Received: 17 June 2019 | Accepted: 18 June 2019
DOI: 10.1002/cpe.5452

Check for updates

WILEY

EDITORIAL

Special Issue: Graph Computing

1 | BACKGROUND

Graph computing now is popular in many areas, including social network and gene sequence alignment. Graph computing system and algorithm have a history prior to the use of graph databases and have a future that is not necessarily entangled with typical database concerns. With the data's increasing size, many distributed graph-computing systems have been developed in recent years to process and analyze massive graphs. Researchers pay more attention on the graph partition schemes on distributed environment. However, other researchers think a single system can avoid the network overhead and may have better performance even if the data size is too big for the memory space. With the rapid development of coprocessors, some researchers think it is promising to build a domain specific computer, just for graph computing.

The proposed special issue of *Concurrency and Computation: Practice and Experience* contains revised and extended versions of selected best papers with respect to graph computing at the 21st IEEE International Conference on Parallel and Distributed Systems (ICPADS'16), which was held at Wuhan, China, on December 13-16, 2016. Established in 1992, ICPADS has been a major international forum for scientists, engineers, and users to exchange and share their experiences, new ideas, and latest research results on all aspects of parallel and distributed computing systems. The purpose of this special issue is to provide a comprehensive view into recent advances in systems software, algorithms, partition schemes, and even graph computer based on new advances in computer architecture and applications.

The five selected papers are summarized as follows.

The first paper, titled "An efficient iterative graph data processing framework based on bulk synchronous parallel model" by Liu et al,[1] presents an efficient computational framework for graph data processing based on the bulk synchronous parallel model. Existing Pregel-like graph processing systems remains in its early stage, and there still exist many challenges with prohibitive superstep-synchronized overhead. Furthermore, the graph data partition strategy in these earlier graph systems fails to support load balancing, therefore causing the increase of network I/O overhead as the scale of graph data grows. Thus, this paper leverages a global synchronization mechanism to enhance the performance of graph computation. Meanwhile, a balanced hash-based graph partition mechanism is presented to optimize the large-scale graph data processing. The work has a real implementation upon on Pregrel system, which can better support a variety of graph analytics applications.

The second paper, titled "An efficient iterative graph data processing framework based on bulk synchronous parallel model" by Linchen Yu,[2] proposes an optimized scheduling system for parallelizing the programs in the Xen. Virtualization challenges the traditional CPU scheduling, leading that the spin lock in virtualized environment can be preempted by the VMM, increasing synchronization overhead and decreasing the performance of parallel programs. Many studies have proposed the co-scheduling to alleviate this problem. However, these earlier attempts are not suitable to non-parallel workloads with the CPU fragmentation problem as well. Therefore, a simultaneous optimization scheduling system, called CCHybrid, is proposed in the Xen virtualized environment. Results show the efficiency of CCHybrid over the traditional Xen Credit scheduler.

The third paper, titled "ms-PoSW: A multi-server aided proof of shared ownership scheme for secure deduplication in cloud" by Xiong et al,[3] introduces a novel concept of the Proof for securing client-side deduplication of the shared files. With the rapid development of cloud computing and big data technologies, collaborative cloud applications are inextricably linked to our daily life and, therefore, produce a large number of shared files, which is challenging for secure access and data duplication in cloud. This paper proposes a novel multiserver-aided PoSW scheme for collaborative cloud applications and propose a hybrid PoSW scheme to reduce the computational cost of the shared owner's client. Furthermore, a hybrid PoSW scheme is constructed to address the secure proof of hybrid cloud architectures.

The fourth paper, titled "Sparse random compressive sensing based data aggregation in wireless sensor networks" by Yin et al,[4] introduces a compressive data aggregation scheme. In wireless sensor networks, the increasingly expanding data volume has high spatial-temporal correlation. Although some earlier studies attempt to eliminate data redundancy, few can handle energy consumption and latency simultaneously. In this paper, the authors a delay-minimum energy-balanced data aggregation method, which can eliminate the redundancy among the readings and prolong the network lifetime. A sparse random matrix is adopted as a measurement matrix to balance communication cost. Particularly, each measurement can form an aggregation tree with minimum delay. Furthermore, a novel scheduling method is used to avoid information interference as well.

The fifth paper, titled "Dynamic cluster strategy for hierarchical rollback-recovery protocols in MPI HPC applications" by Liao et al,[5] proposes a dynamic cluster strategy to adapt to the runtime variation of communication pattern by using a prediction scheme. The idea comes from a fact that Hierarchical rollback-recovery protocols provide failure containment and reduce the amount of message to be logged, making it an attractive and scalable solution for fault tolerance even at a large scale. This paper shows how the communication pattern changes with the stages

© 2019 John Wiley & Sons, Ltd.

of application because MPI HPC applications scale up and become more complex. Therefore, to further increase the efficiency of hierarchical rollback-recovery protocols, the authors propose a dynamic cluster strategy (DCS) to adapt to the change of communication pattern. In contrast to the existing static process partition algorithms, this strategy adopts a prediction mechanism by using the clusters of processes obtained from prior part of applications in the succeeding part. Detailed experiments are then performed to evaluate the effectiveness and efficiency DCS at an extremely large scale.

We hope that the readers would find the contents of this special issue interesting and further inspire them to look ahead into the challenges of designing, exploring, and exploiting graph analytics applications.

ORCID

Hai Jin https://orcid.org/0000-0002-3934-7605
Robert Lovas https://orcid.org/0000-0001-9409-2855

Hai Jin[1]
Xipeng Shen[2]
Robert Lovas[3]
Xiaofei Liao[1]
[1]*School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, China*
[2]*Department of Computer Science, North Carolina State University, Raleigh, North Carolina*
[3]*Hungarian Academy of Sciences, Budapest, Hungary*

Correspondence

Hai Jin, School of Computer Science and Technology, Huazhong University of Science and Technology, Luoyu Road 1037, Hongshan District, Wuhan 430074, China.
Email: hjin@hust.edu.cn

REFERENCES

1. Liu C, Zeng D, Yao H, Yan X, Yu L, Fu Z. An efficient iterative graph data processing framework based on bulk synchronous parallel model. *Concurrency Computat Pract Exper*. 2020;32:e4432. https://doi.org/10.1002/cpe.4432
2. Yu L. CCHybrid: CPU co-scheduling in virtualization environment. *Concurrency Computat Pract Exper*. 2020;32:e4213. https://doi.org/10.1002/cpe.4213
3. Xiong J, Zhang Y, Lin L, Shen J, Li X, Lin M. ms-PoSW: A multi-server aided proof of shared ownership scheme for secure deduplication in cloud. *Concurrency Computat Pract Exper*. 2020;32:e4252. https://doi.org/10.1002/cpe.4252
4. Yin L, Liu C, Guo S, Yang Y. Sparse random compressive sensing based data aggregation in wireless sensor networks. *Concurrency Computat Pract Exper*. 2020;32:e4455. https://doi.org/10.1002/cpe.4455
5. Liao X, Zheng L, Zhang B, Zhang Y, Jin H, Shi X, Lin Y. Dynamic cluster strategy for hierarchical rollback-recovery protocols in MPI HPC applications. *Concurrency Computat Pract Exper*. 2020;32:e4173. https://doi.org/10.1002/cpe.4173