

عنوان فارسی مقاله :

چارچوب تحلیلی داده های بزرگ برای مدیریت داده های علمی

عنوان انگلیسی مقاله :

A big data analytics framework for scientific data management



توجه !

این فایل تنها قسمتی از ترجمه میباشد. برای تهیه مقاله ترجمه شده کامل با فرمت ورد (قابل ویرایش) همراه با نسخه انگلیسی مقاله، [اینجا](#) کلیک نمایید.

IV. CONCLUSIONS AND FUTURE WORK

The Ophidia analytics framework, a core part of the Ophidia research project, has been presented. As discussed in this work, the analytics framework is responsible for atomically processing, transforming and manipulating array-based data, by providing a common way to run on large clusters analytics tasks applied to big datasets. The paper has highlighted the design principles, the algorithm and the most relevant implementation aspects associated to the Ophidia analytics framework. Besides a comprehensive list of the most relevant operators, three of them have been also presented in detail (OPH_APPLY and OPH_PUBLISH for the data part, OPH_CUBE_ELEMENTS for the metadata part). We have also presented some promising experimental results involving two operators (OPH_APPLY and OPH_PUBLISH) executing in a real cluster environment.

We plan in future work to develop an extended set of parallel operators to support new scientific use cases. Array-based primitives extensions, a data analytics query language and an optimized query planner will be considered to support more complex operators and dataflow driven requests. A comprehensive analytics benchmark will be also defined and implemented to further evaluate the performance of the system.

نتایج و کار آتی

Ophidia بخش اصلی پروژه تحقیقاتی، Ophidia چارچوب تحلیلی مطرح شده است. همان گونه که در این کار بحث شد، چارچوب تحلیلی مسئولیت پردازش، تبدیل و دستکاری اتمی داده های مبتنی بر آرایه، با فراهم نمودن راهی معمول و عمومی برای اجرا روی وظایف تحلیلی خوشه های بزرگ را برعهده دارد که در مجموعه داده های بزرگ کاربرد دارند. مقاله حاضر اصول طراحی، الگوریتم و اصلی ترین جنبه راروشن و راجع به Ophidia های اجرای وابسته به چارچوب تحلیلی آن ها توضیح داده است. به علاوه، لیستی جامع از وابسته ترین اپراتورها عرضه می کند که سه مورد از آنها به طور مفصل مطرح شده، برای بخش داده ها (oph_apply و oph_publish) اند (برای بخش ابرداده ها). در اینجا برخی از نتایج oph_cube_elements و oph_apply (آزمایشی امیدوارکننده من جمله اجرای دو اپراتور در محیط خوشه واقعی نیز مطرح شده است) oph_publish قصد داریم در کار آتی مجموعه ای توسعه یافته از اپراتورهای موازی برای پشتیبانی از موارد استفاده علمی جدید را توسعه دهیم. توسعه توابع اولیه مبتنی بر آرایه، زبان پرس و جوی تحلیلی داده ها و یک برنامه ریز پرس و جوی بهینه شده از اپراتورهای پیچیده تر و درخواست های مبتنی بر جریان داده ها (جریان داده مدار) پشتیبانی می کنند. محک تحلیلی جامع تعریف و برای ارزیابی هر چه بیشتر عملکرد سیستم، اجرا می شود



توجه!

این فایل تنها قسمتی از ترجمه میباشد. برای تهیه مقاله ترجمه شده کامل با فرمت

ورد (قابل ویرایش) همراه با نسخه انگلیسی مقاله، [اینجا](#) کلیک نمایید.

همچنین برای مشاهده سایر مقالات این رشته [اینجا](#) کلیک نمایید.