وجود دارد اما نقاط افتراقی نیز در بین آن ها هست که نباید ما را دچار اشتباه کند. در جدول ۲-۱ برخی از این شباهت ها آمده است.

جدول ۲-۱: شباهت های موجود میان گرید و شبکه برق

شبکه برق	گرید
هنگامی که وسیله ای را به برق می زنیم برایمان مهم نیست که برق از کجا تامین شده و چگونه به این محل رسیده است. تنها می دانیم که هر وسیله ای را که متصل کنیم می تواند برق خود را از شبکه برق تامین کند.	هنگامی که از کامپیوتر استفاده می کنیم می دانیم که با اتصال به گرید می توان از خدمات آن بهره گرفت و به عنوان مثال سرویسی را اجرا کرد. اما اینکه این سرویس چگونه منابع خود را تامین می کند و در کجا اجرا می شود برایمان مهم نیست.
بستری که امر بالا را ممکن می سازد را با نام شبکه برق می شناسیم. این بستر امکان انتقال برق بین نقاط مختلف را فراهم می آورد. این بستر با استفاده از ایستگاه ها و خطوط برق و ترانسفورمر ها امکان دسترسی همگان به برق تولید شده در نیروگاه ها را فراهم می آورد.	بستری که کامپیوتر ها را به منظور اشتراک و بهره گیری از منابع به هم متصل می کند را می توان گرید نامید. بدین ترتیب گرید امکان دسترسی به انواع منابع از قبیل منابع محاسباتی و ذخیره سازی را برای استفاده کنندگان فراهم می آورد.

[4] Local Grid

شبکه برق به صورت ناپیدا[5] عمل می کند. بدین صورت که جزیئیات مربوط به تولید برق و انواع مختلف روش های تولید برق را از دید مصرف کننده مخفی نگاه می دارد.	گرید نیز به صورت ناپیدا عمل می کند. کاربر گرید لازم نیست از بسیاری از جزئیات و مشکلات مربوط به سیستم توزیع یافته آگاهی داشته باشد.
شبکه برق به صورت فراگیر[6] وجود دارد. تقریبا در هر جایی و با استفاده از یک پریز برق استاندارد می توان از خدمات آن استفاده کرد.	هدف از گرید نیز آن است که به صورت فراگیر موجود باشد. وسائل مختلف از قبیل کامپیوتر های شخصی و موبایل ها باید بتوانند با استفاده از واسطی استاندارد از خدمات گرید بهره ببرند.
شبکه برق به عنوان یک خدمت عمومی[7] مورد استفاده قرار می گیرد. بدین صورت که برای استفاده از آن درخواست داده می شود و در صورت استفاده از نیز هزینه آن باید پرداخته شود.	برخی از سرویس های گرید نیز ممکن است به صورت خدمات عمومی عرضه شود. کابران در صورت استفاده از این خدمات هزینه آن ها را نیز می پردازند. که این مفهوم با نام Utility Computing نیز شناخته می شود.
شبکه برق این نیاز را که هر کسی باید برای خود ژنراتور برق بخرد و برق خود را تامین کند از بین برد. زیرا این نوع تامین برق هم گران بود و هم سخت.	با استفاده از گرید لازم نیست که برای انجام یک کار تمامی منابع مربوط به آن را به صورت محلی در اختیار داشته باشیم بلکه در صورت نیاز (On Demand) می توان از سایر منابع موجود استفاده کنیم که در نتیجه هزینه این کار کاهش می یابد.

البته همانطور که گفته شد این تشابه اسمی تا حدی مفاهیم مربوط به گرید را مشخص می کند اما در برخی از جنبه ها تفاوت هایی نیز وجود دارد. به عنوان مثال همگان خدمت مشابهی را از شبکه برق می گیرند اما در گرید ممکن است با توجه به ویژگی های کیفی مورد نظر کاربر و یا محدودیت های موجود برای کاربران، سرویس ها با کیفیت مختلفی به کاربران مختلف و یا حتی در زمان های مختلف داده شود [۲۹].

یکی دیگر از راه های شناخت گرید توجه به ویژگی ها و قابلیت های آن و نیازمندی های مورد انتظار از آن است. نیازمندی هایی که یک گرید باید بتواند یا می تواند پوشش دهد متنوع و بسیار بوده اما با جمع آوری لیستی از آن ها می توان به حوزه عملکرد گرید و ارتباط آن با سایر تکنولوژی ها پی برد. در زیر لیستی از این نیازمندی ها و قابلیت ها آورده شده است [۳۰]. مسلما لازم نیست که یک گرید تمامی این نیازمندی ها را در بر داشته باشد اما بدین ترتیب می توان یک گرید ایده آل را، که در واقع هدف اصلی گریدهای امروزی است، بهتر شناخت.

[5] Transparent

[6] Pervasive

[7] Utility

- فراهم آوردن قابلیت اشتراک و دسترسی به منابعی که به صورت توزیع یافته و در سطح گسترده قرار گرفته اند.
- کشف منابع موجود و جدید در گرید. همچنین آگاهی از تغییرات رخ داده در این منابع.
- هر دارنده و عرضه کننده منبع باید بتواند سیاست های[8] استفاده از منبع خود را در گرید مشخص کند. سیاست تعریف شده می تواند مثلا محدودیت استفاده فقط برای اعضاء و یا محدودیت های زمانی باشد. این محدودیت ها می توانند به صورت قواعد و یا سایر مدل ها بیان شوند.
- کاربر گرید باید بتواند شرایط استفاده از یک منبع مثل کیفیت خدماتی[9] که مورد نظرش می باشد را به گرید اعلام کند. به عنوان مثال کاربر ممکن است مشخص کند که فقط می خواهد از منابع امن استفاده کند.
- استفاده هماهنگ و کنترل شده از منابع مشترک. گرید باید با توجه به شرایط درخواستی و با در نظر گرفتن قرار دادهای فی ما بین و همچنین سیاست های تعریف شده بر روی هر منبع، استفاده از منابع را زمان بندی و مدیریت کند تا تمامی شرایط تا حد ممکن بر قرار باشند.
- کیفیت خدمت ارائه شده باید پایدار باشد.
- خدمات گرید باید قابل اتکا[10] باشد. قابیت اتکا شامل تضمین موارد مختلفی همچون انطباق پذیری، در دسترس بودن، و داشتن امنیت کافی می باشد.
- در صورتیکه استفاده از منبع مستلزم هزینه هایی است گرید باید بتواند هزینه استفاده از این سرویس ها را محاسبه و نگه داری کند و به عرضه کننده آن منبع اطلاع دهد.
- گرید باید بتواند از نظر امنیتی تضمین های لازم را فراهم آورده و از بستر های موجود به خوبی بهره برداری کند.
- گرید باید با استفاده از استانداردهای باز[11] توسعه یافته باشد. این امر باعث می شود که گرید مورد توافق همگان باشد و در نتیجه استفاده همگانی و گسترده از آن میسر شود. همچنین وجود استانداردها موجب ایجاد قابلیت همکاری و یا یکپارچه سازی برنامه ها و یا حتی گریدهای مختلف می شود.

---

[8] Policies

[9] Quality of Service (QoS)

[10] Dependable

[11] Open Standard