

معماری مبتنی بر معماری باز سرویس‌های گرید برای داده‌کاوی موازی و توزیع‌شده بر روی گرید

علی غفاری نژاد، محمد کاظم اکبری

دانشکده مهندسی کامپیوتر و فناوری اطلاعات دانشگاه صنعتی امیرکبیر

E-mail: alighaffarin@yahoo.com, akbari@ce.aut.ac.ir.

چکیده - امروزه حجم بسیار بالایی اطلاعات در مخازن داده الکترونیکی ذخیره شده است و به سرعت بر حجم این اطلاعات افزوده می‌شود. زمانی که حجم بالای داده‌ها در کنار توزیع‌شدگی جغرافیایی داده‌ها، کاربران، و سیستم‌ها قرار می‌گیرد، لازم است تا از ترکیب فناوریهای مختلفی با هدف پیاده‌سازی سیستم‌های اکتشاف دانش توزیع‌شده با کارایی بالا استفاده گردد. از طرفی گرید محاسباتی استاندارد برای توسعه و گسترش برنامه‌های کاربردی توزیع‌شده با کارایی بالا است. در این مقاله ما به تشریح اصول طراحی و نحوه پیاده‌سازی یک معماری نرم افزاری جدید مبتنی بر معماری باز سرویس‌های گرید زیرساختاری جدید به منظور انجام داده‌کاوی موازی و توزیع‌شده بر روی گرید پرداخته ایم. این معماری بر روی *Globus Toolkit* پیاده‌سازی شده است و از *OGSA-DAI* برای دسترسی به منابع داده بر روی گرید استفاده می‌کند. تحلیل کارایی انجام گرفته بیانگر کارایی معماری ارائه شده برای انجام داده‌کاوی موازی و توزیع‌شده بر روی حجم بالای داده است.

کلید واژه - داده‌کاوی موازی و توزیع‌شده، گرید محاسباتی، معماری باز سرویس‌های باز گرید، اکتشاف دانش با کارایی بالا.

۱- مقدمه

محاسبات توزیع‌شده با کارایی بالا که نیاز به منابع به اشتراک گذاشته دارند واسطه‌هایی را پیشنهاد داده است که می‌توانند به صورت پویا و با فراخوانی کارخانه هر سرویس گرید از آنها نمونه‌ای حالت‌دار ایجاد کنند [2][4]. یک سرویس گرید شامل PortType هایی است که هرکدام از آنها عملیاتی را که یک سرویس گرید می‌تواند انجام دهد را مشخص می‌کنند. اطلاعات پویای هر سرویس گرید در عناصر داده سرویس ذخیره می‌شود. PortType های مهم یک سرویس گرید عبارتند از:

- کارخانه: این PortType یک سرویس نمونه را به صورت پویا ایجاد می‌کند و بعد از ایجاد سرویس نمونه یک دستگیره سرویس گرید منحصر به فرد که به سرویس نمونه گرید اشاره می‌کند را برمی‌گرداند.

- راهنما: سرویس‌های دیگر گرید در این PortType ثبت نام می‌کنند و با کمک آن می‌توان در سرویس‌های ثبت شده جستجو کرد.

امروزه بسیاری از شرکت‌های تجاری در بازارهای جهانی حضور فعال دارند. این سازمان‌ها برای به دست آوردن دانش از داده‌ها نیازمند انجام داده‌کاوی بر روی حجم عظیمی از داده‌های توزیع‌شده و نامتجانس هستند. داده‌کاوی به معنای استخراج دانش از حجم عظیمی از داده‌ها است. گرید به منظور پشتیبانی از برنامه‌های کاربردی که نیاز به کارایی بالا، توزیع‌شدگی، همکاری، و تعامل‌های پیچیده بین منابع خودمختار و پراکنده طراحی شده است. امروزه نسل جدید فناوری گرید به طرف معماری باز سرویس‌های گرید یا OGSA گرایش پیدا کرده است [1]. معماری باز سرویس‌های گرید یک چهارچوب سطح بالا است که برای پشتیبانی از سازمان‌های مجازی که منبع و داده‌ها با مدیریت مستقل را بر روی شبکه‌ای از کامپیوترهای ناهمگن به اشتراک می‌گذارند طراحی شده است [2]. طراح OGSA، مجمع جهانی گرید [3]، برای پاسخگویی به نیازهای کاربردهای