

بررسی چالش های رایانش ابری در داده های بزرگ

سید محمد جواد مقدم^۱، صالح سلطان پور^۲

۱-عضویت علمی، گروه کامپیوتر، دانشگاه بزرگمهر، قائنات

۲-دانشکده مهندسی کامپیوتر، گروه نرم افزار، دانشگاه آزاد اسلامی، بیرجند

چکیده

داده های بزرگ در حال حاضر یکی از مهم ترین فن آوری های در حال ظهور است. داده های بزرگ به عنوان یک مفهوم استفاده می شود که به ناتوانی معماری داده های سنتی مربوط می شود تا بتواند مجموعه داده های جدید را مدیریت کند. داده های بزرگ ۴ - ۷ حجم، سرعت، تنوع و درستی، مدیریت داده ها و تجزیه و تحلیل را برای انبارهای سنتی داده ها به چالش می کشد. گسترش سند های داده ها و افزایش رسانه های چندرسانه ای، رسانه های اجتماعی و اینترنت افزایش حجم اطلاعات ساختاری، نیمه ساختاری و غیر ساختاری را تولید می کند. در راستای بررسی این حجم زیادی از داده ها، داده های بزرگ و تجزیه و تحلیل داده ها تبدیل شده اند. ابر رایانه به نظر می رسد وسیله ای مناسب برای میزبانی کارهای داده بزرگ است. با این حال، کار بر روی داده های بزرگ در ابر، چالش خود را برای تطبیق دو اصل متناقض طراحی می کند. توانایی ذخیره داده های بزرگی از داده ها در فرم های مختلف و پردازش آن با سرعت بسیار زیاد، اطلاعاتی را به ارمغان می آورد که می تواند موسسات تجاری و موسسات آموزشی را سریعاً در حال توسعه کند. با این وجود، نگرانی زیادی در مورد حریم خصوصی و امنیت در هنگام حرکت به ابر وجود دارد که دلیل اصلی این است که چرا کسب و کار و موسسات آموزشی به ابر حرکت نمی کنند. محاسبات ابر بر پایه مفاهیم تثبیت و تلفیق منابع است، اما سیستم های داده ای بزرگ مانند Hadoop بر اساس اصل مشترک است که هر گره مستقل و خودکفا است. داده های بزرگ یکپارچه با فناوری های محاسبات ابری، کسب و کار و موسسات آموزشی می تواند مسیری بهتر برای آینده داشته باشد.

کلمات کلیدی: داده های بزرگ، محاسبات ابر، Hadoop