

عنوان مقاله:

آناتومی کارهای نگاشت کاهش و بررسی چالش های زمان بندی، کارایی و امنیت در هادوپ

محل انتشار:

کنفرانس ملی فناوری های نوین در مهندسی برق و کامپیوتر (سال: 1396)

تعداد صفحات اصل مقاله: 20

نویسندگان:

محمد علی ترکمانی - گروه مهندسی کامپیوتر، واحد یاسوج، دانشگاه آزاد اسلامی، یاسوج، ایران

کیوان رحیمی زاده - استادیار، دانشکده فنی و مهندسی، دانشگاه یاسوج، یاسوج، ایران

خلاصه مقاله:

هادوپ یک ابزار متن باز مبتنی بر رهیافت نگاشت-کاهش در زمینهی ذخیرهسازی و پردازش داده های کلان است که تحقق انقلاب داده های کلان را بیشتر از پیش ممکن ساخته است. به جای تکیه بر سخت افزارهای اختصاصی گران قیمت، خوشه های هادوپ معمولا از صدها هزار ماشین معمولی چندهسته ای تشکیل شده اند. هادوپ کدها را به جای انتقال داده ها به گره های پردازشی، به ماشینی که داده ها در آنها قرار دارند منتقل می کند که این امر مقیاس پذیری را افزایش میدهد. قابلیت پردازش مقادیر بالای داده های متنوع به شکل توزیع شده و موازی همراه با تحمل پذیری خطا و استفاده از نرم افزارهای رایگان و سخت افزارهای مناسب ارزان قیمت، استفاده از هادوپ را به عنوان سکوی انتخابی داده های کلان برای بیشتر سازمان های تجاری و دولتی توجیه پذیر می کند. با این حال، ساخت یک کار نگاشت کاهش که چند ترابایت داده را در بین صدها ماشین در زمانی معقول خوانده و پردازش کند، چالش بزرگی است. علاوه بر این، روش های حفاظتی سنتی داده ها برای داده های کلان قابل استفاده نیستند، به همان اندازه که حجم داده ها افزایش می یابد، ریسک های امنیتی نیز شدیدتر می شود سکوهایی داده های کلان به طور خاص برای پشتیبانی از فرم پیشرفتهای از تحلیل طراحی شده اند که کارایی دقیق و الزامات مقیاس پذیری را برآورده میکند. با این حال، تاکنون توجه خاصی معطوف حفاظت از داده ها نشده است. در واقع، با وجود آنکه داده های تحلیل شده اغلب حاوی اطلاعات شخصی و حساس هستند و تهدیداتی در ارتباط با امنیت و حریم خصوصی متوجه آنها است. این مقاله ابتدا اکوسیستم هادوپ را توصیف کرده و جزئیاتی را در مورد موتور نگاشت کاهش موجود در هسته هادوپ ارائه می کند. همچنین یک برنامه ساده نگاشت کاهش به زبان جاوا بررسی خواهد شد که این برنامه می تواند رشته ها را معکوس کند. سپس به بحث در مورد برنامه های زمان بندی نگاشت کاهش مختلف و تاثیر آن ها بر کارایی می پردازد. همچنین چالش های امنیتی هادوپ بررسی خواهد شد

کلمات کلیدی:

داده های کلان، فایل سیستم توزیع شده هادوپ، کارایی، زمان بندی، امنیت، نگاشت کاهش هادوپ

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/758759>

